Christoph Nils Bannwarth

Development and Application of Efficient Methods for the Computation of Electronic Spectra of Large Systems

Dissertation

Development and Application of Efficient Methods for the Computation of Electronic Spectra of Large Systems

Dissertation
zur
Erlangung des Doktorgrades (Dr. rer. nat.)
der
Mathematisch-Naturwissenschaftlichen Fakultät
der
Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von
Christoph Nils Bannwarth
aus Köln
–2017–

Dekan:	Prof. Dr. Johannes Beck
Erster Gutachter:	Prof. Dr. Stefan Grimme
Zweiter Gutachter:	Prof. Dr. Thomas Bredow
Tag der Disputation:	08.03.2018
Erscheinungsjahr:	2018

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Rheinischen Friedrich-Wilhelms-Universität Bonn

To my parents

Publications

Parts of this thesis have been published in peer-reviewed journals:

1. Bannwarth, C.; Hansen, A.; Grimme, S. "The Association of Two 'Frustrated' Lewis Pairs by State-of-the-Art Quantum Chemical Methods", *Isr. J. Chem.* **2015**, *55*, 235–242.

2. Gütz, C.; Hovorka, R.; Klein, C.; Jiang, Q.-Q.; Bannwarth, C.; Engeser, M.; Schmuck, C.; Assenmacher, W.; Mader, W.; Topić, F.; Rissanen, K.; Grimme, S.; Lützen, A. "Enantiomerically Pure $[M_6L_{12}]$ or $[M_{12}L_{24}]$ Polyhedra from Flexible Bis(Pyridine) Ligands", *Angew. Chem. Int. Ed.* **2014**, *53*, 1693–1698.

3. Bannwarth, C.; Grimme, S. "A simplified time-dependent density functional theory approach for electronic ultraviolet and circular dichroism spectra of very large molecules", *Comput. Theor. Chem.* **2014**, *1040–1041*, 45–53.

4. Bannwarth, C.; Grimme, S. "Electronic Circular Dichroism of Highly Conjugated π-Systems: Breakdown of the Tamm–Dancoff/Configuration Interaction Singles Approximation", *J. Phys. Chem. A* **2015**, *119*, 3653–3662.

5. Bannwarth, C.; Seibert, J.; Grimme, S. "Electronic Circular Dichroism of [16]Helicene With Simplified TD-DFT: Beyond the Single Structure Approach", *Chirality* **2016**, *28*, 365–369.

6. Grimme, S.; Bannwarth, C. "Ultra-fast computation of electronic spectra for large systems by tight-binding based simplified Tamm-Dancoff approximation (sTDA-xTB)", *J. Chem. Phys.* **2016**, *145*, 054103.

7. Grimme, S.; Bannwarth, C.; Shushkov, P. "A Robust and Accurate Tight-Binding Quantum Chemical Method for Structures, Vibrational Frequencies, and Noncovalent Interactions of Large Molecular Systems Parametrized for All spd-Block Elements (Z = 1–86)", *J. Chem. Theory Comput.* **2017**, *13*, 1989–2009.

Further publications:

1. a) Hounjet, L. J.; Bannwarth, C.; Garon, C. N.; Caputo, C. B.; Grimme, S.; Stephan, D. W. "Combinations of Ethers and $B(C_6F_5)_3$ Function as Hydrogenation Catalysts", *Angew. Chem. Int. Ed.* **2013**, *52*, 7492–7495.
   b) Hounjet, L. J.; Bannwarth, C.; Garon, C. N.; Caputo, C. B.; Grimme, S.; Stephan, D. W. "Combinations of Ethers and $B(C_6F_5)_3$ Function as Hydrogenation Catalysts", *Angew. Chem.* **2013**, *125*, 7640–7643.

2. Gütz, C.; Hovorka, R.; Klein, C.; Jiang, Q.-Q.; Bannwarth, C.; Engeser, M.; Schmuck, C.; Assenmacher, W.; Mader, W.; Topić, F.; Rissanen, K.; Grimme, S.; Lützen, A. "Enantiomerenreine $[M_6L_{12}]$- oder $[M_{12}L_{24}]$-Polyeder aus flexiblen Bis(pyridin)-Liganden", *Angew. Chem.* **2014**, *126*, 1719–1724.

3. Hansen, A.; Bannwarth, C.; Grimme, S.; Petrović, P.; Werlé, C.; Djukic, J.-P. "The Thermochemistry of London Dispersion-Driven Transition Metal Reactions: Getting the 'Right Answer for the Right Reason' ", *ChemistryOpen* **2014**, *3*, 177–189.

4. Yu, J.; Kehr, G.; Daniliuc, C. G.; Bannwarth, C.; Grimme, S.; Erker, G. "Direct synthesis of a geminal zwitterionic phosphonium/hydridoborate system - developing an alternative tool for generating frustrated Lewis pair hydrogen activation systems", *Org. Biomol. Chem.* **2015**, *13*, 5783–5792.

5. Regeta, K.; Bannwarth, C.; Grimme, S.; Allan, M. "Free electrons and ionic liquids: study of excited states by means of electron-energy loss spectroscopy and the density functional theory multireference configuration interaction method", *Phys. Chem. Chem. Phys.* **2015**, *17*, 15771–15780.

6. Chen, G.-Q.; Türkyilmaz, F.; Daniliuc, C. G.; Bannwarth, C.; Grimme, S.; Kehr, G.; Erker, G. "Enamine/butadienylborane cycloaddition in the frustrated Lewis pair regime", *Org. Biomol. Chem.* **2015**, *13*, 10477–10486.

7. Grimme, S.; Brandenburg, J. G.; Bannwarth, C.; Hansen, A. "Consistent structures and interactions by density functional theory with small atomic orbital basis sets", *J. Chem. Phys.* **2015**, *143*, 054107.

8. Jarzebski, A.; Bannwarth, C.; Tenten, C.; Benkhäuser, C.; Schnakenburg, G.; Grimme, S.; Lützen, A. "Synthesis, Chiral Resolution, and Absolute Configuration of Functionalized Tröger's Base Derivatives: Part III", *Synthesis* **2015**, *47*, 3118–3132.

9. Musina, E. I.; Shamsieva, A. V.; Strelnik, I. D.; Gerasimova, T. P.; Krivolapov, D. B.; Kolesnikov, I. E.; Grachova, E. V.; Tunik, S. P.; Bannwarth, C.; Grimme, S.; Katsyuba, S. A.; Karasik, A. A.; Sinyashin, O. G. "Synthesis of novel pyridyl containing phospholanes and their polynuclear luminescent copper(I) complexes", *Dalton Trans.* **2016**, *45*, 2250–2260.

10. Grimme, S.; Hansen, A.; Brandenburg, J. G.; Bannwarth, C. "Dispersion-Corrected Mean-Field Electronic Structure Methods", *Chem. Rev.* **2016**, *116*, 5105–5154.

11. Masnyk, M.; Butkiewicz, A.; Górecki, M.; Luboradzki, R.; Bannwarth, C.; Grimme, S.; Frelek, J. "Synthesis and Comprehensive Structural and Chiroptical Characterization of Enones Derived from (–)-α-Santonin by Experiment and Theory", *J. Org. Chem.* **2016**, *81*, 4588–4600.

12. Petrović, P.; Djukic, J.-P.; Hansen, A.; Bannwarth, C.; Grimme, S. "Non-covalent Stabilization in Transition Metal Coordination and Organometallic Complexes", *Non-covalent Interactions in the Synthesis and Design of New Compounds* **2016**, 115–143.

13. Allan, M.; Regeta, K.; Gorfinkiel, D. J.; Mašín, Z.; Grimme, S.; Bannwarth, C. "Recent research directions in Fribourg: nuclear dynamics in resonances revealed by 2-dimensional EEL spectra, electron collisions with ionic liquids and electronic excitation of pyrimidine", *Eur. Phys. J. D* **2016**, *70*, 123.

14. Ratzke, W.; Schmitt, L.; Matsuoka, H.; Bannwarth, C.; Retegan, M.; Bange, S.; Klemm, P.; Neese, F.; Grimme, S.; Schiemann, O; Lupton, J. M.; Höger, S. "Effect of Conjugation Pathway in Metal-Free Room-Temperature Dual Singlet–Triplet Emitters for Organic Light-Emitting Diodes", *J. Phys. Chem. Lett.* **2016**, *7*, 4802–4808.

15. Frömel, S.; Daniliuc, C. G.; Bannwarth, C.; Grimme, S.; Bussmann, K.; Kehr, G.; Erker, G. "Indirect synthesis of a pair of formal methane activation products at a phosphane/borane frustrated lewis pair", *Dalton Trans.* **2016**, *45*, 19230–19233.

16. Burganov, T. I.; Zhukova, N. A.; Mamedov, V. A.; Bannwarth, C.; Grimme, S.; Katsyuba, S. A. "Benzimidazolylquinoxalines: novel fluorophores with tuneable sensitivity to solvent effects", *Phys. Chem. Chem. Phys.* **2017**, *19*, 6095–6104.

17. a) Struch, N.; Bannwarth, C.; Ronson, T. K.; Lorenz, Y.; Mienert, B.; Wagner, N.; Engeser, M.; Bill, E.; Puttreddy, R.; Rissanen, K.; Beck, J.; Grimme, S.; Nitschke, J. R.; Lützen, A. "An Octanuclear Metallosupramolecular Cage Designed To Exhibit Spin-Crossover Behavior", *Angew. Chem. Int. Ed.* **2017**, *56*, 4930–4935.
b) Struch, N.; Bannwarth, C.; Ronson, T. K.; Lorenz, Y.; Mienert, B.; Wagner, N.; Engeser, M.; Bill, E.; Puttreddy, R.; Rissanen, K.; Beck, J.; Grimme, S.; Nitschke, J. R.; Lützen, A. "Ein achtkerniger metallosupramolekularer Würfel mit Spin-Crossover-Eigenschaften", *Angew. Chem.* **2017**, *129*, 5012–5017.

18. Strelnik, I. D.; Musina, E. I.; Ignatieva, S. N.; Balueva, A. S.; Gerasimova, T. P.; Katsyuba, S. A.; Krivolapov, D. B.; Dobrynin, A. B.; Bannwarth, C.; Grimme, S.; Kolesnikov, I. E.; Karasik, A. A.; Sinyashin, O. G. "Pyridyl Containing 1,5-Diaza-3,7-diphosphacyclooctanes as Bridging Ligands for Dinuclear Copper(I) Complexes", *Z. Anorg. Allg. Chem.* **2017**, *643*, 895–902.

19. Grimme, S.; Bannwarth, C.; Caldeweyher, E.; Pisarek, J.; Hansen, A. "A general intermolecular force field based on tight-binding quantum chemical calculations", *J. Chem. Phys.* **2017**, *147*, 161708.

20. Caldeweyher, E.; Bannwarth C.; Grimme, S. "Extension of the D3 dispersion coefficient model", *J. Chem. Phys.* **2017**, *147*, 034112.

21. Seibert, J.; Bannwarth, C.; Grimme, S. "Biomolecular Structure Information from High-Speed Quantum Mechanical Electronic Spectra Calculation", *J. Am. Chem. Soc.* **2017**, *139*, 11682–11685.

Presentations:

1. Poster: "Excited States of Large Systems using a simplified Tamm-Dancoff Approximation to TDDFT", *European Summerschool in Quantum Chemistry*, **2013**, Palermo, Italy.

2. Poster: "Excited States of Large Systems using a simplified Tamm-Dancoff Approximation to TDDFT", *49th Symposium on Theoretical Chemistry*, **2013**, Erlangen, Germany.

3. Poster: "Excited States of Large Systems using a simplified Tamm-Dancoff Approximation to TDDFT", *Excited States and Complex Environments*, **2013**, Münster, Germany.

4. Poster: "Excited States of Large Systems using a simplified Tamm-Dancoff Approximation to TDDFT", *29th Winter School in Theoretical Chemistry*, **2013**, Helsinki, Finland.

5. Poster: "Electronic Circular Dichroism of very large Systems by the simplified TD-DFT Approach", *Sostrup Summerschool: Quantum Chemistry and Molecular Properties*, **2014**, Ry, Denmark.

6. Talk: "The simplified TD-DFT and TDA approaches: Excited states and spectra for very large systems", *50th Symposium on Theoretical Chemistry*, **2014**, Vienna, Austria.

7. Poster: "The simplified TD-DFT approach for electronic spectra of very large systems", *International Symposium Chemistry at Spin Centers II*, **2014**, Bad Honnef, Germany.

8. Poster: "Broad range electronic spectra of large closed-shell and open-shell systems by simplified TD-DFT and TDA-DFT", *CECAM Workshop - Perspectives of many-particle methods: total energy, spectroscopy and time-dependent dynamics*, **2015**, Bremen, Germany.

9. Talk: "Simplified Time-Dependent DFT: Theory, Application and Recent Developments", *Max-Planck-Institut for Chemical Energy Conversion*, **2015**, Mülheim an der Ruhr, Germany.

10. Talk: "Computation of Electronic Circular Dichroism Spectra and Optical Rotation by Simplified Time-Dependent Density Functional Theory", *15^{th} International Conference on Chiroptical Spectroscopy*, **2015**, Sapporo, Japan.

11. Poster: "Electronic excitation spectra of closed- and open-shell systems employing the simplified time-dependent density functional theory approaches", *51^{st} Symposium on Theoretical Chemistry*, **2015**, Potsdam, Germany.

12. Talk: "Electronic absorption and circular dichroism spectra of large closed-shell and open-shell systems by simplified TD-DFT methods", *Graduate Talks on Chemistry@Spin Centers*, **2015**, Trier, Germany.

13. Poster: "Ultra-fast computation of electronic spectra for large systems by tight-binding based simplified Tamm-Dancoff approximation (sTDA-xTB)", *The 8^{th} Molecular Quantum Mechanics*, **2016**, Uppsala, Sweden.

14. Talk: "Berechnung elektronischer Anregungsspektren mittels vereinfachter TD-DFT Methoden", *GDCh Kolloquium*, **2016**, Bonn, Germany.

15. Poster: "Ultra-fast computation of electronic spectra for large systems by tight-binding based simplified Tamm-Dancoff approximation (sTDA-xTB)", *Computational Chemistry Gordon Research Conference*, **2016**, Girona, Spain.

16. Poster: "An Exciton Coupling Approach Based on Simplified TD-DFT", *52^{nd} Symposium on Theoretical Chemistry*, **2016**, Bochum, Germany.

17. Talk: "Simplified Methods for the Computation of Electronic Absorption and Circular Dichroism Spectra", *Interdisciplinary Center for Scientific Computing*, **2017**, Heidelberg, Germany.

18. Talk: "A robust and accurate tight-binding method for structures and energies of supramolecular complexes", *SupraChem*, **2017**, Aachen, Germany.

19. Talk: "A robust and accurate tight-binding method for structures and non-covalent interaction energies", *CUP XVII*, **2017**, Santa Fe, NM, USA.

20. Talk: "Simplified Methods for the Computation of Electronic Absorption and Circular Dichroism Spectra", *Karlsruhe Institute of Technology*, **2017**, Karlsruhe, Germany.

21. Talk and Poster: "Simplified Methods for the Computation of Electronic Absorption and Circular Dichroism Spectra", *254^{th} ACS National Meeting*, **2017**, Washington, DC, USA.

Abstract

In this thesis, an efficient procedure to compute electronic excitation spectra of molecular systems is presented, focusing particularly on the computation of electronic circular dichroism (ECD) spectra. ECD spectroscopy is commonly used to distinguish between the two enantiomers of a chiral compound. Due to a strong sensitivity to the three-dimensional structure, reliable simulation of ECD spectra of solvated molecules by quantum chemical methods requires the knowledge of the relevant conformers along with the corresponding ECD signals (i.e., the individual transition intensities and energies) and Boltzmann populations.

The latter point can be addressed by an established thermochemical protocol. It combines electronic energies computed in gas phase by dispersion-corrected density functional theory (DFT-D) with nuclear ro-vibrational and solvation contributions to yield the free energies in solution. This model is applied to study the association of two intermolecular frustrated Lewis pairs (FLPs). Though this case study does not aim at computing an ECD spectrum, it provides insight on whether such a scheme could also be suited to rank conformers in solution. Comparison to high-level reference methods and partially available experimental data suggests that the largest uncertainty can be attributed to the implicit solvation model. The errors for different dimer arrangements, however, appear to be within the order of 1 kcal mol^{-1}, which is encouraging for the pursued computation of conformer free energies. In combination with a quadruple-ζ basis set, hybrid DFT-D methods like the PW6B95-D3 are almost converged with respect to a complete basis and provide satisfactory results for the electronic energy contribution. Hence, they are recommended choices for the final electronic structure level to rank different conformers in routine calculations.

The major part of this thesis deals with the development and application of cost-efficient excited state methods. The current state-of-the-art to compute ECD spectra for systems with roughly 100 atoms is the time-dependent density functional theory (TD-DFT) approach. Based on the latter, the simplified TD-DFT (sTD-DFT) method is developed. The excited state treatment is accelerated by at least three orders of magnitude, resulting from semiempirically approximated two-electron integrals and a significant reduction of the involved matrix dimensions. The introduced approximations are in line with the ones in the previously presented simplified Tamm-Dancoff approximated TD-DFT (sTDA-DFT). It is shown that the sTD-DFT and the sTDA-DFT approaches provide roughly the same accuracy for vertical excitation energies, as well as absorption and ECD spectra, as their parental schemes, i.e., TD-DFT and Tamm-Dancoff approximated TD-DFT (TDA-DFT), respectively. Thus, sTD-DFT is an efficient approach that is suitable for the computation of ECD spectra.

Furthermore, sTD-DFT calculations conducted on "snapshots" from molecular dynamics (MD) simulations offer an appealing way to effectively incorporate vibronic effects without a quantum mechanical (QM) treatment of the nuclei. Such a treatment is exemplified for [16]helicene (102 atoms) and a di-substituted derivative (164 atoms).

While the feasibility of applying sTDA-DFT to very large systems is demonstrated for two palladium(II) metallosupramolecular spheres (822 and 1644 atoms, respectively), it is also shown that this method produces ECD spectra of incorrect sign in the origin-independent dipole velocity formalism for extended π-systems. This behavior is due to the Tamm-Dancoff approximation (TDA) and, therefore, it is also present in TDA-DFT and the related configuration interaction singles (CIS) approach. Based on the insights obtained from this study, the A+B/2 correction is developed, which corrects the (simplified) TDA eigenvectors affording origin-independent dipole velocity ECD spectra of roughly (s)TD-DFT quality, while retaining the lower computational cost of the (s)TDA excited state treatment. Combination with a newly developed, purpose-specific extended tight-binding procedure for the ground state yields the ultra-fast sTDA-xTB approach. Due to different adjustments of the atomic orbital basis and the tight-binding Hamiltonian, the method is on a par with TDA-PBE0/def2-SV(P) for vertical excitation energies. The entire computation of an ECD spectrum (< 9 eV) for [16]helicene is finished in 10 s with sTDA-xTB, while the analogous calculation with sTD-BHLYP/def2-SV(P) takes more than one hour. Along with the availability of the required parameters for most elements of the periodic table, the extremely low computational cost of this newly developed method allows routine calculation of spectra for large systems (with roughly 1000 atoms), even if many different conformers need to be considered.

The last part of this thesis reports on another purpose-specific extended tight-binding scheme, GFN-xTB, which provides molecular geometries, harmonic vibrational frequencies, and non-covalent interaction energies with comparable or better accuracy than existing semiempirical methods. Since parameters are available for all elements with $Z \leq 86$, the method offers great potential to sample the conformational space of almost arbitrary molecules with up to a few hundred atoms. In combination with the ultra-fast sTDA-xTB approach, ECD spectra can be computed in an almost "black box" manner, e.g., by computing spectra on MD snapshots. Together with the established thermochemistry protocol mentioned above, the newly developed architecture sets the stage for a fully automatic multi-level ECD procedure to be developed in the near future.

Zusammenfassung

Diese Dissertation stellt einen effizienten Ansatz zur Berechnung von elektronischen Anregungsspektren molekularer Systeme vor, wobei der besondere Fokus auf der Berechnung von elektronischen Circulardichroismus-(ECD-)Spektren liegt. Die ECD-Spektroskopie wird typischerweise verwendet, um zwischen den beiden Enantiomeren einer chiralen Verbindung zu unterscheiden. Aufgrund der hohen Sensibilität für die räumliche Struktur des Moleküls wird zur zuverlässigen Simulation von ECD-Spektren die Kenntnis der relevanten Konformere inklusive ihrer Boltzmann-Populationen und der jeweiligen ECD-Signale (d.h. deren energetische Lage und Intensitäten) benötigt.

Die Populationen können mithilfe eines literaturbekannten Thermochemieprotokolls unter Verwendung der dispersionskorrigierten Dichtefunktionaltheorie (DFT-D) näherungsweise berechnet werden. In der vorliegenden Arbeit wird dieses Modell verwendet, um die Komplexbildung von zwei intermolekularen frustrierten Lewispaaren (FLPs) zu untersuchen. Obwohl diese Fallstudie keine Berechnung eines ECD-Spektrums zum Ziel hat, geben die gewonnenen Erkenntnisse durchaus Aufschluss darüber, ob sich der gewählte Ansatz auch dazu eignet, die Populationen verschiedener Konformere zu bestimmen. Der Vergleich mit hochwertigen Vergleichsrechnungen auf der einen und mit zum Teil verfügbaren experimentellen Daten auf der anderen Seite legt nahe, dass der größte Unsicherheitsfaktor in den Solvatationsbeiträgen vorliegt, welche mithilfe eines impliziten Lösungsmittelmodells bestimmt werden. Allerdings liegen deren geschätzte Fehler für unterschiedliche räumliche Anordnungen des Komplexes, d.h. bei einer gleichbleibenden Systemgröße von ca. 50–100 Atomen, lediglich bei etwa 1 kcal mol^{-1}. Für die Berechnung von freien konformellen Enthalpien ist mit ähnlich großen Fehlern zu rechnen. Kombiniert mit Quadruple-ζ-Basissätzen weisen Hybrid-DFT-Methoden bereits nahezu konvergierte elektronische Energien auf und können bei gleichzeitiger Verwendung einer Dispersionskorrektur relativ genaue Gasphasenenergiebeiträge (so z.B. PW6B95-D3) zu den freien Enthalpien in Lösung beitragen.

Der Großteil dieser Dissertation beschäftigt sich mit der Entwicklung und Anwendung von kosteneffizienten Methoden zur Berechnung angeregter Zustände. Die gegenwärtig am häufigsten verwendete Methode zur Berechnung von ECD-Spektren ist die zeitabhängige Dichtefunktionaltheorie (TD-DFT). Von dieser ausgehend wird die vereinfachte TD-DFT Methode (sTD-DFT) entwickelt. Aufgrund der semiempirischen Näherung der Zweielektronenintegrale und der deutlichen Reduzierung der relevanten Matrixdimensionen wird die Berechnung der angeregten Zustände um mindestens drei Größenordnungen beschleunigt. Diese Näherungen sind konsistent zu jenen, die bereits in dem vereinfachten Tamm-Dancoff-genäherten TD-DFT (sTDA-DFT)

Ansatz eigeführt wurden.

Im Vergleich zu den Ausgangsmethoden, also TD-DFT und seiner Tamm-Dancoff-Näherung (TDA-DFT), ist weder eine signifikante Beeinträchtigung der senkrechten Anregungsenergien noch eine Verschlechterung der Absorptions- und ECD-Intensitäten bemerkbar. Insbesondere die sTD-DFT Methode eignet sich zur effizienten und zuverlässigen Berechnung von ECD-Spektren. Die Effizienz der sTD-DFT Methode ermöglicht unter anderem die Berechnung von Spektren auf Nichtminimumsstrukturen, die aus einer Molekulardynamik-(MD)-Simulation stammen. Somit können vibronische Effekte näherungsweise erfasst werden, ohne dass ein quantenmechanischer (QM) Ansatz für die Kerne verwendet werden muss. Exemplarisch wird dieses Verfahren für das [16]Helicen (102 Atome) und einem disubstituierten Derivat (164 Atome) angewandt.

Die Anwendbarkeit der sTDA-DFT Methode auf sehr große Systeme wird am Beispiel von zwei Palladium(II)-metallosupramolekularen Komplexen (822 und 1644 Atome) verdeutlicht, doch zeigt eine weitere Studie, dass Tamm-Dancoff-genäherte (TDA) Methoden für die ECD Spektren von ausgedehnten, delokalisierten π-Systemen im Impulsformalismus das falsche Vorzeichen liefern. Gleiches gilt für den verwandten Konfigurationswechselwirkungs-Ansatz mit Einfachanregungen (CIS). Basierend auf den Erkenntnissen dieser Studie ist es gelungen, die sogenannte A+B/2-Näherung zu entwickeln, welche die entsprechenden Fehler in den TDA Eigenvektoren behebt, ohne die Kosten der Methode sichtlich zu erhöhen. Durch die Kombination des so korrigierten vereinfachten TDA-Ansatzes mit einer speziell optimierten semiempirischen Tight-Binding-Methode für den Grundzustand wird die äußerst schnelle sTDA-xTB-Methode erhalten. Aufgrund verschiedener Modifikationen der Atomorbitalbasis und des Tight-Binding-Potentials erreicht diese Methode eine ähnliche Genauigkeit für senkrechte Anregungsenergien wie z.B. eine DFT-basierende Rechnung auf TDA-PBE0/def2-SV(P) Niveau. Die beachtliche Effizienz der Methode wird im Vergleich zum bereits effizienten sTD-BHLYP/def2-SV(P) Ansatz für das [16]Helicen (alle Anregungen bis 9 eV) deutlich: Während letzterer Ansatz etwas mehr als eine Stunde Rechenzeit benötigt, ist das ECD-Spektrum mit sTDA-xTB bereits nach 10 s verfügbar. Da die Parametrisierung nahezu das gesamte Periodensystem abdeckt, werden Standardrechnungen von Spektren großer Systeme (mit ca. 1000 Atomen) ermöglicht, selbst wenn mehrere Konformere berücksichtigt werden.

Im letzten Teil der Arbeit wird eine weitere spezialisierte Tight-Binding-Methode vorgestellt (GFN-xTB), die wiederum auf die Berechnung von Geometrien, harmonischen Frequenzen und nichtkovalenten Wechselwirkungen ausgelegt ist und hierfür bessere Ergebnisse liefert als vergleichbare semiempirische Methoden. Die Verfügbarkeit von Parametern für alle Elemente mit $Z \leq 86$ ermöglicht das Absuchen des konformellen Raums für unterschiedliche Systeme mit wenigen hundert Atomen. Zusammen mit sTDA-xTB sind in kürzester Zeit Berechnungen von Sprektren z.B. entlang von MD-Trajektorien möglich. Vereint mit den bereits existierenden Thermochemieprotokollen sind somit die ersten Voraussetzungen für eine völlig automatische Prozedur zur Berechnung von ECD-Spektren geschaffen worden.

Contents

Publications vii

Abstract xiii

Zusammenfassung xv

I. Introduction and Theoretical Background 1

1. Introduction 3

2. Theoretical Background 9

2.1. Methodological Overview . . . 9
2.2. Electronic Structure Methods . . . 10
2.2.1. Definition of the Electronic Hamiltonian . . . 10
2.2.2. Rayleigh-Schrödinger Perturbation Theory . . . 12
2.2.3. Hartree-Fock Theory and Electron Correlation . . . 14
2.2.4. Kohn-Sham Density Functional Theory . . . 17
2.2.5. Roothaan-Hall Formalism . . . 22
2.2.6. Linear Response Theory . . . 24

II. Dispersion-Corrected Density Functional Theory in Free Energy Calculations 31

3. The Association of Two "Frustrated" Lewis Pairs 37

3.1. Introduction . . . 38
3.2. Theoretical Background . . . 39
3.3. Technical Details of the Calculations . . . 40
3.3.1. Geometry Optimizations . . . 40
3.3.2. Single-Point Energy Calculations . . . 40
3.3.3. Ro-Vibrational Contributions . . . 42
3.3.4. Solvent Contributions . . . 43

3.4. Results and Discussion . . . 43
3.4.1. Association of $B(C_6F_5)_3PMes_3$ . . . 43
3.4.2. Comparing the Association of $B(C_6F_5)_3PMes_3$ and $B(C_6F_5)_3PtBu_3$ . . . 45
3.4.3. Structural Implication for the Reactivity . . . 48
3.5. Conclusion . . . 48

III. Development and Application of Simplified Time-Dependent Density Functional Theory 51

4. Enantiomerically Pure $[M_6L_{12}]$ or $[M_{12}L_{24}]$ Polyhedra 55
4.1. Introduction . . . 56
4.2. Methodological Overview . . . 57
4.3. Results and Discussion . . . 58
4.4. Conclusion . . . 64

5. Simplified Time-Dependent Density Functional Theory 65
5.1. Introduction . . . 66
5.2. Theory . . . 68
5.2.1. TD-DFT and TDA-DFT . . . 68
5.2.2. The sTDA-DFT Approach . . . 69
5.2.3. Transition Dipole Moments, Oscillator and Rotatory Strengths . . . 71
5.2.4. The Simplified Time-Dependent DFT (sTD-DFT) Approach . . . 72
5.3. Technical Details of the Calculations . . . 73
5.4. Results and Discussion . . . 74
5.4.1. Rotatory Strengths from Conventional and Simplified TDA-DFT and TD-DFT in the Biphenyl Molecule . . . 74
5.4.2. Comparison of Vertical Excitation Energies in sTDA-DFT and sTD-DFT 76
5.4.3. UV Spectrum of Indigo . . . 78
5.4.4. ECD Spectrum of (*P*)-[11]Helicene . . . 79
5.4.5. ECD Spectrum of $(S,S,^{f,s}A)$-C_{60} . . . 80
5.5. Conclusion . . . 82

6. Electronic Circular Dichroism of Highly Conjugated π-Systems 83
6.1. Introduction . . . 84
6.1.1. Suitable Methods to Compute Full Range ECD Spectra for Large Systems 85
6.1.2. Systems under Consideration . . . 86
6.2. Theoretical Background . . . 87
6.2.1. Time-Dependent DFT and HF . . . 87
6.2.2. Simplified TDA-DFT and TD-DFT . . . 89
6.2.3. Rotatory Strengths: Length and Velocity Form . . . 89

6.3. Computational Details . . . 90
6.4. Results and Discussion . . . 92
6.4.1. The (^{s,f}A)-C_{76} Fullerene . . . 92
6.4.2. The Carbon Nanotube Model . . . 95
6.4.3. (*P*)-[11]Helicene . . . 95
6.4.4. α-Helical Ace-$(Ala)_{19}$-NMe . . . 97
6.5. Conclusion . . . 99

7. The Electronic Circular Dichroism of [16]Helicene with Simplified TD-DFT 101
7.1. Introduction . . . 102
7.2. Results and Discussion . . . 104
7.3. Conclusion . . . 109

IV. Development and Assessment of Purpose-Specific Tight-Binding Based Approaches 111

8. Ultra-Fast Spectra by Tight-Binding Based Simplified Tamm-Dancoff Approximation 115
8.1. Introduction . . . 116
8.2. Theory . . . 119
8.2.1. TD-DFT and TDA-DFT . . . 119
8.2.2. The Simplified TDA-DFT and TD-DFT Approaches . . . 120
8.2.3. Correction for Monopole Approximation Errors in sTDA/sTD-DFT . . . 122
8.2.4. A+B/2-Approximation in TDA for Rotatory Strengths . . . 123
8.2.5. The xTB Composite Procedure . . . 125
8.3. Results and Discussion . . . 134
8.3.1. Charges, Spin Populations, and Bond Orders from the VTB Hamiltonian 134
8.3.2. Excitation Energies for the Fit Sets . . . 136
8.3.3. Rydberg States . . . 137
8.3.4. Cross-Checking of Excitation Energies and Comparison to TD-DFT and Other Semi-Empirical Methods . . . 138
8.3.5. Test of the A+B/2 Vector Approximation for ECD and Optical Rotations 140
8.3.6. UV and ECD Spectra in Comparison to Experiment . . . 142
8.4. Conclusion . . . 148

9. A Tight-Binding Method for Structures, Frequencies, and Non-Covalent Interactions 151
9.1. Introduction . . . 152
9.2. Theory . . . 154
9.2.1. The GFN-xTB Hamiltonian . . . 154
9.2.2. Parametrization and Technical Details . . . 159

9.3. Results and Discussion . . . 162
9.3.1. Structures . . . 162
9.3.2. Non-Covalent Interactions . . . 166
9.3.3. Conformational Energies . . . 175
9.3.4. Potential Energy Curves, Thermochemistry, and Other Properties . . . 178
9.3.5. Known Problems . . . 183
9.3.6. Performance for Large Systems . . . 185
9.4. Conclusion . . . 187

V. Final Summary and Conclusion **189**

Bibliography **197**

VI. Appendix **221**

A1. Supporting Information to Chapter 2 **222**

A2. Supporting Information to Chapter 3 **225**

A3. Supporting Information to Chapter 4 **227**

A4. Supporting Information to Chapter 6 **230**

A5. Supporting Information to Chapter 7 **231**

A6. Supporting Information to Chapter 8 **235**

A7. Supporting Information to Chapter 9 **239**

A8. Supporting Information to Part V **260**

Acknowledgments **261**

Part I.

Introduction and Theoretical Background

1. Introduction

The main purpose of theoretical models in chemistry is to rationalize experimental findings,[1–3] i.e., to explain the reactivity or spectroscopic properties of compounds. Quantum mechanical (QM) models are often employed to complement experimental studies but can also provide predictive power and guide experimental work.[1,3–7] Along with the availability of more powerful computers,[8] the development of efficient methods has established computational chemistry as a fundamental part of chemistry.

Knowledge about the three-dimensional structure of molecular compounds remains an important task, since the geometry is directly linked to the chemical and physical properties of the respective system.[9] Experimentally, the three-dimensional structure information, e.g., of organic and biomolecular compounds is typically obtained from X-ray crystallography of the corresponding molecular crystal. The latter are often not accessible experimentally and, furthermore, it remains questionable whether the conformation in the crystalline sample is also representative for the molecular geometry adopted in solution.[10] In that case, simulations can be used to determine the three-dimensional structure of the respective compound. In particular, accurate structure simulation of large biomolecules like proteins still remains one of the "Holy Grails" of computational chemistry.[1] When attempting to simulate chemical systems in solution, it is important to have some kind of feedback between the simulation and the experiment in order to assess the employed level of approximation in the calculation. More specifically, the following questions need to be answered: does the simulation reflect the experimental conditions properly and is the suggested structure in solution meaningful? The necessary feedback is offered by spectroscopic methods, ideally the combined use of different spectroscopic techniques. Therefore, based on the simulated structures, the spectroscopic quantities are computed and then compared to the experimentally recorded spectra.

A spectroscopic method that is particularly sensitive to the three-dimensional arrangement is electronic circular dichroism (ECD) spectroscopy.[11,12] ECD refers to the differential absorption of left- and right-circularly polarized light in the visible to ultraviolet energy (UV-Vis) range (see Figure 1.1 for schematic depiction).[11] As in UV-Vis absorption spectroscopy, the system undergoes an electronic transition to an excited electronic state upon absorbing the energy of the incoming photon. Different from UV-Vis absorption, ECD is exhibited only by chiral compounds* and ECD spectroscopy is in frequent use to assign their absolute configuration.[11,12,14–19]

Furthermore, supramolecular structures can also be studied by ECD spectroscopy, provided

*Chiral molecules possess no improper rotation axis. A chiral molecule and its mirror image form a pair of enantiomers. Many pharmaceuticals are chiral, and even though both enantiomers have identical chemical composition and bond distances, their physiological and toxicological properties can be tremendously different.[13]

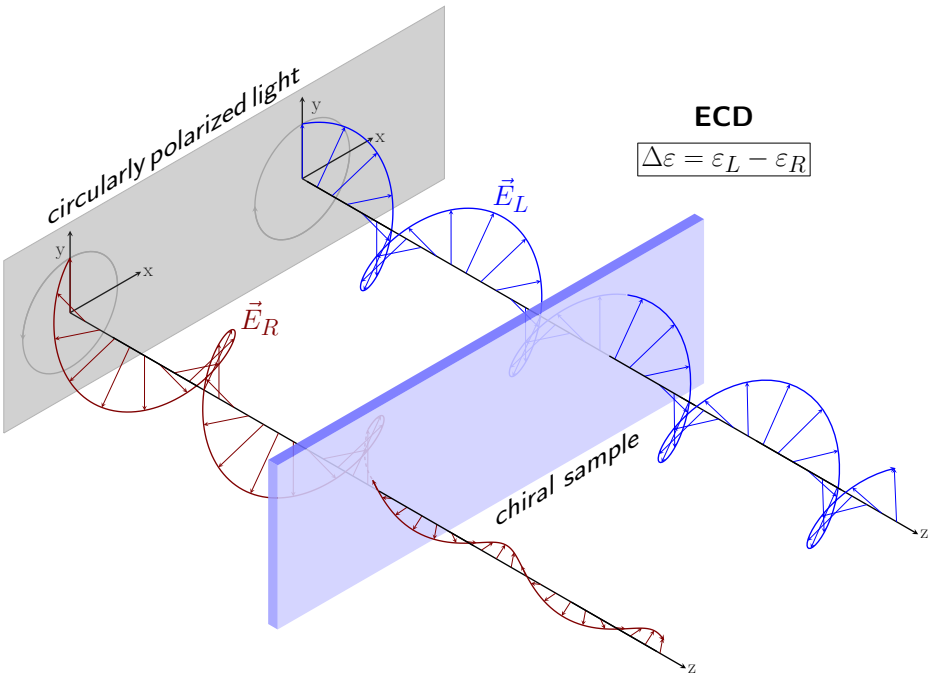


Figure 1.1.: Schematic depiction of electronic circular dichroism (ECD). In this example, the chiral sample exhibits stronger absorption of right- than of left-circularly polarized light (drawn in red and blue, respectively), thus, affording a negative ECD signal at this particular wavelength.

that they comprise chiral components and absorb light in the UV-Vis spectral region. A prominent example from supramolecular biochemistry are proteins, which exhibit different ECD spectra depending on their secondary structure.[20–23] In this context, ECD spectroscopy is very sensitive to changes in the polypeptide backbone structure, e.g., induced by mutations or the formation of complexes.[24] Folding processes can also be monitored by ECD spectroscopy.[25–27]

Since the differential absorption of light occurs instantaneously, the resulting ECD signal can be measured directly. This, along with the high sensitivity to structural changes, has triggered the development of chiral sensoring materials in recent years.[18,28–31] The chiral sensor consists of a non-chiral chromophore that can form a non-covalent complex with a chiral species. Since the formed assembly is also chiral, a so-called induced ECD signal of the chromophore can be recorded. In principle, this allows real-time monitoring of chemical reactions involving chiral compounds, even if the latter themselves do not exhibit strong ECD in the UV-Vis spectral range.

The theory for the simulation of ECD spectra can be divided in three parts, as essentially pointed out in 1975 by Schellman:[32]

1. The connection between the experimentally observed macroscopic ECD and the quantum mechanical quantity needs to be established.

2. Feasible electronic structure methods that provide the QM quantity need to be developed.

3. ECD signals are highly sensitive to the three-dimensional structure of the molecule. Hence, the latter needs to be computed accurately. Possibly, different conformers have to be considered and the individual ECD spectra need to be averaged.

The first problem was essentially solved at that time, i.e., for ECD in the UV-Vis spectral region the rotatory strength R is required, which is computed from electric and magnetic transition dipole moments.[32,33] The other points can be addressed sufficiently well for small and rigid molecules, however, they remain non-trivial until today for many chemical systems, due to the size and complexity of the latter. In particular the aforementioned supramolecular systems pose a huge challenge for theory.

The focus of this thesis is the development and assessment of methods to address the third and in particular the second point mentioned above. Given that ECD involves electronic transitions, a QM treatment is required for their computation.[11] Depending on the size and electronic structure of the considered molecule, a large number of excited states has to be computed to obtain a reasonable electronic spectrum. Highly accurate methods based on correlated wave function theory are restricted to systems containing about 10–15 non-hydrogen atoms, due to the unfavorable scaling of the computational demands with the system size.[11,14,15] To treat larger systems, more approximate methods need to be employed. In particular, the advent of time-dependent density functional theory (TD-DFT) has allowed the routine calculation of ECD spectra for small to moderately sized systems ($<$ 100 atoms) with reasonable accuracy.[11,16,17,19,34] This enabled an increased synergy between theory and experiment in the past 10–20 years.[11,17,19]

If flexible systems are studied, more than just a single structure needs to be considered in the excited state calculation. This becomes prohibitive even with TD-DFT, especially if a full electronic spectrum in the UV-Vis range is to be computed. Meanwhile, the availability of intense synchrotron radiation sources has enabled the measurement of ECD spectra down to about 160 nm (7.75 eV) also for large systems like proteins.[24,26] In order to benefit from this progress in spectroscopy and to be able to study realistic flexible systems (e.g., chiral sensor systems), the development of fast excited state methods is necessary. The mostly sufficient accuracy of TD-DFT[34,35] should be retained, as well as the general applicability to systems containing arbitrary elements. The latter is an issue in so-called semiempirical QM (SQM) methods due to the use of element-specific or even element pair-specific parameters.[36–38]

Furthermore, the structures to be considered in the excited state treatment have to be determined (see also Figure 1.2). In order to sample the conformational space of chemical and biochemical systems, often SQM methods[38,39] or classical force-fields are employed.[40,41] Though being significantly faster than *ab initio* QM methods, the accuracy of SQMs and force-fields is much lower and lacking element parameters can preclude their use. Having obtained an ensemble of structures, these need to be ranked by their free energy, which gives rise to the Boltzmann distribution. From the respective probabilities, the relevant conformers are then determined. The applied protocol needs to be efficient, yet, accurate, because small errors in the conformational free energies can have severe effects on the Boltzmann distribution.

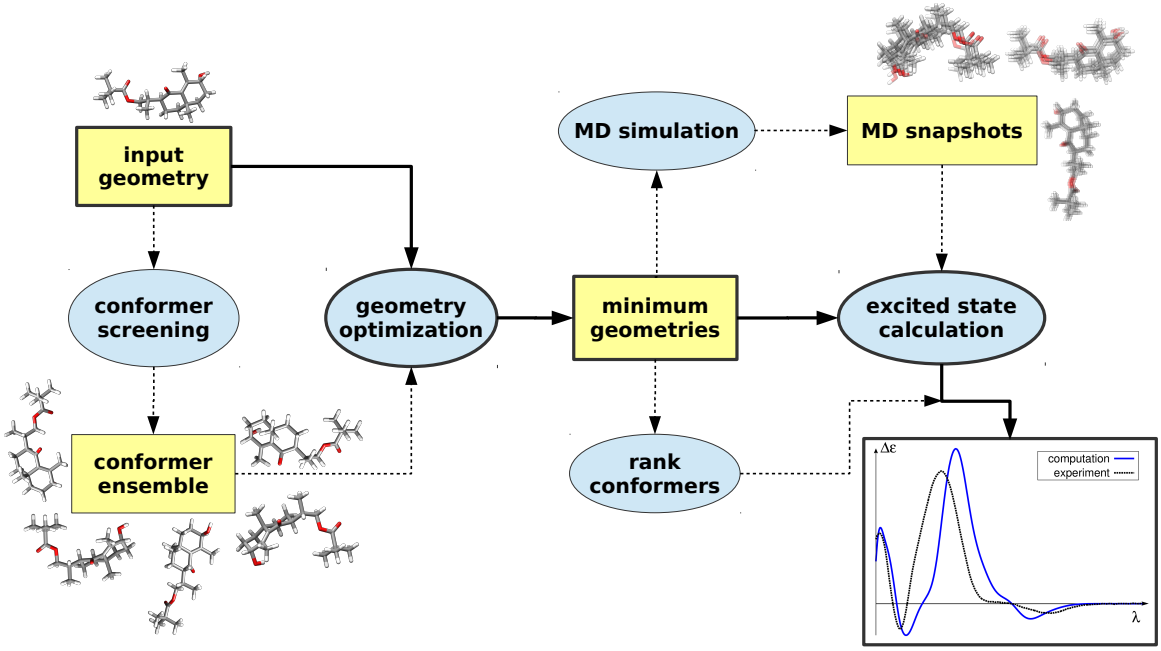


Figure 1.2.: Flowchart for the computation of ECD spectra. The central steps (indicated by bold lines) include obtaining and optimizing the relevant structures and the computation of the ECD spectrum. While these steps are sufficient for rigid molecules, conformer sampling and/or performing molecular dynamics (MD) simulations is required for flexible systems. This rather elaborate task can be performed with fast classical force-field or semiempirical quantum mechanical methods. The final structure optimization, the free energy ranking, and the excited state calculations typically require a sophisticated semiempirical QM or an *ab initio* QM approach.

The overall procedure to face the challenge of computing accurate ECD spectra in solution is schematically depicted in Figure 1.2.

In the upcoming chapter of the present part, an overview of well-established QM methods, which are relevant for the present work, is given. Part II deals with dispersion-corrected density functional theory and its ability to describe free energies in solution in combination with a multi-level scheme that includes a solvation and thermostatistical treatment.[42] This scheme is employed in Chapter 3 to describe the association of two intermolecular frustrated Lewis pair (FLP) systems in solution. FLPs have caught attention in the past ten years due to their catalytic ability to activate small molecules like H_2.[43,44] The investigated intermolecular FLP systems share the same type of Lewis acid, namely $B(C_6F_5)_3$, which then associates with two comparably similar Lewis bases, i.e., trimesitylphosphine ($PMes_3$) and tri-*tert*-butylphosphine ($PtBu_3$), respectively. For the former, an association free energy estimated from nuclear magnetic resonance (NMR) experiments is available,[45] which is used to assess the employed method-

ology. The study highlights the importance of London dispersion interactions and their efficient treatment by means of computationally simple dispersion corrections. Since the investigated systems are non-chiral, the particular study does not aim for the subsequent computation of ECD spectra. Nevertheless, the same procedure can, e.g., be applied to model the aforementioned chiral sensors and also to rank different conformers of flexible systems.

Part III is devoted to the development and application of simplified variants of TD-DFT. In Chapter 4, the previously developed simplified Tamm-Dancoff approximated TD-DFT (sTDA-DFT) method[46] is applied in a collaborative work to compute the ECD spectra of two large metallosupramolecular complexes containing 822 and 1644 atoms, respectively. These complexes are synthesized by the Lützen group at the University of Bonn. The ECD spectra simulated this way serve as an additional proof for the existence of the self-assembled structures in solution. The sTDA-DFT approach incorporates drastic approximations, which accelerate the excited state calculation by several orders of magnitude, enabling this simulation in the first place. Though the two-electron integrals occurring in the excited state treatment are treated in a semiempirical monopole approximation, this method was shown to mostly retain the accuracy of the parental approach, i.e., Tamm-Dancoff approximated TD-DFT (TDA-DFT).[46]

Chapter 5 presents the development, initial implementation and assessment of the simplified TD-DFT (sTD-DFT) method. The approximations leading from TDA-DFT to sTDA-DFT are transferred to the TD-DFT framework. While being computationally slightly more demanding than sTDA-DFT, the method is several orders of magnitude faster than regular TD-DFT, but produces electronic spectra of similar quality. In fact, it is apparent that the Tamm-Dancoff approximation (TDA) has a more severe effect than the integral approximations in sTD(A)-DFT. It is well-known that TDA methods are gauge-variant and less reliable for ECD spectra.[47]

Chapter 6 follows up on this topic: the dramatic failure of TDA methods in the computation of dipole-velocity rotatory strengths for delocalized π-systems is demonstrated. While initially observed for sTDA-DFT, it turns out that this problem is more fundamental and present in all TDA based approaches. By contrast, the sTD-DFT method does not exhibit this problem. In the course of this work, the sTD-DFT approach is implemented into the ORCA program package,[48] and furthermore, sTD-DFT is combined with a range-separated hybrid functional.

Chapter 7 presents an sTD-DFT study of the absorption and ECD spectra of [16]helicene and a di-substituted triisopropylsilyloxy (TIPSO) derivative. This helicene derivative is the largest helicene to be synthesized so far and the computed spectra are compared to the experimentally available data. As a proof-of-principle, we go beyond the minimum structure approximation by running force-field[49] based molecular dynamics (MD) simulations including explicit solvent molecules and periodic boundary conditions. sTD-DFT is then applied on 100 structures to compute MD averaged spectra for these fairly large systems (102 and 164 atoms, respectively). It turns out that a major bottleneck of this procedure is the ground state calculation that precedes the sTD-DFT treatment.

Part IV deals with the development of two purpose-specific SQM methods, more precisely extended tight-binding schemes. Both approaches have in common that a purely element-specific

parametrization strategy is employed. This way, empirical parameters for the largest part of the periodic table can be provided. The first approach presented in Chapter 8 aims at circumventing the current computational bottleneck of sTDA-DFT/sTD-DFT treatments. Instead of a Kohn-Sham density functional ground state, the orbital information from the tight-binding procedure is now used in a subsequent simplified TDA treatment, leading to the sTDA-xTB method. In this study, a correction scheme for the otherwise flawed dipole-velocity rotatory strengths from TDA treatments is developed. This faster excited state treatment in combination with the semiempirically approximated ground state treatment now enables the ultra-fast calculation of electronic spectra including ECD. In Chapter 9, a related semiempirical tight-binding method (GFN-xTB) is presented, which is designed to provide good molecular geometries and reasonable non-covalent interaction energies. This approach provides excellent prospects to be used in multi-level schemes to scan the conformational space of almost arbitrary molecules. The performance of both purpose-specific tight-binding methods is assessed on benchmark sets. In both studies, a superior performance compared to other SQM methods for the respective purpose is observed.

2. Theoretical Background

2.1. Methodological Overview

Spectroscopic measurements are mostly carried out in solution at room temperature. Given that solute concentrations of several mmol L^{-1} are common, even the simulation of all solute molecules in a 1 mL droplet would require to consider about 10^{17} molecules (excluding solvent molecules). Instead of simulating this unmanageable macroscopic system explicitly, one assumes the individual solute molecules to be non-interacting and quantum chemical treatments mostly focus on a single solute molecule (or molecular complex) that is stable under the ambient conditions. The investigation might be restricted to a single molecular geometry or to a set of representative conformations. Based on the assumption that each conformation is clearly separated from the others by an energetic barrier, the considered set of structures represents a microstate ensemble of the thermodynamic system. With the aid of statisticial mechanics, quantum chemical methods can then be used to determine the population of these microstates under the ambient conditions.[50] This is equivalent to knowing the structures that are representative for the ensemble of solute molecules. Observable properties, such as spectra, can be computed for the set of representative structures, which, along with the corresponding microstate populations, give rise to the respective properties of the macroscopic system.

Even though the corresponding experiment is mostly conducted in solution, quantum chemical methods generally treat molecules in gas phase. The importance of the neglected solvation effects on energies and properties can be quite different and depends on the solvent and the solute. A commonly used approach in quantum chemistry is to employ implicit solvation models that incorporate the electrostatic screening effects due to solvation. Furthermore, such approaches can straightforwardly be augmented with empirical corrections, which describe the free energy changes associated with the replacement of a volume fraction of solvent by the solute. The theory of these methods is beyond the scope of this work and an extensive review is given in Ref. 51. Of importance for the present work is the fact that implicit solvation models offer an efficient way to consider changes in the electronic structure due to solvation and, furthermore, to efficiently compute (relative) free energies in solution if combined with an electronic structure approach and a thermostatistical treatment.[42] Such a treatment is employed in Chapter 3 and may also be used for the (free) energetic ranking of different conformers.

In principle, the *ab initio* quantum mechanical (QM) methods, which are summarized in Section 2.2, could be used to scan the conformational space. However, such QM calculations are typically too demanding for chemically interesting systems and one needs to resort to signifi-

cantly faster semiempirical quantum mechanical (SQM) methods or classical force-fields for that purpose. A plethora of different SQMs and force-fields exist, but these are not discussed here. A comprehensive review on SQM methods can be found in Ref. 38. It is however relevant to note that the applicability of SQMs and force-fields is hampered by their lower accuracy compared to *ab initio* calculations and, furthermore, by the limited number of parametrized elements. In Chapter 9, a new SQM method is presented, which is designed to provide good molecular geometries and is parametrized for a large part of the periodic table. This approach provides excellent prospects to be used in multi-level schemes to scan the conformational space of almost arbitrary molecules.

In the next section, an overview of quantum mechanical electronic structure methods is given. The discussed methods include approaches that can be applied in the computation of energies and structures, as well as methods to compute properties like electronic absorption and ECD spectra.

2.2. Electronic Structure Methods

2.2.1. Definition of the Electronic Hamiltonian

Molecular systems consist of bonded atoms, and therefore comprise negatively charged electrons and positively charged nuclei. Indices e and n refer to electrons and nuclei, respectively, in the following. The relevant interaction in such a system is the Coulomb interaction.* In the Born-Oppenheimer approximation the nuclei are described as classical particles, while the significantly lighter electrons are treated as quantum mechanical particles. In the absence of an external potential, the non-relativistic electronic Hamiltonian for the electronic system is

$$\hat{H} \equiv \hat{H}_e = \hat{T}_e + \hat{V}_{ee} + \hat{V}_{ne}\,. \tag{2.1}$$

$\hat{T}_e$ is the electronic kinetic energy operator, while $\hat{V}_{ee}$ and $\hat{V}_{ne}$ denote the Coulomb operator for the electron-electron and electron-nucleus interactions, respectively (see Appendix A1 for the mathematical definition of the operators). If not noted otherwise, atomic units are used throughout. The total energy of the system is obtained by adding the classical Coulomb interaction between the nuclei E_{nn}, which is a constant for a given geometry, to the electronic energy.

While for systems comprising heavier elements relativistic effects of the core electrons become important,[50] such effects are typically small for the valence electrons. The latter play the central role in chemistry, i.e., in low-energy (<8 eV) electronic absorption spectroscopy, as well as in chemical bonding and chemical reactivity. Unless one is interested in properties of the core

*Due to the small mass of the involved particles, gravitational interactions are significantly smaller in magnitude compared to the Coulomb interactions (e.g., for two protons by roughly a factor of 10^{-36}). The strong and weak forces are important within the nuclei themselves, but are too short-ranged ($< 10^{-15}$ m)[50] to be relevant for the interactions within a molecule (cf. the Bohr radius a_0 corresponding to the most probable electron-proton separation in the hydrogen atom has a value of $5.2918 \cdot 10^{-11}$ m).[52]

electrons, the following procedure has become common practice for heavier elements (atomic number $Z > 36$):[53] the core electrons are removed from the quantum mechanical treatment, while the valence electrons are described by the non-relativistic electronic Hamiltonian (see Eq. 2.1). The non-negligible electrostatic and exchange (*vide infra*) interactions between core and valence electrons are then absorbed in an effective core potential (ECP) acting on the valence electrons.

For a time-independent Hamiltonian $\hat{H}$ the electronic system is described in the Schrödinger picture of quantum mechanics[54] as a time-dependent (TD) wave $\Psi_K(t)$ that is given by

$$\Psi_K(t) = \Psi_K e^{-iE_K t} \,. \tag{2.2}$$

Here, the phase factor $e^{-iE_K t}$ refers to the TD part of the wave function, while the amplitude of the wave Ψ_K corresponds to a particular stationary electronic state K.[50] E_K corresponds to the energy of the wave function in state K (see below), while t is the time. Ψ_K is independent of time but depends on the spatial and spin coordinates of the electrons, which are omitted here for clarity. The time evolution of the electrons is given by the TD electronic Schrödinger equation (SE)[50,52]

$$\hat{H}\Psi_K(t) = i\frac{\partial \Psi_K(t)}{\partial t} \,. \tag{2.3}$$

Since $\hat{H}$ is time-independent, the TD part of the wave function vanishes for stationary states.[†] This way, the time-independent Schrödinger equation is obtained

$$\hat{H}\Psi_K = E_K \Psi_K \,. \tag{2.4}$$

E_K is the electronic energy of the stationary electronic state K with corresponding wave function Ψ_K. These are the eigenvalue and eigenfunction, respectively, of the non-relativistic electronic Hamiltonian $\hat{H}$. This yields the electronic energy as the expectation value of the electronic Hamilton operator

$$\int \Psi_K^* \hat{H} \Psi_K \mathrm{d}\mathbf{a} \equiv \left\langle \Psi_K \middle| \hat{H} \middle| \Psi_K \right\rangle \equiv H_{KK} = E_K \left\langle \Psi_K \mid \Psi_K \right\rangle = E_K \,. \tag{2.5}$$

The integration on the left hand side of Eq. 2.5 is done over all variables (condensed within the variable $\mathbf{a}$). Here, the short hand notation for expectation values is introduced (see also Appendix A1). If the functions on the left and the right hand side of the operator are different, this integral is often called matrix element.[52] We adapt a notation that assumes the integration to be performed over the entire space ($-\infty$ to ∞) if not noted differently. The overlap integral on the right hand side of Eq. 2.5 vanishes, since the wave functions Ψ_K are orthonormalized,

[†]The electronic system is still oscillatory in the stationary state K. However, the electronic state, i.e., the corresponding wave amplitude Ψ_K and all observables associated with it remain the same.

i.e.,

$$\langle \Psi_K \mid \Psi_L \rangle = \delta_{KL} \,. \tag{2.6}$$

The Kronecker delta δ_{KL} gives values of one and zero if $K = L$ and $K \neq L$, respectively. The exact eigenfunctions to the electronic Hamiltonian $\hat{H}$ are typically not known *a priori*, but the variational principle may be exploited

$$\tilde{E}_K = \frac{\left\langle \tilde{\Psi}_K \middle| \hat{H} \middle| \tilde{\Psi}_K \right\rangle}{\left\langle \tilde{\Psi}_K \middle| \tilde{\Psi}_K \right\rangle} \geq \left\langle \Psi_K \middle| \hat{H} \middle| \Psi_K \right\rangle = E_K \,. \tag{2.7}$$

This principle states that any trial wave function $\tilde{\Psi}_K$ will result in an energy expectation value $\tilde{E}_K$ that is higher than the energy E_K of the true wave function Ψ_K.

2.2.2. Rayleigh-Schrödinger Perturbation Theory

Since the solution of the SE for a fully interacting many-electron system is not accomplishable in practice, it is important to employ approximate treatments for the many-body problem. Though not restricted to this problem, it is convenient to introduce the concept of perturbation theory (PT) for this purpose. Here, the full Hamiltonian $\hat{H}$ is expressed as[55,56]

$$\hat{H} = \hat{H}_0 + \lambda \hat{H}' \,. \tag{2.8}$$

$\hat{H}_0$ refers to an effective zeroth order Hamiltonian with eigenfunctions $\Phi^{(0)}$. $\hat{H}'$ is the *perturbation operator* and λ the *perturbation parameter*. λ may have a physical meaning, but is typically a parameter ranging from zero (unperturbed case) to one (perturbed case). The basic idea followed in Rayleigh-Schrödinger perturbation theory (RS-PT) is that the dominant part of the total energy is already obtained by solving the time-independent zeroth order problem. Based on this assumption, the following procedure is applied:[50,56]

- A feasible time-independent zeroth order problem $\hat{H}_0 \Phi_K^{(0)} = E_K^{(0)} \Phi_K^{(0)}$ is set up and solved instead of the non-feasible $\hat{H} \Psi_K = E_K \Psi_K$.
- The neglected contributions in $\hat{H}_0$ (with respect to $\hat{H}$) are transferred to $\hat{H}'$, i.e., $\hat{H}' = \hat{H} - \hat{H}_0$.
- The energy E_K and wave function Ψ_K of the fully interacting system are expressed in a power expansion around the unperturbed solution, i.e., at $\lambda = 0$:

$$E_K = \sum_{n=0}^{\infty} \lambda^n E_K^{(n)} \tag{2.9}$$

$$\Psi_K = \sum_{n=0}^{\infty} \lambda^n \Phi_K^{(n)} \tag{2.10}$$

- Within the convergence radius of these expansions, Eqs. 2.9 and 2.10 can be applied in the time-independent SE (Eq. 2.4) in combination with the Hamiltonian in Eq. 2.8.
- Equating terms of identical order of λ yields (shown up to second order):

$$\lambda^0: \left(\hat{H}_0 - E_K^{(0)}\right)\Phi_K^{(0)} = 0 \tag{2.11a}$$

$$\lambda^1: \left(\hat{H}_0 - E_K^{(0)}\right)\Phi_K^{(1)} + \left(\hat{H}' - E_K^{(1)}\right)\Phi_K^{(0)} = 0 \tag{2.11b}$$

$$\lambda^2: \left(\hat{H}_0 - E_K^{(0)}\right)\Phi_K^{(2)} + \left(\hat{H}' - E_K^{(1)}\right)\Phi_K^{(1)} - E_K^{(2)}\Phi_K^{(0)} = 0 \tag{2.11c}$$

- These can be consecutively solved for the energy up to a given order by projecting onto the zeroth order state $\Phi_K^{(0)}$:

$$E_K^{(0)} = \left\langle \Phi_K^{(0)} \middle| \hat{H}_0 \middle| \Phi_K^{(0)} \right\rangle \tag{2.12a}$$

$$E_K^{(1)} = \left\langle \Phi_K^{(0)} \middle| \hat{H}' \middle| \Phi_K^{(0)} \right\rangle \tag{2.12b}$$

$$E_K^{(2)} = \left\langle \Phi_K^{(0)} \middle| \hat{H}' \middle| \Phi_K^{(1)} \right\rangle \tag{2.12c}$$

Here, it is exploited that PT employs an *intermediate normalization* of the wave function: the higher order wave function corrections are generally assumed to be small, such that $\left\langle \Psi_k \middle| \Phi_K^{(0)} \right\rangle \approx 1$. Furthermore, the former are orthogonal to the zeroth order wave function (i.e., $\left\langle \Phi_K^{(0)} \middle| \Phi_K^{(n)} \right\rangle = 0$, for all $n \neq 0$).

Possible variants of $\hat{H}_0$ and $\Phi^{(0)}$ are described in the upcoming Sections 2.2.3 and 2.2.4 on Hartree-Fock (HF) and Kohn-Sham (KS) density functional theory, respectively. In RS-PT, the first order corrections $\Phi_K^{(1)}$, which are, e.g., required for the second order energy correction in Eq. 2.12c, are expressed in terms of linear combinations of the eigenfunctions obtained at zeroth order[56]

$$\Phi_K^{(1)} = \sum_{L \neq K} c_{LK}\Phi_L^{(0)}\,. \tag{2.13}$$

c_{LK} are the linear combination coefficients and determine how much $\Phi_L^{(0)}$ contributes to state K at first order of the perturbation. These can be determined similar to the procedure described above by projection onto a state $\Phi_M^{(0)}$ (with $M \in \{L\}$). This way, the total energy of the perturbed system (i.e., $\lambda = 1$) in state K up to second order is then given as

$$E_K \approx E_K^{(0)} + \left\langle \Phi_K^{(0)} \middle| \hat{H}' \middle| \Phi_K^{(0)} \right\rangle + \sum_{M \neq K} c_{MK} \left\langle \Phi_K^{(0)} \middle| \hat{H}' \middle| \Phi_M^{(0)} \right\rangle \tag{2.14a}$$

$$= E_K^{(0)} + H'_{KK} + \sum_{M \neq K} \frac{H'_{KM}H'_{MK}}{E_M^{(0)} - E_K^{(0)}}\,. \tag{2.14b}$$

Typically, K refers to the electronic ground state (i.e., $K = 0$). The summation then runs over all excited states M, which are solutions to the zeroth order Hamiltonian $\hat{H}_0$. Since most employed perturbation operators $\hat{H}'$ at most act on two electrons (cf. Coulomb operator Eq. A1.8), the summation is typically restricted to singly and doubly excited states.

2.2.3. Hartree-Fock Theory and Electron Correlation

In HF theory, a solvable zeroth order problem is obtained, by using a single Slater determinant[52,57,58] as wave function $\Psi_K \approx \Phi_K^{(0)} \equiv \Phi_K$

$$\Phi_K \equiv \Phi_K(\underline{1}, \underline{2}, \cdots, \underline{\mathrm{N}}) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_{1\sigma}(\underline{1}) & \psi_{2\tau}(\underline{1}) & \cdots & \psi_{N\upsilon}(\underline{1}) \\ \psi_{1\sigma}(\underline{2}) & \psi_{2\tau}(\underline{2}) & \cdots & \psi_{N\upsilon}(\underline{2}) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_{1\sigma}(\underline{\mathrm{N}}) & \psi_{2\tau}(\underline{\mathrm{N}}) & \cdots & \psi_{N\upsilon}(\underline{\mathrm{N}}) \end{vmatrix} . \tag{2.15}$$

Φ_K denotes the wave function of a system containing N electrons, while $\psi_{i\sigma}(\underline{\mathrm{k}})$ denotes the i^{th} molecular orbital (MO) with spin σ. For clarity, the spin of the MO is labeled by the respective extra subscript ($\sigma, \tau, \upsilon \in \{\alpha, \beta\}$). The MO $\psi_{i\sigma}(\underline{\mathrm{k}})$ is occupied by the k^{th} electron, respectively ($i, k \in \{1, 2, \cdots, N\}$). A molecular orbital itself is defined as the wave function of a single particle system.

The electronic many-electron system is a fermionic system (i.e., $\Psi(\underline{1}, \underline{2}) = -\Psi(\underline{2}, \underline{1})$).[50] Hence, the *anti-symmetrized product* of one-particle functions (see Appendix A1), which is an equivalent representation of the determinant, is an appropriate wave function approximation and is exact in the case of non-interacting particles. The orbitals $\psi_{i\sigma}(\underline{1})$ (depending on all variables of electron 1) in turn consist of a spin part σ_i (i.e., an α or β spin function) and a spatial MO $\psi_i(\mathbf{r}_1)$ (depending only on the position in Cartesian space of electron $\mathbf{r}_1$)

$$\psi_{i\sigma}(\underline{1}) = \sigma_i \psi_i(\mathbf{r}_1) \,. \tag{2.16}$$

The spin MO $\psi_{i\sigma}(\underline{1})$ and the spatial MO $\psi_i(\mathbf{r}_1)$ are both expressed by the Greek letter ψ in the following. They are distinguished by the variables they depend on and the additional/missing index σ. In HF theory, the energy of the N-electron wave function Φ_K is determined by the orbitals. Since the orbitals are unknown initially, the variational principle is exploited and a trial wave function $\tilde{\Phi}_K$ (with trial orbitals $\tilde{\psi}_{i\sigma}(\underline{1}) = \sigma_i \tilde{\psi}_i(\mathbf{r}_1)$) needs to be employed. Insertion of $\tilde{\Phi}_K$ into the time-independent SE (as in Eq. 2.7) and minimization of $\tilde{E}_K^{(0)}$ by variation of the orbitals subject to the constraint $\left\langle \tilde{\psi}_{i\sigma} \middle| \tilde{\psi}_{j\tau} \right\rangle = \delta_{\sigma\tau}\delta_{ij}$ (using Lagrange's method of undetermined multipliers) leads to an effective one-electron problem known as the canonical Hartree-Fock equations (see Ref. 52, Chapter 3.2)

$$\hat{f}(\mathbf{r}_1)\psi_{i\sigma}(\underline{1}) = \left[\hat{h}(\mathbf{r}_1) + \sum_j^N \left(\hat{J}_j(\mathbf{r}_1) - \hat{K}_{j\tau}(\mathbf{r}_1) \right) \right] \psi_{i\sigma}(\underline{1}) = \epsilon_{i\sigma}\psi_{i\sigma}(\underline{1}) \,. \tag{2.17}$$

$\hat{h}(\mathbf{r}_1)$ is the one-electron operator, which contains the kinetic energy and the nuclear Coulomb interaction operator, depending on the spatial coordinates $\mathbf{r}_1$ of one electron (see also Appendix A1). $\hat{J}_j(\mathbf{r}_1)$ and $\hat{K}_{j\tau}(\mathbf{r}_1)$ are the electron-electron Coulomb interaction (also called Hartree term) and (Fock-)exchange operator (see Appendix A1 for the definitions). As can be seen from Eq. A1.12, the exchange interaction is only non-zero for electrons of the same spin. Different from the other terms, the exchange term has no classical counterpart and originates from the anti-symmetry of the fermionic wave function. The eigenvalue of the Fock operator $\hat{f}(\mathbf{r}_1)$ corresponds to the energy $\epsilon_{i\sigma}$ of the molecular orbital $\psi_{i\sigma}(\underline{1})$, in the mean field of the other electrons. Since the Fock operators themselves depend on the orbitals (via $\hat{J}_j(\mathbf{r}_1)$ and $\hat{K}_{j\tau}(\mathbf{r}_1)$), the Hartree-Fock equations are solved iteratively in a self-consistent field (SCF) procedure. For this purpose, one starts with a set of trial orbitals $\tilde{\psi}_{i\sigma}$ and minimizes the HF energy, which for a system of N electrons is then given by[50,52]

$$E_{HF} = \sum_i^N \left[\left\langle \psi_i \middle| \hat{h} \middle| \psi_i \right\rangle + \frac{1}{2} \sum_j^N \left(\left\langle \psi_i \middle| \hat{J}_j \middle| \psi_i \right\rangle - \left\langle \psi_{i\sigma} \middle| \hat{K}_{j\tau} \middle| \psi_{i\sigma} \right\rangle \right) \right] \tag{2.18a}$$

$$= \sum_i^N h_{ii} + \frac{1}{2} \sum_{i,j}^N \left[(ii|jj) - \delta_{\sigma\tau}(ij|ji) \right] . \tag{2.18b}$$

Here, we have introduced the chemist's notation[52] for two-electron integrals

$$(ij|kl) = \iint \psi_i^*(\mathbf{r}_1)\psi_j(\mathbf{r}_1)\frac{1}{r_{12}}\psi_k^*(\mathbf{r}_2)\psi_l(\mathbf{r}_2)\mathrm{d}\mathbf{r}_1\mathrm{d}\mathbf{r}_2 \,, \tag{2.19}$$

with $r_{12} = |\mathbf{r}_2 - \mathbf{r}_1|$. For the total energy of the system, the classical nuclear-nuclear repulsion energy E_{nn} needs to be added to E_{HF}. Variation of the orbitals to minimize the energy automatically yields a ground state Slater determinant Φ_0 with orbital occupations that follow the "Aufbau" principle (see Figure 2.1).

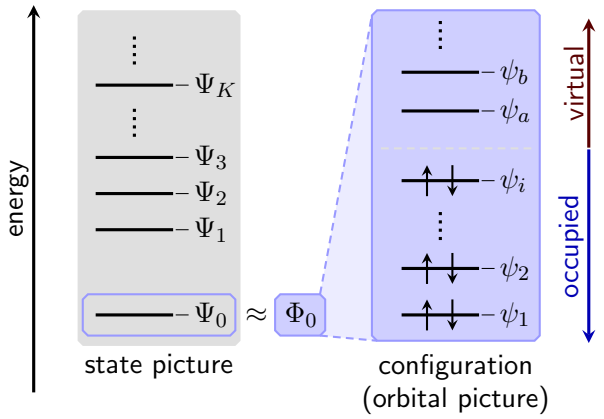


Figure 2.1.: Schematic depiction of the electronic eigenstate spectrum to the electronic Hamiltonian (left). The eigenstates Ψ_K are ranked according to their energy. The energetic ground state Ψ_0 is approximated by a single determinant Φ_0 in HF theory. Here, the many-electron wave function is described as a set of independent one-particle functions, the orbitals. Variational ground state methods such as HF theory and Kohn-Sham density functional theory (see below) lead to an electron configuration that follows the "Aufbau" principle (right).

In HF theory, the N electron system is described as a system of N non-interacting electrons each experiencing the average or mean field of the remaining $(N-1)$ electrons.[‡] In molecular systems, this amounts to about 99% of the true total energy.[55] The difference between the HF energy and the non-relativistic energy E of the fully interacting system is defined as the electron correlation energy

$$E_{\text{corr.}} = E - E_{HF}\,. \tag{2.20}$$

While the correlation contribution is small in terms of total energies, its contribution to relative energies and molecular properties is larger.[55] Thus, the inclusion of electron correlation is highly important to accurately describe covalent binding.[50,52,55] Furthermore, at long interelectronic separations the electron correlation is responsible for the *London dispersion* interaction, which is crucial for the proper description of non-covalently interacting systems. Hence, London dispersion forces are missing in HF theory.

In wave function theory (WFT), electron correlation can, e.g., be included by means of many-body RS-PT (MB-PT) as outlined in Section 2.2.2. The most commonly employed RS-PT variant is the Møller-Plesset (MP) approach, which defines $\hat{H}^{(0)}_{MP} = \sum_i^N \hat{f}(i)$ with the HF Slater determinant as zeroth order eigenfunction. Consequently, the zeroth order in Møller-Plesset perturbation theory (MP-PT) solution then equals the summed energies of all occupied orbitals and *not*

[‡]The exchange integral is identical to the Coulomb integral if $i = j$, thus, canceling the electron interaction with itself.

the HF energy. However, one arrives at the HF solution by choosing the *fluctuation potential*[50,55] to be the perturbation operator (cf. Eq. 2.8) $H'_{MP} = \frac{1}{2}\sum_{i}^{N}\left[\hat{V}_{ee}(\mathbf{r}_i) - \sum_{j}^{N}\left(\hat{J}_j(\mathbf{r}_i) - \hat{K}_{j\tau}(\mathbf{r}_i)\right)\right]$. The fluctuation potential corresponds to the difference of the true and the mean field electron-electron interaction. This way the energy obtained up to first order then coincides with the HF energy (i.e., $E_{HF} \equiv E_{MP1} = E^{(0)}_{MP} + E^{(1)}_{MP}$). Electron correlation first appears at second order and the corresponding MP2 approach[59] is the least involved electron correlation method based on a HF reference. One may consider higher orders in the perturbation, however, the method then becomes computationally more expensive (by an order of N for each additional order of λ). Furthermore, the MP series does not necessarily converge[55] and non-perturbative methods are often used instead. Noteworthy examples for the latter are the configuration interaction (CI)[52,55] and the coupled cluster (CC)[55,60] approaches. In a nutshell, these approaches approximate the wave function Ψ_K as a combination of electron configurations and optimize the respective amplitudes. Starting from the HF determinant, the excited configurations are obtained by transferring electrons from the occupied to the virtual space (see Figure 2.1). Due to the tremendously growing computational demands, the respective expansions generally need to be truncated. Compared to the HF method, all mentioned electron correlation methods (MPx (x≥2), CI, and CC) are formally higher in the computational demands by at least one order of N (i.e., with the number of electrons). For example, the computational costs of a CC treatment including singles, doubles and perturbative triples, shortly denoted as CCSD(T) and often termed the "gold standard" of quantum chemistry,[50,55] grow with N^7 (cf. the N^4 scaling of HF and N^5 of MP2). Hence, the use of correlated WFT methods becomes prohibitive for large scale calculations and it would be beneficial to improve the zeroth order description of the (non-interacting) electronic system. A possible pathway for this is provided by Kohn-Sham density functional theory.

2.2.4. Kohn-Sham Density Functional Theory

The electron density $\rho(\mathbf{r})$ is a three dimensional, observable property. For a system of non-interacting particles, it corresponds to the summed probability densities of the individual orbitals[50,61]

$$\rho(\mathbf{r}) = \sum_{\sigma=\alpha,\beta}\rho_\sigma(\mathbf{r}) = \sum_{\sigma=\alpha,\beta}\sum_{i}^{N_\sigma}\left|\psi_i(\mathbf{r})\right|^2 . \tag{2.21}$$

Here, the density has been split into its different spin components, which in turn are given by the summed probability densities of the occupied orbitals of the corresponding spin. N_σ corresponds to the total number of electrons with spin σ ($N = N_\alpha + N_\beta$). It was established[9,62] that the electron density determines the potential and, therefore, the many-electron wave function. As pointed out by Kohn and Sham,[63] the true electron density of the interacting system can be provided by a system of non-interacting fermionic particles if the corresponding mean field potential was known. Formally, replacing the sum over exchange operators $\hat{K}_{j\tau}(\mathbf{r}_1)$ in Eq. 2.17 by a potential $\hat{v}_{XC}[\rho]$ that solely depends on the electron density yields the corresponding equations

for Kohn-Sham (KS) density functional theory (KS-DFT or simply DFT[§]).[50,63] Likewise, in the corresponding HF energy expression (Eq. 2.18) the sum over $\delta_{\sigma\tau}(ij|ji)$ is replaced by an electron density dependent functional $E_{XC}[\rho]$, which is linked to the potential by $\hat{v}_{XC}[\rho] = \delta E_{XC}[\rho]/\delta\rho$. This exchange-correlation (XC) functional is composed of an exchange ($E_X[\rho]$) and a correlation ($E_C[\rho]$) functional part, though these are not necessarily separable.[69,70] If the exact functional was known, $E_{XC}[\rho]$ would yield the correct exchange (X) and correlation (C) energy.[9,71] In this case, a system of non-interacting fermionic particles would yield the same density and energy as the fully interacting N-electron system.[63,69] This way, KS-DFT offers an appealing alternative to treat electron correlation within a mean field or non-interacting particle approach. Since the exact functional of an arbitrary density is unknown, density functional approximations (DFAs) are used. This fact results in the existence of a plethora of different DFAs,[72,73] purely empirical ones and others that are derived from physical constraints. In this section, no specific functional form is explicitly given, but rather the fundamental categories of DFAs are presented. The local spin density approximation (LSDA), which is derived from a uniform electron gas (UEG, here $\nabla\rho(\mathbf{r}) = 0$ at all possible $\mathbf{r}$),[50,66,74] yields a functional at position $\mathbf{r}$ that solely depends on the electron density at this position, hence, the term "local".

In a general form, the (semi-)local density dependent exchange-correlation energy $E_{XC}[\rho]$ can be given by[50,61]

$$E_{XC}[\rho] \equiv E_{XC}[\rho_\alpha, \rho_\beta] = \int \varepsilon_{XC}\left[\nabla^a\rho(\mathbf{r})\right]\left(\rho_\alpha(\mathbf{r}) + \rho_\beta(\mathbf{r})\right)\mathrm{d}\mathbf{r}\ . \tag{2.22}$$

Typically, $\varepsilon_{XC} = \varepsilon_X + \varepsilon_C$, thus, it consists of different expressions to describe the exchange and correlation part, respectively. Generally, ε_{XC} depends on the individual α and β densities (and their derivatives). For brevity, this is not explicitly denoted here, but implied by the dependence on $\nabla^a\rho(\mathbf{r})$. Both incorporate different order derivatives of the respective (α and β) electron densities, while in the LSDA, ε_X and ε_C only depend on the α and β electron densities at position $\mathbf{r}$ (i.e., $a = 0$ in Eq. 2.22). LSDA performs poorly for non-homogeneous densities, as found in molecules, thus, more sophisticated approximations for E_{XC} have been proposed. Along with the the local density $\rho(\mathbf{r})$, these consider density derivatives:

- local spin density approximation (LSDA): $\rho(\mathbf{r})$
- generalized gradient approximation (GGA): $\rho(\mathbf{r})$, $\nabla\rho(\mathbf{r})$
- meta-GGA: $\rho(\mathbf{r})$, $\nabla\rho(\mathbf{r})$, $\nabla^2\rho(\mathbf{r})$, $\tau(\mathbf{r})$

DFAs belonging to the family of GGAs and meta-GGAs are often denoted as semi-local functionals, as they incorporate some non-local density information via derivatives of the electron

[§]Density functional theory was initially formulated purely in terms of the density.[50,64–66] In the KS approach,[63] the density is constructed from orbitals (see Eq. 2.21), which are furthermore used to compute the kinetic energy in an independent particle model (like HF). The significant improvement over previous orbital-free formulations is partly responsible for the success story of density functional theory (Kohn was awarded the Nobel Prize in 1998)[67] and has become the *de-facto* standard way of employing density functional theory.[68] Hence, DFT refers to its KS formulation in the following, while explicit reference to KS (or KS-DFT) may occur to point out the independent particle nature and conceptual similarity to HF theory.

density, yet, at a single position $\mathbf{r}$. $\tau(\mathbf{r})$ is the kinetic energy density, which is given by

$$\tau(\mathbf{r}) = \frac{1}{2} \sum_{\sigma=\alpha,\beta} \sum_{i}^{N_\sigma} |\nabla \psi_i(\mathbf{r})|^2 \tag{2.23}$$

for a system of non-interacting particles. It is preferentially employed in meta-GGA DFAs instead of the second order derivative (or Laplacian) of the density, which typically introduces some numerical noise to the functional.[72]

Among the appealing features of DFT are the use of the electron density, which is an experimentally observable object, and the theoretical possibility of describing electron correlation in a mean field approach.[50] Unfortunately, local and semi-local DFAs suffer from the self-interaction error (SIE), which refers to the Coulomb interaction of an electron with itself. Note that the sum in Eq. 2.17 runs over all occupied orbitals j including orbital i. In HF theory, the exchange term exactly cancels the corresponding Coulomb contribution if $j = i$ and, hence, removes the self-interaction. In (semi-)local DFT, this is not the case and the SIE can become quite pronounced, generally favoring delocalized electron densities. Becke introduced the concept of hybrid DFAs in 1993.[69] In the so-called global hybrid DFAs, some fraction a_x of the (semi-)local exchange energy is replaced by the Fock exchange contribution from HF theory. Initially this approach was physically motivated to interpolate between the exchange energy of the non-correlated HF system and the correlated DFT system (*adiabatic connection*). In most hybrid DFA developments, however, a_x is simply a fit parameter and varies depending on the property the DFA was fitted to reproduce.[72] Introducing Fock exchange reduces the SIE and hybrid functionals generally perform better than the respective semi-local DFAs for thermochemistry and excitation energies.[34,35,73,75] The expression of the global hybrid KS-DFT energy is then given by

$$E_{KS}^{\text{hybrid}} = \sum_{i}^{N} h_{ii} + \frac{1}{2} \sum_{i,j}^{N} \left[(ii|jj) - \delta_{\sigma\tau} a_x (ij|ji)\right] + (1 - a_x) E_X + E_C \,, \tag{2.24}$$

while the corresponding KS operator is given by

$$\hat{f}(\mathbf{r}_1) = \hat{h}(\mathbf{r}_1) + \sum_{j}^{N} \left[\hat{J}_j(\mathbf{r}_1) - a_x \hat{K}_{j\tau}(\mathbf{r}_1)\right] + (1 - a_x)\hat{v}_X[\rho] + \hat{v}_C[\rho] \tag{2.25a}$$

$$= \hat{h}(\mathbf{r}_1) + \hat{v}_{MF}(\mathbf{r}_1) \,. \tag{2.25b}$$

Eqs. 2.24 and 2.25 are valid for the case of a global hybrid DFA and HF (in that case $\hat{v}_C[\rho] = 0$ and $a_x = 1$, cf. Eqs. 2.18 and 2.17). The corresponding equations for a semi-local functional of LSDA, GGA, or meta-GGA type are obtained if $a_x = 0$. $\hat{v}_{MF}(\mathbf{r}_1)$ is the mean field potential generated by the electrons.

Due to the local nature of the Coulomb operator, the computation of the $(ii|jj)$ terms can generally be accelerated in practical calculations,[76] leading to a formal scaling of the computational costs with N^3 (N being the number of electrons). Since this is not efficiently possible

for the $(ij|ji)$ terms, the inclusion of Fock exchange in hybrid DFAs makes the functionals computationally more involved.

The classification of DFAs and the connection to their accuracy and computational cost is summarized best by the "Jacob's ladder" metaphor coined by Perdew and Schmidt.[61] A schematic representation is given in Figure 2.2.

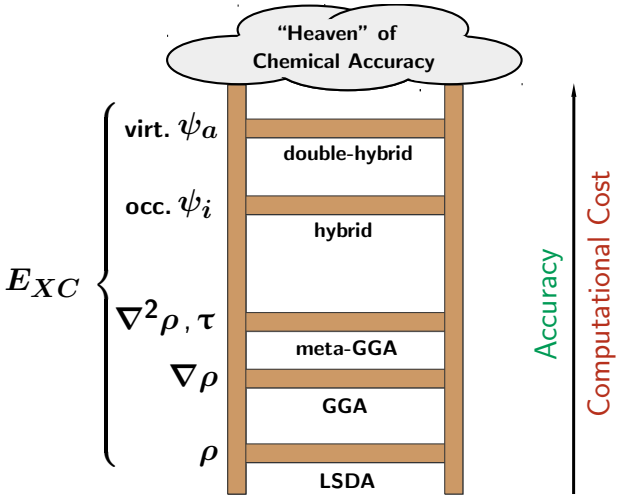


Figure 2.2.: Adapted schematic depiction of Perdew and Schmidt's "Jacob's ladder"[61] of DFAs. The different spacings between the rungs indicate the different degree of improvement in terms of the accuracy.

The least sophisticated approximation for exchange and correlation, the LSDA, shows insufficient accuracy for molecular calculations.[50] Including first (GGA) and second (meta-GGA) local derivatives of the density (or τ) one can climb up the ladder with the expressions for ε_{XC} becoming slightly more involved. Since the expressions for GGA and meta-GGA still depend only on one point in space, they still work in a semi-local approximation and, hence, retain N^3 scaling. Inclusion of Fock exchange (i.e., of non-local occupied orbital information) leads to a N^4 scaling of the method and significantly increased accuracy of the method.[73,77] The highest rung on the ladder is not extensively discussed here. It corresponds to a non-local treatment of both exchange and correlation. The latter is then achieved by also including virtual orbital information,[78–83] e.g., by means of (modified) Møller-Plesset PT[84] as in WFT. Hence, these so-called double-hybrid functionals are no longer pure mean field approaches and go beyond the zeroth order term in the many-body perturbation series (see Section: 2.2.2).[85] Typical variants can be regarded to be the corresponding KS counterpart to MP2, thus, having similar computational demands. Nevertheless, they typically provide a higher accuracy than the respective HF based counterpart.[73,81,82,86]

The best description of the electronic structure by a mean field approach is consequently achieved at the hybrid DFT level (fourth rung on the "Jacob's ladder"), which typically surpasses

the accuracy of HF and semi-local DFAs.[34,35,73,75] For the present work, it is important to consider another variant of the hybrid DFA, more specifically the range-separated hybrid (RSH) DFA. Here, the Coulomb operator used in the exchange part $K_{j\tau,i\sigma}$ of HF is first separated into a short-range (SR) and a long-range (LR) part,[87] which are then treated separately[88–90]

$$\frac{1}{r_{12}} = \underbrace{\frac{a_x' + a_x''\mathrm{erf}(\mu r_{12})}{r_{12}}}_{\text{LR}\rightarrow\text{non-local Fock}} + \underbrace{\frac{1 - a_x' - a_x''\mathrm{erf}(\mu r_{12})}{r_{12}}}_{\text{SR}\rightarrow\text{(semi-)local DFA}} . \tag{2.26}$$

The SR part is replaced by a possibly adjusted[91–93] semi-local DFA from the LSDA, GGA, or meta-GGA class, while the LR part is exactly evaluated, yielding a regular Fock exchange term (scaled by a_x') and a modified one (scaled by a_x''), which includes the error function term $\mathrm{erf}(\mu r_{12})$. The latter is used to smoothly interpolate between the SR and LR part with μ being a parameter controlling the steepness of the interpolation. a_x' defines the amount of constantly or globally employed Fock exchange (note that Eq. 2.26 corresponds to the global hybrid case if $a_x'' = 0$). If $a_x' + a_x'' = 1$, the XC potential correctly decays with $1/r_{12}$ at LR, while it incorrectly decays with a_X/r_{12} in global hybrid DFAs ($\hat{v}_{XC}$ decays exponentially).[92] Thus, SIE related effects are more pronounced in semi-local and global hybrid DFAs with small a_x. Such functionals show underestimated gaps between occupied and virtual orbitals and, hence, lead to an inconsistent description of Rydberg and charge transfer (CT) excitations with respect to valence excitations (see also Section 2.2.6).[34,94–96] RSH functionals reduce the SIE and extenuate this inconsistency, though to a varying degree depending on the parameters a_x', a_x'', and μ.[97] Consequently, they have become of significant importance in the computation of excited states and related properties.[34,98] In Table 2.1, the DFAs mentioned in this thesis are listed.

Table 2.1.: Density functional approximations appearing in this work. The corresponding key references are given in the brackets that follow the DFA abbreviation.

GGA	BLYP [99,100], BP86 [99,101,102], B97[a] [103], PBE [104]
meta-GGA	TPSS [105]
global hybrid	BHLYP [69], B3LYP [106,107], M06 [108,109], PBEh-3c[b] [110], PBE0 [111], PW6B95 [112]
range-separated hybrid	CAM-B3LYP[c] [92], ωB97X [113], ωB97X-D3 [114]
double-hybrid	B2PLYP [84]

[a] Originally a global hybrid DFA,[115] this work refers to the pure GGA reparametrization from Ref. 103.

[b] Composite scheme employing modified PBE0, a double-ζ basis set, and atom pairwise corrections.

[c] This RSH is not asymptotically correct, since $a_x' + a_x'' = 0.65 \neq 1$.

While the self-interaction is exactly canceled for occupied orbitals in HF, unoccupied or virtual

orbitals experience the full potential from the N electrons of the system.[¶] Therefore, occupied and virtual orbitals are not described on equal footing in HF and the latter rather correspond to a configuration with $(N+1)$ electrons. As a result, the energy gap between the highest occupied MO (HOMO) and lowest unoccupied MO (LUMO) is not a good first guess for excitation energies. Since the XC potential in semi-local functionals lacks MO information, occupied and virtual orbitals are treated on equal footing, hence, providing a better guess for valence excitation energies.[116]

2.2.5. Roothaan-Hall Formalism

Hartree-Fock and Kohn-Sham theory simplify the non-solvable multi electron problem by an effective one-electron problem that is solved by minimizing the total HF/KS energy with respect to the orbitals. So far, the HF/KS equations (Eqs. 2.17 and 2.25, respectively) represent a set of integro-differential equations, which remain practically non-solvable for molecular systems.[52] However, they can efficiently be solved in the Roothaan-Hall formalism.[52,117,118] Here, the spatial part of the molecular orbitals is described as an orthonormal linear combination of atom-centered orbitals (LCAO-MO)

$$\psi_i(\mathbf{r}_1) = \sum_{\mu}^{N_{AO}} C_{\mu i}\phi_\mu(\mathbf{r}_1)\,. \tag{2.27}$$

N_{AO} is the number of atomic orbitals (AOs) provided in the calculation. Since the spin part of the spin MOs $\psi_{i\sigma}$ guarantees orthogonality to spin MOs $\psi_{j\tau}$ of different spin ($\sigma \neq \tau$), the same AO expansion in Eq. 2.27 may be employed for the spatial part of α and β spin MOs. Expressing the spatial part of $\psi_{i\sigma}$ in Eqs. 2.16 by Eq. 2.27, the HF/KS equations can be solved as an algebraic general eigenvalue problem by variation of the LCAO-MO coefficients $C_{\mu i}$

$$\mathbf{FC} = \mathbf{SC}\boldsymbol{\epsilon}\,. \tag{2.28}$$

$\mathbf{F}$ is the Fock matrix with elements $F_{\mu\nu} = \delta_{\sigma\sigma'}\left\langle \phi_\mu \middle| \hat{f} \middle| \phi_\nu \right\rangle = \delta_{\sigma\sigma'}F_{\mu\nu\sigma}$, with $\hat{f}$ being the effective one-electron Fock operator from Hartree-Fock or Kohn-Sham theory. It is seen that $\mathbf{F}$ has only non-zero elements if ϕ_μ and ϕ_ν are AOs that are used to span the spatial MO space of spin MOs with identical spin (i.e., $\sigma = \sigma'$). $\mathbf{C}$ is the corresponding LCAO-MO coefficient matrix with one eigenvector per column, and $\mathbf{S}$ is the atomic orbital (AO) overlap matrix with elements $S_{\mu\nu} = \delta_{\sigma\sigma'}\langle\phi_\mu \,|\, \phi_\nu\rangle$. $\boldsymbol{\epsilon}$ is a diagonal matrix containing the MO eigenvalues $\epsilon_{i\sigma}$ on the diagonal. For a fixed set of AO functions (called a *basis*) located on the atoms, this reduces the solution of the HF/KS equations to a variational optimization of the LCAO-MO coefficients. The non-zero elements of the Fock matrix (with ϕ_μ,$\phi_\nu \in \psi_{i\sigma}$, ϕ_κ,$\phi_\iota \in \psi_{j\tau}$) are given in the general form by

$$F_{\mu\nu\sigma} = h_{\mu\nu} + \sum_{\kappa,\iota\in j}^{N_{AO}} P_{\kappa\iota\tau}\left[(\mu\nu|\kappa\iota) - \delta_{\sigma\tau}a_x(\mu\iota|\kappa\nu)\right] + (1-a_x)\left\langle\phi_\mu\left|\hat{v}_X[\rho]\right|\phi_\nu\right\rangle + \left\langle\phi_\mu\left|\hat{v}_C[\rho]\right|\phi_\nu\right\rangle\,. \tag{2.29}$$

[¶] If $\hat{f}(\mathbf{r}_1)$ acts on a virtual orbital $\psi_{a\sigma}(\underline{1})$ (cf. Eq. 2.17), the orbital j in the summation, which runs over all occupied orbitals, can never become equal to a.

Here, the density matrix $\mathbf{P}_\sigma$ is introduced with matrix elements given by

$$P_{\mu\nu\sigma} = \sum_{i}^{N_\sigma} n_i C_{\mu i}^* C_{\nu i} \,. \tag{2.30}$$

n_i is the occupation number and equal to one for an occupied spin MO $\psi_{i\sigma}$. Eq. 2.29 is generally valid for semi-local and global hybrid KS-DFT but also for HF (if $a_x = 1$ and $\hat{v}_C = 0$). Due to the dependence of the Fock matrix elements on the coefficients via the density (in the XC potential) and the density matrix, the Roothaan-Hall equations are solved iteratively until the energy is converged, corresponding to the SCF procedure. The Roothaan-Hall approach has become the standard way to solve the SCF HF and KS-DFT equations. Thus, in any upcoming reference to HF theory or KS-DFT, even if not explicitly mentioned, it is implied that these are treated in the Roothaan-Hall formalism. It should be noted that the Roothaan-Hall formalism is equivalent to the exact HF or KS-DFT solution if an infinitely large and complete basis set (CBS) of AOs is provided. In practice, this is not possible as the matrix sizes and computational demands grow with the number of basis functions.

Since only the coefficients are optimized during the SCF procedure, it is obvious that the quality of the final wave function Φ_0 is constricted by the finite set of employed AOs. This introduces some empiricism, since the proper orbital shape of an "atom-in-a-molecule" remains unknown. Typically identical elements in a molecule are provided with the same set of AOs and a large number of basis sets for different elements has been presented.[50] Without going into too much detail, it should be noted that most AO basis sets are given in terms of normalized linear combinations (*contractions*) of primitive Gaussian-type orbitals (GTOs), which is the preferred choice in molecular systems, due to the analytic form and efficient calculation of the involved integrals.[55] The number of contracted GTOs per "physical" AO shell (the latter corresponding to the occupied orbital shells of the free atom) defines the cardinal number ζ of the basis set. For example, providing one, two, or three s-type AOs per hydrogen atom corresponds to a single-ζ (or minimal), double-ζ, or triple-ζ basis for this atom, respectively. For an accurate description of molecules, AOs that are "formally unoccupied" in the free atom need to be provided. Typically, these have a higher angular momentum l than the "physical" AOs and are called *polarization* functions. They yield an improved description of the polarized atom inside the molecule. Fortunately, the total energies computed by mean field methods are almost converged with respect to the CBS limit with polarized quadruple-ζ basis sets.[55] Correlated WFT methods converge significantly slower to the CBS limit, which additionally hampers their application for large systems. See Refs. 119–122 for reviews on basis sets and basis set related errors.

In passing, it is noted that established semiempirical methods generally work with minimal basis sets and apply approximations to the corresponding HF/KS Roothaan-Hall equations, in particular the computation of the two-electron integrals, while element-specific parameters are typically provided for the one-electron parts.[38]

2.2.6. Linear Response Theory

The SCF procedure in HF theory and KS-DFT generally converges to a variational wave function that corresponds to the ground state. Since absorption of electromagnetic radiation in the visible to ultraviolet (UV-Vis) energy range involves the transition from the electronic ground state to an excited state,[‖] we need a way to compute excited states as well. For this purpose, the concepts of PT can be applied: having determined a time-independent zeroth order wave function, we can compute its *response* to a small (external) perturbation that has initially been neglected in the zeroth order Hamiltonian H_0 (cf. Eq. 2.1). In the case of light, the system is perturbed by an oscillating electric (and/or magnetic) field

$$\hat{V}(\mathbf{r}_1, t) = \hat{V}(\mathbf{r}_1) \underbrace{\frac{1}{2}\left(e^{-i\omega t} + e^{i\omega t}\right)}_{\cos(\omega t)} . \tag{2.31}$$

Here, ω is the energy of the incoming light wave. In principle, one can introduce this external perturbation and follow the work flow as in Section 2.2.2 for the orbitals, now working with the TD-SE. However, this would neglect that the zeroth order Hamiltonian, more precisely the mean field potential $\hat{v}_{MF}$ in the Fock operator, changes upon perturbation of the wave function. Neglecting this dependency corresponds to an *uncoupled* treatment. Attempting to be on an equal footing with the variational ground state procedure, we need to take into account the changes in the zeroth order Hamiltonian due to the changing wave function (*coupled* PT). Thus, the effective one-particle Hamiltonian is then given by

$$\hat{h}^{\text{eff}}(\mathbf{r}_1, t) = \hat{f}(\mathbf{r}_1) + \lambda \hat{v}'(\mathbf{r}_1, t) = \hat{f}(\mathbf{r}_1) + \lambda \left[\hat{V}(\mathbf{r}_1, t) + \hat{v}'_{MF}(\mathbf{r}_1, t)\right] . \tag{2.32}$$

$\hat{v}'_{MF}(\mathbf{r}_1, t)$ corresponds to the change in the mean field potential due to the changed wave function. In analogy to Eqs. 2.2 and 2.13, we can express the TD orbital as

$$\psi_{i\sigma}(\underline{1}, t) = \psi_{i\sigma}(\underline{1}) e^{-i\epsilon_{i\sigma} t} \tag{2.33}$$

and the first order correction of the TD orbital

$$\psi^{(1)}_{i\sigma}(\underline{1}, t) = \sum_a c_{ai}(t) \psi_{a\tau}(\underline{1}, t) . \tag{2.34}$$

The letter a indicates that this summation runs over all virtual orbitals (not explicitly written as an upper limit) with spin τ, while i corresponds to an occupied orbital with spin σ. In principle, the summation would go over all orbitals, but for particle conserving perturbations, such as excitations, the terms involving occupied orbitals will cancel, when density perturbations

[‖] There is also a contribution from the changes in the nuclear wave functions.[15] They are responsible for the vibrational fine structure and broadening of absorption signals in electronic spectroscopy. Their explicit treatment is neglected in the present work and, instead, their effect on spectra is typically simulated by an empirical broadening of the vertically computed intensities.

are to be computed. The coefficients $c_{ai}(t)$ are determined analogously to Section 2.2.2, now by projection of the first order terms in the TD Schrödinger equation onto a possible time-independent solution $\psi_{a\tau}(\underline{1})$ and solving the resulting first order differential equation

$$c_{ai}(t) = -i\delta_{\sigma\tau} \iint \Theta(t-t')\psi_a^*(\mathbf{r}_1)\hat{v}'(\mathbf{r}_1,t)\psi_i(\mathbf{r}_1) \cdot e^{i(\epsilon_a-\epsilon_i)t'} \mathrm{d}\mathbf{r}_1 \mathrm{d}t' \tag{2.35}$$

The integration over t' is performed with upper bound t, as enforced by the Heaviside function $\Theta(t-t')$. The theory outlined here is called *linear response theory*, which gives rise to density (and expectation value) changes to first order of the perturbation (thus the term "linear"), A first order TD change in the electron density is given by

$$\rho^{(1)}(\mathbf{r}_1,t) = \sum_i n_i \left[\psi_{i\sigma}^*(\underline{1},t)\psi_{i\sigma}^{(1)}(\underline{1},t) + \psi_{i\sigma}^{(1)*}(\underline{1},t)\psi_{i\sigma}(\underline{1},t)\right] . \tag{2.36}$$

Here, $n_i = 1$ is the occupation of the occupied orbital i with corresponding spin σ. Investigating the change in the density is equivalent to studying a change of an expectation value. Once we know the change in the density, we can insert the respective operator (e.g., electric dipole operator) acting on $\mathbf{r}_1$ and integrate over $\mathbf{r}_1$ to obtain its response. In fact, derivation of linear response theory often starts from a electric dipole moment response.[11,50,123] Inserting Eqs. 2.33, 2.34, and 2.35 in Eq. 2.36 followed by a Fourier transformation yields the frequency-dependent density change to first order[123,124]

$$\rho^{(1)}(\mathbf{r}_1,\omega) = 2\int \sum_i \sum_a \delta_{\sigma\tau}\psi_i^*(\mathbf{r}_1)\psi_a(\mathbf{r}_1)\psi_a^*(\mathbf{r}_2)\psi_i(\mathbf{r}_2)\frac{\epsilon_{i\sigma}-\epsilon_{a\tau}}{(\epsilon_{i\sigma}-\epsilon_{a\tau})^2-\omega^2}\hat{v}'(\mathbf{r}_2,\omega)\mathrm{d}\mathbf{r}_2 \tag{2.37a}$$

$$= \int \chi_0(\mathbf{r}_1,\mathbf{r}_2,\omega)\,\hat{v}'(\mathbf{r}_2,\omega)\mathrm{d}\mathbf{r}_2 \tag{2.37b}$$

$$\stackrel{!}{=} 2\int \sum_{K\neq 0} \omega_{0K}\frac{\Psi_0^*(\mathbf{r}_1)\Psi_K(\mathbf{r}_1)\cdot\Psi_K^*(\mathbf{r}_2)\Psi_0(\mathbf{r}_2)}{\omega_{0K}^2-\omega^2}\hat{V}(\mathbf{r}_2,\omega)\mathrm{d}\mathbf{r}_2 \tag{2.37c}$$

$$= \int \chi(\mathbf{r}_1,\mathbf{r}_2,\omega)\,\hat{V}(\mathbf{r}_2,\omega)\mathrm{d}\mathbf{r}_2\,. \tag{2.37d}$$

In Eq. 2.37b, the frequency dependent response function of the independent particle system $\chi_0(\mathbf{r}_1,\mathbf{r}_2,\omega)$ is introduced. It describes the change in the electron density at position $\mathbf{r}_1$ due to a TD perturbation at position $\mathbf{r}_2$ with frequency ω. Neglecting changes in the mean field potential (in that case $\hat{v}'_{\mathrm{MF}}(\mathbf{r}_2,\omega)=0$) $\chi_0(\mathbf{r}_1,\mathbf{r}_2,\omega)$ then corresponds to the response function of an uncoupled treatment. If the change in the mean field potential is properly accounted for, the first order density change of the KS system resembles the true density response $\chi(\mathbf{r}_1,\mathbf{r}_2,\omega)$ (assuming the exact functional $E_{XC}[\rho]$ was known).[62,125] Since $\hat{v}'_{\mathrm{MF}}(\mathbf{r}_2,\omega)$ is absent in the interacting system, the true response function is solely induced by the external perturbation $\hat{V}(\mathbf{r}_2,\omega)$ (cf. Eq. 2.37c and 2.37d).** Equating the expressions for the density change of the true interact-

**In the static limit (i.e., $\omega \to 0$), the response coincides with the second order energy correction in RS-PT (compare Eq. 2.37c and 2.14b). The additional factor of two in Eq. 2.37c is absent in Eq. 2.14b, since the power series term $E^{(2)}$ is equivalent to the corresponding second order term of the Taylor expansion, i.e., $E^{(2)} \equiv \frac{1}{2}\frac{\partial^2 E}{\partial\lambda^2}$,

ing and the KS independent particle system (Eq. 2.37d and Eq. 2.37b, respectively) then yields

$$\begin{aligned}\chi\left(\mathbf{r}_1,\mathbf{r}_2,\omega\right) =& \chi_0\left(\mathbf{r}_1,\mathbf{r}_2,\omega\right) + \iint \chi_0\left(\mathbf{r}_1,\mathbf{r}_3,\omega\right)\Big[\frac{1-a_x\hat{P}(\underline{3},\underline{4})}{r_{34}} \\ &+ \left(1-a_x\right) f_{XC}(\mathbf{r}_3,\mathbf{r}_4,\omega)\Big]\chi\left(\mathbf{r}_4,\mathbf{r}_2,\omega\right)\mathrm{d}\mathbf{r}_3\mathrm{d}\mathbf{r}_4.\end{aligned} \tag{2.38}$$

The terms in brackets result from $\hat{v}'_{\mathrm{MF}}(\mathbf{r}_2,\omega)$, with the first term corresponding to the anti-symmetrized Coulomb operator. $\hat{P}(\underline{3},\underline{4})$ exchanges the electron variables of electron 3 and 4 in two orbitals, thus affording the exchange component of the Coulomb interaction. In the global hybrid DFA case, the latter is scaled by a_x. The last term in brackets is a condensed notation for the XC potential response, which in the so-called adiabatic local density approximation (ALDA) is given by

$$\left(1-a_x\right) f_{XC}(\mathbf{r}_3,\mathbf{r}_4,\omega) \approx \delta\left(\mathbf{r}_3-\mathbf{r}_4\right)\left[\left(1-a_x\right)\frac{\hat{v}_X(\mathbf{r}_3)}{\partial\rho(\mathbf{r}_3)} + \frac{\hat{v}_C(\mathbf{r}_3)}{\partial\rho(\mathbf{r}_3)}\right]. \tag{2.39}$$

The ALDA approximation is assumed to work well for slowly oscillating perturbations, i.e., $\omega \to 0$.[124] Eq. 2.38 can be interpreted in the following way: the density response to a TD external perturbation in the true interacting system (left hand side) is also reflected by the independent KS particle system (right hand side). In the latter, the equivalent TD external perturbation at position $\mathbf{r}_2$ leads to an uncoupled density response at $\mathbf{r}_1$ (first term), which is augmented with a coupling response. The same TD external perturbation at $\mathbf{r}_2$ induces a change in the density at position $\mathbf{r}_4$. This density response at position $\mathbf{r}_4$ due to an external field at $\mathbf{r}_2$, is described correctly by the true response function (Eq. 2.37). Integration over $\mathbf{r}_4$, affords the change in the mean field potential, which is experienced at position $\mathbf{r}_3$. The change of the mean field experienced at position $\mathbf{r}_3$ then also affects the KS response at $\mathbf{r}_1$. The induced oscillation at position $\mathbf{r}_1$ is thus equivalently described by the true response function $\chi\left(\mathbf{r}_1,\mathbf{r}_2,\omega\right)$ "directly" reacting to $\hat{V}(\mathbf{r}_2,\omega)$ and also by $\chi_0\left(\mathbf{r}_1,\mathbf{r}_2,\omega\right)$ reacting to $\hat{v}'(\mathbf{r}_2,\omega)$ as given by Eq. 2.38. $\chi_0\left(\mathbf{r}_1,\mathbf{r}_2,\omega\right)$ only depends on the orbitals and their respective energies and is thus easily computable for a given perturbation, which is different for $\chi\left(\mathbf{r}_1,\mathbf{r}_2,\omega\right)$. Since the latter appears on both sides of the equation, a self-consistent treatment is necessary for each perturbation frequency. However, we will find in the following that such a treatment is not required to compute the perturbation independent excited eigenstates of the system. Before this is shown, it is noteworthy that the (true) response function gives rise to any property changes due to an external TD perturbation. For example, if the TD perturbation is an oscillating electric field, the induced first order change of the electric dipole moment μ, i.e., the dynamic

while Eq. 2.37d rather correspond to the double partial derivative without the factor 1/2. Hence, the result obtained in the TD framework equals the one obtained from a time-independent treatment.

polarizability, is accessible

$$\alpha_{uv}(\omega) = 2 \sum_{K \neq 0} \omega_{0K} \frac{\langle \Psi_0 | \hat{\mu}_u | \Psi_K \rangle \cdot \langle \Psi_K | \hat{\mu}_v | \Psi_0 \rangle}{\omega_{0K}^2 - \omega^2} \,. \tag{2.40}$$

Starting from the true response function in Eq. 2.37, we have inserted the electric dipole moment operator (see Eq. 5.16 in Chapter 5) in the Cartesian direction u (acting on the coordinate $\mathbf{r}_1$) and set the external perturbation (acting on $\mathbf{r}_2$) equal to an oscillating electric dipole, e.g., the electric field of a photon oscillating in direction v, followed by an integration over $\mathbf{r}_1$. Being an energy correction term of second order of the perturbation (cf. Eq. 2.14), this sum-over-states (SOS) expression runs over all excited states K. Since $\hat{\mu}(\mathbf{r}_1)$ is a one-electron operator, the summation only includes singly excited states. Absorption of the photon energy occurs if the perturbation frequency ω equals the excitation energy ω_{0K}. The residual of the diagonal elements α_{uu} of the 3×3 polarizability tensor then provides the probability for the electronic transition, which is observable as a signal in the absorption spectra. Since we are mostly interested in systems in solution, the average of the three Cartesian components defines the so-called oscillator strength

$$f_{0K} = \frac{2}{3} \sum_{u=x,y,z} \omega_{0K} \langle \Psi_0 | \hat{\mu}_u | \Psi_K \rangle \cdot \langle \Psi_K | \hat{\mu}_u | \Psi_0 \rangle = \frac{2}{3} \omega_{0K} \vec{\mu}_{0K} \cdot \vec{\mu}_{K0} \,. \tag{2.41}$$

Optical activity is related to a change in the electric dipole moment due to a magnetic dipole perturbation (and vice versa), hence

$$\beta_{uv}(\omega) = 2c \sum_{K \neq 0} \frac{\mathrm{Im} \left(\langle \Psi_0 | \hat{\mu}_u | \Psi_K \rangle \cdot \langle \Psi_K | \hat{m}_v | \Psi_0 \rangle \right)}{\omega_{0K}^2 - \omega^2} \,. \tag{2.42}$$

Here, $\hat{m}$ is the magnetic moment operator (see Eq. 5.15 in Chapter 5). The imaginary part of the residual corresponds to the observed ECD intensity from the particular photon absorption with energy $\omega = \omega_{0K}$.[32] The analogue to the oscillator strength is the so-called rotatory strength (see Eq. 5.14 in Chapter 5).

The eigenstates ω_{0K} and transition moments in Eqs. 2.40 and 2.42 are given for the true system of interacting particles. We have established the connection of its TD response to the response of the non-interacting particle system in Eq. 2.38. Hence, we should be able to compute these entities by means of our non-interacting particle system. To achieve this, we can alternatively express the density change in Eq. 2.37 in terms of changes in the density matrix

$$\rho^{(1)}(\mathbf{r}_1, \omega) = \sum_{ia} P'_{ia}(\omega) \psi_i^*(\mathbf{r}_1) \psi_a(\mathbf{r}_2) \,. \tag{2.43}$$

Just like the response function, $P'_{ia}(\omega)$ is again a function of itself. Equating with Eq. 2.37, assuming real orbitals only, and solving for the expectation value of the external TD perturbation yields two sets of linear equations, one for $V_{ia}(\omega) = \langle \psi_i \left| \hat{V}(\omega) \right| \psi_a \rangle$ and another one for its adjoint

$W_{ia}(\omega) = V_{ai}(\omega)$. These can be written in a matrix notation as

$$\left[\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B^*} & \mathbf{A^*} \end{pmatrix} - \omega \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{-1} \end{pmatrix}\right] \begin{pmatrix} \mathbf{X}(\omega) \\ \mathbf{Y}(\omega) \end{pmatrix} = \begin{pmatrix} \mathbf{V}(\omega) \\ \mathbf{W}(\omega) \end{pmatrix}, \tag{2.44}$$

with matrix elements

$$\begin{aligned} A_{ia\sigma,jb\tau} = \delta_{\sigma\tau}\delta_{ij}\delta_{ab}(\epsilon_{a\tau} - \epsilon_{i\sigma}) + (ia|jb) - a_x\delta_{\sigma\tau}(ij|ab) \\ + (1 - a_x)(ia|f_{XC}|jb) \end{aligned} \tag{2.45}$$

and

$$B_{ia\sigma,jb\tau} = (ia|bj) - a_x\delta_{\sigma\tau}(ib|aj) + (1 - a_x)(ia|f_{XC}|bj)\,. \tag{2.46}$$

$\mathbf{X}(\omega)$ and $\mathbf{Y}(\omega)$ describe the excitation and de-excitation contributions to the density matrix change, since $X_{ia}(\omega) = P'_{ia}(\omega)$ and $Y_{ia}(\omega) = P'_{ai}(\omega)$. From Eq. 2.44 it becomes obvious that the response of the electronic system depends only on the perturbation frequency, not on the type of perturbation itself. If we let the strength of the external perturbation strength tend to zero, the right hand side of Eq. 2.44 vanishes. Apart from the trivial solutions, i.e., $\mathbf{X}(\omega) = \mathbf{Y}(\omega) = \mathbf{0}$, the left hand side becomes zero for all perturbation frequencies that correspond to the excitation energies of the electronic system, i.e., if $\omega = \omega_{0K}$. Solving the resulting eigenvalue problem (generically referred to as LR-TD in the following) affords the excitation energies ω_{0K} for the transition from the ground to an excited state K of the coupled HF/KS system. This is equivalent to knowing the poles of the response function of the coupled HF/KS system in Eq. 2.38 and, in turn, the poles of the true response function, provided the correct density functional (and its TD response) were known. The eigenvectors $\mathbf{X}$ and $\mathbf{Y}$ then allow the computation of the respective transition moments. For a purely real perturbation operator $\hat{V}^{(r)}$ (e.g., the electric dipole moment), the respective transition moment is given by $V^{(r)}_{0K} = \sum_{ia} V^{(r)}_{ia}(X^K_{ia} + Y^K_{ia})$, while $V^{(i)}_{0K} = \sum_{ia} V^{(i)}_{ia}(X^K_{ia} - Y^K_{ia})$ is the expression for a purely imaginary perturbation operator $\hat{V}^{(i)}$, e.g., a magnetic field perturbation. In the HF case, the LR-TD problem is referred to as time-dependent Hartree-Fock (TD-HF), while it is termed time-dependent density functional theory (TD-DFT) if the underlying orbitals come from a KS-DFT calculation. A more general derivation of linear response theory based on variational perturbation theory is given in Ref. 126 (chapters 2, 11, and 12). The matrix elements of $\mathbf{A}$ connect two singly excited configurations Φ^a_i and Φ^b_j, which are obtained from an $i \to a$ and $j \to b$ orbital excitation, respectively, based on the ground state determinant Φ_0 (see right-hand side of Figure 2.1). In the case of a HF reference, the matrix $\mathbf{A}$ is identical to the one employed in configuration interaction singles (CIS) calculations.[50,126] Matrix $\mathbf{B}$ contains elements that connect the ground state determinant Φ_0 with the corresponding doubly excited configuration Φ^{ab}_{ij}[50] as a result from the mean field response.[123] The computational bottlenecks in TD-DFT and TD-HF calculations are typically the computation of the two-electron integrals and the solution of the non-standard eigenvalue problem, which is complicated by the huge matrix dimensions of $N_{\text{occ.}} \times N_{\text{virt.}}$. Neglecting the

matrix **B** corresponds to the so-called Tamm-Dancoff approximation (TDA)[127] and leads to the CIS problem in HF. In the DFT context,[128] this will be abbreviated by TDA-DFT. If Fock exchange is present, this approximation speeds up the calculation by roughly a factor of two. Typically, it has only minor effects on excitation energies, but can have severe effects on the transition moments as will be discussed in Chapter 6. Part III of this thesis is devoted to the development and application of simplified TD-DFT and TDA-DFT variants in which the matrix dimensions are reduced and the involved computation of two-electron integrals is approximated.

Due to the smaller HOMO-LUMO gap, TD(A)-DFT generally provides better valence excitation energies than TD-HF or CIS.[15,34,35,129] As mentioned above, its performance for Rydberg or CT states depends primarily on the amount of Fock exchange in the DFA. In particular, asymptotically correct RSH DFAs have become popular to avoid CT-related problems, since they provide the asymptotically correct Coulomb attraction (resulting from the Fock exchange response in Eq. 2.45) in the excited state.[15,129]

Part II.

Dispersion-Corrected Density Functional Theory in Free Energy Calculations

This part deals with the application of dispersion-corrected density functional theory in a multi-level scheme to compute free energies in solution. Since HF and semi-local DFAs are devoid of long-range Coulomb correlation, they fail to describe the non-negligible London dispersion (also called "van der Waals") interactions.[50,56,130] In particular the preferential orientation of flexible or rotatable groups, which can lead to visible bands in the UV-Vis and ECD spectra, is influenced by these interactions. The importance for ECD spectra is exemplified in Figure 2.3 for two conformers of an Ace-Ala-Gly-Ala-NMe tetrapeptide. It is obvious that the ECD spectra of both structures are quite different and reasonable comparisons with experimental data can only be made if it is known which one is more stable. The hybrid density functional B3LYP[69,99,100,106,107] erroneously describes the unfolded structure to be the energetically more stable one, but this situation is remedied if London dispersion interactions are taken into account.[130] Hence, their proper description is of vital importance also for the simulation of ECD spectra.

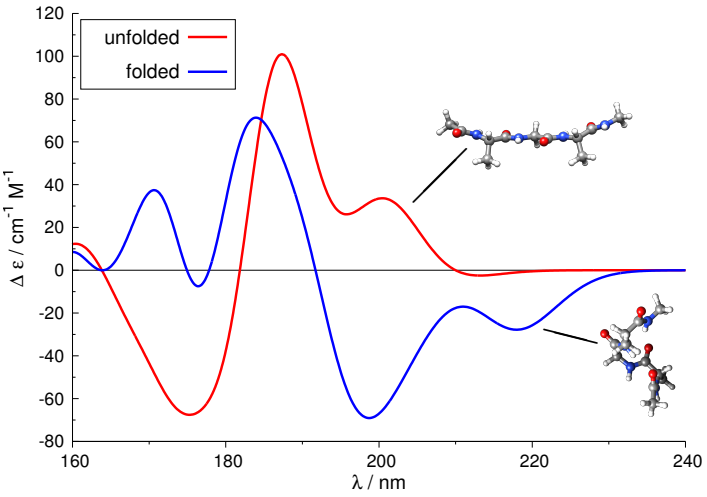


Figure 2.3.: Computed ECD spectra of the Ace-Ala-Gly-Ala-NMe tetrapeptide in an unfolded (red) and folded (blue) conformation. The geometries from Ref. 130 are used and the spectra are computed with the sTDA-xTB method that is presented in Chapter 8. Both spectra are obtained by broadening of the vertically computed ECD signals by Gaussian functions with a full width at $1/e$ maximum of 0.4 eV.

The inherent local nature of the Coulomb operator, which dominates the asymptotic region (i.e., $r_{12} \to \infty$) enabled the development of efficient correction schemes to account for dispersion interactions[130] without having to resort to elaborate correlated wave function methods. The basic procedure can be derived from perturbation theory (see Section 2.2.2). At long-range, the correlation energy from second order many-body perturbation theory (MB-PT, cf. Eq. 2.14) is

given by

$$\lim_{r_{12}\to\infty} E^{(2)}_{\text{MB-PT}} = -\sum_{i,a\in A}\sum_{j,b\in B}\frac{(ia|jb)^2}{E^{(0)}_{ij\to ab}-E^{(0)}_0}\,. \tag{2.47}$$

Here, ij denote occupied and ab denote virtual molecular orbitals, respectively, whereas the denominator corresponds to the excitation energy for the double excitation $ij \to ab$. While the orbitals i and a belong to fragment A, the orbital j and b are located on fragment B. For convenience, we will assume these fragments to be two atoms from the very beginning, as it is a commonly employed partitioning of the dispersion energy.[130] The atoms themselves may be part of the same, as well as of different molecules. If the integral $(ia|jb)$ is expressed in a multipole expansion, the dipole-dipole term is obtained at lowest order. Assuming spherical particles and averaging over all possible orientations, the well-known Casimir-Polder equation[56,131,132] is obtained

$$E^{AB}_{\text{disp}} \approx -\frac{3}{\pi R^6_{AB}}\int_0^\infty \alpha_A(i\omega)\,\alpha_B(i\omega)\,\mathrm{d}\omega = -\frac{C^{AB}_6}{R^6_{AB}}\,. \tag{2.48}$$

We have introduced the dispersion coefficient C^{AB}_6 of the two atoms A and B. R_{AB} is the interatomic distance and $\alpha_A(i\omega)$ is the isotropic dynamic dipole-dipole polarizability at imaginary frequencies $i\omega$ (cf. Eq. 2.40). Many dispersion corrections simply approximate the dispersion coefficient C^{AB}_6 in Eq. 2.48 and furthermore, introduce a damping function to avoid double counting of correlation effects, as well as divergence of E^{AB}_{disp} at short distances.

One of the most widely employed dispersion corrections is the D3 scheme.[133,134] For each element pair, the C^{AB}_6 are pre-computed for reference systems (free atoms and hydrides) employing *first principles* time-dependent density functional theory. The reference systems reflect the atom within a given molecular environment corresponding to different hybridizations. In addition to the dipole-dipole contribution, the D3 model also includes pairwise dipole-quadrupole, and optionally the Axilrod-Teller-Muto dipole-dipole-dipole three-body contributions. If the latter is included, the superscript "ATM" will be used for the method abbreviation (i.e., D3^{ATM}) in the following. The respective dispersion coefficients C^{AB}_8 and C^{ABC}_9 are empirically derived from the non-empirical C^{AB}_6 coefficients.[133]

The total D3^{ATM} dispersion energy is then given by

$$\begin{aligned} E^{\text{D3}^{\text{ATM}}}_{\text{disp}} = &-\sum_{n=6,8}\sum_{A<B}^{\text{atoms}}\frac{C^{AB}_n}{R^n_{AB}}f^{(n)}_{\text{damp}}(R_{AB}) \\ &-\sum_{A<B<C}^{\text{atoms}}\frac{C^{ABC}_9(3cos\theta_a cos\theta_b cos\theta_c+1)}{(R_{AB}R_{AC}R_{BC})^3}f^{(9)}_{\text{damp}}(\overline{R}_{ABC})\,. \end{aligned} \tag{2.49}$$

$f^{(n)}_{\text{damp}}(R_{AB})$ is a damping function that prevents the dispersion energy from diverging at small

R_{AB} ($\overline{R}_{ABC} = \sqrt[3]{R_{AB}R_{AC}R_{BC}}$). The damping function is a key ingredient of any dispersion correction[130,134] and the damping parameters typically need to be adjusted for any DFA that is to be combined with the respective dispersion correction. The initial version of D3, denoted as D3(0) in the following, employs damping functions that damp the individual contributions to zero for small R_{AB}.[133] In the most recent version of D3, the two-body terms are rationally damped.[134] The latter damping function[130,134] is more physically sound and inspired by the work of Johnson and Becke.[135] Its use is abbreviated by either D3(BJ) or simply D3. If the three-body term (last term in Eq. 2.49) is used, it is always combined with the original zero damping function.

The D3 method is among the fastest and, at the same time, most accurate dispersion corrections.[130,136,137] Other conceptually different, as well as atom pairwise, dispersion corrections exist, which are not discussed here (see Ref. 130 and 138 for recent reviews). The performance of DFAs in combination with the D3 dispersion correction for reaction energies, reaction barriers, and non-covalent interactions was previously assessed on different benchmark sets.[73]

In this part, the performance of dispersion-corrected density functional theory is demonstrated in the computation of association energies of two bimolecular frustrated Lewis pair (FLP) employing a supramolecular approach.[42] The obtained results are compared to contemporary correlated wave function theory and if possible to binding energies derived from experimental data. The computed association (free) energies imply a slightly different reactivity of the two FLPs for the reaction with small molecules such as H_2. It is shown that dispersion-corrected DFAs can properly describe the important non-covalent interactions along with intramolecular relaxation effects. Furthermore, these methods suffer much less from the basis set superposition error (BSSE) than the computationally more demanding correlated WFT approaches. In agreement with other studies,[139–141] the largest uncertainty in the computation of binding free energies seems to arise from the solvation contribution. The methodology applied in this part to FLPs may likewise be employed in a multi-level scheme to determine the absolute configuration of chiral compounds. In this context, dispersion-corrected DFT serves as a suitable method to obtain reasonable molecular geometries and conformational energies. The efficient computation of excited states (including absorption and ECD spectra) is discussed in Part III. In Part IV, two semiempirical low-cost methods are presented: one is designed to further accelerate the computation of electronic spectra, while the other method can be used in the context of free energy calculations for the conformational sampling, fast geometry optimizations, and the computation of ro-vibrational contributions.

3. The Association of Two "Frustrated" Lewis Pairs by State-of-the-Art Quantum Chemical Methods

Christoph Bannwarth,* Andreas Hansen,* and Stefan Grimme*

Received 29th of September 2014, Published online 10th of February 2015

Reprinted (adapted) with permission from†
Bannwarth, C.; Hansen, A.; Grimme, S. *Isr. J. Chem.* **2015**, *55*, 235–242.

DOI 10.1002/ijch.201400138

Own manuscript contribution

- Performing the calculations except for the HF, MP2, and DLPNO-CCSD(T) calculations with the aug-cc-pVDZ and aug-cc-pVTZ basis sets
- Interpretation of the results
- Writing the manuscript

*Mulliken Center for Theoretical Chemistry, Institut für Physikalische und Theoretische Chemie, Rheinische Friedrich-Wilhelms-Universität Bonn, Beringstraße 4, 53115 Bonn, Germany

Abstract State-of-the-art quantum chemical methods have been applied to describe the association of two frustrated Lewis pairs (FLPs) $B(C_6F_5)_3/PR_3$ (**1**: R=2,4,6-$Me_3C_6H_2$; **2**: R=CMe_3) with different steric demands of the base component. Interaction energies are calculated at the dispersion-corrected DFT, MP2 (second-order Møller-Plesset), and DLPNO-CCSD(T) (domain-based local pair natural orbital-coupled cluster, including single, double and perturbative triple excitations) levels of theory, combined with extended triple- or quadruple-ζ AO (atom-centered orbital) basis sets. Thermostatistical contributions to the free binding energy are calculated from harmonic frequencies at the efficient HF-3c (minimal basis Hartree-Fock with three corrections) level, while solvation effects in benzene are accounted for by the COSMO-RS (conductor-like screening model for realistic solvents) continuum model. Comparison to the recently measured experimental value for the free association energy of the FLP **1** reveals agreement between theory and experiment within the estimated error bars. The computed gas phase interaction energies for both FLPs are similar (about -13 kcal mol^{-1}) with only small variations (about ± 3 kcal mol^{-1}) for various quantum chemical methods when London dispersion interactions are accounted for properly. The association of the more "frustrated" FLP **1** is mainly driven by non-directional dispersion forces resulting in non-preferential orientations which is in agreement with experimental results. On the other hand, in FLP **2** with the "smaller" base the boron and phosphorus atoms face each other in the favored complex structure indicating a weak P-B donor-acceptor interaction. This conformation of **2** seems to be more suitable for small molecule (e.g., H_2) activations.

3.1. Introduction

The electrophilic borane $B(C_6F_5)_3$ (**3**) is known for about 50 years[142,143] and has found tremendous attention in the past decade with the upcoming of the so-called "frustrated" Lewis pair (FLP) chemistry.[44,144] This is the chemistry of Lewis acid/base pairs (intra- or intermolecular) which are hindered from forming covalent bonds, e.g., due to "sterical hindrance". An avalanche of studies in this field has been triggered by a landmark paper in 2006, where the ability of a FLP to activate molecular hydrogen was discovered.[43] As a result, a variety of intra- and intermolecular FLPs and their ability to activate small molecules, such as H_2, CO_2, and NO, have been studied both experimentally[145–147] and theoretically.[148–155]

The catalytic activity of intermolecular FLPs containing **3** is generally thought to be related to the "degree of frustration" of the FLP, i.e., the reduced ability to form the covalently bound Lewis pair. Recently, the association of FLP **1** (alternatively dubbed $B(C_6F_5)_3/PMes_3$) in d_6-benzene was studied by nuclear Overhauser effect spectroscopy (NOESY) measurements.[45] The association was found to be slightly endergonic ($\Delta G_{\text{exptl.}}$=+0.4 kcal mol^{-1}, at T=298 K) and no preferential orientation of $B(C_6F_5)_3$ and $PMes_3$ in the complex was observed. This experimental reference opens up the possibility to assess the quality of different quantum chemical methods in describing this FLP association. Since many of these approaches were previously used to theoretically describe, e.g., H_2 activation by FLPs, the present study is also important to judge

their applicability to FLPs in general.[150,156,157] Aside from **1**, the respective formation of FLP **2** (dubbed $B(C_6F_5)_3/PtBu_3$) will be studied and the associations of both systems are compared to each other.

3.2. Theoretical Background

For the association of the FLPs, which is assumed to occur under equilibrium conditions, the calculated free energy $\Delta G_R^T(X)$ is given in our physically motivated partitioning scheme[42] as

$$\Delta G_R^T(X) = \Delta E + \Delta G_{RRHO}^T + \Delta\delta G_{\text{solv.}}^T(X)\,. \tag{3.1}$$

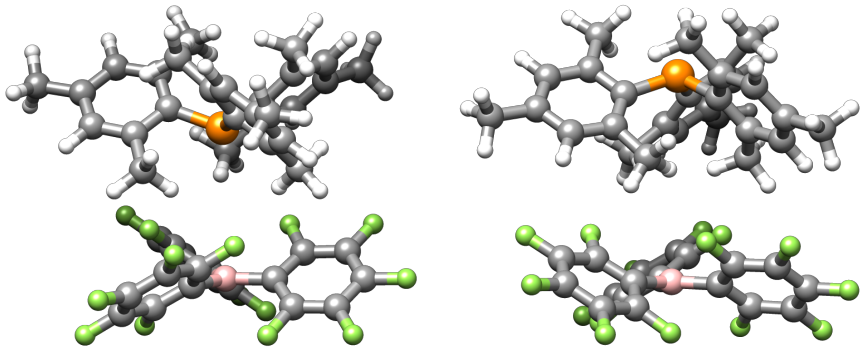
Here, for the FLP complex and its separate acid and base components, ΔG_{RRHO}^T is the difference of their ro-vibrational gas phase free energies and $\Delta\delta G_{\text{solv.}}^T(X)$ is the difference in solvation free energies. T is a given temperature and X a given solvent, i.e., in the present study 298 K and benzene (bz), respectively. ΔE is the difference in electronic, zero point vibrational exclusive gas phase energies obtained by the respective quantum chemical (QC) method (including fragment structure relaxation and optionally dispersion corrections).

Instead of calculating free energies and to compare them to experimental values, it is also possible to back-correct experimentally measured free reaction energies to "experimental" energy differences, $\Delta E_{\text{exptl.}}$

$$\Delta E_{\text{exptl.}} = \Delta G_{\text{exptl.}}^T(X) - \Delta G_{RRHO}^T - \Delta\delta G_{\text{solv.}}^T(X)\,. \tag{3.2}$$

This allows to evaluate exclusively the description of the changing electronic structure by the different methods and to conveniently benchmark approximate electronic structure methods. Since an experimental value for the association of FLP **1** has been measured recently, we perform a back-correction with the methods presented below and compare the calculated ΔE values to the "experimental" $\Delta E_{\text{exptl.}}$. We use the experimentally determined $\Delta G_{\text{exptl.}}^{298\,\text{K}}(\text{bz})$ by Rocchigiani *et al.* to derive $\Delta E_{\text{exptl.}}$ according to Eq. 3.2.[45]

For the interaction energies of supramolecular systems, which have some similarities to the association of the FLPs considered here, this procedure provided a valuable benchmark data set which agrees well with high-level quantum chemical results.[42,158] In the 2013 SAMPL4 blindtest on predicting free energies of supramolecular host-guest binding reactions,[159] this protocol performed excellent for cucurbit[7]uril host-guest associations (among the top three out of 20 submissions by all statistical measures employed).[140] Furthermore, the methods to obtain ΔG_{RRHO}^T and $\Delta\delta G_{\text{solv.}}^T(X)$ have successfully been applied in previous studies to study the thermochemistry of different reactions in solution, including reactions catalyzed by FLPs.[42,152,157,160,161]



1a **1b**

Figure 3.1.: Ball-and-stick depiction of **1a** and **1b**. Geometries are optimized at the TPSS-D3/def2-TZVP level (color code: B=pink, C=gray, F=green, H=white, P=orange).

3.3. Technical Details of the Calculations

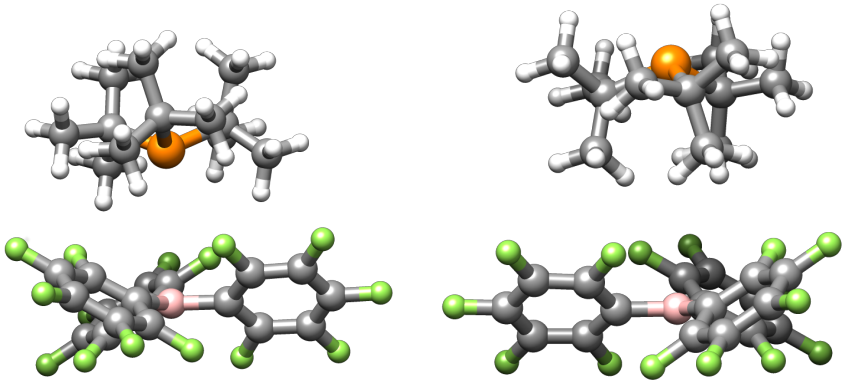
3.3.1. Geometry Optimizations

All geometries were fully optimized at the TPSS-D3/def2-TZVP level[105,133,134,162,163] using the TURBOMOLE suite of programs.[164a,165] The appended "-D3" denotes our standard atom-pairwise D3 dispersion correction combined with the Becke-Johnson damping scheme.[133,134] The resolution-of-the-identity (RI) approximation has been used to speed up the calculation of Coulomb integrals (dubbed RI-J)[76,166] using matching auxiliary basis sets.[167,168] For the numerical integration of the exchange-correlation potential, the *m5* grid was applied.

For both complexes **1** and **2**, two orientations of the Lewis base PR_3 are considered here: The phosphorus of the Lewis base may point "toward" (appended **a**) or "away from" (appended **b**) the boron of the Lewis acid. Each of these were optimized at the above-mentioned level of theory and studied by the methods described below. The resulting structures of **1** and **2** are depicted in Figure 3.1 and 3.2, respectively. Every geometry was verified as a minimum structure by performing a normal coordinate analysis at the HF-3c[169] level (*vide infra*) after re-optimizing the geometries at this level of theory.

3.3.2. Single-Point Energy Calculations

These geometries were used to perform single-point energy calculations at higher levels of theory in combination with the extended def2-QZVP atom-centered orbital (AO) basis set.[163,170] Among these are the hybrid (meta) functionals PBE0[104,111], PW6B95[112], B3LYP[99,100,106,107] and M06,[108,109] as well as the B2PLYP[84] double-hybrid functional. Dispersion interactions are accounted for by applying the D3 correction (with Becke-Johnson damping, zero-damping for M06).[73,133,134] In addition, the B3LYP functional was combined with the density-dependent, non-local (NL) van-der-Waals correction by Vydrov and Van Voorhis (VV10, dubbed B3LYP-NL



2a **2b**

Figure 3.2.: Ball-and-stick depiction of **2a** and **2b**. Geometries are optimized at the TPSS-D3/def2-TZVP level (color code: B=pink, C=gray, F=green, H=white, P=orange).

here) in the post-SCF fashion.[171,172] The applied wave function methods include (dispersion-corrected) Hartree-Fock (HF), as well as second-order Møller-Plesset theory (MP2)[59] and its spin-component-scaled (SCS-MP2)[173] variant which performs better for thermochemistry,[174] as well as for aromatic stacking interactions.[173]

The method of highest accuracy applied here is domain-based local pair-natural-orbital coupled-cluster including single, double and perturbative triple excitations, dubbed DLPNO-CCSD(T).[175] The default settings for the local cutoff thresholds were used. Since the computations on the complexes with the def2-QZVP basis are too demanding with this method, a triple-zeta basis set was used instead. In order to ensure that London dispersion interactions are treated properly, the larger and diffuse aug-cc-pVTZ (aTZ) basis set[176,177] was preferred over the smaller def2-TZVP. Unfortunately, the aTZ basis set is not properly converged with respect to the complete basis set (CBS) limit and is particularly prone to basis set superposition error (BSSE).[178,179] To account for these, an aDZ/aTZ basis set extrapolation according to Halkier *et al.*[180] at the counterpoise (CP) corrected[181] MP2 level was performed (aDZ=aug-cc-pVDZ). Using CP corrected values, the D/T extrapolation is more robust and expected to approach the CBS limit faster. Unfortunately, a HF calculation using the larger aug-cc-pVQZ basis set is not practicable due to near linear dependencies (for a similar problem already encountered with a TZVP basis set in solids see Ref. 182). Given that the association hardly changes the geometry of the individual components, the deformation contribution has been neglected in the CP treatment.

Based on this extrapolation, the association energy at the DLPNO-CCSD(T)/δCBS was obtained from a focal point analysis according to Eq. 3.3:

$$\begin{aligned}\Delta E_{\text{DLPNO-CCSD(T)}/\delta\text{CBS}} =& \Delta E_{\text{MP2/CBS-CP}} \\ &+ \left(\Delta E^{\text{corr.}}_{\text{DLPNO-CCSD(T)/aTZ}} - \Delta E^{\text{corr.}}_{\text{MP2/aTZ-CP}}\right).\end{aligned} \tag{3.3}$$

Here $\Delta E_{\text{MP2/CBS-CP}}$ is the association energy at the CP corrected aDZ/aTZ-extrapolated

MP2 level. For the association reaction, $\Delta E^{\text{corr.}}_{\text{MP2/aTZ-CP}}$ refers to the difference in the MP2 correlation energies with the aTZ basis set. $\Delta E^{\text{corr.}}_{\text{DLPNO-CCSD(T)/aTZ}}$ is the respective difference in the non-CP corrected DLPNO-CCSD(T)/aTZ correlation energies. This scheme was employed since BSSE is not well defined for the local correlation treatment of DLPNO-CCSD(T). It shall be pointed out that the underlying HF part of $\Delta E_{\text{DLPNO-CCSD(T)}/\delta\text{CBS}}$, for which BSSE is well defined, is counterpoise-corrected as this is contained within $\Delta E_{\text{MP2/CBS-CP}}$. By adding the difference written in parentheses in Eq. 3.3, we corrected the correlation contribution of the MP2/CBS-CP energy by DLPNO-CCSD(T), assuming that when combined with aTZ, the BSSE of DLPNO-CCSD(T) is less pronounced than in canonical MP2. The discrepancy between the CP corrected MP2 association energy with aTZ and the extrapolated value (rounded to the nearest larger multiple of 0.5 kcal mol^{-1}) will be used as an estimate for the basis set errors in $\Delta E_{\text{DLPNO-CCSD(T)/aTZ}}$. Adding to this 0.5 kcal mol^{-1} to take into account possible locality errors, this yields the error estimate of $\Delta E_{\text{DLPNO-CCSD(T)}/\delta\text{CBS}}$.

The above mentioned RI-J approximation[76,166] has been used in all SCF treatments employing the def2-QZVP basis set. Calculation of two-electron integrals in MP2 (also as part of the B2PLYP functional) were sped up by exploiting density fitting (dubbed RI-C).[76,183] In the correlation part of (SCS)-MP2 and DLPNO-CCSD(T), the core orbitals were kept frozen. The ORCA suite of programs (version 3.0) was used for the M06, B3LYP-NL as well as the MP2/aDZ, MP2/aTZ and DLPNO-CCSD(T)/aTZ calculations.[48] All other calculations were done with TURBOMOLE.[164a,165] In the hybrid DFT calculations with ORCA, i.e., for M06 and B3LYP-NL, the exchange-correlation energy was evaluated using *grid7*. Furthermore, calculations of exchange integrals were accelerated using the chain-of-sphere approximation (COSX) with *gridx5*.[184,185]

In all single-point calculations, the SCF convergence criterion was set to $10^{-7}\,E_h$. In all TURBOMOLE calculations, C_3 point group symmetry of each system studied here (D_3 for B(C_6F_5)$_3$ (**3**)) has been exploited. As an error estimate for the employed dispersion-corrected density functionals, HF as well as (SCS)-MP2, we used the weighted total mean absolute deviation from Ref. 73 (Ref. 172 for B3LYP-NL) for non-covalent interactions rounded to the nearest larger multiple of 0.5 kcal mol^{-1}. The differences in the association energies obtained from plain (i.e., without dispersion correction) B3LYP/def2-QZVP in TURBOMOLE (RI-J) and ORCA (RI-J-COSX) were in good agreement (maximum deviation of 0.16 kcal mol^{-1} for FLP **1a**) and thus, we consider the introduced errors by COSX to be negligible in the present study.

3.3.3. Ro-Vibrational Contributions

The thermostatistical contribution to the free energy was calculated by a modified ideal gas, rigid-rotor, harmonic-oscillator (RRHO) approach.[42] Therein, entropy contributions arising from vibrational frequencies smaller than about 100 cm^{-1} are obtained from an interpolated scheme that treats them partially as a free rotor and a harmonic oscillator. This has proven to improve results for thermochemistry, particularly for supramolecular complexes where such small vibrational frequencies are common.[42] The elaborate calculation of the required harmonic frequencies

is carried out with an efficient, corrected small basis set HF approach, termed HF-3c,[169] on geometries re-optimized at this level of theory. For the frequencies, we use the scaling factor of 0.86 as recommended in Ref. 169. Since the (scaled) frequencies are relatively insensitive to the chosen method,[42,186,187] this should yield sufficiently accurate results within about $\pm 1\,\mathrm{kcal\ mol^{-1}}$ for ΔG^{T}_{RRHO}.

3.3.4. Solvent Contributions

Calculating accurate solvation free energies for large systems is still a challenging task. Typically employed methods based on continuum solvation models are the so-called COSMO-RS[188] and SMD approaches.[189] The solvation contribution to the free energy is generally considered to be less accurate compared to ΔG^{T}_{RRHO}[42,141] although they are in the present case numerically much smaller. Here we applied the COSMO-RS(2012) approach which uses the conductor-like-screening model (COSMO)[190] (based on BP86/TZVP[99,101,102,162] data) as implemented in the COSMOtherm program.[191]

For further comparison, we have applied the COSMO-based SMD implementation within the development version of ORCA[48,189] at the BP86/def2-TZVP[99,101,102,162,163] level.
From experience and given that SMD was originally developed for the polarizable continuum model,[51,189] we consider the COSMO-RS values to be more reliable. However, we will use the difference between both corrections (rounded to the nearest larger multiple of $0.5\,\mathrm{kcal\ mol^{-1}}$) as error estimates for $\Delta\delta G^{298\,\mathrm{K}}_{\mathrm{solv.}}(\mathrm{bz})$.

3.4. Results and Discussion

3.4.1. Association of $B(C_6F_5)_3PMes_3$

The NMR experiments by Rocchigiani *et al.* do not indicate a preferential occurrence of the **1a** or **1b** form.[45] Instead their results are compatible with a rather fluxional system with no well-defined intermolecular orientation. Hence to estimate the "experimental" association energy $\Delta E_{\mathrm{exptl.}}$ according to Eq. 3.2, the same experimental value $\Delta G^{298\,\mathrm{K}}_{\mathrm{exptl.}}(\mathrm{bz}) = +0.4\,\mathrm{kcal\ mol^{-1}}$ has been used for both structures **1a** and **1b** in the back correction scheme. The two structures represent models with "extreme" orientations of the B and P atoms, respectively.

Table 3.1 contains the ΔE values for the association reaction of $B(C_6F_5)_3$ and $PMes_3$ to **1a** and **1b** obtained at different levels of theory along with the back-corrected experimental value.

From the results of plain and dispersion-corrected HF and B3LYP, it is obvious that the formation of FLP **1** in both orientations **a** and **b** is mainly driven by London dispersion forces. Without dispersion correction, both methods compute an unrealistic, endothermic formation in the gas phase of at least $+4.8\,\mathrm{kcal\ mol^{-1}}$.

On the other hand, the experimental values suggest an exothermic reaction with $\Delta E < -10\,\mathrm{kcal\ mol^{-1}}$. Applying dispersion corrections in an atom-pairwise (-D3) or a density-dependent (-NL) flavor to the B3LYP functional improves the results leading to agreement with

Table 3.1.: Theoretical and "experimental" ΔE for the association of **1a** and **1b** in kcal mol^{-1}. Unless noted otherwise, all theoretical results are obtained with the def2-QZVP basis set.

Method	**1a**	**1b**	estimated error[d]
exptl.[a]	-13.8 ± 2.5[b]	-13.2 ± 3.0[c]	
B3LYP	5.5	4.8	n.a.
B3LYP-D3	−10.5	−10.7	±1.5
B3LYP-NL	−11.7	−12.2	±1.5[e]
M06	−6.3	−6.1	n.a.
M06-D3(0)	−13.3	−13.4	±1.0
PW6B95-D3	−9.8	−10.9	±1.0
PBE0-D3	−9.7	−10.2	±1.5
B2PLYP-D3	−11.6	−11.7	±1.0
HF	7.7	6.0	±7.0
HF-D3	−10.2	−10.6	±1.0
MP2	−14.9	−14.6	±1.0
SCS-MP2	−10.3	−10.3	±1.5
DLPNO-CCSD(T)/aTZ	−15.0	−15.9	n.a.
MP2/CBS(aDZ/aTZ)-CP	−13.4	−12.9	n.a.
DLPNO-CCSD(T)/δCBS	−15.4	−15.0	±1.0[f]

[a] Back-corrected from $\Delta G^{298\,\mathrm{K}}_{\mathrm{exptl.}}(\mathrm{bz}) = +0.4\,\mathrm{kcal\ mol^{-1}}$ [45] using Eq. 3.2.
[b] $\Delta G^{298\,\mathrm{K}}_{RRHO} = 14.6 \pm 1.0\,\mathrm{kcal\ mol^{-1}}$, $\Delta\delta G^{298\,\mathrm{K}}_{\mathrm{solv.}}(\mathrm{bz}) = -0.4 \pm 1.5\,\mathrm{kcal\ mol^{-1}}$. For comparison: $\Delta\delta G^{298\,\mathrm{K}}_{\mathrm{solv.}}(\mathrm{bz})$ from SMD is $+0.9\,\mathrm{kcal\ mol^{-1}}$.
[c] $\Delta G^{298\,\mathrm{K}}_{RRHO} = 14.8 \pm 1.0\,\mathrm{kcal\ mol^{-1}}$, $\Delta\delta G^{298\,\mathrm{K}}_{\mathrm{solv.}}(\mathrm{bz}) = -1.2 \pm 2.0\,\mathrm{kcal\ mol^{-1}}$. For comparison: $\Delta\delta G^{298\,\mathrm{K}}_{\mathrm{solv.}}(\mathrm{bz})$ from SMD is $+0.8\,\mathrm{kcal\ mol^{-1}}$.
[d] Estimated error for non-covalent interactions from Ref. 73.
[e] Estimated error for non-covalent interactions from Ref. 172.
[f] Estimated error from discrepancy according to Section 3.3.
n.a. = not attempted.

experiment within the estimated error bars. In the association reaction of FLP **1** and FLP **2** (*vide infra*), neither the Lewis base nor the Lewis acid undergo large conformational changes or electronic reorganization. Therefore, it is mainly the dispersion interactions that make up for the largest part of the overall association energy. This is the reason why even dispersion-corrected HF provides good results for these reactions. Nevertheless, the latter should not be the method of choice if conformational changes and bond breaking or formation occur as it does not explicitly account for generally important short- and medium-range correlation effects.

M06, which was constructed to also account for medium-range correlation effects, describes the reaction as being exothermic, but still quantitatively wrong. Including long-range correlation by means of the D3(0) correction accounts for the missing London contributions to the association energy, which seem to be crucial in these "flexibly" bound FLPs. Methods that have previously proven to be quite reliable in thermochemistry applications (i.e., dispersion-corrected

density functionals, SCS-MP2),[73,172,174] yield association energies that lie reasonably within (or negligibly above as, e.g., for **1a** with PW6B95-D3) the upper error range of the experimental value.

MP2, which generally overestimates π-π interactions, yields association energies that also agree with the experiment within the error range. Yet, it shall be noted that the system is overbound compared to the other methods applied with this AO basis set. Further increase of the basis will make the binding in MP2 even stronger (roughly estimated by $1-2$ kcal mol^{-1}) thereby increasing the deviation from the reference value, while the opposite would hold for SCS-MP2. Notably, in particular the methods without a known tendency to overbind π-π stacking, are systematically closer to the upper error range. This indicates that in this association reaction, the errors in $\Delta G^{298\,\mathrm{K}}_{RRHO}$ and $\Delta\delta G^{298\,\mathrm{K}}_{\mathrm{solv.}}$(bz) do not cancel each other but may add up instead. Tentatively from the behavior of the electronic structure methods, we would favor a value more at the lower side of the experimental ΔE value. Being the highest-level wave function method applied, the DLPNO-CCSD(T)/δCBS results match nicely (within the error ranges) with the results of the other well-performing QC methods, as well as with the back-corrected experimental value. Yet, the somewhat larger association energies might suggest some residual BSSE contained in the correlation part of DLPNO-CCSD(T)/δCBS*.

In agreement to the other QC methods, the association energies of both orientations **1a** and **1b** are predicted to be identical within < 1 kcal mol^{-1}. The good performance of the recently proposed DLPNO-CCSD(T) methodology is in agreement with recently reported results for an organometallic reaction.[141]

Comparison of the computed association energies of **1a** and **1b** shows that with any of the employed QC methods, there is no orientational preference. This is in agreement with the recently obtained data from NOESY experiments[45] and explained by the fact that the association is predominantly London dispersion-driven. This also demonstrates that "frustration" in this sterically hindered system is almost complete and no significant directional P-B interaction occurs. To achieve agreement with the experimental association energy, correctly accounting for London dispersion interactions, in particular for atomic separations > 5 Å (*vide infra*), is crucial which becomes obvious from comparing the pure HF, B3LYP and M06 results to their dispersion-corrected counterparts.

3.4.2. Comparing the Association of B(C$_6$F$_5$)$_3$PMes$_3$ and B(C$_6$F$_5$)$_3$P*t*Bu$_3$

A similar, inter-molecular FLP that has been used in hydrogen activation is B(C$_6$F$_5$)$_3$P*t*Bu$_3$ (**2**).[44] FLP **1** and **2** differ only in their Lewis base. Rocchigiani *et al.* performed NOESY experiments also on FLP **2** which indicate a more directive association.[45] Nevertheless, no quantitative data for the association of this system exists and hence, no comparison with experimental (free) energies is possible.

*Recent DLPNO-CCSD(T) studies[130,192] suggest that the DLPNO-CCSD(T)/δCBS estimate in the present work is too large in magnitude by about 4–5 kcal mol^{-1} for ΔE of **1a**. This overestimation could be due to BSSE and may contribute to the residual deviation from the DFT-D3 results.

Table 3.2.: Association energies (free energies in benzene solution in parentheses) of **2a** and **2b** in kcal mol^{-1}. Unless noted otherwise, all theoretical results are obtained with the def2-QZVP basis set.

Method	**2a**	**2b**	estimated error[c]
$\Delta G^{298\,\mathrm{K}}_{RRHO}$ [a]	14.3	13.5	±1.0
$\Delta\delta G^{298\,\mathrm{K}}_{\mathrm{solv.}}(\mathrm{bz})$[b]	-0.9 ± 2.0	-2.2 ± 3.0	
B3LYP	2.8	2.7	n.a.
B3LYP-D3	−11.2 (2.3)	−7.7 (3.6)	±1.5
B3LYP-NL	−12.7 (0.7)	−8.7 (2.6)	±1.5[d]
M06-D3(0)	−13.7 (−0.2)	−8.8 (2.4)	±1.0
PW6B95-D3	−12.0 (1.4)	−7.6 (3.7)	±1.0
PBE0-D3	−10.5 (2.9)	−7.3 (4.0)	±1.5
B2PLYP-D3	−11.9 (1.6)	−8.2 (3.1)	±1.0
HF-D3	−11.3 (2.1)	−7.9 (3.4)	±1.0
MP2	−14.1 (−0.6)	−9.8 (1.5)	±1.0
SCS-MP2	−10.3 (3.2)	−7.1 (4.2)	±1.5
DLPNO-CCSD(T)/aTZ	−13.4 (0.0)	−10.5 (0.8)	n.a.
MP2/CBS(aDZ/aTZ)-CP	−13.0 (0.4)	−8.7 (2.6)	n.a.
DLPNO-CCSD(T)/δCBS	−13.1 (0.3)	−9.9 (1.4)	±1.0[e]

[a] Harmonic frequencies are obtained at the HF-3c level.
[b] $\Delta\delta G^{298\,\mathrm{K}}_{\mathrm{solv.}}(\mathrm{bz})$ is calculated by the COSMO-RS solvation model.
For comparison: $\Delta\delta G^{298\,\mathrm{K}}_{\mathrm{solv.}}(\mathrm{bz})$ from SMD is +1.0 kcal mol^{-1}.
For comparison: $\Delta\delta G^{298\,\mathrm{K}}_{\mathrm{solv.}}(\mathrm{bz})$ from SMD is +0.4 kcal mol^{-1}.
[c] Estimated error for non-covalent interactions from Ref. 73.
[d] Estimated error for non-covalent interactions from Ref. 172.
[e] Estimated error from discrepancy according to Section 3.3.
n.a. = not attempted.

In this section, we apply the same levels of theory as in Section 3.4.1 (apart from plain HF and M06) to calculate the association (free) energy of FLP **2** in both orientations **a** and **b** (see Figure 3.2). The results are listed in Table 3.2.

All methods predict the orientation in FLP **2a** to be more favorable than the one in **2b** by approximately 3−5 kcal mol^{-1} (1−2 kcal mol^{-1} in free energies). This is different compared to the association of FLP **1** indicating that in the case of **2**, the process is slightly less dispersion driven and some P-B bonding interaction is present (this is discussed in Ref. 193 in terms of frontier orbitals). Given that the corrections to the free energy are of similar magnitude for **1** and **2**, the ΔE values are discussed in the following. The preferred association to complex **2a** is about as exothermic as is the association to **1** (either **a** or **b**). The DLPNO-CCSD(T)/δCBS values for the association energy agree well with the other QC methods exhibiting a similar difference between **2a** and **2b**. Furthermore, the association to **2a** is described as being about

as exothermic as the formation of **1a**/**1b** which is also in agreement with the other QC methods[†].

Therefore, also the computationally less expensive, dispersion corrected QC methods have proven to work quite well in describing the association of this FLP. Furthermore, the good mutual agreement of the conceptually different NL(VV10) and D3 dispersion correction schemes is noteworthy. *Vice versa*, DLPNO-CCSD(T) provides high-level reference energies for a system that is beyond the applicability of canonical CCSD(T).

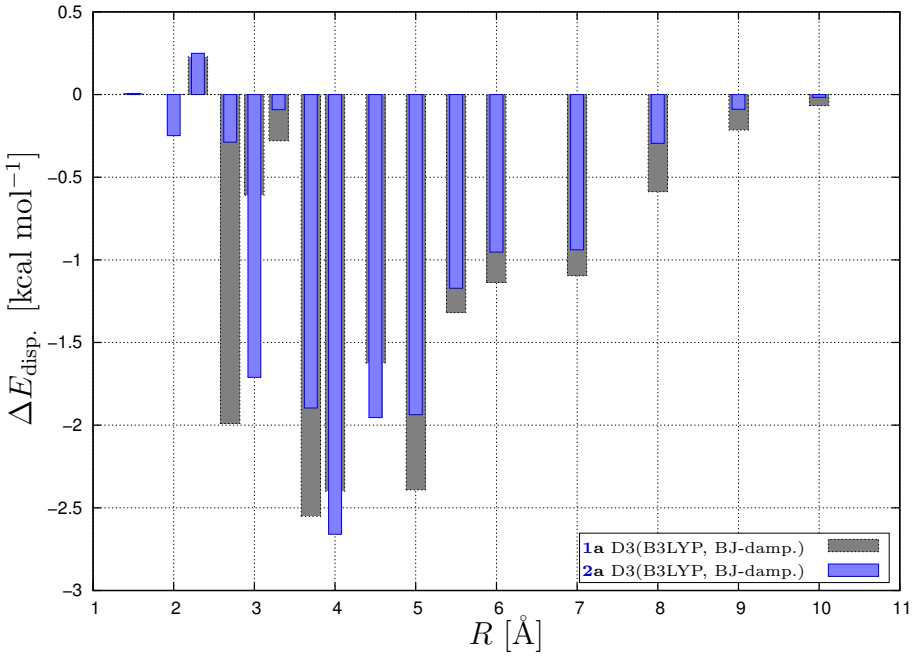


Figure 3.3.: Comparison of the distance-dependent, London dispersion contributions to the association energy of **1a** and **2a**. The dispersion energies were calculated using the atom pair-wise D3 method (B3LYP parametrization for Becke-Johnson damping) and distributed to the different distances.

In Figure 3.3, the contributions of the dispersion interaction to the association energy of the complexes **1a** and **2a** are plotted against the inter-atomic distance regime they originate from. While in the range up to 5 Å, the London dispersion interactions are not systematically larger for either of the two, it can clearly be seen that for FLP **1a**, the long-range regime (> 5 Å) contributes more to the association energy, summing up to additional -1.5 kcal mol^{-1} compared to **2a**. The total contribution to the association energy from this part is approximately -7 kcal mol^{-1}. From Table 3.1, it is evident that this more or less corresponds to the contribution that is missing in the plain M06 functional (which only accounts for the short- and medium-range correlation effects). These observations also suggest the use of an augmented basis in the DLPNO-CCSD(T) calculations. A test calculation on the association energy of **1a** at the DLPNO-CCSD(T)/def2-TZVP level results in association energies severely underestimated in

[†]Recent DLPNO-CCSD(T) studies[130,192] suggest that the DLPNO-CCSD(T)/δCBS estimate in the present work is too large in magnitude by about 2–3 kcal mol^{-1} for ΔE of **2a**. This overestimation could be due to BSSE and may contribute to the residual deviation from the DFT-D3 results.

magnitude ($-5.4\,\mathrm{kcal\ mol^{-1}}$).
From Figure 3.3, one can understand that this def2-TZVP basis set does not provide enough diffuse AOs to allow an asymptotically correct description of London dispersion interactions that make up a significant part of the association energy.

With respect to the performance of routinely employed QC methods to study FLPs,[150] we emphasize that these interactions were correctly described by either the NL or the D3 dispersion-corrected DFT methods at lower cost. In particular, this allowed the use of a polarized quadruple-zeta basis sets but at the same time, the delicate handling of diffuse basis functions was avoided in such realistic systems without sacrificing the accuracy in the long-range regime.

3.4.3. Structural Implication for the Reactivity

The distance between the phosphorus and the boron atoms in complexes **1a** and **2a** are significantly different (4.74 Å compared to 4.22 Å). Taking the phosphorus-boron separation as a measure, the FLP **2a** can be considered as being less "frustrated" than **1a**, although we note in passing that P-B distances > 4 Å are rather large in absolute values compared to those in similar but non-frustrated Lewis pairs ($2 - 2.5$ Å, see Ref. 194).

A possible implication of our results might be a different activity of the two systems in reactions with H_2 or other small molecules. In order to cleave the H_2 bond, the phosphorus and boron atoms need to be oriented towards each other,[148] i.e., as is the case in the **1a** and **2a**. For this orientation, both FLPs show the same, slightly positive association free energy in solution. While the **b** orientation of FLP **1** is just as stable as the **a** oriented one, this is not the case for **2**. All of the well-suited QC methods yield the same relative stabilities therefore drawing a uniform picture of a more directed association of FLP **2** in contrast to a more random orientation in **1**.

If a hydrogen molecule approaches a FLP complex **2**, it is more likely that it is present in the **a** orientation enabling a fast reaction with H_2. In the case of hydrogen approaching complex **1**, the Lewis base and Lewis acid of this FLP may be oriented to each other in any possible way (with the two extrema **1a** and **1b**). Thus, the probability of a reaction with hydrogen would be lower compared to FLP **2**. To our knowledge, no quantitative experiments were carried out so far that studied the different reaction rates of both systems. Some qualitative support of our conclusions can be taken from the different yields in the reaction of H_2 with either FLP **1** or **2**. Under identical reaction conditions, 74% of the former reacted to the respective zwitterion, while the reaction with the latter yielded 90% of the hydrogenated species.[144]

3.5. Conclusion

In this work, we were able to reproduce the experimentally determined association free energy of $B(C_6F_5)_3/PMes_3$ (**1**) in benzene solution using state-of-the-art quantum chemistry methods from the realms of density functional and wave function theory. According to Eq. 3.2, an

"experimental" value for the gas phase association energy was obtained from back-correcting the experimentally obtained association free energy from Ref. 45. High-level CCSD(T) based methods (the current quantum chemistry "gold standard") extrapolated to the basis set limit were applied to these large FLP systems for the first time and the results agree within the estimated error limits with the experimental value.

All electronic structure methods which describe London dispersion interactions asymptotically correct, gave very similar results for the association energies of $B(C_6F_5)_3/PMes_3$ (**1**) and good agreement with the experimental value. Additionally, these QC methods compute the complexes **1a** and **1b** to be equally stable which is in agreement with recent NOESY experiments.[45]

The association of $B(C_6F_5)_3/P\mathit{t}Bu_3$ (**2**) was studied as well. While the association energy to form complex **2a** appears to be similar in magnitude to the association energy of $B(C_6F_5)_3/PMes_3$, this is different to the formation of the oppositely oriented complex **2b** which is disfavored by a few kcal mol^{-1}. This implies a more directive complexation involving some covalent P-B bonding interaction in the case of $B(C_6F_5)_3/P\mathit{t}Bu_3$ than for $B(C_6F_5)_3/PMes_3$ where London dispersion interactions are the major driving force for binding. A possible consequence might be a different activity in the reaction with small substrate molecules like H_2. Analyzing the validity of this hypothesis, e.g., by molecular dynamics simulations will be the subject of upcoming efforts.

It was also pointed out that for wave function-based local correlation methods like DLPNO-CCSD(T), the use of a sufficiently large AO basis set including diffuse functions is necessary for a correct description of the association. It is encouraging that in DFT calculations of such systems, the correct physical behavior is captured at lower computational cost by the additive dispersion corrections allowing replacement of the technically inconvenient diffuse basis sets by semi-diffuse but larger basis sets such as def2-QZVP. Nevertheless, the DLPNO-CCSD(T) method has proven to be applicable to chemically realistic problems and therefore seems to qualify as a reference method when systems are beyond the applicability of canonical CCSD(T).

Acknowledgments

The authors like to express their gratitude to Prof. Dr. Douglas W. Stephan and Prof. Dr. Gerhard Erker for their fruitful collaborations over the years and for "opening" this interesting research area also for quantum chemistry.

Part III.

Development and Application of Simplified Time-Dependent Density Functional Theory

In Part II, dispersion-corrected DFT was introduced and its good performance in the computation of non-covalent interaction energies, as demonstrated in Chapter 3, suggests its usefulness for the ranking of conformers in future electronic circular dichroism (ECD) studies of flexible systems.

The present part deals with the development and application of simplified methods for the fast computation of electronic absorption and circular dichroism spectra. In 2013, the simplified Tamm-Dancoff approximated time-dependent density functional theory (sTDA-DFT) approach was presented. Starting from a TDA-DFT treatment (see Section 2.2.6), the involved computation of two-electron integrals is approximated in a semiempirical fashion and the matrix dimensions are drastically reduced. As described in Refs. 46 and 195, these simplifications hardly affect the accuracy compared to regular TDA-DFT.

In Chapter 4, the sTDA-DFT method is used in a collaborative study with the Lützen group (University of Bonn) to compute ECD spectra for large metallosupramolecular complexes with more than 800 atoms. The computed spectra are composed of more than 5000 states and, along with the experimental spectra, provide additional support for the presence of the self-assembled structures, which have been suggested based on other analytical techniques.

The approximations of the sTDA-DFT method are transferred to a linear response time-dependent (LR-TD) DFT treatment in Chapter 5, affording the simplified time-dependent density functional theory (sTD-DFT) approach. While being computationally only slightly more involved compared to its Tamm-Dancoff approximated counterpart, sTD-DFT provides more accurate absorption and, in particular, ECD intensities than sTDA-DFT. In this respect, both approaches behave very similar to their parental methods.

ECD spectra of very delocalized π-systems are identified as worst cases for Tamm-Dancoff approximated methods in Chapter 6. (s)TDA-DFT and CIS methods compute ECD intensities of incorrect sign in the origin-independent velocity formalism and a full LR-TD treatment as in sTD-DFT is inevitable for these systems.

Due to the drastically reduced cost of the excited state treatment compared to a regular TD-DFT calculation, sTD-DFT is suitable to be applied on structures (or "snapshots") extracted from a molecular dynamics (MD) simulation. The latter may be used for conformational sampling and also to go beyond the commonly employed minimum structure approach. Such a treatment is conducted in Chapter 7 to compute the electronic absorption and ECD spectra of the [16]helicene and a substituted derivative. For the latter, the computed spectrum is compared to an experimental one. While the spectra computed on the minimum geometries show slight differences between [16]helicene and the derivative, the two spectra become almost identical in the sTD-DFT//MD treatment. The bias due to a single, possibly arbitrary structure is "washed out" in the latter treatment, which provides a more realistic simulation of the dynamic system in solution.

4. Enantiomerically Pure $[M_6L_{12}]$ or $[M_{12}L_{24}]$ Polyhedra from Flexible Bis(Pyridine) Ligands

Christoph Gütz,[*] Rainer Hovorka,[*] Christoph Klein,[*] Qian-Qian Jiang,[†] Christoph Bannwarth,[‡] Marianne Engeser,[*] Carsten Schmuck,[†] Wilfried Assenmacher,[§] Werner Mader,[§] Filip Topić,[¶] Kari Rissanen,[¶] Stefan Grimme,[‡] and Arne Lützen[*]

Received 4th of October 2013, Published online 22nd of January 2014

Reprinted (adapted) with permission from[‖]
Gütz, C.; Hovorka, R.; Klein, C.; Jiang, Q.-Q.; Bannwarth, C.; Engeser, M.; Schmuck, C.; Assenmacher, W.; Mader, W.; Topić, F.; Rissanen, K.; Grimme, S.; Lützen, A. *Angew. Chem. Int. Ed.* **2014**, *53*, 1693–1698.

DOI 10.1002/anie.201308651

Own manuscript contribution

- Performing all quantum chemical calculations
- Interpretation of the computed absorption and ECD spectra
- Writing the parts of the manuscript that deal with the quantum chemical calculations
- Writing the detailed description of the computational study (in Appendix A3)

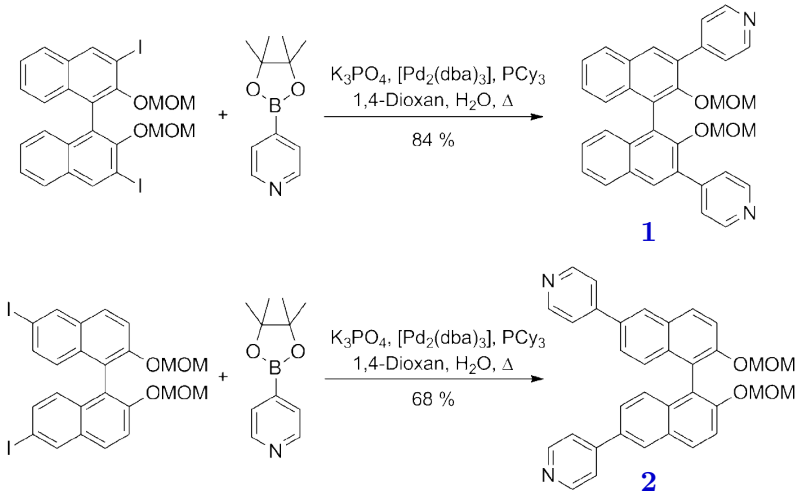
[*]Kekulé-Institut für Organische Chemie und Biochemie, Rheinische Friedrich-Wilhelms-Universität Bonn, Gerhard-Domagk-Straße 1, 53121 Bonn, Germany
[†]Institut für Organische Chemie, Universität Duisburg-Essen, Universitätsstraße 7, 45117 Essen, Germany
[‡]Mulliken Center for Theoretical Chemistry, Institut für Physikalische und Theoretische Chemie, Rheinische Friedrich-Wilhelms-Universität Bonn, Beringstraße 4, 53115 Bonn, Germany
[§]Institut für Anorganische Chemie, Rheinische Friedrich-Wilhelms-Universität Bonn, Römerstraße 1, 53117 Bonn, Germany
[¶]Department of Chemistry, Nanoscience Center, University of Jyväskylä, Finland
[‖]Permission requests to reuse material from this chapter should be directed to WILEY-VCH.

Abstract Coordination-driven self-assembly is one of the most powerful strategies to prepare nanometer-sized discrete (supra)molecular assemblies. Herein, we report on the use of two constitutionally isomeric BINOL-based bis(pyridine) ligands for this purpose. Upon coordination to Pd^{II} ions these self-assemble into enantiomerically pure *endo*- and *exo*-functionalized hexa- and dodecanuclear metallosupramolecular spheres with a chiral skeleton depending on the substitution pattern of the BINOL core. These aggregates were characterized by nuclear magnetic resonance (NMR), mass spectrometry (MS), dynamic light scattering (DLS), transmission electron microscopy (TEM), and electron energy loss spectroscopy (EELS), as well as ECD. Furthermore, experimental ECD data could be compared to those obtained from theoretical simulations using a simplified Tamm-Dancoff approximation to time-dependent DFT to rationalize the extraordinary high molar circular dichroisms. Despite the rotational freedom around the central aryl-aryl bond of these ligands, the self-assembly process happens completely selective in a "narcissistic" self-recognition manner.

4.1. Introduction

Coordination-driven self-assembly has proven to be one of the best strategies to prepare discrete (supra)molecular assemblies with dimensions in the 1–10 nm regime.[196–219] Usually these are highly symmetrical metal-organic polygons or polyhedra that are formed from highly directional bridging organic ligands and geometrically prefixed metal containing nodes. Among the overwhelming number of different combinations of metal centers and ligands that has been used for this purpose to date,[196–219] the combination of palladium and platinum metal centers and ligands containing pyridine groups has become one of the most successful and popular coordination motifs in supramolecular coordination chemistry.[220–227] The use of tetravalent Pd^{II} and Pt^{II} acceptors and rigid ditopic bridging N-donor ligands, for instance, has been demonstrated to give rise to three dimensional $[M_nL_{2n}]$ assemblies[228–246] with n ranging from two up to even twenty four in a relatively reliable manner following the general assumption that the smallest assembly forms in which all the coordination sites on the metal centers and the ligands are occupied (the maximum occupancy rule) that does not experience too much steric strain. According to the beautiful examples of Fujita and co-workers the outcome of the self-assembly process in terms of both composition and size of the closed 3D-assemblies critically depends on the bend angle of the bridging ligand.[242–246] However, this approach is not only interesting from a conceptual point of view but it is also very attractive for a number of applications because functionalized ligands can lead to *endo*- or *exo*-functionalized assemblies with interesting properties.[247–260]

What has not been established yet is the construction of a palladium- or platinum-containing metallosupramolecular assembly with six or more metal atoms and an inherently chiral ligand leading to a chiral spherical metallosupramolecular framework rather than a non-chiral sphere with chiral groups on the *endo*- or *exo*-side of the assembly.[250,254,256–259] Hence, we decided to prepare two constitutionally isomeric bis(pyridine) ligands **1** and **2** based on a chiral 2,2'-



Scheme 4.1: Synthesis of the optically pure ligands **1** and **2**. dba = dibenzylidene acetone, MOM = methoxymethyl, Cy = cyclohexyl.

dihydroxy1,1'-binaphthyl (BINOL) core, both in enantiomerically pure *P*- and *M*-forms and racemic forms, and study their self-assembly to metal–organic $[M_nL_{2n}]$ polyhedra upon coordination to $[Pd(CH_3CN)_4](BF_4)_2$. Knowing that the angle between the two pyridine nitrogen atoms is greater than or equal to 90° we expected the formation of metallosupramolecular assemblies that contain at least six metal atoms. However, BINOL derivatives can adopt several conformations due to the rather unrestricted rotation around the aryl–aryl bond. Thus, another aim of this study was to elucidate how this rotational freedom would affect the selectivity of the self-assembly process in terms of the assemblies' composition.

4.2. Methodological Overview

The synthesis of the chiral ligands starts from (*M*)- or (*P*)-3,3'- or -6,6'-diiodo-2,2'-bis(methoxymethoxy)-1,1'-binaphthyl, which were prepared according to literature methods.[261,262] From these compounds the target ligands could be obtained by Suzuki cross coupling in good yield (68–84%) (Scheme 4.1).[263]

The enantiomerically pure ligands were then mixed with $[Pd(CH_3CN)_4](BF_4)_2$ in a 2:1 ratio in dimethyl sulfoxide (DMSO) or acetonitrile and heated for 3 h at 80°C and the resulting oligonuclear complexes were examined by NMR and ECD spectroscopic, mass spectrometric, dynamic light scattering (DLS), and transition electron microscopic means as well as theoretical density functional theory (DFT) methods.

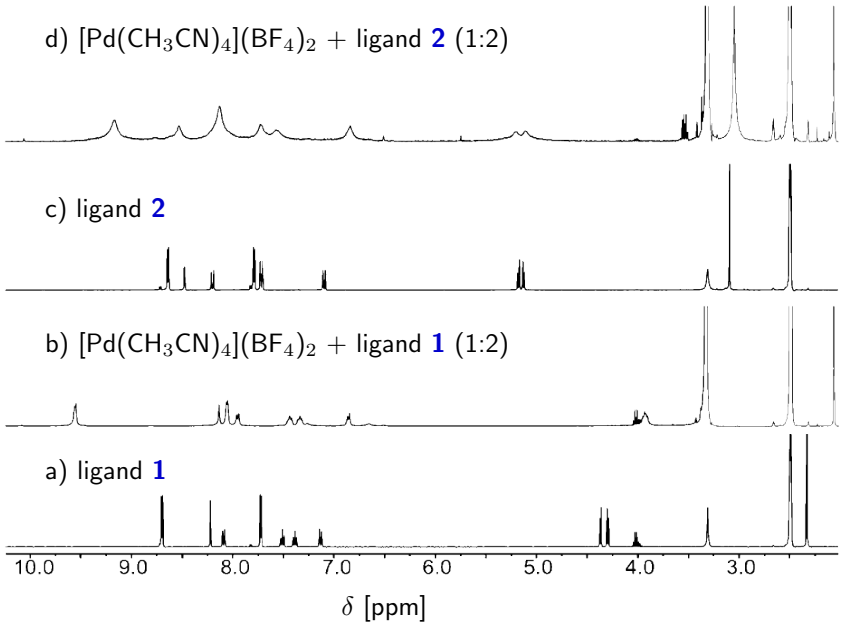


Figure 4.1.: ^{1}H NMR spectra (400.1 MHz in $[D_6]$DMSO at 293 K) of a) **1**, b) a 2:1 mixture of **1** and $[Pd(CH_3CN)_4](BF_4)_2$, c) **2**, and d) a 2:1 mixture of **2** and $[Pd(CH_3CN)_4](BF_4)_2$.

4.3. Results and Discussion

First we measured the ^{1}H NMR spectra (Figure 4.1) and observed a strongly shifted set of signals for both cases. The signals show the same symmetry as the free ligand indicating the formation of symmetrical coordination compounds. Although, signal broadening was observed in both cases, the broadening was rather small in case of the complex from ligand **1** indicating the formation of a discrete species. On comparison of the data with that of similar $[M_6L_{12}]$ or even larger assemblies, even the more severe signal broadening in the spectrum of ligand **2** complex is still in accordance with a discrete oligonuclear species.

To obtain more evidence for the postulated large assemblies we performed diffusion-ordered NMR spectroscopy, i.e., 2D-^{1}H-DOSY measurements and compared the results with the theoretically calculated size of the molecules. For the ligand **1** complex, we obtained diffusion coefficients $D = 6.00 \times 10^{-11}\,m^2\,s^{-1}$ in $[D_6]$DMSO (Figure 4.2) and $D = 4.29 \times 10^{-10}\,m^2\,s^{-1}$ in $[D_3]$acetonitrile (see Supporting Information*), which corresponds to an object with a radius of 13.0–15.1 Å according to the Stokes–Einstein equation. This size fits very well to the calculated dimensions of a truncated cube-shaped structure of a $[Pd_6(\mathbf{1})_{12}](BF_4)_{12}$ aggregate of $10.5 < r_{calcd.} < 16.2$ Å.†

*Supporting information for the quantum chemical study is included in Appendix A3. The complete supporting information is available online at http://dx.doi.org/10.1002/anie.201308651.

†Please note that the truncated cube-like shape of the $[Pd_6(\mathbf{1})_{12}]$ assembly deviates considerably from an ideal sphere. Hence, we gave the calculated distances between two opposite palladium atoms and between the two most-remote hydrogen atoms of opposite naphthyl units as minimum and maximum distances.

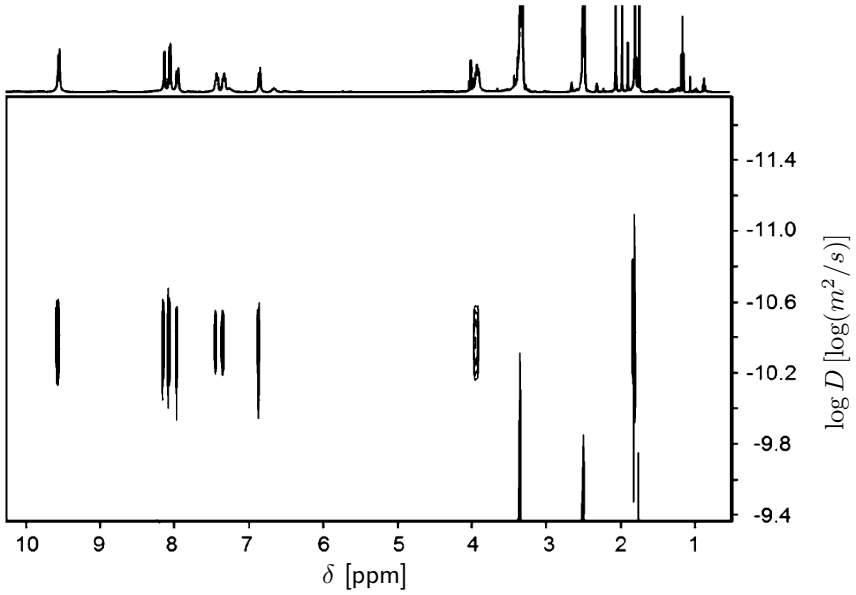


Figure 4.2.: 2D-^{1}H-DOSY-NMR (500.1 MHz in [D$_6$]DMSO at 293 K) of the Pd^{II} complex of ligand **1**.

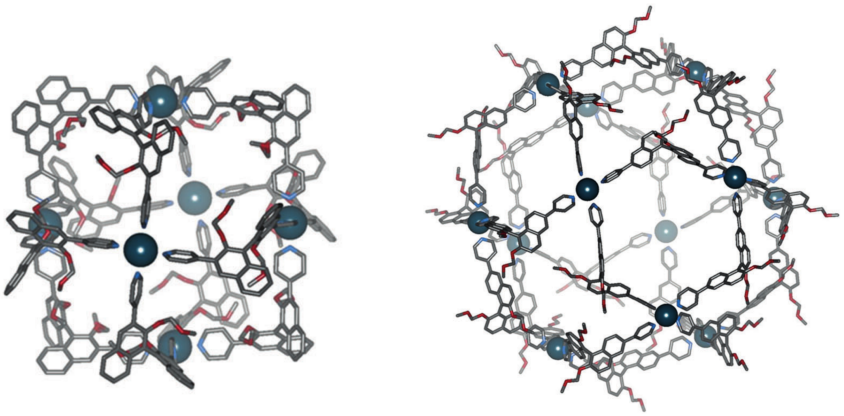
The DOSY spectra (see Supporting Information*) of the Pd^{II} complex of ligand **2**, however, gave considerably lower diffusion coefficients of $D = 3.29 \times 10^{-11}\mathrm{m}^2\,\mathrm{s}^{-1}$ in [D$_6$]DMSO and $D = 2.04 \times 10^{-10}\mathrm{m}^2\,\mathrm{s}^{-1}$ in [D$_3$]acetonitrile which both indicate a considerably larger object with a hydrodynamic radius of $r_H = 27.6$ Å. Such a large size matches the calculated size of an almost spherical [Pd$_{12}$(**2**)$_{24}$](BF$_4$)$_{24}$ aggregate of ligand **2** ($r_{\mathrm{calcd.}} = 26.7$ Å) extremely well. These two assemblies are depicted in Figure 4.3.

Both structures were optimized imposing O symmetry at the dispersion- and geometrically counterpoise-corrected DFT level using the PBE functional and the def2-SV(P) basis set.[104,133,134,163,264] The optimizations were performed with the TURBOMOLE suite[164a,165] and applying the COSMO solvation model ($\epsilon = 35.7$).[190] We will refer to this level of computation as PBE-D3-gCP/def2-SV(P) (see Supporting Information for details).

To get further support for the composition of our assemblies, electrospray ionization (ESI) mass spectra of the two complex solutions were recorded. Unfortunately, we did not succeed in obtaining an ESI mass spectrum under ambient conditions showing an ion of an intact metallo-supramolecular assembly of ligand **2** with accurately resolved isotopic pattern.[‡]

However for the complex of ligand **1**, we were able to get an ESI spectrum that clearly shows the pattern expected for a hexanuclear complex in six different charge states with different number of counterions, which confirms our NMR spectroscopic results (Figure 4.4). In fact, this is one of the very first ESI mass spectra of such a large kind of oligonuclear complexes that could

[‡]We observed a series of lower molecular-weight fragments of our assembly such as [Pd$_6$(**2**)$_{11}$]F$_6{}^{6+}$, [Pd$_7$(**2**)$_{12}$](BF$_4$)$_7{}^{7+}$, [Pd$_7$(**2**)$_{13}$]F$_7{}^{7+}$, [Pd$_8$(**2**)$_{13}$]F$_8{}^{8+}$, [Pd$_9$(**2**)$_{17}$]F$_9{}^{9+}$, or [Pd$_{11}$(**2**)$_{20}$]F$_{11}{}^{11+}$ (see Supporting Information*).



$[Pd_6\{(M)\text{-}\mathbf{1}\}_{12}]^{12+}$ $[Pd_{12}\{(P)\text{-}\mathbf{2}\}_{24}]^{24+}$

Figure 4.3.: DFT-optimized structures of the *O*-symmetric complexes $[Pd_6\{(M)\text{-}\mathbf{1}\}_{12}]^{12+}$ (left) and $[Pd_{12}\{(P)\text{-}\mathbf{2}\}_{24}]^{24+}$ (right; color code: petrol=Pd, red=O, blue=N, and gray=C; hydrogen atoms are omitted for clarity).

be recorded under ambient conditions without using cold-spray ionization.

Additional information about the size of nanoscopic objects in solution can be obtained by dynamic light scattering (DLS) experiments.[265] These DLS experiments confirmed the results of the DOSY NMR and ESI-MS studies for both complexes, and hence provide additional independent support of the size, and therefore, the composition of our assemblies (see Supporting Information*), especially because these techniques could also be used to correctly characterize a $[M_{12}L_{24}]$ aggregate from Fujita *et al.*[§]

We were also able to grow quite well formed crystals from our complexes but owing to the very large voids in the crystal filled with disordered anions and solvent molecules, the X-ray scattering power of the crystals was so weak that not even a unit cell could be obtained.[¶]

We also performed transmission electron microscopy (TEM) measurements to visualize the metallosupramolecular spheres. We prepared samples of the metallosupramolecular aggregates on a perforated carbon foil supported by a Cu grid by wetting the grid with solutions of the complexes. Quickly recorded TEM bright field images of the dried samples revealed particles with a weak contrast embedded in the thin film formed by the dried solvent (Figure 4.5). Given the fact that the metallosupramolecular spheres contain only 6 or 12 palladium atoms, respectively, we expected such low contrasting objects and their size of about 3 or 5 nm fits the size of the hexa- and dodecanuclear $[Pd_6(\mathbf{1})_{12}]$ and $[Pd_{12}(\mathbf{2})_{24}]$ complexes very well.

When the sample is exposed to the electron beam for a longer time the film breaks and does not protect the assemblies anymore. We observe a fast reduction of the palladium(II) complexes

[§]Note that we referenced our analytical methodology by characterizing one of Fujita's $[M_{12}L_{24}]$ metallosupramolecular assemblies (see Ref. 243) whose composition has unambiguously been assigned by different analytical tools. In fact our DOSY, ESI-MS, and DLS approach perfectly matched these results (see Supporting Information*).

[¶] Note that the unknown amount of solvent molecules that fill the voids of our spheres also prevented adequate elemental analysis data from being obtained.

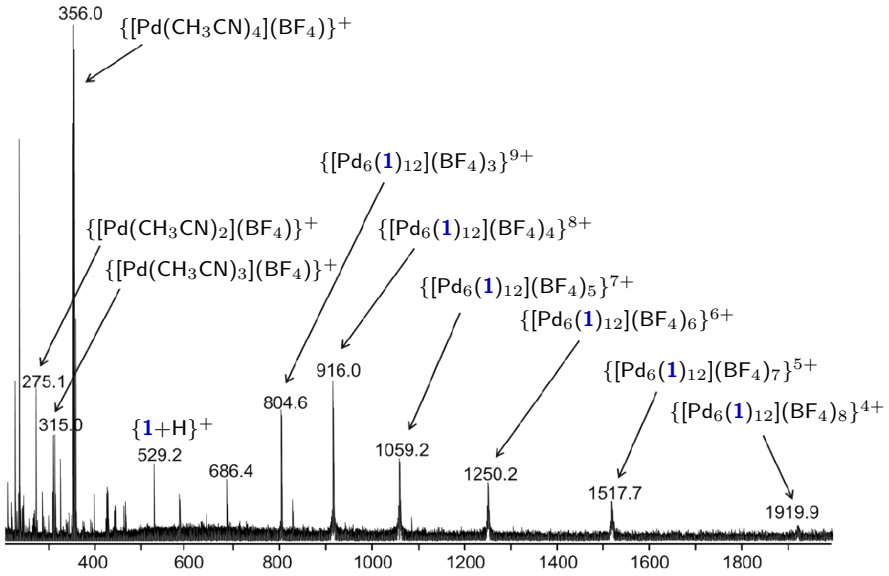


Figure 4.4.: Positive ESI-MS spectrum of the $[Pd_6(\mathbf{1})_{12}](BF_4)_{12}$ complex solution in acetonitrile.

to crystalline palladium(0) nanoparticles in a couple of minutes (see Supporting Information*). This effect has also been described for other palladium complexes under TEM conditions.[266]

To confirm that these are really the palladium complexes we measured EELS spectra, which showed the characteristic ionization edges for palladium, boron, and fluorine as well as oxygen, nitrogen and omnipresent carbon (see Supporting Information*).

All of the results obtained from the different and complementary analytical techniques clearly demonstrate the selective formation of two different types of metallosupramolecular spheres by using the constitutional isomeric ligands **1** and **2**. This result is especially interesting because both ligands contain a rather freely rotatable aryl–aryl bond that allows both compounds to markedly change the relative orientation of the two naphthyl groups. Nevertheless, it seems that the 6,6'-disubstituted ligand **2** adopts a conformation with a larger nominal bend angle between the coordinating pyridine moieties than the more concave one of its 3,3'-disubstituted isomer **1**. Clearly, the difference has to be substantial since we observe a completely size-selective assembly in both cases. To study this phenomenon further we studied the behavior of a 1:1:1 mixture of $[Pd(CH_3CN)_4](BF_4)_2$, (*P*)-**1**, and (*P*)-**2**.[‖] Interestingly, the ^{1}H NMR spectrum of the equilibrated mixture (3 h at 80°C in [D$_6$]DMSO) revealed the completely selective formation of the two homoleptic complexes $[Pd_6(\mathbf{1})_{12}](BF_4)_{12}$ and $[Pd_{12}(\mathbf{2})_{24}](BF_4)_{24}$ but no mixed species. Hence, the self-assembly process proceeds with complete self-sorting in a "narcissistic" self-recognition manner.[267–269]

[‖] We also studied the stereoselectivity of the self-assembly process by using the ligands in racemic form. However, the processes were not diastereoselective in a self-sorting manner although the overall composition of the aggregates did not change.

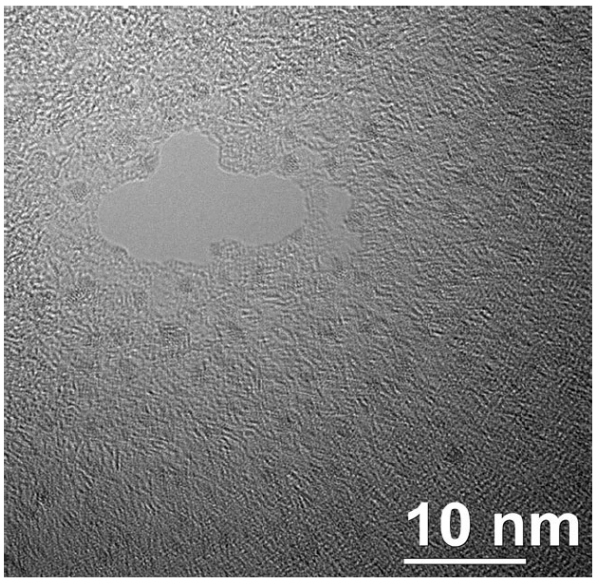


Figure 4.5.: TEM image of $[Pd_6(\mathbf{1})_{12}]$ complexes.

Finally, we turned our attention to ECD spectroscopy (Figure 4.6 and Supporting Information*) since our assemblies are the first metallosupramolecular spheres of this kind with an enantiomerically pure *O*-symmetric skeleton. Besides some new bands that result from the formation of metal–ligand bonds, the intensities of the bands are considerably larger than the sum of 12 or 24 bands of the pure ligands. To our knowledge, the ECD spectra of these assemblies show the highest $\Delta\varepsilon$ values (up to 4000 $\mathrm{L\,mol^{-1}\,cm^{-1}}$) recorded for well-defined molecular entities. Clearly, the relative orientation of the chromophores changes and gets more ordered in the metallosupramolecular complexes which is another excellent indication for the formation of structurally very-well defined assemblies. Interestingly, this effect is more pronounced in the smaller assembly indicating that the conformational change of the binaphthyl core seems to be greater in **1** than in **2**. This assumption is also supported by the comparison of the calculated dihedral angles between the two naphthyl groups in the free ligands **1** and **2** and in their metal complexes (see Supporting Information).

For further confirmation of our structural assignment, we performed a computational study at the hybrid DFT level to calculate the ECD spectra of the $[Pd_6\{(M)\text{-}\mathbf{1}\}_{12}]^{12+}$ and $[Pd_12\{(P)\text{-}\mathbf{2}\}_{24}]^{24+}$ ions. These highly positively charged complexes contain 822 and 1644 atoms, respectively, and hence, represent the largest systems of this kind that have been modeled at a high *first principles* level of theory up to date. Even though the huge size of the molecules forced us to make certain approximations in the DFT treatments (see Supporting Information), the spectra calculated using a simplified Tamm–Dancoff approximation to time-dependent DFT (sTDA-DFT)[46] are in very good agreement both in terms of position and absolute intensity of the individual bands with the experimental ones (see Figure 4.6b). The observed blue-shift for bands below 250 nm is attributed to various effects (used BHLYP hybrid functional, incom-

plete basis set and solvation treatment; see Supporting Information for a thorough discussion also of geometry and conformational effects and UV/CD spectra). However, the deviations are systematic and acceptable concerning the size and complexity of the aggregates. These sTDA-DFT calculations set new standards in the field of theoretical ECD spectroscopy what can be investigated nowadays. They provide further support for the aggregates composition and their structure and the self-assembly of just one enantiomer.

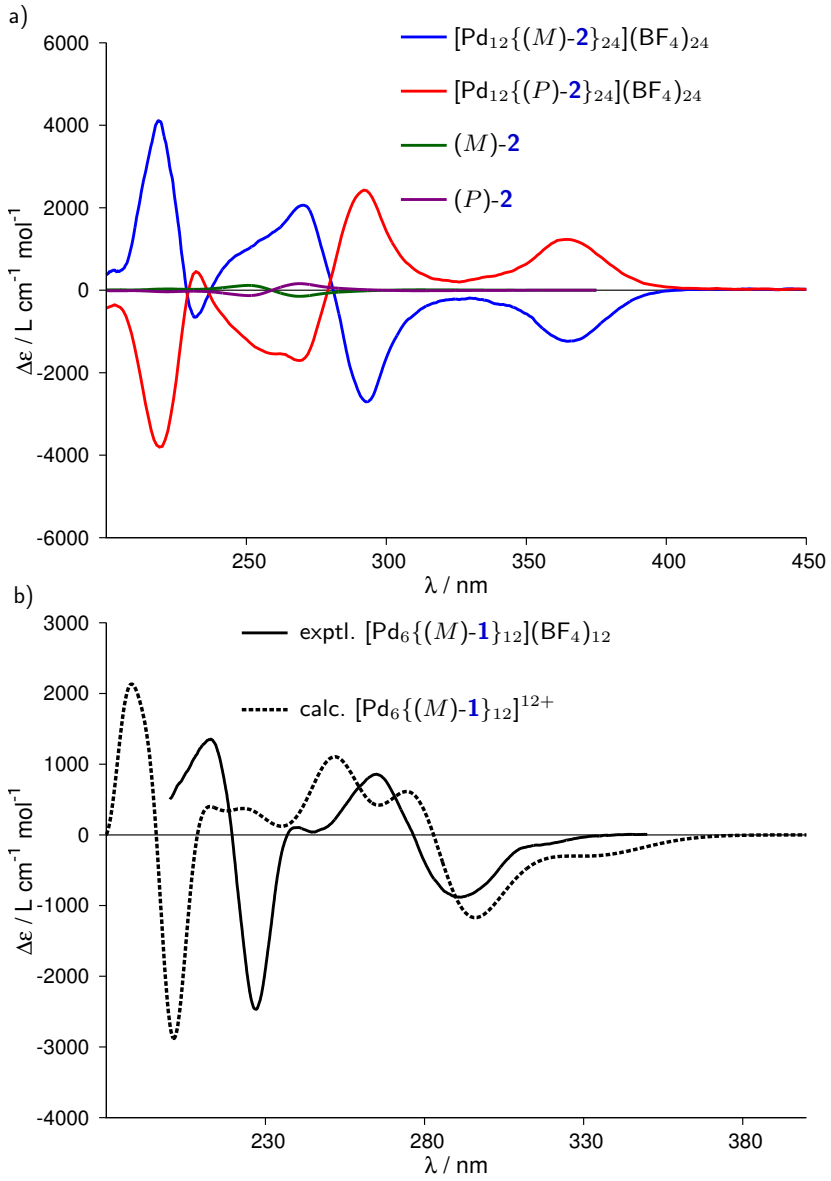


Figure 4.6.: a) ECD spectra of the enantiopure complexes of ligands **2** compared to the ones of (*M*)- and (*P*)-**2** in acetonitrile. (for a better view of the ligand spectra see Supporting Information*). b) Experimental and calculated ECD spectra of $[Pd_6\{(M)\text{-}\mathbf{1}\}_{12}](BF_4)_{12}$ in acetonitrile.

4.4. Conclusion

In conclusion we have shown that BINOL-based bis(pyridine) ligands form hexa- and dodecanuclear enantiomerically pure metallosupramolecular spheres with a chiral skeleton upon coordination to Pd^{II} ions depending on the substitution pattern of the BINOL core. Despite the relatively high rotational freedom around the central aryl–aryl bond of these ligands the self-assembly process happens completely selective in a narcissistic self-recognition manner. In this way we are able to prepare *endo*- and *exo*-functionalized chiral nanoscopic supramolecular objects that show extraordinary high molar circular dichroisms.

Acknowledgments

This work was supported by the DFG (SFB 624 and 813) and the Academy of Finland (KR. grant no. 265328 and 263256).

5. A Simplified Time-Dependent Density Functional Theory Approach for Electronic Ultraviolet and Circular Dichroism Spectra of Very Large Molecules

Christoph Bannwarth* and Stefan Grimme*

Received 15th of December 2013, Published online 1st of March 2014

Reprinted (adapted) with permission from†
Bannwarth, C.; Grimme, S. *Comput. Theor. Chem.* **2014**, *1040–1041*, 45–53.

DOI 10.1016/j.comptc.2014.02.023

Own manuscript contribution

- Implementation of the sTD-DFT method into the `stda` standalone program
- Performing the calculations
- Interpretation of the results
- Writing the manuscript

*Mulliken Center for Theoretical Chemistry, Institut für Physikalische und Theoretische Chemie, Rheinische Friedrich-Wilhelms-Universität Bonn, Beringstraße 4, 53115 Bonn, Germany
†Permission requests to reuse material from this chapter should be directed to Elsevier.

Abstract We present a simplified time-dependent density functional theory approach (sTD-DFT) that allows fast computation of electronic ultraviolet (UV) or circular dichroism (ECD) spectra of molecules with 500–1000 atoms. The matrix elements are treated in the same way as in the recently proposed simplified Tamm-Dancoff approach (sTDA-DFT: Grimme, S. *J. Chem. Phys.* **2013**, *138*, 244104) but instead of applying the Tamm-Dancoff approximation, the standard linear-response density functional theory problem is solved. Compared to sTDA-DFT, the method leads to an increase in computation time (typically a factor of 2–5 compared to the corresponding sTDA-DFT) which is justified since the resulting transition dipole moments are in general of higher quality. This becomes important if spectral intensities (e.g., single-photon oscillator and rotatory transition strengths) are of interest. Comparison of electronic UV absorption and ECD spectra obtained from sTD-DFT and sTDA-DFT for some typical systems employing standard hybrid functionals shows that both yield very similar excitation energies but the advantage of using the former approach for transition moments. In order to show the applicability of sTD-DFT to systems which are far beyond the scope of conventional TD-DFT, we present the ECD spectrum of a substituted, chiral fullerene over a range of almost 1200 excited states. We propose this method as a more reliable alternative for the prediction especially of the more challenging ECD spectra.

5.1. Introduction

Kohn-Sham density functional theory (KS-DFT) is now the most widely used method for electronic structure calculations of larger molecules in the electronic ground state. For the calculation of excited state properties and electronic spectra of fairly large systems (about 100 atoms), time-dependent density functional theory (TD-DFT)[124,125,270,271] based on a KS-DFT ground state determinant has become the most important method (see, e.g., Refs. 15,129,272–274 for reviews). For TD-DFT calculations of ground state properties like dispersion coefficients, see, e.g., Refs. 275 and 276.

It is generally assumed in TD-DFT that the density changes only slowly with time. For this reason, the time-dependent exchange-correlation (XC) kernel is replaced by a time-independent one which is then evaluated using the time-dependent density. This so-called adiabatic approximation[125,270] allows one to make use of ground state XC potentials in TD-DFT. The states of common interest are often valence states far below the first ionization potential for which this approximation works fairly well.[124]

The accuracy of TD-DFT for vertical excitation energies is roughly comparable to that of KS-DFT for relative ground state energies (about 0.2–0.3 eV for TD-DFT, 2–5 kcal $mol^{-1}$for KS-DFT).[73,75] There are some well-known deficiencies of TD-DFT such as the description of excitations with double excitation or multiplet character which are problematic for any single-reference method. Density functionals of the general gradient approximation (GGA) type, have an XC potential of an incorrect asymptotic form[129] and suffer also from the self-interaction error (SIE)[277–280] and the related integer discontinuity problem.[280,281] General implications are

underestimated ionization potentials, overestimated electron affinities and too small energy gaps between Kohn-Sham orbitals. Due to these deficiencies TD-DFT/GGA severely underestimates charge-transfer (CT) and Rydberg excitations.[94,95,129,281,282] This artificially introduces many states with low oscillator strength ("ghost states") to the low energetic part of the spectrum. Even though they are often not observed in a simulated ultraviolet-visible (UV-Vis) spectrum, they may cause artificial mixing of configurations and "contaminate" the bright states which can then corrupt spectra considerably.

Since Hartree-Fock (HF) does not suffer from SIE and exhibits the correct asymptotic potential, admixing certain amounts of non-local Fock exchange to the GGA exchange, as done in global hybrids (e.g., B3LYP[69,106,107]), partially alleviates the SIE as well as the CT problem in TD-DFT. Such functionals are more reliable in an excited state treatment than pure GGAs on the one hand and HF based single-excitation methods on the other. The advantage of (hybrid) TD-DFT over the latter are the implicitly "correlated" Kohn-Sham orbitals. Since semiempirical methods are typically less reliable and wave-function based methods, which include double and higher excitations, are too costly to be applied to large systems, hybrid TD-DFT has become one of the most widely used methods to describe excited states of systems up to about 100 atoms.

The amount of non-local Fock exchange for typical hybrid functionals lies in the range of 10-25%, but can be as high as 50% like in the BHLYP functional.[69] In a systematic study, it has been found that on average admixing 40% of non-local Fock exchange yields the best excitation energies of fairly large organic molecules,[283] although the correct treatment of some larger systems required higher amounts of Fock exchange up to 50% or more. A commonly applied class of functionals in TD-DFT calculations nowadays are the range-separated-hybrid (RSH) functionals that, starting with no or low amounts of Fock exchange, asymptotically employ 100% non-local Fock exchange (65% in the case of CAM-B3LYP).[88–90,92] With these functionals, the correct asymptotic behavior is achieved. Nevertheless, we will employ exclusively global hybrids in this work and make use of the BHLYP functional in large systems where the description of CT states may become problematic. The extension of the here proposed method to RSH functionals is straightforward (see below) and will be discussed elsewhere.

Even though TD-DFT can deal with systems beyond the scope of traditional wave function based methods, the theoretical treatment of an entire UV-Vis electronic spectrum in a typical excitation energy range from 2 to 7 eV for systems with several hundreds up to about 1000 atoms remains a challenge. Recently, a simplified Tamm-Dancoff approach to time-dependent density functional theory (sTDA-DFT) has been proposed which allows routine computations of electronic absorption or circular dichroism (ECD) spectra of such large systems.[46] The drastic simplifications are on the one hand, the evaluation of the two-electron integrals as short-range damped Coulomb interactions between (transition) charge density monopoles and a massive truncation of the single excitation expansion space on the other. Solving the Tamm-Dancoff approximated (TDA) problem[127,128] instead of the TD-DFT equation requires the solution of only one eigenvalue problem and, along with the simplifications mentioned above, this makes

the sTDA-DFT approach extremely fast even for large molecules.

While the TDA-DFT typically gives quite similar excitation energies as TD-DFT, it suffers from the fact that it is not gauge invariant and oscillator and rotatory strengths obtained from TDA-DFT do not satisfy the respective sum rules.[47,129,271] These shortcomings can become particularly problematic if one is interested in calculating rotatory strengths for ECD spectra. TD-DFT, on the other hand, does not suffer from these shortcomings. Therefore, a simplified time-dependent density functional theory approach (sTD-DFT) is presented here which makes use of the same simplifications as made in sTDA-DFT while the full TD-DFT problem is solved.

This paper is structured in the following way: after a brief outline of the basic theory followed by a short summary of the sTDA-DFT method, we will recapitulate the dipole length and the dipole velocity formalisms for transition moments. Then the sTD-DFT is discussed. Finally, we will compare the performance of sTD-DFT to sTDA-DFT and conventional TD-DFT for the computation of excitation energies, as well as absorption and ECD spectra, for various systems.

5.2. Theory

5.2.1. TD-DFT and TDA-DFT

The full TD-DFT response problem is given by the following non-Hermitian eigenvalue problem[125,129]

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B^*} & \mathbf{A^*} \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} = \begin{pmatrix} \omega & 0 \\ 0 & -\omega \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix}, \tag{5.1}$$

where $\mathbf{A}$ and $\mathbf{B}$ are the so-called orbital rotation Hessian matrices with eigenfunctions $\mathbf{X}$ and $\mathbf{Y}$ and ω is a vector with the dimension of the number of roots (nroots) that contains the respective eigenvalues. In wave function theory, this equation corresponds to time-dependent Hartree-Fock (TD-HF, also called random-phase approximation, RPA). For a global hybrid density functional in the spin-restricted case, the elements of the matrices $\mathbf{A}$ and $\mathbf{B}$ take the form

$$\begin{aligned} A_{ia,jb} = \delta_{ij}\delta_{ab}(\epsilon_a - \epsilon_i) + 2(ia|jb) - a_x(ij|ab) \\ + (1 - a_x)(ia|f_{XC}|jb) \end{aligned} \tag{5.2}$$

and

$$B_{ia,jb} = 2(ia|bj) - a_x(ib|aj) + (1 - a_x)(ia|f_{XC}|bj). \tag{5.3}$$

Here, ϵ_a and ϵ_i are the orbital energies of the virtual and occupied orbitals obtained from the respective ground state calculation. The two-electron integrals are given in the so-called chemists' notation, i.e., $(ia|jb) = \iint \psi_i(\mathbf{r}_1)\psi_a(\mathbf{r}_1)\frac{1}{r_{12}}\psi_j(\mathbf{r}_2)\psi_b(\mathbf{r}_2)\mathrm{d}\mathbf{r}_1\mathrm{d}\mathbf{r}_2$ (ij stand for occupied, ab for virtual orbitals), a_x is the amount of non-local Fock exchange that is mixed into the XC functional[69] such that

$$E_{XC}^{hybrid} = (1 - a_x)E_X^{GGA} + a_x E_X^{Fock} + E_C^{GGA} \tag{5.4}$$

where E_X^{GGA} and E_X^{Fock} denote semi-local GGA and non-local Fock exchange terms, respectively and E_C^{GGA} is the GGA correlation energy. For $a_x = 1$ the above equations correspond to the HF/RPA case.[124] The other extreme (i.e., $a_x = 0$) corresponds to semi-local TD-DFT equations.[125] The terms $2(ia|jb)$ and $2(ia|bj)$ in the matrix elements of **A** and **B**, respectively, are of exchange type and result from the response of the Coulomb integrals in the ground state. The response of the XC functional is given by the terms $(ia|f_{XC}|jb)$ and $(ia|f_{XC}|bj)$, respectively, and its contribution is scaled by the amount of (semi-)local density functional exchange used in the ground state (first term on the right-hand side of Eq. 5.4). Accordingly, the terms scaled by a_x correspond to the response of the non-local Fock exchange. While this term is of exchange type in the **B** matrix elements, it is of Coulomb type in the **A** matrix elements. This term alleviates the incorrect description of CT states by GGA-based TD-DFT.

Since the used orbitals are usually real, instead of the non-Hermitian eigenvalue problem in Eq. 5.1, the Hermitian one

$$(\mathbf{A}-\mathbf{B})^{\frac{1}{2}}\,(\mathbf{A}+\mathbf{B})\,(\mathbf{A}-\mathbf{B})^{\frac{1}{2}}\,\mathbf{Z} = \omega^2\mathbf{Z} \tag{5.5}$$

with

$$\mathbf{Z} = (\mathbf{A}-\mathbf{B})^{-\frac{1}{2}}\,(\mathbf{X}+\mathbf{Y}) \tag{5.6}$$

can be solved.[124,125] The matrix $(\mathbf{A}-\mathbf{B})$ is diagonal in the case of a pure GGA[129,284] but not in TD-HF and hybrid TD-DFT. Thus, in order to take the square-roots a unitary transformation needs to be performed which involves a diagonalization in the full configuration space. To circumvent this step, Hirata and Head-Gordon have applied the Tamm-Dancoff approximation[127] to TD-DFT in which the contribution of the **B** matrix is neglected.[128] Instead of solving two eigenvalue problems as in hybrid TD-DFT, only one eigenvalue problem needs to be solved:

$$\mathbf{A}\mathbf{t} = \omega_{TDA}\mathbf{t} \tag{5.7}$$

The solution vector **X** has been replaced by **t** and ω by ω_{TDA} to point out that the eigenfunctions and eigenvalues, respectively, are different from the ones in Eq. 5.1. In the case of $a_x = 1$, TDA-DFT corresponds to the configuration interaction approach for single excitations (CIS).[129] It has been shown that excitation energies obtained from TDA-DFT are usually only slightly larger than the respective ones obtained from TD-DFT but come at a much lower cost.[128] For a recent comparison of TDA-DFT and TD-DFT treatments for valence states see Ref. 285, for an early precursor to TDA-DFT derived from the CI formalism (DFT/CIS) see Ref. 286.

5.2.2. The sTDA-DFT Approach

Starting from the TDA-DFT equation (Eq. 5.7), three simplifications lead to the recently published sTDA-DFT approach.[46] The response term of the (semi)-local density functional (last term in Eq. 5.2) is neglected to avoid expensive numerical integration.

The second simplification concerns the remaining two-electron integrals of the **A** matrix. They

are approximated by short-range damped Coulomb interactions of (transition) charge density monopoles. These (transition) charge densities q_{pq}^A are obtained from a Löwdin population analysis[287]

$$q_{pq}^A = \sum_{\mu \in A} C'_{\mu p} C'_{\mu q} \tag{5.8}$$

where the sum is over AO functions (index μ) centered on atom A (indices $pqrs$ denote molecular orbitals that can be either occupied or virtual). The matrix $\mathbf{C}'$ denotes orthogonalized MO coefficients obtained from $\mathbf{C}' = \mathbf{S}^{1/2}\mathbf{C}$ where $\mathbf{C}$ are the coefficients in the original basis. It should be noted that $\sum_{\mu}^{N_{bf}} |C'_{\mu p}|^2 = 1$ for all p (N_{bf} is the number of AOs). The two-electron integrals are then approximated by

$$(pq|rs) \approx \sum_{A}^{N} \sum_{B}^{N} q_{pq}^A q_{rs}^B \gamma_{AB} \tag{5.9}$$

with γ_{AB} being the Mataga-Nishimoto-Ohno-Klopman[288–290] damped Coulomb operator. If Eq. 5.9 is used to replace a Coulomb type integral (superscript J), γ_{AB} is given by:

$$\gamma_{AB}^J = \left(\frac{1}{(R_{AB})^\beta + (a_x \eta)^{-\beta}} \right)^{\frac{1}{\beta}} \tag{5.10}$$

Here, R_{AB} is the interatomic distance, β is a parameter, and η is the arithmetic average of the chemical hardness of the two atoms A and B, $\eta(A) = \frac{\partial^2 E(A)}{\partial n^2}$, where n is the number of electrons and E is the total atomic electronic energy. The tabulated $\eta(A)$ values consistent for all elements of the periodic table from Ref. 291 are used. The respective expression of γ_{AB} for exchange type integrals is then:

$$\gamma_{AB}^K = \left(\frac{1}{(R_{AB})^\alpha + \eta^{-\alpha}} \right)^{\frac{1}{\alpha}} \tag{5.11}$$

The values of the parameters β in Eq. 5.10 and α in Eq. 5.11 are taken from Ref. 46. These are proportional to the amount of non-local Fock exchange a_x, but independent of the chosen XC functional.

The approximated $\mathbf{A}$ matrix (denoted as $\mathbf{A}'$) then has elements of the form

$$A'_{ia,jb} = \delta_{ij}\delta_{ab}(\epsilon_a - \epsilon_i) + \sum_{A,B}^{N_{atoms}} (2 q_{ia}^A \gamma_{AB}^K q_{jb}^B - q_{ij}^A \gamma_{AB}^J q_{ab}^B). \tag{5.12}$$

Note that the prefactor a_x is dropped, but its effect is accounted for in the γ_{AB}^J. Since the parameters α and β are empirically adjusted, it can be argued that the response of the XC functional is actually not completely neglected. Rather, its contribution is absorbed into the fitted parameters. In the orignal paper, it has been shown that sTDA-DFT yields the correct asymptotic behavior that is required to correctly describe CT states.[46,195]

The last simplification applied in sTDA-DFT concerns the truncation of the single excitation space. Truncation of CI spaces has been applied successfully at only minor loss of accuracy already in other approaches.[292–295] We refer to the original sTDA-DFT paper for details and only state here that configuration spaces are typically reduced to about 1% of the full value and that $\mathbf{A}'$ matrices of dimension of a few thousand are treated even in large cases.

These three simplifications along with the simple eigenvalue problem (Eq. 5.7) allow extremely fast computations for a fairly large number of roots which is desirable when one is interested in predicting spectra.[46] Due to its correct description of CT states, it is also applicable to very large systems as long as $a_x = 1$ or RSH functionals are applied.[195]

5.2.3. Transition Dipole Moments, Oscillator and Rotatory Strengths

The sTDA-DFT method uses transition dipole moments which are not integral-approximated. This implies that transition dipole integrals between CSF are identical in sTDA-DFT and conventional TDA-DFT. On the other hand, the transition dipole moments between states are different, since the simplifications in the $\mathbf{A}'$ matrix compared to $\mathbf{A}$ lead to different amplitudes t_i^a. Except for very local excitations (e.g., nπ^*), the differences are comparably small.[46]

A more important problem can be the TDA formalism itself as it is not gauge invariant[271] which can lead to severe inaccuracies. For a state ν to be visible in the UV-Vis absorption spectrum, it must have a non-zero oscillator strength[296]

$$f_{0\nu} = \frac{2}{3}\omega_{0\nu}\vec{\mu}_{0\nu} \cdot \vec{\mu}_{\nu 0} \tag{5.13}$$

with $\vec{\mu}_{0\nu}$ being the electric dipole transition moment. The respective condition for a state to show ECD, is a non-vanishing rotatory strength[33]

$$R_{0\nu} = Im\{\vec{\mu}_{0\nu} \cdot \vec{m}_{\nu 0}\} \tag{5.14}$$

with $\vec{m}_{\nu 0}$ being the magnetic dipole transition moment (atomic units are used throughout)

$$\vec{m}_{\nu 0} = \frac{i}{2}\left\langle \nu \left| \vec{r} \times \vec{\nabla} \right| 0 \right\rangle \tag{5.15}$$

There are two formalisms to describe the electric dipole transition moment which then also lead to different expressions for the oscillator and particularly the rotatory strength. One is the so-called dipole length formalism

$$\vec{\mu}_{0\nu}^{L} = \langle 0 \left| \vec{r} \right| \nu \rangle \ , \tag{5.16}$$

the other one is the dipole velocity formalism

$$\vec{\mu}_{0\nu}^{V} = \frac{\vec{\nabla}_{0\nu}}{\omega_{0\nu}} = \frac{\left\langle 0 \left| \vec{\nabla} \right| \nu \right\rangle}{\omega_{0\nu}} \ . \tag{5.17}$$

For exact wave functions, Eqs. 5.16 and 5.17 are equivalent and the oscillator and rotatory

strengths obtained from either of these formalisms will be identical.[54] In the complete basis set limit, this will be the case for TD-DFT, but not for the TDA-DFT.

To understand the implications of this, it is necessary to understand the effect of shifting the origin (i.e., $\vec{r} \rightarrow \vec{r} + \vec{R}$). For orthonormal MOs, both expressions of the electric dipole transition moments (Eqs. 5.16 and 5.17) will remain unaltered, i.e., they are origin-independent.[297] This is not the case for the magnetic dipole transition moment which then is given by

$$\begin{aligned} \vec{m}_{\nu 0} &= \frac{i}{2} \left\langle \nu \left| \left(\vec{r} + \vec{R}\right) \times \vec{\nabla} \right| 0 \right\rangle \\ &= \frac{i}{2} \vec{m}'_{\nu 0} + \frac{i}{2} \left(\vec{R} \times \vec{\nabla}_{\nu 0}\right) , \end{aligned} \tag{5.18}$$

with $\vec{m}'_{\nu 0}$ being the origin-independent part of the magnetic dipole transition moment which is given in Eq. 5.15 (prefactor excluded). The second term carries the origin dependence. Plugging this into Eq. 5.14 yields

$$R^L_{0\nu} = \frac{1}{2} \vec{\mu}^L_{0\nu} \cdot \vec{m}'_{\nu 0} + \frac{1}{2} \vec{\mu}^L_{0\nu} \cdot \left(\vec{R} \times \vec{\nabla}_{\nu 0}\right) \tag{5.19}$$

for the dipole length formalism. In the dipole velocity formalism, the origin-dependent term vanishes:

$$\begin{aligned} R^V_{0\nu} &= \frac{1}{2\omega_{0\nu}} \vec{\nabla}_{0\nu} \cdot \vec{m}'_{\nu 0} + \frac{1}{2\omega_{0\nu}} \underbrace{\vec{\nabla}_{0\nu} \cdot \left(\vec{R} \times \vec{\nabla}_{\nu 0}\right)}_{=0} \\ &= \frac{1}{2\omega_{0\nu}} \vec{\nabla}_{0\nu} \cdot \vec{m}'_{\nu 0} \end{aligned} \tag{5.20}$$

This expression is therefore always origin-independent.[298,299] One can easily see that in the dipole length case, $R^L_{0\nu}$ is only origin-independent if $\vec{\mu}^L_{0\nu}$ is parallel to $\vec{\nabla}_{\nu 0}$. That is the case if the dipole length and dipole velocity formalisms are equivalent. Since this does not hold in the TDA formalism, the calculated values of $R_{0\nu}$ will always be flawed in a way, either by the origin-dependence (in $R^L_{0\nu}$) or the incorrect description of the electric dipole transition moment in the dipole velocity formalism. Since the oscillator strength is always positive and due to the fact that both expressions for the electric dipole transition moment are origin-independent, oscillator strengths are in general less affected. Nevertheless, the magnitude of $f_{0\nu}$ can be quite different for both formalisms. Since we are interested in producing reliable UV-Vis absorption and ECD spectra for the large systems that are targeted by sTDA-DFT, we apply the same simplifications made in the sTDA-DFT method to the TD-DFT formalism (Eq. 5.1) in order to improve the description of transition moments in general.

5.2.4. The Simplified Time-Dependent DFT (sTD-DFT) Approach

The starting point for the sTD-DFT approach is Eq. 5.1. The matrix $\mathbf{A}$ is replaced by the approximate matrix $\mathbf{A}'$ from sTDA-DFT (with elements given by Eq. 5.12). The configuration

selection is the same as in sTDA-DFT (see Ref. 46). Because the TD-DFT problem needs to be solved, the matrix **B** has to be set up in a consistent, simplified manner. Just like in sTDA-DFT, the response term of the density functional (last term in Eq. 5.3) is neglected and the first term of the right-hand side will be approximated by the exchange-type Mataga-Nishimoto-Ohno-Klopman damped Coulomb interaction (see Eq. 5.11). As mentioned above, the term originating from the non-local Fock exchange is of Coulomb type in matrix **A**, but of exchange type in matrix **B**. Since we do not want to fit another set of parameters, we will also use the exchange-type Mataga-Nishimoto-Ohno-Klopman damped Coulomb interaction for this integral but keep the scaling factor a_x. The elements of the approximated matrix $\mathbf{B}'$ are then given by

$$B'_{ia,jb} = \sum_{A,B}^{N_{atoms}} (2q_{ia}^{A}\gamma_{AB}^{K}q_{bj}^{B} - a_x q_{ib}^{A}\gamma_{AB}^{K}q_{aj}^{B}). \tag{5.21}$$

We stress here that neither any of the parameters presented in Ref. 46 has been refitted nor any additional parameters have been introduced. Solving the Hermitian eigenvalue problem in Eq. 5.5 yields the solution vectors **X** and **Y**. In the spin-restricted case, the transition moments from the ground to an excited state are then obtained from[47]

$$\vec{\mu}_{0\nu}^{L} = \sqrt{2}\sum_{ia}^{N_{CSF}} \langle \psi_i \left|\vec{r}\right| \psi_a \rangle \left(X_{ia,\nu} + Y_{ia,\nu}\right) \tag{5.22}$$

$$\vec{\nabla}_{0\nu} = \sqrt{2}\sum_{ia}^{N_{CSF}} \left\langle \psi_i \left|\vec{\nabla}\right| \psi_a \right\rangle \left(X_{ia,\nu} - Y_{ia,\nu}\right) \tag{5.23}$$

$$\vec{m}_{0\nu} = \sqrt{2}\sum_{ia}^{N_{CSF}} \left\langle \psi_i \left|\vec{r}\times\vec{\nabla}\right| \psi_a \right\rangle \left(X_{ia,\nu} - Y_{ia,\nu}\right) \tag{5.24}$$

based on excitations from orbitals ψ_i to ψ_a.

5.3. Technical Details of the Calculations

We used the TURBOMOLE suite of programs[164a,165] (version 6.5) to perform all ground state DFT calculations. The sTD-DFT was implemented into our group-intern `sTDA` code which is a stand-alone program that is compatible with TURBOMOLE. All of the sTDA-DFT or sTD-DFT excited state calculations presented here were carried out with this version of the program. We employ standard integration grids (*m4* if not noted otherwise) and typical SCF convergence criteria ($10^{-7}\,E_h$). With the exception of Section 5.4.2, the resolution of the identity (RI) integral approximation[76,166–168] was used in all SCF calculations. Structures were fully optimized at the TPSS-D3/def2-TZVP level.[105,133,134,163] In the excited state treatments, either the atom-centered SV(P)[300] or TZVP[162] basis sets are used that were successfully applied in the sTDA-DFT treatments.[46] Since Rydberg states are not of interest here, we do not use diffuse

basis functions. For large systems, augmented basis sets or basis sets with cardinal number higher than three can become problematic as they might lead to near linear dependencies (for the related problem even with TZVP in solids see Ref. 182).

Since the effect of the chosen XC functional is by far less important than the amount of non-local Fock-exchange in the hybrid density functional, we will restrict our calculations to standard hybrid functionals such as PBE0[104,111] and BHLYP[69] (see also footnote*). All excitation energies refer to vertical singlet states for the optimized ground state geometry (TPSS-D3/def2-TZVP). The default configuration selection threshold of $t_p = 10^{-4}\,E_h$ is used.[46] The E_{max} value which specifies desired spectral range will be individually given for each calculation. The active occupied and virtual MOs are automatically selected according to the chosen E_{max} value and their number is typically only about 20-30% of the full MO space. For comparison, conventional TDA-DFT or TD-DFT calculations were performed using the `escf` module[301] from TURBOMOLE.

As stated above, the one-electron transition moments are calculated without any further approximation from the exact moment integrals in the Cartesian AO basis.

In order to simulate the vibrationally broadened experimental UV absorption or ECD bands, we summed oscillator or rotatory strengths weighted Gaussian curves with a full width at $1/e$-height of $\sigma = 0.4\,\mathrm{eV}$ for each calculated electronic transition. If not stated otherwise, we will use the dipole length formalism for oscillator strengths in UV absorption spectra and the origin-independent dipole velocity formalism for rotatory strengths in ECD spectra.

5.4. Results and Discussion

5.4.1. Rotatory Strengths from Conventional and Simplified TDA-DFT and TD-DFT in the Biphenyl Molecule

In this section, the inherent deficiency of TDA-DFT in obtaining reliable rotatory strengths as well as oscillator strengths will be demonstrated using the biphenyl molecule as a test case. As a measure of internal consistency of each method, we consider the difference of the rotatory strengths obtained from the length and velocity formalism (Eqs. 5.19 and 5.20). We will thus use the modulus of the difference $\Delta R = \left|R^V_{0\nu} - R^L_{0\nu}\right|$, which is zero for exact wave functions.

For biphenyl as a test case (Figure 5.1), single-point calculations were done at the BHLYP/TZVP and BHLYP/SV(P) levels and used in the corresponding sTDA-DFT/TDA-DFT and sTD-DFT/TD-DFT treatments.

Based on these Kohn-Sham determinants, in the conventional TD-DFT and TDA-DFT 50 roots were taken within the B_1, B_2 and B_3 irreducible representations which have non-vanishing rotatory strengths. For TDA-DFT and TD-DFT, the Table 5.1 shows the mean absolute devi-

*From our experience, among the global hybrid density functionals, BHLYP is the most reliable for larger systems. Test calculations on a photoactive protein with almost 2000 atoms yield a large number of ghost states with the PBE0 and especially the B3LYP functional. BHLYP was still able to correctly yield the dipole allowed state as the lowest one and is hence our global hybrid default for very extended systems.

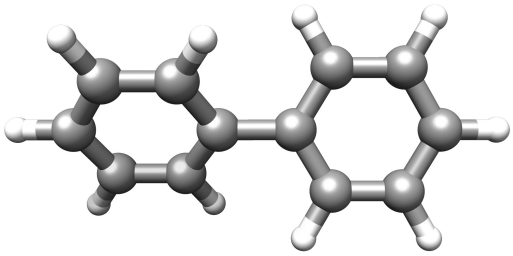


Figure 5.1.: Ball-and-stick depiction of the biphenyl (D_2 symmetry) molecule used in the calculations.

ation $\Delta\overline{R}$ for the considered 150 states.

It can be seen that TDA-DFT yields relative deviations that are an order of magnitude larger than in the case of TD-DFT (see column 5 in entries 1 and 2 or 5 and 6, respectively). The large deviations between the dipole length and the velocity expressions in TDA-DFT indicate that this approach is not very reliable for the computation of ECD spectra. This is different in TD-DFT which gives small deviations between both formalisms even for the comparably small SV(P) basis set (entry 6 in Table 5.1).

Table 5.1.: Mean absolute deviations (MAD) between length and velocity formalism rotatory strengths ($\Delta\overline{R}$), average magnitude of R^V (in 10^{-40} erg $\cdot$ cm^3) and relative MADs $\Delta\overline{R}_{rel.}$ for conventional and simplified TDA-DFT and TD-DFT for 150 excited states of biphenyl.

entry	method	$\Delta\overline{R}$	$\lvert\overline{R}^V\rvert$	$\Delta\overline{R}_{rel.}$
		BHLYP/TZVP		
1	TDA-DFT	14.64	23.74	0.62
2	TD-DFT	2.13	39.04	0.05
3	sTDA-DFT	9.16	41.66	0.22
4	sTD-DFT	2.16	49.11	0.04
		BHLYP/SV(P)		
5	TDA-DFT	10.62	25.63	0.42
6	TD-DFT	1.84	35.85	0.05
7	sTDA-DFT	8.33	29.10	0.29
8	sTD-DFT	1.75	34.95	0.05

Entries 3 and 7 show the results obtained for sTDA-DFT and 4 and 8 for sTD-DFT, respectively. To be comparable to the conventional methods, 150 lowest states which showed non-zero rotatory strengths were selected. An interesting point is that on average both sTDA-DFT and sTD-DFT show similar (the former even smaller) absolute $\Delta\overline{R}$ than their respective conventional counterparts. For sTDA-DFT, even the relative deviation $\Delta\overline{R}_{rel.} = \Delta\overline{R}/\left|\overline{R}^V\right|$ is smaller by a

factor of approximately 1.5–3 compared to TDA-DFT.

It can be seen that conventional TD-DFT and sTD-DFT show comparably larger R^V values than the respective TDA approaches. This behavior seems to be the result of underestimated magnitudes of $\vec{\mu}_{0\nu}^V$ in the TDA formalism since it is also observable for the oscillator strengths, see, e.g., the UV-Vis absorption spectrum of indigo in Figure 5.2(a). The magnitudes of $\vec{\mu}_{0\nu}^L$ on the other hand seem to be overestimated in the TDA formalism. This difference in magnitudes then leads to the comparably large deviations $\Delta\overline{R}$ for TDA-DFT and sTDA-DFT which, in addition, can increase if $\vec{\mu}_{0\nu}^L$ and $\vec{\mu}_{0\nu}^V$ are differently oriented.

The very encouraging result from this study is the extremely small deviation $\Delta\overline{R}$ obtained from the sTD-DFT which are almost identical to the deviations for TD-DFT and thus lower than $\Delta\overline{R}$ from either of the TDA approaches. Taking into account the different magnitudes of R^V in all four approaches, a fairer comparison is the relative deviation $\Delta\overline{R}_{rel.}$. For this measure, sTD-DFT also outperforms TDA-DFT and sTDA-DFT showing relative deviations that are equivalent to conventional TD-DFT. Therefore, it seems that sTD-DFT can function as a reliable method to produce spectra even for larger molecules. In the following sections, this will be demonstrated for ECD, but also for UV absorption spectra.

5.4.2. Comparison of Vertical Excitation Energies in sTDA-DFT and sTD-DFT

In Ref. 46, the empirical parameters α and β have been fitted to 26 valence singlet excitation energies. Here we have reduced the number to 20 valence states since some systems seemed less well suited for comparison due to various reasons (e.g., multi-reference character, questionable reference data). The used molecules along with the respective vertical excitation energies are given in Table 5.2. We calculated the respective vertical excitation energy (VEE) by both sTDA-DFT and sTD-DFT. The results and the respective mean signed deviation (MSD) and mean absolute deviation (MAD) are also given in Table 5.2.

As can be seen, the excitation energies obtained by sTD-DFT are always slightly smaller than the respective sTDA-DFT ones. This is expected and known from a comparison of conventional TDA-DFT and TD-DFT.[128] In comparison to sTDA-DFT, the VEEs from sTD-DFT show slightly larger deviations. This results from the fact that the parameters α and β have been fitted partially on this test set but are not adjusted here. Nevertheless, the MAD of 0.26 eV for sTD-DFT is negligibly larger than the MAD for sTDA-DFT (0.21 eV) on this test set and very small regarding the drastic approximations made. These MAD values are of about the sizes as changes induced by varying the amount of Fock exchange in the functional by 10-20%.[46]

Table 5.2.: Comparison of calculated[a] sTDA-DFT and sTD-DFT singlet excitation energies (in eV) with respective experimental values.

entry	molecule	state	sTDA-DFT	sTD-DFT	exptl.
1	CH_2S	1^1A_2	2.09	2.09	2.24
2	tetrazine	1^1B_{1u}	2.08	2.08	2.29
3	CH_2O[b]	1^1A_2	4.01	4.01	3.88
4	uracil	1^1A"	4.70	4.70	4.8
5	P_4	1^1T_1	5.31	5.31	5.6
6	acetamide	1^1A"	5.73	5.73	5.69
7	porphyrine	1^1B_{2u}	2.10	2.08	2.0
8	"	1^1B_{3u}	2.30	2.26	2.4
9	azulene[b]	1^1B_1	2.45	2.39	2.19
10	perylene	2^1B_{2u}	2.87	2.69	3.44
11	coumarin153[b]	2^1A	3.17	3.04	3.51
12	anthracene	1^1B_{3u}	3.35	3.18	3.7
13	*t*-azobenzene	1^1B_u	3.84	3.68	3.9
14	DMABN[b]	1^1B_1	4.47	4.41	4.3
15	"	1^1A_1	4.84	4.67	4.6
16	octatetraene	1^1B_u	4.39	4.01	4.66
17	hexatriene	1^1B_u	5.24	4.83	5.10
18	norbornadiene	1^1A_2	5.21	5.11	5.34
19	benzene	1^1B_{1u}	5.52	5.51	5.08
20	"	1^1B_{2u}	6.61	6.44	6.54
	MSD[c]		−0.05	−0.15	–
	max-min[d]		1.00	1.18	–
	MAD[e]		0.21	0.26	–

[a] Based on PBE0/TZVP using a TPSS-D3/def2-TZVP optimized geometry.
[b] Deviation of sTDA-DFT VEEs to values from Ref. 46 results from a different geometry.
[c] Mean (signed) deviation. A negative value corresponds to a systematic underestimation of the excitation energy.
[d] Difference between largest and smallest deviation.
[e] Mean absolute deviation.

The excitation energies for local excitations (entries 1−6) are identical in sTDA-DFT and sTD-DFT while differences appear for rather delocalized excitations (entries 7−20). This behavior is reasonable as the elements of matrix $\mathbf{B}$ consist of exchange type terms only (see Eq. 5.3). Within the monopole approximation as applied in sTDA-DFT and sTD-DFT, the transition charge densities q_{ia}^A in Eq. 5.21 will vanish for very local transitions such as nπ^*. For the exchange terms in matrix $\mathbf{A}'$, this has already been discussed in the original sTDA-DFT paper.[46] In the case of sTD-DFT, this means that the respective elements of matrix $\mathbf{B}'$ are vanishingly small and thus, the description of the resulting states is more or less identical in sTD-DFT and sTDA-DFT. Even in the conventional, non-approximated case, the exchange type terms are in fact comparably small. Therefore, a similar behavior is already observed for conventional TDA-DFT

and TD-DFT (see Table 1 in Ref. 46 for the VEEs from the respective conventional approaches). At this point, it should be noted that the MADs as well as the differences between largest and smallest deviation (max-min) for sTD-DFT on this test set are quite similar to TD-DFT on the slightly larger test set (from Ref. 46: MAD = 0.28 eV, max-min = 1.39 eV). Only the negative MSD is slightly more pronounced for sTD-DFT than for conventional TD-DFT (from Ref. 46: MSD = −0.08 eV) which results from using parameters that were fitted for sTDA-DFT.

From the comparison of excitation energies, one can draw the conclusion that sTD-DFT can be employed without a re-fit of the parameters α and β. This is a useful result since the sTD-DFT approach can be combined with RSH functionals and the respective sTDA-DFT parameters (see Ref. 195) in the same straightforward way.

Concerning the computation time, there is hardly any noticeable difference between sTD-DFT and sTDA-DFT for the comparably small systems treated so far. This changes with system size and the increased number of configurations and will be noticeable, e.g., for the system treated in Section 5.4.5.

5.4.3. UV Spectrum of Indigo

The quite close resemblance of UV spectra obtained from sTDA-DFT to TDA-DFT spectra has been shown for both, indigo and 9-(N-carbazolyl)-anthracene.[46] Nevertheless for both molecules, there have been some differences concerning the intensities of certain bands in the sTDA-DFT (and TDA-DFT) spectra and the respective ones obtained from TD-DFT. As has been discussed in Section 5.4.1, sTD-DFT and TD-DFT showed comparable reliability in calculating rotatory strengths. It is thus expected, that a similar behavior is observed for oscillator strengths. We will use the indigo dye in order to examine the performance of sTD-DFT in calculating UV spectra. In Figure 5.2 (a), the electronic absorption spectrum of indigo is given for sTDA-DFT and sTD-DFT obtained from both the length and velocity form for the oscillator strength. We used the PBE0 functional to allow easy comparison with the spectrum from Ref. 46 (please see footnote[†] for a comment on the intensities). For sTDA-DFT, the intensities of bands throughout the whole spectrum are quite different for the two representations while sTD-DFT on the other hand, exhibits excellent agreement between both representations. This is in agreement with the results in Section 5.4.1 for the rotatory strengths. As expected, the agreement of sTD-DFT with TD-DFT is better than with sTDA-DFT. Especially in the region between 7 and 9 eV one can find an improved description compared to sTDA-DFT. As can be seen in Ref. 46, the spectrum obtained from sTDA-DFT resembles more the conventional TDA-DFT spectrum. Thus, both simplified approaches show good agreement with their respective conventional counterparts. The

[†]In the previous version (1.1) of the `stda` program, there was an error in the conversion factors for the oscillator and rotatory strengths. Oscillator strengths were too small by a factor of $\sqrt{2}$, while rotatory strength were too small by a factor of two. Hence, correct absolute intensities of the spectra published in Refs. 46,195, and 302 require a multiplication by the respective factors. Since the relative intensities were correct, the interpretations drawn in these publications are still valid. In Ref. 46, a mixed form for the rotatory strength was used in which R^V is scaled by the ratio f^L/f^V. Due to their different magnitudes in the sTDA-DFT approach (see Section 5.4.3), this compensated somewhat artificially for the missing factor of two.

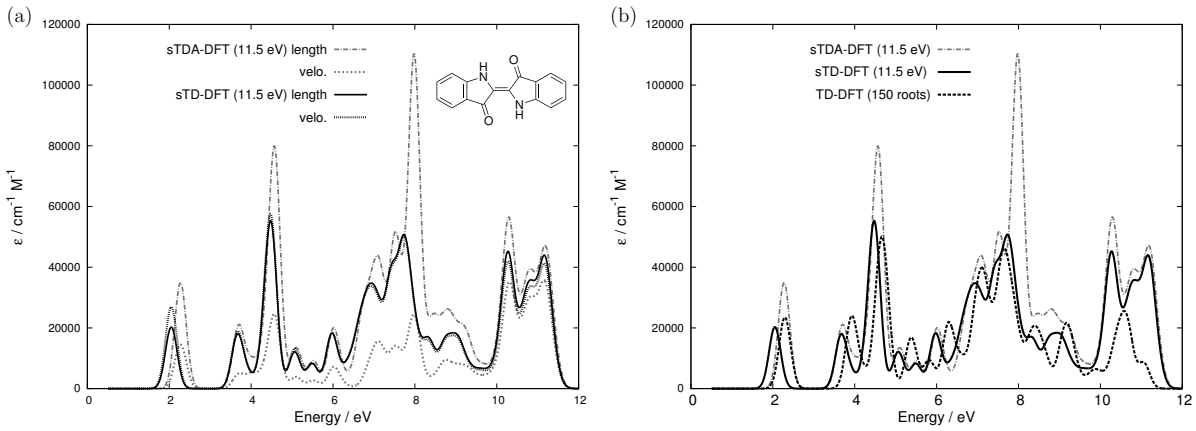


Figure 5.2.: (a) Comparison of computed (PBE0/TZVP) UV-Vis spectra for indigo at sTDA-DFT (gray) and sTD-DFT levels (black). Dashed lines refer to the velocity representation. (b) Respective sTDA-DFT (gray, solid) and sTD-DFT (black, solid) spectra compared to conventional TD-DFT spectrum (black, dashed, from Ref. 46). Spectra are in the length representation.

better agreement between sTDA-DFT and TD-DFT which is observed for excitation energies (see Figure 5.2 (b)) is due to the fit of the α and β parameters to experimental excitation energies. The sTD-DFT spectrum is slightly red-shifted which has already been observed for the VEEs in Section 5.4.2.

From the observations made so far, the sTD-DFT combines the advantages of both sTDA-DFT and TD-DFT which will be further demonstrated for the calculation of ECD spectra.

5.4.4. ECD Spectrum of (*P*)-[11]Helicene

Helicenes are interesting systems for theoretical ECD spectroscopy as they are chiral on the one hand, but have a comparably rigid structure on the other. This allows the calculation of ECD spectra based on a single geometry without averaging over different conformations and has been done for [4]-, [5]-, [6]-, [7]- and [12]helicene at the GGA/TD-DFT level of theory some time ago.[303] Except for the missing negative bands at 280–290 nm, the sTDA-DFT method successfully reproduced the experimental ECD spectrum of (*P*)-[7]helicene. To our knowledge, the largest helicene for which an experimental ECD spectrum has been measured so far is the (*P*)-[11]helicene.[304] Its spectrum exhibits a similar negative band as (*P*)-[7]helicene in the region of 290–340 nm (band II in Figure 5.3). In a very recent publication, it could be shown that sTDA-DFT is not able to produce this part of the spectrum in combination with a global hybrid functional (PBE0) while for a RSH functional the negative ECD was obtained in this region of the spectrum.[195]

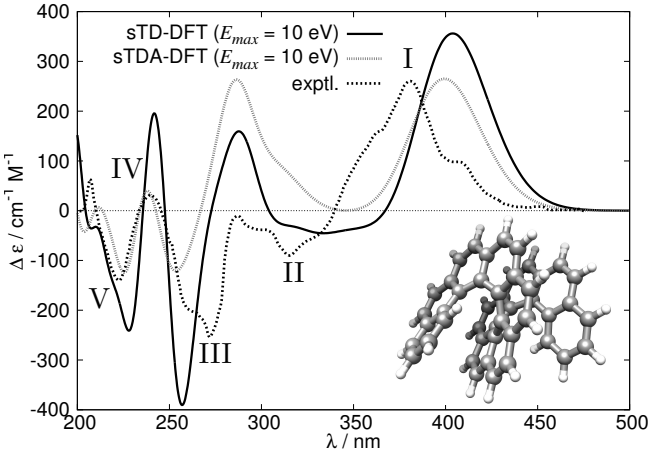


Figure 5.3.: Computed sTDA-DFT (dotted,gray) and sTD-DFT (solid,black) ECD spectra of (*P*)-[11]helicene at the BHLYP/TZVP level. The dashed black line shows the experimental spectrum from Ref. 304. The calculated spectra are blue-shifted by 0.2 eV.

Here, we want to approach this problem from a different direction. In Section 5.4.2, it was pointed out that VEEs from sTDA-DFT and sTD-DFT differ more for strongly delocalized states. One can thus infer that sTD-DFT will also affect and, based on the preliminary results, improve the ECD spectra for a large helicene with a conjugated π-system. We will use the BHLYP functional together with the TZVP basis set[69,162] which should provide similar results as RSH functionals for such a system (see also footnote*). The ECD spectra obtained from sTDA-DFT and sTD-DFT are compared to the experimental one[304] in Figure 5.3.

The computed spectra exhibit a red-shift and are therefore blue-shifted by 0.2 eV for a convenient comparison with the experimental spectrum. For sTDA-DFT, we obtain a similar result as published in Ref. 195. Just like with PBE0, the sTDA-DFT together with the global hybrid functional BHLYP also fails to describe the negative band between 290–340 nm. However, by applying sTD-DFT the negative ECD in the respective region is observed although still with insufficient intensity. This finding suggests that especially for delocalized states, the sTD-DFT leads to better results than the respective sTDA-DFT calculation and that improved transition moments are of particular importance in ECD spectroscopy.

5.4.5. ECD Spectrum of (S,S,$^{f,s}A$)-C_{60}

As a final "show case" for the importance of using the gauge-invariant TD-DFT formalism instead of the TDA formalism, we consider the ECD spectrum of a functionalized, chiral C_{60} fullerene (labeled as (S,S,$^{f,s}A$)-C_{60}, see Figure 5.4), synthesized by Diederich and coworkers.[305]

In the previous section, the improvement of spectra involving delocalized excitations has been demonstrated already and similar states are expected to be important here. In Figure 5.5, we

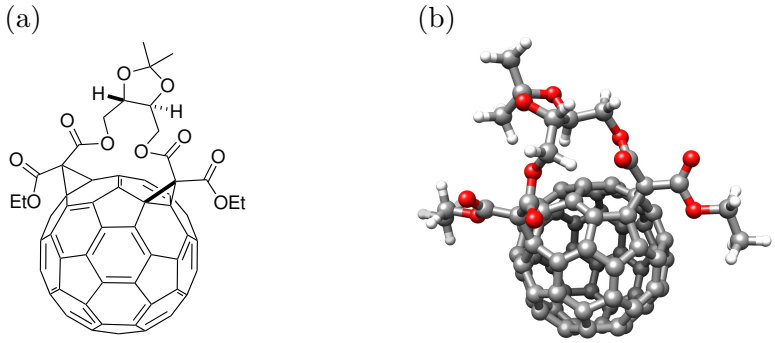


Figure 5.4.: (a) Lewis structure of (S,S,$^{f,s}A$)-C_{60} and (b) ball-and-stick depiction of the geometry used in the computation of the ECD spectra.

compare the computed ECD spectra obtained from a BHLYP/SV(P)//TPSS-D3/def2-TZVP treatment to the experimental one from Ref. 305.

Albeit from a slight blue-shift in the high energy region of the spectrum, we find very good agreement between the sTD-DFT and the experimental spectrum concerning both position and the relative intensities. sTDA-DFT does not only fail to reproduce the relative intensities in the high energy region, but even exhibits bands of incorrect sign for the transitions above 400 nm. When comparing to the experimental spectrum, the bands from sTD-DFT are overpronounced which is likely the result of using only a single geometry for the calculation of the ECD spectrum. Nevertheless, the excellent agreement concerning the positions and relative intensities of bands, which is observed between the sTD-DFT spectrum and the experimental one, on the one hand and the much larger deviations for sTDA-DFT on the other, is a convincing support of our approximations in general and sTD-DFT in particular.

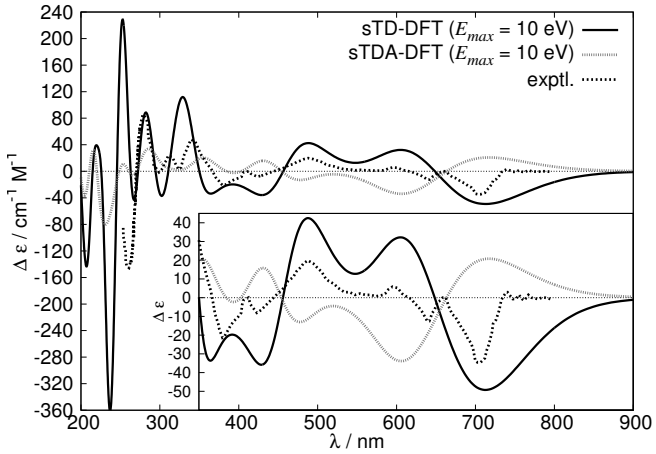


Figure 5.5.: Computed sTDA-DFT (dotted,gray) and sTD-DFT (solid,black) ECD spectra for (S,S,$^{f,s}A$)-C_{60} at the BHLYP/SV(P) level. The dashed black line shows the experimental spectrum (redrawn from Ref. 305). The inset displays the enlarged spectrum above 350 nm.

While possible for the non-functionalized, achiral C_{60} (I_h symmetry), applying conventional TD-DFT becomes very expensive for (S,S,$^{f,s}A$)-C_{60} since this derivative does not exhibit any symmetry. The sTD-DFT calculation itself only takes about 6 minutes on a single processor and the SCF procedure remains the time-determining step. However, it is more expensive than the sTDA-DFT method which only took 0.5 minutes on the same machine but the improved results justify the additional time needed, especially since a conventional TD-DFT calculation with the same number of required roots (1190) is not feasible in a reasonable period of time.

5.5. Conclusion

A simplified approach to TD-DFT has been presented which is based on the same simplifications as in the recently published sTDA-DFT method.[46] Without refitting of parameters, reliable UV and ECD spectra can now routinely be obtained for systems with up to 1000 atoms. The compatibility of sTDA-DFT parameters with sTD-DFT was shown by examining VEEs obtained from sTD-DFT. This makes sTD-DFT a valuable and (parameter) consistent extension to sTDA-DFT, which we recommend to be used in order to obtain higher quality absorption and ECD spectra, i.e., when intensities are of paramount interest.

The sTD-DFT is computationally more expensive than sTDA-DFT, but even the computation of 1200 roots as done for (S,S,$^{f,s}A$)-C_{60} finished in a few minutes of computation time, which is still extremely fast given that the respective conventional calculation would be out of reach on a conventional desktop computer.

Just like its conventional counterpart, sTD-DFT is only weakly dependent on the basis set used and yields reliable oscillator and rotatory strengths even with a small SV(P) basis set. This has been demonstrated by comparing rotatory strengths obtained from the dipole length and dipole velocity representations. This is the major improvement over the sTDA-DFT method which shows comparably large deviations between both representations as a result of the gauge-variant Tamm-Dancoff formalism.

Since our program is developed as an add-on for the TURBOMOLE suite of programs in which RSH functionals are not implemented yet, our calculations were restricted to global hybrid functionals. While for small systems, hybrid functionals with small amounts of non-local Fock exchange typically yield good results, this is not the case when going to larger systems. In the near future, the sTD-DFT code will be implemented into the ORCA program and can then be combined with RSH functionals as has been done with sTDA-DFT already[195] and which might lead to further improvement of computed spectra.

Acknowledgments

This work was supported by the Fonds der Chemischen Industrie and the DFG in the framework of the SFB 813 ("Chemistry at Spin-Centers"). C. B. thanks P. Shushkov for helpful discussions during the implementation of the sTD-DFT code, which is implemented in a stand-alone program and which can be obtained from our website.[306]

6. Electronic Circular Dichroism of Highly Conjugated π-Systems: Breakdown of the Tamm-Dancoff/Configuration Interaction Singles Approximation

Christoph Bannwarth* and Stefan Grimme*

Received 18th of February 2015, Published online 23rd of March 2015

Reprinted (adapted) with permission from†
Bannwarth, C.; Grimme, S. *J. Phys. Chem. A* **2015**, *119*, 3653–3662.

DOI 10.1021/acs.jpca.5b01680

Own manuscript contribution

- Implementation of the sTD-DFT method into the ORCA program package
- Performing the calculations
- Interpretation of the results
- Writing the manuscript

*Mulliken Center for Theoretical Chemistry, Institut für Physikalische und Theoretische Chemie, Rheinische Friedrich-Wilhelms-Universität Bonn, Beringstraße 4, 53115 Bonn, Germany

Abstract In the present work we show that the electronic circular dichroism (ECD) of delocalized π-systems represents a worst case scenario for Tamm-Dancoff approximated (TDA) linear response methods. We mainly consider density functional theory (TDA-DFT) variants together with range-separated hybrids but the conclusions also apply for other functionals, as well as the configuration interaction singles (CIS) approaches. We study the effect of the TDA for the computation of ECD spectra in some prototypical extended π-systems. The C_{76} fullerene, a chiral carbon nanotube fragment and [11]helicene serve as model systems for inherently chiral, π-chromophores. Solving the full linear response problem is inevitable in order to obtain accurate ECD spectra for these systems. For the C_{76} fullerene and the nanotube fragment, TDA/CIS approximated methods yield spectra in the origin-independent velocity gauge formalism of incorrect sign which would lead to the assignment of the opposite (wrong) absolute configuration. As a counterexample we furthermore study the ECD of an α-helix polypeptide chain. Here, the lowest energy transitions are dominated by localized excitations within the individual peptide units and TDA methods perform satisfactorily. The results may have far-reaching implications for simple semiempirical methods which often employ TDA/CIS for huge molecules. Our recently presented simplified TD-DFT (sTD-DFT) approach proves to be an excellent low-cost linear response method, which together with range-separated density functionals like ωB97X-D3, produces ECD spectra in very good agreement with experiment.

6.1. Introduction

Kohn-Sham density functional theory (KS-DFT) is now the most widely used method for electronic ground state calculations of molecules and solids. Due to the moderate computational cost and reasonable accuracy, time-dependent density functional theory (TD-DFT) [124,125,270,271] based on a KS-DFT reference has become the leading method used to calculate excited state properties and electronic spectra (see, e.g., Refs. 15,129,272–274 for reviews, for TD-DFT calculations of ground state properties like dispersion coefficients see, e.g., Refs. 276 and 275). It is routinely applicable to large systems (about a few hundred atoms) where correlated wave function based methods of similar accuracy are not feasible. The calculation of a few excited states by TD-DFT in very large systems with about 2000 atoms applying special software on graphics processing units has also been reported. [307]

However, the theoretical treatment of an entire electronic spectrum in a typical excitation energy range (e.g., up to 7 eV) for systems with several hundred atoms is still too demanding, in particular if hybrid TD-DFT is used. Applying the Tamm-Dancoff approximation to TD-DFT (TDA-DFT) reduces the computational cost but sacrifices the gauge-invariance of TD-DFT. It is well known that the sum rules for the oscillator and rotatory strengths are no longer fulfilled by TDA-DFT. [11,47] Oscillator strengths are unsigned quantities, therefore only their magnitude changes when going from TD-DFT to TDA-DFT. The resulting discrepancy between TD-DFT and TDA-DFT is often beyond the accuracy necessary for comparison of experimental and theoretical UV absorption spectra. For rotatory strengths, which are the signed quantities

required for electronic circular dichroism (ECD) spectra, the consequence of the Tamm-Dancoff approximation (TDA) is more severe. Thus, TDA-DFT is generally not used to compute ECD spectra, as long as a TD-DFT calculation is still feasible. For an overview on current state-of-the-art computation of ECD spectra, see Refs. 11,14,308,309.

To our knowledge, it has not been studied yet where TDA-DFT might still be applicable or how bad its performance actually is for so-called "worst cases". Similar considerations also hold for the configuration interaction singles (CIS) approach, which is equivalent to Tamm-Dancoff approximated time-dependent Hartree Fock (TD-HF).[129] It is often used in combination with semiempirical variants of Hartree-Fock (HF). Typically applied semiempirical approaches employ the zero differential overlap (ZDO) approximation[310,311] and have been particularly parametrized for the calculation of electronic spectra such as CNDO/S,[312–314] INDO/S,[315,316] or MSINDO-sCIS.[317–319] For a very recent variant (dubbed INDO/X) particularly designed for electronic excited states see Ref. 320. The CIS formalism has been used in previous studies to calculate the ECD spectra of various systems either in combination with an independent chromophore approach (then often termed "matrix method")[12,298,321–323] or in combination with the above mentioned semiempirical methods.[324–326] However, the limited applicability of these approaches to particular systems due to the use of the TDA/CIS approximation is rarely discussed. It is shown here that TDA/CIS methods can lead to computed ECD spectra of reasonable shape but with an almost mirror-image relation to the correct ones which would lead to incorrect absolute configuration assignments.

6.1.1. Suitable Methods to Compute Full Range ECD Spectra for Large Systems

There are only a few gauge-invariant approaches that are capable of computing a whole spectrum up to, e.g., 7 eV, in particular for large molecules with several hundred or even more than a thousand atoms. The fastest variant is the time-dependent density functional tight-binding (TD-DFTB) approach by Elstner and coworkers.[327] In terms of the computational cost, it is comparable to the semiempirical ZDO approaches, due to the use of a minimal basis set as well as integral approximations. In contrast to the HF-based semiempirical CIS approaches, the linear-response time-dependent (LR-TD) eigenvalue problem is solved instead of the Tamm-Dancoff approximated (TDA) one, retaining the gauge-invariance. The benefit from using a DFT-derived (i.e., from density functional tight-binding, dubbed DFTB) reference is the diagonal nature of the **A** - **B** matrix (see below), considerably reducing the computational cost. Some significant drawbacks are however, the small basis set (as in most semiempirical methods) and particularly the lack of exact Fock-exchange. The semi-local exchange results in a wrong description of charge-transfer (CT) and Rydberg states. For a discussion of the same problem in semi-local TD-DFT see Refs. 94 and 95. Attempts have been made to remove these spurious "ghost states" by removing configuration state functions (CSFs) with small electric transition dipole moments.[328] Such an approach may lead to removal of electrically forbidden, yet physically important CSFs as well (e.g., $n\pi^*$ transitions). It should also be noted that such a criterion is not applicable for ECD as the rotatory strength depends on both the electric and magnetic

transition dipole moment. As for most semiempirical methods, the parametrization of (TD-)DFTB is non-trivial and is thus restricted to a certain small set of elements.

Another method suited to compute entire ECD spectra for large systems is the simplified TD-DFT (sTD-DFT) approach that has recently been proposed by our group.[329] Similar to regular TD-DFT, it uses a (hybrid) KS-DFT reference but is orders of magnitude faster in the post-SCF procedure than TD-DFT, due to a drastic simplification of the two-electron integrals as well as a massive reduction of the single excitation configuration space (*vide infra*). The computational bottleneck is the preceding KS-DFT procedure to obtain the ground-state orbitals. It could be shown that this method produces electronic excitation spectra of similar quality as spectra from TD-DFT while the earlier reported simplified TDA-DFT variant (sTDA-DFT)[46] behaves like its non-simplified counterpart TDA-DFT.[329] Due to its good and representative performance and the possibility to combine it with common hybrid density functionals, this method will be used in this present work. For another approach to facilitate the computation of entire absorption spectra see Ref. 330.

6.1.2. Systems under Consideration

In the present work, we study the effect of the Tamm-Dancoff approximation (TDA) on the computed ECD spectra of three inherently chiral and structurally rigid chromophores. The first is the D_2 symmetric C_{76} fullerene, the second is a cutout from a the chiral (11,7) carbon nanotube. The smallest chromophore studied is the [11]helicene. All systems are highly conjugated π-systems with delocalized orbitals.

We first noticed a discrepancy between ECD spectra from sTDA-DFT and sTD-DFT for a substituted, chiral C_{60} derivative in our original paper on the sTD-DFT method.[329] Therein, the spectra (obtained from the velocity form, *vide infra*) computed by both approaches differed in sign over a wide energy range. Therefore, a further study of this phenomenon seems appropriate. In the present study, C_{76} was chosen because, due to its rigidity, we can ignore conformational effects on the ECD spectrum. In that respect, substituted chiral C_{60} derivatives (see, e.g., Ref. 305) are less suited. Furthermore, we may exploit its symmetry allowing us to additionally perform standard TD-DFT/TDA-DFT and TD-HF/CIS calculations on this system. According to the isolated pentagon rule,[331] there is only one chiral isomer of the C_{76} fullerene allowing direct comparison to the experimental ECD spectrum based on a single structure. This is a big advantage compared to the previously studied C_{84} fullerene where the occurrence of other isomers might have caused differences in the calculated and experimental ECD intensities.[332]

The second model system is a cutout from a chiral (11,7) carbon nanotube which has a diameter of about 13 Å.[333] It contains 180 carbon atoms and the edges are capped by hydrogens (18 on each edge), we will shortly denote it as (11,7)-hCNTh. This system is different compared to fullerenes since the π-system is non-continuous in one direction. The second difference is the connectivity within the π-system. Fullerenes contain hexagonic as well as pentagonic structural motifs, while carbon nanotubes contain only hexagons.

[11]helicene is the third inherently chiral chromophore to be studied here. An experimental

ECD spectrum has been presented a few years ago.[304] In our previous work, we have shown that most of the ECD bands are reproduced correctly by either sTDA-DFT with a range-separated hybrid (RSH)[195] or sTD-DFT with a global hybrid (BHLYP).[329] Even though the negative bands in the range of 280−340 nm were qualitatively reproduced, the ECD at 250−280 nm was improperly described. In the present study, we investigate this system by combining a gauge-invariant method (i.e., sTD-DFT) with an RSH functional.

Another conceptually different case to be studied is a polypeptide. Our model system is Ace-$(Ala)_{19}$-NMe, i.e., an *N*-terminus acylated, *C*-terminus methylated *poly*-Ala chain (in a right-handed α-helix conformation) containing in total 20 peptide chromophores. Here, the chromophores are not conjugated and instead separated by the α-carbons. The low-lying excitations may therefore be constructed from coupled, localized states.[334] This system with rather high-lying excitations (<300 nm) shall serve as a counterexample to highly delocalized π-systems.

We apply the sTDA-DFT and sTD-DFT approaches to these four systems. It has been shown that these methods produce spectra which are of similar quality as their conventional counterparts.[46,329] For further validation we compare the sTDA-DFT/sTD-DFT results to conventional TDA-DFT/TD-DFT for C_{76}. The ECD of C_{76} has previously been studied by means of semiempirical CIS,[325,326,335] as well as semi-local[336,337] and hybrid TD-DFT.[338] However, in the mentioned studies it was neither compared how well the length (R^L) and velocity (R^V) representations of the rotatory strengths match, nor was it studied how the TDA influences these entities. This is the focus of the present work. Before discussing the results, a brief summary of the underlying theory is given, followed by computational details.

6.2. Theoretical Background

6.2.1. Time-Dependent DFT and HF

The linear response TD-DFT equation is given by the following non-Hermitian eigenvalue problem[125,129]

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B^*} & \mathbf{A^*} \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} = \begin{pmatrix} \omega & 0 \\ 0 & -\omega \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix}, \tag{6.1}$$

where **A** and **B** are the orbital rotation Hessian matrices. **X** and **Y** are their eigenfunctions, and ω is a diagonal matrix with the respective excitation energy on the diagonal. In wave function theory, this equation corresponds to time-dependent Hartree-Fock (TD-HF, also called random-phase approximation, RPA). Whenever we discuss general features of methods relying on Eq. 6.1, i.e., being valid for both TD-HF and TD-DFT, we will simply refer to these methods as LR-TD (for time-dependent DFT/HF from linear-response theory). The matrix elements of **A** and **B** in the spin-restricted case are given by

$$\begin{aligned} A_{ia,jb} = \delta_{ij}\delta_{ab}(\epsilon_a - \epsilon_i) + 2(ia|jb) - a_x(ij|ab) \\ + (1 - a_x)(ia|f_{XC}|jb) \end{aligned} \tag{6.2}$$

and

$$B_{ia,jb} = 2(ia|bj) - a_x(ib|aj) + (1 - a_x)(ia|f_{XC}|bj). \tag{6.3}$$

Here, ϵ_a and ϵ_i are the orbital energies of the virtual and occupied orbitals, respectively, as they are obtained from the HF or KS-DFT ground state treatment. The two-electron integrals are abbreviated by the Mulliken notation, i.e., $(ia|jb) = \iint \psi_i(\mathbf{r}_1)\psi_a(\mathbf{r}_1)\frac{1}{r_{12}}\psi_j(\mathbf{r}_2)\psi_b(\mathbf{r}_2)\mathrm{d}\mathbf{r}_1\mathrm{d}\mathbf{r}_2$. Here, ij represent occupied and ab virtual orbitals. a_x is the amount of non-local Fock exchange that is mixed into the XC functional.[69] If $a_x = 1$, the above equations correspond to the TD-HF case.[129] Otherwise if $a_x = 0$, it corresponds to the semi-local TD-DFT equation.[125] The term $2(ia|jb)$ in $A_{ia,jb}$ ($2(ia|bj)$ in $B_{ia,jb}$) results from the response of the Coulomb part in the ground state and is thus identical in the HF and KS-DFT case. The terms in $A_{ia,jb}$ and $B_{ia,jb}$ which are scaled by a_x, correspond to the response of the non-local Fock exchange. Within the adiabatic approximation, the response of the density functional is given by the terms $(ia|f_{XC}|jb)$ and $(ia|f_{XC}|bj)$, in $A_{ia,jb}$ and $B_{ia,jb}$, respectively. Due to the missing $(ij|ab)$ term in $A_{ia,jb}$, pure semi-local TD-DFT severely underestimates Rydberg and CT state energies.

For real orbitals, the non-Hermitian LR-TD eigenvalue problem in Eq. 6.1 can be solved as the following Hermitian one[124,125]

$$(\mathbf{A} - \mathbf{B})^{\frac{1}{2}} (\mathbf{A} + \mathbf{B}) (\mathbf{A} - \mathbf{B})^{\frac{1}{2}} \mathbf{Z} = \omega^2 \mathbf{Z}\,. \tag{6.4}$$

Here, $\mathbf{Z}$ is given by

$$\mathbf{Z} = \sqrt{\omega}\,(\mathbf{A} - \mathbf{B})^{-\frac{1}{2}} (\mathbf{X} + \mathbf{Y})\,. \tag{6.5}$$

In semi-local DFT, the matrix $(\mathbf{A} - \mathbf{B})$ is diagonal significantly reducing the computational cost. Including non-local Fock exchange, as done in hybrid density functionals, generally improves the accuracy of excitation energies but at the expense of the non-diagonal nature of $(\mathbf{A} - \mathbf{B})$.

Hirata and Head-Gordon introduced the TDA[127] to TD-DFT (TDA-DFT) into the quantum chemistry community[128] (for a related precursor see Ref. 286). Here, the $\mathbf{B}$ matrix is neglected. This then leads to the simple eigenvalue problem

$$\mathbf{A}\mathbf{X}_{TDA} = \omega_{TDA}\mathbf{X}_{TDA}\,. \tag{6.6}$$

The eigenvector $\mathbf{X}$ and ω from Eq. 6.1 are replaced by $\mathbf{X}_{TDA}$ and ω_{TDA}, respectively, to clarify that these are not identical in TDA and LR-TD. For $a_x = 1$ with a HF reference, TDA is identical to the configuration interaction singles approach (CIS).[129] Whenever we discuss general features of methods relying on Eq. 6.6, i.e., CIS and TDA-DFT, we will simply refer to this as TDA.

The excitation energies obtained from TDA are in general similar to the ones from LR-TD, in particular for KS-DFT. In the hybrid DFT and HF case, TDA leads to some speed-up over solving the LR-TD equation since the matrix $(\mathbf{A} - \mathbf{B})$ does not need to be diagonalized.[128] For hybrids with a large amount of Fock exchange (and HF), TDA is more robust to triplet instabilities.[97,339] While TDA-DFT and TD-DFT are both widely used to simulate UV-spectra,

the former is typically not applied in computations of ECD spectra due to the missing gauge invariance. This is different in most semiempirical ZDO approaches. Due to the large computational cost associated with solving the LR-TD equation when Fock exchange is present, CIS has mostly replaced TD-HF as low-cost excited state method and it has also been used to compute ECD spectra in the past.[324–326] For an old semiempirical π-SCF based LR-TD approach for large chromophores see Ref. 340.

6.2.2. Simplified TDA-DFT and TD-DFT

In the recently proposed simplified approaches to TDA-DFT (sTDA-DFT)[46] and TD-DFT (sTD-DFT),[329] the following approximations lead to a significant reduction in the computational cost:

1. The response of the density functional (last term in Eqs. 6.2 and 6.3, respectively) is neglected,
2. the remaining two-electron integrals are approximated by atom-centered charge/transition density monopole interactions, and
3. the single excitation space is restricted to include all excitations up to a certain energy threshold.

The matrix elements of the approximated (denoted by a prime) matrices $\mathbf{A}'$ and $\mathbf{B}'$ (for sTD-DFT only) are given by (spin-restricted case)

$$A'_{ia,jb} = \delta_{ij}\delta_{ab}(\epsilon_a - \epsilon_i) + \sum_{C,D}^{N_{atoms}} (2q_{ia}^C \gamma_{CD}^K q_{jb}^D - q_{ij}^C \gamma_{CD}^J q_{ab}^D) \tag{6.7}$$

$$B'_{ia,jb} = \sum_{C,D}^{N_{atoms}} (2q_{ia}^C \gamma_{CD}^K q_{bj}^D - a_x q_{ib}^C \gamma_{CD}^K q_{aj}^D). \tag{6.8}$$

The transition/charge densities q_{pq}^C (located on atom C, the indices pq may refer to occupied or virtual orbitals) are obtained from a Löwdin population analysis.[46,287] $\gamma_{AB}^{K,J}$ are the Mataga-Nishimoto-Ohno-Klopman[288–290] damped Coulomb operators used to replace exchange (K) and Coulomb (J) type integrals, respectively. These contain the global parameters of the method (in total four general parameters for global hybrids and three functional-specific parameters for RSH functionals).[46,195] For more details on these methods, we refer to the original papers on sTDA-DFT,[46] its extension to RSH functionals[195] and on sTD-DFT.[329]

6.2.3. Rotatory Strengths: Length and Velocity Form

The condition for a transition to be visible in an ECD spectrum, is a non-vanishing rotatory strength[33]

$$R_{0\nu} = -\mathrm{Im}\{\vec{\mu}_{0\nu} \cdot \vec{m}_{0\nu}\} \tag{6.9}$$

with $\vec{\mu}_{0\nu}$ being the electric and $\vec{m}_{0\nu}$ the magnetic dipole moment (atomic units are used throughout) for the transition $0 \rightarrow \nu$.

$$\vec{m}_{0\nu} = \frac{i}{2} \left\langle 0 \left| \vec{r} \times \vec{\nabla} \right| \nu \right\rangle \tag{6.10}$$

For the electric dipole transition moment, two formalisms are in use. The dipole length form is given by

$$\vec{\mu}^L_{0\nu} = \langle 0 \left| \vec{r} \right| \nu \rangle \ , \tag{6.11}$$

while the dipole velocity form is

$$\vec{\mu}^V_{0\nu} = \frac{\vec{\nabla}_{0\nu}}{\omega_{0\nu}} = \frac{\left\langle 0 \left| \vec{\nabla} \right| \nu \right\rangle}{\omega_{0\nu}} \ . \tag{6.12}$$

Here, $\omega_{0\nu}$ is the excitation energy of the transition $0 \rightarrow \nu$. For gauge-invariant wave functions, Eqs. 6.11 and 6.12 are identical (e.g., see p. 427 of Ref. 54). This holds in the complete basis set (CBS) limit for LR-TD, however, not for TDA. Consequently, rotatory strengths from both formalisms are different in TDA. Being a signed property, the effect of TDA on the rotatory strengths may be severe.

The length form is generally thought to converge faster to the CBS limit, yet unfortunately, the rotatory strength derived from it (dubbed R^L) is not necessarily origin independent. The velocity form (R^V) ensures an origin independent expression of the rotatory strength (for both TDA/LR-TD with any basis set, for a proof see Refs. 297 or 329). Clearly, origin independence is a prerequisite for studying molecules with large spatial dimensions like proteins. Therefore, the velocity form of the rotatory strength is typically preferred.[297,299]

The quality of R^V is improved if the LR-TD equation is solved,[11,329] but the quality of R^V in the TDA case may highly depend on the molecule or excitation of interest. To assess the quality of R^V, a sole comparison with the experimental ECD spectrum is insufficient. For "smaller" systems, in which the origin dependence of R^L is not pronounced, R^V needs to be compared to R^L as well. Relying purely on comparisons with experiment might lead to an assignment of the incorrect enantiomer which will be demonstrated in Section 6.4.

6.3. Computational Details

We used the TURBOMOLE suite of programs[164a,165] (version 6.5) for all geometry optimizations employing symmetry restrictions if possible.

For the D_2 symmetric C_{76} fullerene, we used the (^{s,f}A)-C_{76} enantiomer ((^{f}C)-C_{76} in the old nomenclature)[341,342] such that our computed spectra are directly comparable to the ones from Refs. 325,326,336–338. In the starting geometry of Ace-$(Ala)_{19}$-NMe, we used the Ramachandran angles reported in Ref. 343 for the right-handed α-helix ($\phi = -62°$, $\phi = -41°$). Both structures were fully optimized at the TPSS level employing the atom-centered, spherical def2-TZVP basis set.[105,163] London dispersion interactions were treated by means of the atom-pairwise D3 dispersion correction[133] in combination with the rational damping scheme according

to Becke and Johnson.[134] This combined level will be denoted as TPSS-D3(BJ)/def2-TZVP.

The (*P*)-[11]helicene structure from Ref. 329 had already been optimized at the TPSS-D3(BJ)/def2-TZVP level and was taken as is. The geometry of (11,7)-hCNTh was fully optimized at the efficient, minimal basis Hartree-Fock level with three corrections (HF-3c).[169] Since the HF-3c and TPSS-D3(BJ)/def2-TZVP geometries for π-systems tend to be quite similar,[169] we do not expect any significant effect on the ECD spectrum from using an HF-3c geometry. All structures used can be found in the Supporting Information*.

The atom-centered, spherical def2-SV(P)[163,300] basis set has been used in the single-point calculations on (^{s,f}A)-C_{76}, (11,7)-hCNTh and Ace-$(Ala)_{19}$-NMe. This basis set was successfully employed in different studies,[46,302] and yields similar accuracy as, e.g., TZVP[162] for rotatory strengths in large molecules.[329] Due to the size of the studied systems, we prefer the use of RSH density functionals or global hybrid functionals with a large amount of Fock-exchange. In this study, we employ the ωB97X-D3 functional,[114] which is available in the ORCA program.[48] This functional shows the correct long-range behavior (i.e., 100% Fock-exchange at large inter-electronic separations). For the sTDA-DFT and sTD-DFT calculations with this functional, a development version of the ORCA program is used with the parameters previously reported in Ref. 195 (note however that $\alpha^{(1)}$ given in Table 1 therein should actually be $\beta^{(1)}$ and *vice versa*).

Comparison with conventional TDA-DFT and TD-DFT was done for (^{s,f}A)-C_{76} using the global hybrid functional BHLYP[69,99,100] (i.e., 50% of Fock-exchange) as implemented in the TURBOMOLE suite of programs.[164a,165] The `escf` module[301] allows to exploit the symmetry of the D_2-symmetric fullerene. Within each of the irreducible representations B_1, B_2, and B_3, 30 roots were calculated. The simplified methods use the same ground-state reference determinant as the conventional ones, but symmetry is not exploited in the sTDA-DFT/sTD-DFT calculations. Similarly, for (^{s,f}A)-C_{76}, *ab initio* CIS and TD-HF computations were conducted. The behavior of sTDA-DFT and sTD-DFT for (11,7)-hCNTh is studied at the ωB97X-D3/def2-SV(P) level. For Ace-$(Ala)_{19}$-NMe, sTDA-DFT and sTD-DFT calculations are performed on a ωB97X-D3/def2-SV(P) as well as a BHLYP/def2-SV(P) reference. The computed ECD spectra of the somewhat smaller (*P*)-[11]helicene were obtained from sTDA-DFT and sTD-DFT on a ωB97X-D3/def2-TZVP reference. For comparison, we also give the respective ECD spectra at the ωB97X-D3/def2-SV(P) level in the Supporting Information.

In all sTDA-DFT/sTD-DFT calculations, the primary configuration space (P-CSF space, see Ref. 46 for details) included all configurations with uncoupled one-electron excitation energies smaller than 10 eV. Beyond that energy, configuration state functions (CSF) were included if their perturbatively estimated net coupling to the P-CSFs is larger than 10^{-4} Hartree.

The computed rotatory strengths were always broadened by Gaussian functions with a full width at $1/e$ maximum of 0.4 eV to empirically simulate vibronic effects in the computed spectra.[11] All computations were performed with the origin located at the geometric center.

*Cartesian coordinates of all studied compounds and the ECD spectrum of (*P*)-[11]helicene at the ωB97X-D3/def2-SV(P) level: this material is available online at http://dx.doi.org/10.1021/acs.jpca.5b01680. The ECD spectrum of (*P*)-[11]helicene at the ωB97X-D3/def2-SV(P) level is included also in Appendix A4.

6.4. Results and Discussion

6.4.1. The (^{s,f}A)-C_{76} Fullerene

Figure 6.1a shows the ECD spectra as obtained from the length (R^L) and velocity (R^V) expressions computed by TD-DFT[125] and TDA-DFT[128] employing the BHLYP[69,99,100] functional. The respective spectra from the simplified variants, sTD-DFT[329] and sTDA-DFT,[46] respectively, are given in Figure 6.1b. Our TD-DFT and sTD-DFT spectra agree well with the one obtained from TD-BHLYP/6-31G*[344,345] as reported in Ref. 338.

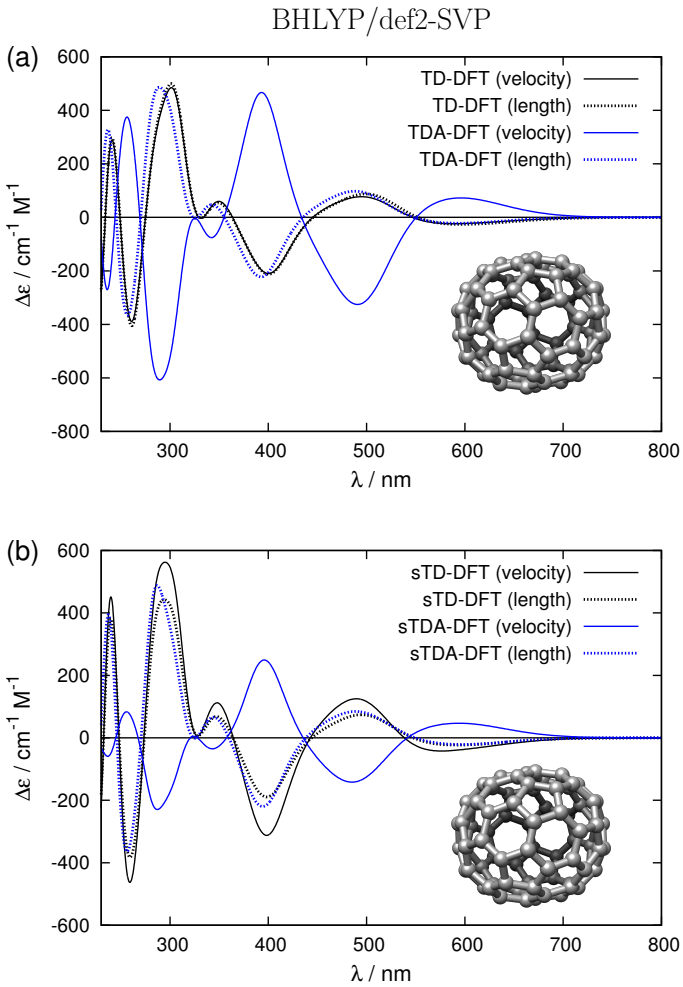


Figure 6.1.: (a) Comparison of calculated ECD spectra obtained from the velocity (R^V, solid lines) and length (R^L, dashed lines) forms of the rotatory strengths. The black lines refer to the TD-DFT and the blue lines to the TDA-DFT results obtained for (^{s,f}A)-C_{76}. The spectrum is blue-shifted by 0.35 eV. (b) The same as in (a) but for sTD-DFT (black) and sTDA-DFT (blue) with a blue-shift of 0.90 eV.

For both TD-DFT and sTD-DFT, the spectra obtained via R^L and R^V are almost identical. With TDA-DFT and sTDA-DFT, R^L leads to qualitatively the same ECD spectrum as either

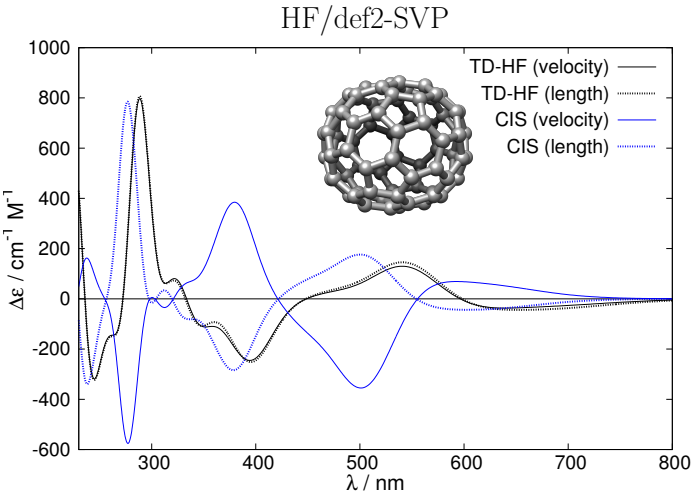


Figure 6.2.: Comparison of calculated ECD spectra obtained from the velocity (R^V, solid lines) and length (R^L, dashed lines) forms of the rotatory strengths. The black lines refer to the TD-HF and the blue lines to the CIS results obtained for (^{s,f}A)-C_{76}. The computed spectra are red-shifted by $-0.25\,\mathrm{eV}$.

form of the rotatory strength obtained from an LR-TD procedure. R^V, however, yields a spectrum with almost exactly the opposite sign in the studied spectral range. Since TDA-DFT and sTDA-DFT behave very similar in that aspect, this error can solely be attributed to the TDA. This result would lead to an incorrect assignment of the absolute configuration and is thus considered as a disastrous TDA failure.

Since semiempirical CIS variants were used for this system in the past,[325,326] we also study the ECD of this fullerene by the corresponding Hartree-Fock based methods (CIS and TD-HF). As semiempirical variants like CNDO, INDO, etc. are approximations to HF, the insights gained from this comparison may apply in the same way to the semiempirical approaches like the former. In Figure 6.2, the ECD spectra (both R^L and R^V) calculated by TD-HF and CIS are given. It is apparent that similar to the DFT case, solving the LR-TD equation is necessary to obtain conclusive results. Like TD-DFT, TD-HF produces almost identical ECD spectra from either R^L or R^V. CIS exhibits the same problem as TDA-DFT, namely prediction of the incorrect enantiomer by R^V.

Even though the R^L formalism works well in this particular case, we emphasize that its use is generally not recommended to study the ECD of extended systems as it is not origin-independent. An inherently chiral chromophore like C_{76} is apparently a worst case scenario where the matrix $\mathbf{B}$ is non-negligible and the TDA is not valid. This can be explained by considering the matrix elements of $\mathbf{B}$ (Eq. 6.3) and $\mathbf{B}'$ (Eq. 6.8) containing two-electron integrals exclusively of exchange-type. Only if the orbitals i and a are of similar type (e.g., π and π^*) and spatially close, the respective exchange-type integrals are significant. This is the case for systems with delocalized orbitals.

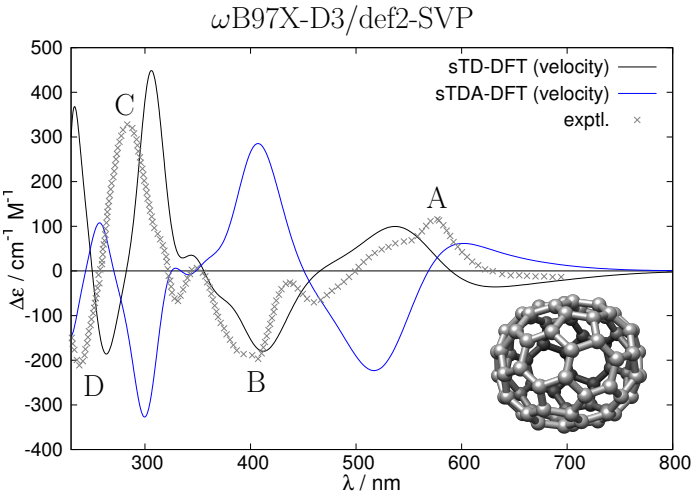


Figure 6.3.: Comparison of calculated ECD spectra of (^{s,f}A)-C_{76} computed with ωB97X-D3. The spectra are obtained from the velocity (R^V, solid lines) representation of the rotatory strengths. The black line refers to sTD-DFT and the blue line to sTDA-DFT. Both computed spectra are blue-shifted by 0.25 eV. The experimental curve (gray points) has been redrawn from Ref. 325.

In Figure 6.3, the ECD spectra of (^{s,f}A)-C_{76} computed by sTD-DFT and sTDA-DFT (both R^V) based on a ωB97X-D3 ground-state are presented together with the experimental spectrum measured by Diederich and coworkers.[325] Over the whole spectral range, the major experimental bands are nicely reproduced by sTD-DFT. Only the computed bands C and D are slightly red-shifted compared to the experiment. Thus, sTD-DFT in combination with a RSH functional like ωB97X-D3 is well-suited to determine the absolute configuration of inherently chiral, conjugated chromophores. The sTDA-DFT spectrum with the incorrect, almost opposite ECD is shown for completeness.

Diederich and coworkers[325] assigned the absolute configuration by comparing with computed ECD spectra via the π-electron SCF-CI-DV MO (self consistent field-configuration interaction-dipole velocity molecular orbital) method,[335] i.e., a semiempirical CIS variant relying on R^V. Somewhat unexpected regarding the observations made above, their assignment of the absolute configuration agrees with ours based on TD-DFT/TD-HF/sTD-DFT.

Our spectra are shifted such that the energetically lowest experimental and computed bands align energetically. This way, the negative ECD beyond 650 nm as well as the bands A, B, and C in Figure 6.3 are reproduced by the respective LR-TD spectrum (i.e., from sTD-DFT/TD-DFT/TD-HF). The energetically lowest band is always displayed in our spectra. It can only be speculated why the original assignment is correct, despite of the fact that a TDA approach relying on R^V has been used. Just to mention one possible explanation, we note that the calculated spectrum in Ref. 325 is not shifted and the lowest energy band is either truncated or might not be visible at all. It is therefore difficult to trace whether the positive band at 281 nm,

which seems to be used in Ref. 325 for the assignment, does correspond to the experimental band C (in Figure 6.3) or whether it is a neighboring band produced with opposite sign due to the use of R^V from CIS.

In any case, based on the LR-TD results, our assignment of the absolute configuration agrees with the original assignment made in Ref. 325. Our study is, however, supposed to serve as a guide for future works and as a reminder that TDA-DFT/CIS based ECD spectra need to be judged carefully. sTD-DFT proves to be a reasonable alternative as long as the preceding KS-DFT ground-state calculation is still feasible.

6.4.2. The Carbon Nanotube Model

The computed ECD spectra for the carbon nanotube model, dubbed (11,7)-hCNTh, are given in Figure 6.4. As there is no experimental spectrum to compare with, we plot the computed spectra on a linear energy scale.

The lowest vertical excitation energy is very small (sTD-DFT: 0.2 eV, sTDA-DFT: 0.5 eV), thus we applied a shift of 0.5 eV to the sTD-DFT spectrum (0.2 eV to the sTDA-DFT spectrum) to make the Gaussian-broadened, first band entirely visible (see Figure 6.4a). Notably, the length and velocity forms obtained from sTD-DFT yield almost identical spectra over the whole studied energy range. Only the two lowest bands show differences in intensity, in particular the lowest one. The velocity form of the electric dipole moment $\vec{\mu}^V_{0\nu}$ is reciprocally proportional to the respective excitation energy $\omega_{0\nu}$ (see Eq. 6.12). The small excitation energies for the two lowest transitions lead to an overestimated magnitude of $\vec{\mu}^V_{0\nu}$, and consequently of $R^V_{0\nu}$ as well. Since $\vec{\mu}^V_{0\nu}$ is just scaled by $\frac{1}{\omega_{0\nu}}$, its direction remains unchanged. Therefore, the length and velocity forms from sTD-DFT both allow a definite absolute configuration assignment.

sTDA-DFT in the velocity form (see Figure 6.4b) yields an opposite sign for the two lowest bands similar to the behavior observed for the C_{76} fullerene. Up to about 4 eV, the R^V spectrum is also considerably different from the R^L-based and the sTD-DFT spectra. Beyond 4 eV, the mutual agreement between the length and velocity forms from sTDA-DFT is good, likely due to the fact that these transitions are more localized involving σ/σ^* orbitals. However, the agreement with the sTD-DFT spectra especially beyond 6 eV is still not satisfactory.

The ECD of the chiral nanotube is badly described by TDA, especially in the energy range where experimental measurements are typically carried out (>200 nm, <6 eV). It is therefore necessary to apply methods that solve the full LR-TD problem to obtain reliable results.

6.4.3. (*P*)-[11]Helicene

Another class of inherently chiral, aromatic chromophores are the helicenes. Their ECD was studied in detail by semi-local TD-DFT[303] and INDO/S.[324] The largest helicene with an experimental ECD spectrum available is [11]helicene (comprising eleven phenyl rings). Previous studies of our group showed that most of its ECD bands are well reproduced by either sTDA-DFT or sTD-DFT.[195,329] However, the overall negative ECD in the region of 250−340 nm could

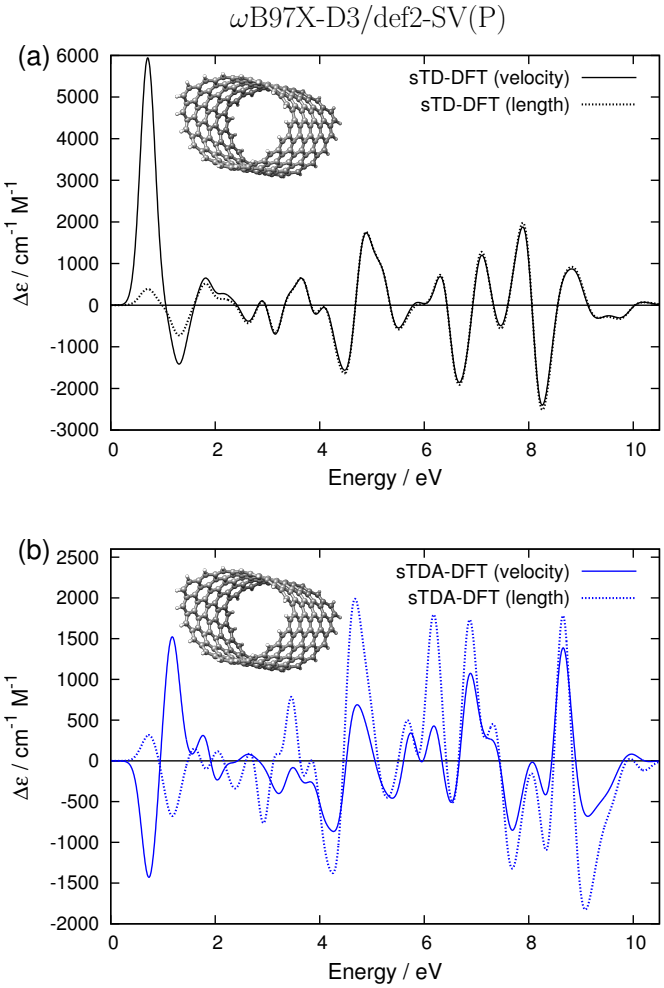


Figure 6.4.: (a) Calculated sTD-DFT ECD spectra (black) obtained from the velocity (R^V, solid line) and length (R^L, dashed line) forms of the rotatory strengths. These spectra are blue-shifted by 0.5 eV. (b) The sTDA-DFT ECD spectra (blue) from the velocity (R^V, solid) and length (R^L, dashed line) forms are blue-shifted by 0.2 eV.

not be reproduced entirely. Up to this point, the best qualitative agreement was obtained from sTD-DFT based on the global hybrid functional BHLYP. In a recent sTDA-DFT study, RSH functionals in the length representation (R^L) showed improved spectra in this region compared to global hybrids.[195]

Here, we re-compute the ECD spectrum of this helicene by sTD-DFT/ωB97X-D3. In Figure 6.5, the experimental ECD spectrum is given along with the calculated spectra from sTD-DFT (and sTDA-DFT for comparison).

The sTD-DFT spectrum is in very good agreement with the experiment over the complete displayed range. Only band D (in Figure 6.5) is not produced with a slightly positive ECD but as a distinguishable local maximum between two neighboring negative bands which may be identified as band C and E, respectively. Comparing with the corresponding sTDA-DFT

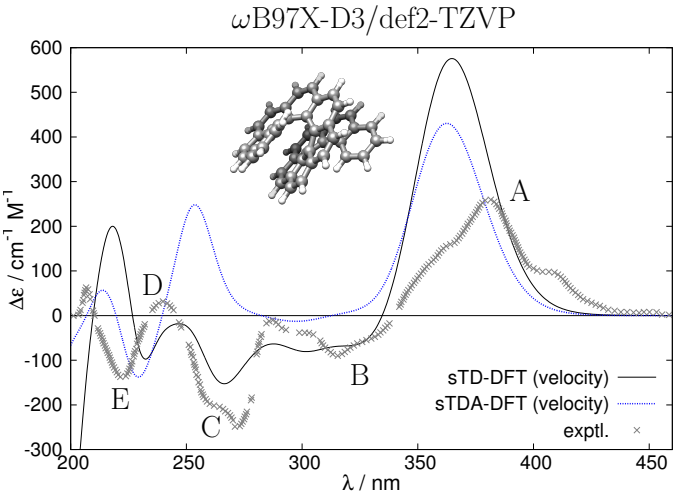


Figure 6.5.: Calculated ECD spectra of (*P*)-[11]helicene computed with ωB97X-D3 in the velocity representation (R^V). The black line refers to sTD-DFT and the dotted blue line to sTDA-DFT. The experimental curve (gray points) has been redrawn from Ref. 304.

spectrum, it is apparent that band C is of opposite sign leading to significant positive ECD (enhanced by the neighboring band D). Even though the effect of the TDA is less severe compared to the previous two cases, a large part of the spectrum is incorrectly described by sTDA-DFT. Thus, apart from employing an asymptotically correct functional, solving the LR-TD eigenvalue problem is necessary to obtain reliable ECD spectra of this inherently chiral chromophore. To our knowledge, the computed sTD-DFT/ωB97X-D3 ECD spectrum shows the best agreement to the experimental spectrum outperforming previous approaches.[195,329]

6.4.4. α-Helical Ace-(Ala)$_{19}$-NMe

The last system considered is Ace-(Ala)$_{19}$-NMe in the right-handed α-helix conformation. The ECD spectrum of polypeptides in the range of 180−250 nm is governed by coupled $n\pi^*$ and $\pi\pi^*$ excitations localized on the inherently non-chiral peptide chromophores.[26,297,334] Being a representative example for polypeptides in general, the ECD spectrum of Ace-(Ala)$_{19}$-NMe in the range of 180−250 nm can therefore be considered to be conceptually different situation compared to the previously discussed, inherently chiral compounds.

For two singly excited electronic configurations which are located on two different chromophores, the exchange type integrals found in the off-diagonal elements of matrix **B** (and **B′**), are expected to be vanishingly small. In such a case, the TDA should be reasonable.

Comparison of the sTDA-DFT and sTD-DFT spectra (R^V as well as R^L) obtained from a ωB97X-D3 (Figure 6.6a) or BHLYP (Figure 6.6b) reference shows that this is indeed true. All bands (corresponding to the experimental bands A, B, and C) are well reproduced within both representations by sTD-DFT as well as sTDA-DFT. With sTDA-DFT and ωB97X-D3 (Figure

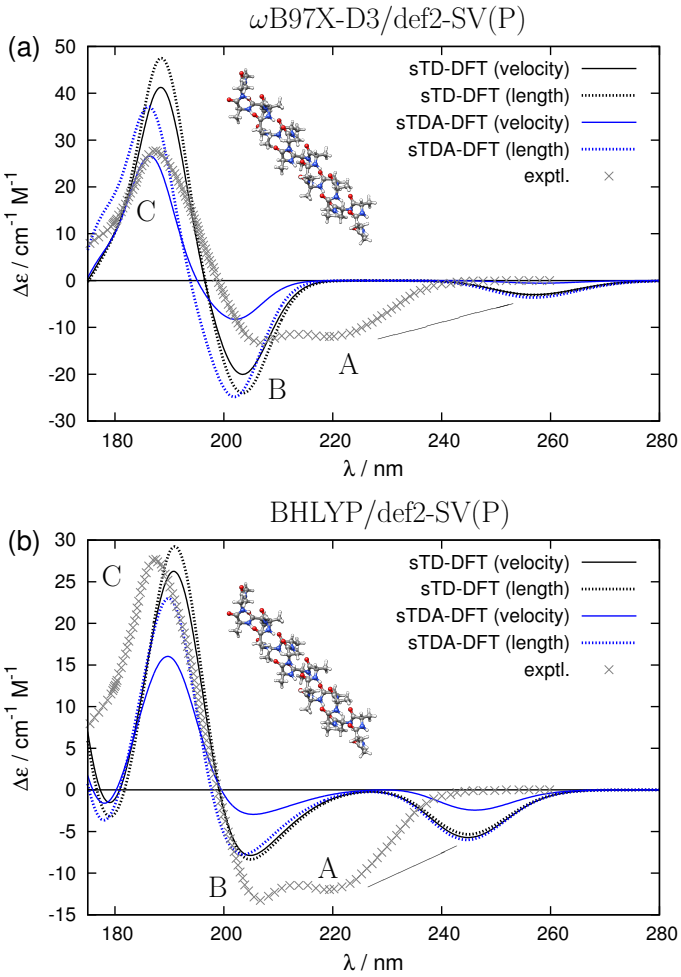


Figure 6.6.: Calculated ECD spectra of Ace-$(Ala)_{19}$-NMe computed with (a) ωB97X-D3 and (b) BHLYP. The computed spectra are red-shifted by 1 eV. The black lines refer to spectra obtained from sTD-DFT and the blue lines from sTDA-DFT, respectively. The experimental spectrum (gray points) is from Ref. 346.

6.6a) however, the intensity of band A is too small if R^V is used.

The experimental spectrum given in Figure 6.6 does not refer to the Ace-$(Ala)_{19}$-NMe model used in the computations but was obtained from α-helical poly-γ-methyl glutamate measured in hexafluoro-propanol.[346] As the three bands A, B, and C are quite distinctive and common for any α-helical polypeptide (see Ref. 26 for the ECD spectrum of myoglobin), this spectrum serves as a reference for the ECD exhibited by a perfect α-helix in this region. The computed intensities were scaled by the number of peptides in Ace-$(Ala)_{19}$-NMe yielding the mean peptide ECD.

This example shows that the ECD of polypeptides in the considered energy regime is well described by TDA methods. That explains why independent chromophore approaches[298] applied in a TDA manner (often termed "matrix method" in that context), were successfully applied for

decades[322,323,347,348] to study the ECD of polypeptides (for a very recent review on protein ECD see Ref. 349). Nevertheless, these methods introduce further approximations and empiricism. As long as the KS-DFT single-point calculation is still feasible, sTDA-RSH/SV(P) based studies on related systems may be a viable alternative to independent chromophore approaches.

6.5. Conclusion

In this work we demonstrated that for large π-systems, Tamm-Dancoff approximated (TDA) variants of TD-DFT or TD-HF (i.e., TDA-DFT and CIS, respectively) fail dramatically. The computed ECD spectra obtained from the length (R^L) and velocity (R^V) forms of the rotatory strengths with TDA methods are more or less of opposite sign. In particular, the origin independent R^V based results are incorrect. This was demonstrated by comparison with results obtained from solving the full time-dependent linear response (LR-TD) equation as well as by comparison with experimental spectra. Due to its origin-dependence, the R^L form from TDA-DFT/CIS is no alternative when spatially extended systems are of interest. For LR-TD methods like TD-HF and TD-DFT, both R^L and R^V lead to almost identical ECD spectra and one can infer that the contribution from the **B** matrix, which is neglected in TDA-DFT/CIS, is non-negligible.

The simplified TD-DFT approach (sTD-DFT)[329] shows the same behavior as regular TD-DFT. In combination with the range-separated hybrid (RSH) functional ωB97X-D3, sTD-DFT provides excellent agreement between the computed and experimental spectra at drastically reduced computational cost.

Three inherently chiral π-chromophores were studied by means of sTD-DFT and its TDA variant sTDA-DFT:[46] The (^{s,f}A)-enantiomer of the C_{76} fullerene, a cutout from the chiral (11,7) carbon nanotube and (*P*)-[11]helicene. In the experimentally accessible energy range (>200 nm, <6 eV), the ECD spectra (R^V form) of the former two systems are computed mostly with opposite (wrong) sign by sTDA-DFT. sTD-DFT provided excellent spectra with its R^L and R^V forms always in agreement to each other. For the C_{76} fullerene, our findings were verified against conventional TD-DFT/TD-HF and TDA-DFT/CIS calculations. While TDA-DFT/CIS exhibited the same dramatic failure as sTDA-DFT, the TD-DFT/TD-HF showed the same good performance as sTD-DFT.

For the (*P*)-[11]helicene, the effect of TDA is less pronounced. However, by combining a range-separated hybrid functional (ωB97X-D3) with sTD-DFT, the ECD spectrum of (*P*)-[11]helicene was reproduced with an unprecedented agreement to the experimental spectrum.[304] In disagreement to the experiment, the respective ECD spectrum computed by sTDA-DFT shows bands of opposite sign in the range of 250−280 nm.

For comparison, the ECD of a long α-helical polypeptide chain was studied by sTD-DFT and sTDA-DFT. In contrast to the previous compounds, the low-lying excitations are relatively localized on the peptide chromophores and the TDA is valid. Hence, the ECD spectra obtained from sTDA-DFT by either R^L or R^V are quite similar to each other and also similar to the

respective sTD-DFT ones.

Whenever solving the LR-TD equation is prohibitive, one should be very careful when using TDA approaches to study ECD. The discrepancy between the length and velocity formalisms of the rotatory strengths must be considered, for example, if semiempirical variants of CIS are applied.

For broad range spectra of the systems studied here, the sTD-DFT[329] approach has proven to yield reliable results similar to regular TD-DFT. Being significantly faster, sTD-DFT is also applicable to quite large systems as long as the preceding KS-DFT ground-state calculation is computationally feasible.

Acknowledgments

This work was supported by the Fonds der Chemischen Industrie and the DFG in the framework of the SFB 813 ("Chemistry at Spin-Centers"). C. B. thanks Dr. A. Hansen and U. Becker for their help related to the ORCA implementation of sTD-DFT, as well as J. G. Brandenburg for helpful comments on the manuscript.

7. The Electronic Circular Dichroism of [16]Helicene with Simplified TD-DFT: Beyond the Single Structure Approach

Christoph Bannwarth,* Jakob Seibert,* and Stefan Grimme*

Received 12th of January 2016, Published online 12th of April 2016

Reprinted (adapted) with permission from†
Bannwarth, C.; Seibert, J.; Grimme, S. *Chirality* **2016**, *28*, 365–369.

DOI 10.1002/chir.22594

Own manuscript contribution

- Supervising all computations
- Performing the PBEh-3c geometry optimizations and harmonic frequency calculations
- Calculation of the absorption and ECD spectra on the minimum geometries
- Interpretation of the computed data
- Writing the manuscript

*Mulliken Center for Theoretical Chemistry, Institut für Physikalische und Theoretische Chemie, Rheinische Friedrich-Wilhelms-Universität Bonn, Beringstraße 4, 53115 Bonn, Germany

Abstract The electronic circular dichroism (ECD) spectrum of the recently synthesized [16]helicene and a derivative comprising two triisopropylsilyloxy protection groups is computed by means of the very efficient simplified time-dependent density functional theory (sTD-DFT) approach. Different from many previous ECD studies of helicenes, non-equilibrium structure effects are accounted for by computing ECD spectra on "snapshots" obtained from a molecular dynamics (MD) simulation including solvent molecules. The trajectories are based on a molecule specific classical potential as obtained from the recently developed quantum chemically derived force field (QMDFF) scheme. The reduced computational cost in the MD simulation due to the use of the QMDFF (compared to *ab initio* MD), as well as the sTD-DFT approach make realistic spectral simulations feasible for these compounds that comprise more than 100 atoms. While the ECD spectra of [16]helicene and its derivative computed vertically on the respective gas phase, equilibrium geometries show noticeable differences, these are "washed" out when non-equilibrium structures are taken into account. The computed spectra with two recommended density functionals (ωB97X and BHLYP) and extended basis sets compare very well with the experimental one. In addition we provide an estimate for the missing absolute intensities of the latter. The approach presented here, could also be used in future studies to capture non-equilibrium effects, but also to systematically average ECD spectra over different conformations in more flexible molecules.

7.1. Introduction

Time-dependent density functional theory (TD-DFT)[124,125,270,271] is the most popular method for the computation of excited states in medium sized molecules (up to $\approx$ 100 atoms). In order to compute electronic spectra in a broad energy range for these and even larger molecules, our group developed the simplified TD-DFT (sTD-DFT) approach.[46,195,329] Like regular TD-DFT, it computes excited states from a (hybrid) Kohn-Sham reference state, but is orders of magnitude faster due to a reduced matrix eigenvalue problem size and drastic integral simplifications (monopole approximation) in the TD-DFT treatment. In the present work, we apply the sTD-DFT approach to compute the electronic circular dichroism (ECD) spectrum of a recently reported [16]helicene[350] derivative.

Helicenes, or more precisely carbohelicenes,[351] are screw-symmetric, chiral isomers of acenes. They represent a class of inherently chiral, conjugated chromophores and consequently, their ECD is of particular interest. The very first one, [5]helicene, was already synthesized in 1918.[352] Newman *et al.* were the first to synthesize the chirally resolved [6]helicene and report on its optical rotation[353,354] and ECD.[355] Since then, the optical properties of different carbohelicenes were studied both experimentally[356–358] and theoretically.[303,324,359] We refer the interested reader to recent review articles on helicenes.[360,361]

The ECD spectrum of [11]helicene (one of the largest experimentally described ones) was reported in 2009.[304] It has been studied by sTD-DFT[329,362] and its Tamm-Dancoff approxi-

mated variant (sTDA-DFT)[195,362] in our group. It turned out that for good agreement with the experimental ECD spectrum, sTD-DFT needs to be combined with a range-separated hybrid (RSH) density functional with asymptotically 100% of non-local Fock exchange (e.g., the ωB97 class).[362] With sTDA-DFT or global hybrid functionals, the ECD spectrum between 250 and 340 nm (3.6–5.0 eV) was not or at most only qualitatively reproduced. Very recently, the synthesis of [16]helicene (**[16]H**) has been reported[350] which – to our knowledge – represents the largest synthesized carbohelicene. For a derivative comprising two triisopropylsilyloxy (TIPSO) protection groups (at positions 3 and 34), an experimental absorption as well as an ECD spectrum were recorded. However, for the latter, only the relative intensities were given. In this study, we compute the ECD spectrum of **[16]H** (and its TIPSO derivative, denoted as **[16]H-tipso**) with sTD-DFT.[329] Together with a RSH functional, this approach already performed very well for the smaller [11]helicene.[362] The **[16]H** consists of 102 and the **[16]H-tipso** of 164 atoms. They comprise 16 annulated benzene rings and a screw with roughly 2.5 turns (see Figure 7.1).

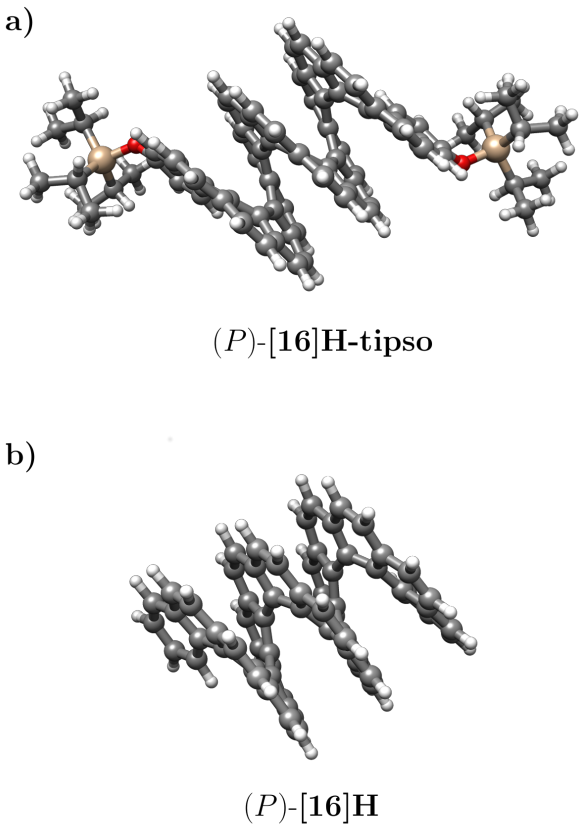


Figure 7.1.: Ball-and-stick representations of the minimum structures of a) (*P*)-**[16]H-tipso** and b) (*P*)-**[16]H** optimized at the PBEh-3c level.

7.2. Results and Discussion

One aspect of this study is to provide an estimate for the lacking experimental absolute intensities by comparison of the relative ECD spectrum from Ref. 350 with our computed spectrum. For this purpose we go beyond the single (minimum) structure approach that is typically applied to rigid π-systems like helicenes.[46,195,303,324,329,362] This is accomplished by calculating ECD spectra on non-equilibrium structures ("snapshots") from a molecular dynamics (MD) trajectory. For earlier works in which ECD was computed on a MD trajectory, see Refs. 363–369. Since the helicene framework itself shows only one energetically low-lying minimum, the intention is not to consider other conformers but to study the effect of non-equilibrium structures (i.e., elongated bonds and ring torsions/twisting).

First, the effect of the employed density functional for the sTD-DFT calculation is to be tested. This is done for a single minimum geometry. Therefore, the ground state geometry of (*P*)-**[16]H-tipso** is optimized at the accurate and efficient PBEh-3c level.[110] In this geometry, the carbons at position 3 and 34 are separated by 11.79 Å. After verifying the geometry as a minimum, the absorption and ECD-spectra including all singly excited states up to 10 eV are computed vertically by sTD-ωB97X/def2-TZVP[162,163,370] and sTD-BHLYP/def2-TZVP[69,99,100,162,163] ($\approx$ 1000 and 2100 states, respectively). The spectra are given in Figure 7.2 along with the experimental ones (note that the absolute intensities for the experimental ECD spectrum were estimated as described below). Throughout this work, the length form of the electric dipole transition moment is used for oscillator strengths (i.e., absorption spectra), while the velocity form is used the rotatory strengths (ECD spectra) always guaranteeing gauge origin independence.[47,297] The TURBOMOLE[164c,165,371] suite of programs (version 7.0) is used for all geometry optimizations and harmonic frequency calculations. A development version (based on the 3.0 release) of the ORCA[48] program is used for the sTD-DFT (and preceding ground state) single-point calculations.

Both computed spectra are shifted such that the computed distinctive absorption bands roughly match the experimental one at 4.8 eV. The computed intensities for this band agree nicely with the experiment. In the shifted sTD-BHLYP spectrum, the lower lying bands are red-shifted compared to the experiment (at 3.2 and 3.9 eV). For sTD-ωB97X, the same can be seen for the lowest lying band, however, the absorption band corresponding to the experimental one at 3.9 eV occurs at slightly higher energies. Therefore the sTD-ωB97X absorption spectrum below 4.4 eV appears to be "stretched" compared to the experiment. This might be a consequence of the range-separation parameter (0.3 a.u.) which was found to be suboptimal for large π-conjugated systems.[372,373]

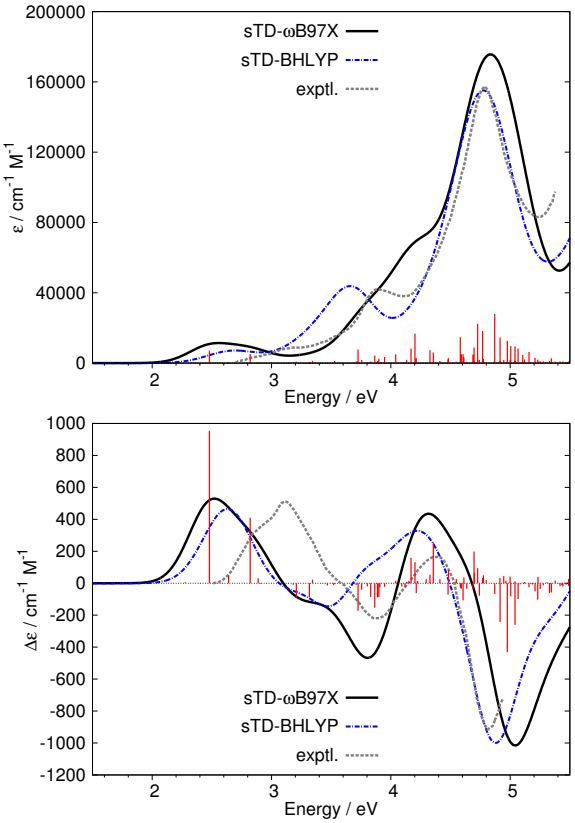


Figure 7.2.: Calculated absorption and ECD spectra of (*P*)-**[16]H-tipso** computed with sTD-DFT on an ωB97X (black solid line) and BHLYP (blue dashed line) reference. The def2-TZVP basis set has been used throughout. The individual transition strengths are broadened by Gaussians with a full width at $1/e$ maximum of 0.5 eV and the spectra are red-shifted by 1 eV and 0.3 eV, respectively. The oscillator and rotatory strengths from the sTD-ωB97X calculation are depicted as red sticks. The experimental curves (gray dotted line) are taken from Ref. 350 and refer to a solution of (*P*)-**[16]H-tipso** in $CHCl_3$. The absolute intensities of the experimental ECD spectrum are not given therein but are estimated in this work.

The observations made for the absorption spectra also hold for the ECD spectra. In addition, it can be seen that the energetically lowest, positive feature in the experimental ECD spectrum seems to consist of two bands, a shoulder towards lower energies (around 2.8 eV) and a maximum around 3.2 eV. While the sTD-BHLYP spectrum shows only one band in this region, a double band feature is present for sTD-ωB97X, however, with reversed relative intensities compared to the experiment. Overall, the agreement with the experimental ECD spectrum is better for sTD-ωB97X, which is used in the following. However, it should be noted that apart from shifts sTD-BHLYP qualitatively reproduces all bands in the spectrum, which is different from the previously studied [11]helicene.[329] In Figure 7.2, we have also plotted the individual oscillator

and rotatory strengths from the sTD-ωB97X/def2-TZVP calculation which reveal two dominant transitions in the low energy range of the ECD spectrum. It is in principle possible to analyze these in more detail (e.g., plotting transition densities)[374], however, this is beyond the scope of this study.

To go beyond the minimum structure approach, the PBEh-3c computed Hessian matrix (scaled by 0.95[110]) of (*P*)-**[16]H-tipso** is used to parametrize a force field following the procedure described in Ref. 49. This quantum mechanically derived force field (QMDFF) is then used in a standard NPT MD simulation of (*P*)-**[16]H-tipso** inside a box of 250 $CHCl_3$ molecules with periodic boundary conditions. The in-house program QMSIM[375] is used for this purpose. After equilibrating at 300 K, an MD run is performed for 0.5 ns with a timestep of 0.5 fs (more details can be found in the Supporting Information*). All C–H bond lengths are kept constrained during the simulation using the SHAKE algorithm.[376,377] From this MD trajectory, 100 equidistantly selected structures are considered for sTD-DFT calculations (vertical excitations) at the ωB97X/def2-TZVP level. Because the vibrational line broadening is explicitly included in these calculations, the width of the Gaussian intensity distributions used to simulate the ECD spectra are reduced to 0.35 eV in this case. The resulting absorption and ECD spectra are given in Figure 7.3 along with the computed minimum structure and experimental spectra.

It can be seen that in the computed absorption spectrum based on the MD structures all bands are red-shifted (0.1–0.3 eV) relative to the spectrum on the minimum geometry. The intense absorption band around 4.8 eV is split into two separate bands. Concerning the ECD spectrum, no severe changes (i.e., in sign) are observed. Yet, the splitting of the positive double band at low energies is enhanced by the dynamics. Furthermore, the relative ratio of the positive ECD band at 4.3 eV to the intense negative one at 5 eV decreases, which is in better agreement with the relative ratio in the experiment. An overlay of the individual spectra for all MD structures can be found in the Supporting Information.

*The supporting information (excluding the Cartesian coordinates of **[16]H-tipso** and **[16]H**) is included in Appendix A5. The complete supporting information may be found online at http://dx.doi.org/10.1002/chir.22594.

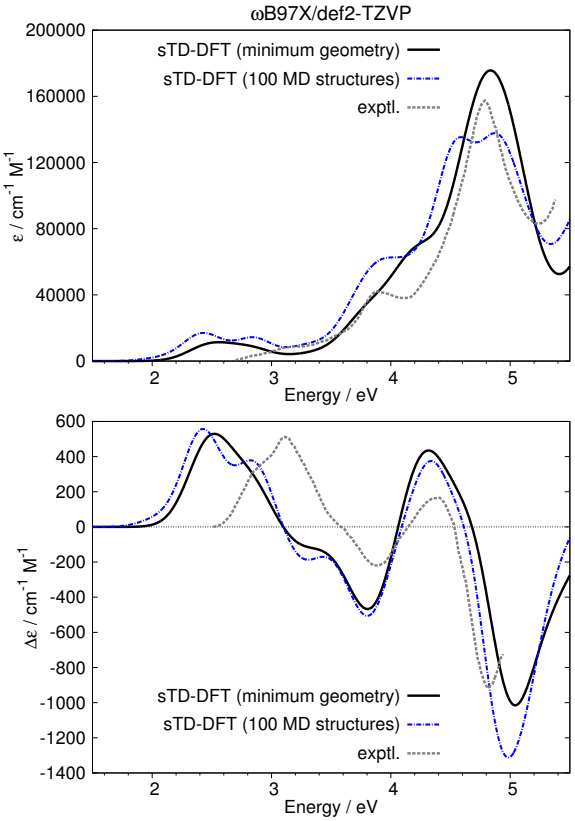


Figure 7.3.: Calculated absorption and ECD spectra of (*P*)-**[16]H-tipso** computed with sTD-ωB97X/def2-TZVP on the minimum geometry (black solid line) and 100 structures from an MD trajectory (blue dashed line). All individual transition strengths are broadened by Gaussians with a full width at $1/e$ maximum of 0.5 eV (0.35 eV in the MD case) and the spectra are red-shifted by 1 eV. The experimental curves (gray dotted line) are taken from Ref. 350. The absolute intensities of the experimental ECD spectrum are not given therein but are estimated in this work.

Given that the computed and experimental intensities in the absorption spectrum match, the absolute intensities of the ECD spectrum may be estimated using our calculations. In the computation, however, one does observe an overestimation of the absolute absorption intensities of about 30–40% and our ECD estimate needs to be corrected for this factor.[†] For this purpose, the very first, positive ECD signal at low energies is numerically integrated (as the summed area of trapezoids between the data points) for both the computed and the experimental spectrum. Then the experimental intensities are scaled by the ratio of the two integrals (corrected by the overestimation factor observed in the absorption spectrum[†]).This procedure is chosen for

[†]By numerical integration of the experimental and computed (sTD-ωB97X on the MD "snapshots") absorption spectra up to the local minimum above 5 eV, we obtained a factor to correct for overestimation. The intensities in the computed absorption spectrum are overestimated by 37%. To arrive at the final estimate for the ECD spectrum, we therefore scaled it by a factor of 0.73.

two reasons: the low-lying part of the spectrum results only from a small number of states (only two very intense ones in the minimum geometry, see Figure 7.2) and is most likely to be free from spurious states arising from errors in the calculation (e.g., from the applied density functional, missing double excitations) which are present in energetically higher lying parts of the spectrum.[129] Furthermore, the reversed intensity ordering for the two lowest positive ECD bands does not allow for a clear assignment of the individual bands. Thus, it is more appropriate to take the summed intensities in the lowermost part of the spectrum.

The resulting scaled experimental ECD spectrum is the one which is plotted as experimental reference curve throughout this work (Figures 7.2, 7.3, and 7.4). This estimate seems to be reasonable when compared to the absolute intensities of other well-studied helicenes. For example, the maximum ECD found in the lowest energy band for [6]helicene and [7]helicene is 190 and 240 $\mathrm{cm^{-1}\,M^{-1}}$,[303] respectively, in magnitude. Extrapolating by the number of benzene rings, a value of 500–550 $\mathrm{cm^{-1}\,M^{-1}}$ for the ECD maximum of (*P*)-[16]helicene can be expected, which is in agreement with our estimate for this band (513 $\mathrm{cm^{-1}\,M^{-1}}$). We have checked for a few differently sized helicenes (see Figure A5.2 in the Supporting Information) that a linear relationship between the number of benzene rings and the (computed) ECD seems is reasonable. Nevertheless, it should be stressed that we can only give an estimate for the absolute ECD intensity of an enantiopure solution.

The negative ECD signal at 4.8 eV is very large in magnitude (about 900 $\mathrm{cm^{-1}\,M^{-1}}$) for a system of this size and if a similar intensity was present also for the TIPSO devoid **[16]H**, this would – to the best of our knowledge – correspond to one of the most intense ECD signals per number of atoms (102 for the latter) ever observed. To investigate this, the ECD spectrum is computed by sTD-ωB97X/def2-TZVP on the PBEh-3c minimum structure of (*P*)-**[16]H** as well as on 100 structures from an MD trajectory which is obtained as described above. The resulting ECD spectra are given in Figure 7.4 (the absorption spectra are found in the Supporting Information).

As anticipated, the spectra of **[16]H-tipso** and **[16]H** resemble each other already for the respective minimum geometries (Figure 7.4a). Nevertheless, the spectrum of **[16]H** is slightly shifted to lower energies, in particular in the energy range of 3.4–4.2 eV. Furthermore, all computed ECD intensities are smaller by up to 10% compared to **[16]H-tipso**. Aside from an electronic substituent effect (e.g., oxygen lone pairs participating in the π-system) this can be at least partially attributed to different minimum geometries of the π-framework. Attractive London dispersion forces with the TIPSO protection groups at the far ends of **[16]H-tipso** (see Figure 7.1) lead to a slightly compressed gas phase minimum structure compared to the TIPSO devoid derivative **[16]H** (C3-C34 distance of 11.79 vs. 11.86 Å) which may cause the slight change in the excitation energies and ECD intensities. From Figure 7.4b, it becomes obvious that these effects are "washed" out on the MD trajectories and the ECD spectra for both compounds are remarkably similar.

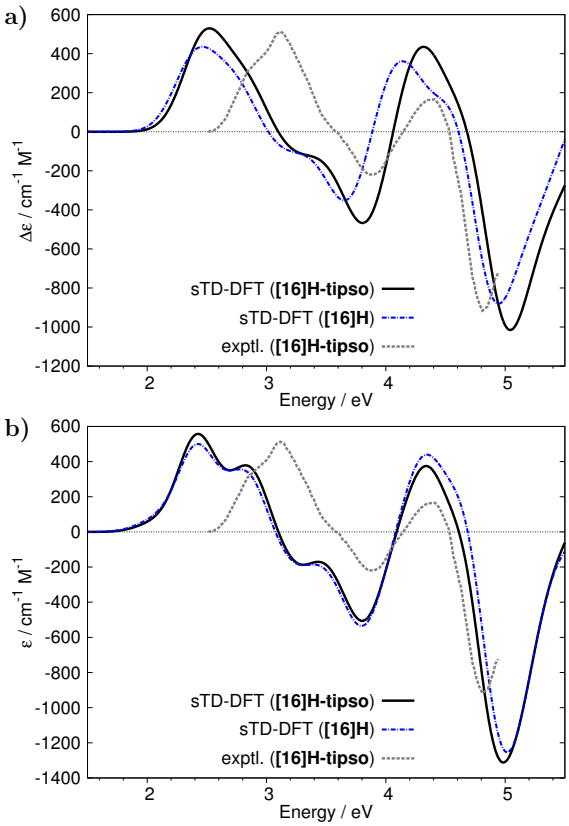


Figure 7.4.: Calculated ECD spectra of (*P*)-**[16]H-tipso** (black solid line) and (*P*)-**[16]H** (blue dashed line) computed with sTD-ωB97X/def2-TZVP on a) the minimum geometry and b) 100 structures from a MD trajectory. All rotatory strengths are broadened by Gaussians with a full width at $1/e$ maximum of 0.5 eV (0.35 eV in the MD case) and the spectra are red-shifted by 1 eV. The experimental curves (gray dotted line) for (*P*)-**[16]H-tipso** are taken from Ref. 350.

7.3. Conclusion

In this study, the ECD spectra of the largest, experimentally resolved [16]helicene and its derivative containing two TIPSO protection groups were computed. By employing efficient, state of the art quantum chemical approaches, non-equilibrium structure effects on the ECD spectrum of these moderately sized systems (102 and 164 atoms, respectively) could be studied. It was found that the ECD spectra of [16]helicene and its TIPSO derivative differed slightly at the respective minimum geometries but turned out to be almost identical when derived from the MD trajectories. As a by-product of these investigations, an estimate for the missing absolute experimental ECD intensities of **[16]H-tipso**/**[16]H** was given.

The approach presented makes use of the recently presented QMDFF[49] in the MD simulation. This is advantageous as it mimics the underlying accurate quantum chemical method (PBEh-3c[110] in our case) but energies and gradients come at a drastically lower cost compared

to an *ab initio* MD treatment. Furthermore, the sTD-DFT approach allows for a highly efficient calculation of a large number of excited states. It was shown in previous studies[46,329,362] that particularly for π-chromophores the absorption and ECD spectra computed via sTDA-DFT and sTD-DFT resemble, apart from a small overall energy shift, the spectra obtained by regular Tamm-Dancoff approximated and full TD-DFT, respectively, while being significantly faster. For a single **[16]H-tipso** geometry, the excited state treatment by sTD-DFT for about 1000 states takes 20–30 min. The bottlenecks of the overall approach presented here are the calculation of the Hessian (which is required to parametrize the QMDFF), the MD simulation itself (including solvent, equilibration etc.), and the ground state self-consistent field calculations (with ωB97X/def2-TZVP) on each of the selected MD structures. Once the MD trajectory is obtained, however, the latter step is straightforwardly parallelized. At this point we accounted for solvent effects by explicit inclusion of $CHCl_3$ molecules in the MD simulation (and hence their effects on structure) but excluded their electronic effect on the spectra which will be the topic of future work. An approach like the one presented here is in principle applicable to similarly sized and moderately more flexible systems to determine their ECD spectrum without manually sorting conformers. Such a (semi)automatic procedure will be investigated further in the near future.

Acknowledgments

This work was supported by the Fonds der Chemischen Industrie and by the DFG in the framework of the SFB 813 ("Chemistry at Spin-Centers").

Part IV.

Development and Assessment of Purpose-Specific Tight-Binding Based Approaches

Part III was devoted to the development and application of simplified excited state approaches with the focus on fast, yet accurate computation of electronic circular dichroism (ECD) spectra. The simplified TD-DFT (sTD-DFT) method was introduced in Chapter 5 and was shown to yield accurate ECD spectra, at a fraction of the computational cost required for regular TD-DFT calculations. The efficiency of the sTD-DFT scheme allowed to go beyond the single rigid structure approximation and to compute the ECD spectrum of a [16]helicene derivative (**[16]H-tipso**, with 164 atoms) on 100 structures ("snapshots") obtained from a classical molecular dynamics (MD) simulation (Chapter 7). The overall computational cost is mostly dominated by the ground state SCF calculations, each of which takes more than 15 hours, while the subsequent sTD calculation takes less than one hour.[‡]

The present part deals with the development and assessment of purpose-specific, semiempirical tight-binding (TB) methods. TB methods make use of solved electronic structures (orbitals and orbital energies) of the free atoms, which are stored as empirical input parameters. The covalent binding is then entirely described by an overlap-dependent term in the TB Hamiltonian, while electrostatic interactions are treated in the monopole approximation.

In the first chapter of this part, a purpose-specific extended tight-binding (xTB) procedure is presented, which provides orbitals for the simplified excited state calculation. With this method, the computational bottleneck of sTDA-DFT and sTD-DFT, i.e., the KS-DFT ground state calculation, is replaced by a significantly faster xTB approach. The subsequent excited state treatment, which now becomes the rate-determining step, is then conducted with the orbitals from the xTB procedure instead of the KS-DFT orbitals. To achieve an extremely fast scheme for the computation of electronic absorption and ECD spectra, the simplified Tamm-Dancoff approximation (sTDA) is used in the excited state calculation and the final method is called sTDA-xTB in the following. Already in combination with a preceding DFT calculation, (i.e., sTDA-DFT), the sTDA scheme was shown to be applicable to very large systems with more than 1000 atoms (see Chapter 4). However, as demonstrated in Chapter 6, Tamm-Dancoff approximated methods may yield unreliable rotatory strengths in the origin-independent dipole velocity formalism R^V_{TDA}, which prohibits their general use in the computation of ECD spectra. Therefore, a significant contribution to sTDA-xTB is the development of a novel scheme to correct the flawed R^V_{TDA}. The development of this so-called A+B/2 correction along with the semiempirical xTB procedure enables ultra-fast calculations (20 s in total for one **[16]H-tipso** structure) of ECD spectra with only minor loss of accuracy compared to hybrid sTD-DFT treatments. The sTDA-xTB method is presented and assessed in Chapter 8.

In Part II, the good applicability of dispersion-corrected DFT in a multi-level scheme was demonstrated for the computation of free energies in solution. However, for larger and more complex systems, conformational sampling and even MD simulations (e.g., done with a classical force-field as shown in Chapter 7) may be required. Furthermore, full geometry optimizations for systems like very large metallosupramolecular complexes, e.g., which are studied in Chapter 4

[‡]The timings refer to sTD-BHLYP/def2-TZVP calculations on 8 CPUs in ORCA (development version based on the stable version 3.0).

(with up to 1644 atoms), take several days even with a GGA functional and comparably small (double-ζ) basis sets. Therefore, a much more efficient, yet, still reliable method is needed for such studies. In Chapter 9, a second tight-binding approach (GFN-xTB) is presented, which has the specific purpose to yield reasonable geometries, harmonic frequencies, and non-covalent interaction energies in a fraction of the computation time that is typically required for low-cost DFT methods. For the aforementioned purposes, GFN-xTB can have a significant impact on multi-level procedures, including MD simulations in solutions, whereas the sTDA-xTB method presented in Chapter 8 can then be used to compute spectra on the respective MD structures.

Both TB approaches presented in this part have in common that no element pair-specific parametrization is employed. As a result, parameters for most chemical elements are available. Along with their higher accuracy for the respective target properties, the broad applicability distinguishes both approaches from existing state-of-the-art semiempirical methods which, for comparison, are considered in the respective assessments of both methods.

8. Ultra-Fast Computation of Electronic Spectra for Large Systems by Tight-Binding Based Simplified Tamm-Dancoff Approximation (sTDA-xTB)

Stefan Grimme* and Christoph Bannwarth*

Received 24th of May 2016, Published online 1st of August 2016

Reproduced (adapted) with permission from†
Grimme, S.; Bannwarth, C. *J. Chem. Phys.* **2016**, *145*, 054103.

DOI 10.1063/1.4959605

Own manuscript contribution

- Development of the A+B/2 correction
- Contributing to the development of the sTDA-xTB procedure
- Implementation of the sTDA-related features into the `stda` standalone program and interfacing to the `xtb` standalone program
- Cross-check calculations for excitation energies with sTDA-xTB, TDA-PBE0, OM2-CIS, and MSINDO-sCIS (see Table 8.5) and interpretation of the respective results
- Calculation of the specific optical rotations and interpretation of the respective results
- Calculation of the absorption and ECD spectra and interpretation of the respective results
- Writing large parts of the manuscript

*Mulliken Center for Theoretical Chemistry, Institut für Physikalische und Theoretische Chemie, Rheinische Friedrich-Wilhelms-Universität Bonn, Beringstraße 4, 53115 Bonn, Germany

Abstract The computational bottleneck of the extremely fast simplified Tamm-Dancoff approximated (sTDA) time-dependent density functional theory procedure (Grimme, S. *J. Chem. Phys.* **2013**, *138*, 244104) for the computation of electronic spectra for large systems is the determination of the ground state Kohn-Sham orbitals and eigenvalues. This limits such treatments to single structures with a few hundred atoms and hence, e.g., sampling along molecular dynamics trajectories for flexible systems or the calculation of chromophore aggregates is often not possible. The aim of this work is to solve this problem by a specifically designed semiempirical tight-binding (TB) procedure similar to the well established self-consistent-charge density functional TB scheme. The new special purpose method provides orbitals and orbital energies of hybrid density functional character for a subsequent and basically unmodified sTDA procedure. Compared to many previous semiempirical excited state methods, an advantage of the ansatz is that a general eigenvalue problem in a non-orthogonal, extended atomic orbital basis is solved and therefore correct occupied/virtual orbital energy splittings as well as Rydberg levels are obtained. A key idea for the success of the new model is that the determination of atomic charges (describing an effective electron-electron interaction) and the one-particle spectrum is decoupled and treated by two differently parametrized Hamiltonians/basis sets. The three-diagonalization-step composite procedure can routinely compute broad range electronic spectra (0–8 eV) within minutes of computation time for systems composed of 500–1000 atoms with an accuracy typical of standard time-dependent density functional theory (0.3–0.5 eV average error). An easily extendable parametrization based on coupled-cluster and density functional computed reference data for the elements H–Zn including transition metals is described. The accuracy of the method termed sTDA-xTB is first benchmarked for vertical excitation energies of open- and closed-shell systems in comparison to other semiempirical methods and applied to exemplary problems in electronic spectroscopy. As side products of the development, a robust and efficient valence electron TB method for the accurate determination of atomic charges as well as a more accurate calculation scheme of dipole rotatory strengths within the Tamm-Dancoff approximation (TDA) are proposed.

8.1. Introduction

Electronically excited molecules play important roles in many areas of chemistry and physics. Large organic molecules which absorb or emit in the visible region of the electromagnetic radiation spectrum are used as dyes or fluorescent markers in biological processes. The photochemistry of living systems, namely photosynthesis and the vision process, is currently under intense investigation and chiral electronic spectroscopy is used to elucidate protein structure. All these phenomena require at least initially a detailed consideration of the spectral properties of excited states. The ability of electronic structure theory to make reliable predictions for excitation energies (ΔE) and transition moments (spectral intensities) has advanced extraordinarily in recent years. By using time-dependent density functional theory (TD-DFT), linear-response or equation-of-motion (EOM) coupled-cluster (CC), or multi-reference perturbation theory (MR-

PT) even relatively large molecules now can be investigated routinely.[378–384]

For fairly large systems (about 100 atoms or more) for which traditional wave function based methods are usually not feasible, TD-DFT[124,125,270,271] has emerged as the "work-horse" of computational chemistry for the calculation of electronic spectra (see, e.g., Refs. 15,129,273, and 274 for reviews). However, even with TD-DFT, the computation of an ultraviolet-visible (UV-Vis) electronic spectrum in a typical energy range from 0 to 7 eV for several hundreds of atoms still remains a challenge. Even with efficient algorithms[385,386] routine calculation of absorption spectra of small proteins with about 1000 atoms is currently out of reach (see Ref. 307 for corresponding calculations of a few low-lying states). When averaging along a molecular dynamics (MD) trajectory, possibly including explicit solvent molecules, is necessary, the required computational resources grow tremendously so that such studies on larger systems in fact do no exist. This problem has been partially addressed in Ref. 46 where the simplified Tamm-Dancoff approximation (sTDA*) has been introduced. This method employs a regular Kohn-Sham ground state determinant as in standard Tamm-Dancoff approximated TD-DFT[127,128] (TDA-DFT or just TDA) but truncates the singles expansion space and uses drastic two-electron integral approximations (for details see below) leading to computational savings of more than two to three orders in magnitude. It has recently been extended to the full linear response case (sTD-DFT[329]), combined with range-separated hybrid functionals[195], and applied to various chemical problems[387–391] including electronic circular dichroism (ECD).[302,362,392–397].

The computational bottleneck of sTDA-DFT (or sTD-DFT) calculations is the solution of the self-consistent field (SCF) equations for the electronic ground state which provide the molecular orbital (MO) and orbital energy input for sTDA/sTD-DFT. This critical step is addressed in the present work by introducing a modified tight-binding (TB) Hamiltonian based semiempirical method. For the closely related time-dependent calculations of vertical excitation energies via the self-consistent charge (SCC) density functional tight-binding (DFTB) method, see Refs. 327 and 398, and a ΔSCF extension has been described very recently in Ref. 399.

The new method proposed here termed sTDA-xTB, where x stands for "extended" (or sTD-xTB when the sTD-DFT equations are solved) uses an extended atomic orbital (AO) basis set, provides higher accuracy than existing semiempirical approaches for a wider range of systems, and is computationally faster than any existing scheme with similar accuracy. Within a TB treatment, the sTDA procedure becomes the computation time determining step of the whole procedure. The necessary empirical parameters are easily determined by fitting to reasonably accurate SCS-CC2[400,401] or (TD-)DFT/PBE0[111] data so that already in this first paper the most important chemical elements H–Zn including the 3d-transition metals are covered.

The method described herein has the following fundamental characteristics:

- Extended AO basis set: in order to simultaneously and consistently describe energetically low-lying valence states in, e.g., π-systems and saturated molecules with dominantly Rydberg (valence shell expanding) states, diffuse basis functions (i.e., an sp-set for hydrogen

*If based on a Kohn-Sham DFT ground state reference, the abbreviation sTDA-DFT is preferred in this thesis.

and main group non-metals) are added to a standard minimal valence basis. In order to keep the method fast and robust (i.e., avoiding SCC convergence problems and unphysical solutions), a "one-shot" diagonalization procedure is proposed. The atomic charge dependent term in the SCC Hamiltonian is evaluated only once for given, preferably accurate input atomic charges. We previously had very good experience with the so-called CM5 charge model[402] which seems to provide the most physical atomic charges currently available[49]. They are derived approximately from another special purpose TB Hamiltonian that employs geometric Gasteiger type model charges[403] as input.

- The spectra of open-shell systems are considered from the very beginning and are treated with a standard unrestricted formalism. The unrestricted (U) input orbitals can be obtained following the corresponding U-SCC-DFTB scheme[404] which requires as additional input only atomic spin constants. They can straightforwardly be computed by spin-polarized atomic DFT calculations as occupation number derivatives of valence orbital energies.

- We follow a minimal empirical strategy regarding the number of method parameters. Pair-specific potentials are entirely avoided and only a few global as well as element specific parameters are employed. For example for the second row elements B–Ne only three atomic energy levels (2s/2p/3s3p) and two Slater atomic orbital (AO) exponents need to be determined (2s/2p, Rydberg exponents are taken from Ref. 405). The method is thus easy to parametrize and from the very beginning designed to cover to many chemical elements.

- The method should be applicable on MD "snapshots", i.e., for non-equilibrium geometries. For consistency we therefore try to approximately include the correct curvature around equilibrium (R_e) structures of excited state PES. This can only be achieved if theoretical reference data (excitation energies) for well defined molecular geometries are used to determine the empirical parameters. Beside mainly ground state R_e structures, also excited state R_e as well as bond compressed or bond elongated model geometries are employed in the parametrization process. Because of the inherent accuracy limitations of the method, the supplied reference data have to have only medium accuracy (excitation energy errors of about 0.3 eV or less). Thus, efficient excited state methods like resolution-of-the-identity (RI), spin-component-scaled (SCS), coupled-cluster second-order model (CC2)[174,380,400,401] or TD-DFT/hybrid-functional with (for Rydberg states augmented) triple-zeta quality AO basis sets can be applied. This allows the inclusion of molecules with realistic size (e.g., typical organic dyes) in the reference data set.

- It has been shown that sTDA yields the correct asymptotic electronic potential that is required to correctly describe charge-transfer (CT) states.[46,195] Because the new xTB Hamiltonian does not suffer from the well-known self-interaction error of typical semi-local density functionals[95,129,278,279], the proposed sTDA-xTB combination is asymptotically

correct for large molecule-electron distances and hence is able to yield reasonable CT as well as Rydberg transitions.

After a brief recapitulation of the sTDA/sTD-DFT approach[46,329] in sections 8.2.1 and 8.2.2, the details of the xTB Hamiltonian and its parametrization will be given in Section 9.2.1. Because the method should be applied routinely in the context of chiral spectroscopy, e.g., for ECD, the computation of coordinate origin independent transition moments is important. For this purpose we furthermore introduce here a new, accurate, and efficient so-called A+B/2 correction to TDA (or sTDA) which is described in section 8.2.4. Comparisons of the performance for excitation energies are made with four semiempirical excited state methods, namely TD-DFTB,[327] OM2-CIS,[406,407] PM6-CI,[408] as well as MSINDO-sCIS.[317,319]

8.2. Theory

8.2.1. TD-DFT and TDA-DFT

The full TD-DFT response problem is given by the following non-Hermitian eigenvalue problem[125,129]

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B^*} & \mathbf{A^*} \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} = \begin{pmatrix} \omega & 0 \\ 0 & -\omega \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} \tag{8.1}$$

where $\mathbf{A}$ and $\mathbf{B}$ are the so-called orbital rotation Hessian matrices with eigenfunctions $\mathbf{X}$ and $\mathbf{Y}$ and the corresponding eigenvalues ω. For a global hybrid density functional in the singlet, spin-restricted case (see Ref. 128 for extensions to the unrestricted formalism), the elements of the matrices $\mathbf{A}$ and $\mathbf{B}$ take the form

$$\begin{aligned} A_{ia,jb} = \delta_{ij}\delta_{ab}(\epsilon_a - \epsilon_i) + 2(ia|jb) - a_x(ij|ab) \\ + (1 - a_x)(ia|f_{XC}|jb) \end{aligned} \tag{8.2}$$

and

$$B_{ia,jb} = 2(ia|bj) - a_x(ib|aj) + (1 - a_x)(ia|f_{XC}|bj). \tag{8.3}$$

Here, the indices a and i refer to the virtual and occupied orbitals obtained from the respective ground state calculation and a_x is the amount of non-local Fock exchange that is mixed into the XC functional[69]. Two-electron integrals are given in the charge-cloud (or Mulliken) notation. The terms $(ia|jb)$ and $(ia|bj)$ in the matrix elements of $\mathbf{A}$ and $\mathbf{B}$, respectively, are of exchange type and result from the response of the Coulomb integrals in the ground state. The response of the XC functional is given by the terms $(ia|f_{XC}|jb)$ and $(ia|f_{XC}|bj)$, respectively, and its contribution is scaled by the amount of (semi-)local density functional exchange used in the ground state. Accordingly, the terms scaled by a_x correspond to the response of the non-local Fock potential. While this term is of exchange type in the $\mathbf{B}$ matrix elements, it is of Coulomb type in the $\mathbf{A}$ matrix elements.

Since the orbitals used are usually real, instead of the non-Hermitian eigenvalue problem in

Eq. 8.1, the Hermitian one

$$(\mathbf{A}-\mathbf{B})^{\frac{1}{2}}\,(\mathbf{A}+\mathbf{B})\,(\mathbf{A}-\mathbf{B})^{\frac{1}{2}}\,\mathbf{Z}=\omega^2\mathbf{Z} \tag{8.4}$$

with

$$\mathbf{Z}=\sqrt{\omega}\,(\mathbf{A}-\mathbf{B})^{-\frac{1}{2}}\,(\mathbf{X}+\mathbf{Y}) \tag{8.5}$$

can be solved.[124,125] Thus, in order to take the square-roots a unitary transformation needs to be performed which involves a diagonalization of the full configuration singles space for hybrid functionals[129,284]. To circumvent this step, Hirata and Head-Gordon have applied the Tamm-Dancoff approximation[127] to TD-DFT in which the contribution of the **B** matrix is neglected.[128] Instead of solving two eigenvalue problems as in hybrid TD-DFT, only one eigenvalue problem needs to be solved

$$\mathbf{At}=\omega_{TDA}\mathbf{t}\ . \tag{8.6}$$

Here, the solution vector **X** has been replaced by **t** and ω by ω_{TDA} to emphasize that the eigenfunctions and eigenvalues, respectively, are different to the ones in Eq. 8.1. In the case of $a_x = 1$, TDA corresponds to the configuration interaction singles (CIS) approach.[129] It has been shown that excitation energies obtained from TDA-DFT are usually only slightly larger than the respective ones obtained from TD-DFT but come at a much lower cost.[128] For a recent comparison of TDA-DFT and TD-DFT treatments for valence states, see Ref. 285, for an early, related precursor to TDA-DFT termed DFT/CIS see Ref. 286.

8.2.2. The Simplified TDA-DFT and TD-DFT Approaches

Starting from the TDA-DFT equation (Eq. 8.6), three simplifications lead to the recently published sTDA-DFT approach.[46] First, the response term of the (semi)-local density functional (last term in Eq. 8.2) is neglected to avoid expensive numerical integration. The second simplification concerns the two-electron integrals in the **A** matrix which are approximated by short-range damped Coulomb interactions of transition density monopoles. The transition charge densities q_{pq}^A are obtained from a Löwdin population analysis[287]

$$q_{pq}^A=\sum_{\mu\in A}C'_{\mu p}C'_{\mu q} \tag{8.7}$$

where the sum is over AO functions (index μ) centered on atom A (indices $pqrs$ denote molecular orbitals that can be either occupied or virtual). The matrix $\mathbf{C}'$ denotes orthogonalized MO coefficients obtained from $\mathbf{C}'=\mathbf{S}^{1/2}\mathbf{C}$ where **C** are the coefficients in the original basis and **S** is the AO overlap matrix. The two-electron integrals are then approximated by

$$(pq|rs)\approx\sum_{A}^{N}\sum_{B}^{N}q_{pq}^{A}q_{rs}^{B}\Gamma_{AB} \tag{8.8}$$

with the Mataga-Nishimoto-Ohno-Klopman[288–290] damped Coulomb operator Γ_{AB}. If Eq. 8.8 is used to replace a Coulomb type integral (superscript J), Γ_{AB} is given by

$$\Gamma^J_{AB} = \left(\frac{1}{(R_{AB})^{y_J} + (a_x \eta)^{-y_J}} \right)^{\frac{1}{y_J}} \tag{8.9}$$

where R_{AB} is the interatomic distance, y_J is a parameter, and η is the arithmetic average of the chemical hardness for the two atoms A and B, $\eta(A) = \frac{\partial^2 E(A)}{\partial N^2}$, where N is the number of electrons and E is the total atomic energy. Tabulated $\eta(A)$ values consistent for all elements of the periodic table from Ref. 291 are used. While the Fock-exchange mixing parameter a_x in standard sTDA-DFT is defined by the basic density functional used, it is a free parameter in sTDA-xTB. Interestingly, the best performance of the entire model is obtained for values around $a_x = \frac{1}{2}$ which is the theoretical value from the adiabatic connection theorem[69]. This intriguing observation is related to the fact that globally best excitation energies from TD-DFT are obtained by hybrid functionals with a large amount (40-50%) of Fock-exchange[133,283].

The respective expression of Γ_{AB} for exchange type integrals is given by

$$\Gamma^K_{AB} = \left(\frac{1}{(R_{AB})^{y_K} + \eta^{-y_K}} \right)^{\frac{1}{y_K}} \quad . \tag{8.10}$$

The values of the parameters y_J in Eq. 8.9 and y_K in Eq. 8.10 (named β and α, respectively, in the original publication) are determined here specifically for the xTB case as in Ref. 195 for range-separated functionals, i.e., each functional in sTDA/sTD is specified by these two parameters plus an "effective" Fock-exchange mixing parameter a_x.

In summary, the approximated $\mathbf{A}$ matrix (denoted as $\mathbf{A}'$) is given by

$$A'_{ia,jb} = \delta_{ij}\delta_{ab}(\epsilon_a - \epsilon_i) + \sum_{A,B}^{N_{atoms}} (2q^A_{ia}\Gamma^K_{AB}q^B_{jb} - q^A_{ij}\Gamma^J_{AB}q^B_{ab}). \tag{8.11}$$

The last simplification applied in sTDA concerns the reduction of the single excitation space. Truncation of CI spaces has been applied successfully at only minor loss of accuracy already in other approaches.[292,294,295] We refer to the original sTDA-DFT paper for details and only state here that the space is typically reduced to about 1% of the full size and that $\mathbf{A}'$ matrices of dimension of a few thousand are typically diagonalized.

These three simplifications along with the simple eigenvalue problem (Eq. 8.6) allow extremely fast computations for a fairly large number of roots as required for broad energy range spectra. Due to the inherently correct description of CT states (see Ref. 46 for details), it is also applicable to very large systems.

The modifications described above are consistently applied also in the sTD-DFT approach[329]. The matrix $\mathbf{A}$ in Eq. 8.1 is replaced by the approximate matrix $\mathbf{A}'$ from sTDA-DFT (with elements given by Eq. 8.11) and the matrix $\mathbf{B}$ is set up in a consistent, simplified manner. The configuration selection is the same as in sTDA-DFT (see Ref. 46). Just like in sTDA-DFT, the

response term of the density functional (last term in Eq. 8.3) is neglected and the first terms of the right-hand side are approximated by the exchange-type Mataga-Nishimoto-Ohno-Klopman damped Coulomb interaction (see Eq. 8.10). As mentioned above, the term originating from the non-local Fock exchange is of Coulomb type in matrix **A**, but of exchange type in matrix **B**. The same Eq. 8.10 for this integral is used but keeping the scaling factor a_x. The elements of the approximated matrix $\mathbf{B}'$ are then given by

$$B'_{ia,jb} = \sum_{A,B}^{N_{atoms}} (2q^A_{ia}\Gamma^K_{AB}q^B_{bj} - a_x q^A_{ib}\Gamma^K_{AB}q^B_{aj}). \tag{8.12}$$

Solving the Hermitian eigenvalue problem in Eq. 8.4 yields the solution vectors **X** and **Y**. The electronic transition moments from the ground to an excited state are then obtained from [47]

$$\vec{\mu}^L_{0\nu} = \sqrt{2}\sum_{ia}^{N_{CSF}} \langle\psi_i\,|\vec{r}|\,\psi_a\rangle\,(X_{ia,\nu} + Y_{ia,\nu}) \tag{8.13}$$

$$\vec{\nabla}_{0\nu} = \sqrt{2}\sum_{ia}^{N_{CSF}} \left\langle\psi_i\left|\vec{\nabla}\right|\psi_a\right\rangle (X_{ia,\nu} - Y_{ia,\nu}) \tag{8.14}$$

$$\vec{m}_{0\nu} = \sqrt{2}\sum_{ia}^{N_{CSF}} \left\langle\psi_i\left|\vec{r}\times\vec{\nabla}\right|\psi_a\right\rangle (X_{ia,\nu} - Y_{ia,\nu}) \tag{8.15}$$

in terms of orbital excitations from ψ_i to ψ_a.

Except for one xTB specific modification, which is described in the next subsection, the above sTDA and sTD methods are applied in an unaltered fashion with xTB input using our existing code.

8.2.3. Correction for Monopole Approximation Errors in sTDA/sTD-DFT

Within the monopole approximation applied in Eq. 8.8, one center exchange integrals $(ia|ia)$ (i, a are atomic or localized molecular orbitals) are incorrectly very small or even zero and hence localized, atom-like transitions of, e.g., $n \rightarrow \pi^*$ type in carbonyl compounds or $d \rightarrow d$ type in transition metal complexes are computed systematically too low in energy. The corresponding integrals are 0.5-1 eV for light atoms [409] and such errors are actually observed in the corresponding excitation energies [46]. This problem can be cured in principle by inclusion of higher-order dipole and quadrupole terms in the multipole expansion but, however, at a significantly increased cost. Instead we propose a simple correction term $\Delta_K(ia)$ which is added to the diagonal elements of Eq. 8.2 and that should correct for the missing $(ia|ia)$ term. It reads

$$\Delta_K(ia) = \frac{\Delta_{max}}{1 + [(ia|ia)/\sigma_K]^4} \tag{8.16}$$

where Δ_{max} is the maximum value of the correction and σ_K is a damping parameter which determines up to which magnitude of the exchange integral the correction is applied. The idea is that the problematic configurations are indicated by a vanishing value of $(ia|ia)$, while, e.g., delocalized $\pi \rightarrow \pi*$ excitations for which $(ia|ia)$ is large are essentially unaffected. At this point we propose to apply this correction only for sTD(A)-xTB. Whether this correction is also beneficial for DFT-based sTDA/sTD will be the subject of future work and is not considered here. In the course of the xTB parametrization process for singlet excitation energies (see below), we determined optimum values of $\Delta_{max} = 0.5$ eV and $\sigma_K = 0.1$ eV, respectively. In the xTB case the correction also compensates for errors in the Hamiltonian which appear to be larger for localized than for de-localized orbitals. If orbitals i or a belong to the manifold of open-shell (OS) orbitals (see below), an additional shift Δ_K^{OS} with two other parameters Δ_{max}^{OS} and σ_K^{OS} is used.

8.2.4. A+B/2-Approximation in TDA for Rotatory Strengths

The sTDA method uses transition dipole moments which are not approximated. Consistently, such approximations are neither applied for xTB and all magnetic and electric dipole integrals in the AO basis are computed in analogy to standard DFT. As mentioned above, the time-determining step in sTDA/sTD-DFT calculations is not the excited state treatment but the preceding SCF procedure. Due to the negligible cost of the xTB calculation the situation is different for sTD(A)-xTB where the excited state treatment dominates the computational cost. It is thus preferable to combine xTB with the computationally more efficient sTDA instead of solving the full linear response problem as in sTD. There remains, however, the problem of the lacking gauge invariance of Tamm-Dancoff approximated wave functions resulting in non-equivalent expressions of the rotatory strengths in the length (R^L) and the strictly origin independent velocity representation (R^V).[11,47] In a recent study,[362] we demonstrated the drastic failure of using R^V values within TDA approaches in the calculation of ECD spectra of inherently chiral, extended π-systems.

Based on the satisfactory performance of the origin dependent R^L expression in TDA, a correction scheme for R^V is presented in the following which fixes the problem for TDA, while retaining the origin independence. This yields an origin independent, yet well performing representation for the rotatory strength in (s)TDA. In Eqs. 8.13–8.15, the expressions for the electric position, velocity, and magnetic dipole transition moments in the (s)TD framework are given, which lead to the following expressions of the rotatory strengths:

$$R_{TD}^L = \frac{1}{2} \sum_{a=x,y,z} \langle \vec{\mu}_a^L | X+Y \rangle \langle X-Y | \vec{m}_a \rangle \tag{8.17}$$

$$R_{TD}^V = \frac{1}{2\omega} \sum_{a=x,y,z} \left\langle \vec{\nabla}_a | X-Y \right\rangle \langle X-Y | \vec{m}_a \rangle \tag{8.18}$$

For each state, the inner products in Eqs. 8.17 and 8.18 refer to the contraction of the con-

figurational transition moments (for electronic excitations $i \rightarrow a$, see Eqs. 8.13–8.15) with the respective linear combination coefficients. Different from (s)TDA where $\mathbf{X} = \mathbf{t}$ and $\mathbf{Y} = 0$, the eigenvector combinations differ for the length and velocity representation. Since the Kohn-Sham reference is the same in TDA and TD, the observed differences in the rotatory strengths are solely due to the different contraction with the eigenvectors. From Ref. 362 we gained the experience that there occur cases where R^V_{TDA} differs even in sign compared to R^V_{TD} (and R^L_{TD}) while R^L_{TDA} performs reasonable. Thus, the length expression from (s)TD is mimicked much better by the (s)TDA eigenvectors and consequently, we approximate $R^L_{TDA} \approx R^L_{TD}$ or by

$$|t\rangle \langle t| \approx |X+Y\rangle \langle X-Y| \ . \tag{8.19}$$

The right hand side (rhs) of the above equation refers to the outer product of the eigenvectors in (s)TD used for R^L (see Eq. 8.17) while the left hand side (lhs) corresponds to the (s)TDA case. Since the lhs is symmetric, the same approximate equivalency is true for a transposed rhs. It is well known[271] from TD-DFT that

$$|X-Y\rangle = \frac{(\mathbf{A}+\mathbf{B})}{\omega} |X+Y\rangle \ . \tag{8.20}$$

If applied to the rhs of Eq. 8.19, this yields the outer product used for R^V_{TD}. Due to the comparable excitation energies, one can furthermore assume that $\omega \approx \omega_{TDA}$ and transform $|t\rangle \langle t|$ with $(\mathbf{A}+\mathbf{B})/\omega_{TDA}$. This would, however, result in a non-symmetric matrix in (s)TDA, i.e., the outer product is formed of two different vectors. To maintain the strict origin independence of R^V, the outer product must yield a symmetric matrix and thus, the transformation in Eq. 8.20 is applied symmetrically.

$$\begin{aligned} &\frac{1}{2}\left(\frac{(\mathbf{A}+\mathbf{B})}{\omega_{TDA}} |t\rangle \langle t| + |t\rangle \langle t| \frac{(\mathbf{A}+\mathbf{B})^T}{\omega_{TDA}}\right) \\ \approx &\frac{1}{2}\left(\frac{(\mathbf{A}+\mathbf{B})}{\omega} |X+Y\rangle \langle X-Y| + |X-Y\rangle \langle X+Y| \frac{(\mathbf{A}+\mathbf{B})^T}{\omega}\right) \end{aligned} \tag{8.21}$$

Here, we have used the fact that the assumption in Eq. 8.19 should also hold for the transposed rhs. By factorizing out and exploiting that $|t\rangle$ is an eigenfunction of $\mathbf{A}$, we obtain

$$|t\rangle \langle t| + \frac{1}{2\omega_{TDA}}\mathbf{B} |t\rangle \langle t| + |t\rangle \langle t| \mathbf{B}^T \frac{1}{2\omega_{TDA}} \approx |X-Y\rangle \langle X-Y| \ . \tag{8.22}$$

Using the short hand notation $|t'\rangle = -\frac{1}{2\omega_{TDA}}\mathbf{B} |t\rangle$ we can rewrite this as

$$|t\rangle \langle t| - |t'\rangle \langle t| - |t\rangle \langle t'| \approx |X\rangle \langle X| - |X\rangle \langle Y| - |Y\rangle \langle X| + |Y\rangle \langle Y| \ . \tag{8.23}$$

Since the last contribution of the rhs (i.e., $|Y\rangle \langle Y|$) is comparably small, the henceforth called A+B/2 transformation yields an expression that resembles the outer product used for R^V in

(s)TD. The corrected version R'^{V}_{TDA} is the given as

$$R'^{V}_{TDA} = \frac{s(\omega_{TDA})}{2} \sum_{a=x,y,z} \left[\left\langle \vec{\nabla}_a \right| t \rangle \langle t | \vec{m}_a \rangle - \left\langle \vec{\nabla}_a \right| t' \rangle \langle t | \vec{m}_a \rangle - \left\langle \vec{\nabla}_a \right| t \rangle \langle t' | \vec{m}_a \rangle \right] \tag{8.24}$$

with

$$s(\omega_{TDA}) = \left(\frac{1}{\omega_{TDA}} \right)^{1-\exp(-150 \cdot \omega^2_{TDA})} . \tag{8.25}$$

The factor $1/\omega_{TDA}$ is replaced by $s(\omega_{TDA})$ to counteract the overestimation of ECD intensities for small excitation energies. $s(\omega_{TDA})$ is constructed such that $s(0) = 1$ (resembling the length expression of the rotatory strength), while it yields the standard factor of $1/\omega_{TDA}$ for excitation energies ω_{TDA} >3 eV. An alternative way of writing Eq. 8.24 more closely resembling Eq. 8.18 is given by

$$R'^{V}_{TDA} = \frac{s(\omega_{TDA})}{2} \sum_{a=x,y,z} \left\langle \vec{\nabla}_a \right| t - t' \rangle \langle t - t' | \vec{m}_a \rangle \text{ , with } |t'\rangle \langle t'| = 0. \tag{8.26}$$

It should be noted that the A+B/2 correction does not deteriorate the results for cases where the Tamm-Dancoff approximation is valid, since $\mathbf{B}$ and likewise $\mathbf{t}'$ are then in fact close to zero. The above described approximation is the default for the computation of ECD spectra with sTDA-xTB while for UV spectra, the uncorrected TDA vector and the dipole length formalism (Eq. 8.13 with $\mathbf{X} = \mathbf{t}$ and $\mathbf{Y} = 0$) is applied.

8.2.5. The xTB Composite Procedure

General

The basic idea of the present approach is to take orbitals and orbital energies for sTDA/sTD from a fast TB calculation which should mimic the ones from a hybrid DFT treatment in a reasonable AO basis set. This similarity requires that the general eigenvalue problem

$$\mathbf{FC} = \mathbf{SC}\boldsymbol{\epsilon} \tag{8.27}$$

has to be solved where the matrix $\mathbf{F}$ denotes the Fock (Kohn-Sham) matrix representation in the AO basis, $\mathbf{S}$ is the AO overlap matrix, and $\mathbf{C}$ are the MO coefficients with eigenvalues ϵ. We first outline the procedure for the restricted closed-shell case and extend it to the open-shell case later.

The valence MOs ψ are expanded as usual in atom centered Slater type orbitals ϕ (STO) which are approximated by Stewart's Gaussian expansions[410] with usually $m = 4$ primitives per AO (three for diffuse functions) according to

$$\psi_i = \sum_{\mu}^{N_{AO}} C_{\mu i} \phi_\mu(\zeta, \text{STO-}m\text{G}) \tag{8.28}$$

Table 8.1.: Description of the AO basis sets used. n denotes the principal quantum number of the valence shell of the element.

	treatment	
element	VTB	XTB
H-He	ns	ns, (n+1)sp
group I/II	nsp	nsp
B-Ne	nsp	nsp, (n+1)sp
Al,Ga,In,Zn,Cd,Hg	nsp	nsp
remaining group		
IV-VII non-metals	nsp, (n+1)d	nsp, (n+1)sp
d-block elements	nd, (n+1)sp	nd, (n+1)sp

where C is an MO coefficient and ζ is the STO exponent of the corresponding AO. Core electrons are neglected as usual in semiempirical theories and spherical functions are used throughout. These AOs are made partially flexible in a sense that depending on the geometric molecular structure, they are spatially contracted or expanded. The geometric D3 coordination number[133] CN of the corresponding atom in the molecule is used for this purpose and the element specific base exponent ζ^0 changes linearly according to

$$\zeta = \zeta^0(1 + k_\zeta CN) \tag{8.29}$$

where k_ζ is a global parameter. This scaling it not applied to the exponents ζ_D of diffuse AO functions.

Two different basis sets are used in the xTB composite scheme: a minimal valence (plus polarization functions on the third and higher row main group non-metals) termed VBS for the atomic charge calculation and a minimal valence plus diffuse functions on hydrogen and main group non-metals (termed XBS). The corresponding Hamiltonians which employ these sets are termed XTB and VTB in the following. The complete xTB method uses both XTB and VTB as outlined below. The type and composition of the functions is given in Table 8.1.

Note that for electronegative main group elements of third and higher rows which can form hypervalent electronic structures an additional single d-polarization function is provided within the VBS (but not for the XBS). In the XBS, diffuse basis functions are added only to non-metal main group elements for which Rydberg states are most important. These diffuse functions are atom-wise Schmidt-orthonormalized to the existing valence functions in order to minimize artificial couplings by the approximate Hamiltonian (i.e., pure atomic levels enter the XTB calculation).

The matrix elements of the XTB and VTB Hamiltonians in their respective bases are computed similar to the second-order SCC-DFTB method[411]

$$\langle\phi_\mu|F|\phi_\nu\rangle = \langle\phi_\mu|H_0|\phi_\nu\rangle + k_q\frac{1}{2}S_{\mu\nu}\sum_\kappa(\Gamma_{\alpha\kappa} + \Gamma_{\beta\kappa})q_\kappa \quad (\mu \in \alpha, \nu \in \beta) \tag{8.30}$$

where H_0 is the zeroth-order part, q is an atomic charge, k_q an empirical scaling parameter, and the inter-electronic repulsion function Γ is similar to the one used in sTDA

$$\Gamma_{\alpha\beta} = \frac{1}{R_{\alpha\beta} + \frac{2}{\eta(\alpha)+\eta(\beta)}} \,. \tag{8.31}$$

In the following, the small Greek symbols α, β, and κ index atoms while μ and ν denote AO basis functions. The atomic charges q in the above equation refer to CM5 charges[402] based on Mulliken populations from the VTB calculation. The VTB treatment uses geometry only dependent electronegativity charges as starting point and then two SCC iterations are conducted. The H_0 part is given by a sum of overlap proportional terms and scaled kinetic energy integrals according to

$$\langle\phi_\mu|H_0|\phi_\nu\rangle = \frac{1}{2}(k_\mu^l + k_\nu^{l'})\frac{1}{2}(h_\mu^l + h_\nu^{l'})S_{\mu\nu} - k_T\langle\phi_\mu|\hat{T}|\phi_\nu\rangle \tag{8.32}$$

where k_l are the usual "Hückel" constants taken here as free parameter for each angular momentum l, k_T is another global scaling parameter, and h^l are the effective atomic energy levels. For diffuse-diffuse (D-D) and diffuse-valence (D-V) AO interactions, two special parameters k_{D-D} and k_{D-V} replace the average $\frac{1}{2}(k_\mu^l + k_\nu^{l'})$ in the above equation. The h^l also depend linearly on the geometric D3 coordination number CN[133] according to

$$h_\mu^l = H_\mu^l(1 + k_{CN}CN_\alpha) \quad (\mu \in \alpha\,) \tag{8.33}$$

where k_{CN} is a global scaling parameter. This modification is applied only for the valence levels of non-metals. For all other diagonal elements, just the base atomic levels H^l are taken. For the diffuse levels, only one element parameter H_D ("D" indicating a diffuse AO) needs to be specified while the s-p level splitting is globally described by the the parameter ΔH_D, i.e., $H_D^s = H_D - \Delta H_D$ and $H_D^p = H_D + \Delta H_D$.

In order to start the partial SCC procedure, approximate non-electronic charges of Gasteiger type[403] are employed. They are based on distance weighted atomic electronegativities EN

$$N_\alpha^{EN} = Z_\alpha^{val} + \sum_\beta^{atoms} \frac{EN_\alpha - EN_\beta}{R_{\alpha\beta}^6} \tag{8.34}$$

where Z^{val} is the valence atomic charge, and R is an interatomic distance. The electron numbers N computed this way are properly normalized to yield the desired atomic charges according to

$$q_\alpha = (N_\alpha^{EN} - Z_\alpha^{val})\frac{N_{tot}}{\sum_\beta N_\beta^{EN}} \tag{8.35}$$

where N_{tot} is the total number of electrons in the system.

At this point, we can already outline the complete algorithm to obtain xTB orbitals and eigenvalues for sTDA:

1. For given input coordinates, approximate atomic charges are calculated by Eqs. 8.34 and

8.35.

2. They are used to construct an initial VTB Hamiltonian matrix using Eq. 8.30. With the resulting density matrix, Mulliken atomic charges are computed which accounts for basic and important quantum effects on the charge distribution. In order to handle orbital near-degeneracy cases automatically, the Fermi-smearing technique (fractional occupation numbers obtained at finite electronic temperature[412], T=300 K in our case) is used in the VTB scheme.

3. The Mulliken charges are taken as approximate substitute for Hirshfeld charges[413] which serve as input to compute CM5 charges[402] that are used in the next step.

4. A second VTB Hamiltonian matrix is constructed and diagonalized. The SCC procedure is stopped at this (first-order correction) point and the resulting new CM5 charges and orbitals correspond to the final VTB result. The global parameters of the VTB Hamiltonian including the atomic electronegativities are determined by minimizing the root-mean-square difference between VTB based CM5 and "true" CM5 charges (based on charges from a PBE0/TZVP[162] Hirshfeld analysis) for a large number of reference molecules.

5. The CM5(VTB) charges are used to construct the XTB Hamiltonian matrix from Eq. 8.30 which is diagonalized only once and the resulting MOs and eigenvalues are supplied to the existing (slightly modified) `stda` code. The empirical global and element-specific parameters in XTB are different from the ones in VTB and fitted to reproduce reference excitation energies.

Already at this point it is important to mention that the computed sTDA-xTB excitation energies are sensitive to the quality of the input atomic charges. We speculate that the relative large excitation energy errors of many semiempirical methods (see below) are directly related to errors in the charge distribution. This problem is alleviated in our approach because the two steps (charge generation and use of their potential) are decoupled and handled by two different special purpose Hamiltonians in two different basis sets (for previous special purpose semiempirical methods see, e.g., Refs. 414,415, and 416). Mulliken charges computed in the XBS are ill-defined due to the use of diffuse basis functions and hence they cannot be taken which furthermore motivates the proposed two-step procedure. Note that it requires only three diagonalizations (two in the smaller VBS, one in the XBS) which is much less than in normal self-consistent procedures. Furthermore, the basically non-iterative nature of the method introduces robustness as convergence failures or convergence to unphysical solutions is avoided. In fact, for hundreds of molecules tested in the development stage including radicals, biradicals, and open-shell transition metal complexes, practically no fundamentally incorrect MO occupations and derived charges were observed. Electronically difficult, multi-reference type systems are normally detected as such and indicated by significant non-integer occupation numbers in the VTB calculation and resulting small or negative excitation energies in sTDA-xTB. Hence, the method (in particular the VTB part) is recommended as a quick routine or even automatic

check for the correctness of geometry, number of electrons, or electronic state in large and complex molecules with hundreds or thousands of atoms. As will be discussed later in the results section, not only are the CM5(VTB) charges extremely accurate but also, e.g., computed Wiberg/Mayer[417] bond orders or atomic spin densities compare very well with corresponding DFT data. A disadvantage of the scheme is that very polar (ionic) situations may not be well described compared to regular SCC-DFTB because of the incomplete SCF procedure.

The HOMO-LUMO gaps from the XTB procedure somewhat resemble the gaps obtained from (semi-)local density functional calculations. In order to obtain accurate excitation energies in the XTB step, the procedure needs further modification. Because of the incomplete account of inter-electronic repulsion and correlation in a TB treatment, good absolute and relative excitation energies can only be obtained if the virtual and singly occupied orbitals are shifted relative to the doubly occupied orbital space. The virtual orbitals obtained from XTB are generally, i.e., in restricted as well as unrestricted treatments shifted by a constant $\Delta\epsilon_V$

$$\epsilon_{virt}^{xTB} = \epsilon_{virt}^{XTB} + \Delta\epsilon_V \ . \tag{8.36}$$

This yields the final xTB orbitals which then enter the sTDA treatment. In the case of open-shell (OS) systems (see below) another shift $\Delta\epsilon_{OS}$ is used for the OS orbitals

$$\epsilon_{OS}^{xTB} = \epsilon_{OS}^{XTB} \pm \Delta\epsilon_{OS} \tag{8.37}$$

where the minus sign holds for up-spin ($\uparrow$) levels (we assume $N_{tot}^{\uparrow} > N_{tot}^{\downarrow}$) which are energetically down-shifted while the opposite applies to down-spin. The correction in 8.37 is applied only to open-shell orbitals, i.e., those which would be singly occupied in a restricted OS formalism. In our unrestricted treatment, the up-spin MOs become singly occupied while they are empty in the down-spin manifold. Note that Eq. 8.36 is not applied to the latter case.

For OS systems as usual two eigenvalue problems are solved

$$\mathbf{F}_{\uparrow}\mathbf{C} = \epsilon_{\uparrow}\mathbf{S}\mathbf{C}_{\uparrow} \, , \tag{8.38}$$

$$\mathbf{F}_{\downarrow}\mathbf{C} = \epsilon_{\downarrow}\mathbf{S}\mathbf{C}_{\downarrow} \ . \tag{8.39}$$

The elements of the two Fock matrices are given by[404]

$$\langle\phi_\mu|F^{\uparrow,\downarrow}|\phi_\nu\rangle = \langle\phi_\mu|F|\phi_\nu\rangle \pm \frac{1}{2}S_{\mu\nu}\sum_{l\in\kappa}(w_{\kappa ll'} + w_{\kappa ll'})p_l \quad (\mu,\nu\in\kappa) \tag{8.40}$$

where the plus and minus signs correspond to spin-up and spin-down electrons, respectively, AO μ has angular momentum l, AO ν has angular momentum l', and p_l is a diagonal element of the AO spin-density matrix $P_\uparrow - P_\downarrow$. The last term on the right-hand side of the above equation splits only intra-atomic levels between up- and down-spin electrons. We exclude contributions

from diffuse as well as polarization functions in the sum. The atomic spin constants w are computed as described in Ref. 404 from spin-polarized, atomic PBE[104]/QZV[170] treatments as numerical derivatives of Kohn-Sham (KS) eigenvalues with respect to occupation number. These constants increase with increasing amount of Fock-exchange, e.g., by a factor of 2–3 when going from PBE to BHLYP[69]. We found best agreement between VTB computed atomic spin densities and atomic charges in comparison to DFT calculations when the PBE computed $w_{ll'}$ values are scaled by a factor k_{sc}. In the VTB calculation where two diagonalizations are conducted, the p values are initially zero which hence corresponds to a restricted open-shell treatment, i.e., the spin density is the simple sum over all OS orbitals. The corresponding p values are then used in the second diagonalization. For the unrestricted XTB part, we just take the p values from the preceding VTB treatment similar to the treatment of the atomic charges.

All global empirical parameters for sTDA-xTB as obtained from non-linear fitting to reference data (see next section) are gathered in Table 8.2. For the VTB part there are only eight free parameters (seven for closed-shell cases) and 27 parameters are required for the entire sTDA-xTB approach (23 for closed-shell cases). This intermediate degree of empiricism is typical for many semiempirical methods like PM6[408] or MSINDO[317] which rely on atomic parametrization. The element-specific parameters are listed in the supplementary material[†].

Note that no pair potential (effective nuclear-nuclear repulsion) is defined, and hence the method cannot be used to compute reasonable energy differences for different molecular geometries. For excited state treatments, it is assumed that these contributions are the same in ground and excited states and hence cancel. A qualitatively correct description of the excitation energy for structural distortions not too far away from the equilibrium situation can thus be expected which is actually observed when spectra are averaged along MD trajectories (see Section 8.3.6).

A central and critical issue of the whole approach is its speed in typical applications and we want to give a brief impression of the necessary computational effort for a few examples in Table 8.3 (timings of the corresponding (s)TD-DFT treatments for the three smallest systems are given in the supplementary material).

As can be seen, the computation of the VTB(CM5) charges is extremely fast and takes an insignificant fraction of the entire computation time. It is dominated by the sTDA part and strongly dependent on the required energy range as discussed in the original reference[46]. Note that the above timings are upper limits because some parts of the code are not well parallelized and the others are not fully optimized. In any case from the above data, it is clear that averaging spectra along MD trajectories with typically 100–300 "snapshots" is routinely possible for systems with several hundreds of atoms.

[†]The supplementary material is available online under http://dx.doi.org/10.1063/1.4959605. Except for the element-specific VTB and XTB parameters and the Cartesian coordinate data, this supporting information is also included in Appendix A6.

Table 8.2.: Global empirical parameters defining the sTDA-xTB method. Parameters are dimensionless if not noted otherwise.

parameter	VTB	XTB
k_s	1.80000	1.600
k_p	2.83500	2.180
k_d	2.08600	1.950
k_{D-V}	–	2.275
k_{D-D}	–	1.765
k_T	0.3300	0.2580
k_q	0.3910	1.0
k_ζ	0.0650	0.0343
k_{CN}	0.0030	0.0270
k_{sc}	2.6000	2.600
$\Delta\epsilon_V$	–	3.1 eV
$\Delta\epsilon_{OS}$	–	0.20 eV
ΔH_D	–	0.70 eV [a]
sTDA/sTD part		
$y_J = 4$		
$y_K = 2$		
$a_x^{sTDA} = \frac{1}{2}$		
$\sigma_K = 0.1$ eV		
$\Delta_{max} = 0.5$ eV		
$\sigma_K^{OS} = 0.5$ eV		
$\Delta_{max}^{OS} = 1.0$ eV		

[a] $\Delta H_D = 0$ and 2.25 eV for H and He, respectively.

Table 8.3.: Wall-clock timings t in seconds for the VTB/XTB parts and the entire sTDA-xTB calculation on a (single-core) laptop computer. The spectral energy range E_{max} (in eV) considered is also given.

molecule	number of atoms	number of states	E_{max} /eV	t / s VTB	XTB	sTDA
indigo	30	84	9	0.02	0.1	0.4
[16]helicene	102	628	9	0.1	2	7
poly-alanine$_{20}$	203	466	9	0.3	3	13
poly-thiophene$_{40}$	282	1902	9	1.9	5	127
poly-methylglu$_{20}$	383	683	9	1.9	12	69
protein-ligand complex[a]	624	1022	7	5.5	36	232
$(H_2O)_{265}$	795	1222	9	4.6	49	547
DNA fragment	1052	298	7	37	156	568

[a] 6 Å cut-out from PDB structure (code 2W26) with coagulation enzyme Factor Xa. [418]

Parametrization and Technical Details

The parametrization involves determination of the global parameters in Table 8.2 and the atomic parameters H_s, H_p, H_d, ζ_s^0, ζ_p^0, ζ_d^0, ζ_D, H_D, and EN. The actual number depends on the element and type of basis set (VBS or XBS) and varies between four and six. In order to keep the total number of unknowns minimal, the values of the p-exponents were also used for the d-polarization functions. All atomic parameters of the transition metal blocks were obtained by linear interpolation with nuclear charge Z where only the start ($Z = 21$, Sc) and end points ($Z = 29$, Cu) are freely fitted which reduces the number of unknowns from 54 to 12. For the group IIb elements Zn-Hg the d-electrons were not treated as valence (sp-shell only).

Generally, a minimization of the root-mean-square deviation (RMSD) between the calculated and reference data using the Levenberg-Marquardt (LM) algorithm[419,420] was conducted. All molecular structures used were optimized at the PBEh-3c level[110] which yields very accurate equilibrium (R_e) structures for a wide range of systems and which is the geometry optimization default level in our group. The following procedure was applied:

1. Global and atomic parameters were simultaneously determined for the elements H, C, N, O. The other atomic parameters were optimized element-wise while keeping all other already existing parameters fixed. Typically, 5–10 (50–100 for VTB) reference data points are used per parameter.

2. All initial fits were conducted for closed-shell molecules and the four additional parameters for open-shell cases were determined subsequently without re-adjusting already fitted values. A special set of open-shell main-group molecules was used for this purpose.

3. The VTB part was determined first by fitting on PBE0/TZVP computed CM5 charges.

4. The XTB part was subsequently determined by fitting mainly on SCS-CC2/aug-cc-pVTZ computed vertical excitation energies. For some larger molecules, the diffuse f (d on H) functions were discarded in the SCS-CC2 calculations. For some open-shell or metal containing systems, TD-DFT(PBE0)/def2-TZVP[163] data were used as reference.

5. Only lowest (or second-lowest), clearly assignable electronic states with definite correspondence between reference and sTDA-xTB calculation were taken. Computed dipole oscillator strengths were partially used for assignment.

6. Parameters with small influence on the RMSD were fixed at some point of the entire procedure and their numerical value rounded to a significant figure.

The quality of the reference data is crucial for any empirical scheme. The SCS-CC2 or TD-DFT-PBE0 excitation energies are typically accurate to 0.2–0.3 eV for closed-shell systems[401,421,422] (about 0.3–0.5 eV for open-shell cases[423,424]). This is sufficient for sTDA-xTB which inherently cannot be more accurate due to the various approximations employed. Note that we fit to ΔE values for the same ground state geometry in the gas phase. Thus, if the

reference data have no systematic deviation and the final mean signed deviation (MSD) is small (which is the case), the excitation energies from sTDA-xTB should be correct (on average) for a given gas phase structure. Solvent as well as geometrical relaxation effects are *not* included implicitly and in comparison with experimental spectra where those effects are present, one can expect blue-shifted theoretical bands when gas phase vertical transition energies are used in the comparison. These typical shifts are 0.2–0.4 eV. However, in MD runs and when including explicit solvent molecules in the treatment, more or less the absolutely correct band positions should be obtained. Note that systematic variations in covalent bond lengths on the order of 0.01–0.02 Å which are typical when different quantum chemical methods are used (e.g., DFT vs. HF-3c[169] or DFTB[411] optimized structures) can lead to changes in sTDA-xTB excitation energies of about 0.1–0.2 eV.

The VTB fit to the CM5 charges is unproblematic because no state assignment has to be made and an order of magnitude of more reference data points is available. The human working time required to parametrize a new element is about 1–2 working days.

We used the TURBOMOLE suite of programs[164c,165,371] (version 7.0) to perform most of the ground state DFT calculations. For the calculation of the CM5 charges with PBE0 the ORCA code[48] was used. All of the sTDA or sTD excited state calculations presented here were carried out with a stand-alone program which is available on our website[306]. We employ standard exchange-correlation functional integration grids (*m4* if not noted otherwise), typical SCF convergence criteria ($10^{-7}\,E_h$), and the resolution of the identity (RI) integral approximation[76,167,168] in the DFT calculations. Structures were fully optimized at the PBEh-3c composite level if not mentioned otherwise.[110]

Comparative calculations were conducted with the MOPAC12[425] (PM6-CI[408]), ADF[426,427] (TD-DFTB[327,428]), MSINDO[317] (MSINDO-sCIS[319]), and MNDO99[407] (OM2-CIS[406]) codes. In the PM6-CI treatments, a full CI within an active orbital window of eight around the Fermi level was used while the CIS treatments as well as the TD-DFTB procedure are expanded in the full orbital space.

All excitation energies refer to vertical singlet states for a given (usually ground state equilibrium) geometry. The default configuration selection threshold of $t_p = 10^{-4}\,E_h$ is used in sTDA/sTD.[46] The E_{max} value which specifies the desired spectral range was set to 10 eV if not mentioned otherwise. The active occupied and virtual MOs are automatically selected according to the chosen E_{max} value and their number is typically only about 10-20% of the full MO space. Conventional TDA-DFT or TD-DFT calculations were performed using the `escf` module[301] of TURBOMOLE. For the SCS-CC2 calculations, the efficient `ricc2` module from TURBOMOLE was used[400] utilizing matching auxiliary basis sets in the RI treatment[429] of the CC2 equations but not using any RI in the SCF step.

As noted above, the one-electron transition moments are calculated without any further approximation from the exact electronic moment integrals in the Cartesian AO basis. In order to simulate the vibrationally broadened experimental UV or ECD bands, we summed oscillator or rotatory strengths weighted Gaussian curves with a full width at $1/e$-height of $\sigma = 0.4\,\text{eV}$

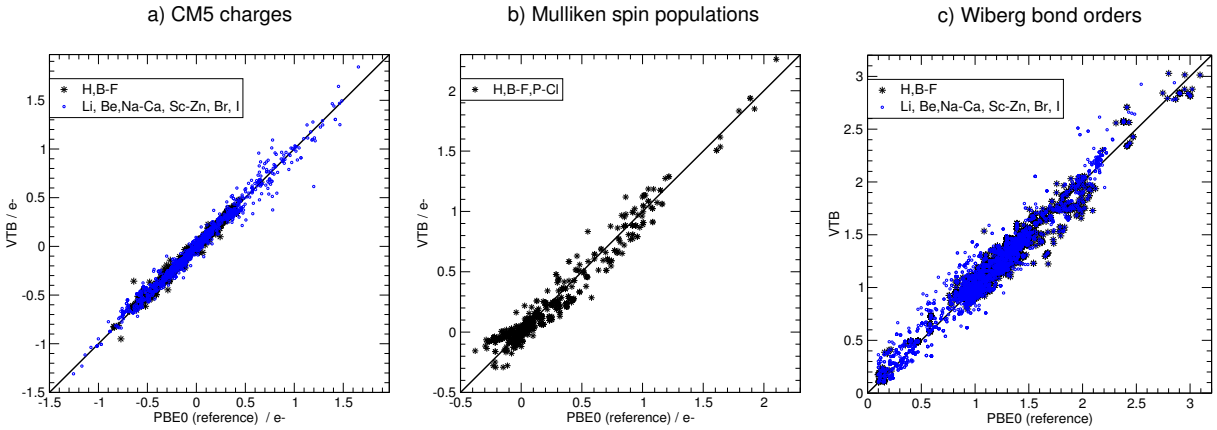


Figure 8.1.: Comparison of VTB and PBE0/TZVP data for a) CM5 atomic charges (2863 and 2585 comparisons, respectively), Mulliken atomic spin populations for open-shell systems (952 comparisons), and c) Wiberg bond orders (4761 and 2590 comparisons, respectively). For bond orders, only values >0.1 are compared. The solid line denotes a one to one correspondence of the two data sets.

for each calculated electronic transition. In MD runs, this line-width is reduced to 0.1 eV. We employ the dipole length formalism for oscillator strengths for UV absorption spectra and the origin-independent A+B/2 corrected dipole velocity formalism for rotatory strengths in ECD spectra.

8.3. Results and Discussion

8.3.1. Charges, Spin Populations, and Bond Orders from the VTB Hamiltonian

The fitting sets of molecules comprise 203 closed-shell systems for H,B-F (termed "organic" elements), 179 for Li,Be,Na-Cl,K,Ca,Zn,Br,I, and 86 molecules for the 3d-block elements Sc–Cu. In total 76 open-shell molecules (with elements H,B-F,Si-S and one- to three unpaired electrons) were used in the determination of the open-shell related parameters. In Figure 8.1 we show the results for these fit sets regarding properties which are computed with VTB, i.e., CM5 charges, spin populations (SP), and Wiberg bond orders (WBO)[417]. Note that only the CM5 charges were used in the fitting process as reference data. Therefore, comparison of SP and WBO values from VTB with the reference (PBE0/TZVP always) corresponds to a detailed cross-checking of the empirical model.

The reference charges are very well reproduced by the VTB approach and this holds for the electronically simpler organic as well as for the third-row elements and metals for which only slightly worse results are found. Because the mean deviations in all cases are close to zero (i.e., no systematic deviations are found), we only discuss mean absolute deviation (MAD) and standard deviation (SD) as statistical performance measure in units of electron charge. For the organic CM5 charges we observe tiny deviations (MAD=0.01, SD=0.02) with a maximum

absolute deviation (MAXD) of only 0.28. The corresponding values for the rest of the elements (MAD=0.03, SD=0.05, MAXD=0.58) are somewhat larger but one has to consider that the charges are also larger on average for these typically more polar molecules. In the course of the method development, we observed that the second VTB diagonalization significantly improves in particular polar and charged molecules while further SCC steps sometimes lead to divergence of the procedure or incorrect charge distributions. The present scheme is in our opinion the best compromise between accuracy and robustness at a minimum of computational effort.

The open-shell molecules have similarly small deviations for the charges (MAD=0.02, SD=0.03) and in addition, the spin populations are described almost equally well (MAD=0.05, SD=0.07). This indicates a reasonable parametrization of the unrestricted part of the Hamiltonian (see supplementary material for a correlation plot of $\langle S^2 \rangle$ expectation values). The WBO values, which indicate the quality of the covalent bonding patterns, are typically accurate to about 0.1 or 5–10% (organic: MAD=0.05, SD=0.07, MAXD=0.46, other elements: MAD=0.08, SD=0.11, MAXD=0.55) which is also acceptable. Overall, the simple VTB Hamiltonian can provide rather accurate charges and bond orders robustly for chemically diverse systems including general main group and transition metal complexes. This is intriguing because TB Hamiltonians with an incomplete account of inter-electronic repulsion and no "true" SCF are expected to be problematic for polar/metallic cases. Although there are indeed a few outliers, they are not large and relatively rare so that we can recommend the VTB method as a general tool in quantum chemistry for the extremely fast generation of atomic charges and bonding information. By construction it can handle any system including charged and open-shell species. Furthermore, for the several hundreds of molecules tested so far, we basically never observed an incorrect electronic state (wrongly occupied orbitals) which qualifies its use for SCF start orbital generation. Another application are automatic tests for correctness of electronic (charge) state and geometry of complex molecules as they are clearly indicated by large occupied/virtual orbital energy gaps and close to integer occupation numbers. This will be investigated in future studies.

Because other semiempirical methods are not constructed to yield CM5 charges, a direct method comparison is not possible. A few comparisons of Mulliken type PM6 charges with CM5(PBE0) ones, however, indicate much larger systematic as well as randomly distributed deviations compared to VTB.

In the original paper presenting the D3 dispersion correction,[133] a hypothetical molecule was used to illustrate the concept of coordination numbers. In Figure 8.2, the same molecule is used for a direct comparison of the CM5 charges derived from either VTB(Mulliken) or PBE0/TZVP(Hirshfeld). For most atoms in this example, the differences compared to PBE0/TZVP are less than 0.1 electrons and only for the central iron atom a larger difference of 0.26 electrons is obtained. However, even for such a model system with a variety of elements and a relatively complicated electronic structure no real outliers (i.e., incorrect sign) are found.

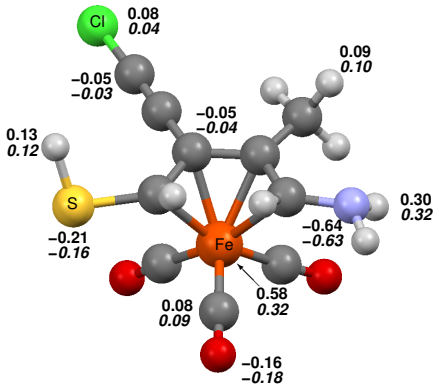


Figure 8.2.: Comparison of VTB (bottom, in italics) and PBE0/TZVP (top) derived CM5 atomic charges for the hypothetical molecule from Ref. 133.

8.3.2. Excitation Energies for the Fit Sets

In this section we first want to discuss the results for the fit sets using statistical measures. In a second step we evaluate sTDA-xTB in comparison to other semiempirical methods for molecules that were not considered in the fit (i.e., cross-checking, see Section 8.3.4). The molecules used in the element wise parametrization of the XTB part are basically the same for XTB and VTB except that in the former case roughly 10–20 % had to be removed because of unclear excited state assignment or large deviations between SCS-CC2 and TD-DFT results which can be taken as a reliability criterion of the reference value. The final data are shown graphically in Figure 8.3 for organic and the other elements as well as for the open-shell set. As noted before, the accuracy of the reference excitation energies is about 0.2–0.3 eV which limits the ultimately achievable accuracy MAD/SD values to this range.

Overall, the sTDA-xTB method performs excellently with MAD and SD values of only 0.34 and 0.44 eV, respectively, for the organic set. For the other elements (see part b) in Figure 8.3) the deviations are somewhat larger because of the electronically more complicated structures (MAD=0.40 eV, SD=0.52 eV) but still we consider these results as reasonably good and useful for many applications. This also holds for the open-shell cases for which unrestricted sTDA-xTB performs well (MAD=0.48 eV, SD=0.59 eV). The MAXD value in all sets is about 1.0–1.9 eV. One should note, however, that the method covers a wide energy range from 0 to 11 eV. For an average excitation energy of about 5 eV the typical deviation of 0.3–0.5 eV corresponds to less than 10% error. As will be shown below, the errors of sTDA-xTB for low energy transitions (<5 eV) in organic chromophores and dyes are often only about 0.2–0.3 eV which is similar to that of TD-DFT with hybrid functionals. We think that this is really encouraging considering the inherent speed of the method which is at least 10^5 times faster in the orbital generating part compared to a medium basis set, hybrid DFT calculation.

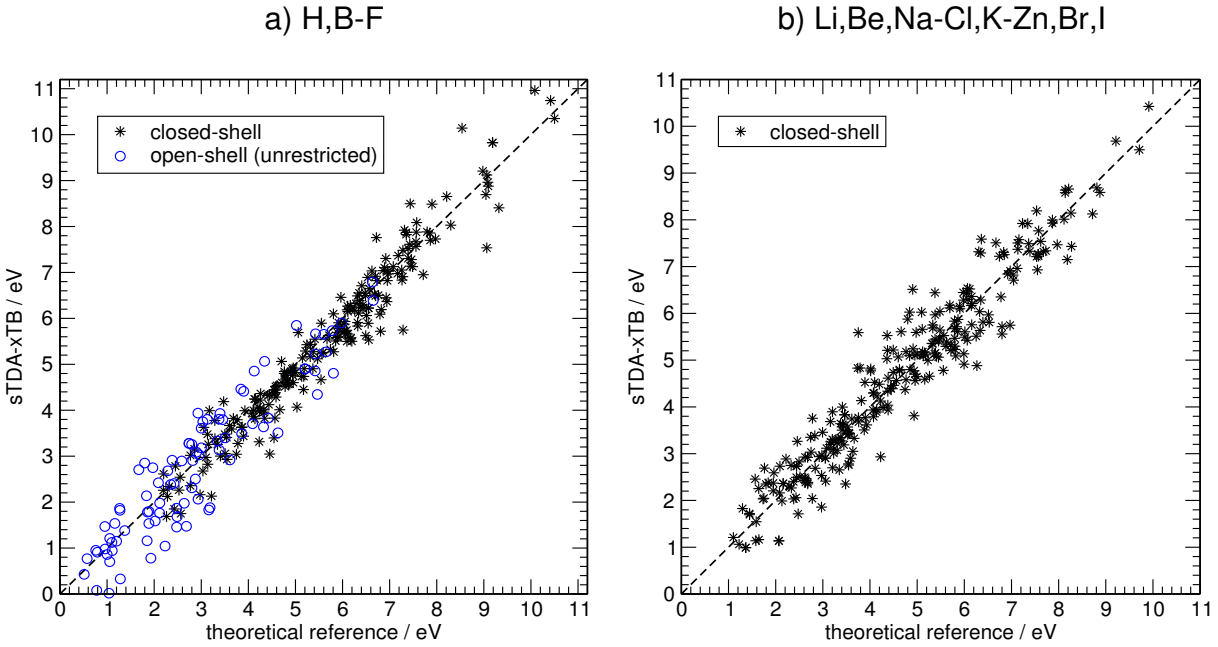


Figure 8.3.: Comparison of sTDA-xTB and reference excitation energies for organic (a) and other elements (b) and for closed- as well as open-shell molecules. The data correspond to the fit sets and the solid line denotes a one to one correspondence of the two data sets.

8.3.3. Rydberg States

One of the special features of the xTB Hamiltonian is the inclusion of the small exponent basis functions which allows consistent treatment of diffuse high energy states. They are often characterized as Rydberg, i.e., atom-like valence shell-extending states. They dominate the absorption spectra of small molecules as well as medium sized systems if no clear chromophoric groups (e.g., π or carbonyl groups) are present [430]. In Table 8.4 we compare sTDA-xTB results for a few typical cases with corresponding SCS-CC2/aug-cc-pVTZ results which have an error of less than 0.2 eV for such states.

As can be seen from Table 8.4, the agreement with the reference excitation energies is excellent. Furthermore, the relatively small oscillator strengths which are typical for Rydberg transitions are well reproduced by the semiempirical method. This is very encouraging and supports the use of the specially orthogonalized diffuse AO basis functions. Note, however, that diffuse d as well as higher principal quantum number describing functions are missing in xTB, and hence, the results (in particular for smaller molecules) are expected to deteriorate with increasing excitation energies. The only notable difference for the oscillator strengths is observed for methane (TD-DFT calculations also yield f values around 0.1).

Table 8.4.: Comparison of sTDA-xTB and reference vertical singlet-singlet excitation energies for a few states with dominant Rydberg character and corresponding oscillator strengths f (dipole lengths form).

	SCS-CC2/aTZ		sTDA-xTB	
state	ΔE / eV	f	ΔE / eV	f
		H_2O		
n-3s	7.39	0.057	7.52	0.042
n-3p	9.08	0	9.06	0
		N_2H_4		
n_1-3s	6.10	0.022	6.19	0.016
n_2-3s	6.16	0.024	6.28	0.026
		CF_4		
n-3s	12.59	0	12.25	0
		C_2F_4		
π-3s	7.53	0	7.29	0
π-3s	8.24	0.033	7.30	0.007
		acetone		
n-3s	6.41	0.031	6.71	0.019
		trans-1,2-dimethyloxirane		
n-3s	7.08	0.012	7.12	0.009
n-3p	7.35	0.014	7.87	0.041
		CH_4		
σ-3s[a]	10.73	0.14	10.35	0.020
		propane		
σ-3s	9.04	0.0006	9.09	0.0013

[a] Only one component of the T state given for f.

8.3.4. Cross-Checking of Excitation Energies and Comparison to TD-DFT and Other Semi-Empirical Methods

Here we consider a more or less randomly selected set of closed-shell molecules for cross-checking purposes of the new method (see supplementary material for the respective chemical structures and Cartesian coordinates[†]) The excited states considered always correspond to the lowest excited singlet state in the SCS-CC2 reference calculation, and in unclear cases, state assignments based on computed oscillator strengths are made.

In the cross-check, sTDA-xTB performs similarly well as for the fitting set with small MSD, MAD, and SD values (see Table 8.5). For this set which includes one charge-transfer complex and some larger dyes, the semiempirical method actually performs better than TD-DFT with a semi-local functional. At this TD-PBE level, the excitation energies are systematically underestimated (MSD=-0.66 eV) but also the scatter of the data is somewhat larger than with sTDA-xTB (SD of 0.44 vs. 0.41 eV). TD-DFTB which represents an approximation to TD-PBE does not exhibit the systematic underestimation of the excitation energies (MSD=0.09 eV), and hence yields a better MAD of 0.42 which is, however, higher than that of sTDA-xTB (MAD=0.27 eV). Because of the considerable number of larger deviations for TD-DFTB, the SD value is worse compared to TD-PBE or sTDA-xTB. For a few cases, we also tested another more organic molecule specific

DFTB parametrization[431]. This leads to somewhat different values but overall not to better statistical data.

Table 8.5.: Comparison of vertical singlet-singlet excitation energies (in eV) for states with dominant valence character. SCS-CC2 values (oscillator strengths f for assignment of transitions in parentheses) are taken as reference and all ground state structures were optimized at the PBEh-3c level.

entry	molecule	state	SCS-CC2[a]	sTDA-xTB	TD-PBE[b]	TDA-PBE0[c]	TD-DFTB	OM2-CIS[d]	PM6-CI	MSINDO-sCIS
1	$MePC_2H_4$	1A"	6.53 (0.0399)	6.56	6.28	7.11	6.23	–	5.27	5.92
2	P_2H_4	2A	6.25 (0.0498)	5.91	5.71	6.53	5.66	–	6.19	6.09
3	CF_3COOH	1A"	5.95 (0.0003)	5.95	5.45	5.80	7.12	4.53	4.83	4.80
4	Si_4H_8	A_2	5.22 (0.0000)	6.65	4.63	5.09	4.26	–	2.86	6.19
5	saccharin	2A'	4.91 (0.0045)	4.86	4.33	5.18	4.21	–	4.03	3.73
6	hexatriyne	$1\Sigma_u^-$	4.85 (0.0000)	5.04	3.76	3.93	5.89	4.18	3.71	4.78
7	ASS	2A	4.80 (0.0219)	4.62	3.98	4.93	4.56	4.68	3.90	4.93
8	purine	1A"	4.69 (0.0029)	4.49	3.60	4.32	5.20	4.19	3.31	3.85
9	dithiacyclohexane	1B	4.52 (0.0088)	4.50	3.99	4.50	4.17	–	1.61	4.09
10	fluorisochinoline	2A'	4.50 (0.0282)	4.44	4.07	4.68	5.23	4.61	3.46	4.61
11	bisthiophene	$1B_u$	4.48 (0.3949)	4.21	3.81	4.49	4.41	–	3.27	4.37
12	c-propenone	$1B_2$	4.42 (0.0001)	3.89	3.40	4.25	4.47	4.41	2.57	4.04
13	terpyridine	$1A_2$	4.23 (0.0000)	4.69	2.84	3.86	4.41	4.45	3.62	4.13
14	silabenzene	$1B_1$	4.23 (0.0595)	4.64	4.39	4.80	4.90	–	2.67	4.68
15	$B(C_6F_5)_3$	$1A_2$	4.10 (0.0224)	4.30	2.78	3.69	3.00	–	4.25	4.76
16	corannulene	$1E_2$	3.86 (0.0000)	3.56	3.25	3.77	4.21	4.09	3.40	3.67
17	S0904 dye	2A'	3.81 (0.7251)	3.72	2.97	3.70	3.99	3.87	3.52	4.10
18	S2127 dye	2A	3.66 (0.9637)	3.49	2.62	3.41	3.50	3.78	3.24	3.99
19	acenaphthene	$1B_1$	3.65 (0.0002)	2.74	2.89	3.34	3.66	4.91	2.84	4.59
20	proflavine	$1B_1$	3.54 (0.2342)	3.30	3.01	3.61	3.85	3.85	2.93	3.85
21	S2408 dye	2A	3.64 (0.8379)	3.52	3.18	3.90	4.21	3.63	3.08	4.15
22	HCSOH	1A"	3.57 (0.0000)	3.49	3.34	3.59	3.56	–	1.52	3.77
23	S2153 dye	2A	3.48 (1.0045)	3.36	2.86	3.54	3.75	–	3.04	3.91
24	CT5 system	2A	3.33 (0.0170)	2.91	1.58	2.19	3.53	3.98	3.19	5.03
25	S0491 dye	1B	2.70 (1.5478)	2.75	2.56	3.09	3.25	2.69	2.69	3.18
26	S2084 dye	$1B_u$	2.37 (1.4233)	2.60	2.42	2.98	2.33	2.11	2.22	1.94
27	thioindigo	$1B_u$	1.84 (0.3255)	1.92	1.64	2.29	1.96	–	1.78	2.71
	MSD		–	−0.04	−0.66	−0.02	0.09	0.00	−0.89	0.10
	MAD		–	0.27	0.67	0.31	0.42	0.37	0.90	0.52
	SD		–	0.41	0.44	0.42	0.55	0.58	0.76	0.66

[a] aug-cc-VXZ AO basis (X=T, X=D for the larger molecules). [b] def2-TZVP AO basis. [c] def2-SV(P) AO basis. [d] Incomplete data set due to missing parameters for B, Si, P, and S.

Furthermore encouraging is the fact that on average sTDA-xTB performs remarkably similar to Tamm-Dancoff approximated TD-DFT based on a PBE0/def2-SV(P) reference. The standard ground state oriented PM6 method performs badly for excitation energies showing systematically too small ΔE values together with a large error spread (SD=0.76 eV). This behavior is typical for standard zero-differential-overlap (ZDO) methods like MNDO, AM1, or PM6 and well documented in the literature[432]. In fact it is clear from the above data that overlap based or corrected methods (xTB, DFT, DFTB, OM2) perform better than the ZDO schemes like PM6 (or MSINDO which partially employs orthogonalization corrections). Due to missing parameters for the third row elements, OM2-CIS can only be applied to a subset of the molecules. On this subset, the method performs quite well with no systematic deviation and an SD of

0.58 eV. The MSINDO-sCIS approach[319] is a modified and partially re-parametrized version of MSINDO,[317] designed to calculate excited states. While the method shows on average only a slight overestimation tendency (MSD=0.1 eV), the error spread is the second largest of the methods considered (SD=0.66 eV). Note that of all the methods tested sTDA-xTB yields the best agreement with the reference. Even on this test set (which was not used for fitting) it shows almost no systematic underestimation tendency and the smallest error spread of all methods.

As any semiempirical method, sTDA-xTB is, however, not free of outliers. Known cases are described here and some general comments are given. Systematic deviations are observed for anions with relatively localized electronic structure, e.g., OH^-, acetate, or phenolate for which the sTDA-xTB ΔE values are too large by 1–2 eV compared to SCS-CC2. This is not unexpected because the further simplified SCC scheme using just atomic charges as well as the monopole approximation in sTDA should be more problematic for highly charged systems. In zwitter-ionic amino acids, however, we do not observe such errors. Not unexpectedly, small molecules with a somewhat complicated electronic structure seem to be problematic, notably Me-NO, SF_6, or O_3. The first allowed $\sigma \rightarrow \sigma^*$ excited state of H_2 also has a large error (ΔE =9.5 eV instead of 12.8 eV with SCS-CC2) but this is attributed to an inaccurate valence-Rydberg mixing. A related problem is observed for N_2 (error of -2 eV), but not for CO. In general, fluorine containing molecules show somewhat larger errors than expected from the MAD/SD but except for FHF^-, no "true" outliers (errors > 2 eV) have been observed so far. Empirically it seems that the method overall performs very well for the large target systems while small molecules are described less accurately.

8.3.5. Test of the A+B/2 Vector Approximation for ECD and Optical Rotations

Before ECD spectra are to be discussed, the quality of the A+B/2 correction scheme described in Section 8.2.4 is first examined in combination with Kohn-Sham DFT orbitals in order to remove effects from the semiempirical approximations in xTB. For this purpose, the ECD spectrum of C_{76} in the velocity formalism is computed from sTD-DFT, regular and the A+B/2 corrected sTDA-DFT (Figure 8.4a). Furthermore, specific optical rotations (ORs) based on a sum-over-states treatment and a PBE0/TZVP reference are computed and given in Figure 8.4b.

In Ref. 362, C_{76} served as a prototypical example where Tamm-Dancoff approximated methods in the velocity representation fail dramatically in the computation of ECD spectra. Thus, this system is used here to assess the quality of the newly proposed A+B/2 correction. The BHLYP/def2-SVP[69,163,300] Kohn-Sham ground state determinant (based on a TPSS-D3(BJ)/def2-TZVP[105,133,134,162,163] geometry) already used in Ref. 362 is used as reference. To be comparable with Refs. 362 and 325, the computed spectra in Figure 8.4 are blue-shifted by 0.9 eV. Compared to sTD-DFT, the ECD spectrum from sTDA-DFT is of opposite sign over the entire energy range considered, thus suggesting an incorrect assignment of the absolute configuration. Employing the A+B/2 correction yields an origin-independent, sTDA-DFT derived ECD spectrum showing the overall correct sign in close resemblance to the sTD-DFT one. The former calculation is, however, faster by a factor of about 10 which becomes very significant when xTB is

used instead of a Kohn-Sham DFT reference. As a cross-check for a system where non-corrected sTDA-DFT already performs well, the ECD spectrum of an α-helical oligopeptide is given in the supplementary material.

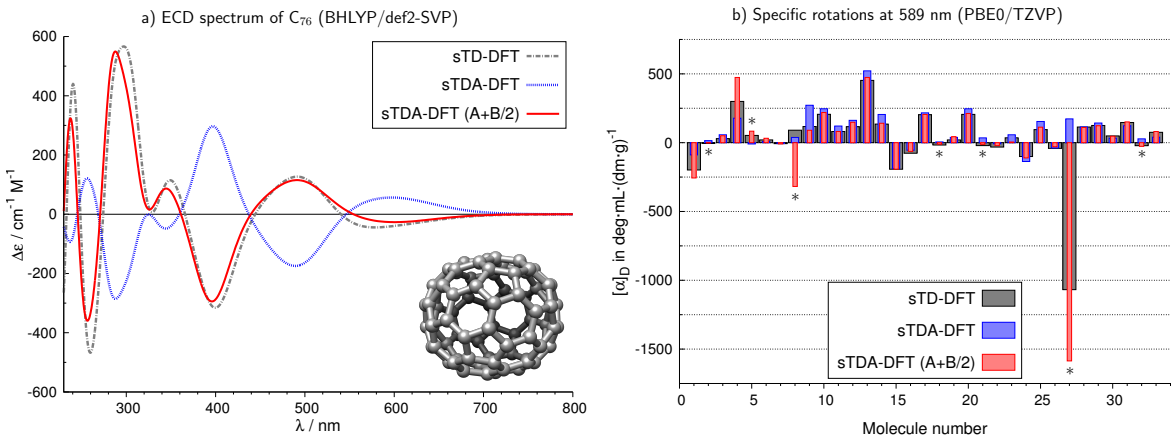


Figure 8.4.: a) Computed ECD spectra of (^{s,f}A)-C_{76} based on a BHLYP/def2-SVP reference. The sTD-DFT (gray dashed) and sTDA-DFT (blue dotted) spectra are identical to the ones in Ref. 362. The A+B/2 corrected sTDA-DFT spectrum is given as red solid line. The origin independent velocity representation is used throughout and all spectra are shifted by 0.9 eV. b) Specific optical rotations at 589 nm (excluding the Lorentz pre-factor) of 33 molecules computed via the sum-over-states expression for the isotropic optical rotation. The sTD(A) calculations are based on a PBE0/TZVP reference. The specific rotations from sTD-DFT are given as gray bars, while the ones from non-corrected sTDA-DFT are given in blue. Specific rotations computed from A+B/2 corrected sTDA-DFT are given in red. The origin independent velocity representation is used throughout and all excitations up to 60 eV were considered. Changes in sign due to the A+B/2 correction are indicated by asterisks.

Based on a PBE0/TZVP reference, the OR at 589 nm is computed for 33 molecules using the sum-over-states expression and by considering all excitations up to 60 eV (Figure 8.4b, see supplementary material for the molecules considered). The Lorentz factor $(n^2+2)/3$ is excluded in these values. Here sTD-DFT always yields the same sign of the OR as regular TD-DFT and only differences in the absolute values are observed (see supplementary material). To investigate the effect of the Tamm-Dancoff approximation and the new correction scheme only, we therefore compare the latter to specific rotations from sTD-DFT. It is seen that for the selected systems, sTDA-DFT yields the correct sign for most systems as well as reasonable magnitudes compared to sTD-DFT. However, for six cases, the Tamm-Dancoff approximation results in a sign change and thus to qualitative differences with (s)TD-DFT. A+B/2 corrected sTDA-DFT yields the correct sign in these six cases. Furthermore, the otherwise good performance of sTDA-DFT for the other systems is retained, except for the case of *E*-cyclooctene (system 8 in Figure 8.4b). As

the ECD spectrum of *E*-cyclooctene reveals (see supplementary material), the agreement with sTD-DFT is actually much better for A+B/2 corrected sTDA-DFT and the only significant difference is the slightly overestimated excitation energy of the lowest transition which results from the Tamm-Dancoff approximation. Non-corrected sTDA-DFT seems to yield the correct (positive) sign in this case, simply due to error cancellation (i.e., underestimated magnitudes for the rotatory strengths between 7 and 9 eV).

While the original eigenvectors from the Tamm-Dancoff approximated eigenvalue problem serve as a good approximation to linear-response R^L intensities (see section 8.2.4), they require an appropriate correction to compute reasonable R^V. The A+B/2 correction to the (s)TDA (or CIS) eigenvectors is such an approximation and it performs very well at a drastically lower cost compared to solving the full (s)TD problem as demonstrated in this section.

8.3.6. UV and ECD Spectra in Comparison to Experiment

In this section direct comparisons of sTDA-xTB computed and experimental UV and ECD spectra are given to illustrate the applicability for typical problems.

Large organic dyes play an important role in the construction of light-emitting diodes and recently promising new phenazine chromophores for efficient electroluminescence have been proposed[433]. A related pyrazine group containing system[434] (for structure see Figure 8.5) consisting of 192 atoms is investigated here as a first example. The geometry optimization of this compound was conducted at the DFTB level with third-order charge density fluctuation correction and self-consistent charge redistribution[435] as implemented in the DFTB+ program.[436] Dispersion interactions were accounted for by means of the atom pair-wise DFT-D3 correction[133,437]. This ground state structure was taken for the xTB treatment which took about 10 seconds of computation time for the orbital generation step on a single-core desktop computer. For comparison the sTDA-DFT computed UV spectrum with the range-separated hybrid functional ωB97X-D3[114]/TZVP[162] is shown in Figure 8.5 together with the experimental one.[434] The SCF for this DFT calculation runs for about six hours on a single-core desktop computer.

Inspection of Figure 8.5 reveals a very close similarity between the DFT and xTB spectra which both agree well with the experiment. The lowest-lying broad signal actually consists of two overlapping bands A and B corresponding to two transitions with strongly mixed excited configurations. This mixing is somewhat underestimated by sTDA-xTB leading to a lower state with too small intensity. However, small changes in the structure can change such a mixing in large multi-chromophoric systems considerably and hence we do not consider this as a major error. Note that no artificial low-lying CT states appear in sTDA-xTB, which plague TD-DFT with semi-local (or low Fock-exchange containing hybrid) density functionals for such large systems (see below and, e.g., Ref. 438 for another example).

As mentioned above, sTDA-xTB is free of the charge-transfer problem which occurs in TD-DFT with approximate functionals[129,439] and likewise in TD-DFTB.[328] We consider here a dual-chromophore perylene-zinc-porphyrine compound[440] as an exemplary system (Figure 8.6). In order to illustrate the problem, we conducted TD-DFT calculations (30 states) with the

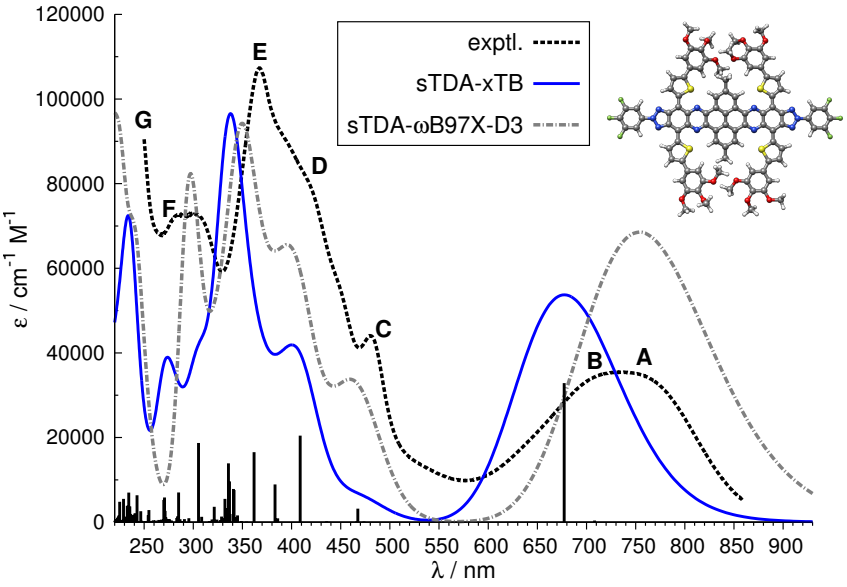


Figure 8.5.: Comparison of sTDA-xTB, sTDA-ωB97X-D3/TZVP, and experimental UV spectra (in dichloromethane) for the dye shown. The calculated intensity was scaled by a factor of 0.25 (0.5 for ωB97X-D3) and the theoretical spectra are shifted by −0.35 eV (−0.2 eV for ωB97X-D3). The black sticks indicate the relative intensity of the individual transitions (oscillator strengths) obtained by sTDA-xTB.

semi-local PBE functional[104] for comparison.

As can be seen from the experimental spectrum, there are three distinct low-lying electronic transitions with substantial absorption probability in the 2–3 eV range. They are well reproduced by sTDA-xTB with the reservation that the splitting between A and B is too small and that transition B is too low in intensity. Nevertheless, the computed spectrum is qualitatively correct and no artificial "ghost" states occur. This is different in TD-DFT(PBE) which incorrectly computes 11 low-lying states with vanishing intensity below the first bright one. Furthermore, the three bands A–C in fact occur as higher-lying roots but obviously are too weakly coupled and hence their splitting is worse than with sTDA-xTB.

The computation of electronic spectra for open-shell species is less common in the literature and in particular semiempirical methods are less well tested. Nevertheless, such compounds are interesting as reaction intermediates or low-energy absorber/emitter and electronic spectra are often an important tool for characterization. Figure 8.7 shows measured low-energy UV spectra for two typical cases (one cation and anion, experimental data from Refs. 441 and 442) in comparison to results of a vertical sTDA-xTB treatment.[443]

As can be seen, the agreement between theory and experiment is excellent regarding band positions and intensities. For the oligothiophene radical cation, we observe a somewhat larger absolute error of about 1 eV (theoretically blue-shifted bands) while the value of 0.4 for the anionic dye is typical considering the neglected solvent as well as geometry relaxation effects.

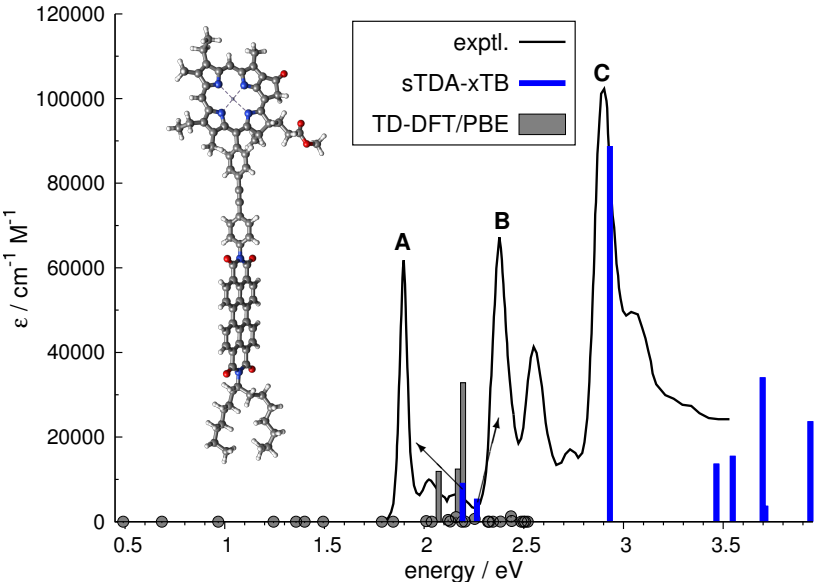


Figure 8.6.: Comparison of sTDA-xTB, TD-DFT(PBE/def2-TZVP) and experimental UV spectra for a perylene-Zn-porphyrine type charge-transfer system. The positions and relative intensity of the calculated transitions are indicated by bars (if $f > 0.1$) and circles. The theoretical energies are not shifted. Note the absence of any artificial low-lying "ghost" states for sTDA-xTB.

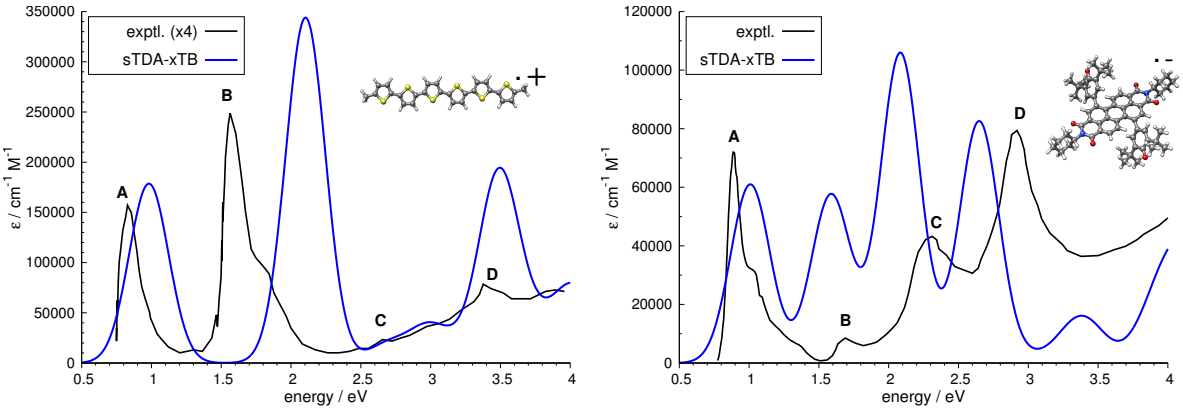


Figure 8.7.: Comparison of sTDA-xTB and experimental UV spectra for a hexathiophene radical cation (left, in dichloromethane) and a radical anion of the perylene bisimide derivative shown (right, in dichloromethane). The absolute experimental intensities are scaled by a factor of four for the former. For the perylene bisimide derivative, absolute experimental intensities were not measured and the values are scaled to approximately match the theoretical data. The theoretical spectra are shifted by -1 eV and -0.4 eV, respectively.

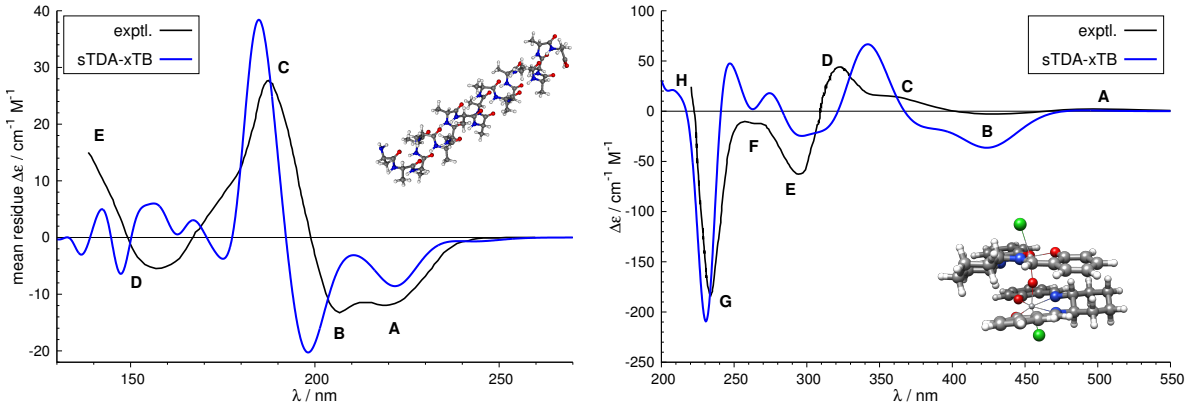


Figure 8.8.: Comparison of sTDA-xTB and experimental ECD spectra for an Ala_{20} polypeptide model (left) and the bis-titanium complex shown (right, in dichloromethane). The theoretical spectra are shifted by −0.4 eV. The experimental polypeptide ECD spectrum does not refer to Ala_{20}, but to poly-γ-methyl glutamate (in hexafluoro-2-propanol) showing the prototypical ECD features of an α-helix.

The good performance of sTDA-xTB in the unrestricted form for these examples supports the good results found for the smaller open-shell molecules in the fitting set.

Chiral molecules are important in nucleic acid, peptide, and sugar chemistry and the determination of the absolute configuration of, e.g., natural products is an important task for quantum chemistry. Electronic circular dichroism (ECD) is sensitive to molecular conformation and furthermore a routine tool for the characterization of the secondary structure of proteins. For an overview of recent theoretical work for very large systems which usually refers to a semiempirical exciton-coupling model, see Refs. 20,444, and 349. The different interaction of the left- and right-handed enantiomers with left- and right-handed circularly polarized light, respectively, is theoretically more complicated because the relative orientation of electric and magnetic transition moments is involved.[11,15] It represents a challenging test for any quantum chemical method.

Two closed-shell examples are shown in Figure 8.8. Two oligo-alanine$_n$ (n=10, 40) chains in the α-helix conformation have already been investigated in Ref. 46 at the sTDA-BHLYP level. The experimental spectrum corresponding to poly-γ-methyl glutamate in hexafluoro-propanol[346] is used here for comparison but, e.g., the one of myoglobin[445] (large α-helix content) in water looks rather similar. The structure of Ala_{20} was optimized at the HF-3c level[169]. As usual the intensities are normalized to the number of residues. The second example is a helical bis-titanium complex from Ref. 446.

The theoretical description of the poly-peptide ECD spectrum is almost perfect and only the band D is computed with too small intensity. Further studies will have to show if this holds also for other secondary structure types and if this excellent performance persists on an MD level (see below), it would open new possibilities for the elucidation of protein structure by combined

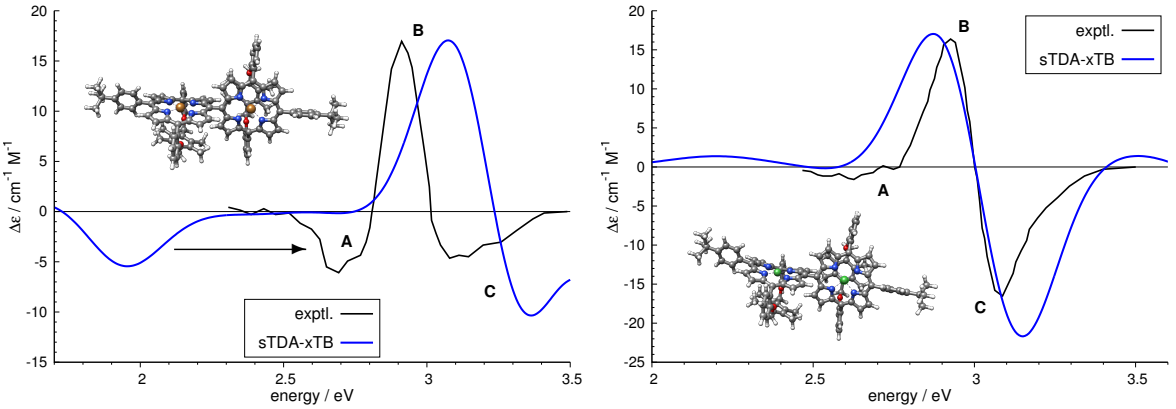


Figure 8.9.: Comparison of sTDA-xTB and experimental ECD spectra for the M-helical copper (left) and nickel complex shown (right). The theoretical spectra are shifted by -1 and -0.8 eV, respectively, and the experimental intensities were scaled to approximately match the computed data.

experimental and theoretical ECD. The agreement of the experimental and sTDA-xTB spectra for the bis-titanium complex (optimized at the TPSS-D3/def2-TZVP level)[105,133,134,162,163] is also remarkable considering its more complicated electronic structure. In particular, local transitions involving the metal d-orbitals with large magnetic but weak electric moments represent a challenge for any theoretical method. Although the computed spectrum is not fully correct (bands C and F), a clear assignment of absolute configuration could still be made on the basis of the sTDA-xTB calculation.

The last ECD examples, i.e., two large, chiral bis-metallo-porphyrin compounds (M-enantiomers) as recently synthesized in the group of Bringmann[397] are even more challenging (see Figure 8.9). Here we investigate the same ligand framework but with different metals in the porphyrin core (Cu vs. Ni) leading to an open-shell triplet case (two doublet-coupled copper-porphyrins) and a closed-shell system for nickel, respectively. This tests the consistency of our modifications of the unrestricted Hamiltonian in a complex setting. Unfortunately, the experimental spectra were recorded only in the small energy range between 2 and 3.5 eV and absolute intensities are not available so that more detailed comparisons are not possible at this point.

The two experimental spectra show some similarities but also distinct differences which are qualitatively reproduced by the sTDA-xTB calculations. The relative splittings of the individual bands B and C are correctly described. The excitations in this range originate from coupled $\pi\pi^*$ transitions on the porphyrin moieties (corresponding to the Soret bands). The energy of band A in the copper complex is underestimated relative to bands B and C, but produced with correct sign. In the nickel complex, negative rotatory strengths are present around band A, however, they are quenched by neighboring positive bands. For both systems we obtain a few more lower-lying states in the (non-shifted) range 1.3–2 eV which are, however, not observed because of tiny rotatory strengths ($< 1-2\cdot10^{-40}$ cgs units). Considering the complex molecular

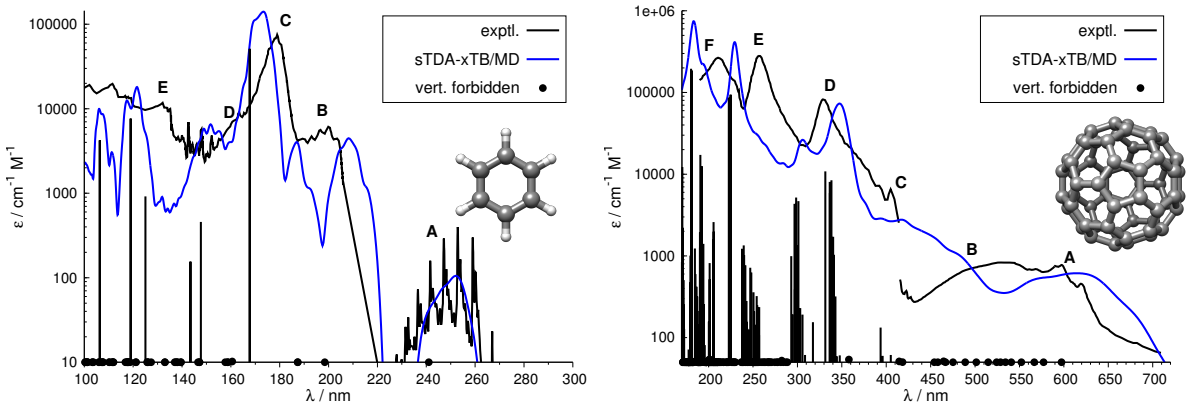


Figure 8.10.: Comparison of sTDA-xTB/MD and experimental UV spectra for benzene (gas phase[449]) and C_{60} (in *n*-hexane[450]). The black sticks indicate the relative oscillator strengths of the vertical transitions from a single ground state equilibrium structure. Dipole forbidden transitions are indicated by circles. The 0.25 ns MD trajectories refer to gas phase, NVT conditions (at about 300 K) from which 250 equidistant snapshots were taken. For C_{60}, a shift of −0.35 eV was applied to the computed transition energies.

structure and the difficulties arising from the coupling of two large chromophores in a close to achiral (orthogonal) arrangement, we consider this example, and in particular, the copper case (with the difficulty attributes: open-shell, large, 3d-transition metal, ECD) as a convincing show case for the method.

In the last examples, we want to demonstrate the effect of averaging spectra along molecular dynamics trajectories. For highly symmetric molecules, this should provide realistic absorption intensities also for vertically forbidden transitions. Note, however, that such treatments do not involve the proper quantum mechanical, vibronic coupling which can routinely be calculated on a TD-DFT level[283,447,448] for smaller systems. Instead it should provide a reasonable account of the nuclear dynamics for low-frequency modes (conformational sampling) and eventually should be used to automatically treat rather flexible systems. Here we consider benzene and the rather rigid aromatic C_{60} fullerene with I_h symmetry. For both molecules the low-energy part of the vertical UV spectrum is dominated by dipole-forbidden transitions. In the case of benzene, all excitations up to 20 eV were computed. The MD (250 ps production run time) is based on a PBEh-3c derived QMDFF force-field.[49]

The comparison to the experimental spectra in Figure 8.10 shows that sTDA-xTB/MD reproduces all major features of the measurements regarding relative intensity and band position fairly well. For C_{60}, the computed spectrum is of similar quality as that obtained at the computationally much more involved sTDA-PBE0 level.[46] In general, the peak maxima are red-shifted by 0.2–0.3 eV in the MD compared to the vertical treatment mostly improving agreement between theory and experiment. This is consistent because our reference structures used in the parameter fitting refer to the ground state equilibrium structures. The lower energy part of the C_{60}

spectrum seems to consist of two close-lying bands A and B which are dipole-forbidden vertically and that gain the correct intensity by the MD averaging over many distorted structures. The intensities of the so-called L_b and L_a vertically dipole forbidden transitions in benzene (bands A and B) are reproduced well in the MD. For the L_a band (B) we note some red-shift of 0.3 eV compared to the experiment while the intense band C (transition to E_{1u} state) improves its position considerably in the MD. Energetically low-lying artifacts which may occur by excited state potentials with incorrect shape are not observed in both simulations.

8.4. Conclusion

The main aim of this work was to develop a robust and broadly applicable semiempirical quantum chemical method for the computation of electronic absorption spectra. The size of the target systems is around 500-1000 atoms for which reasonably accurate spectra should be computable in minutes on standard desktop computers. High computational efficiency is also mandatory for the envisaged spectral averaging along molecular dynamics trajectories which requires hundreds of computations.

The goals could be achieved by combining the well established and extremely fast sTDA procedure with a newly developed semiempirical tight-binding based approach. The entire method is dubbed sTDA-xTB. The existing sTDA scheme was modified slightly only in two aspects, namely the more accurate calculation of dipole rotatory strengths applying a newly developed A+B/2 vector correction and an exchange integral correction for monopole approximation induced errors for very localized states. Importantly, the new approach does not suffer from the well-known charge-transfer state problem in TD-DFT with semi-local (or low Fock-exchange fraction) density functionals.

The xTB is represented by a composite scheme which employs two different basis sets and differently parametrized Hamiltonians for the calculation of atomic charges and the one-particle spectrum, respectively. It uses slightly modified matrix element expressions from the well-known self-consistent charge (SCC) density functional tight-binding (DFTB) scheme in restricted as well as unrestricted (open-shell) implementations. Empirical parameters for 32 elements including 3d-transition metals are provided. The remaining atomic parameters for the rest of the periodic system can be determined in a simple and straightforward way and will be reported in future work. As input for the parametrization, only theoretical reference data (CM5 atomic charges and vertical excitation energies) from DFT, TD-DFT, and SCS-CC2 treatments were employed.

Two new key features make the method accurate and robustly applicable. First, in order to simultaneously and consistently describe energetically low-lying valence states in, e.g., π-systems and saturated molecules with dominantly Rydberg states, diffuse (small exponent) basis functions are added to a standard minimal valence basis. This ensures that the empirical parametrization process does not artificially need to describe diffuse components in the excited states as they are correctly covered by the AO basis set. In this way, electronic spectra of, e.g., water, alkanes, or per-fluorinated compounds can be investigated even at high excitation energies

around 10 eV. Comparisons with SCS-CC2 calculations show that also the dipole oscillator strengths for these Rydberg transitions are described reasonably well by sTDA-xTB. The second point concerns the calculation of the atomic charges which are required in the TB Hamiltonian and that describe the effective electronic Coulomb interaction in the system. Here we follow a new strategy which involves a specially parametrized TB version in a partially polarized minimal valence basis set to compute just these charges which are then used as subsequently to define a second, extended basis set Hamiltonian. In total 23 global empirical parameters define the method while each element requires 9-15 additional parameters. As usual only a molecular structure, the molecular charge, and total spin state are required for an sTDA-xTB calculation which is furthermore specified by the spectral excitation energy range of interest (usually 0–8 eV).

We presented results for a few hundred molecules in comparison to *ab initio* reference data for atomic charges, spin populations, bond orders, and most importantly vertical excitation energies which is our main target quantity. Relatively simple organic and main group molecules as well as electronically more complicated metal and transition metal complexes were thoroughly tested. Overall, the new semiempirical method performs excellently providing robust atomic charges and bond orders of roughly DFT quality with typical deviations of only 5–10%. We propose (independently from the sTDA excited state treatment) to use these data in various large scale theoretical chemistry applications (e.g., as reactivity or polarity descriptors).

The sTDA-xTB excitation energies are of good quality with typical deviations of 0.3–0.5 eV. The consistently good performance for low-energy valence (0–5 eV) as well as high-energy (5–11 eV) Rydberg transitions is impressive and unprecedented at the semiempirical level. For electronically not too complicated systems, the accuracy of the model is roughly of TD-DFT/hybrid functional quality or even better for Rydberg or charge-transfer transitions. Other semiempirical methods are clearly outperformed. The computed UV and ECD spectra for a few examples compare well with experimental spectra indicating that the transition moments (which were not included in the reference data sets/fitting procedure) are also evaluated rather accurately. In two cases, we conducted an averaging of spectra along molecular dynamics trajectories. This yields realistic absorption intensities also for vertically forbidden transitions and shows that the sTDA-xTB excited state nuclear potential near the equilibrium structure is reasonable. A combined sTDA-xTB/MD approach thus opens new possibilities for the computation of ECD spectra, e.g., for peptides and DNA/RNA which require a dynamical treatment due to their flexibility.

Further work will complete the xTB parametrization for the largest part of periodic system of elements (i.e., up to $Z = 86$). We hope that this provides a robust and useful routine computational tool for the quantum chemistry of large and very large systems. Executables of the `xtb` and `stda` computer codes can be downloaded from the authors' web page.[306]

Acknowledgments

This work was supported by the DFG in the framework of the "Gottfried-Wilhelm-Leibniz" prize and the Sonderforschungsbereich 813 ("Chemistry at spin centers"). The authors thank Dr. M. Steinmetz for the spin-constant and TD-DFTB computations, J. Seibert for MD related programming work, S. Steiniger for setting up parts of the test set (Table 8.5), and E. Fernandes Lima for initial tests. Furthermore, both authors are deeply grateful to Prof. Dr. Tom Ziegler for fruitful discussions on the ECD problem in TDA which led to the development of the A+B/2 correction scheme.

9. A Robust and Accurate Tight-Binding Quantum Chemical Method for Structures, Vibrational Frequencies, and Non-Covalent Interactions of Large Molecular Systems Parametrized for All spd-Block Elements (Z=1–86)

Stefan Grimme,[*] Christoph Bannwarth,[*] and Philip Shushkov[*]

Received 4th of February 2017, Published online 18th of April 2017

Reprinted (adapted) with permission from[†]
Grimme, S.; Bannwarth, C.; Shushkov, P. *J. Chem. Theory Comput.* **2017**, *13*, 1989–2009.

DOI 10.1021/acs.jctc.7b00118

Own manuscript contribution

- Contributing to the development of the GFN-xTB Hamiltonian
- Implementation of the coordination number dependence (see Eq. 9.12), including gradients, into the `xtb` code
- Performing all GFN-xTB calculations and all benchmark calculations with other semiempirical methods except for the GFN-xTB fit procedure, the atomization energy calculations (G3/99 set), and the enthalpy of formation calculations
- Interpretation of the results
- Writing major parts of the manuscript

[*]Mulliken Center for Theoretical Chemistry, Institut für Physikalische und Theoretische Chemie, Rheinische Friedrich-Wilhelms-Universität Bonn, Beringstraße 4, 53115 Bonn, Germany

Abstract We propose a novel, special purpose semiempirical tight-binding (TB) method for the calculation of structures, vibrational frequencies, and non-covalent interactions of large molecular systems with 1000 or more atoms. The functional form of the method is related to the self-consistent density functional TB scheme and mostly avoids element pair-specific parameters. The parametrization covers all spd-block elements and the lanthanides up to Z=86 using reference data at the hybrid density functional theory level. Key features of the Hamiltonian are the use of partially polarized Gaussian-type orbitals, a double-ζ orbital basis for hydrogen, atomic-shell charges, diagonal third-order charge fluctuations, coordination number-dependent energy levels, a non-covalent halogen-bond potential, and the well-established D3 dispersion correction. The accuracy of the method, called Geometry, Frequency, Non-covalent, eXtended TB (GFN-xTB), is extensively benchmarked for various systems in comparison with existing semiempirical approaches, and the method is applied to a few representative structural problems in chemistry.

9.1. Introduction

In quantum chemistry, geometry optimizations, vibrational frequency calculations, and molecular dynamics simulations are currently dominated by Kohn-Sham density functional theory (DFT) for system sizes up to a few hundred atoms. These computations serve as a starting point for investigations of various spectroscopic and/or thermochemical properties, possibly with higher-level wave function theory (WFT) methods. For fairly large systems with 1000 or more atoms, however, even simplified schemes such as the PBEh-3c hybrid DFT[110,451] and Hartree-Fock/minimal basis set (HF-3c[169]) composite methods, which we have proposed in the past few years, become computationally unfeasible. For instance, many interesting proteins with 2000–5000 atoms are out of reach for routine DFT optimizations and such investigations are conducted at the force-field (FF) level.[452–454] The disadvantages of the FF treatment are numerous, including difficulties with metals, proton transfer and protonation states, polarization and chemical reactions, and can only be circumvented by a quantum mechanical (QM) treatment of the electrons.

Semi-empirical quantum-mechanical (SQM) methods[37,38,455] provide a well-known alternative route because they are at least two orders of magnitude faster than conventional DFT treatments. SQM methods approximate single-reference Hartree-Fock (HF) or *first principles* DFT theory and have been investigated extensively in the 1980s and 1990s. The development of SQM methods has recently seen a renewed interest triggered primarily by the advent of the density functional tight-binding (DFTB) method pioneered by Seifert, Elstner, and Frauenheim.[327,404,411,435,456,457] In the most recent self-consistent-charge (SCC) form with third-order charge fluctuation terms,[435] the method is relatively accurate for a wide range of chemical systems[437,458–463] and has been implemented in several quantum chemical codes.[426,427,436,464,465] SQM methods derived from HF theory neglect to a varying degree the differential overlap between atomic basis functions, which results in vanishing classes of one- and two-electron integrals

and thus substantial computational savings. Depending on the extent of approximations, these SQM methods fall into various categories such as complete or intermediate neglect of differential overlap (CNDO or INDO). We refer the reader to Ref. 38 for an excellent review of SQM methods, including non-covalent interactions (NCI). Other notable recent developments in the field of SQM methods include modifications of PM6[408,466–470] and OM2,[406,471–474] and the development of PM7,[475] INDO/X,[320] and MSINDO-sCIS.[319] For a recently proposed semiempirical minimal-basis scheme in between DFT and DFTB employed in embedding approaches, see Ref. 476. Recently we have also proposed an extended special purpose TB variant, which uses an augmented atomic orbital (AO) basis set, for the fast computation of electronic spectra.[477] The encouraging performance of this method, called sTDA-xTB, has stimulated the current work.

The objective of the paper is to propose a new semiempirical TB method, called GFN-xTB, where GFN indicates the design of the approach to yield reasonable Geometries, vibrational Frequencies, and Non-covalent interactions, and "x" stands for extensions in the AO basis set and the form of the Hamiltonian. In general, GFN-xTB provides for molecules from the whole periodic table higher accuracy for the target properties than existing "general-purpose" semiempirical approaches. It is applicable to a wider range of systems and is computationally and numerically more robust than other schemes with comparable accuracy. The premise is that SQM methods cannot provide the same level of accuracy for a broad range of chemical properties, such as molecular structures and chemical reaction energies. Thus, the compromise established in conventional SQM methods is often inaccurate and likely limits their use in many applications. This has led to the poor reputation of SQM methods in the 1990s, which were eventually surpassed by the successful performance of DFT. Special purpose SQM methods that aim at a specific chemical or physical property from the outset have been pioneered by Truhlar and coworkers for computing reaction dynamics,[416] atomic charges[414] (see also Ref. 478), and solvation free energies.[415] Along these lines, we found recently that certain SQM methods can be useful in gas phase reaction dynamics to simulate electron-impact mass spectra (QCEIMS[474,479]), and the applications to this problem have fueled the current development.

The method described herein has the following characteristics:

- Structural properties around equilibrium, such as vibrational frequencies and non-covalent interactions, are the main target quantities. These are difficult to obtain with *ab initio* procedures for systems with more than 1000 atoms and are important in many chemical problems. Chemical reaction energies are not part of the training set and are thus less accurate than with general purpose SQM methods. However, some relevant thermochemical properties outside of the training set (dissociation energies, relative proton affinities (PA), and relative ionization potentials (IP)) are obtained reasonably accurately such that potential energy surfaces (PES) are useful for high-temperature molecular dynamics calculations in modeling mass spectra. Furthermore, reasonable enthalpies of formation can be obtained by correcting the computed atomization energies with a sum over atom-specific terms. The accuracy for conformational energies is comparable or better than existing SQM methods.

- Most semiempirical methods employ minimal AO basis sets, which limit the accuracy and can lead sometimes to unphysical results, especially for heavier elements. The GFN-xTB basis set consists of a minimal basis set of atom centered, approximate Slater functions augmented with a second s-function for hydrogen to improve the description of hydrogen bonding and d-polarization functions for higher row elements to facilitate the description of hyper-valent bonding arrangements. The Slater functions are approximated as contractions of standard primitive Gaussian functions (STO-mG).[410]

- The number of parameters is kept to a minimum. Thus, pair-specific potentials, which are characteristic of DFTB, are avoided, and mainly global and element-specific parameters are employed. For example, for each of the second row elements B–F only 11 parameters need to be determined, the two atomic energy levels (2s/2p) and corresponding Slater AO exponents, two repulsion parameters, two off-diagonal Hamiltonian polynomial factors, two chemical hardness parameters, and one shell-charge scaling parameter. The method is therefore easy to parametrize using reference data generated by hybrid DFT calculations. This parametrization scheme is flexible and can be adjusted to specific systems or chemical properties.

- The method is applicable to diverse chemical systems composed of elements from different corners of the periodic table and ranging from organic molecules to inorganic main group compounds and (transition-)metal complexes. Biochemical systems like (metallo-)proteins, RNA and DNA, and their inter- and intramolecular interactions are of primary interest. Electronically degenerate (multi-configurational) situations are handled by a finite electronic temperature approach,[412] which allows the qualitatively correct description of biradicals, orbital-symmetry forbidden reactions, and dissociation processes.[474,480]

The paper is organized as follows: A detailed description of the GFN-xTB Hamiltonian is provided in Section 9.2.1, and the parametrization procedure is outlined in Section 9.2.2. The GFN-xTB performance for structures and NCIs is compared to contemporary semiempirical methods in Section 9.3.

9.2. Theory

9.2.1. The GFN-xTB Hamiltonian

In the following, atomic units are used, if not stated otherwise. The total GFN-xTB energy expression consists of electronic (el), atom pair-wise repulsion (rep), dispersion (disp), and halogen-bonding (XB) terms

$$E = E_{el} + E_{rep} + E_{disp} + E_{XB}\,. \tag{9.1}$$

The electronic energy E_{el} is given by

$$E_{el} = \sum_{i}^{occ.} n_i \langle \psi_i | H_0 | \psi_i \rangle + \frac{1}{2} \sum_{A,B} \sum_{l(A)} \sum_{l'(B)} p_l^A p_{l'}^B \gamma_{AB,ll'} + \frac{1}{3} \sum_{A} \Gamma_A q_A^3 - T_{el}\, S_{el}\,, \tag{9.2}$$

where ψ_i are the valence molecular orbitals (MOs) with occupation numbers n_i (which may be fractional, see Eq. 9.6), and H_0 is the zero-order Hamiltonian. The electronic free energy $T_{el}\, S_{el}$ term is discussed below. The SCC contributions are represented by a second-order term and a third-order diagonal contribution (the second and third term in Eq. 9.2), where q_A is the Mulliken charge of atom A and Γ_A is the charge derivative of the atomic Hubbard parameter η_A. The summation in the second-order term is over all shells l and l' located on atoms A and B with p_l^A the charge distributed over the orbital shells with angular momenta l on atom A

$$p_l^A = p_l^{A_0} - \sum_{\nu}^{N_{AO}} \sum_{\mu \in A, \mu \in l} S_{\mu\nu} P_{\mu\nu}\ . \tag{9.3}$$

Here, N_{AO} is the total number of AOs and $p_l^{A_0}$ is the reference shell occupation of the free atom (e.g., for carbon $p_{2s}^{A_0} = 2$, $p_{2p}^{A_0} = 2$). Instead of the ground state shell occupations of the free atoms, the canonical ones are chosen that result from the "Aufbau" principle together with ignoring well-known exceptions (e.g., for Cr we take a $3d^4 4s^2$ instead of the ground state $3d^5 4s^1$ occupation). The shell atomic charges sum to the partial atomic charges, $\sum_{l \in A} p_l^A = q_A$. Please refer to Ref. 481,482 for more details of the DFTB treatment of d-block elements. The third-order density fluctuations terms are described in Refs. 456 and 435. Our approach is restricted to a diagonal third-order term, which reduces the numerical complexity but keeps the most essential part. The combination of l-dependent second-order and diagonal third-order density fluctuation terms is used for the first time in a DFTB context.

The distance dependence of the Coulomb interaction is described by the generalized Mataga-Nishimoto-Ohno-Klopman[46,288–290] formula, which is given by:

$$\gamma_{AB,ll'} = \left(\frac{1}{R_{AB}^{k_g} + \eta^{-k_g}} \right)^{1/k_g}\ . \tag{9.4}$$

Here, R_{AB} is the interatomic distance, k_g is a global parameter, and the average chemical hardness η of the two atoms A and B is defined as

$$\eta = 2 \left(\frac{1}{(1+\kappa_A^l)\eta_A} + \frac{1}{(1+\kappa_B^{l'})\eta_B} \right)^{-1}\ , \tag{9.5}$$

where η_A and η_B are taken as element-specific fit parameters, and the parameters κ_A^l and $\kappa_B^{l'}$ are element-specific scaling factors for spd-levels.

In case of fractional orbital occupations, an electronic temperature times entropy term, $-T_{el}S_{el}$ enters the electronic energy (see Eq. 9.2). The electronic temperature serves as an adjustable

Table 9.1.: Description of the Slater type AO basis sets. n denotes the principal quantum number of the valence shell of the element.

element	basis functions
H	ns, (n+1)s
He	ns
Be-F,group I, Zn, Cd, Hg, Tl-Bi	nsp
group II-VIII	nsp, (n+1)d
transition metals and lanthanides	(n−1)d,nsp

parameter (default: $T_{el} = 300\,\mathrm{K}$), and the so-called Fermi smearing is employed to achieve fractional occupations for systems with near-degenerate orbital levels. The orbital occupations for the spin orbital ψ_i are given by[412]

$$n_i(T_{el}) = \frac{1}{\exp[(\epsilon_i - \epsilon_F)/(k_B T_{el})] + 1} \ , \tag{9.6}$$

where ϵ_i is the orbital energy of the spin orbital ψ_i and ϵ_F – the Fermi level. Higher temperatures may be used in certain cases, e.g., to investigate thermally forbidden reaction pathways (see Section 9.3.4). In this way, an electronic free energy, including an entropy contribution, is variationally minimized in the SCC procedure. This should not be confused with a true thermostatistical treatment because the Fermi smearing technique is used only as a way to take static correlation into account.

Expressing the molecular orbitals ψ_i in Eq. 9.2 as a linear combinations of atom-centered orbitals (LCAO),[117,118]

$$\psi_i = \sum_{\mu}^{N_{AO}} C_{\mu i} \phi_\mu(\zeta, \text{STO-}m\text{G}) \ , \tag{9.7}$$

and variational minimization of the energy expression in Eq. 9.2 with respect to the linear coefficients $C_{\mu i}$ leads to the general eigenvalue problem

$$\mathbf{FC} = \mathbf{SC}\boldsymbol{\epsilon}, \tag{9.8}$$

where $\mathbf{F}$ denotes the approximate TB Fock/Kohn-Sham matrix in the AO basis, $\mathbf{S}$ is the AO overlap matrix, and $\mathbf{C}$ are the MO coefficients with eigenvalues ϵ. For reasons of computational efficiency and robustness, open-shell systems are treated in a restricted manner by a fractional occupation of a single set of spatial orbitals (see below). The spherical Slater type AOs ϕ are approximated by Stewart's Gaussian expansions[410] with $m = 4 - 6$ primitives per AO. They are available up to principal quantum number $n = 6$ (for $l = 2$ up to $n = 5$). The exponents ζ are element-specific, fitted parameters. A description of the AO basis is given in Table 9.1. For most main group elements of third and higher rows, which can form hypervalent electronic structures, an additional single d-polarization function is provided. The same holds

for earth alkaline metals, which require d-functions in certain bonding situations (e.g., bent MX_2 molecules where X is a halogen). Polarization functions on elements H to F do not substantially improve the results and increase the computational cost considerably (by a factor of 5–10). For the group IIb elements, Zn-Hg, the d-electrons are not treated as valence (sp-shell only). As we do not focus on magnetic and electronic spectroscopic properties, the lanthanides are treated as 4d transition metals without explicit consideration of the f-electrons (i.e., all lanthanides have three valence electrons). This "f-in-core" approximation is rather accurate according to *ab initio* pseudo-potential calculations.[483] The actinides, for which this approximation is not possible, are presently not covered.

The matrix elements of the GFN-xTB Hamiltonian are computed similarly to the second- and third-order SCC-DFTB3 methods[411,456,481]

$$\begin{aligned}\langle\phi_\mu|F|\phi_\nu\rangle =&\langle\phi_\mu|H_0|\phi_\nu\rangle + \frac{1}{2}S_{\mu\nu}\sum_C\sum_{l''}(\gamma_{AC,ll''}+\gamma_{BC,l'l''})p^C_{l''}\\ &+\frac{1}{2}S_{\mu\nu}(q_A^2\Gamma_A+q_B^2\Gamma_B),\quad (\mu\in l(A),\nu\in l'(B)),\end{aligned} \tag{9.9}$$

where indices μ and ν label the AOs with corresponding angular momenta l and l' and the second sum runs over all atoms C and their shells l''. The H_0 elements are given by

$$\begin{aligned}\langle\phi_\mu|H_0|\phi_\nu\rangle =&K_{AB}\frac{1}{2}(k_l+k_{l'})\frac{1}{2}(h^l_A+h^{l'}_B)S_{\mu\nu}(1+k_{EN}\Delta EN^2_{AB})\\ &\times\Pi(R_{AB,ll'}),\quad (\mu\in l(A),\nu\in l'(B)),\end{aligned} \tag{9.10}$$

where k_l and $k_{l'}$ are the well-known "Hückel" constants taken as free parameters for each angular momentum (l for A, l' for B). h^l_A and $h^{l'}_B$ are effective atomic energy levels. For s-p interactions and those with the hydrogen 2s function, two parameters k_{sp} and $k_{2s,H}$ replace the average $\frac{1}{2}(k_l+k_{l'})$ in Eq. 9.10. The electronegativity difference of the two atoms, $\Delta EN_{AB}=EN_A-EN_B$, computed with standard Pauling values modifies the matrix elements with the global parameter k_{EN} as a coefficient of proportionality. K_{AB} is a scaling constant, which differs from unity only for a limited number of element pairs. We observed that after the optimization of element-specific and global parameters, the bond strength between certain elements was slightly unbalanced, and we introduced eight element-pair specific and intermetallic K_{AB} parameters (see Table 9.2). We would like to emphasize that we avoided element pair-specific parameters in general and introduced the ones listed in Table 9.2 to fine tune the accuracy for properties such as conformational energies, which were significantly improved by reducing the value of K_{HH}. In our opinion, the element-specific parameters form a reasonably accurate parametrization suitable for general chemistry and users can customize the Hamiltonian by varying certain K_{AB} parameters.

$\Pi(R_{AB,ll'})$ is a distance and l-dependent function given by:

$$\Pi(R_{AB,ll'})=\left(1+k^{poly}_{A,l}\left(\frac{R_{AB}}{R_{cov,AB}}\right)^{\frac{1}{2}}\right)\left(1+k^{poly}_{B,l'}\left(\frac{R_{AB}}{R_{cov,AB}}\right)^{\frac{1}{2}}\right) \tag{9.11}$$

where $R_{cov,AB} = r_{cov,A} + r_{cov,B}$ is the covalent distance, r_{cov} is the covalent atomic radius,[484] and $k^{poly}_{A,l}$ and $k^{poly}_{B,l'}$ are element-specific parameters. Finally, the effective atomic energy levels of non-metals, h^l_A, depend linearly on the geometric D3 coordination number, CN,[133] of atom A

$$h^l_A = H^l_A(1 + k_{CN,l}CN_A), \quad (l \in A) \tag{9.12}$$

where $k_{CN,l}$ are three global scaling parameters for spd-shells. The dependence of the zero-order Hamiltonian on $\Pi(R_{AB,ll'})$, ΔEN and CN_A/CN_B, with the latter two contributions specific to our TB scheme and at variance with the widely used DFTB3, introduces more flexibility in the covalent binding part without the need for element pair-specific parametrization. A larger ΔEN will reduce the extent of "covalent" binding and effectively increase the ionic character for element pairs with a non-zero ΔEN. The dependence of h^l_A on the coordination number CN_A changes the energetic gap between the different l shells of atom A. This is of particular importance for the s and p levels as this affects the hybridization in different bonding situations (thus the dependency on the CN_A).

For the repulsion energy in Eq. 9.1, we employ an atom pairwise potential similar to the one proposed in Ref. 49.

$$E_{rep} = \sum_{AB} \frac{Z^{eff}_A Z^{eff}_B}{R_{AB}} e^{-(\alpha_A\alpha_B)^{\frac{1}{2}}(R_{AB})^{k_f}}, \tag{9.13}$$

where Z^{eff} are effective nuclear charges, k_f is a global parameter, and α are element-specific parameters. Z^{eff} are fitted parameters and the optimized values deviate by <20–30% from the nuclear charge, Z, for most elements. An exception from this general trend are the lanthanides, for which the difference is about 50%.

The dispersion energy in Eq. 9.1 is computed by the well-established D3 method[133] in the BJ-damping scheme[135] without three-body terms (see Ref. 130 for a recent review). The two short-range damping parameters a_1, a_2, and also s_8 are adjusted to yield good non-covalent interaction energies and geometries.

Halogen-bonding (XB) interactions are described poorly[38] by the point charge approximation of electrostatic interactions in all standard DFTB schemes. This holds also, e.g., for PM6, for which corrections have been developed to account for the deficiencies.[469,485] We use a pairwise repulsive correction of a modified Lennard-Jones form

$$E_{XB} = \sum_{XB} f^{AXB}_{dmp} k_X \left(1 + \left(\frac{R_{cov,AX}}{R_{AX}}\right)^{12} - k_{X2}\left(\frac{R_{cov,AX}}{R_{AX}}\right)^{6}\right) / \left(\frac{R_{cov,AX}}{R_{AX}}\right)^{12}, \tag{9.14}$$

where $R_{cov,AX} = k_{XR}(r_{cov,A} + r_{cov,X})$ is an effective covalent distance and k_{XR} and k_{X2} are global parameters. The correction takes into account the halogens Br, I, and At, and N or O atoms are the donors B. The atom A closest to the halogen defines the AXB angle θ. The

damping function f_{dmp}^{AXB} is given by

$$f_{dmp}^{AXB} = \left(\frac{1}{2} - \frac{1}{4}\cos\theta_{AXB}\right)^6 \tag{9.15}$$

such that the correction vanishes for non-linear arrangements, and the magnitude of the interaction is determined by the parameter k_X.

The values of 16 global parameters for GFN-xTB, as obtained from a non-linear fit to reference data (see Section 9.2.2) are gathered in Table 9.2. The roughly 1000 element-specific parameters for the periodic table up to $Z = 86$ are listed in the Supporting Information*. This intermediate degree of parametrization is typical for many SQM methods, such as PM6[408] or MSINDO,[317] which rely on element-specific parameters. The K_{AB} values deviating from one are given in Table 9.2.

The new method is implemented in a standalone general quantum chemical program, called `xtb`. It features analytical nuclear gradients, shared memory parallelization, an efficient geometry optimizer[486] based on approximate normal coordinates (ANC) from Lindh's model Hessian,[487] molecular dynamics, conformational searches, and reaction path optimization modules. It has been tested in full protein structure optimizations with up to 3000 atoms, a few hundred ps dynamics for about 500 atoms, and in single-point computations of the fractional orbital density (FOD[480]) of about 10000 atoms. In addition, an efficient and computationally inexpensive solvation model supplemented with analytical nuclear gradients is developed. It is based on the generalized Born model for the dielectric response of the solvent,[488] a term proportional to the solvent accessible surface area (SA),[489] and a short-range hydrogen bonding contribution. Details of the new solvation model together with applications to (bio)chemical systems will be provided in a separate publication.[490] The complete program can be obtained upon request.[491]

9.2.2. Parametrization and Technical Details

The parametrization involved the determination of the global parameters in Table 9.2 and the element-specific parameters. To keep the total number of parameters to a minimum, the values for the lanthanides were obtained by linear interpolation with the nuclear charge Z, where only the start ($Z = 58$, Ce) and end points ($Z = 71$, Lu) were fitted.

The parameters were determined by a minimization of the root-mean-square deviation (RMSD) between calculated and reference data using the Levenberg-Marquardt algorithm.[419,420] Most molecular structures were optimized at the PBEh-3c hybrid DFT level,[110] which yields very accurate equilibrium (R_e) structures for a wide range of systems. PBE0-D3(BJ)/def2-TZVP[111,163] was the reference structure level for the rare gas systems. Global and atomic parameters were simultaneously optimized for the elements H, C, N, O. The rest of the atomic parameters were optimized element-wise while keeping all existing parameters fixed.

*The full supporting information is available online at http://dx.doi.org/10.1021/acs.jctc.7b00118. The detailed results obtained on the considered test sets (except for the G3/99 and the heat of formation set from Figure 9.6) are also included in Appendix A7.

Table 9.2.: Global empirical parameters defining the GFN-xTB method. The parameters are either dimensionless or in atomic units.

parameter	value
k_s	1.85
k_p	2.25
k_d	2.00
k_{sp}	2.08
$k_{2s,H}$	2.85
k_g	2.0
k_f	1.5
$k_{CN,s}$	0.006
$k_{CN,p}$	−0.003
$k_{CN,d}$	−0.005
k_{EN}	−0.007
a_1(D3)	0.63
a_2(D3)	5.0
s_8(D3)	2.4
k_{XR}	1.3
k_{X2}	0.44
K_{HH} [a]	0.96
K_{BH} [a]	0.95
K_{NH} [a]	1.04
K_{PB} [a]	0.97
K_{SiB} [a]	1.01
K_{NiH} [a]	0.90
K_{ReH} [a]	0.80
K_{PtH} [a]	0.80
K_{3d-3d} [a,b]	1.10
$K_{4d/5d/4f-4d/5d/4f}$ [a,b]	1.20

[a] Special H_0 scaling parameters K_{AB} (see Eq. 9.10) for certain element pairs and inter transition metal interactions (for 3d and 4d/5d/4f block elements).

[b] $K_{3d-4d/5d/4f} = \frac{1}{2}(K_{4d/5d/4f-4d/5d/4f} + K_{3d-3d})$.

The fitting for the periodic table was thus continued with the halogens, after which elements were treated in more or less canonical order while trying to minimize inter-dependencies. About

50–100 reference data points (see below) were used on average per fitted parameter. The reference molecules were predominantly of closed-shell character and covered common bonding situations. The human work time required to parametrize an element, including the generation of reference molecular structures, was about 1–2 days.

Five types of reference data were employed in the parametrization: a) equilibrium structures, b) distorted geometries with energies with respect to the equilibrium of a few kcal mol^{-1}, c) harmonic vibrational frequencies (mainly for small molecules with <10 atoms), d) CM5[402] atomic charges, and e) non-covalent interaction (NCI) energies and structures. For a) and b) a force matching procedure was applied where all atomic force components of the reference molecule were fitted. This procedure was not sufficient to determine the NCI related parameters. Therefore, full structure optimizations were carried out for a few small complexes from the S22 benchmark set[492,493] and the all-atom-structure RMSD from the reference was included in the fit. Spline-interpolated CCSD(T) data (minima and interaction energies) for the S66x8 NCI benchmark set[494] and interaction energies for the L7 set[495] were also used. The Mulliken charge-derived CM5 atomic charges were obtained by GFN-xTB as described previously.[477] The charges were used to stabilize the fit and to detect physically inconsistent parameter sets. Their relative contribution to the training set RMSD was within about 10% or less, which was smaller than the contribution of parts a)-c). It was observed during optimization that reasonably accurate charges (mean relative deviation from absolute PBE0/def-TZVP based reference values <10–20%) were obtained when reference structures and frequencies were described well. All global parameters in Table 9.2 except those related to halogen bonding were obtained by fitting to the HCNO training set. It comprised about 260 molecules ranging from diatomics to systems with about 100 atoms. The k_{XR} and k_{X2} halogen bond parameters were determined by fitting to energies and forces of the XB18 benchmark set.[496] The GFN-xTB harmonic vibrational frequencies were not scaled and were fitted to the corresponding PBEh-3c values scaled by 0.95.[110]

We used the TURBOMOLE suite of programs[164c,165,371] (version 7.0) to conduct most of the ground state DFT calculations and geometry optimizations. For the calculation of the Hirshfeld[413] based CM5 charges at the PBE0/def-TZVP level the ORCA code[48] was used. We employed standard exchange-correlation functional integration grids ($grid2$ and $finalgrid4$ in ORCA and $m4$ in TURBOMOLE if not noted otherwise), typical self-consistent field (SCF) convergence criteria ($10^{-7}\,E_h$), and the resolution of the identity (RI) integral approximation[76,167,168] in the DFT calculations.

Calculations for comparisons were conducted with the `DFTB+`[436] (DFTB3[435] with the 3OB parametrization[431,497,498]), `MOPAC16`[499] (PM6-D3H4X[408,468,469]), `MSINDO`[317] (version 3.6), and `MNDO99`[407] (OM2[406]) codes. The DFTB3 method was used in its self-consistent version, which we refer to as DFTB3. Other DFTB parametrizations[500] comprising some 3d-transition metal elements[501] are available; however, since we are focusing on organic systems, we employed the 3OB parametrization[431,497,498] throughout. PM6-D3H4X is applied as implemented (with the identical keyword) in the `MOPAC16` code with the D3 correction in the zero-damping scheme.[133,468]

The electronic energies obtained from DFTB3, MSINDO, and OM2 were corrected by the D3[133] approach using the rational/Becke-Johnson damping function.[134] For clarity, we use the suffix "-D3(BJ)" for these rationally damped schemes. The same correction is used in GFN-xTB, which is an integral part of the new method, and therefore no suffix is used here. For MSINDO, we used the neglect of diatomic differential overlap (NDDO) Hamiltonian and augmented it with the H+ hydrogen bond correction[502] (denoted by the "H+" suffix) as implemented in the `MSINDO` code. The D3(BJ) and H+ parameters for DFTB3-D3(BJ), OM2-D3(BJ), and MSINDO-D3(BJ)H+ were taken from Refs. 503, 474, and 437, and are listed in the Supporting Information*. Our standalone `dftd3` code[306] was used for the calculations of the D3(BJ) corrections. As in GFN-xTB, the Fermi smearing technique ($T_{el} = 300\,\mathrm{K}$) was employed in the DFTB3 and MSINDO procedures, which did not affect energies or structures of most of the systems considered here. Structures of systems that do not belong to established benchmark sets were compared to PBEh-3c,[110] if not stated otherwise. In these cases, the grid in TURBOMOLE was reduced to *m3*, and *g*-functions were removed from the auxiliary basis sets for the PBEh-3c calculations.

9.3. Results and Discussion

In this section, the performance of the GFN-xTB approach is benchmarked for structures, non-covalent interaction energies, and conformational energies. Existing SQM methods are assessed on the same sets and compared to the performance and computational cost of GFN-xTB. The discussion continues with an exploration of characteristic features of a number of potential energy surfaces and concludes with a demonstration of the efficiency of GFN-xTB in geometry optimizations of large chemical systems.

9.3.1. Structures

The geometry optimization of structures is one of the most important tasks for SQM methods, in particular for large systems, where it becomes prohibitively expensive at a *first principles* level such as DFT. We assess the performance of GFN-xTB for the computation of structures of systems consisting of main group elements and some organometallic complexes.

Organic and Main Group Molecules

We choose ROT34, LB12, and HMGB11 test sets to assess the performance for molecular structures of systems containing only main group elements. The ROT34[77,504,505] set consists of the experimentally derived rotational constants, B_e, of 12 small organic molecules.

Table 9.3.: Comparison to experimental data of optimized geometries with different semiempirical methods. Relative deviations in the rotational constants of medium sized molecules (ROT34)[a], as well as deviations in the bond lengths of very long intramolecular bonds (LB12)[c] and covalent bonds of heavy main group elements (HMGB11)[f] are considered.

	GFN-xTB	PM6-D3H4X	MSINDO-D3(BJ)H+	OM2-D3(BJ)	DFTB3-D3(BJ)	HF-3c
			ROT34 (deviations in %)[a]			
MRD:	0.4	−1.6[b]	1.8	1.6[b]	−1.3	−1.4
MURD:	1.1	2.5[b]	3.1	2.3[b]	1.5	1.4
SRD:	1.5	2.5[b]	3.7	2.8[b]	1.3	1.0
MAXR:	5.8	6.1[b]	11.7	8.5[b]	4.6	4.5
			LB12 (deviations in pm)[c,g]			
MSD:	−10.9	−5.5[d]	−17.6[e]	-	-	−5.2
MAD:	13.2	20.5[d]	27.7[e]	-	-	14.1
SD:	18.8	30.0[d]	31.9[e]	-	-	23.6
MAX:	55.7	64.2[d]	74.6[e]	-	-	61.1
			HMGB11 (deviations in pm)[f,g]			
MSD:	−0.5	−3.8	−12.0	-	-	5.5
MAD:	3.0	10.1	18.3	-	-	5.8
SD:	3.9	17.3	25.7	-	-	5.3
MAX:	8.5	42.6	71.7	-	-	15.7

[a] Rotational constants B_e (excluding vibrational effects) from Ref. 504 with an estimated reference error of 0.2%.

[b] Statistical data discarding the isoamyl-acetate for which a wrong conformer is obtained.

[c] Bond lengths of long bonds as used in Ref. 110.

[d] Statistical data discarding the transition metal containing systems HAPPOD[110] and KAMDOR and . In both cases, the optimization resulted in cleavage of the metal-metal bond.

[e] Statistical data discarding the transition metal containing system HAPPOD[110] where the optimization yielded a structure with dissociated (and reordered) ligands.

[f] Bond lengths of covalent bonds comprising heavy main group elements ($n_{val} \geq 3$) as used in Ref. 110.

[g] The reference bond lengths have an estimated uncertainty of 2 pm possibly resulting from crystal packing and/or vibrational effects.

MRD=mean relative deviation, MURD=mean unsigned relative deviation, SRD=standard relative deviation, MAXR=maximum unsigned relative deviation, MSD=mean signed deviation, MAD=mean absolute deviation, SD=standard deviation, MAX=maximum absolute deviation.

The ROT34 set is a sensitive measure for the accuracy of chemical structures because small changes in the bond lengths and angles can result in significant deviations from accurately measured rotational constants. Overestimated values indicate underestimated bond lengths,

owing to the reciprocal dependence of B_e on the moment of inertia. The performance of GFN-xTB is compared to other dispersion-corrected semiempirical methods, including DFTB3-D3(BJ),[431,435,503] HF-3c,[169] MSINDO-D3(BJ)H+,[317,437] PM6-D3H4X,[408,468,469] and OM2-D3(BJ),[406,474] and the results are listed in Table 9.3. Well-performing density functional methods, such as PBE0-D3(BJ)/def2-TZVP,[104,111,133,134,162,163] yield a mean unsigned relative deviation (MURD) and standard relative deviation (SRD) of less than 0.5%.[77,110,504] The error obtained for semiempirical methods is naturally larger. The performance of GFN-xTB is very good, owing partly to the fact that a few systems of this set were part of the training set. The MURD of GFN-xTB is the lowest one (1.1%), even outperforming HF-3c. There is no large systematic shift of the bond lengths with a mean relative deviation (MRD) of only 0.4%. DFTB3-D3(BJ) gives a slightly lower SRD than GFN-xTB at the expense of systematically elongated bonds. This behavior has been discussed before[77,110] and is shared by DFTB, (semi-)local density functionals, and HF-3c. The NDDO-based methods PM6-D3H4X, MSINDO-D3(BJ)H+, and OM2-D3(BJ) show MURDs and SRDs $\geq$ 2.5% and give less reliable geometries compared to DFTB3-D3(BJ), HF-3c, and GFN-xTB.

The LB12[110] (a set of 12 molecules each containing a single long bond between two atoms) and HMGB11[110] (11 heavy main group bond) sets are sensitive to the performance for heavier elements and bonding features that were not included in the training set. The molecules in each set can be found in Ref. 110 (see Supporting Information for detailed results). We restrict the comparison to GFN-xTB, PM6-D3H4X, MSINDO-D3(BJ)H+, and HF-3c, which have available parameters for all elements. The semiempirical methods underestimate the bond lengths in both LB12 and HMGB11 (see Table 9.3) sets. GFN-xTB produces the smallest mean absolute deviation (MAD) in both sets (MAD=13.2 pm and 3.0 pm) surpassing PM6-D3H4X, MSINDO-D3(BJ)H+, and even HF-3c. The GFN-xTB mean signed deviation (MSD) is close to zero in the HMGB11 set, and the bond lengths in the LB12 set are on average underestimated by 10.9 pm. The errors are systematic because the standard deviation (SD) is the lowest of all methods (SD=18.8 pm and 3.9 pm). HF-3c performs similarly and only slightly worse than GFN-xTB (MADs and SDs are larger by 1–3 pm and 1–5 pm). PM6-D3H4X gives two outliers in the LB12 set, where the metal-metal bond dissociates during optimization, and these values are excluded from the statistical analysis.

Overall, GFN-xTB performs similarly or even better in all three sets compared to the significantly more expensive HF-3c approach. GFN-xTB furthermore has the advantage of being applicable to transition metals, which we demonstrate in the following subsection.

Transition Metal Complexes

The computation of transition metal and organometallic complex geometries is a challenge for both single reference QM and SQM methods. As a benchmark for 3d-transition metal systems, we use a set of 32 complexes with 50 bond distances compiled by Bühl and Kabrede, dubbed TMC32.[506] The complexes are: $Sc(acac)_3$, $TiCl_4$, $TiMeCl_3$, $TiMe_2Cl_2$, $Ti(BH_4)_3$, VOF_3, VF_5, $VOCl_3$, $V(NMe_2)_4$, $V(Cp)(CO)_4$, CrO_2F_2, CrO_2Cl_2, $CrO_2(NO_3)_2$, $Cr(C_6H_6)_2$, $Cr(C_6H_6)(CO)_3$,

Table 9.4.: Comparison of calculated and experimental ground state equilibrium bond distances R_e (in pm) for 3d-Transition metal complexes (TMC32)[a].

	GFN-xTB	PM6-D3H4X	MSINDO-D3(BJ)H+
MSD:	−2.7	−1.0	−12.7
MAD:	5.1	6.6	12.7
SD:	5.9	10.7	7.4
MAX:	16.5	44.8	28.3

[a] The set is taken from Ref. 506.
MSD=mean signed deviation, MAD=mean absolute deviation,
SD=standard deviation, MAX=maximum absolute deviation.

$Cr(O)_4$, MnO_3F, $MnCp(CO)_3$, $Fe(CO)_5$, $Fe(CO)_3$, $Fe(CO)_2(NO)_2$, $FeCp_2$, $Fe(C_2H_4)(CO)_4$, $Fe(C_5Me_5)$, $CoH(CO)_4$, $Co(CO)_3(NO)$, $Ni(CO)_4$, $Ni(acac)_2$, $Ni(PF_3)_4$, $CuCH_3$, CuCN, and $Cu(acac)_2$. OM2-D3(BJ) and DFTB3-D3(BJ) are excluded in the comparison because of missing parameters, and HF-3c is not considered either due to SCF convergence problems.

Table 9.4 contains the statistical data on the TMC32 set obtained with GFN-xTB, PM6-D3H4X, and MSINDO-D3(BJ)H+. All methods have a negative MSD indicating underestimated bond lengths. The MSDs and MADs of PM6-D3H4X and GFN-xTB are much smaller in magnitude compared to MSINDO-D3(BJ)H+. GFN-xTB gives on average slightly shorter bonds with more systematic deviations than PM6-D3H4X. This good performance of GFN-xTB is not due to the finite temperature Fermi smearing because orbital occupations remain integral.

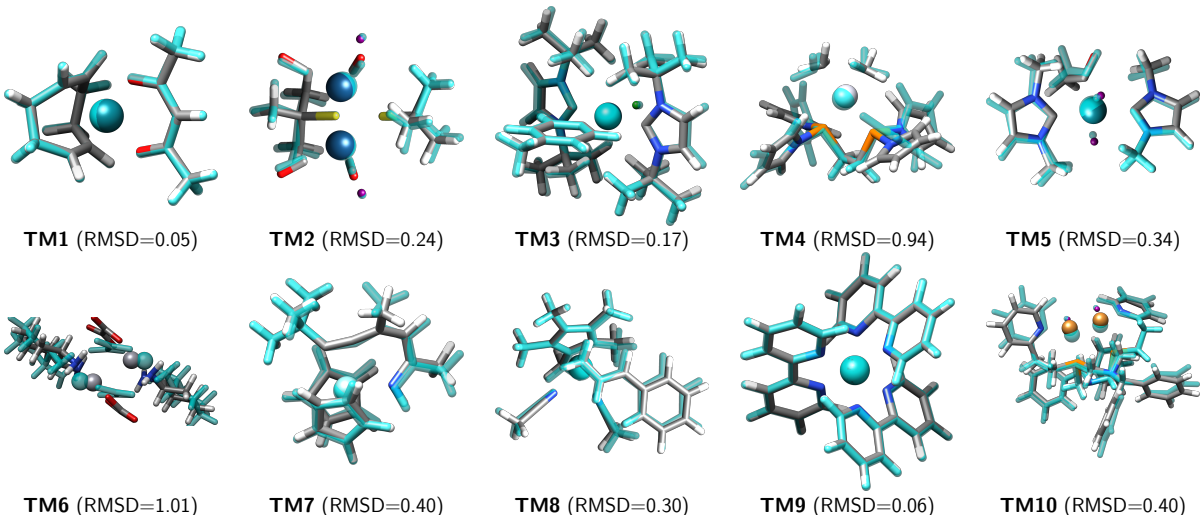


Figure 9.1.: Overlay of 10 larger transition metal complex[49] structures optimized with PBEh-3c (transparent cyan) and GFN-xTB (color coded atoms). The root-mean-square-deviation (RMSD) for an all-atom best fit is given in Å.

The structures of 10 transition metal complexes[49] computed with GFN-xTB and PBEh-

3c[110] as reasonably accurate DFT reference are overlaid in Figure 9.1. The agreement between both methods is very good with root-mean-square-deviations (RMSDs) ranging between 0.05–0.40 Å. Two structures (TM4 and TM6) show larger RMSDs of about 1 Å . In TM4, this is due to rotated phenyl groups. TM6 is a flexible binuclear complex with separate, non-covalently interacting ligands. The hydrogen bonds between the oxalate and diaminocyclohexane make this a challenging case. PM6-D3H4X shows smaller RMSDs for TM4 and TM6 (0.51 and 0.54 Å), but yields larger RMSDs than GFN-xTB for all the other complexes (see Supporting Information). Notable cases are TM5 and TM10. Here the RMSDs with PM6-D3H4X are larger by a factor of four and five, respectively, compared to GFN-xTB. In all complexes, the overall structure of the transition metal complexes is reproduced by GFN-xTB without significant reorganization or dissociation.

Together with the TMC32 set (Table 9.4), this shows that GFN-xTB is able to capture the bonding situations in organometallic systems rather reasonably and significantly better than NDDO-type methods without being specifically modified for such systems. The Fermi smearing technique (see Eq. 9.6) also allows the treatment of systems with small HOMO-LUMO gaps, which are problematic for standard single reference treatments. More detailed tests on transition metal complexes will follow in a separate paper. We also discuss a few cases where static electron correlation is important in Section 9.3.4.

9.3.2. Non-Covalent Interactions

Non-Covalent Interaction Energies

A number of test sets to assess the performance of quantum chemistry methods for London dispersion interactions,[130,507] which are missing in mean-field methods like HF and semilocal DFT, have been developed in the past decade. Here, we apply uniformly all SQM methods in combination with the atom-pairwise D3(BJ)[133,134] correction (D3 with zero damping[133] for PM6-D3H4X) for London dispersion interactions. The SQM methods, however, differ in the way they account for the electrostatic, Pauli exchange repulsion, and induction interactions, which contribute together to the total non-covalent interaction. Different approaches to improve the treatment of hydrogen bonds are also applied. PM6-D3H4X and MSINDO-D3(BJ)H+ use a force field-type hydrogen bonding potential, and DFTB3 and GFN-xTB incorporate modifications in the electronic Hamiltonian, such as modified Coulomb interaction for hydrogen atoms in DFTB3 and additional 2s function for hydrogen in GFN-xTB.

The S22[492] set consists of 22 non-covalently bound, small size complexes and serves as a measure for the accuracy of quantum chemistry methods to describe non-covalent interactions that are dominated by London dispersion and hydrogen bonding. We use the most recent reference values for association energies.[493]

Well-performing, dispersion-corrected DFT methods typically give MADs of $\leq$0.5 kcal mol^{-1}.[130] HF-3c and PM6-D3H4X provide similar accuracies with no significant systematic deviations (Table 9.5). Most of the other studied SQM methods yield MADs close to 1.0 kcal mol^{-1}.

Table 9.5.: Association energies of non-covalently bound complexes. The sets are S22[a] and S12L[b], as well as sets with emphasis on hydrogen bonding, such as the WATER27[e], C15[f], and PL24[g]. Deviations of association energies are given in kcal mol^{-1}.

	GFN-xTB	PM6-D3H4X	MSINDO-D3(BJ)H+	OM2-D3(BJ)	DFTB3-D3(BJ)	HF-3c
			S22[a]			
MSD:	1.3	0.3	0.4	0.8	0.7	−0.1
MAD:	1.3	0.6	1.6	0.9	1.0	0.5
SD:	1.1	0.7	2.0	1.2	1.5	0.7
MAX:	3.6	2.1	3.6	3.9	5.0	1.4
			S12L[b]			
MSD:	−2.5	−6.0	−7.9	−4.4[c]	−3.6[d]	−3.8
MAD:	7.5	6.8	8.6	4.5[c]	4.7[d]	5.3
SD:	9.4	8.5	7.5	4.3[c]	5.4[d]	5.5
MAX:	23.2	21.8	19.1	12.8[c]	13.0[d]	10.2
			WATER27[e]			
MSD:	−0.4	0.1	−4.0	0.2	0.1	−4.4
MAD:	1.4	0.9	4.0	0.9	0.8	4.4
SD:	1.9	1.3	3.3	1.2	1.0	4.3
MAX:	6.1	4.2	12.1	4.3	1.9	14.8
			C15[f]			
MSD:	2.1	0.7	0.2	1.8	3.9	−3.7
MAD:	2.1	1.3	1.8	1.8	4.1	3.7
SD:	1.5	1.5	2.3	1.2	3.5	2.5
MAX:	5.2	3.6	4.0	4.6	12.5	8.3
			PL24[g]			
MSD:	−1.7	−7.9	−4.6	-	−1.0	1.2
MAD:	3.5	8.2	5.5	-	3.0	1.7
SD:	4.5	7.8	6.1	-	3.8	1.9
MAX:	12.5	23.2	16.7	-	8.8	5.3

[a] Reference data taken from Ref. 493.

[b] Reference structures are taken from Ref. 42, while the reference energies are taken from Refs. 130,139.

[c] Excluding systems 3b and 7a due to missing parameters.

[d] Excluding system 7a due to missing parameters.

[e] Reference energies taken from Refs. 508 and 509. Energies are given per water (hydronium/hydroxide) molecule.

[f] Charged H-bonded dimers. Reference data taken from Ref. 468.

[g] 24 protein-ligand interactions with reference data taken from Ref. 510.

This holds also for GFN-xTB where the deviation indicates a systematic underbinding tendency. GFN-xTB shows the largest relative deviations for π-stacked structures (see Supporting Information), such as benzene·cyanide and ethene dimer, which results from the monopole approximation and the poor description of benzene and ethene quadrupole moments. The NDDO-based methods PM6-D3H4X and OM2-D3(BJ), and also HF-3c, perform much better for these systems because of the more accurate electrostatic interactions. GFN-xTB describes rather well the hydrogen bonded systems with largest deviation observed for the stacked uracil dimer (3.5 kcal mol^{-1}, 16%) and the best overall performer of GFN-xTB – the formic acid dimer (>0.2 kcal mol^{-1}, >1%). Therefore, the description of hydrogen bonds facilitated by an additional s-basis function on hydrogen performs reasonably well. The S22 (or the related S66[494]) set was used as a fitting set for most of the considered methods and their performance has to be assessed for larger system.[437]

To this end, we use the S12L[42] set, which consists of large, experimentally accessible supramolecular complexes. The magnitudes of the association energies, derived from experimental free association energies, are in the range of 20–35 kcal mol^{-1} for the neutral complexes, and the values for the charged complexes are significantly larger – about 80 kcal mol^{-1} for 6a and 6b, and 133 kcal mol^{-1} for 7a. Complex 7a (see Supporting Information) is excluded from the statistical analysis for DFTB3-D3(BJ) and OM2-D3(BJ) due to missing parameters for iron. The chlorine containing complex 3b is also neglected for OM2-D3(BJ) for the same reason. Given that the S12L energies are close to 5 times larger than the S22 ones, the performance of DFTB3-D3(BJ) and OM2-D3(BJ) with MADs of 4.7 and 4.5 kcal mol^{-1} is consistent throughout the two sets. When the 7a complex is included in the statistics, the average S12L energy becomes 6 to 7 times larger than the S22 energies. Therefore, the performance of the remaining SQM methods is also uniform throughout, except PM6-D3H4X and HF-3c, which show somewhat larger deviations for S12L. GFN-xTB demonstrates the smallest MSD of all methods, which approaches zero if system 7a is excluded from the analysis. The MAD values indicate consistent deviations for the S22 and S12L sets. The relatively larger standard deviation of GFN-xTB is due to complex 7a, which is overstabilized by ≈23 kcal mol^{-1} (14%). If we exclude this system the SD goes down to 7.2 kcal mol^{-1}. The effect of the monopole approximation is less pronounced for the more charge delocalized systems of S12L, and neither HF-3c nor the NDDO methods perform systematically better than GFN-xTB or DFTB3-D3(BJ).

We consider next the association energies of neutral and also charged water clusters. The well established WATER27 set[508] is used for this purpose, including the revised reference energies for the $(H_2O)_{20}$ subset.[509] Due to the varying cluster sizes, we normalize the association energies to the number of water (and hydronium/hydroxide) molecules (see Supporting Information). The electrostatic interactions and hydrogen bonding are very important for these highly polar complexes. MSINDO-D3(BJ)H+ and the computationally more demanding HF-3c approach show relatively large systematic overestimation of the binding energy, as the magnitudes of the MSDs are equal to the MADs throughout. The largest part of these errors result from the subset of charged clusters incorporating either hydronium or hydroxide. In agreement with Ref. 437,

DFTB3-D3(BJ) performs best with an MAD and SD less or equal to one. PM6-D3H4X and OM2-D3(BJ) yield negligibly larger MAD and SD values. The respective values are larger by about 0.5 kcal mol^{-1} for GFN-xTB, which performs worse for the anionic clusters. Excluding these systems from the statistical analysis, GFN-xTB shows similar deviations as PM6-D3H4X and OM2-D3(BJ) on this subset (MAD and SD of $0.6 - 0.7$ kcal mol^{-1}).

The C15 set contains charged hydrogen bonded systems. PM6-D3H4X shows the best overall performance with an MAD of 1.3 kcal mol^{-1} and a relatively small MSD and SD. OM2-D3(BJ) gives the smallest SD but with MSD and MAD equal to 1.8 kcal mol^{-1} is only slightly better than GFN-xTB. GFN-xTB demonstrates a bit larger systematic deviation, MSD and MAD of 2.1 kcal mol^{-1}, and an error spread that is the same as PM6-D3H4X, and thus the method describes the hydrogen bonds of charged systems quite well. Overall, NDDO-type methods perform relatively better than TB methods, especially if hydrogen bonding corrections are used as in PM6-D3H4X and MSINDO-D3(BJ)H+.

The last organic benchmark set listed in Table 9.5 is the protein-ligand binding motif set PL24.[510] HF-3c gives the smallest MAD of 1.7 kcal mol^{-1} and the smallest SD. DFTB3-D3(BJ) ranks second followed by GFN-xTB with an MAD of 3.5 kcal mol^{-1}. The TB methods demonstrate better accuracy than the NDDO methods PM6-D3H4X and MSINDO-D3(BJ). PM6-D3H4X, which performs very well on the hydrogen bonded sets, is worse than MSINDO-D3(BJ)H+ here. The good performance of GFN-xTB on PL24 is very encouraging and justifies further applications of GFN-xTB to biological systems.

We continue the analysis of the SQM methods accuracy for NCI energies with a few more challenging and unconventional sets listed in Table 9.6. OM2-D3(BJ) is not included in the discussion due to missing parameters. The first set is a collection of 3 subsets compiled by Truhlar and coworkers,[511] which contain systems with predominantly dipolar (DI6/04), charge-transfer interactions (CT7/04), and hydrogen-bonded complexes (HB6/04). This benchmark set contains the elements F, Cl, and S, apart from the standard elements H, C, N, and O. GFN-xTB is the front runner here with the smallest MAD and SD of 0.9 and 1.1 kcal mol^{-1}, very close to "chemical accuracy". This is remarkable because the magnitudes of the binding energies range from 1 to 16 kcal mol^{-1}, and no other SQM method tested yields comparable accuracy. For instance, DFTB3-D3(BJ), which ranks second, shows deviations that are larger by 1 kcal mol^{-1}. Both TB methods have no systematic deviations, unlike the NDDO methods and HF-3c.

The X40[512] set contains 40 halogen bonded systems and presents a challenge to the approximate electrostatics and element pair-specific parametrization in SQM methods. GFN-xTB and PM6-D3H4X employ specific correction terms for halogen bonds (see Eq. 9.14), which are parametrized for related systems. This is reflected by the smallest MSDs of 0.7 kcal mol^{-1} of both methods in Table 9.6. MSINDO-D3(BJ)H+, DFTB3-D3(BJ), and HF-3c have MSDs $>$ 1 kcal mol^{-1}. While the MADs of GFN-xTB and PM6-D3H4X are rather similar, the latter gives outliers and a relatively large SD of 2.2 kcal mol^{-1}), which is close to the one of DFTB3-D3(BJ) and HF-3c. GFN-xTB performs best with an MSD, MAD, and SD below 1 kcal mol^{-1}.

Table 9.6.: Association energies of non-covalently bound complexes, including the HB6/04, CT7/04, and DI6/04 sets of Truhlar and coworkers[a], the X40[b] set, a set of metal-organic framework fragment interactions with CO_2 (MOF-CO_2)[d], and a set of rare gas·cucurbituril complexes (RG-CB_x)[e]. Deviations of association energies are given in kcal mol^{-1}.

	GFN-xTB	PM6-D3H4X	MSINDO-D3(BJ)H+	DFTB3-D3(BJ)	HF-3c
			Non-bonded Interactions[a]		
MSD:	0.3	6.7	2.5	0.0	2.4
MAD:	0.9	7.1	2.7	1.8	2.8
SD:	1.1	11.2	2.1	2.4	3.1
MAX:	2.1	31.4	6.0	5.9	9.1
			X40[b]		
MSD:	0.7	0.7	−1.6[c]	−1.1	−1.2
MAD:	0.9	1.1	2.5[c]	1.8	1.9
SD:	0.9	2.2	5.0[c]	2.4	2.3
MAX:	3.7	11.4	21.5[c]	6.9	6.9
			MOF-CO_2[d]		
MSD:	0.5	−0.5	−2.8	-	−1.4
MAD:	0.7	0.8	2.8	-	1.4
SD:	0.7	1.0	2.2	-	1.4
MAX:	1.7	2.7	11.8	-	6.3
			RG-CB_x[e]		
MSD:	−0.1	13.5 (4.1)[f]	-	-	−8.1 (−0.1)[g]
MAD:	0.3	14.3 (5.0)[f]	-	-	9.8 (2.2)[g]
SD:	0.3	21.2 (6.0)[f]	-	-	18.0 (2.8)[g]
MAX:	0.6	61.8 (15.2)[f]	-	-	52.7 (4.5)[g]

[a] This is a collection of subsets: hydrogen bonded (HB6/04), charge-transfer (CT7/04), and dipole interacting (DI6/04) systems. Reference data taken from Ref. 511.

[b] Halogen-bonded systems with reference data taken from Ref. 512.

[c] Excluding iodine containing systems due to missing parameters.

[d] Interactions between CO_2 and metal-organic framework (MOF) fragments. Reference data taken from Ref. 513.

[e] Association energies of rare gas atoms with cucurbituril hosts (CB5 and CB6). Reference structures at the PBEh-3c level and reference energies from counterpoise-corrected DLPNO-CCSD(T)/def2-QZVPP[163,170,175] calculations (data will be published elsewhere).

[f] Data without krypton containing systems.

[g] Data without xenon containing systems.

Thus, shortcomings of the monopole approximation are compensated by the halogen bond correction in GFN-xTB.

MOF-CO_2 is a test set for non-covalent interactions of CO_2 with functional groups of metal-organic frameworks.[513] DFTB3 is not tested due to missing parameters for Li and Zn in the 3OB parametrization.[431,497,498] GFN-xTB performs the best here showing the smallest magnitudes of MSD (together with PM6-D3H4X), MAD, and SD. GFN-xTB shows a slight underbinding tendency supported by a similar MAD and a small spread of errors. PM6-D3H4X ranks second with a larger SD, whereas HF-3c and especially MSINDO-D3(BJ)H+ perform worse.

The last set in Table 9.6 contains unpublished results for the non-covalently bound complexes of rare gas atoms and cucurbituril[x] hosts (x=5,6). The reference energies are obtained by domain-based local pair natural orbital approximated coupled cluster calculations (DLPNO-CCSD(T)/def2-QZVPP),[163,170,175] including counterpoise correction (more details will be published elsewhere). MSINDO and DFTB3 calculations are not discussed due to missing parameters. The performance of GFN-xTB is exceptionally good with a MSD close to zero and MAD and SD of only 0.3 kcal mol^{-1}. HF-3c and PM6-D3H4X provide large MADs around 10 kcal mol^{-1}. The largest outliers result from problems for specific elements, krypton for PM6-D3H4X and xenon for HF-3c. The performance of both methods is improved significantly if we exclude the corresponding complexes from the statistical analysis. The densely packed cucurbituril[5] (CB5) complexes show large errors that indicate systematic errors at short distances for these element-method combinations. While the reasons for the drastic destabilization of the krypton complexes with PM6-D3H4X (see Supporting Information) are unclear, the problems of HF-3c with xenon are attributed to the geometrical counterpoise correction (gCP). Both methods are outperformed by GFN-xTB, even when the problematic cases are excluded.

Finally, we summarize the performance on all test sets with emphasis on the range of applicability and consistent performance of the methods. In Figure 9.2, we plot the mean unsigned relative deviation (MURD) of all methods, which parametrization covers all test sets. Consequently, OM2-D3(BJ) and DFTB3-D3(BJ) are not included because even though they perform well for some sets, they do not have parameters for all elements in the sets. Figure 9.2 indicates that the 3 remaining methods give similar MURDs, less than 30%, for the first 5 sets, these are the sets in Table 9.5. GFN-xTB provides MURDs smaller than 30% for the last four sets as well, which correspond to the sets in Table 9.6. Unlike the consistent accuracy of GFN-xTB for all test sets, PM6-D3H4X and HF-3c demonstrate larger errors for some of the sets and therefore are not consistently accurate for NCI energies.

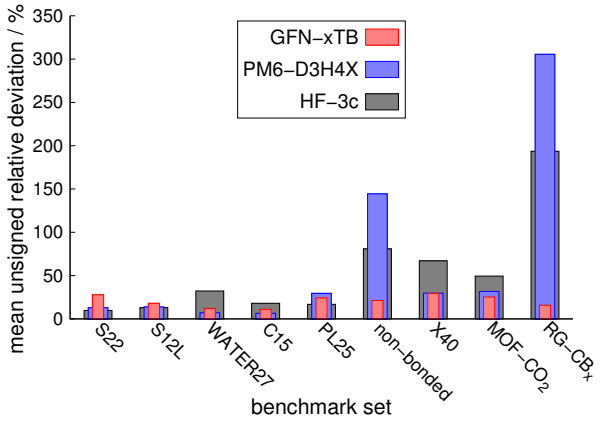


Figure 9.2.: Mean unsigned relative deviations of the semiempirical methods GFN-xTB (red), PM6-D3H4X (blue), and HF-3c (gray) on the non-covalent interaction benchmark sets considered in this work. The deviations are in percent (%).

Structures of Non-Covalently Bound Systems

This section is devoted to the performance of SQM methods for geometries of non-covalently bound systems. This is of practical significance because SQM methods are likely to be used for geometry optimizations followed by single-point calculations at a higher level of theory in a multi-level scheme. Because even HF-3c can become too computationally expensive for such applications, we focus on the performance of the cheaper SQM methods.

We optimize the dimers of S22[492] and give the deviations of the center-of-mass (CMA) distances in Table 9.7. HF-3c gives the smallest deviations, followed by DFTB3-D3(BJ), which underestimates the CMA distances on average, and GFN-xTB and PM6-D3H4X show comparable MADs and SDs. In general, methods that perform well for S22 energies provide reasonable structures, except the outlier OM2-D3(BJ). The vanishing MSD and the small SD of the CMA distances observed for GFN-xTB suggest that this method produces potential energy surfaces resembling the high-level reference curves.

Hobza and coworkers have presented a benchmark set, which contains structures of small isolated peptides with aromatic side chains (P26).[514] The relative stabilities of these structures depend significantly on the intramolecular non-covalent interactions. The average heavy atom RMSD is the best for DFTB3-D3(BJ) and GFN-xTB, followed by HF-3c. PM6-D3H4X and OM2-D3(BJ) show larger $\overline{\text{RMSD}}$s and perform better than MSINDO-D3(BJ)H+. We plot the heavy atom RMSDs of the individual systems for GFN-xTB, PM6-D3H4X, and DFTB3-D3(BJ) in Figure 9.3a (see Supporting Information for system abbreviations and results with the other SQM methods). GFN-xTB produces a few larger RMSDs (>0.2 Å), which arise from deviations in the dihedral angle of the aromatic side chains, the case for system 5 (FGG_470). For most systems, GFN-xTB shows RMSDs that are smaller than the ones of DFTB3-D3(BJ). PM6-D3H4X gives larger RMSD values, including larger outliers compared to the TB models.

Table 9.7.: Comparison of structures of non-covalently bound systems. The center-of-mass (CMA) distances of the S22[a] complexes and the root mean square deviation (RMSD) of the heavy atom positions of the small peptides from P26[b] are compared for fully optimized geometries. The CMA distances of the S66[c] complexes are estimated from cubic spline interpolations of energies computed on the S66x8 set.

	GFN-xTB	PM6-D3H4X	MSINDO-D3(BJ)H+	OM2-D3(BJ)	DFTB3-D3(BJ)	HF-3c
			S22[a] (CMA distance in pm)			
MSD:	−5	3	−25	−17	−11	−2
MAD:	15	14	35	26	14	7
SD:	18	21	47	29	16	11
MAX:	50	58	186	66	45	34
			P26[b] (heavy atom position RMSD of small peptides in pm)			
$\overline{\text{RMSD}}$:	12	30	48	28	11	20
			S66x8[c] (CMA distance in %)			
MRD:	−0.6	0.6	−2.5	−5.0	−2.9	−1.2
MURD:	2.0	2.0	5.3	5.0	3.1	1.2
SRD:	2.7	2.6	6.0	2.7	2.8	1.1
MAXR:	7.0	6.7	14.5	12.4	10.9	4.2

[a] Reference structures taken from Ref. 492.

[b] Reference structures are taken from Ref. 514.

[c] CMA distances determined from cubic spline interpolation (see Ref. 110) based on reference structures and energies from Ref. 494.

The S66x8[494] by Hobza and coworkers has been used to compare non-covalently bound complex equilibrium geometries at high theoretical level. The reference minimum CMA distances are determined by interpolation from the CCSD(T)/CBS energies. The same procedure is applied to the computations with the SQM methods, and the statistical results for the comparison of the equilibrium CMA distances are compiled in Table 9.7 (see Supporting Information for details). HF-3c performs the best with the smallest MURD and SRD and a slight tendency to underestimate the equilibrium distances by 1.2%. GFN-xTB and PM6-D3H4X rank second, both providing distances with MURDs of 2.0% – PM6-D3H4X slightly overestimating, and GFN-xTB slightly underestimating the distances. OM2-D3(BJ), which performs quite well for energies (see Table 9.5), underestimates significantly the distances (−5%) similarly to the performance on S22. Together with the results on ROT34[504] (see Table 9.3), OM2-D3(BJ) does not confer enough accuracy for geometry optimizations in multi-level approaches.

Provided the encouraging performance of GFN-xTB for fast geometry optimizations, we test the method in a particular multi-level scheme. To this end, we re-optimize the PBEh-3c structures of the S30L[139] supramolecular test set with PM6-D3H4X, DFTB3-D3(BJ), and GFN-xTB,

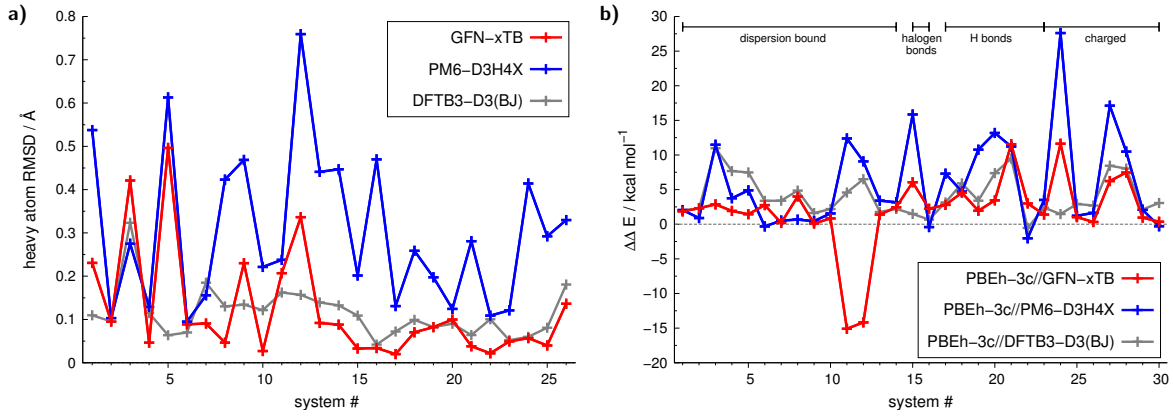


Figure 9.3.: a) Root-mean-square-deviations (RMSD) of heavy atom positions relative to the MP2/aug-cc-pVTZ[59,176,177] reference structures for the P26[514] peptide benchmark set. The RMSDs obtained for geometries optimized with GFN-xTB, PM6-D3H4X, and DFTB3-D3(BJ) are given in red, blue, and gray, respectively. b) Deviations in the association energies for 30 large non-covalent complexes from S30L[139] computed with PBEh-3c on geometries obtained with the semiempirical methods GFN-xTB, PM6-D3H4X, and DFTB3-D3(BJ). The deviations are given in kcal mol^{-1} as the difference $\Delta\Delta E = \Delta E_{\text{PBEh-3c//SQM}} - \Delta E_{\text{PBEh-3c//PBEh-3c}}$.

and after that re-compute single point energies with PBEh-3c. Because the geometries of the molecular fragments often do not differ much between the free and the associated state, deviations of the association energies, $\Delta\Delta E$, reflect changes of the intermolecular PES of SQM methods with respect to PBEh-3c. These deviations are then positive indicating underbinding and effective cancellation of intramolecular contributions. The results are plotted in Figure 9.3b and absolute ΔE values can be found in the Supporting Information. The $\Delta\Delta E$ values are mostly positive, except two negative $\Delta\Delta E$ (systems 11 and 12) for GFN-xTB due to large relaxation of the pentakis(1,4-benzodithiino)-corannulene host upon binding. In fact, the free host with GFN-xTB deviates more from the PBEh-3c structure than the bound state leading to the erroneous stabilization. Apart from this outlier, GFN-xTB gives intermolecular PES that resemble quite well the PBEh-3c surface (excluding systems 11 and 12 results in MRD=MURD=SRD=9.4%). The deviations with PM6-D3H4X are roughly twice as large, and the performance of DFTB3-D3(BJ) is quite good with MRD=13.2%, MURD=13.3%, SRD=11.1%. We note that the results may be different for other higher level methods and we choose PBEh-3c for the comparison because similarly to GFN-xTB, the approach is designed to provide geometries and NCIs that are comparable to the results of dispersion-corrected hybrid-functional calculations in large basis sets. It is also a sound choice because of the computational cost of the method and the large system sizes considered. Possible problematic systems can be straightforwardly diagnosed by comparison of the association energies computed at the SQM geometries with the higher level method (PBEh-3c) and the SQM method (GFN-xTB). The typical GFN-xTB error is $< 30\%$ for NCI energies (Figure 9.2) and the multi-level PBEh-3c//GFN-xTB approach gives errors

below 10% for association energies with respect to a full PBEh-3c//PBEh-3c computation. For system 11 and 12, the ΔE from GFN-xTB//GFN-xTB are -38.8 and -40.5 kcal mol^{-1} but are -54.1 and -53.9 kcal mol^{-1}, respectively, with PBEh-3c//GFN-xTB and can be identified as problematic for the multi-level treatment.

9.3.3. Conformational Energies

SQM methods find applications in the sampling and energetic ranking of conformers, which are relevant to biophysical chemistry, organic supramolecular chemistry, and organic crystal polymorph prediction.[515] We use a number of well established benchmark sets for conformational energetics to gauge the performance of GFN-xTB. ACONF[516] is a set of conformers of butane, propane, and hexane, CYCONF[517] consists of gas phase conformers of cysteine, PCONF[518] contains conformers of a small Phe-Gly-Gly tripeptide, and MCONF[519] has 51 conformers of melatonine. We consider in addition three hairpin conformers of long alkane chains, called the hairpin set,[520] and 46 diuracilphosphate conformers that comprise the UpU46 set[521] as an example for a charged system. The statistical data for all methods are given in Table 9.8, and details are deferred to the Supporting Information.

The ACONF set tests mainly the balance between repulsive Pauli exchange and attractive London dispersion interactions and is characterized by relatively small conformational energies ranging between 0.5 and 5.0 kcal mol^{-1}. The NDDO methods OM2-D3(BJ) and PM6-D3H4X perform very well for this set showing small MADs and nearly no systematic deviation. The worst performer is the MSINDO-D3(BJ)H+ method, which strongly overestimates the conformational energies. The accuracy of GFN-xTB and DFTB3-D3(BJ) is much better than of MSINDO-D3(BJ)H+, somewhat worse than that of the other NDDO approaches, and on par with the much more computationally expensive HF-3c. The TB models underestimate the conformational energies on average with mean signed deviations of -0.7 and -0.8 kcal mol^{-1}, which points to softer short-range repulsion forces. GFN-xTB and DFTB3-D3(BJ), however, capture correctly the energetic ordering of the conformers as evidenced by Figure 9.4a, unlike HF-3c, which exhibits changes in the conformer energetic ordering (see Supporting Information).

CYCONF presents a more difficult test set than ACONF for SQM methods because cysteine conformational energies are sensitive to the description of electrostatic interactions and hydrogen bonds. OM2 lacks parameters for sulfur and is not included in the discussion. The best performing method is HF-3c showing hardly any systematic deviation and a spread less than 1.0 kcal mol^{-1}. GFN-xTB and DFTB3-D3(BJ) perform the best from the group of the more empirical models. Both methods underestimate the conformational energies with GFN-xTB demonstrating slightly better accuracy than DFTB3-D3(BJ). The NDDO methods PM6-D3H4X and MSINDO-D3(BJ)H+ cannot compete with the accuracy of the TB models on this set.

Table 9.8.: Comparison of conformational energies computed with different semiempirical methods relative to high-level reference calculations. The sets are ACONF[a], CYCONF[b], MCONF[c], and PCONF[d] from the GMTKN30 benchmark database, and the UpU46[e] set and a recently introduced set of 3 large, linear alkanes, called hairpin[f]. Deviations in the conformational energies are relative to the lowest energy conformer calculated at the reference level and are given in kcal mol^{-1}.

	GFN-xTB	PM6-D3H4X	MSINDO-D3(BJ)H+	OM2-D3(BJ)	DFTB3-D3(BJ)	HF-3c
			ACONF[a]			
MSD:	−0.7	0.0	1.2	0.1	−0.8	−0.9
MAD:	0.7	0.5	1.2	0.2	0.8	0.9
SD:	0.4	0.6	0.6	0.3	0.5	0.4
MAX:	1.4	1.3	2.5	0.8	1.8	1.8
			CYCONF[b]			
MSD:	−1.2	−2.8	−4.4	-	−1.7	−0.3
MAD:	1.4	3.3	4.7	-	1.7	0.7
SD:	1.3	2.7	3.5	-	1.1	0.7
MAX:	2.7	5.5	8.0	-	2.9	1.1
			MCONF[c]			
MSD:	−1.6	−1.2	−2.6	−2.5	−1.1	−0.3
MAD:	1.6	1.5	2.8	2.6	1.8	0.9
SD:	0.9	1.2	2.3	1.8	1.9	1.1
MAX:	2.7	3.2	7.0	5.4	4.1	2.2
			PCONF[d]			
MSD:	2.5	2.5	5.6	0.9	0.6	2.6
MAD:	2.5	2.6	5.7	1.1	1.2	2.6
SD:	1.9	1.8	3.1	0.9	1.4	1.1
MAX:	5.9	6.0	10.0	2.0	2.8	3.6
			UpU46[e]			
MSD:	−0.1	2.2	1.9	-	−0.8	2.7
MAD:	1.4	2.6	2.3	-	1.4	3.0
SD:	1.8	2.3	2.0	-	1.6	2.1
MAX:	5.6	7.0	6.0	-	6.4	6.1
			hairpin[f]			
MSD:	−0.2	0.6	7.4	0.0	−2.4	−2.9
MAD:	0.2	0.6	7.4	0.2	2.4	2.9
SD:	0.2	0.2	0.6	0.2	0.3	0.3
MAX:	0.4	0.7	8.1	0.2	2.6	3.1

[a] Reference data taken from Ref. 516. [b] Reference data taken from Ref. 517.

[c] Reference data taken from Ref. 519. [d] Reference data taken from Ref. 518.

[e] Reference data taken from Ref. 521. [f] Energetic difference of the folded and linear, *all-anti* conformer.[520]

MSD=mean signed deviation, MAD=mean absolute deviation, SD=standard deviation,

MAX=maximum absolute deviation.

MCONF tests similar interactions as CYCONF that have a stronger dispersion component due to the indole residue of melatonine, and lacks 3rd row elements. HF-3c retains the very good accuracy shown with CYCONF. GFN-xTB, PM6-D3H4X, and DFTB3-D3(BJ) compete closely for the second best position, whereas OM2-D3(BJ) ranks this time with the worst performer, which is MSINDO-D3(BJ)H+. The three second best performers underestimate the melatonine conformational energies on average, and GFN-xTB shows somewhat larger (MAD = -1.6 kcal mol^{-1}) but more systematic deviations (SD=0.9 kcal mol^{-1}) in the group.

The best performing methods on the tripeptide PCONF[518] set are OM2-D3(BJ) and DFTB3-D3(BJ) with MADs close to 1 kcal mol^{-1}. HF-3c, PM6-D3H4X, and GFN-xTB rank at the second place, each of them giving an MAD of about 2.5 kcal mol^{-1}. HF-3c, however, shows a smaller standard deviation by 1 kcal mol^{-1} than PM6-D3H4X and GFN-xTB. Because the conformational energies of this set range from 0 to 2.5 kcal mol^{-1}, the overall accuracy of all SQM methods is insufficient, which makes this set a challenging test for approximate quantum chemical methods.

The UpU46[521] set contains charged molecules and reference energies that vary from 0.5 to 16.5 kcal mol^{-1}. DFTB3-D3(BJ) and GFN-xTB perform very well for this set with MADs of only 1.4 kcal mol^{-1} and MSDs of -0.8 kcal mol^{-1} and -0.1 kcal mol^{-1}. PM6-D3H4X and HF-3c are among the worst performers here together with MSINDO-D3(BJ)H+. They systematically overestimate the conformational energies and show a large spread of 2.6 and 3.0 kcal mol^{-1}. A more detailed analysis of the individual conformational energies in Figure 9.4b shows that conformer 36, which contains a hydrogen bond between a phosphate oxygen and a ribose hydroxyl group, presents a difficult case. Despite performing similarly to GFN-xTB on average, DFTB3-D3(BJ) predicts different energetic ordering of the conformers exhibiting an alternative energetic minimum. This solidifies the leading position of GFN-xTB, which reproduces the trend of conformational energies with a very small MSD and reasonable MAD.

The hairpin[520] set is an extension of ACONF to much larger alkanes. The NDDO methods OM2-D3(BJ) and PM6-D3H4X are among the front runners again. The performance of the TB models, however, shows differences. The accuracy of GFN-xTB becomes better than PM6-D3H4X and approaches the one of OM2-D3(BJ), and the accuracy of DFTB3-D3(BJ) deteriorates. HF-3c similarly to DFTB3-D3(BJ) underestimates the relative conformational energies of the folded, hairpin structures by about 2–3 kcal mol^{-1}. MSINDO-D3(BJ)H+ shows large systematic errors (>7 kcal mol^{-1}). The small standard deviations of all methods, consistent with the results on the ACONF set, show that the observed deviations have systematic character and are due to imbalance of short-range repulsion and long-range dispersion interactions.

In summary, GFN-xTB performs rather well reproducing the trends of the conformational energies and providing reasonable absolute estimates. A possible application of GFN-xTB is therefore the conformational sampling of large flexible systems as part of a multi-level scheme for structure prediction.

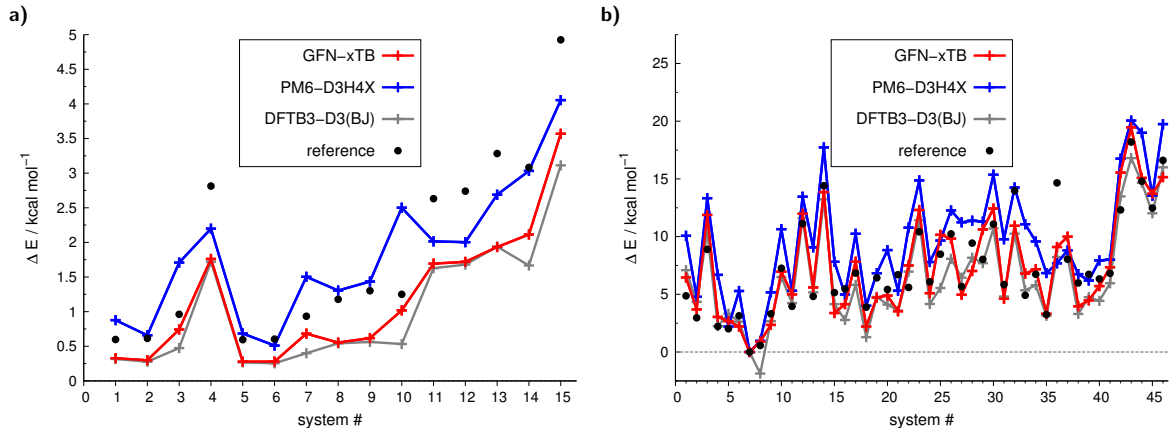


Figure 9.4.: a) Conformational energies of the alkane conformer (ACONF) set computed with GFN-xTB (red), PM6-D3H4X (blue), and DFTB3-D3(BJ) (gray). The reference energies from Ref. 516 are given as black dots. b) Conformational energies of the diuracilphosphate (UpU46) set computed with GFN-xTB (red), PM6-D3H4X (blue), and DFTB3-D3(BJ) (gray). The reference energies from Ref. 521 are given as black dots.

9.3.4. Potential Energy Curves, Thermochemistry, and Other Properties

In the preceding sections, we demonstrated the excellent performance of GFN-xTB for a broad range of systems for the target properties, molecular structures and non-covalent interaction energies. In this section, we discuss the performance of GFN-xTB for potential energy surfaces and their characteristics, such as covalent bond energies, harmonic vibrational frequencies, and reaction barriers.

We compare the accuracy of GFN-xTB for harmonic frequencies using the systems in the ROT34,[77,504] LB12, and HMGB11[110] sets, for which we computed the zero point vibrational energies E_{ZPVE} and free energy (translational, rotational, and vibrational) contributions G_{TRV} from the nuclear degrees of freedom.

The agreement of GFN-xTB with the PBEh-3c reference is excellent for both E_{ZPVE} and G_{TRV} (see Figure 9.5). GFN-xTB, similar to GGAs, does not require any scaling of the harmonic frequencies to obtain good agreement with PBEh-3c, which requires scaling by 0.95. These results suggest that GFN-xTB can be used in multi-level schemes for the computation of harmonic frequencies.

We tested GFN-xTB on the ISO34[522] and ISOL22[523] sets, which contain isomerization energies of organic molecules, yielding MADs of 6.5 and 9.8 kcal mol^{-1}. The performance for ISOL22 is comparable to the dispersion-corrected GGAs BLYP-D3(BJ)[99,100] and B97-D3(BJ)[103] and inferior to dispersion-corrected hybrid functionals whose MADs are 2–3 times smaller.[73,110] The performance of GFN-xTB for ISO34 is worse compared to dispersion-corrected GGAs with an MAD around 2 kcal mol^{-1}.[73] In passing, we mention that PM6-D3H4X and DFTB3-D3(BJ) yield smaller MADs of 7.0 and 8.0 kcal mol^{-1}, respectively, for ISO34 and 5.1 and 3.4 kcal mol^{-1}

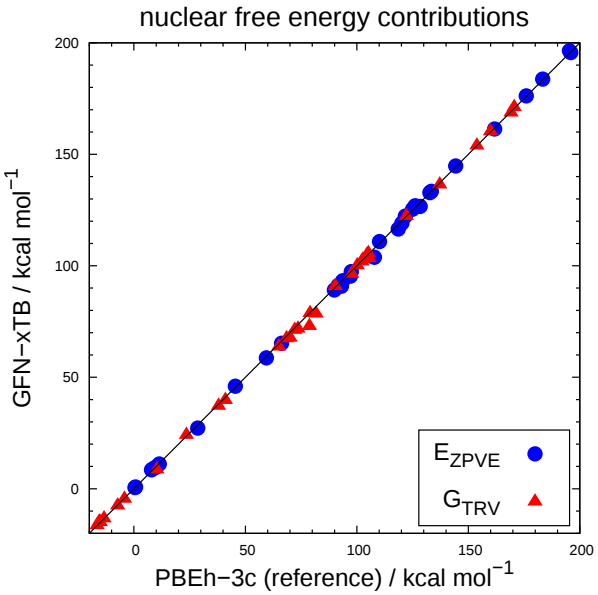


Figure 9.5.: Comparison of zero point vibrational energies E_{ZPVE} and free energy contributions G_{TRV} from the translational, rotational, and vibrational degrees of freedom computed with GFN-xTB and PBEh-3c for the ROT34,[77,504] LB12, and HMGB11[110] systems (G_{TRV} at $T = 298.15\,\mathrm{K}$ according to Ref. 42). All values are given in kcal mol^{-1} and harmonic frequencies from PBEh-3c are scaled by a factor of 0.95.[110] For a more condensed representation, the values for the three LB12 systems 2-(1-diamantyl)[121]tetramantane (DIAD), the FLP, and in,in-bis(hydrosilane) (BHS), which are significantly larger than 200 kcal mol^{-1}, are not shown here and given in the Supporting Information.

for ISOL22 (see Supporting Information for detailed results also of other SQM methods). Thus, these methods, which were parametrized also to energies, perform better for bond energies than GFN-xTB, but are still less accurate than dispersion-corrected density functionals. GFN-xTB provides good molecular geometries but in general does not yield very accurate (covalent) thermochemistry and is not recommended for routine reaction energy calculations. However, in our opinion, this holds for SQM methods in general, because accurate molecular energies are difficult to obtain even at sophisticated DFT or WFT levels, and the effect of semiempirical approximations is most severe for covalent bond energies. Hence, SQM methods should rather be used for rough estimates of covalent thermochemistry or as initial part of multi-level, large-scale screening applications. Note that this is different for non-covalent interaction energies, which are of reasonable to good quality as shown above.

For the purpose of standard molecular thermochemistry estimates, we employ a semiempirical protocol for the computation of standard heats of formation at 298 K (HOF). It introduces element-specific parameters to correct the computed atomization energy E_{at} for offsets in the

atomic GFN-xTB energy levels of atom A, i.e.,

$$E'_{at} = E_{at}(\text{GFN-xTB}) + \sum_{A}^{atoms} \delta E_{el}^{A} + CN_A \; f_{el}^{A} \; . \tag{9.16}$$

Here δE_{el}^{A} is an element-specific correction for atom A and f_{el}^{A} is an additional element-specific parameter, which scales a contribution from hybridization, as taken into account via the coordination number CN_A. The value of E'_{at} is then used as usual to derive the HOF at 298 K by adding the molecular thermostatistical enthalpy including zero point vibrational contributions. The empirical parameters δE_{el}^{A} and f_{el}^{A} are determined by fitting E'_{at} to the G3/99 set[524] of reference HOF. We note in passing that also for DFTB, the use of additional element-specific fit parameters for the calculation of heats of formation has been presented,[462] while the conventional NDDO-based SQM methods do not apply such corrections.[525] Including only the CN_A independent correction leads to an MAD for the G3 set of 12.2 kcal mol^{-1}, while the addition of the CN_A dependent term reduces the deviation to 9.8 kcal mol^{-1}. This approach is cross-validated on an independent HOF benchmark set taken from Table 3 of Ref. 526 (only neutral systems included) and compared to PM6 data (see Figure 9.6).

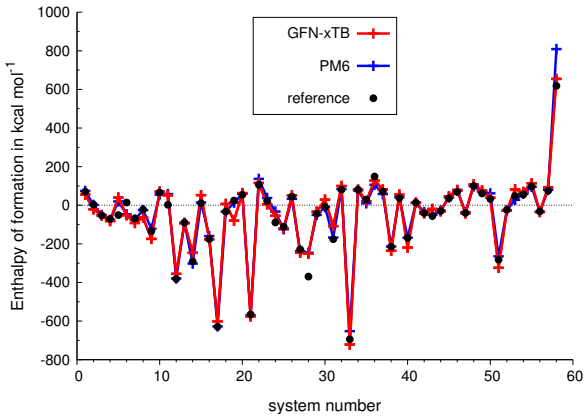


Figure 9.6.: Enthalpies of formation for "difficult" molecules (including elements up to chlorine) computed with GFN-xTB and PM6 along with experimental values (see Ref. 526). The detailed values are given in the Supporting Information*.

Both methods employ fully optimized structures. The GFN-xTB approach to HOF is further simplified by computing not the true GFN-xTB Hessian for the thermostatistical terms but employing Lindh's model Hessian,[487] which is also employed in the optimizer of the xtb code (see above), for this purpose. This approximation introduces additional errors of only about 1 kcal mol^{-1} on average for the G3/99 set of molecules, and it may be beneficial for large-scale screening applications (our code also allows usage of the GFN-xTB Hessian if this is preferred). As can be seen from Figure 9.6, GFN-xTB yields HOFs for this rather difficult set of molecules, which are of comparable quality (MAD=24.9 kcal mol^{-1}) as the ones from PM6 (MAD=17.2 kcal mol^{-1}). Though not designed for that purpose, the corrected GFN-

xTB scheme can thus be used to obtain reasonable estimates for molecular thermochemistry. For example, the MADs obtained with GFN-xTB on the ISO34 and ISOL22 sets decrease by slightly more than 2 kcal mol^{-1}, if the atomic correction in Eq. 9.16 is included.

We study the behavior of GFN-xTB for covalent bonding to gain further insight. In Figure 9.7, we give the potential energy curves computed with GFN-xTB for the diatomic molecules H_2, N_2, F_2, and LiH where we also depict the positions of the minima computed at higher electronic structure level. The data demonstrate that GFN-xTB describes very well bond lengths in agreement with the results in Section 9.3.1 but the method overestimates significantly the covalent bond energies. In general it seems very difficult in a SQM context to fit simultaneously covalent bond lengths and bond energies by varying the fit weights of the properties because focusing the fit on bonding energies resulted in overestimated bond lengths and vice versa. Related observations have been recently reported for DFT energies and densities in Ref. 527. This correlation, however, does not affect the good non-covalent energies because they depend weakly on the H_0 part of the TB Hamiltonian.

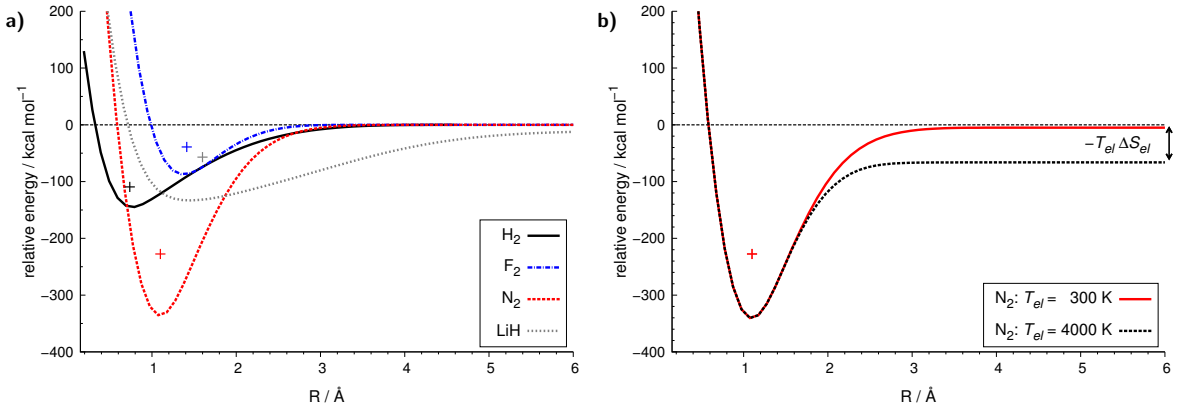


Figure 9.7.: a) Potential energy curves for the dissociation of H_2, F_2, N_2, and LiH computed with GFN-xTB (T_{el} = 300 K). The points mark the position of the minimum obtained from high-level calculations for F_2,[528] N_2,[528] and LiH.[529] The H_2 reference point is computed at the FCI/aug-cc-pV6Z//MP2/aug-cc-pV6Z level of theory (this work). The energies are given relative to the free atoms (S=3/2 for nitrogen, S=1/2 for the others). b) The potential energy curve for N_2 computed at two different electronic temperatures T_{el} = 300 K and T_{el} = 4000 K. The difference in relative energy from zero at $R \to \infty$ results from the difference in the electronic entropy ΔS_{el} for dissociated N_2 (S=0) and free atoms (S=3/2).

As is shown in Figure 9.7a, GFN-xTB is able to dissociate molecules correctly. This is made possible by the finite temperature electron smearing approach, which mimics the incorporation of static correlation in an independent electron framework. The value of the electronic temperature controls the amount of electron smearing between the occupied and virtual orbital spaces, and we use it as a parameter. The curves in Figure 9.7a are plotted with T_{el} = 300 K, and a N_2 curve with a much larger electronic temperature of T_{el} = 4000 K is also shown in Figure 9.7b.

It follows from the comparison in Figure 9.7b that Fermi smearing leaves the equilibrium part of the dissociation curve unaltered. In fact, Fermi smearing dissociates N_2 to N atoms with the correct value of spin, S= 0, and the difference in the dissociation asymptote is completely accounted for by the change in electronic entropy with temperature, $-T_{el}\Delta S_{el}$, with respect to the free atoms, each with spin S=3/2. We include $-T_{el}\Delta S_{el}$ in the electronic energy because it enforces the correct spin and spatial symmetry and fractional occupations along the curve. Note that this only holds for a restricted treatment with identical spatial orbitals for the two spin parts. The downside is, however, that GFN-xTB in the present, restricted version cannot properly distinguish different spin states (see Ref. 477 for a unrestricted TB treatment). For $T_{el} > 0\,\mathrm{K}$ and degenerate HOMO/LUMO levels, a low spin configuration is always more stable (larger S_{el}) than a high spin one.

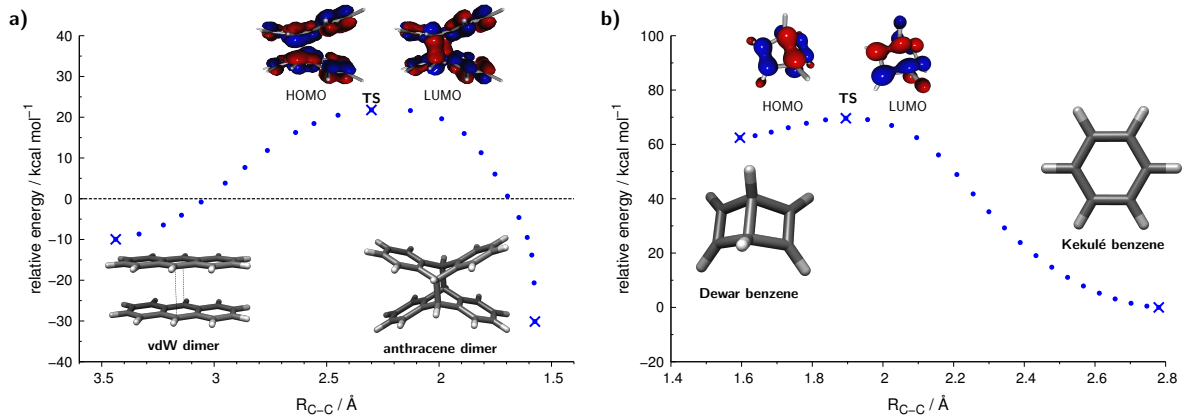


Figure 9.8.: a) Potential energy curve for the anthracene dimer formation (T_{el} = 4000 K). The energy is given relative to the free anthracene molecules. The frontier molecular orbitals are plotted at the transition state (TS) geometry as well. b) Potential energy curve for the conversion of Dewar benzene to Kekulé benzene (T_{el} = 14000 K). The frontier molecular orbitals are plotted at the estimated transition state (TS) geometry.

The use of Fermi smearing reduces barriers of chemical processes, especially with methods that underestimate the HOMO-LUMO gap such as GGA DFT functionals and TB methods. GFN-xTB yields HOMO-LUMO gaps that are only slightly smaller (by 0.2 eV) than the ones from GGA DFT calculations (see Table A7.26 in the Supporting Information). The use of the finite temperature technique allows also the investigation of reaction pathways before turning to computationally demanding multi-reference methods and provides stable scans of the PES without convergence problems during the SCC step. Two examples of orbital symmetry forbidden reaction pathways are given in Figure 9.8. GFN-xTB gives a smooth PES for the thermally activated (formally forbidden) dimerization of anthracene (Figure 9.8a) with a transition state region characterized by fractionally occupied HOMO and LUMO orbitals. For the benzene reac-

tion, the HOMO-LUMO gap is larger, and higher temperature needs to be applied to populate the relevant levels and to obtain a smooth surface around the transition state geometry.

Since GFN-xTB overestimates bond energies, the method may, however, provide reasonable reaction energetics at larger electronic temperatures, which effectively lowers the bond energies (see Figure 9.7b). This will facilitate the implementation of GFN-xTB in QCEIMS[530] for the simulation of electron impact mass spectra in the future.

Finally, the dissociation potential energy curve of the ion pair [BMIM][PF_6], which is a typical ionic liquid, is shown in Figure 9.9. GFN-xTB produces an intermolecular potential energy curve that agrees very well with the high-level coupled cluster reference suggesting use of the approach in molecular dynamics studies of ionic liquids.

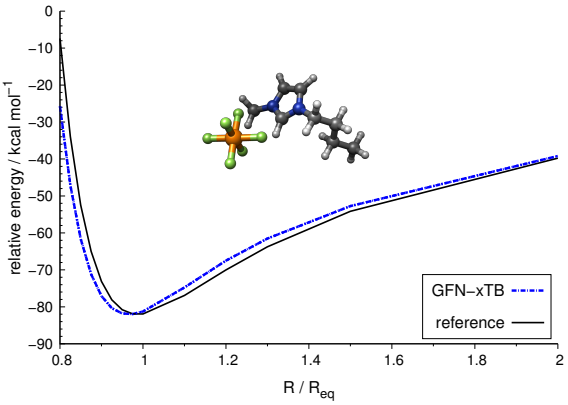


Figure 9.9.: Intermolecular potential energy curve for the dissociation of the [BMIM][PF_6] ion pair computed with GFN-xTB (T_{el} = 300 K) on PBEh-3c geometries. The reference curve is taken from counterpoise corrected DLPNO-CCSD(T)/TightPNO/aug-cc-pVQZ single-point calculations.

9.3.5. Known Problems

SQM methods compromise accuracy and computational speed. In a TB method, the sources of error mainly are self-interaction error, monopole description of the electrostatic interactions, AO basis set deficiencies, and parametrization errors. Similarly to GGA DFT functionals, TB methods tend to overestimate the delocalization of electrons and produce too small orbital energy gaps. In turn, they implicitly include static correlation effects and hence work well for metallic systems and the dissociation of covalent bonds. At variance from GGAs, the damped Coulomb law (Eq. 9.4) provides an asymptotically correct (exchange-correlation) potential. The monopole approximation is a weak part of the approach, and the errors originating from it are difficult to estimate. From our experience, the accuracy diminishes the more polar and more directional the electrostatic interactions are. A good example are the halogen bonds, which require an atom pairwise correction for reasonable performance. Empirical evidence seems to indicate that strong directional electrostatic effects are somewhat emulated by “covalent” interactions in

the H_0 part, which is reflected by relatively large sp-shell occupations in the transition metal complexes compared to *first principles* DFT results.

Approximate electrostatic interactions and/or small AO basis sets may lead sometimes to qualitatively incorrect molecular geometries. Extensive experience with GFN-xTB in our lab has singled out the following problematic systems for the new method. Cyclobutane has a folded, D_{2d} symmetric, non-planar structure (C-C-C-C torsional angle of about 20°), which becomes planar in GFN-xTB and virtually all SQM methods tested. The related and chemically more relevant case of cyclopentane with a C_s envelope structure is less problematic. The ring C-C-C-C torsion angle, which is around 40° at PBEh-3c (or PBE0-D3/def2-TZVP) level, is described well by OM2-D3(BJ) (36.5°), GFN-xTB (35.5°), and DFTB3-D3(BJ) (34.1°), while MSINDO-D3(BJ)H+ and PM6-D3H4X give geometries closer to a planar structure (25.7° and 27.5°). Polyynes are common building blocks in supramolecular structures and their C-C-C bending potentials determine the shape persistence of these complexes.[531] For 1,3-butadiyne, all methods provide reasonable C-C-C bending potentials with harmonic frequencies around 200 cm^{-1}, close to the PBEh-3c value of 240 cm^{-1}. Problems appear with almost all TB- and NDDO-type methods for C_2F_2 (or C_2OMe_2) yielding bent instead of linear minima. The error is more or less absent for PM6-D3H4X and less pronounced for GFN-xTB, which provides linear HC_2F (DFTB3-D3(BJ): bent) and a smaller bending angle for C_2F_2 (21.5° instead of 36.4° for DFTB3-D3(BJ)). This problem can only be repaired by parametrization at the expense of sacrificing accuracy for other hydrocarbons. Another typical example of self-interaction error is C_2 symmetric cis-1,3-butadiene, which has a C-C-C-C torsion angle of 35° at the PBE0-D3(BJ)/def2-TZVP level. This diminishes to 32° at the GGA (PBE) level and essentially to zero with GFN-xTB and DFTB3-D3(BJ). The OM2-D3(BJ), MSINDO-D3(BJ)H+, and PM6-D3H4X methods which are closer to HF in character give correctly the non-planar structure with reasonable torsion angles (29°, 40°, and 46°, respectively).

Related to the self-interaction error problem in (meta-)GGAs and TB methods is the vanishing gap obtained from gas phase calculations of large polynucleotide, polypeptide chains, or zwitter ions. This way, the electronic structure of these biologically very important systems is as badly described as in (semi-)local DFT.[532,533] In GFN-xTB this problem is less severe when the GBSA implicit solvation model[490] is applied. For example, without GBSA, GFN-xTB produces a HOMO-LUMO gap of 0.7 eV for the α-helical $Ace(Ala)_{19}NMe$ oligopeptide (20 peptide units, geometry taken from Ref. 362). Similar observations have been reported for GGAs.[533] GFN-xTB including implicit GBSA(water) solvation yields a much more realistic gap of 4.5 eV at no significant additional cost. We thus recommend to always use the GBSA solvation model for solvated organic and biomolecular systems and in particular as a workaround for zwitterionic systems where the SCC is otherwise not convergent. We demonstrate the applicability of GFN-xTB/GBSA for a larger, realistic case in the next section. The theory as well as more detailed tests on the GBSA approach will be presented elsewhere.[490]

9.3.6. Performance for Large Systems

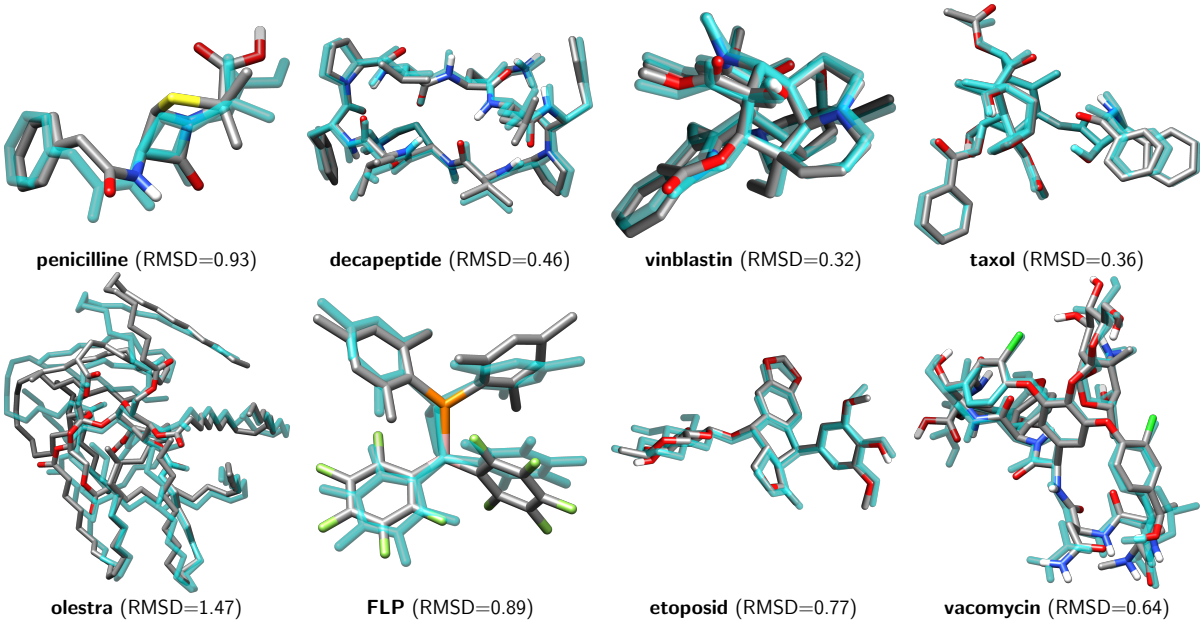


Figure 9.10.: Overlay of eight moderately sized, organic molecules. The structures are optimized with PBEh-3c (transparent cyan) and GFN-xTB (color coded atoms) The root-mean-square-deviation (RMSD) for an all atom best fit is given in Å. Hydrogen atoms bonded to carbon are not depicted for clarity.

In the previous sections, we discussed the accuracy of the GFN-xTB method for standard benchmark sets containing many small cases, and in the present section, we gauge the speed and applicability of the method and its implementation in the `xtb` code to more extended chemical systems.

We compare GFN-xTB and PBEh-3c structures from a set of relatively large and flexible organic molecules taken from Ref. 49 in Figure 9.10. To facilitate the analysis, we provide the all-atom position RMSD below each structure overlay in the same figure. GFN-xTB closely resembles the PBEh-3c structures for the more rigid molecules like taxol and the compounds with significant stabilization from hydrogen bonds like the decapeptide. We see larger deviations for flexible structures, where the PES for dihedral distortions is flat, such as the frustrated Lewis pair (FLP) with rotating aryl substituents, and olestra, which comprises long aliphatic fatty acid chains. Olestra, in particular, contains 453 atoms, which makes DFT calculations very time-consuming; one optimization step with the efficient PBEh-3c method takes about 20 minutes on 16 CPUs, whereas the full GFN-xTB optimization comprising 454 steps takes only 14 minutes on 4 CPUs. The GFN-xTB geometries are very reasonable and can be used for single-point calculations or refinements at higher levels of theory.

Systems with several hundreds or thousands of atoms are ubiquitous in biochemistry and supramolecular chemistry. To demonstrate the applicability of GFN-xTB to such systems, we

consider the optimization of a large dodecanuclear organometallic polyhedron $[Pd_{12}\mathbf{L}_{24}]^{24+}$ comprising 1644 atoms,[302] where **L** stands for a chiral (*P*)-6,6'-dipyridinyl-2,2'-bis(methoxymethoxy)-1,1'-binaphthyl ligand. We have previously computed the electronic circular dichroism spectrum of this molecule using only the *O*-symmetric minimum geometry, for which we could afford a DFT optimization with the TURBOMOLE program suite.[164c,165,371] Here we show the energy during the optimization in Figure 9.11 together with an overlay of the final GFN-xTB and initial HF-3c structures. The number of optimization steps is 133 pointing to a fast convergence for a system with more than 1600 atoms. Despite the flexibility of the ligands, the all-atom position RMSD between the GFN-xTB and HF-3c geometry is only 0.74 Å.

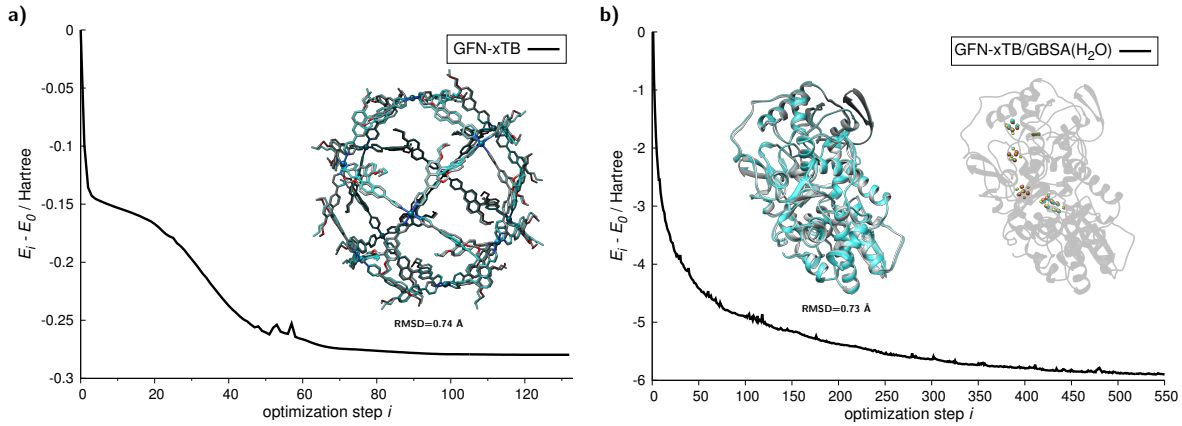


Figure 9.11.: a) Energy along a full GFN-xTB optimization (133 steps) of the large $[Pd_{12}\mathbf{L}_{24}]^{24+}$ complex starting from the *O*-symmetric HF-3c structure. An overlay of the final geometry is shown (hydrogen atoms are omitted for clarity) together with the HF-3c structure (cyan) and the all atom position RMSD. b) Energy along a GFN-xTB/GBSA(H_2O) optimization (550 steps) of the *Clostridium pasteurianum* [FeFe]-hydrogenase (PDB ID: 3C8Y) with loose convergence thresholds.[534,535] The entire protein contains 8951 atoms (charge: −23). An overlay of the X-ray (cyan) and the optimized GFN-xTB geometry is shown for the protein backbone (left) and the organometallic centers (right). The heavy atom position RMSD is given below the left overlay.

Finally, we apply GFN-xTB in combination with the GBSA(H_2O) solvation model to optimize an entire [FeFe]-hydrogenase protein complex, which comprises close to 9000 atoms. We took the initial structure (3C8Y) from the protein data bank (PDB),[534,536] removed crystal water molecules, and added hydrogen atoms[537] using the Maestro 11[538] program suite. We performed a combined GFN-xTB/GBSA(H_2O) geometry optimization with loose convergence thresholds for this protein complex. For this protein, which comprises a charge of −23, this optimization with 550 steps took two months on 32 CPUs. We provide a plot of the energy during the optimization along with an overlay of the initial and final structures in Figure 9.11b. The RMSD of the heavy atom positions of the obtained structure and the starting X-ray structure is only 0.73 Å, which is an indication that the ground state PES of this system is reasonably well described

with GFN-xTB/GBSA(H_2O). The embedded Fe_xS_x (x=2,4) clusters complicate the quantum chemical description of the hydrogenase due a vanishing HOMO-LUMO gap and demand finite temperature Fermi smearing. The overlay of the inorganic centers in Figure 9.11b affirms the robustness and consistency of this treatment and the code. Thus, GFN-xTB/GBSA(H_2O) successfully captures the electronic structure of both the organometallic sites and the organic polypeptide.

9.4. Conclusion

We developed a robust and broadly applicable semiempirical quantum chemical method for the computation of structures, vibrational frequencies, and non-covalent interactions for molecules across the periodic table. It is connected to an efficient implicit solvation model (GBSA), which we intend to publish in the near future, enabling the simulation of bulk electrostatic screening and molecular surface effects occurring in condensed phases. The size of the typical target system is around 1000 atoms, and initial tests show that the method can be successfully applied to proteins with 3000 atoms or more. Although the focus is on bio/organic systems, an advantageous characteristic of the method is the robust and consistent treatment of inorganic main group and transition metal complexes. This opens new avenues to the computation of large metallo-protein structures and exploration of the dynamics of large transition metal complexes. We envision numerous applications in the future, including and not limited to studies of protein structure, protein-protein and protein-ligand interactions, and supramolecular complexes ranging from esoteric molecular machines to giant shape-persistent nano-architectures.

A key premise of the present method is its special purpose character. In our view, low-cost semiempirical QM methods cannot describe simultaneously very different chemical properties, such as structures and chemical reaction energies, and the GFN-xTB method (as the name conveys) focuses on structural properties. The efficiently computed structures and vibrational frequencies or the conformers obtained from global search procedures can (and should) be used subsequently for more accurate DFT or WFT refinements. We hope that the method can serve as a general tool in quantum chemistry and in particular recommend GFN-xTB optimized structures (and thermostatistical corrections) in a multi-level scheme together with PBEh-3c single-point energies. Large-scale molecular dynamics, screening of huge molecular spaces (libraries), parametrization of force-fields, or providing input for novel machine learning techniques are obvious other fields of application. A strength of the new approach is that the Hamiltonian contains physically interpretable, element-wise parameters, which can straightforwardly be adjusted to other properties or specific substance classes. Our own current efforts include a reparametrization for ionization potential and electron affinity calculations for electrochemical problems and the use of GFN-xTB electronic structure information in an intermolecular potential. We believe that the new method can help bridge different molecular scales by making possible the consistent computation of reasonable structures and non-covalent interactions for medium-sized ($\leq$100 atoms) and large chemical systems (up to 10000 atoms) at the same level

of theory.

Acknowledgments

This work was supported by the DFG in the framework of the "Gottfried-Wilhelm-Leibniz" prize. P. S. would also like to kindly acknowledge the financial support of the Alexander von Humboldt foundation for a postdoctoral fellowship. The authors thank V. Ásgeirsson, C. Bauer, P. Burger, M. Bursch, M. P. Checinski, S. Ehrlich, and J. Seibert for many initial tests of the method.

Part V.

Final Summary and Conclusion

In this thesis, different low-cost electronic structure methods for the fast and still accurate computation of electronic spectra were developed and tested. Particular attention was paid to the simulation of electronic circular dichroism (ECD) spectra.

Since the accurate simulation of such spectra does not only require the computation of excited states, but also the proper knowledge about the relevant molecular conformers, a previously presented thermochemical protocol[42] for the computation of free energies in solution has been applied to study the formation of two supramolecular frustrated Lewis pair (FLP) complexes. The employed electronic structure method was of particular focus. For the complex of $B(C_6F_5)_3$ and $PMes_3$, comparison of the computed and experimentally determined association free energies[†] revealed the satisfactory performance of dispersion corrected hybrid density functionals. Though the study was not aimed at computing ECD spectra, the challenge to describe their association correctly in solution is very similar to the problem, which needs to be solved if the spectra of supramolecular aggregates, such as chiral sensor materials, are to be simulated. Furthermore, the final ranking of a conformer ensemble in solution can be performed with the same protocol. Taking into account the performance on established benchmark sets,[73] PW6B95-D3[112,133,134] and B3LYP-NL[106,107,172] together with a quadruple-ζ basis set can be recommended. In particular, the fast basis set convergence of these mean field approaches in combination with the fast and accurate D3[133,134] dispersion model[136] works well and allows their "black box" use, while the in principle more accurate correlated wave function methods are hampered by the significantly worse scaling of the computation time with the system size and persisting errors due to the use of finite basis sets.

Taking into account other related studies,[42,73,139,141] the largest uncertainty for these neutral systems is attributed to the solvation model. For the studied association processes, which involves a large reduction in the solvent-accessible surface area (SASA), the error in the solvation free energy contribution is estimated to be 2–3 kcal mol^{-1} by comparison of the two implicit solvation models SMD[189] and the standard 2012 parametrization of COSMO-RS.[188,539] This corresponds to an error of about 10–20% compared to the corresponding electronic energy contribution ΔE. If different arrangements or conformers are to be compared, the SASA changes only slightly and errors due to the solvation model rarely exceed 1 kcal mol^{-1} for the FLP systems considered. Though there definitely is room for improvement, the existing procedure should be sufficient for the conformer ranking of neutral systems, provided a reliable electronic structure method is used for the electronic energy contribution. Charged systems are known to be more problematic,[139,140] and further developments will be necessary to improve their description.

The dominant part of this thesis has been dedicated to the cost-efficient calculation of the spectra themselves, i.e., the computation of a large number of excited states along with the respective ECD intensities. In this context, the recently developed simplified Tamm-Dancoff approximated time-dependent density functional theory (sTDA-DFT) scheme[46] was successfully applied in a collaborative work with the Lützen group from the University of Bonn to compute

[†]These values were obtained in a different study from nuclear magnetic resonance (NMR) measurements.[45]

the ECD spectra of large palladium(II) metallosupramolecular spheres, each comprising 822 and 1644 atoms, respectively. These spectra provided important support for the formation of the supposed structures. Furthermore, the sTDA-DFT gives rise to the computation of ECD spectra for these large systems, which would not be feasible, e.g., with regular time-dependent density functional theory (TD-DFT).

In line with the approximations employed in the sTDA-DFT scheme, a simplified variant of TD-DFT, termed sTD-DFT, was proposed. While the density functional response is entirely neglected, the two-electron integrals in the excited state calculation are then simplified by Coulomb interactions between transition-density monopoles that are damped at short range. The semiempirical damping[288–290] introduces one element-specific and three functional-dependent empirical parameters, which are identical in sTDA-DFT and sTD-DFT. Like sTDA-DFT,[46,195] it can in principle be combined with any kind of global or range-separated hybrid functional. Furthermore, the matrix dimensions are kept small in both treatments by considering only the relevant singly excited configurations up to a user-specified threshold. As a consequence, the excited state treatment in sTD-DFT is faster by approximately three orders of magnitude compared to TD-DFT, but at the same time, the former provides a significantly larger number of excited states. As for sTDA-DFT, the ground state self-consistent field procedure is now the computational bottleneck in the sTD-DFT procedure. Since the simplified and original TD-DFT approaches share the same non-simplified DFT ground state reference, a one-to-one comparison is possible to exclusively determine the effect of the simplified excited state treatment. It was found that the accuracy of vertical excitation energies and also of the computed ECD intensities are very similar to the original methods, thus the introduced errors are negligible for most type of valence excitations. Hence, the presented development has enabled the routine calculation of ECD spectra for arbitrary molecular systems in the entire UV-Vis spectral range, which coincides with the computation of several thousand excited states.

The ECD of the recently synthesized [16]helicene and a di-substituted triisopropylsilyloxy (TIPSO) derivative[350] was studied by sTD-DFT. In this particular study, the efficiency of the sTD-DFT approach was exploited to compute spectra on 100 “snapshots” extracted from a classical molecular dynamics (MD) simulation. Non-minimum structure effects could be effectively included by averaging the snapshot spectra without resorting to the cumbersome task of explicitly computing the nuclear vibrational contributions of each electronic excitation. While this study served rather as a proof-of-principle, the presented treatment may be relevant for flexible systems, in particular if non-minimum structure effects are important.[367,540]

Though the sTD-DFT marks a significant speed-up compared to TD-DFT, the excited state treatment is slower than in sTDA-DFT by about a factor of 10. Tamm-Dancoff approximation (TDA) methods like sTDA-DFT are, however, not gauge-invariant, which erroneously results in different numerical values if rotatory strengths are computed in the dipole length or in the dipole velocity formalism.[47] This has practical implications, because only the latter form guarantees origin-independent results for the rotatory strengths. Like regular TD-DFT, sTD-DFT does not suffer from this deficiency. By exploiting the efficiency of the sTDA-DFT and sTD-DFT

methods, the effect of the TDA for the ECD was assessed for some larger systems, including the chiral C_{76} fullerene and an α-helical oligopeptide. This study revealed a tremendous failure of TDA methods in the computation of ECD spectra of delocalized π-systems (e.g., C_{76}). While the ECD spectra computed in the dipole length form are in good agreement with the ECD spectra from sTD-DFT[‡], the origin-independent dipole velocity formalism yields spectra of opposite sign, suggesting the presence of the incorrect enantiomer upon comparison with experimental data. The error could be ascribed to the magnitude of the two-electron integrals in the **B** matrix which is neglected in TDA methods. For the oligopeptide, the corresponding integrals are small and the TDA methods work well.

In order to accelerate the computation of absorption and ECD spectra even further, the rate-determining DFT ground state treatment has been replaced by a newly developed, purpose-specific semiempirical extended tight-binding (xTB) treatment. This method is similar to the well-established density functional tight-binding (DFTB),[411] but only employs global and element-specific parameters and does not perform a full self-consistent field treatment. By including diffuse functions to the typically employed minimal basis set, the method is capable of providing molecular orbital (MO) input that resembles the one from a hybrid DFT treatment, but at a fraction of the computational time. The orbitals from this xTB treatment then enter the mostly unchanged simplified excited state treatment, as if they were obtained from a Kohn-Sham DFT calculation. While the self-consistent field treatment for the aforementioned [16]helicene takes more than one hour at the BHLYP/def2-SV(P)[69,163,300] level, the newly developed xTB treatment takes only two seconds. Hence, the excited state procedure is now the rate-determining step and combination with the much faster simplified TDA is preferable. In this context, the so-called A+B/2 correction of TDA eigenvectors has been developed, which corrects the aforementioned failures of TDA methods at negligible costs. This way, the newly developed sTDA-xTB scheme can be applied for the ultra-fast calculation of ECD, as well as electronic absorption spectra. The accuracy of sTDA-xTB for excitation energies is comparable (for valence states) or even better (for Rydberg states) compared to a TDA-DFT treatment based on a global hybrid functional with a double-ζ basis set and clearly surpassing other semiempirical quantum mechanical (SQM) methods. The necessity of applying shifts to the computed spectrum (typically <0.5 eV) to achieve agreement with the experiment is not considered to be a significant problem. The responsible energetic discrepancy typically originates from neglected solvation or vibrational effects. Furthermore, shifting is also necessary and common practice in applications of regular TDA-DFT and TD-DFT. Moreover, an experimental absorption spectrum is typically available, which should be used to determine the shift. Compared to DFT based methods, sTDA-xTB does not suffer from the well-known charge-transfer problem and for this reason, is expected to provide more consistent results if applied to extended systems. This should be investigated in the future, with the focus on a consistent performance in the computation of excited states (and spectra) of small and extended systems.

[‡]sTD-DFT provides excellent mutual agreement for the rotatory strengths in the dipole length and dipole velocity form.

Another purpose-specific extended tight-binding scheme has been developed in analogy to the one mentioned above. Here, a partially polarized minimal basis set is combined with a self-consistent field procedure (for simple gradients) including third order charge fluctuations.[456] This GFN-xTB approach is parametrized to provide good molecular geometries, harmonic vibrational frequencies, and non-covalent interaction energies. Assessment for these target properties shows that the method is as accurate or more accurate than other competing SQM methods. Due to the availability of parameters for practically the entire periodic table of elements ($Z \leq 86$) along with the very low cost of this method, it can be used to sample the conformational space of almost arbitrary molecules with up to a few hundred atoms. The availability of a separately developed generalized Born implicit solvation model with SASA terms (GBSA)[490] furthermore provides the possibility to carry out such calculations in solution. In addition, the method was shown to provide ro-vibrational contributions to the free energies in good agreement to the ones computed with the significantly more elaborate PBEh-3c[110] approach. Hence, its usage for this purpose in a thermochemical protocol (as the one applied to the FLP systems) is obvious.

Another option is to apply GFN-xTB for MD simulations and directly compute spectra on the snapshots with sTDA-xTB (cf. the aforementioned sTD-DFT//MD study for the [16]helicene). The immense potential offered by this synergy of the newly developed GFN-xTB and sTDA-xTB approaches is demonstrated in Figure 9.12. For the large $[Pd_6\{(M)\text{-}\mathbf{1}\}_{12}]^{12+}$ complex (synthesized in the Lützen group), which has been studied in Chapter 4, the agreement of the simulated and experimental ECD spectra improves significantly if the sTDA-xTB computation is carried out on MD snapshots and not restricted to the single O-symmetric geometry.

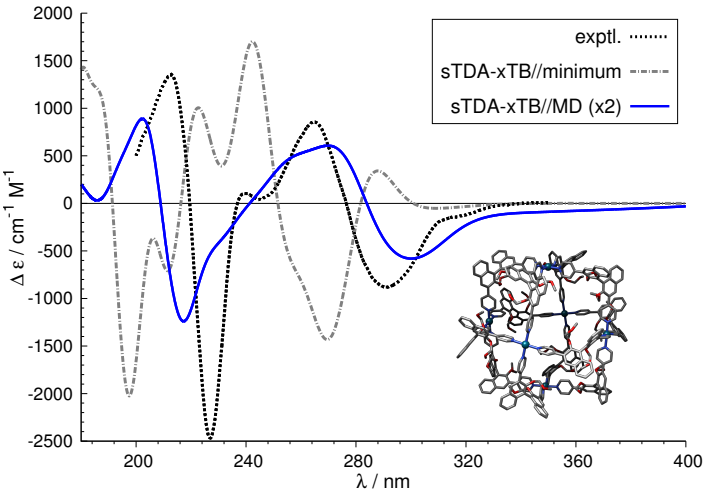


Figure 9.12.: ECD spectra of $[Pd_6\{(M)\text{-}\mathbf{1}\}_{12}]^{12+}$ (see Chapter 4) computed with the newly developed sTDA-xTB method (see Chapter 8) on a single O-symmetric minimum structure and on snapshots from an MD simulation employing the GFN-xTB method (presented in Chapter 9). See Appendix A8 for details.

While large parts of the necessary architecture is now available, further testing will certainly be required in the future to determine an optimal multi-level procedure that is able to find and correctly weight the relevant conformers. This is a very challenging task as different methods need to be employed ranging from classical force-fields over to SQM and *ab initio* QM methods. In addition, solvation effects need to be included, ideally on equal footing, at all levels of theory. Due to the non-parallel energy surfaces of the different methods, there remains the risk of erroneously discarding relevant conformers. This may occur whenever a method is applied in a single-point fashion (typically a high-level method) on structures that have been optimized at some other level of theory. In this regard, the problems are quite similar to the ones faced in the field of crystal structure prediction.[515,541] In Section 9.3.2, non-covalent interaction energies were computed by the PBEh-3c[110] composite scheme at its own potential energy surface (PES) minimum and at the PES minima obtained with SQM methods. Apart from two easily identifiable and well-understood outliers, employing a GFN-xTB structure introduces errors on the order of 10%. Hence, those deviations, which are simply due to slightly different geometries, need to be taken into account as well, whenever conformers are to be re-ranked by a higher-level method. Obviously, the insights obtained from the above mentioned FLP association study and the development of the GFN-xTB model are relevant not only for simulating ECD spectra, but are certainly applicable in general for thermodynamic studies of supramolecular complexes and possibly even for more quantitative spectroscopy techniques such as nuclear magnetic resonance (NMR). In fact, NMR spectroscopy can provide useful experimental reference data, such as conformer populations or equilibrium constants for the formation of supramolecular complexes,[12] which can then serve as benchmarks for entire multi-level schemes in the future, starting with the Lewis structure of the respective compound as input.

All in all, the methodology developed in this thesis sets the stage for routine calculation of electronic absorption and in particular ECD spectra. The sTDA-xTB approach marks an almost optimal procedure for the computation of broad range spectra. In critical situations, cross-checks or refinement by sTD-DFT calculations should still be able in many cases. Since the former behaves much like its parental approach TD-DFT, the experience obtained from well-known benchmarks[34] for the latter in combination with different density functionals is mostly transferable to sTD-DFT.

At this point, we should recall the yet unresolved challenges in the computation of ECD spectra, as stated initially (i.e., point 2 and 3 following the ideas of Schellman):[32]

2. Feasible electronic structure methods that provide the QM quantity (i.e., excitation energies and rotatory strengths) need to be developed.

3. ECD signals are highly sensitive to the three-dimensional structure of the molecule. Hence, the latter needs to be computed accurately. Possibly, different conformers have to be considered and the individual ECD spectra need to be averaged.

Having the highly efficient and still reliable sTDA-xTB and sTD-DFT methods at hand, it seems fair to say that for systems with less than 2000 atoms, limitations due to the excited

state treatment, e.g., due the computational cost or restricted sets of parameters, could be removed. Consequently, the second point has essentially been solved and for a majority of rigid compounds[§], assignment of the absolute configuration by the ECD spectrum is already within reach. Due to the extremely low computational cost of the newly developed methods, in particular of sTDA-xTB, even considering many different conformers does not represent an obstacle anymore.

Though the third point mentioned above remains an open issue for many chemical systems, bright prospects to treat more flexible and supramolecular systems also exist here. Since conformers can now be sampled by the fast and robust GFN-xTB scheme (including solvation) and subsequently ranked employing established thermochemistry protocols[42] in combination with dispersion-corrected DFT, the first steps towards a fully automatic ECD procedure are already taken.

[§]This holds, if the compound shows sufficiently large ECD intensities in the UV-Vis spectral range. Furthermore, the ground state wave function needs to be sufficiently well described by a single reference determinant (as in TD-DFT).

Bibliography

[1] Houk, K. N.; Liu, F. *Acc. Chem. Res.* **2017**, *50*, 539–543.

[2] Thiel, W. *Angew. Chem. Int. Ed.* **2011**, *50*, 9216–9217; *Angew. Chem.* **2011**, *123*, 9382–9384.

[3] Cavalleri, M. *Int. J. Quantum Chem.* **2013**, *113*, 1.

[4] Rostamikia, G.; Janik, M. J. *Energy Environ. Sci.* **2010**, *3*, 1262–1274.

[5] Jain, A.; Shin, Y.; Persson, K. A. *Nat. Rev. Mater.* **2016**, *1*, 15004.

[6] Baik, M.-H.; Friesner, R. A. *J. Phys. Chem. A* **2002**, *106*, 7407–7412.

[7] Cernak, T. A.; Gleason, J. L. *J. Org. Chem.* **2008**, *73*, 102–110.

[8] Waldrop, M. M. *Nature* **2016**, *530*, 144–147.

[9] Hohenberg, P.; Kohn, W. *Phys. Rev.* **1964**, *136*, B864–B871.

[10] Garbuzynskiy, S. O.; Melnik, B. S.; Lobanov, M. Y.; Finkelstein, A. V.; Galzitskaya, O. V. *Proteins* **2005**, *60*, 139–147.

[11] Warnke, I.; Furche, F. *WIREs Comput. Mol. Sci.* **2012**, *2*, 150–166.

[12] Berova, N.; Di Bari, L.; Pescitelli, G. *Chem. Soc. Rev.* **2007**, *36*, 914–931.

[13] Smith, S. W. *Toxicol. Sci.* **2009**, *110*, 4–30.

[14] Crawford, T. D.; Tam, M. C.; Abrams, M. L. *J. Phys. Chem. A* **2007**, *111*, 12057–12068.

[15] Grimme, S. In *Reviews in Computational Chemistry*; Lipkowitz, K. B., Boyd, D. B., Eds.; Wiley-VCH: New York, 2004; Vol. 20; pp 153–218.

[16] Autschbach, J. *Chirality* **2009**, *21*, E116–E152.

[17] Autschbach, J.; Nitsch-Velasquez, L.; Rudolph, M. *Top. Curr. Chem.* **2011**, *298*, 1–98.

[18] Pescitelli, G.; Di Bari, L.; Berova, N. *Chem. Soc. Rev.* **2014**, *43*, 5211–5233.

[19] Pescitelli, G.; Bruhn, T. *Chirality* **2016**, *28*, 466–474.

[20] Woody, R. W. In *Comprehensive Chiroptical Spectroscopy, Volume 2: Applications in Stereochemical analysis of Synthetic Compounds*; N. Berova, K. N., P. L. Polavarapu, Woody, R. W., Eds.; Wiley: New York, 2012; pp 475–497.

[21] Ranjbar, B.; Gill, P. *Chem. Biol. Drug Des.* **2009**, *74*, 101–120.

[22] Miles, A. J.; Wallace, B. A. *Chem. Soc. Rev.* **2016**, *45*, 4859–4872.

[23] Kelly, S. M.; Jess, T. J.; Price, N. C. *Biochim. Biophys. Acta* **2005**, *1751*, 119–139.

[24] Wallace, B.; Janes, R. W. *Biochem. Soc. Trans.* **2010**, *38*, 861–873.

[25] Riley, M. L.; Wallace, B. A.; Flitsch, S. L.; Booth, P. J. *Biochemistry* **1997**, *36*, 192–196.

[26] Wallace, B. A.; Janes, R. W. *Curr. Opin. Chem. Biol.* **2001**, *5*, 567–571.

[27] Hoffmann, A.; Kane, A.; Nettels, D.; Hertzog, D. E.; Baumgärtel, P.; Lengefeld, J.; Reichardt, G.; Horsley, D. A.; Seckler, R.; Bakajin, O.; Schuler, B. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 105–110.

[28] Allenmark, S. *Chirality* **2003**, *15*, 409–422.

[29] Wolf, C.; Bentley, K. W. *Chem. Soc. Rev.* **2013**, *42*, 5408–5424.

[30] Biedermann, F.; Nau, W. M. *Angew. Chem. Int. Ed.* **2014**, *53*, 5694–5699; *Angew. Chem.* **2014**, *126*, 5802–5807.

[31] Labuta, J.; Hill, J. P.; Ishihara, S.; Hanyková, L.; Ariga, K. *Acc. Chem. Res.* **2015**, *48*, 521–529.

[32] Schellman, J. A. *Chem. Rev.* **1975**, *75*, 323–331.

[33] Rosenfeld, L. *Z. Phys.* **1929**, *52*, 161–174.

[34] Laurent, A. D.; Jacquemin, D. *Int. J. Quantum Chem.* **2013**, *113*, 2019–2039.

[35] Jacquemin, D.; Wathelet, V.; Perpète, E. A.; Adamo, C. *J. Chem. Theory Comput.* **2009**, *5*, 2420–2435.

[36] Bredow, T.; Jug, K. *Theor. Chem. Acc.* **2005**, *113*, 1–14.

[37] Thiel, W. *WIREs Comput. Mol. Sci.* **2014**, *4*, 145–157.

[38] Christensen, A. S.; Kubař, T.; Cui, Q.; Elstner, M. *Chem. Rev.* **2016**, *116*, 5301–5337.

[39] Kamachi, T.; Yoshizawa, K. *J. Chem. Inf. Model.* **2016**, *56*, 347–353.

[40] Wang, W.; Donini, O.; Reyes, C. M.; Kollman, P. A. *Annu. Rev. Biophys. Biomol. Struct.* **2001**, *30*, 211–243.

[41] Mackerell, A. D. *J. Comput. Chem.* **2004**, *25*, 1584–1604.

[42] Grimme, S. *Chem. Eur. J.* **2012**, *18*, 9955–9964.

[43] Welch, G. C.; San Juan, R. R.; Masuda, J. D.; Stephan, D. W. *Science* **2006**, *314*, 1124–1126.

[44] Stephan, D. W.; Erker, G. *Angew. Chem. Int. Ed.* **2010**, *49*, 46–76; *Angew. Chem.* **2010**, *122*, 50–81.

[45] Rocchigiani, L.; Ciancaleoni, G.; Zuccaccia, C.; Macchioni, A. *J. Am. Chem. Soc.* **2014**, *136*, 112–115.

[46] Grimme, S. *J. Chem. Phys.* **2013**, *138*, 244104.

[47] Bouman, T. D.; Hansen, A. E. *J. Chem. Phys.* **1977**, *66*, 3460–3467.

[48] Neese, F. *WIREs Comput. Mol. Sci.* **2012**, *2*, 73–78.

[49] Grimme, S. *J. Chem. Theory Comput.* **2014**, *10*, 4497–4514.

[50] Jensen, F. *Introduction to Computational Chemistry*; Wiley: Chichester, 207; Vol. 2.

[51] Tomasi, J.; Mennucci, B.; Cammi, R. *Chem. Rev.* **2005**, *105*, 2999–3093.

[52] Szabo, A.; Ostlund, N. S. *Modern Quantum Chemistry*; Dover Publications: New York, 1996.

[53] Schwerdtfeger, P. *ChemPhysChem* **2011**, *12*, 3143–3155.

[54] Bohm, D. *Quantum Theory*; Dover Publications: New York, 1989.

[55] Helgaker, T.; Jørgensen, P.; Olsen, J. *Molecular Electronic-Structure Theory*; J. Wiley: New York, 2000.

[56] Stone, A. J. *The Theory of Intermolecular Forces: Second Edition*; Oxford University Press: Oxford, 2013.

[57] Hartree, D. R. *Math. Proc. Cambridge* **1928**, *24*, 89–110.

[58] Fock, V. *Z. Physik* **1930**, *61*, 126–148.

[59] Møller, C.; Plesset, M. S. *Phys. Rev.* **1934**, *46*, 618–622.

[60] Bartlett, R. J. *Ann. Rev. Phys. Chem.* **1981**, *32*, 359–401.

[61] Perdew, J. P.; Schmidt, K. *AIP Conf. Proc.* **2001**, *577*, 1–20.

[62] Runge, E.; Gross, E. K. U. *Phys. Rev. Lett.* **1984**, *52*, 997–1000.

[63] Kohn, W.; Sham, L. J. *Phys. Rev.* **1965**, *140*, A1133–A1138.

[64] Thomas, L. H. *Math. Proc. Cambridge Philos. Soc.* **1927**, *23*, 542–548.

[65] Fermi, E. *Z. Physik* **1928**, *48*, 73–79.

[66] Dirac, P. A. M. *Math. Proc. Cambridge* **1930**, *26*, 376–385.

[67] Kohn, W. *Rev. Mod. Phys.* **1999**, *71*, 1253–1266.

[68] Ziegler, T. *Chem. Rev.* **1991**, *91*, 651–667.

[69] Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 1372–1377.

[70] Peverati, R.; Truhlar, D. G. *J. Chem. Theory Comput.* **2012**, *8*, 2310–2319.

[71] Becke, A. D. *J. Chem. Phys.* **2014**, *140*, 18A301.

[72] Sousa, S. F.; Fernandes, P. A.; Ramos, M. J. *J. Phys. Chem. A* **2007**, *111*, 10439–10452.

[73] Goerigk, L.; Grimme, S. *Phys. Chem. Chem. Phys.* **2011**, *13*, 6670–6688.

[74] Slater, J. C. *Phys. Rev.* **1951**, *81*, 385–390.

[75] Fabian, J.; Diaz, L. A.; Seifert, G.; Niehaus, T. *J. Mol. Struct. (Theochem)* **2002**, *594*, 41–53.

[76] Vahtras, O.; Almlöf, J.; Feyereisen, M. W. *Chem. Phys. Lett.* **1993**, *213*, 514–518.

[77] Grimme, S.; Steinmetz, M. *Phys. Chem. Chem. Phys.* **2013**, *15*, 16031–16042.

[78] Görling, A.; Levy, M. *Phys. Rev. B* **1993**, *47*, 13105–13113.

[79] Facco Bonetti, A.; Engel, E.; Schmid, R. N.; Dreizler, R. M. *Phys. Rev. Lett.* **2001**, *86*, 2241–2244.

[80] Mori-Sánchez, P.; Wu, Q.; Yang, W. *J. Chem. Phys.* **2005**, *123*, 062204.

[81] Brémond, E.; Ciofini, I.; Sancho-Garc'ia, J. C.; Adamo, C. *Acc. Chem. Res.* **2016**, *49*, 1503–1513.

[82] Goerigk, L.; Grimme, S. *WIREs Comput Mol Sci* **2014**, *4*, 576–600.

[83] Eshuis, H.; Bates, J. E.; Furche, F. *Theor. Chem. Acc.* **2012**, *131*, 1084.

[84] Grimme, S. *J. Chem. Phys.* **2006**, *124*, 034108.

[85] Görling, A.; Levy, M. *Phys. Rev. A* **1994**, *50*, 196–204.

[86] Grimme, S.; Neese, F. *J. Chem. Phys.* **2007**, *127*, 154116.

[87] Adamson, R. D.; Dombroski, J. P.; Gill, P. M. W. *J. Comput. Chem.* **1999**, *20*, 921–927.

[88] Gill, P. M. W.; Adamson, R. D.; Pople, J. A. *Mol. Phys.* **1996**, *88*, 1005–1010.

[89] Leininger, T.; Stoll, H.; Werner, H.-J.; Savin, A. *Chem. Phys. Lett.* **1997**, *275*, 151–160.

[90] Iikura, H.; Tsuneda, T.; Yanai, T.; Hirao, K. *J. Chem. Phys.* **2001**, *115*, 3540–3544.

[91] Savin, A. In *Recent Advances In Density Functional Methods*; Chong, D. P., Ed.; World Scientific: Singapore, 1995; Vol. 1; pp 129–154.

[92] Yanai, T.; Tew, D. P.; Handy, N. C. *Chem. Phys. Lett* **2004**, *393*, 51–57.

[93] Henderson, T. M.; Janesko, B. G.; Scuseria, G. E. *J. Phys. Chem. A* **2008**, *112*, 12530–12542.

[94] Dreuw, A.; Weisman, J. L.; Head-Gordon, M. *J. Chem. Phys.* **2003**, *119*, 2943–2946.

[95] Dreuw, A.; Head-Gordon, M. *J. Am. Chem. Soc.* **2004**, *126*, 4007–4016.

[96] Baer, R.; Livshits, E.; Salzner, U. *Annu. Rev. Phys. Chem.* **2010**, *61*, 85–109.

[97] Sears, J. S.; Koerzdoerfer, T.; Zhang, C.-R.; Brédas, J.-L. *J. Chem. Phys.* **2011**, *135*, 151103.

[98] Toulouse, J.; Rebolini, E.; Gould, T.; Dobson, J. F.; Seal, P.; Angyán, J. G. *J. Chem. Phys.* **2013**, *138*, 194106.

[99] Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098–3100.

[100] Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785–789.

[101] Perdew, J. P. *Phys. Rev. B* **1986**, *33*, 8822–8824.

[102] Perdew, J. P. *Phys. Rev. B* **1986**, *34*, 7406–7406.

[103] Grimme, S. *J. Comput. Chem.* **2006**, *27*, 1787–1799.

[104] Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868; erratum: *Phys. Rev. Lett.* **1997**, *78*, 1396.

[105] Tao, J.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E. *Phys. Rev. Lett.* **2003**, *91*, 146401.

[106] Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648–5652.

[107] Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623–11627.

[108] Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, *120*, 215–241.

[109] Zhao, Y.; Truhlar, D. G. *Acc. Chem. Res.* **2008**, *41*, 157–167.

[110] Grimme, S.; Brandenburg, J. G.; Bannwarth, C.; Hansen, A. *J. Chem. Phys.* **2015**, *143*, 054107.

[111] Adamo, C.; Barone, V. *J. Chem. Phys.* **1999**, *110*, 6158–6170.

[112] Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A* **2005**, *109*, 5656–5667.

[113] Chai, J.-D.; Head-Gordon, M. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615–6620.

[114] Lin, Y.-S.; Li, G.-D.; Mao, S.-P.; Chai, J.-D. *J. Chem. Theory Comput.* **2013**, *9*, 263–272.

[115] Becke, A. D. *J. Chem. Phys.* **1997**, *107*, 8554–8560.

[116] Baerends, E. J.; Gritsenko, O. V.; van Meer, R. *Phys. Chem. Chem. Phys.* **2013**, *15*, 16408–16425.

[117] Roothaan, C. C. J. *Rev. Mod. Phys.* **1951**, *23*, 69–89.

[118] Hall, G. G. *Proc. R. Soc. Lond. A* **1951**, *205*, 541–552.

[119] Davidson, E. R.; Feller, D. *Chem. Rev.* **1986**, *86*, 681–696.

[120] van Duijneveldt, F. B.; van Duijneveldt-van de Rijdt, J. G. C. M.; van Lenthe, J. H. *Chem. Rev.* **1994**, *94*, 1873–1885.

[121] Sure, R.; Brandenburg, J. G.; Grimme, S. *ChemistryOpen* **2016**, *5*, 94–109.

[122] Kestner, N. R.; Combariza, J. E. In *Reviews in Computational Chemistry*; Lipkowitz, K. B., Boyd, D. B., Eds.; Wiley-VCH, 1999; Vol. 13; pp 99–132.

[123] Autschbach, J.; Ziegler, T. *Coord. Chem. Rev.* **2003**, *238*, 83–126.

[124] Bauernschmitt, R.; Ahlrichs, R. *Chem. Phys. Lett* **1996**, *256*, 454–464.

[125] Casida, M. E. In *Recent Advances In Density Functional Methods*; Chong, D. P., Ed.; World Scientific: Singapore, 1995; Vol. 1; pp 155–192.

[126] McWeeny, R. *Methods of molecular quantum mechanics*, 2nd ed.; Academic Press, 1992.

[127] Fetter, A. L.; Walecka, J. D. *Quantum Theory of Many-Particle Systems*; McGraw-Hill: New York, 1971.

[128] Hirata, S.; Head-Gordon, M. *Chem. Phys. Lett.* **1999**, *314*, 291–299.

[129] Dreuw, A.; Head-Gordon, M. *Chem. Rev.* **2005**, *105*, 4009–4037.

[130] Grimme, S.; Hansen, A.; Brandenburg, J. G.; Bannwarth, C. *Chem. Rev.* **2016**, *116*, 5105–5154.

[131] Casimir, H. B. G.; Polder, D. *Phys. Rev.* **1948**, *73*, 360–372.

[132] Stanton, J. F. *Phys. Rev. A* **1994**, *49*, 1698–1703.

[133] Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. *J. Chem. Phys.* **2010**, *132*, 154104.

[134] Grimme, S.; Ehrlich, S.; Goerigk, L. *J. Comput. Chem.* **2011**, *32*, 1456–1465.

[135] Johnson, E. R.; Becke, A. D. *J. Chem. Phys.* **2005**, *123*, 024101.

[136] Grimme, S. *WIREs Comput. Mol. Sci.* **2011**, *1*, 211–228.

[137] Risthaus, T.; Grimme, S. *J. Chem. Theory Comput.* **2013**, *9*, 1580–1591.

[138] DiLabio, G. A.; Otero-de-la-Roza, A. *Reviews in Computational Chemistry*; John Wiley & Sons, Inc, 2016; pp 1–97.

[139] Sure, R.; Grimme, S. *J. Chem. Theory. Comput.* **2015**, *11*, 3785–3801.

[140] Sure, R.; Antony, J.; Grimme, S. *J. Phys. Chem. B* **2014**, *118*, 3431–3440.

[141] Hansen, A.; Bannwarth, C.; Grimme, S.; Petrović, P.; Werlé, C.; Djukic, J.-P. *ChemistryOpen* **2014**, *3*, 177–189.

[142] Massey, A. G.; Park, A. J.; Stone, F. G. A. *Proc. Chem. Soc.* **1963**, 212.

[143] Massey, A. G.; Park, A. J. *J. Organomet. Chem.* **1964**, *2*, 245–250.

[144] Welch, G. C.; Stephan, D. W. *J. Am. Chem. Soc.* **2007**, *129*, 1880–1881.

[145] Mömming, C. M.; Otten, E.; Kehr, G.; Fröhlich, R.; Grimme, S.; Stephan, D. W.; Erker, G. *Angew. Chem. Int. Ed.* **2009**, *48*, 6643–6646; *Angew. Chem.* **2009**, *121*, 6770–6773.

[146] Paradies, J. *Synlett* **2013**, *24*, 777–780.

[147] Chen, D.; Wang, Y.; Klankermayer, J. *Angew. Chem. Int. Ed.* **2010**, *49*, 9475–9478; *Angew. Chem.* **2010**, *122*, 9665–9668.

[148] Grimme, S.; Kruse, H.; Goerigk, L.; Erker, G. *Angew. Chem. Int. Ed.* **2010**, *49*, 1402–1405; *Angew. Chem.* **2010**, *122*, 1444–1447.

[149] Cardenas, A. J. P.; Culotta, B. J.; Warren, T. H.; Grimme, S.; Stute, A.; Fröhlich, R.; Kehr, G.; Erker, G. *Angew. Chem. Int. Ed.* **2011**, *50*, 7567–7571; *Angew. Chem.* **2011**, *123*, 7709–7713.

[150] Schirmer, B.; Grimme, S. *Top. Curr. Chem.* **2013**, *332*, 213–230.

[151] Rokob, T. A.; Hamza, A.; Pápai, I. *J. Am. Chem. Soc.* **2009**, *131*, 10701–10710.

[152] Pereira, J. C. M.; Sajid, M.; Kehr, G.; Wright, A. M.; Schirmer, B.; Qu, Z.-W.; Grimme, S.; Erker, G.; Ford, P. C. *J. Am. Chem. Soc.* **2014**, *136*, 513–519.

[153] Erõs, G.; Mehdi, H.; Pápai, I.; Rokob, T.; Király, P.; Tárkányi, G.; Soós, T. *Angew. Chem. Int. Ed.* **2010**, *49*, 6559–6563; *Angew. Chem.* **2010**, *122*, 6709–6713.

[154] Rokob, T. A.; Bakó, I.; Stirling, A.; Hamza, A.; Pápai, I. *J. Am. Chem. Soc.* **2013**, *135*, 4425–4437.

[155] Chernichenko, K.; Madarász, A.; Pápai, I.; Nieger, M.; Leskelä, M.; Repo, T. *Nat. Chem.* **2013**, *5*, 718–723.

[156] Greb, L.; Oña-Burgos, P.; Schirmer, B.; Grimme, S.; Stephan, D. W. *Angew. Chem. Int. Ed.* **2012**, *51*, 10164–10168; *Angew. Chem.* **2012**, *124*, 10311–10315.

[157] Hounjet, L. J.; Bannwarth, C.; Garon, C. N.; Caputo, C. B.; Grimme, S.; Stephan, D. W. *Angew. Chem. Int. Ed.* **2013**, *52*, 7492–7495; *Angew. Chem.* **2013**, *125*, 7640–7643.

[158] Ambrosetti, A.; Alfè, D.; DiStasio, R. A.; Tkatchenko, A. *J. Phys. Chem. Lett.* **2014**, *5*, 849–855.

[159] Muddana, H. S.; Fenley, A. T.; Mobley, D. L.; Gilson, M. K. *J. Comput.-Aided Mol. Des.* **2014**, *28*, 305–317.

[160] Sajid, M.; Lawzer, A.; Dong, W.; Rosorius, C.; Sander, W.; Schirmer, B.; Grimme, S.; Daniliuc, C. G.; Kehr, G.; Erker, G. *J. Am. Chem. Soc.* **2013**, *135*, 18567–18574.

[161] Gansäuer, A.; Kube, C.; Daasbjerg, K.; Sure, R.; Grimme, S.; Fianu, G. D.; Sadasivam, D. V.; Flowers, R. A. *J. Am. Chem. Soc.* **2014**, *136*, 1663–1671.

[162] Schäfer, A.; Huber, C.; Ahlrichs, R. *J. Chem. Phys.* **1994**, *100*, 5829–5835.

[163] Weigend, F.; Ahlrichs, R. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297–3305.

[164] TURBOMOLE, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989–2007, TURBOMOLE GmbH, since 2007; available from `http://www.turbomole.com`.; (a) Version 6.5, **2013**; (b) Version 6.6, **2014**; (c) Version 7.0, **2015**.

[165] Ahlrichs, R.; Bär, M.; Häser, M.; Horn, H.; Kölmel, C. *Chem. Phys. Lett.* **1989**, *162*, 165–169.

[166] Eichkorn, K.; Treutler, O.; Öhm, H.; Häser, M.; Ahlrichs, R. *Chem. Phys. Lett.* **1995**, *240*, 283–289.

[167] Eichkorn, K.; Weigend, F.; Treutler, O.; Ahlrichs, R. *Theor. Chem. Acc.* **1997**, *97*, 119–124.

[168] Weigend, F. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1057–1065.

[169] Sure, R.; Grimme, S. *J. Comput. Chem.* **2013**, *34*, 1672–1685.

[170] Weigend, F.; Furche, F.; Ahlrichs, R. *J. Chem. Phys.* **2003**, *119*, 12753–12762.

[171] Vydrov, O. A.; Van Voorhis, T. *J. Chem. Phys.* **2010**, *133*, 244103.

[172] Hujo, W.; Grimme, S. *J. Chem. Theory Comput.* **2011**, *7*, 3866–3871.

[173] Grimme, S. *J. Chem. Phys.* **2003**, *118*, 9095–9102.

[174] Grimme, S.; Goerigk, L.; Fink, R. F. *WIREs Comput. Mol. Sci.* **2012**, *2*, 886–906.

[175] Riplinger, C.; Sandhoefer, B.; Hansen, A.; Neese, F. *J. Chem. Phys.* **2013**, *139*, 134101.

[176] Dunning, T. H., Jr. *J. Chem. Phys.* **1989**, *90*, 1007–1023.

[177] Kendall, R. A.; Dunning, T. H., Jr.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6796–6806.

[178] Woon, D. E. *J. Chem. Phys.* **1994**, *100*, 2838–2850.

[179] Van Mourik, T.; Wilson, A. K.; Dunning, T. H. *Mol. Phys.* **1999**, *96*, 529–547.

[180] Halkier, A.; Helgaker, T.; Jørgensen, P.; Klopper, W.; Koch, H.; Olsen, J.; Wilson, A. K. *Chem. Phys. Lett.* **1998**, *286*, 243–252.

[181] Boys, S.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553–566.

[182] Peintinger, M. F.; Oliveira, D. V.; Bredow, T. *J. Comput. Chem.* **2013**, *34*, 451–459.

[183] Weigend, F.; Häser, M. *Theor. Chem. Acc.* **1997**, *97*, 331–340.

[184] Neese, F.; Wennmohs, F.; Hansen, A.; Becker, U. *Chem. Phys.* **2009**, *356*, 98–109.

[185] Izsak, R.; Neese, F. *J. Chem. Phys.* **2011**, *135*, 144105.

[186] Scott, A. P.; Radom, L. *J. Phys. Chem.* **1996**, *100*, 16502–16513.

[187] Castellá-Ventura, M.; Kassab, E. *Spectrochimica Acta A* **1994**, *50*, 69–86.

[188] Klamt, A. *J. Phys. Chem.* **1995**, *99*, 2224–2235.

[189] Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. B* **2009**, *113*, 6378–6396.

[190] Klamt, A.; Schüürmann, G. *J. Chem. Soc., Perkin Trans. 2* **1993**, 799–805.

[191] F. Eckert and A. Klamt, COSMOtherm, Version C3.0, Release 12.01; COSMOlogic GmbH & Co. KG, Leverkusen, Germany, 2012.

[192] Bistoni, G.; Auer, A. A.; Neese, F. *Chem. Eur. J.* **2017**, *23*, 865–873.

[193] Kim, H.-W.; Rhee, Y.-M. *Chem. Eur. J.* **2009**, *15*, 13348–13355.

[194] Wiegand, T.; Eckert, H.; Grimme, S. *Top. Curr. Chem.* **2013**, *332*, 291–345.

[195] Risthaus, T.; Hansen, A.; Grimme, S. *Phys. Chem. Chem. Phys.* **2014**, *16*, 14408–14419.

[196] Leininger, S.; Olenyuk, B.; Stang, P. J. *Chem. Rev.* **2000**, *100*, 853–908.

[197] Holliday, B. J.; Mirkin, C. A. *Angew. Chem. Int. Ed.* **2001**, *40*, 2022–2043; *Angew. Chem.* **2001**, *113*, 2076–2097.

[198] Albrecht, M. *Chem. Rev.* **2001**, *101*, 3457–3498.

[199] Cotton, F. A.; Lin, C.; Murillo, C. A. *Acc. Chem. Res.* **2001**, *34*, 759–771.

[200] Sun, S.-S.; Lees, A. J. *Coord. Chem. Rev.* **2002**, *230*, 170–191.

[201] Würthner, F.; You, C.; Saha-Möller, C. R. *Chem. Soc. Rev.* **2004**, *33*, 133–146.

[202] Hofmeier, H.; Schubert, U. S. *Chem. Soc. Rev.* **2004**, *33*, 373–399.

[203] Fujita, M.; Tominaga, M.; Hori, A.; Therrien, B. *Acc. Chem. Res.* **2005**, *38*, 369–378.

[204] Kaiser, A.; Bäuerle, P. *Top. Curr. Chem.* **2005**, *249*, 127–201.

[205] You, C.-C.; Dobrawa, R.; Saha-Möller, C. R.; Würthner, F. *Top. Curr. Chem.* **2005**, *258*, 39–82.

[206] Nitschke, J. R. *Acc. Chem. Res.* **2007**, *40*, 103–112.

[207] Georgiev, I. G.; MacGillivray, L. R. *Chem. Soc. Rev.* **2007**, *36*, 1239–1248.

[208] Pitt, M. A.; Johnson, D. W. *Chem. Soc. Rev.* **2007**, *36*, 1441–1453.

[209] Lee, S. J.; Lin, W. *Acc. Chem. Res.* **2008**, *41*, 521–537.

[210] Dalgarno, S. J.; Power, N. P.; Atwood, J. L. *Coord. Chem. Rev.* **2008**, *252*, 825–841.

[211] Constable, E. C. *Coord. Chem. Rev.* **2008**, *252*, 842–855.

[212] Kumara, A.; Suna, S.-S.; Lees, A. J. *Coord. Chem. Rev.* **2008**, *252*, 922–939.

[213] Glasson, C. R. K.; Lindoy, L. F.; Meehan, G. V. *Coord. Chem. Rev.* **2008**, *252*, 940–963.

[214] Aromí, G.; Gamez, P.; Reedijk, J. *Coord. Chem. Rev.* **2008**, *252*, 964–989.

[215] Northrop, B. H.; Yang, H.-B.; Stang, P. J. *Chem. Commun.* **2008**, 5896–5908.

[216] Northrop, B. H.; Zheng, Y.-R.; Chi, K.; Stang, P. J. *Acc. Chem. Res.* **2009**, *42*, 1554–1563.

[217] Han, Y.-F.; Jia, W.-G.; Yu, W.-B.; Jin, G.-X. *Chem. Soc. Rev.* **2009**, *38*, 3419–3434.

[218] De, S.; Mahata, K.; Schmittel, M. *Chem. Soc. Rev.* **2010**, *39*, 1555–1575.

[219] Wiester, M. J.; Ulmann, P. A.; Mirkin, C. A. *Angew. Chem. Int. Ed.* **2011**, *50*, 114–137; *Angew. Chem.* **2011**, *123*, 118–142.

[220] Ward, M. D. *Chem. Commun.* **2009**, 4487–4499.

[221] Yoshizawa, M.; Klosterman, J. K.; Fujita, M. *Angew. Chem. Int. Ed.* **2009**, *48*, 3418–3438; *Angew. Chem.* **2009**, *121*, 3470–3490.

[222] Jain, V. K.; Jain, L. *Coord. Chem. Rev.* **2010**, *254*, 2848–2903.

[223] Yoshizawa, D.; Fujita, M. *Bull. Chem. Soc. Jpn.* **2010**, *83*, 609–618.

[224] Inokuma, Y.; Kawano, M.; Fujita, M. *Nat. Chem.* **2011**, *3*, 349–358.

[225] Chakrabarty, R.; Mukherjee, P. S.; Stang, P. J. *Chem. Rev.* **2011**, *111*, 6810–6918.

[226] Amouri, H.; Desmarets, C.; Moussa, J. *Chem. Rev.* **2012**, *112*, 2015–2041.

[227] Debata, N. B.; Tripathy, D.; Chand, D. K. *Coord. Chem. Rev.* **2012**, *256*, 1831–1945.

[228] Sahoo, H. S.; Chand, D. K. *Chem. Commun.* **2010**, *46*, 7223–7225.

[229] Tripathy, D.; Pal, A. K.; Hanan, G. S.; Chand, D. K. *Dalton Trans.* **2012**, *41*, 11273–11275.

[230] Liao, P.; Langloss, B. W.; Johnson, A. M.; Knudsen, E. R.; Tham, F. S.; Julian, R. R.; Hooley, R. J. *Chem. Commun.* **2010**, *46*, 4932–4934.

[231] Johnson, A. M.; Hooley, R. J. *Inorg. Chem.* **2011**, *50*, 4671–4673.

[232] Kishi, N.; Li, Z.; Yoza, K.; Akita, M.; Yoshizawa, M. *J. Am. Chem. Soc.* **2011**, *133*, 11438–11441.

[233] Li, Z.; Kishi, N.; Hasegawa, K.; Akita, M.; Yoshizawa, M. *Chem. Commun.* **2011**, *47*, 8605–8607.

[234] Li, Z.; Kishi, N.; Yoza, K.; Akita, M.; Yoshizawa, M. *Chem. Eur. J.* **2012**, *18*, 8358–8365.

[235] Scott, S. Ø.; Gavey, E. L.; Lind, S. J.; Gordon, K. C.; Crowley, J. D. *Dalton Trans.* **2011**, *40*, 12117–12124.

[236] Clever, G. H.; Kawamura, W.; Shionoya, M. *Inorg. Chem.* **2011**, *50*, 4689–4691.

[237] Freye, S.; Hey, J.; Torras-Galán, A.; Stalke, D.; Herbst-Irmer, R.; John, M.; Clever, G. H. *Angew. Chem. Int. Ed.* **2012**, *51*, 2191–2194; *Angew. Chem.* **2012**, *124*, 2233–2237.

[238] Clever, G. H.; Kawamura, W.; Tashirao, S.; Shiro, M.; Shionoya, M. *Angew. Chem. Int. Ed.* **2012**, *51*, 2606–2609; *Angew. Chem.* **2012**, *124*, 2660–2663.

[239] Chand, D. K.; Biradha, K.; Kawano, M.; Sakamoto, S.; Yamaguchi, K.; Fujita, M. *Chem. Asian J.* **2006**, *1*, 82–90.

[240] Suzuki, K.; Kawano, M.; Fujita, M. *Angew. Chem. Int. Ed.* **2007**, *46*, 2819–2822; *Angew. Chem.* **2007**, *119*, 2877–2880.

[241] Chand, D. K.; Fujita, M.; Biradha, K.; Sakamoto, S.; Yamaguchi, K. *J. Chem. Soc. Dalton Trans.* **2003**, 2750–2756.

[242] Suzuki, K.; Tominaga, M.; Kawano, M.; Fujita, M. *Chem. Commun.* **2009**, 1638–1640.

[243] Tominaga, M.; Suzuki, K.; Kawano, M.; Kusukawa, T.; Ozeki, T.; Sakamoto, S.; Yamaguchi, K.; Fujita, M. *Angew. Chem. Int. Ed.* **2004**, *43*, 5621–5625; *Angew. Chem.* **2004**, *116*, 5739–5743.

[244] Fujita, D.; Takahashi, A.; Sato, S.; Fujita, M. *J. Am. Chem. Soc.* **2011**, *133*, 13317–13319.

[245] Su, Q.-F.; Iwasa, J.; Ogawa, D.; Ishido, Y.; Sato, S.; Ozeki, T.; Sei, Y.; Yamaguchi, K.; Fujita, M. *Science* **2010**, *328*, 1144–1147.

[246] Bunzen, J.; Iwasa, J.; Bonakdarzadeh, P.; Numata, E.; Rissanen, K.; Sato, S.; Fujita, M. *Angew. Chem. Int. Ed.* **2012**, *51*, 3161–3163; *Angew. Chem.* **2012**, *124*, 3215–3217.

[247] Sota, S.; Iida, J.; Suzuki, K.; Kawano, M.; Ozeki, T.; Fujita, M. *Science* **2006**, *313*, 1273–1276.

[248] Murase, T.; Sato, S.; Fujita, M. *Angew. Chem. Int. Ed.* **2007**, *46*, 1083–1085; *Angew. Chem.* **2007**, *119*, 1101–1103.

[249] Murase, T.; Sato, S.; Fujita, M. *Angew. Chem. Int. Ed.* **2007**, *46*, 5133–5136; *Angew. Chem.* **2007**, *119*, 5225–5228.

[250] Suzuki, K.; Kawano, M.; Sato, S.; Fujita, M. *J. Am. Chem. Soc.* **2007**, *129*, 10652–10653.

[251] Kikuchi, T.; Murase, T.; Sato, S.; Fujita, M. *Supramol. Chem.* **2008**, *20*, 81–94.

[252] Suzuki, K.; Iida, J.; Sato, S.; Kawano, M.; Fujita, M. *Angew. Chem. Int. Ed.* **2008**, *47*, 5780–5782; *Angew. Chem.* **2008**, *120*, 5864–5866.

[253] Sato, S.; Ishido, Y.; Fujita, M. *J. Am. Chem. Soc.* **2009**, *131*, 6064–6065.

[254] Suzuki, K.; Sato, S.; Fujita, M. *Nat. Chem.* **2010**, *2*, 25–29.

[255] Suzuki, K.; Takao, K.; Sato, S.; Fujita, M. *J. Am. Chem. Soc.* **2010**, *132*, 2544–2545.

[256] Suzuki, K.; Takao, K.; Sato, S.; Fujita, M. *Angew. Chem. Int. Ed.* **2011**, *50*, 4858–4861; *Angew. Chem.* **2011**, *123*, 4960–4963.

[257] Kamiya, N.; Tominaga, M.; Sato, S.; Fujita, M. *J. Am. Chem. Soc.* **2007**, *129*, 3816–3817.

[258] Ikemi, M.; Kikuchi, T.; Matsumura, S.; Shiba, K.; Sato, S.; Fujita, M. *Chem. Sci.* **2010**, *1*, 68–71.

[259] Kikuchi, T.; Sato, S.; Fujita, M. *J. Am. Chem. Soc.* **2010**, *132*, 15930–15932.

[260] Takao, K.; Suzuki, K.; Ichijo, T.; Sato, S.; Asakura, H.; Teramura, K.; Kato, K.; Ohba, T.; Morita, T.; Fujita, M. *Angew. Chem. Int. Ed.* **2012**, *51*, 5893–5896; *Angew. Chem.* **2012**, *124*, 5995–5998.

[261] Gütz, C.; Hovorka, R.; Schnakenburg, G.; Lützen, A. *Chem. Eur. J.* **2013**, *19*, 10890–10894.

[262] Ishitani, H.; Ueno, M.; Kobayashi, S. *J. Am. Chem. Soc.* **2000**, *122*, 8180–8186.

[263] Gütz, C.; Hovorka, R.; Stobe, C.; Struch, N.; Topić, F.; Schnakenburg, G.; Rissanen, K.; Lützen, A. *Eur. J. Org. Chem.* **2014**, 206–216.

[264] Kruse, H.; Grimme, S. *J. Chem. Phys.* **2012**, *136*, 154101.

[265] Braun, J.; Renggli, K.; Razumovitch, J.; Vebert, C. In *Dynamic Light Scattering in Supramolecular Materials Chemistry in Supramolecular Chemistry–From Molecules to Nanomaterials*; P. A. Gale, J. W. S., Ed.; Wiley: Chichester, 2012; Vol. 2.

[266] Lei, P.; Hedlund, M.; Lomoth, R.; Rensmo, H.; Johansson, O.; Hammarström, L. *J. Am. Chem. Soc.* **2007**, *129*, 26–27.

[267] Osowska, K.; Miljanic, O. S. *Synlett* **2011**, 1643–1648.

[268] Safont-Sempere, M. M.; Fernández, G.; Würthner, F. *Chem. Rev.* **2011**, *111*, 5784–5814.

[269] Saha, M. L.; Schmittel, M. *Org. Biomol. Chem.* **2012**, *10*, 4651–4684.

[270] Gross, E. K. U.; Dobson, J. F.; Petersilka, M. *Top. Curr. Chem.* **1996**, *181*, 81–172.

[271] Furche, F. *J. Chem. Phys.* **2001**, *114*, 5982–5992.

[272] Peyerimhoff, S. D. In *Encyclopedia of Computational Chemistry*; v. Ragué Schleyer, P., Ed.; Wiley: New York, 1998; Vol. 4; pp 2646–2664.

[273] Harbach, P. H. P.; Dreuw, A. In *Modeling of Molecular Properties*; Comba, P., Ed.; Wiley-VCH: Weinheim, 2011; pp 29–47.

[274] *Modeling of Molecular Properties*; See also the special TD-DFT issue in *Phys. Chem. Chem. Phys*, **2009**, *11*, issue 22.

[275] Osinga, V. P.; van Gisbergen, S. J. A.; Snijders, J. G.; Baerends, E. J. *J. Chem. Phys.* **1997**, *106*, 5091–5101.

[276] van Gisbergen, S. J. A.; Snijders, J. G.; Baerends, E. J. *J. Chem. Phys.* **1995**, *103*, 9347–9354.

[277] Bally, T.; Sastry, G. N. *J. Phys. Chem. A* **1997**, *101*, 7923–7925.

[278] Zhang, Y.; Yang, W. *J. Chem. Phys.* **1998**, *109*, 2604–2608.

[279] Gritsenko, O.; Ensing, B.; Schipper, P. R. T.; Baerends, E. J. *J. Phys. Chem. A* **2000**, *104*, 8558–8565.

[280] Cohen, A. J.; Mori-Sanchez, P.; Yang, W. *Chem. Rev.* **2012**, *112*, 289–320.

[281] Tozer, D. J. *J. Chem. Phys.* **2003**, *119*, 12697–12699.

[282] Tozer, D. J.; Amos, R. D.; Handy, N. C.; Roos, B. O.; Serrano-Andrés, L. *Mol. Phys.* **1999**, *97*, 859–868.

[283] Dierksen, M.; Grimme, S. *J. Phys. Chem. A* **2004**, *108*, 10225–10237.

[284] Eshuis, H.; Yarkony, J.; Furche, F. *J. Chem. Phys.* **2010**, *132*, 234114.

[285] Isegawa, M.; Truhlar, D. G. *J. Chem. Phys.* **2013**, *138*, 134111.

[286] Grimme, S. *Chem. Phys. Lett.* **1996**, *259*, 128–137.

[287] Löwdin, P.-O. *J. Chem. Phys.* **1950**, *18*, 365–375.

[288] Nishimoto, K.; Mataga, N. *Z. Phys. Chem.* **1957**, *12*, 335–338.

[289] Ohno, K. *Theor. Chim. Acta* **1964**, *2*, 219–227.

[290] Klopman, G. *J. Am. Chem. Soc.* **1964**, *86*, 4550–4557.

[291] Ghosh, D. C.; Islam, N. *Int. J. Quantum Chem.* **2010**, *110*, 1206–1213.

[292] Buenker, R. J.; Peyerimhoff, S. D. *Theor. Chim. Acta* **1974**, *35*, 33–58.

[293] Buenker, R. J.; Peyerimhoff, S. D. *Theor. Chim. Acta* **1975**, *39*, 217–228.

[294] Grimme, S.; Waletzke, M. *J. Chem. Phys.* **1999**, *111*, 5645–5655.

[295] Neese, F. *J. Chem. Phys.* **2003**, *119*, 9428–9443.

[296] Bethe, H. A.; Salpeter, E. E. *Quantum Mechanics of One- and Two-Electron Atoms*; Dover Publications: New York, 2008; p 250.

[297] Woody, R. W. In *Circular Dichroism and the Conformational Analysis of Biomolecules*; Fasman, G. D., Ed.; Plenum Press: New York, 1996; Chapter 2, pp 25–67.

[298] Raabe, G.; Fleischhauer, J.; Woody, R. W. In *Comprehensive Chiroptical Spectroscopy, Volume 1: Instrumentation, Methodologies, and Theoretical Simulations*; Berova, N., Polavarapu, P. L., Nakanishi, K., Woody, R. W., Eds.; Wiley: New York, 2012; Chapter 20, pp 543–591.

[299] Goerigk, L.; Kruse, H.; Grimme, S. In *Comprehensive Chiroptical Spectroscopy, Volume 1: Instrumentation, Methodologies, and Theoretical Simulations*; Berova, N., Polavarapu, P. L., Nakanishi, K., Woody, R. W., Eds.; Wiley: New York, 2012; Chapter 22, pp 643–673.

[300] Schäfer, A.; Horn, H.; Ahlrichs, R. *J. Chem. Phys.* **1992**, *97*, 2571–2577.

[301] Furche, F.; Rappoport, D. In *Theoretical and Computational Chemistry: Computational Photochemistry*; Olivucci, M., Ed.; Elsevier: Amsterdam, 2005; Vol. 16; Chapter III, pp 93–128.

[302] Gütz, C.; Hovorka, R.; Klein, C.; Jiang, Q.-Q.; Bannwarth, C.; Engeser, M.; Schmuck, C.; Assenmacher, W.; Mader, W.; Topić, F.; Rissanen, K.; Grimme, S.; Lützen, A. *Angew. Chem. Int. Ed.* **2014**, *53*, 1693–1698; *Angew. Chem.* **2014**, *126*, 1719–1724.

[303] Furche, F.; Ahlrichs, R.; Wachsmann, C.; Weber, E.; Sobanski, A.; Vögtle, F.; Grimme, S. *J. Am. Chem. Soc.* **2000**, *122*, 1717–1724.

[304] Sehnal, P.; Stará, I. G.; Šaman, D.; Tichý, M.; Míšek, J.; Cvačka, J.; Rulíšek, L.; Chocholoušová, J.; Vacek, J.; Goryl, G.; Szymonski, M.; Císařová, I.; Starý, I. *Proc. Nat. Acad. Sci.* **2009**, *106*, 13169–13174.

[305] Kessinger, R.; Thilgen, C.; Mordasini, T.; Diederich, F. *Helv. Chim. Acta* **2000**, *83*, 3069–3096.

[306] See `http://www.thch.uni-bonn.de/`.

[307] Isborn, C. M.; Luehr, N.; Ufimtsev, I. S.; Martínez, T. J. *J. Chem. Theory Comput.* **2011**, *7*, 1814–1823.

[308] Autschbach, J. In *Comprehensive Chiroptical Spectroscopy, Volume 1: Instrumentation, Methodologies, and Theoretical Simulations*; Berova, N., Polavarapu, P. L., Nakanishi, K., Woody, R. W., Eds.; Wiley: New York, 2012; Chapter 22, pp 593–642.

[309] Crawford, T. *Theor. Chem. Acc.* **2006**, *115*, 227–245.

[310] Pople, J. A. *Trans. Faraday Soc.* **1953**, *49*, 1375–1385.

[311] Pariser, R.; Parr, R. G. *J. Chem. Phys.* **1953**, *21*, 767–776.

[312] Del Bene, J.; Jaffé, H. H. *J. Chem. Phys.* **1968**, *48*, 1807–1813.

[313] Del Bene, J.; Jaffé, H. H. *J. Chem. Phys.* **1968**, *48*, 4050–4055.

[314] Del Bene, J.; Jaffé, H. H. *J. Chem. Phys.* **1969**, *50*, 1126–1129.

[315] Ridley, J.; Zerner, M. *Theoret. Chim. Acta* **1976**, *42*, 223–236.

[316] Ridley, J.; Zerner, M. *Theoret. Chim. Acta* **1973**, *32*, 111–134.

[317] Bredow, T.; Jug, K. In *Encyclopedia of Computational Chemistry (Online Wdition)*; von Ragué Schleyer, P., Ed.; Wiley: New York, 2004.

[318] Ahlswede, B.; Jug, K. *J. Comput. Chem.* **1999**, *20*, 563–571.

[319] Gadaczek, I.; Krause, K.; Hintze, K. J.; Bredow, T. *J. Chem. Theory Comput.* **2011**, *7*, 3675–3685.

[320] Voityuk, A. A. *J. Chem. Theory Comput.* **2014**, *10*, 4950–4958.

[321] Linnanto, J.; Korppi-Tommola, J. E. I.; Helenius, V. M. *J. Phys. Chem. B* **1999**, *103*, 8739–8750.

[322] Moffitt, W. *J. Chem. Phys.* **1956**, *25*, 467–478.

[323] Gilbert, A. T. B.; Hirst, J. D. *J. Mol. Struct. (Theochem)* **2004**, *675*, 53–60.

[324] Botek, E.; Champagne, B. *J. Chem. Phys.* **2007**, *127*, 204101.

[325] Goto, H.; Harada, N.; Crassous, J.; Diederich, F. *J. Chem. Soc., Perkin Trans. 2* **1998**, 1719–1724.

[326] Orlandi, G.; Poggi, G.; Zerbetto, F. *Chem. Phys. Lett.* **1994**, *224*, 113–117.

[327] Niehaus, T. A.; Suhai, S.; Della Sala, F.; Lugli, P.; Elstner, M.; Seifert, G.; Frauenheim, T. *Phys. Rev. B* **2001**, *63*, 085108.

[328] Rüger, R.; van Lenthe, E.; Lu, Y.; Frenzel, J.; Heine, T.; Visscher, L. *J. Chem. Theory Comput.* **2015**, *11*, 157–167.

[329] Bannwarth, C.; Grimme, S. *Comput. Theor. Chem.* **2014**, *1040–1041*, 45–53.

[330] Lopata, K.; Govind, N. *J. Chem. Theory Comput.* **2012**, *8*, 3284–3292.

[331] Fowler, P. W.; Manolopoulos, D. E. *An Atlas of Fullerenes*; Dover Publications: New York, 2007; pp 88–89.

[332] Furche, F.; Ahlrichs, R. *J. Am. Chem. Soc.* **2002**, *124*, 3804–3805.

[333] Wildöer, J. W. G.; Venema, L. C.; Rinzler, A. G.; Smalley, R. E.; Dekker, C. *Nature* **1998**, *391*, 59–62.

[334] Goldmann, E.; Asher, S. A.; Mukamel, S. *Phys. Chem. Chem. Phys.* **2001**, *3*, 2893–2903.

[335] Harada, N.; Nakanishi, K. *Circular Dichroic Spectroscopy - Exciton Coupling in Organic Stereochemistry*; University Science Books, 1983; pp 406–436.

[336] Furche, F. Dichtefunktionalmethoden für Elektronisch Angeregte Moleküle. Theorie - Implementierung - Anwendung. Ph.D. thesis, Universität Karlsruhe, 2002.

[337] Furche, F.; Burke, K. *Annu. Rep. Comput. Chem.* **2005**, *1*, 19–30.

[338] Polavarapu, P. L.; He, J.; Crassous, J.; Ruud, K. *ChemPhysChem* **2005**, *6*, 2535–2540.

[339] Peach, M. J. G.; Williamson, M. J.; Tozer, D. J. *J. Chem. Theory Comput.* **2011**, *7*, 3578–3585.

[340] Vögtle, F.; Hünten, A.; Vogel, E.; Buschbeck, S.; Safarowsky, O.; Recker, J.; Parham, A.-H.; Knott, M.; Müller, W. M.; Müller, U.; Okamoto, Y.; Kubota, T.; Lindner, W.; Francotte, E.; Grimme, S. *Angew. Chem. Int. Ed.* **2001**, *40*, 2468–2471; *Angew. Chem.* **2001**, *113*, 2534–2537.

[341] Cozzi, W. H. P., F.; Thilgen, C. *Pure Appl. Chem.* **2009**, *77*, 801–923.

[342] Thilgen, C.; Herrmann, A.; Diederich, F. *Helv. Chim. Acta* **1997**, *80*, 183–199.

[343] Barlow, D. J.; Thornton, J. M. *J. Mol. Biol.* **1988**, *201*, 601–619.

[344] Hehre, W. J.; Ditchfield, R.; Pople, J. A. *J. Chem. Phys.* **1972**, *56*, 2257–2261.

[345] Ditchfield, R.; Hehre, W. J.; Pople, J. A. *J. Chem. Phys.* **1971**, *54*, 724–728.

[346] Johnson, Jr., W. C.; Tinoco, Jr., I. *J. Am. Chem. Soc.* **1972**, *94*, 4389–4390.

[347] Tinoco, I., Jr.; Woody, R. W. *J. Chem. Phys.* **1967**, *46*, 4927–4945.

[348] Woody, R. W.; Koslowski, A. *Biophys. Chem.* **2002**, *101–102*, 535–551.

[349] Woody, R. W. *Biomed. Spectrosc. Imaging* **2015**, *4*, 5–34.

[350] Mori, K.; Murase, T.; Fujita, M. *Angew. Chem. Int. Ed.* **2015**, *54*, 6847–6851; *Angew. Chem.* **2015**, *127*, 6951–6955.

[351] Martin, R. H. *Angew. Chem. Int. Ed.* **1974**, *13*, 649–660; *Angew. Chem.* **1974**, *86*, 727–738.

[352] Weitzenböck, R.; Klingler, A. *Monatsh. Chem.* **1918**, *39*, 315–323.

[353] Newman, M. S.; Lutz, W. B.; Lednicer, D. *J. Am. Chem. Soc.* **1955**, *77*, 3420–3421.

[354] Newman, M. S.; Lednicer, D. *J. Am. Chem. Soc.* **1956**, *78*, 4765–4770.

[355] Newman, M. S.; Darlak, R. S.; Tsai, L. L. *J. Am. Chem. Soc.* **1967**, *89*, 6191–6193.

[356] Lightner, D. A.; Hefelfinger, D. T.; Frank, G. W.; Powers, T. W.; Trueblood, K. N. *Nature-Phys. Sci.* **1971**, *232*, 124–125.

[357] Brown, A.; Kemp, C. M.; Mason, S. F. *J. Chem. Soc. A* **971**, 751–755.

[358] Brickell, W. S.; Brown, A.; Kemp, C. M.; Mason, S. F. *J. Chem. Soc. A* **1971**, 756–760.

[359] Autschbach, J.; Ziegler, T.; van Gisbergen, S. J. A.; Baerends, E. J. *J. Chem. Phys.* **2002**, *116*, 6930–6940.

[360] Shen, Y.; Chen, C.-F. *Chem. Rev.* **2012**, *112*, 1463–1535.

[361] Gingras, M.; Felix, G.; Peresutti, R. *Chem. Soc. Rev.* **2013**, *42*, 1007–1050.

[362] Bannwarth, C.; Grimme, S. *J. Phys. Chem. A* **2015**, *119*, 3653–3662.

[363] Hirst, J. D.; III, C. L. B. *J. Mol. Biol.* **1994**, *243*, 173–178.

[364] Sreerama, N.; Woody, R. W. *Proteins: Struct., Funct., Bioinf.* **1999**, *36*, 400–406.

[365] Blauer, G.; Sreerama, N.; Woody, R. W. *Biochemistry* **1993**, *32*, 6674–6679.

[366] Jensen, L.; Swart, M.; van Duijnen, P. T.; Autschbach, J. *Int. J. Quantum Chem.* **2006**, *106*, 2479–2488.

[367] Frelek, J.; Kowalska, P.; Masnyk, M.; Kazimierski, A.; Korda, A.; Woźnica, M.; Chmielewski, M.; Furche, F. *Chem. Eur. J.* **2007**, *13*, 6732–6744.

[368] Kundrat, M. D.; Autschbach, J. *J. Chem. Theory Comput.* **2009**, *5*, 1051–1060.

[369] Kaminský, J.; Kubelka, J.; Bouř, P. *J. Phys. Chem. A* **2011**, *115*, 1734–1742.

[370] Chai, J.-D.; Head-Gordon, M. *J. Chem. Phys.* **2008**, *128*, 084106.

[371] Furche, F.; Ahlrichs, R.; Hättig, C.; Klopper, W.; Sierka, M.; Weigend, F. *WIREs Comput. Mol. Sci.* **2014**, *4*, 91–100.

[372] Körzdörfer, T.; Sears, J. S.; Sutton, C.; Brédas, J.-L. *J. Chem. Phys.* **2011**, *135*, 204107.

[373] Srebro, M.; Autschbach, J. *J. Chem. Theory Comput.* **2012**, *8*, 245–256.

[374] Head-Gordon, M.; Grana, A. M.; Maurice, D.; White, C. A. *J. Phys. Chem.* **1995**, *99*, 14261–14270.

[375] P. Shushkov, S. Grimme. QMSIM, Version 1.0. Universität Bonn (2015). `http://www.thch.uni-bonn.de/tc/`.

[376] Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. *J. Comput. Phys.* **1977**, *23*, 327–341.

[377] van Gunsteren, W.; Berendsen, H. *Mol. Phys.* **1977**, *34*, 1311–1327.

[378] Koch, H.; ; Jørgensen, P. *J. Chem. Phys.* **1990**, *93*, 3333.

[379] Comeau, D. C.; Bartlett, R. J. *Chem. Phys. Lett.* **1993**, *207*, 414.

[380] Christiansen, O.; Koch, H.; Jørgensen, P. *Chem. Phys. Lett.* **1995**, *243*, 409–418.

[381] Angeli, C.; Cimiraglia, R.; Evangelisti, S.; Leininger, T.; Malrieu, J.-P. *J. Chem. Phys.* **2001**, *114*, 10252.

[382] Kats, D.; Schütz, M. *Z. Phys. Chem.* **2010**, *224*, 601–616.

[383] Barone, V.; Polimeno, A. *Chem. Soc. Rev.* **2007**, *36*, 1724–1731.

[384] Jacquemin, D.; Perpete, E. A.; Scalmani, G.; Frisch, M. J.; Kobayashi, R.; Adamo, C. *J. Chem. Phys.* **2007**, *126*, 144105.

[385] Parrish, R. M.; Hohenstein, E. G.; Martínez, T. J. *J. Chem. Theory Comput.* **2016**, *12*, 3003–3007.

[386] Furche, F.; Krull, B. T.; Nguyen, B. D.; Kwon, J. *J. Chem. Phys.* **2016**, *144*, 174105.

[387] Wiebeler, C.; Bader, C. A.; Meier, C.; Schumacher, S. *Phys. Chem. Chem. Phys.* **2014**, *16*, 14531–14538.

[388] Giansante, C.; Infante, I.; Fabiano, E.; Grisorio, R.; Suranna, G. P.; Gigli, G. *J. Am. Chem. Soc.* **2015**, *137*, 1875–1886.

[389] Azpiroz, J. M.; Infante, I.; Angelis, F. D. *J. Phys. Chem. C* **2015**, *119*, 12739–12748.

[390] Azpiroz, J. M.; Angelis, F. D. *ACS Appl. Mater. Interfaces* **2015**, *7*, 19736–19745.

[391] Chmela, J.; Harding, M. E.; Matioszek, D.; Anson, C. E.; Breher, F.; Klopper, W. *ChemPhysChem* **2016**, *17*, 191–191.

[392] Gütz, C.; Hovorka, R.; Struch, N.; Bunzen, J.; Meyer-Eppler, G.; Qu, Z.-W.; Grimme, S.; Topić, F.; Rissanen, K.; Cetina, M.; Engeser, M.; Lützen, A. *J. Am. Chem. Soc.* **2014**, *136*, 11830–11838.

[393] Meyer-Eppler, G.; Sure, R.; Schneider, A.; Schnakenburg, G.; Grimme, S.; Lützen, A. *J. Org. Chem.* **2014**, *79*, 6679–6687.

[394] Jarzebski, A.; Bannwarth, C.; Tenten, C.; Benkhäuser, C.; Schnakenburg, G.; Grimme, S.; Lützen, A. *Synthesis* **2015**, *47*, 3118–3132.

[395] Frost, J. R.; Huber, S. M.; Breitenlechner, S.; Bannwarth, C.; Bach, T. *Angew. Chem. Int. Ed.* **2015**, *54*, 691–695; *Angew. Chem.* **2015**, *127*, 701–705.

[396] Bannwarth, C.; Seibert, J.; Grimme, S. *Chirality* **2016**, *28*, 365–369.

[397] Gehrold, A. C.; Bruhn, T.; Bringmann, G. *J. Org. Chem.* **2016**, *81*, 1075–1088.

[398] Trani, F.; Scalmani, G.; Zheng, G.; Carnimeo, I.; Frisch, M. J.; Barone, V. *J. Chem. Theory Comput.* **2011**, *7*, 3304–3313.

[399] Kowalczyk, T.; Le, K.; Irle, S. *J. Chem. Theo. Comp.* **2016**, *12*, 313–323.

[400] Hättig, C.; Weigend, F. *J. Chem. Phys.* **2000**, *113*, 5154–5161.

[401] Hellweg, A.; Grün, S. A.; Hättig, C. *Phys. Chem. Chem. Phys.* **2008**, *10*, 4119–4127.

[402] Marenich, A. V.; Jerome, S. V.; Cramer, C. J.; Truhlar, D. G. *J. Chem. Theory Comput.* **2012**, *8*, 527–541.

[403] Gasteiger, J.; Marsili, M. *J. Comput. Chem.* **1978**, *19*, 3181–3184.

[404] Frauenheim, T.; Seifert, G.; Elstner, M.; Niehaus, T.; Köhler, C.; Amkreutz, M.; Sternberg, M.; Hajnal, Z.; Carlo, A. D.; Suhai, S. *J. Phys. Condens. Matter* **2002**, *14*, 3015.

[405] Chong, D. P. *Mol. Phys.* **2005**, *103*, 749–761.

[406] Weber, W.; Thiel, W. *Theor. Chem. Acc.* **2000**, *103*, 495–506.

[407] MNDO2005 Version 7.0, W. Thiel, MPI für Kohlenforschung, Mülheim, Germany.

[408] Stewart, J. J. P. *J. Mol. Model.* **2007**, *13*, 1173.

[409] Bingham, R. C.; Dewar, M. J. S.; Lo, D. H. *J. Am. Chem. Soc.* **1975**, *97*, 1285–1293.

[410] Hehre, W. J.; Stewart, R. F.; Pople, J. A. *J. Chem. Phys.* **1969**, *51*, 2657–2664.

[411] Elstner, M.; Porezag, D.; Jungnickel, G.; Elsner, J.; Haugk, M.; Frauenheim, T.; Suhai, S.; Seifert, G. *Phys. Rev. B* **1998**, *58*, 7260–7268.

[412] Mermin, N. D. *Phys. Rev. A* **1965**, *137*, 1441–1443.

[413] Hirshfeld, F. *Theor. chim. acta* **1977**, *44*, 129–138.

[414] Storer, J. W.; Giesen, D. J.; Cramer, C. J.; Truhlar, D. G. *J. Comput. Aid. Mol. Des.* **1995**, *9*, 87–110.

[415] Cramer, C. J.; Truhlar, D. G. *J. Comput. Aid. Mol. Des.* **1992**, *6*, 629–666.

[416] Gonzalez-Lafont, A.; Truong, T. N.; Truhlar, D. G. *J. Phys. Chem.* **1991**, *95*, 4618–4627.

[417] Wiberg, K. B. *Tetrahedron* **1968**, *24*, 1083–1096.

[418] Roehrig, S.; Straub, A.; Pohlmann, J.; Lampe, T.; Pernerstorfer, J.; Schlemmer, K.-H.; Reinemer, P.; Perzborn, E. *J. Med. Chem.* **2005**, *48*, 5900–5908.

[419] Levenberg, K. *Q. Appl. Math.* **1944**, *2*, 164–168.

[420] Marquardt, D. *J. Soc. Ind. Appl. Math.* **1963**, *11*, 431–441.

[421] Goerigk, L.; Grimme, S. *J. Chem. Phys.* **2010**, *132*, 184103.

[422] Brückner, C.; Engels, B. *J. Phys. Chem. A* **2015**, *119*, 12876–12891.

[423] Li, Z.; Liu, W. *J. Chem. Theor. Comput.* **2016**, *12*, 238–260.

[424] Turecek, F. *J. Phys. Chem. A* **2015**, *119*, 10101–10111.

[425] Stewart, J. J. P. *MOPAC2012*; Stewart Computational Chemistry: Colorado Springs, CO, USA, 2012; http://OpenMOPAC.net (August 9, 2014).

[426] te Velde, G.; Bickelhaupt, F. M.; Baerends, E. J.; Fonseca Guerra, C.; van Gisbergen, S. J. A.; Snijders, J. G.; Ziegler, T. *J. Comp. Chem.* **2001**, *22*, 931–967.

[427] ADF DFTB 2013, SCM, Theoretical Chemistry, Vrije Universiteit, Amsterdam, The Netherlands, `http://www.scm.com`.

[428] Wahiduzzaman, M.; Oliveira, A. F.; Philipsen, P.; Zhechkov, L.; van Lenthe, E.; Witek, H. A.; Heine, T. *J. Chem. Theory Comput.* **2013**, *9*, 4006.

[429] Weigend, F.; Köhn, A.; Hättig, C. *J. Chem. Phys.* **2002**, *116*, 3175–3183.

[430] Robin, M. B. *Higher Excited States of Polyatomic Molecules, Vol. 1-2*; Academic Press: New York, 1975.

[431] Gaus, M.; Goez, A.; Elstner, M. *J. Chem. Theory Comput.* **2013**, *9*, 338–354.

[432] Grimme, S. *J. Comp. Chem.* **1994**, *15*, 424–432.

[433] Chaudhuri, D.; Sigmund, E.; Meyer, A.; Röck, L.; Klemm, P.; Lautenschlager, S.; Schmid, A.; Yost, S. R.; Voorhis, T. V.; Bange, S.; Höger, S.; Lupton, J. M. *Angew. Chem. Int. Ed.* **2013**, *52*, 13449–13452; *Angew. Chem.* **2013**, *125*, 13691–13694.

[434] Claus, S.; Höger, S. *private communication.*

[435] Gaus, M.; Cui, Q.; Elstner, M. *J. Chem. Theory Comput.* **2011**, *7*, 931–948.

[436] Frauenheim, T. *DFTB+ (Density Functional based Tight Binding)*; DFTB.ORG, Universität Bremen: Bremen, Germany, 2008; http://www.dftb.org (August 9, 2014).

[437] Brandenburg, J. G.; Hochheim, M.; Bredow, T.; Grimme, S. *J. Phys. Chem. Lett.* **2014**, *5*, 4275–4284.

[438] Goerigk, L.; Grimme, S. *Chem. Phys. Chem.* **2008**, *9*, 2467–2470.

[439] Rüger, R.; van Lenthe, E.; Heine, T.; Visscher, L. *The Journal of Chemical Physics* **2016**, *144*, 184103.

[440] Huang, G.-J.; Harris, M. A.; Krzyaniak, M. D.; Margulies, E. A.; Dyar, S. M.; Lindquist, R. J.; Wu, Y.; Roznyatovskiy, V. V.; Wu, Y.-L.; Young, R. M.; Wasielewski, M. R. *J. Phys. Chem. B* **2016**, *120*, 756–765.

[441] Nakanishi, H.; Sumi, N.; Ueno, S.; Takimiya, K.; Aso, Y.; Otsubo, T.; Komaguchi, K.; Shiotani, M.; Ohta, N. *Synt. Met.* **2001**, *119*, 413–414.

[442] Schmidt, D.; Son, M.; Lim, J. M.; Lin, M.-J.; Krummenacher, I.; Braunschweig, H.; Kim, D.; Würthner, F. *Angew. Chem. Int. Ed.* **2015**, *54*, 13980–13984; *Angew. Chem.* **2015**, *127*, 14186–14190.

[443] Bannwarth, C.; Shushkov, P.; Grimme, S. *manuscript in preparation*.

[444] Toniolo, C.; Formaggio, F.; Woody, R. W. In *Comprehensive Chiroptical Spectroscopy, Volume 2: Applications in Stereochemical analysis of Synthetic Compounds*; N. Berova, K. N., P. L. Polavarapu, Woody, R. W., Eds.; Wiley: New York, 2012; pp 499–574.

[445] Wallace, B. A.; Gekko, K.; Hoffmann, S. V.; Lin, Y.; Sutherland, J. C.; Tao, Y.; Wien, F.; Janes, R. W. *Nuc. Inst. Method. Phys. Res. A* **2011**, *649*, 177–178.

[446] Sasaki, C.; Nakajima, K.; Kojima, M.; Fujita, J. *Bull. Chem. Soc. Jpn.* **1991**, *64*, 1318–1324.

[447] Dierksen, M.; Grimme, S. *J. Chem. Phys.* **2004**, *120*, 3544.

[448] Barone, V.; Baiardi, A.; Biczysko, M.; Bloino, J.; Cappelli, C.; Lipparini, F. *Phys. Chem. Chem. Phys.* **2012**, *14*, 12404–12422.

[449] Pantos, E.; Philis, J.; Bolovinos, A. *J. Mol. Spectrosc.* **1978**, *72*, 36–43.

[450] Leach, S.; Vervloet, M.; Despres, A.; Breheret, E.; Hare, J. P.; Dennis, T. J.; Kroto, H. W.; Taylor, R.; Walton, D. R. M. *Chem. Phys.* **1992**, *160*, 451–466.

[451] Brandenburg, J. G.; Caldeweyher, E.; Grimme, S. *Phys. Chem. Chem. Phys.* **2016**, *18*, 15519–15523.

[452] Dror, R. O.; Dirks, R. M.; Grossman, J.; Xu, H.; Shaw, D. E. *Annu. Rev. Biophys.* **2012**, *41*, 429–452.

[453] Salomon-Ferrer, R.; Case, D. A.; Walker, R. C. *WIREs Comput. Mol. Sci.* **2013**, *3*, 198–210.

[454] Zhu, X.; Lopes, P. E. M.; MacKerell, A. D. *WIREs Comput. Mol. Sci.* **2012**, *2*, 167–185.

[455] Yilmazer, N. D.; Korth, M. *Comput. Struct. Biotechnol. J.* **2015**, *13*, 169–175.

[456] Yang, Y.; Yu, H.; York, D.; Cui, Q.; Elstner, M. *J. Phys. Chem. A* **2007**, *111*, 10861–10873.

[457] Cui, Q.; Elstner, M.; Kaxiras, E.; Frauenheim, T.; Karplus, M. *J. Phys. Chem. B* **2001**, *105*, 569–585.

[458] Elstner, M. *Theor. Chem. Acc.* **2006**, *116*, 316–325.

[459] Elstner, M.; Cui, Q.; Munih, P.; Kaxiras, E.; Frauenheim, T.; Karplus, M. *J. Comput. Chem.* **2003**, *24*, 565–581.

[460] Elstner, M.; Jalkanen, K. J.; Knapp-Mohammady, M.; Frauenheim, T.; Suhai, S. *Chem. Phys.* **2001**, *263*, 203–219.

[461] Zheng, G.; Irle, S.; Morokuma, K. *Chem. Phys. Lett.* **2005**, *412*, 210–216.

[462] Sattelmeyer, K. W.; Tirado-Rives, J.; Jorgensen, W. L. *J. Phys. Chem. A* **2006**, *110*, 13551–13559.

[463] Zhu, H.; Tajkhorshid, E.; Frauenheim, T.; Suhai, S.; Elstner, M. *Chem. Phys.* **2002**, *277*, 91–103.

[464] Frisch, M. J. et al. *Gaussian 09 Revision E.01*; Gaussian Inc. Wallingford CT 2009.

[465] Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. *J. Comp. Chem.* **1993**, *14*, 1347–1363.

[466] Korth, M.; Pitoňák, M.; Řezáč, J.; Hobza, P. *J. Chem. Theory Comput.* **2010**, *6*, 344–352.

[467] Kromann, J. C.; Christensen, A. S.; Steinmann, C.; Korth, M.; Jensen, J. H. *PeerJ* **2014**, *2*, e449.

[468] Řezáč, J.; Hobza, P. *J. Chem. Theory Comput.* **2012**, *8*, 141–151.

[469] S. Brahmkshatriya, P.; Dobeš, P.; Fanfrlík, J.; Řezáč, J.; Paruch, K.; Bronowska, A.; Lepšík, M.; Hobza, P. *Curr. Comput.-Aid. Drug.* **2013**, *9*, 118–129.

[470] Saito, T.; Kitagawa, Y.; Takano, Y. *J. Phys. Chem. A* **2016**, *120*, 8750–8760.

[471] Tuttle, T.; Thiel, W. *Phys. Chem. Chem. Phys.* **2008**, *10*, 2159–2166.

[472] Dral, P. O.; Wu, X.; Spörkel, L.; Koslowski, A.; Weber, W.; Steiger, R.; Scholten, M.; Thiel, W. *J. Chem. Theory Comput.* **2016**, *12*, 1082–1096.

[473] Koslowski, A.; Beck, M. E.; Thiel, W. *J. Comput. Chem.* **2003**, *24*, 714–726.

[474] Grimme, S. *Angew. Chem. Int. Ed.* **2013**, *52*, 6306–6312; *Angew. Chem.* **2013**, *125*, 6426–6433.

[475] Stewart, J. *J. Mol. Model.* **2013**, *19*, 1–32.

[476] Miyamoto, K.; Miller, T. F.; Manby, F. R. *J. Chem. Theory Comput.* **2016**, *12*, 5811–5822.

[477] Grimme, S.; Bannwarth, C. *J. Chem. Phys.* **2016**, *145*, 054103.

[478] Jakalian, A.; Jack, D. B.; Bayly, C. I. *J. Comput. Chem.* **2002**, *23*, 1623–1641.

[479] Bauer, C. A.; Grimme, S. *J. Phys. Chem. A* **2016**, *120*, 3755–3766.

[480] Grimme, S.; Hansen, A. *Angew. Chem. Int. Ed.* **2015**, *54*, 12308–12313; *Angew. Chem.* **2015**, *127*, 12483–12488.

[481] Köhler, C.; Seifert, G.; Frauenheim, T. *Chem. Phys.* **2005**, *309*, 23–31.

[482] Gaus, M.; Jin, H.; Demapan, D.; Christensen, A. S.; Goyal, P.; Elstner, M.; Cui, Q. *J. Chem. Theory Comput.* **2015**, *11*, 4205–4219.

[483] Hülsen, M.; Weigand, A.; Dolg, M. *Theor. Chem. Acc.* **2009**, *122*, 23–29.

[484] Mantina, M.; Valero, R.; Cramer, C. J.; Truhlar, D. G. In *CRC Handbook of Chemistry and Physics, 91nd edition*; Haynes, W. M., Ed.; CRC Press: Boca Raton, FL, 2010; pp 9–49–9–50.

[485] Řezáč, J.; Hobza, P. *Chem. Phys. Lett.* **2011**, *506*, 286–289.

[486] Eckert, F.; Pulay, P.; Werner, H.-J. *J. Comput. Chem.* **1997**, *18*, 1473–1483.

[487] Lindh, R.; Bernhardsson, A.; Karlström, G.; Malmqvist, P.-Å. *Chem. Phys. Lett.* **1995**, *241*, 423–428.

[488] Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T. *J. Am. Chem. Soc.* **1990**, *112*, 6127–6129.

[489] Ooi, T.; Oobatake, M.; Nemethy, G.; Scheraga, H. A. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 3086–3090.

[490] Shushkov, P.; Grimme, S. *manuscript in preparation*.

[491] Please contact `xtb@thch.uni-bonn.de` for the program.

[492] Jurečka, P.; Šponer, J.; Černý, J.; Hobza, P. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1985–1993.

[493] Marshall, M. S.; Burns, L. A.; Sherrill, C. D. *J. Chem. Phys.* **2011**, *135*, 194102.

[494] Řezáč, J.; Riley, K. E.; Hobza, P. *J. Chem. Theory Comput.* **2011**, *7*, 2427–2438.

[495] Sedlak, R.; Janowski, T.; Pitoňák, M.; Řezáč, J.; Pulay, P.; Hobza, P. *J. Chem. Theory. Comput.* **2013**, *9*, 3364–3374.

[496] Kozuch, S.; Martin, J. M. L. *J. Chem. Theory Comput.* **2013**, *9*, 1918–1931.

[497] Gaus, M.; Lu, X.; Elstner, M.; Cui, Q. *J. Chem. Theory Comput.* **2014**, *10*, 1518–1537.

[498] Kubillus, M.; Kubař, T.; Gaus, M.; Řezáč, J.; Elstner, M. *J. Chem. Theory Comput.* **2015**, *11*, 332–342.

[499] Stewart, J. J. P. *MOPAC2016*; Stewart Computational Chemistry: Colorado Springs, CO, USA, 2016; http://OpenMOPAC.net (August 16, 2016).

[500] Oliveira, A. F.; Philipsen, P.; Heine, T. *J. Chem. Theory Comput.* **2015**, *11*, 5209–5218.

[501] Zheng, G.; Witek, H. A.; Bobadova-Parvanova, P.; Irle, S.; Musaev, D. G.; Prabhakar, R.; Morokuma, K.; Lundberg, M.; Elstner, M.; Köhler, C.; Frauenheim, T. *J. Chem. Theory Comput.* **2007**, *3*, 1349–1367.

[502] Korth, M. *J. Chem. Theory Comput.* **2010**, *6*, 3808–3816.

[503] Brandenburg, J. G.; Grimme, S. *Top. Curr. Chem.* **2014**, *345*, 1–23.

[504] Risthaus, T.; Steinmetz, M.; Grimme, S. *J. Comput. Chem.* **2014**, *35*, 1509–1516.

[505] Puzzarini, C.; Heckert, M.; Gauss, J. *J. Chem. Phys.* **2008**, *128*, 194108–194108.

[506] Bühl, M.; Kabrede, H. *J. Chem. Theory Comput.* **2006**, *2*, 1282–1290.

[507] Wagner, J. P.; Schreiner, P. R. *Angew. Chem. Int. Ed.* **2015**, *54*, 12274–12296; *Angew. Chem.* **2015**, *127*, 12446–12471.

[508] Bryantsev, V. S.; Diallo, M. S.; van Duin, A. C. T.; Goddard, W. A. *J. Chem. Theory Comput.* **2009**, *5*, 1016–1026.

[509] Anacker, T.; Friedrich, J. *J. Comput. Chem.* **2014**, *35*, 634–643.

[510] Antony, J.; Grimme, S.; Liakos, D. G.; Neese, F. *J. Phys. Chem. A* **2011**, *115*, 11210–11220.

[511] Zhao, Y.; Truhlar, D. G. *J. Chem. Theory Comput.* **2005**, *1*, 415–432.

[512] Řezáč, J.; Riley, K. E.; Hobza, P. *J. Chem. Theory Comput.* **2012**, *8*, 4285–4292.

[513] Vogiatzis, K. D.; Klopper, W.; Friedrich, J. *J. Chem. Theory Comput.* **2015**, *11*, 1574–1584.

[514] Valdes, H.; Pluhackova, K.; Pitonák, M.; Řezáč, J.; Hobza, P. *Phys. Chem. Chem. Phys.* **2008**, *10*, 2747–2757.

[515] Brandenburg, J. G.; Grimme, S. *Acta Crystallogr. Sect. B* **2016**, *B72*, 502–513.

[516] Gruzman, D.; Karton, A.; Martin, J. M. L. *J. Phys. Chem. A* **2009**, *113*, 11974–11983.

[517] Wilke, J. J.; Lind, M. C.; Schaefer, H. F.; Császár, A. G.; Allen, W. D. *J. Chem. Theory Comput.* **2009**, *5*, 1511–1523.

[518] Řeha, D.; Valdés, H.; Vondrášek, J.; Hobza, P.; Abu-Riziq, A.; Crews, B.; de Vries, M. S. *Chem. Eur. J.* **2005**, *11*, 6803–6817.

[519] Kozuch, S.; Martin, J. M. L. *J. Comput. Chem.* **2013**, *34*, 2327–2344.

[520] Liakos, D. G.; Neese, F. *J. Chem. Theory Comput.* **2015**, *11*, 2137–2143.

[521] Kruse, H.; Mladek, A.; Gkionis, K.; Hansen, A.; Grimme, S.; Sponer, J. *J. Chem. Theory Comput.* **2015**, *11*, 4972–4991.

[522] Grimme, S.; Steinmetz, M.; Korth, M. *J. Org. Chem.* **2007**, *72*, 2118–2126.

[523] Huenerbein, R.; Schirmer, B.; Moellmann, J.; Grimme, S. *Phys. Chem. Chem. Phys.* **2010**, *12*, 6940–6948.

[524] Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Pople, J. A. *J. Chem. Phys.* **2000**, *112*, 7374–7383.

[525] Otte, N.; Scholten, M.; Thiel, W. *J. Phys. Chem. A* **2007**, *111*, 5751–5755.

[526] Grimme, S. *J. Phys. Chem. A* **2005**, *109*, 3067–3077.

[527] Medvedev, M. G.; Bushmarinov, I. S.; Sun, J.; Perdew, J. P.; Lyssenko, K. A. *Science* **2017**, *355*, 49–52.

[528] Bytautas, L.; Ruedenberg, K. *J. Chem. Phys.* **2005**, *122*, 154110.

[529] Maniero, A. M.; Acioli, P. H. *Int. J. Quantum Chem.* **2005**, *103*, 711–717.

[530] Ásgeirsson, V.; Bauer, C. A.; Grimme, S. *Chem. Sci.* **2017**, *8*, 4879–4895.

[531] Schweez, C.; Shushkov, P.; Grimme, S.; Höger, S. *Angew. Chem. Int. Ed.* **2016**, *55*, 3328–3333; *Angew. Chem.* **2016**, *128*, 3389–3394.

[532] Grimme, S.; Hujo, W.; Kirchner, B. *Phys. Chem. Chem. Phys.* **2012**, *14*, 4875–4883.

[533] Rudberg, E.; Rubensson, E. H.; Sałek, P. *J. Chem Theory Comput.* **2011**, *7*, 340–350.

[534] http://www.rcsb.org/pdb.

[535] Pandey, A. S.; Harris, T. V.; Giles, L. J.; Peters, J. W.; Szilagyi, R. K. *J. Am. Chem. Soc.* **2008**, *130*, 4533–4540.

[536] Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2000**, *28*, 235–242.

[537] Sastry, G. M.; Adzhigirey, M.; Day, T.; Annabhimoju, R.; Sherman, W. *J. Comput.-Aided Mol. Des.* **2013**, *27*, 221–234.

[538] **Schrödinger Release 2016-4:** MS Jaguar, Schrödinger, LLC, New York, NY, 2016.

[539] Eckert, F.; Klamt, A. *AIChE J.* **2002**, *48*, 369–385.

[540] Masnyk, M.; Butkiewicz, A.; Górecki, M.; Luboradzki, R.; Bannwarth, C.; Grimme, S.; Frelek, J. *J. Org. Chem.* **2016**, *81*, 4588–4600.

[541] Reilly, A. M. et al. *Acta Crystallogr. Sect. B* **2016**, *72*, 439–459.

[542] Sierka, M.; Hogekamp, A.; Ahlrichs, R. *J. Chem. Phys.* **2003**, *118*, 9136–9148.

Part VI.

Appendix

A1. Supporting Information to Chapter 2

Appendix A1 contains:

- Short-hand notation for integrals
- Definition of operators
- A note on the Slater determinant

Short-Hand Notation

When computing the expectation value of an operator, the operator first acts on a function (to the right of the operator). Then the overlap of the resulting function with another function is computed, i.e., one multiplies by this function from the left and integrates over all variables of electron 1, which are condensed in the variable $\underline{1}$ here. We exemplify this for an arbitrary one-electron operator $\hat{O}(\mathbf{r}_1)$. The operator acting only on the Cartesian space variable $\mathbf{r}_1$ of the electron 1, this is denoted by

$$O_{(i\sigma)(j\tau)} = \int \psi^*_{i\sigma}(\underline{1})\hat{O}(\mathbf{r}_1)\psi_{j\tau}(\underline{1})\mathrm{d}\underline{1} = \left\langle \psi_{i\sigma} \middle| \hat{O} \middle| \psi_{j\tau} \right\rangle = \langle \sigma_i \mid \tau_j \rangle \left\langle \psi_i \middle| \hat{O} \middle| \psi_j \right\rangle \tag{A1.1a}$$

$$= \delta_{\sigma\tau} \int \psi_i(\mathbf{r}_1)\hat{O}(\mathbf{r}_1)\psi_j(\mathbf{r}_1)\mathrm{d}\mathbf{r}_1 = \delta_{\sigma\tau} \left\langle \psi_i \middle| \hat{O} \middle| \psi_j \right\rangle = \delta_{\sigma\tau} O_{ij} \,. \tag{A1.1b}$$

Here, we have introduced a short-hand notation for the expectation value or matrix element[52] of an operator $\hat{O}(\mathbf{r}_1)$. Since $\hat{O}(\mathbf{r}_1)$ acts only on the spatial part, the spin part (see Eq. 2.16) can be factorized and the expectation value reduces to an expectation value over spatial orbitals $\psi_i(\mathbf{r}_1)$ and $\psi_j(\mathbf{r}_1)$ times the overlap of the spin function, hence the Kronecker delta $\delta_{\sigma\tau}$. While an operator may in principle act on all variables, the operators in the Hamiltonian (see Eq. 2.1) at most act on two electron variables at a time. The Fock operator and all elements thereof act only on one electron and only on the spatial variables.

The notation introduced above is often called the "bra-ket" notation.[52]. The "ket" corresponds to

$$|\psi_{i\sigma}\rangle \equiv \psi_{i\sigma}(\underline{1}) = \sigma_i\psi_i(\mathbf{r}_1) \equiv |\sigma_i\rangle\,|\psi_i\rangle \,, \tag{A1.2}$$

while the "bra" is corresponds to the complex conjugate

$$\langle\psi_{i\sigma}| \equiv \psi^*_{i\sigma}(\underline{1}) = \sigma^*_i\psi^*_i(\mathbf{r}_1) \equiv \langle\sigma_i|\,\langle\psi_i| \,. \tag{A1.3}$$

Whenever the "bra-ket" product is formed, possibly with an operator in between as in Eq. A1.1, the short-hand notation implies integration over all variables. In case the "ket" denotes a matrix, then the respective "bra" is given by the adjoint (i.e., complex conjugate of the transposed "ket" matrix).

Definition of Operators

Terms of the All-Electron Hamilton Operator

Atomic units are used throughout, i.e., all equations are given in terms of the electron mass m_e, the elemental charge $|e|$, the Bohr radius a_0, and $\hbar$, which is Planck's constant divided by 2π. The kinetic energy operator is given as:

$$\hat{T}_e \equiv \sum_i^N \hat{T}_e(\mathbf{r}_i) = \sum_i^N \frac{\nabla_i^2}{2} \tag{A1.4}$$

with

$$\nabla_i^2 = \frac{\partial^2}{\partial x_i^2} + \frac{\partial^2}{\partial y_i^2} + \frac{\partial^2}{\partial z_i^2} \tag{A1.5}$$

and the electron position vector

$$\mathbf{r}_i = \begin{pmatrix} x_i \\ y_i \\ z_i \end{pmatrix} . \tag{A1.6}$$

The Coulomb interaction operator between the nuclei and the electrons is given by:

$$\hat{V}_{ne} \equiv \sum_i^N \hat{V}_{ne}(\mathbf{r}_i) = -\sum_i^N \sum_A^{N_n} \frac{Z_A}{|\mathbf{r}_i - \mathbf{r}_A|} . \tag{A1.7}$$

The electron-electron Coulomb operator is defined as:

$$\hat{V}_{ee} \equiv \frac{1}{2} \sum_i^N \hat{V}_{ee}(\mathbf{r}_i) = \frac{1}{2} \sum_i^N \sum_{j \neq i}^N \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} . \tag{A1.8}$$

The Coulomb interaction between the nuclei is given by:

$$E_{nn} = \frac{1}{2} \sum_A^{N_n} \sum_{B \neq A}^{N_n} \frac{Z_A Z_B}{|\mathbf{r}_A - \mathbf{r}_B|} . \tag{A1.9}$$

Terms of the Fock Operator

In the following, the elements of the Fock operator acting on an MO $\psi_{i\sigma}(\underline{1})$ (see Eq. 2.17) are listed. The one-electron operator consists of the kinetic energy operator and the nuclear Coulomb interaction operator

$$\hat{h}(\mathbf{r}_1)\psi_{i\sigma}(\underline{1}) = \sigma_i \hat{h}(\mathbf{r}_1)\psi_i(\mathbf{r}_1) = \sigma_i \left[\hat{T}_e(\mathbf{r}_1) + \hat{V}_{ne}(\mathbf{r}_1)\right] \psi_i(\mathbf{r}_1) . \tag{A1.10}$$

The Coulomb operator is given by

$$\begin{aligned} \hat{J}_{j\tau}(\mathbf{r}_1)\psi_{i\sigma}(\underline{1}) &= \langle \tau | \tau \rangle \int \frac{\psi_j^*(\mathbf{r}_2)\psi_j(\mathbf{r}_2)}{|\mathbf{r}_2 - \mathbf{r}_1|} \mathrm{d}\mathbf{r}_2 \psi_{i\sigma}(\underline{1}) \\ &= \int \frac{\psi_j(\mathbf{r}_2)\psi_j(\mathbf{r}_2)}{|\mathbf{r}_2 - \mathbf{r}_1|} \mathrm{d}\mathbf{r}_2 \psi_{i\sigma}(\underline{1}) \equiv \hat{J}_j(\mathbf{r}_1)\psi_{i\sigma}(\underline{1}) \end{aligned} \tag{A1.11}$$

The exchange operator vanishes if the two orbitals i and j describe MOs of different spin

$$\hat{K}_{j\tau}(\mathbf{r}_1)\psi_{i\sigma}(\underline{1}) = \delta_{\sigma\tau} \int \frac{\psi_j^*(\mathbf{r}_2)\psi_i(\mathbf{r}_2)}{|\mathbf{r}_2 - \mathbf{r}_1|} \mathrm{d}\mathbf{r}_2 \psi_{j\tau}(\underline{1}). \tag{A1.12}$$

A Note on the Slater Determinant

$$\Phi_K(\underline{1},\underline{2},\cdots,\underline{\mathrm{N}}) = \frac{1}{\sqrt{N!}}\sum_{m}^{N!}(-1)^{p_m}\hat{P}_m\prod_{i}^{N}\psi_{i\sigma}(\underline{\mathrm{i}})\,. \tag{A1.13}$$

Another way to express the HF wave function (i.e., the Slater determinant in Eq. 2.15) is given by a sum over products of one-particle functions.[52] Here, the term on the right corresponds to the naturally ordered product, that is the product of diagonal entries of the matrix corresponding to the Slater determinant. The preceding sum runs over all $N!$ possible permutations of electron variables. $\hat{P}_m$ is the permutation operator, which generates the m^{th} permutation by interchanging the corresponding electron variables. p_m is the number of transpositions (exchange of two-electrons) required to arrive at the naturally ordered product, which then defines the sign of that respective term in the sum. The expression in Eq. A1.13 is referred to as an anti-symmetrized product and equivalent to the determinant representation.

A2. Supporting Information to Chapter 3

Appendix A2 contains:

- Energy data

Energy data (in a.u.)

Table A2.1.: Electronic energy contributions for the association to FLP **1**.

method	$B(C_6F_5)_3$	$PMes_3$	**1a**	**1b**
def2-QZVP				
B3LYP[a]	−2208.39036	−1389.83868	−3598.22025	−3598.22139
-D3	−0.08533	−0.15809	−0.26900	−0.26819
B3LYP[b]	−2208.39033	−1389.83860	−3598.21988	−3598.22107
-NL	1.30288	1.02444	2.29963	2.30005
M06	−2208.34668	−1389.77122	−3598.12759	−3598.12799
-D3(0)	−0.00738	−0.01570	−0.03466	−0.03423
PW6B95	−2211.84647	−1392.25623	−3604.10433	−3604.10637
-D3	−0.02767	−0.05352	−0.09516	−0.09491
PBE0	−2207.09265	−1389.17201	−3596.26162	−3596.26285
-D3	−0.04486	−0.08587	−0.14919	−0.14884
B2PLYP	−2208.34668	−1389.83237	−3598.18389	−3598.18452
-D3	−0.04047	−0.07542	−0.12949	−0.12911
HF	−2198.72701	−1382.80809	−3581.52281	−3581.52547
-D3	−0.31955	−0.54275	−0.89079	−0.88886
MP2	−2205.77045	−1387.71245	−3593.50669	−3593.50618
SCS-MP2	−2205.57897	−1387.65853	−3593.25392	−3593.25398
aug-cc-pVDZ				
HF	−2198.00259	−1382.49161	−3580.49117	−3580.49328
δCP	-	-	+0.00642	+0.00784
$MP2_{corr.}$[c]	−5.44402	−3.98945	−9.48361	−9.47937
δCP	-	-	+0.02077	+0.01814
aug-cc-pVTZ				
HF	−2198.55710	−1382.73504	−3581.28255	−3581.28520
δCP	-	-	−0.00014	+0.00185
$MP2_{corr.}$[c]	−6.68477	−4.71015	−11.43807	−11.43478
δCP	-	-	+0.01272	+0.01100
$DLPNO\text{-}CCSD(T)_{corr.}$[c]	−6.91708	−5.04638	−11.99700	−11.99574

[a] RI-J approximation employed. [b] RI-J-COSX approximation employed.

[c] Only correlation energy is given.

Table A2.2.: Electronic energy contributions for the association to FLP **2**.

method	$B(C_6F_5)_3$	$PtBu_3$	**2a**	**2b**
def2-QZVP				
B3LYP[(a)]	−2208.39036	−814.71685	−3023.10270	−3023.10295
-D3	−0.08533	−0.08650	−0.19414	−0.18830
B3LYP[(b)]	−2208.39033	−814.71685	−3023.10252	−3023.10285
-NL	1.30288	0.55756	1.83559	1.84225
M06	−2208.34668	−814.73353	−3023.09254	−3023.08640
-D3(0)	−0.00738	−0.00774	−0.02456	−0.02302
PW6B95	−2211.84647	−816.00355	−3027.85717	−3027.85296
-D3	−0.02767	−0.02747	−0.06716	−0.06424
PBE0	−2207.09265	−814.37794	−3021.47135	−3021.47011
-D3	−0.04486	−0.04635	−0.10720	−0.10326
B2PLYP	−2208.34668	−814.69815	−3023.05193	−3023.04907
-D3	−0.04047	−0.04053	−0.09283	−0.08978
HF	−2198.72701	−811.04822	−3009.76855	−3009.76984
-D3	−0.31955	−0.29301	−0.63724	−0.63055
MP2	−2205.77045	−813.50702	−3019.29988	−3019.29307
SCS-MP2	−2205.57897	−813.52699	−3019.12229	−3019.11724
aug-cc-pVDZ				
HF	−2198.00259	−810.88385	−3008.88722	−3008.88680
δCP	-	-	+0.00595	+0.00520
$MP2_{corr.}$[(c)]	−5.44402	−1.99621	−7.48056	−7.47051
δCP	-	-	+0.01603	+0.01268
aug-cc-pVTZ				
HF	−2198.55710	−811.01031	−3009.56306	−3009.56393
δCP	-	-	+0.00092	+0.00134
$MP2_{corr.}$[(c)]	−6.68477	−2.36325	−9.08302	−9.07386
δCP	-	-	+0.00944	+0.00753
$DLPNO\text{-}CCSD(T)_{corr.}$[(c)]	−6.91708	−2.59207	−9.53486	−9.52940

[(a)] RI-J approximation employed. [(b)] RI-J-COSX approximation employed.
[(c)] Only correlation energy is given.

Table A2.3.: Corrections to the free energy of the association to FLP **1**.

	$B(C_6F_5)_3$	$PMes_3$	**1a**	**1b**
$G^{298\,K}_{RRHO}$ [(a)]	0.08987	0.46489	0.57807	0.57839
$\delta G^{298\,K}_{solv.}(bz)$ [(b)]	−0.01221	−0.02349	−0.03627	−0.03757
$\delta G^{298\,K}_{solv.}(bz)$ [(c)]	0.00150	−0.01488	−0.01200	−0.01214

[(a)] Ro-vibrational correction (including zero-point vibrational energy) at the HF-3c level.
[(b)] Obtained from COSMO-RS(2012) with BP86/TZVP.
[(c)] Obtained from SMD(COSMO) with BP86/def2-TZVP.

Table A2.4.: Corrections to the free energy of the association to FLP **2**.

	$B(C_6F_5)_3$	$PtBu_3$	**2a**	**2b**
$G^{298\,K}_{RRHO}$ [(a)]	0.08987	0.33294	0.44555	0.44429
$\delta G^{298\,K}_{solv.}(bz)$ [(b)]	−0.01221	−0.00704	−0.02060	−0.02277
$\delta G^{298\,K}_{solv.}(bz)$ [(c)]	0.00150	−0.00589	−0.00288	−0.00378

[(a)] Ro-vibrational correction (including zero-point vibrational energy) at the HF-3c level.
[(b)] Obtained from COSMO-RS(2012) with BP86/TZVP.
[(c)] Obtained from SMD(COSMO) with BP86/def2-TZVP.

A3. Supporting Information to Chapter 4

Appendix A3 contains:

- Computational Details
- UV-Vis and ECD spectra of the metallosupramolecular spheres
- Dihedral angles of the naphthyl groups of **1** and **2** and their metal complexes

Computational Details

The geometry optimizations and the single-point calculations have been performed on the DFT level with the TURBOMOLE suite of programs.[164a,165] We exploited the *O*-symmetry of both complexes and used the COSMO solvation model ($\epsilon = 35.7$) in both, geometry optimizations and the single-point calculations.[190] In order to speed up the computations of the Coulomb integrals, the multipole-accelerated resolution of the identity (MARI-J) approximation has been used.[166–168,542] The numerical quadrature grid *m4* has been employed for the integration of the exchange-correlation contribution of the density functional. The SCF convergence criterion was set to 10^{-6}. For the geometry optimizations, we used the PBE density functional[104] along with the split-valence def2-SV(P) basis set of Ahlrichs and coworkers.[163] Additionally, the atom pair-wise DFT-D3 (with BJ-damping) correction to account for medium- and long-range (London) dispersion interactions[133,134] was applied. Due to the size of the systems, and since we are not interested in, e.g., vibrational frequencies, the convergence criterion for geometries was increased to 10^{-2} E_h. The geometries were then sufficiently optimized for our purposes of UV/ECD-spectra computation. We will refer to the method described above as PBE-D3/def2-SV(P). For each complex, one additional geometry was optimized by applying the methods described above but at the same time, we included the geometrical counterpoise correction (gCP).[264] The calculation of the gCP correction and its gradient was performed with the stand-alone program, which was developed in our group.[306] We will refer to this structure by PBE-D3-gCP/def2-SV(P). While for $[Pd_6\{(M)\text{-}\mathbf{1}\}_{12}]^{12+}$, both optimizations lead to quite similar geometries, this is not the case for $[Pd_{12}\{(P)\text{-}\mathbf{2}\}_{24}]^{24+}$. Here, the structures differ significantly, especially in the orientation of the acetal sidechains and the dihedral angles of the pyridinyl ligands. We consider both structures as two extrema of possible conformations in solution. Due to the size and complexity of the systems, a further extensive conformational search was impossible.

The two geometries obtained for each complex were used to conduct single-point calculations using the BHLYP functional[69] along with the def2-SV(P) basis set. Even though the high amount of Fock exchange (50%) will likely cause overestimated excitation energies, we prefer BHLYP over other density functionals as higher Fock exchange will reduce the typical functional errors in treating charge-transfer states. Range-separated functionals as an alternative are currently not implemented in TURBOMOLE but should be applied in the future for such huge systems.

Excitation energies and the spectroscopic data* were calculated using the recently published simplified

*In the computed absorption (plotted in Figures A3.1 and A3.2) and ECD spectra (plotted in Figures 4.6b, A3.3, and A3.4), the magnitudes of the respective intensities are decreased by a factor of $\sqrt{2}$ and 2, respectively.

Tamm-Dancoff approximation to TD-DFT (sTDA-DFT).[46] The energy cutoff to select the primary configurations (see Ref. 46) was set to 7.5 eV and all configurations with diagonal elements higher than 15 eV have been neglected completely. The calculated rotatory strengths refer to the length formalism. The calculated spectrum given in Figure 4.6b (see main text) is then obtained by a 1:1 average of the spectra obtained on the PBE-D3/def2-SV(P) and PBE-D3-gCP/def2-SV(P) geometries of $[Pd_6\{(M)\text{-}\mathbf{1}\}_{12}]^{12+}$ (see Figure A3.4). The same procedure was done for the respective $[Pd_{12}\{(P)\text{-}\mathbf{2}\}_{24}]^{24+}$ geometries and the resulting spectrum is given in Figure A3.3. In the same manner, we calculated UV-Vis absorption spectra which are given for both complexes in Figures A3.1 and A3.2. Comparison with the experimental UV-Vis absorption spectra shows the same shifting of bands as was already observed in the ECD spectra. This further supports our statement that calculated bands below 250 nm are blue-shifted which needs to be taken into account when assigning the experimental bands to the calculated ones.

UV-Vis and ECD spectra of the metallosupramolecular spheres

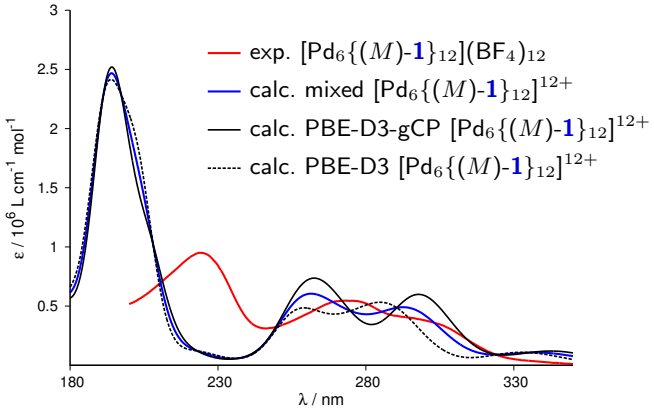


Figure A3.1.: Experimental UV-Vis spectrum of $[Pd_6\{(M)\text{-}\mathbf{1}\}_{12}](BF_4)_{12}$ in acetonitrile and simulated UV-Vis spectra of $[Pd_6\{(M)\text{-}\mathbf{1}\}_{12}]^{12+}$.

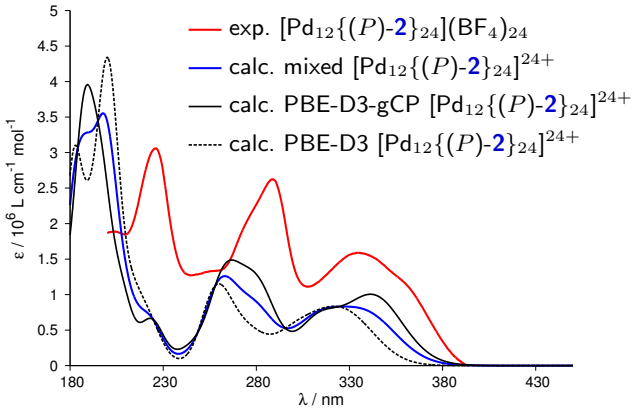


Figure A3.2.: Experimental UV-Vis spectrum of $[Pd_{12}\{(P)\text{-}\mathbf{2}\}_{24}](BF_4)_{24}$ in acetonitrile and simulated UV-Vis spectra of $[Pd_{12}\{(P)\text{-}\mathbf{2}\}_{24}]^{24+}$.

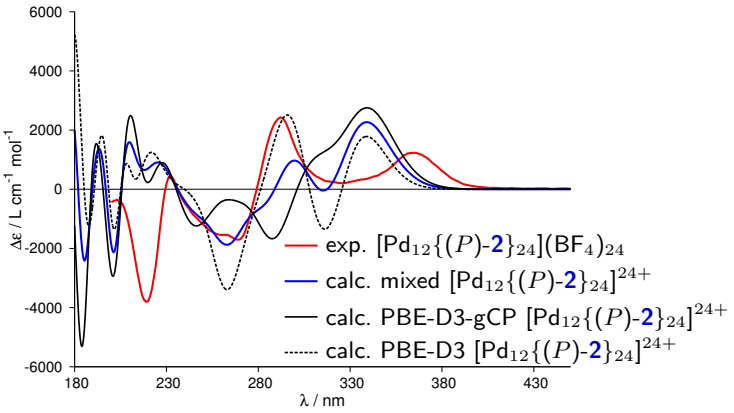


Figure A3.3.: Experimental ECD spectrum of $[Pd_{12}\{(P)\text{-}\mathbf{2}\}_{24}](BF_4)_{24}$ in acetonitrile and simulated ECD spectra of $[Pd_{12}\{(P)\text{-}\mathbf{2}\}_{24}]^{24+}$.

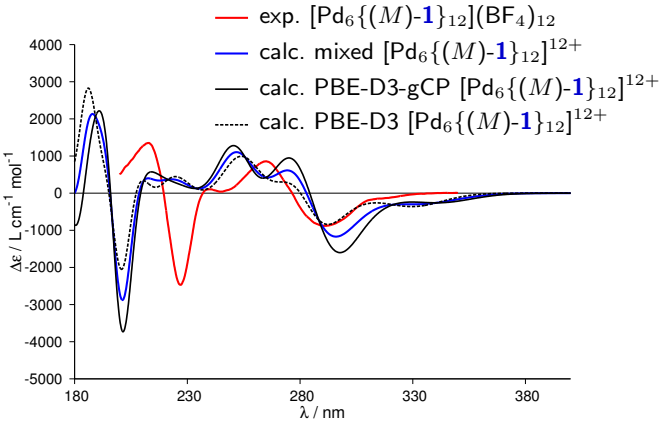


Figure A3.4.: Experimental ECD spectrum of $[Pd_6\{(M)\text{-}\mathbf{1}\}_{12}](BF_4)_{12}$ in acetonitrile and simulated ECD spectra of $[Pd_6\{(M)\text{-}\mathbf{1}\}_{12}]^{12+}$.

Dihedral angles of the naphthyl groups of 1 and 2 and their metal complexes

Table A3.1.: Dihedral angles of the two naphthyl groups along the axis of chirality of the "free" binaphtyl-type ligands (M)-**1** and (P)-**2** in acetonitrile ($\epsilon = 35.7$) and in the respective complexes (differences are shown in parentheses).

Method	(M)-**1**	$[Pd_6\{(M)\text{-}\mathbf{1}\}_{12}]^{12+}$	(P)-**2**	$[Pd_{12}\{(P)\text{-}\mathbf{2}\}_{24}]^{24+}$
PBE-D3/def2-SV(P)	−120.9°	−77.0° (43.9°)	112.5°	109.5° (−3.0°)
PBE-D3-gCP/def2-SV(P)	−120.9°	−78.5° (42.4°)	112.5°	101.7° (−10.8°)

A4. Supporting Information to Chapter 6

Appendix A4 contains:

- ECD spectrum of (*P*)-[11]helicene computed with sTD(A)-ωB97X-D3/def2-SV(P)

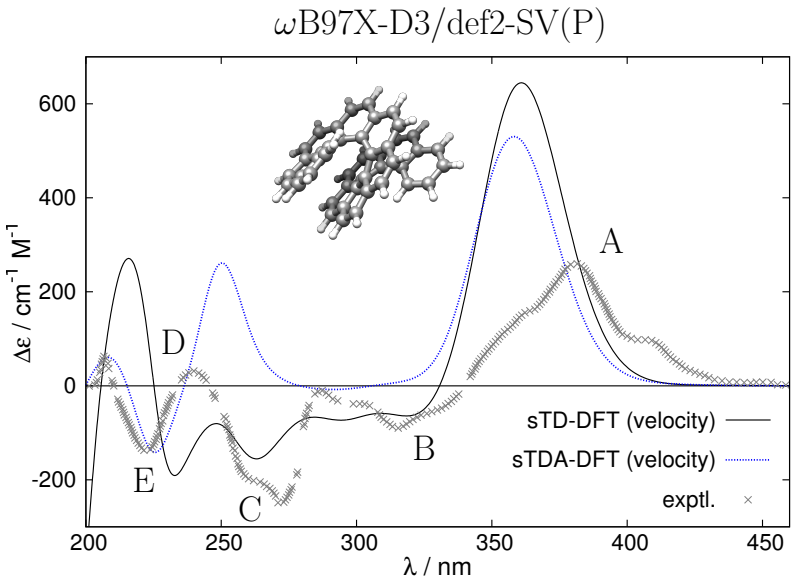


Figure A4.1.: Calculated ECD spectra of (*P*)-[11]helicene computed with the range-separated hybrid functional ωB97X-D3. The spectra are obtained from the velocity representation of the rotatory strengths (R^V). The black line refers to sTD-DFT and the dotted blue line to sTDA-DFT. The experimental curve (gray points) has been redrawn from Ref. 304.

A5. Supporting Information to Chapter 7

Appendix A5 contains:

- Technical details of the calculations
- Supplementary figures
  - The absorption spectra of **[16]H**
  - ECD intensity versus number of annulated benzene rings
  - Individual spectra on MD geometries

Technical Details of the Calculations

The geometries of (*P*)-[16]helicene ((*P*)-**[16]H**) and its derivative with two triisopropylsilyloxy protection groups (*P*)-**[16]H-tipso** were optimized at the efficient PBEh-3c level[110] using the TURBOMOLE suite of programs (v7.0).[164c,165,371] The resolution of the identity approximation with corresponding auxiliary basis sets was used throughout for the Coulomb integrals[76,167,168] and the exchange-correlation energy was evaluated employing the numerical quadrature grid *m4*. The structures are given below and were verified as minima by harmonic frequency calculations. The C_2 symmetry of the compounds was used in all TURBOMOLE calculations.

For the MD runs, a quantum mechanical derived force field (QMDFF)[49] was for the helicenes as well as the solvent molecule ($CHCl_3$) at the same level of theory. The QMDFF parameters were fitted on the scaled Hessian (by 0.95) obtained at the PBEh-3c level.[110] Then for each (*P*)-**[16]H** and (*P*)-**[16]H-tipso**, a molecular dynamics (MD) simulation under particle-conserving, isobaric, and isothermal conditions (NPT) was conducted. The solute was placed in a cubic box with 250 $CHCl_3$ molecules and periodic boundary conditions were employed. This was done with the QMSIM program.[375] The Pauli repulsion terms were scaled up by a factor of 0.85, since this value yielded the correct experimental density for a periodic box of 50 $CHCl_3$ at 300 K.

After equilibrating at 300 K, an MD simulation was performed for 0.5 ns with a timestep of 0.5 fs. All C–H bond lengths are kept constrained during the simulation using the SHAKE algorithm.[376,377] A structure was extracted every 5 ps and used for the subsequent simplified time-dependent density functional theory calculation (sTD-DFT).

The sTD-DFT calculations on both the minimum geometries as well as the structures from the MD trajectory were performed either on a ωB97X[370] or BHLYP (i.e., Becke's *half-and-half* functional)[69,99,100] reference using a development version of the ORCA program (based on v3.0).[48] The def2-TZVP basis set was used throughout[162,163] and the numerical quadrature *grid4* was employed.

Supplementary Figures

The Absorption Spectra of [16]H

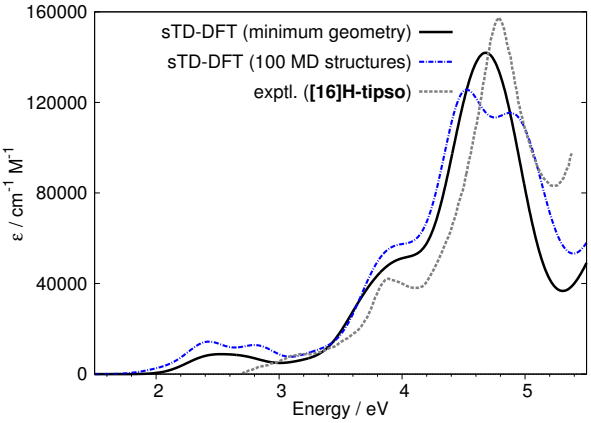


Figure A5.1.: Calculated absorption spectrum of **[16]H** computed with sTD-ωB97X/def2-TZVP[162,163,370] on the minimum geometry (solid black line) and 100 structures obtained from a MD trajectory (blue dashed line). Oscillator strengths are broadened by Gaussians with a full width at $1/e$ maximum of 0.5 eV (0.35 eV in the MD case) and the spectra are red-shifted by 1 eV. The experimental absorption spectrum (gray dotted line) for **[16]H-tipso** is given as well and taken from Ref. 350.

ECD intensity versus number of annulated benzene rings

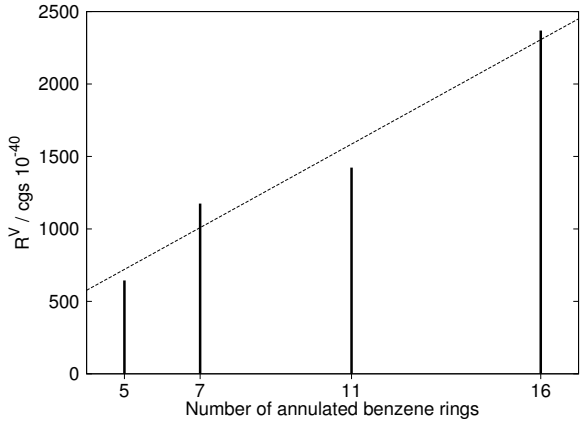


Figure A5.2.: The rotatory strength (velocity representation) in 10^{-40} erg·cm^3 for the ECD dominant ($R^V > 100 \cdot 10^{-40}$ erg·cm^3) transition in the respective lowermost part of the spectrum of (*P*)-[5]helicene, (*P*)-[7]helicene, and (*P*)-[11]helicene. They are computed by sTD-ωB97X/def2-TZVP on PBEh-3c structures. For (*P*)-[16]helicene, this region is dominated by two transitions (like for (*P*)-**[16]H-tipso**, cf. Figure 7.2). The linear regression (root at the origin) yields a slope of $144.1 \pm 6.9 \cdot 10^{-40}$ erg·cm^3.

Individual Spectra on MD Geometries

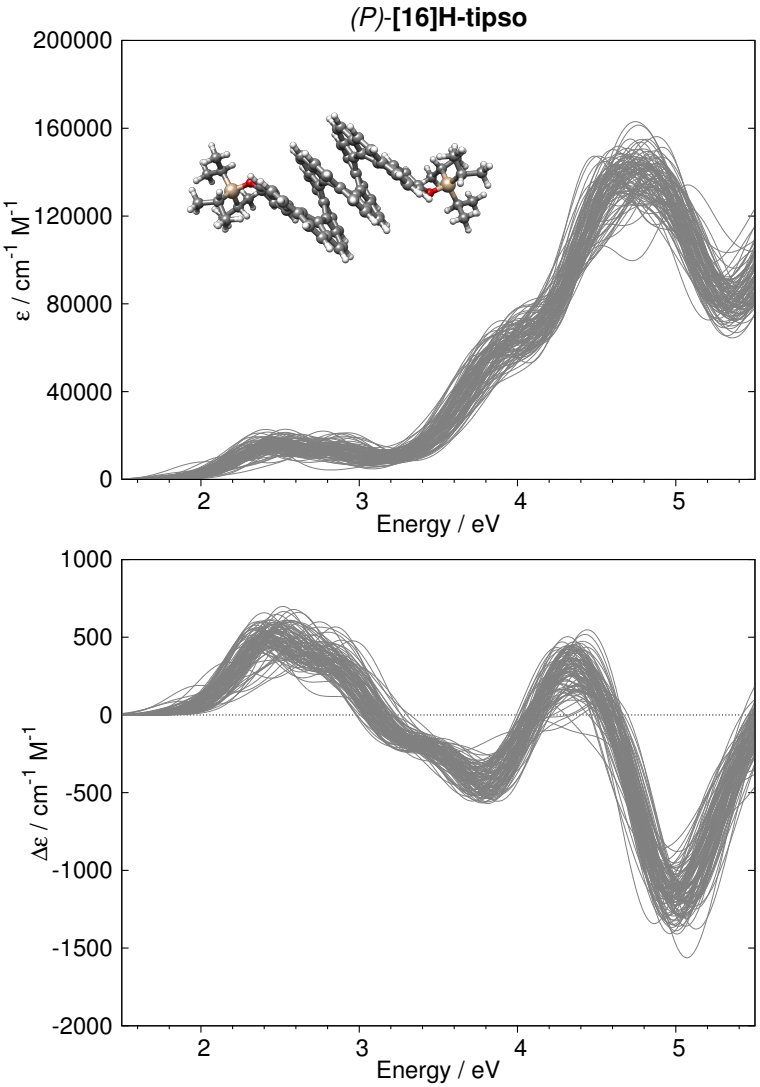


Figure A5.3.: Overlay of individual absorption (top) and ECD (bottom) spectra computed with sTD-ωB97X/def2-TZVP[162,163,370] on the MD geometries of (*P*)-**[16]H-tipso**. The individual transition strengths are broadened by Gaussians with a full width at $1/e$ maximum of 0.5 eV and the spectra are red-shifted by 1 eV.

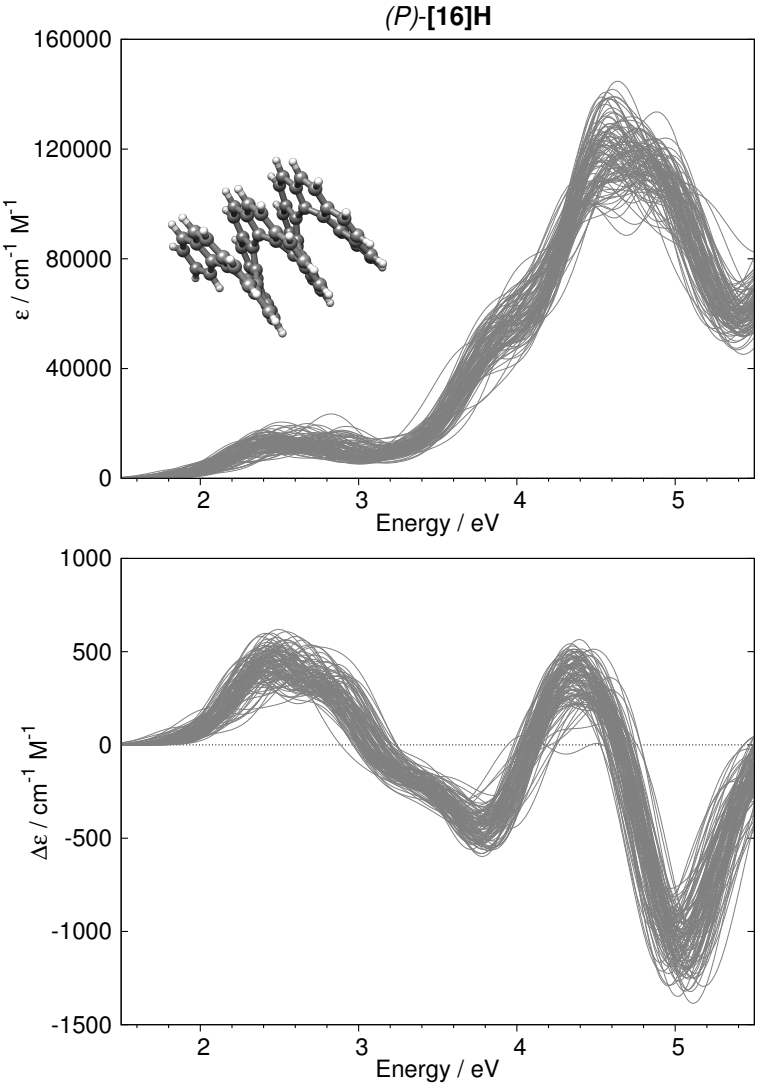


Figure A5.4.: Overlay of individual absorption (top) and ECD (bottom) spectra computed with sTD-ωB97X/def2-TZVP [162,163,370] on the MD geometries of (*P*)-**[16]H**. The individual transition strengths are broadened by Gaussians with a full width at $1/e$ maximum of 0.5 eV and the spectra are red-shifted by 1 eV.

A6. Supporting Information to Chapter 8

Appendix A6 contains:

- Supplementary figures
- Wall-clock timings for (s)TD-DFT procedures

Supplementary Figures

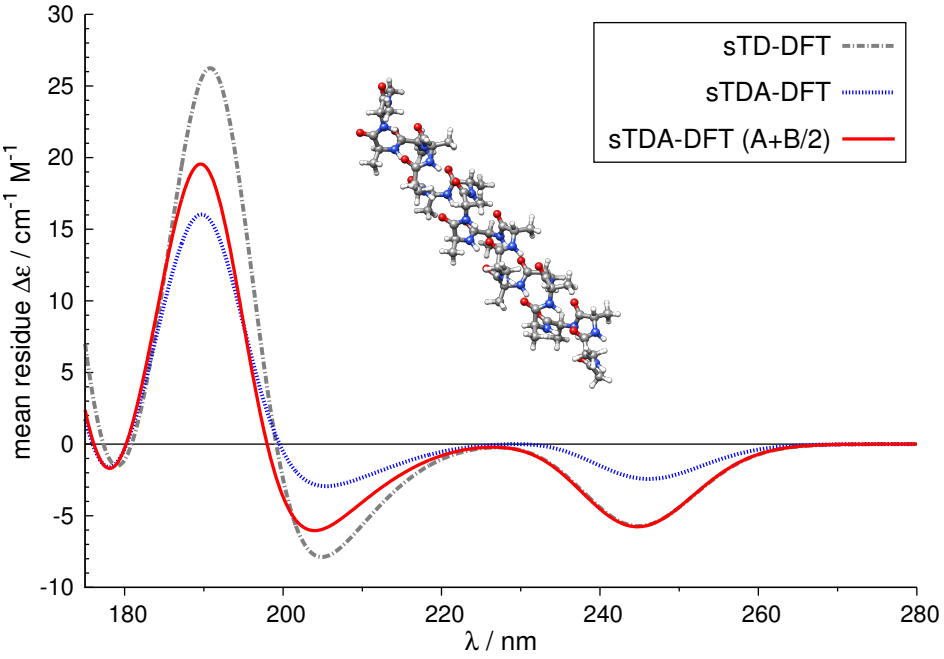


Figure A6.1.: Calculated electronic circular dichroism (ECD) spectrum of α-helical Ace-Ala$_{19}$-NMe computed with sTD-DFT[329] (gray dashed), sTDA-DFT[46] (blue dotted), and A+B/2 corrected sTDA-DFT (red solid) based on a BHLYP/def2-SV(P)[69,163,300] reference (TPSS-D3/def2-TZVP geometry[105,133,134,162,163] from Ref.[362]). All spectra are obtained by convoluting the rotatory strengths (velocity representation) with Gaussian curves with full width at $1/e$ maximum of 0.4 eV and shifted by -1 eV.

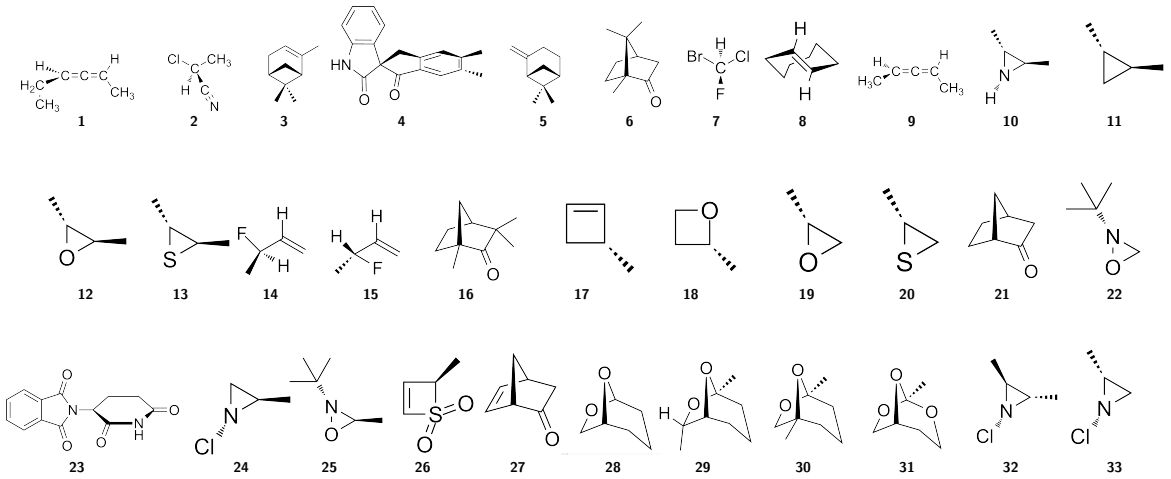


Figure A6.2.: Lewis structures of the systems considered in the computation of the specific rotation at 589 nm (see Figures A6.4 and 8.4b). Note that systems 14 and 15 correspond to different conformers of the same molecule. The Cartesian coordinates (optimized at the PBEh-3c level)[110] are given below.

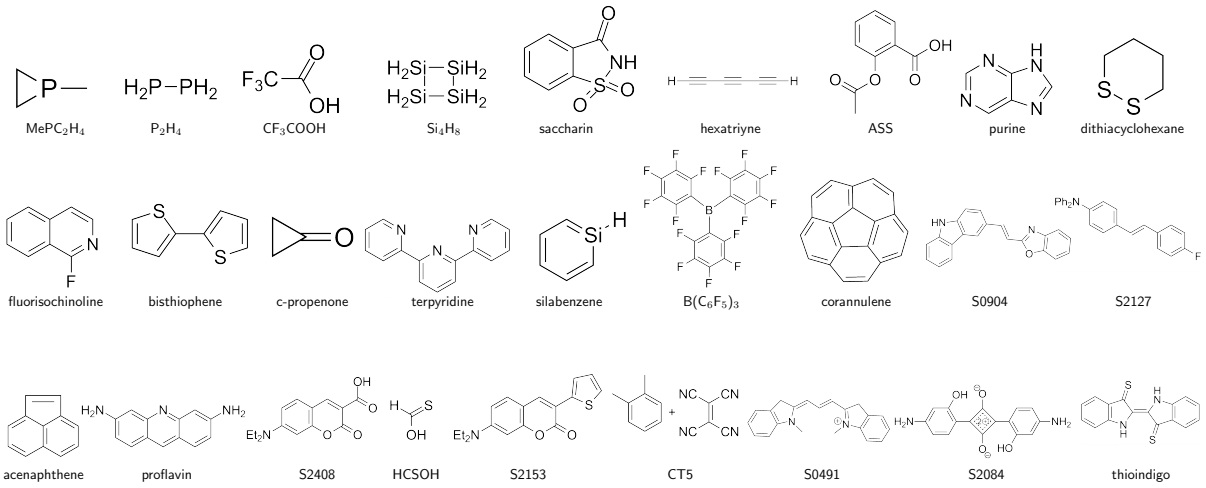


Figure A6.3.: Lewis structures of the systems considered in the cross-check set (see Table 8.5). The Cartesian coordinates (optimized at the PBEh-3c level)[110] are given below.

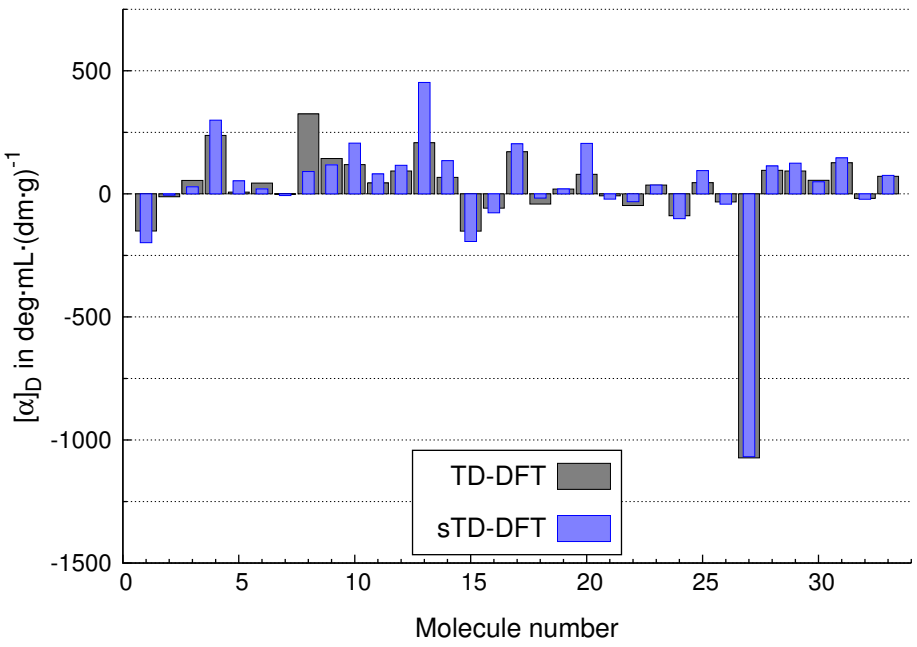


Figure A6.4.: Specific rotations at 589 nm (excluding the Lorentz prefactor) of 33 molecules computed via sTD-DFT (blue, velocity representation in the sum-over-states expression including all states up to 60 eV) and TD-DFT (gray, length representation). The same PBE0/TZVP [104,111,162] reference has been used for both.

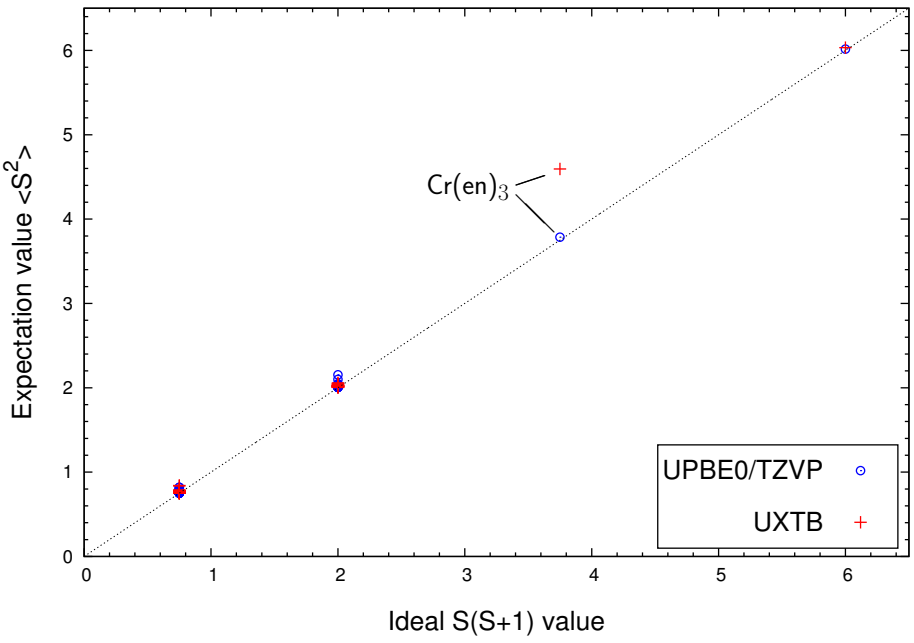


Figure A6.6.: Correlation plot of computed expectation value $\langle S^2 \rangle$ from unrestricted PBE0/TZVP and XTB vs. the ideal value of S(S+1).

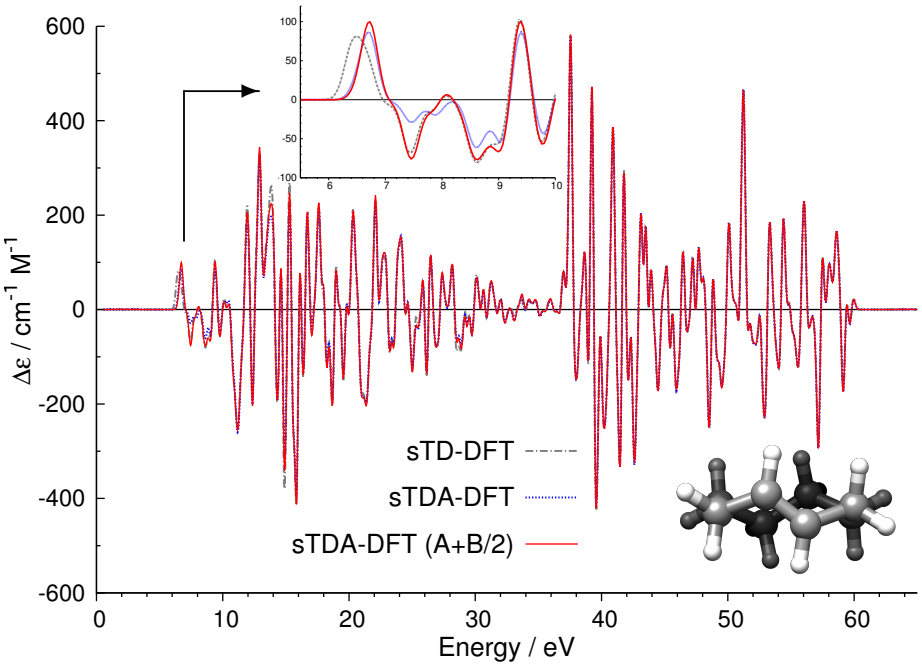


Figure A6.5.: Electronic circular dichroism (ECD) spectrum of *E*-cyclooctene computed by sTD-DFT (gray dashed), sTDA-DFT (blue dotted), and A+B/2 corrected sTDA-DFT (red solid) based on a PBE0/TZVP//PBEh-3c reference.

Wall-clock timings for (s)TD-DFT calculations

Table A6.1.: Wall-clock timings t in seconds for the parts and the entire (s)TD-BHLYP/def2-SV(P) calculation on a (single-core) laptop computer. For sTD-DFT excited state treatments, the spectral energy range E_{max} (in eV) considered is also given. Symmetry was not exploited.

molecule	number of atoms	number of states	E_{max} /eV	t / s SCF	TD-DFT
indigo	30	30	-	94	1369
[16]helicene	102	5	-	4044	32007
poly-alanine$_{20}$	203	5	-	3774	7650
				SCF	sTD-DFT
indigo	30	71	9	94	1.6
[16]helicene	102	723	9	4044	63
poly-alanine$_{20}$	203	741	9	3774	32

A7. Supporting Information to Chapter 9

Appendix A7 contains:

- Tables with detailed results
- Overlay of 10 larger transition metal complexes optimized by PM6-D3H4X and PBEh-3c

Detailed Results

Table A7.1.: Comparison of computed rotational constants of twelve medium sized molecules to experimentally derived ones (ROT34)[a] for different semiempirical methods. The values are given in MHz.

		GFN-xTB	PM6-D3H4X	MSINDO-D3(BJ)H+	OM2-D3(BJ)	DFTB3-D3(BJ)	HF-3c	ref.
1	A	4264.6	4190.4	4298.4	4266.8	4250.8	4218.0	4293.9
	B	1401.7	1395.0	1423.5	1419.5	1383.1	1380.4	1395.9
	C	1134.7	1125.4	1149.7	1151.9	1118.8	1115.8	1130.2
2	A	3130.3	2577.5[b]	3255.3	2511.1[b]	3196.4	3172.3	3322.5
	B	733.1	869.7[b]	783.1	979.0[b]	714.6	715.6	719.8
	C	702.8	852.0[b]	737.2	940.9[b]	690.8	691.4	698.0
3	A	3067.5	3017.8	3127.5	3151.7	3009.6	2991.7	3071.1
	B	1289.1	1306.1	1359.0	1353.4	1251.4	1262.1	1285.0
	C	1252.9	1211.8	1215.4	1187.9	1223.1	1207.2	1248.7
4	A	2771.2	2735.0	2823.3	2798.3	2725.0	2710.6	2755.9
	B	2682.7	2647.0	2708.1	2688.9	2654.7	2648.5	2675.6
	C	2665.2	2608.6	2691.8	2654.6	2637.3	2647.7	2653.3
5	A	2332.9	2272.9	2365.7	2372.4	2299.5	2297.4	2336.9
6	A	1452.7	1374.7	1416.4	1438.4	1432.1	1427.9	1464.2
	B	774.9	811.6	826.7	830.0	761.7	756.6	768.2
	C	587.2	599.3	602.3	615.8	575.5	572.0	580.6
7	A	1171.6	1167.4	1200.4	1195.8	1164.4	1161.4	1165.7
	B	668.3	628.4	628.0	651.2	642.6	657.0	661.2
	C	460.5	442.5	448.5	458.1	449.3	454.8	454.0
8	A	1204.6	1145.9	1302.8	1265.1	1204.0	1164.5	1166.3
	B	757.9	744.5	764.0	767.0	748.2	749.8	767.6
	C	520.5	497.3	532.0	526.4	512.5	504.0	513.0
9	A	878.3	881.1	899.2	900.8	861.0	863.3	862.5
	B	744.9	730.5	750.3	752.8	734.7	737.4	754.2
	C	513.8	511.7	525.4	527.1	506.1	508.2	513.7
10	A	3074.9	3037.5	3127.3	3072.6	3062.3	3060.6	3086.2
	B	727.3	728.0	738.3	738.5	718.7	717.0	723.7
	C	684.2	674.6	682.0	695.5	678.4	677.1	685.0
11	A	1449.3	1390.2	1463.5	1476.1	1417.2	1424.6	1432.1
	B	823.2	792.9	813.4	821.2	809.0	807.4	820.5
	C	689.1	656.9	692.7	695.1	668.6	678.4	679.4
12	A	1500.6	1436.9	1427.4	1493.3	1452.6	1495.6	1523.2
	B	1089.9	1049.3	1137.4	1081.8	1066.1	1049.0	1070.5
	C	739.9	691.2	732.6	717.6	702.3	710.2	719.9

[a] Rotational constants B_e (excluding vibrational effects) from Ref. 504 with an estimated reference error of 0.2%.

[b] A different conformer than in the experiment is obtained. Thus this value is neglected in the statistical analysis of the data set presented in the manuscript.

1: ethynyl-cyclohexane, **2**: isoamyl-acetate, **3**: diisopropyl-ketone, **4**: bicyclo[2.2.2]octadiene, **5**: triethylamine, **6**: vitamin C, **7**: serotonine, **8**: aspirin, **9**: cassyrane, **10**: proline, **11**: lupinene, and **12**: limonene.

Table A7.2.: Untypically long intramolecular bonds (LB12)[a] obtained by geometry optimizations with different semiempirical methods in comparison to experimental values. The values are given in pm.

system	bond	GFN-xTB	PM6-D3H4X	MSINDO-D3(BJ)H+	HF-3c	ref.
DIAD	C–C	167.3	165.8	162.5	166.3	171
FLP	P–B	210.2	214.4	197.6	209.9	212
DTFS	Si–N	207.3	276.6	235.1	271.6	227
MESITRAN	Si–N	226.7	212.4	273.3	233.1	245
S_8^{2+}	S–S	230.3	221.8	211.4	224.8	286
HAPPOD	Rh–Cr	308.2	555.5[b]	265.3[b]	311.4	308
KAMDOR	Os–Cr	313.1	414.6[b]	275.6	300.5	310
PP	C–C	312.6	324.6	313.7	307.8	312
BRCLNA	Br–Cl	305.7	305.7	273.9	311.0	313
C_2Br_6	Br–Br	340.3	344.5	322.5	345.1	342
RESVAN	S–S	382.5	397.9	360.6	399.2	419
BHS	Si–Si	452.8	450.9	460.8	445.2	443

[a] Reference bond lengths of long bonds as used in Ref. 110.
[b] Bonds are dissociated and/or replaced by other bonds. Thus, this value is neglected in the statistical analysis of the data set presented in the manuscript.

Table A7.3.: Covalent bonds of heavy main group elements (HMGB11)[a] from experiment and computed with differend semiempirical methods. The values are given in pm.

system	bond	GFN-xTB	PM6-D3H4X	MSINDO-D3(BJ)H+	HF-3c	ref.
Cl_2	Cl–Cl	201.1	198.0	192.7	203.7	198.8
S_2H_2	S–S	206.1	206.6	198.0	212.2	205.5
P_2Me_4	P–P	218.9	221.7	212.9	222.7	221.2
Br_2	Br–Br	230.3	232.8	223.8	229.7	228.1
Se_2H_2	Se–Se	230.7	241.3	223.8	233.6	234.6
Ge_2H_6	Ge–Ge	232.5	248.8	233.0	244.2	241.0
As_2Me_4	As–As	242.5	242.4	238.1	242.2	242.9
Te_2Me_2	Te–Te	267.6	236.0	196.9	280.1	268.6
Sn_2Me_6	Sn–Sn	280.4	235.0	290.7	287.6	277.6
Sb_2Me_4	Sb–Sb	278.6	293.7	236.4	288.6	281.8
Pb_2Me_6	Pb–Pb	293.6	290.1	309.4	303.7	288.0

[a] Reference bond lengths are the same as used in Ref. 110.

Table A7.4.: Covalent bonds in transition metal complexes (TMC32)[a] from experiment and computed with differend semiempirical methods. The values are given in pm.

	GFN-xTB	PM6-D3H4X	MSINDO-D3(BJ)H+	ref.
1	202.5	211.3	197.4	207.6
2	213.4	217.8	207.5	216.9
3	206.4	195.3	196.4	204.7
4	213.4	216.4	207.8	218.5
5	206.0	197.2	196.6	205.8
6	214.0	213.3	208.5	219.6
7	232.4	210.7	213.8	217.5
8	181.9	166.5	194.7	198.4
9	154.5	156.3	155.4	157.0
10	163.7	173.9	163.4	172.9
11	169.2	174.1	164.9	173.4
12	167.9	173.4	163.8	170.8
13	153.3	155.8	154.5	157.3
14	211.0	217.0	204.6	213.8
15	182.9	185.2	176.3	187.9
16	182.7	200.3	179.0	196.3
17	152.5	158.6	154.9	157.4
18	159.8	160.4	163.5	171.9
19	153.1	159.8	154.6	157.7
20	203.5	195.9	204.6	212.2
21	152.8	157.7	155.4	158.4
22	179.8	196.4	180.1	195.4
23	217.2	220.9	205.9	215.0
24	228.6	228.3	210.5	220.8
25	183.4	184.6	176.3	186.3
26	168.5	168.6	164.7	175.0
27	155.0	159.7	152.1	158.6
28	163.9	171.7	161.7	172.4
29	217.0	218.8	203.6	214.7
30	180.6	181.0	167.3	180.6
31	178.4	178.3	159.6	182.9
32	178.1	176.7	156.4	181.0
33	193.4	186.7	180.0	193.8
34	210.6	206.8	194.1	212.3
35	182.4	175.4	162.7	187.2
36	160.8	159.6	145.4	167.4
37	206.5	205.9	189.3	206.4
38	213.9	221.2	192.1	211.7
39	177.5	226.3	156.1	181.5
40	178.5	171.1	160.8	180.6
41	239.6	222.8	238.9	237.7
42	179.3	179.3	153.5	181.8
43	160.0	159.8	139.9	165.8
44	181.3	212.7	155.4	183.0
45	179.7	181.3	158.5	182.5
46	185.8	185.4	173.9	187.6
47	211.1	220.8	199.6	209.9
48	198.9	193.3	174.3	188.4
49	190.5	180.4	167.7	183.2
50	199.0	188.2	179.1	191.4

[a] Reference bond lengths are from Ref. 506.

Table A7.5.: Association energies of 22 non-covalently interacting systems (S22)[a] computed with different semiempirical methods. The values are given in kcal mol^{-1}.

	GFN-xTB	PM6-D3H4X	MSINDO-D3(BJ)H+	OM2-D3(BJ)	DFTB3-D3(BJ)	HF-3c	ref.
1	-1.77	-3.82	-0.35	-2.51	-1.64	-2.64	-3.13
2	-4.76	-4.86	-6.47	-4.53	-4.71	-5.93	-4.99
3	-18.59	-18.22	-18.41	-14.84	-18.68	-18.05	-18.75
4	-14.50	-16.88	-13.77	-14.35	-15.07	-15.47	-16.06
5	-17.28	-18.51	-17.41	-19.39	-18.15	-20.75	-20.64
6	-14.40	-16.33	-13.63	-13.31	-12.65	-18.07	-16.93
7	-14.15	-15.51	-14.53	-13.51	-11.71	-17.27	-16.66
8	-0.39	-0.70	-0.45	-0.45	-0.68	-0.56	-0.53
9	-0.38	-1.07	-0.43	-1.49	-1.56	-1.73	-1.47
10	-1.04	-1.54	-1.68	-1.42	-1.66	-1.57	-1.45
11	-3.26	-2.96	-5.82	-2.49	-3.74	-3.47	-2.65
12	-3.52	-4.81	-6.67	-4.50	-4.43	-4.64	-4.25
13	-8.54	-9.14	-9.33	-9.29	-9.16	-8.40	-9.80
14	-4.33	-4.03	-8.08	-3.42	-5.37	-4.82	-4.52
15	-9.95	-10.76	-13.07	-10.89	-11.81	-10.43	-11.73
16	-0.91	-1.05	-1.00	-1.64	-1.33	-1.64	-1.50
17	-1.65	-3.30	-2.40	-3.54	-3.08	-3.52	-3.27
18	-1.28	-2.67	-1.96	-2.54	-2.21	-2.54	-2.31
19	-0.98	-2.84	-2.56	-4.53	-3.25	-3.92	-4.54
20	-1.82	-2.37	-2.67	-2.83	-2.55	-2.82	-2.72
21	-3.03	-4.75	-3.10	-5.26	-4.67	-5.57	-5.63
22	-6.05	-7.42	-7.64	-6.28	-6.09	-7.98	-7.10

Reference energies taken from Ref. 493. Structures taken from Ref. 492. Running number as in Ref. 130.
1: ammonia dimer, 2: water dimer, 3: formic acid dimer, 4: formamide dimer, 5: uracil dimer, 6: 2-pyridoxine · 2-aminopyridine, 7: adenine · thymine, 8: methane dimer, 9: ethene dimer, 10: benzene · methane, 11: benzene dimer, 12: pyracine dimer, 13: uracil dimer, 14: indole · benzene, 15: adenine · thymine (stack), 16: ethene · ethine, 17: benzene · water, 18: benzene · ammonia, 19: benzene · cyanide, 20: benzene dimer, 21: indole · benzene (T-shape), 22: phenol dimer.

Table A7.6.: Association energies of 12 large non-covalently host-guest systems (S12L)[a] computed with different semiempirical methods. The values are given in kcal mol^{-1}.

		GFN-xTB	PM6-D3H4X	MSINDO-D3(BJ)H+	OM2-D3(BJ)	DFTB3-D3(BJ)	HF-3c	ref.
2a	TCNA@tweezer	-21.97	-28.83	-30.87	-32.28	-28.62	-30.89	-29.0
2b	DCB@tweezer	-14.51	-18.63	-22.54	-20.47	-19.88	-21.26	-20.8
3a	3c@pincer	-21.36	-25.25	-33.67	-27.15	-25.79	-24.91	-23.5
3b	3d@pincer	-16.52	-19.18	-24.80	–	-18.79	-19.22	-20.3
4a	C_{60}@catcher	-33.02	-27.66	-46.60	-32.72	-36.37	-37.37	-28.4
4b	C_{70}@catcher	-34.86	-29.13	-47.88	-33.68	-38.06	-39.25	-29.8
5a	GLH@mcycle	-25.18	-39.73	-29.45	-36.47	-31.14	-28.56	-33.4
5b	BQ@mcycle	-20.49	-29.11	-25.24	-26.67	-23.09	-20.46	-23.3
6a	$[BuNBH4@CB6]^{1+}$	-94.76	-104.03	-91.02	-94.98	-95.17	-92.41	-82.2
6b	$[PrNH4@CB6]^{1+}$	-90.49	-101.22	-84.26	-91.20	-89.88	-88.76	-80.1
7a	$[FECP@CB7]^{2+}$	-155.87	-145.21	-151.78	–	–	-142.82	-132.7
7b	ADOH@CB7	-28.42	-31.39	-34.93	-23.84	-28.22	-27.61	-24.2

[a] Structures are taken from Ref. 42, reference energies from Ref. 130.

Table A7.7.: Association energies of 19 non-bonded systems[a] computed with different semiempirical methods. This is a collection of the HB6/04, CT7/04, and DI6/04 sets. The values are given in kcal mol^{-1}.

	GFN-xTB	PM6-D3H4X	MSINDO-D3(BJ)H+	DFTB3-D3(BJ)	HF-3c	ref.
$(H_2S)_2$	-2.34	-1.47	-1.10	-1.94	-0.84	-1.66
$(HCl)_2$	-1.89	-1.26	-0.61	-1.65	-0.66	-2.01
$HCl \cdot H_2S$	-4.80	-1.59	0.20	-2.29	-1.92	-3.35
$CH_3Cl \cdot HCl$	-2.40	-1.92	-0.75	-2.19	-1.38	-3.55
$HCN \cdot CH_3SH$	-1.52	-2.12	-2.41	-1.50	-1.75	-3.59
$CH_3SH \cdot HCl$	-5.10	-1.93	0.49	-2.75	-2.50	-4.16
$C_2H_4 \cdot F_2$	-1.08	-0.08	-1.33	-0.61	0.43	-1.06
$NH_3 \cdot F_2$	-2.08	0.57	4.15	0.07	7.33	-1.81
$C_2H_2 \cdot ClF$	-2.25	-0.44	-1.74	-0.29	4.44	-3.81
$HCN \cdot ClF$	-3.54	23.90	-0.18	-6.91	-7.11	-4.86
$NH_3 \cdot Cl_2$	-3.92	26.48	-4.37	-9.09	-1.78	-4.88
$H_2O \cdot ClF$	-5.01	23.51	0.55	-11.24	-1.54	-5.36
$NH_3 \cdot ClF$	-12.25	8.84	-5.31	-13.90	-2.35	-10.62
$(NH_3)_2$	-1.77	-4.21	-0.86	-1.82	-2.80	-3.15
$(HF)_2$	-4.21	3.01	-3.26	-4.40	-3.50	-4.57
$(H_2O)_2$	-4.89	-5.13	-5.33	-4.82	-5.66	-4.97
$NH_3 \cdot H_2O$	-6.47	-7.71	-6.91	-4.08	-7.96	-6.41
$(HCONH_2)_2$	-12.97	-15.06	-11.39	-13.57	-12.23	-14.94
$(HCOOH)_2$	-17.44	-17.46	-12.32	-17.16	-14.31	-16.15

Reference data taken from Ref. 511.

Table A7.8.: Interaction energies of the X40[a] set computed with different semiempirical methods. The values are given in kcal mol^{-1}.

#	system	GFN-xTB	PM6-D3H4X	MSINDO-D3(BJ)H+	DFTB3-D3(BJ)	HF-3c	ref.
1	$CH_4 \cdot F_2$	-0.31	0.01	-0.57	-0.30	0.38	-0.49
2	$CH_4 \cdot Cl_2$	-0.53	-1.32	-1.31	-1.02	-0.50	-1.08
3	$CH_4 \cdot Br_2$	-0.39	-1.98	-3.15	-0.63	-1.71	-1.30
4	$CH_4 \cdot I_2$	-0.64	-2.04	–	-0.98	-2.27	-1.34
5	$CH_3F \cdot CH_4$	-0.38	-0.79	-0.66	-0.74	-0.39	-0.75
6	$CH_3Cl \cdot CH_4$	-0.57	-1.38	-0.82	-1.35	-0.72	-0.98
7	$CHF_3 \cdot CH_4$	-0.58	-0.62	-1.22	-0.80	0.02	-0.69
8	$CHCl_3 \cdot CH_4$	-1.29	-1.65	-1.97	-2.50	-0.50	-1.14
9	CH_3F dimer	-1.88	-0.53	-1.94	-2.08	-0.49	-1.65
10	CH_3Cl dimer	-0.98	-1.45	-2.27	-1.41	-0.76	-1.34
11	$C_6H_3F_3 \cdot C_6H_6$	-3.47	-3.94	-6.49	-5.06	-4.49	-4.40
12	$C_6F_6 \cdot C_6H_6$	-3.97	-5.00	-7.47	-6.82	-5.74	-6.12
13	$CH_3Cl \cdot CH_2O$	-0.43	-0.61	-0.84	-2.55	-0.09	-1.17
14	$CH_3Br \cdot CH_2O$	-1.72	-1.45	-2.84	-3.73	-3.10	-1.72
15	$CH_3I \cdot CH_2O$	-2.73	-2.13	–	-5.37	-6.87	-2.38
16	$CF_3Cl \cdot CH_2O$	-1.05	-1.68	-1.03	-5.93	-1.63	-2.25
17	$CF_3Br \cdot CH_2O$	-3.37	-2.70	-4.82	-6.01	-6.34	-3.10
18	$CF_3I \cdot CH_2O$	-4.53	-4.22	–	-7.59	-10.97	-4.08
19	$C_6H_5Cl \cdot C(CH_3)_2O$	-0.45	-1.22	-0.07	-4.60	-0.59	-1.49
20	$C_6H_5Br \cdot C(CH_3)_2O$	-2.39	-2.55	-2.21	-5.48	-4.86	-2.42
21	$C_6H_5I \cdot C(CH_3)_2O$	-3.87	-3.44	–	-7.37	-9.77	-3.46
22	$C_6H_5Cl \cdot N(CH_3)_3$	-0.93	-1.66	-0.91	-7.73	-1.64	-2.12
23	$C_6H_5Br \cdot N(CH_3)_3$	-3.77	-3.96	-6.96	-9.29	-7.01	-3.78
24	$C_6H_5I \cdot N(CH_3)_3$	-5.25	-5.64	–	-12.30	-11.98	-5.80
25	$C_6H_5Br \cdot CH_3SH$	-2.51	-0.78	-4.13	-3.34	-2.40	-2.31
26	$C_6H_5I \cdot CH_3SH$	-3.24	-3.10	–	-4.81	-4.64	-3.08
27	$CH_3Br \cdot C_6H_6$	-1.20	-2.61	-4.36	-3.27	-3.17	-1.82
28	$CH_3I \cdot C_6H_6$	-1.60	-3.48	-21.27	-3.83	-5.97	-2.49
29	$CF_3Br \cdot C_6H_6$	-1.58	-3.64	-6.29	-4.27	-5.18	-3.11
30	$CF_3I \cdot C_6H_6$	-2.15	-5.02	-25.42	-4.84	-9.28	-3.91
31	$CF_3OH \cdot H_2O$	-9.31	-8.49	-12.86	-9.01	-11.96	-9.67
32	$CCl_3OH \cdot H_2O$	-8.58	-9.72	-11.80	-9.11	-12.40	-10.41
33	$HF \cdot CH_3OH$	-7.77	-3.71	-8.41	-8.58	-9.72	-9.59
34	$HCl \cdot CH_3OH$	-4.51	-5.94	-6.92	-7.79	-7.25	-6.30
35	$HBr \cdot CH_3OH$	-1.61	-5.51	-5.83	-5.45	-7.52	-5.35
36	$HI \cdot CH_3OH$	-0.83	-5.08	–	-4.94	-7.22	-3.97
37	$HF \cdot N(CH_3)H_2$	-14.10	-2.93	-10.62	-7.38	-13.04	-14.32
38	$HCl \cdot N(CH_3)H_2$	-10.21	-8.26	-9.55	-8.62	-13.10	-11.42
39	$CH_3OH \cdot CH_3F$	-3.04	0.86	-2.98	-3.44	-2.64	-3.90
40	$CH_3OH \cdot CH_3Cl$	-2.96	-2.25	-2.18	-2.32	-1.34	-3.78

[a] Structures and reference energies from Ref. 512.

Table A7.9.: Interaction energies of water clusters (WATER27[a] set) computed with different semiempirical methods. The values are given in kcal mol^{-1}per water (and hydronium/hydroxide) molecule.

	GFN-xTB	PM6-D3H4X	MSINDO-D3(BJ)H+	OM2-D3(BJ)	DFTB3-D3(BJ)	HF-3c	ref.
H_2O_2	-2.39	-2.59	-3.21	-2.19	-2.44	-2.97	-2.50
H_2O_3	-4.88	-5.00	-5.29	-4.41	-4.92	-5.61	-5.27
H_2O_4	-6.07	-6.56	-8.12	-6.21	-6.48	-8.06	-6.85
H_2O_5	-6.20	-6.80	-9.26	-6.60	-6.75	-8.66	-7.18
H_2O_6	-7.56	-7.61	-8.61	-6.91	-7.48	-8.59	-7.67
H_2O_6c	-7.22	-7.50	-8.52	-6.74	-7.32	-8.59	-7.63
H_2O_6b	-6.79	-7.27	-9.17	-6.81	-7.15	-8.84	-7.55
H_2O_6c2	-6.30	-6.89	-10.16	-6.95	-7.02	-9.06	-7.38
H_2O_8d2d	-8.52	-9.01	-10.42	-8.12	-8.74	-10.34	-9.07
H_2O_8s4	-8.53	-8.98	-10.42	-8.13	-8.72	-10.32	-9.07
H_2O_{20}	-8.84	-9.79	-12.57	-8.89	-9.31	-11.61	-10.01[b]
H_2O_{20}fc	-10.06	-10.36	-11.90	-9.35	-10.22	-11.61	-10.63[b]
H_2O_{20}fs	-9.87	-10.35	-12.55	-9.44	-10.13	-11.77	-10.75[b]
H_2O_{20}es	-9.85	-10.32	-12.57	-9.45	-10.16	-11.92	-10.89[b]
$H_3O^+H_2O$	-16.60	-14.23	-25.00	-17.86	-15.27	-25.75	-16.75
$H_3O^+H_2O_2$	-18.11	-16.62	-27.25	-19.10	-17.47	-26.91	-18.97
$H_3O^+H_2O_3$	-18.15	-17.04	-26.72	-18.96	-17.83	-26.02	-19.12
$H_3O^+H_2O_6$3d	-16.11	-15.77	-21.66	-16.05	-15.50	-20.93	-16.83
$H_3O^+H_2O_6$2d	-15.69	-15.55	-22.25	-16.27	-15.53	-21.51	-16.41
OH^-H_2O	-19.45	-17.54	-25.40	-17.64	-15.17	-28.15	-13.30
$OH^-H_2O_2$	-20.01	-18.12	-26.13	-18.14	-17.68	-28.71	-16.13
$OH^-H_2O_3$	-19.83	-18.48	-25.10	-17.91	-18.13	-27.86	-16.90
$OH^-H_2O_4$c4	-19.26	-17.64	-23.18	-17.07	-18.22	-25.71	-16.96
$OH^-H_2O_4$cs	-19.45	-18.07	-23.11	-17.08	-18.17	-25.60	-16.96
$OH^-H_2O_5$	-19.14	-17.21	-21.02	-16.28	-18.07	-23.97	-16.78
$OH^-H_2O_6$	-18.43	-17.00	-20.35	-15.82	-17.49	-23.21	-16.53
$H_3O^+H_2O_6OH^-$	-5.36	-2.64	-6.11	-2.64	-3.87	-5.14	-3.56

[a] Reference energies from Ref. 508.
[b] Reference energies from Ref. 509.

Table A7.10.: Interaction energies of 15 charged H-bonded systems (C15)[a] computed with different semiempirical methods. The values are given in kcal mol^{-1}.

	GFN-xTB	PM6-D3H4X	MSINDO-D3(BJ)H+	OM2-D3(BJ)	DFTB3-D3(BJ)	HF-3c	ref.
$CH_3COO^- \cdot CH_3NH_2$	-9.715	-12.041	-9.385	-8.384	-8.671	-12.313	-11.455
$CH_3COO^- \cdot CH_3OH$	-18.057	-19.051	-22.179	-15.111	-17.993	-23.842	-19.753
$CH_3COO^- \cdot H_2O$	-21.157	-21.816	-23.603	-17.939	-22.918	-25.393	-21.062
$CH_6N_3^+ \cdot CH_2O$	-16.507	-18.888	-17.938	-16.579	-14.446	-17.684	-18.090
$CH_6N_3^+ \cdot CH_3NH_2$	-15.008	-19.077	-16.218	-17.384	-11.978	-23.016	-20.196
$CH_6N_3^+ \cdot CH_3OH$	-16.327	-18.828	-18.756	-17.696	-16.246	-21.428	-19.788
$CH_6N_3^+ \cdot H_2O$	-15.799	-16.849	-19.505	-16.185	-15.148	-20.344	-17.467
$C_3H_5N_2^+ \cdot CH_2O$	-14.671	-16.065	-15.020	-15.425	-12.179	-17.065	-16.410
$C_3H_5N_2^+ \cdot CH_3NH_2$	-21.838	-22.672	-22.365	-24.485	-13.445	-32.981	-25.977
$C_3H_5N_2^+ \cdot CH_3OH$	-14.997	-15.359	-15.962	-17.593	-14.147	-23.245	-18.914
$C_3H_5N_2^+ \cdot H_2O$	-14.382	-13.984	-16.918	-16.206	-13.295	-21.700	-16.485
$CH_6N^+ \cdot CH_2O$	-18.861	-20.474	-19.596	-18.229	-16.788	-20.370	-19.096
$CH_6N^+ \cdot CH_3NH_2$	-26.319	-29.002	-28.885	-26.449	-20.156	-36.823	-28.560
$CH_6N^+ \cdot CH_3OH$	-19.211	-19.652	-21.564	-19.960	-19.461	-26.679	-21.225
$CH_6N^+ \cdot H_2O$	-18.206	-18.137	-22.412	-18.353	-17.955	-24.902	-18.514

[a] Reference energies from Ref. 468.

Table A7.11.: 24 protein-ligand (PL24)[a] interaction energies computed with different semiempirical methods. The values are given in kcal mol^{-1}.

	GFN-xTB	PM6-D3H4X	MSINDO-D3(BJ)H+	DFTB3-D3(BJ)	HF-3c	ref.
1	-3.51	1.35	0.04	-1.49	5.35	1.8
2	-6.06	-6.19	-7.06	-5.86	-2.79	-4.6
3	-7.63	-7.62	-7.96	-7.27	-4.11	-7.5
4	-11.68	-14.23	-13.77	-12.01	-13.51	-12.0
5	-12.68	-16.41	-17.28	-14.13	-15.00	-12.7
6	-12.21	-17.30	-15.74	-14.32	-14.11	-15.7
7	-12.42	-19.13	-16.11	-14.10	-14.17	-16.3
8	-12.84	-19.70	-16.93	-14.59	-15.04	-16.7
9	-18.75	-28.81	-26.13	-19.47	-17.34	-17.6
10	-26.60	-39.65	-21.85	-23.93	-18.57	-18.4
11	-21.59	-27.05	-26.95	-17.88	-21.09	-22.7
12	-24.88	-30.47	-25.04	-27.83	-24.13	-24.6
13	-40.78	-49.25	-43.04	-36.82	-28.57	-28.3
14	-32.24	-42.62	-41.41	-32.29	-32.60	-32.1
15	-35.03	-40.95	-33.27	-31.61	-36.36	-36.3
16	-40.89	-43.08	-38.27	-40.11	-33.16	-37.2
17	-43.89	-57.75	-47.87	-40.25	-35.60	-40.2
18	-44.39	-50.43	-50.94	-47.75	-42.72	-43.3
19	-48.93	-49.16	-40.90	-48.37	-41.83	-47.1
20	-57.57	-70.50	-64.92	-52.48	-48.35	-49.2
21	-55.22	-71.99	-60.50	-52.11	-46.54	-48.8
22	-62.30	-68.61	-68.48	-64.72	-58.93	-59.7
23	-57.05	-60.55	-62.56	-55.90	-65.45	-64.7
24	-73.47	-82.36	-85.82	-71.82	-69.13	-69.1

[a] Reference structures and energies from Ref. 510.

Table A7.12.: Interaction energies of carbondioxide with metalorganic framework fragments (MOF-CO_2)[a] computed with different semiempirical methods. The values are given in kcal mol^{-1}.

		GFN-xTB	PM6-D3H4X	MSINDO-D3(BJ)H+	HF-3c	ref.
1	bdc	-1.26	-2.14	-3.54	-2.74	-2.39
2	bdc-li	-2.80	-2.41	-4.75	-4.11	-2.77
3	btc	-3.32	-2.42	-5.06	-4.82	-2.54
4	dpt	-0.80	-1.32	-3.25	-0.33	-1.27
5	bdpd	-2.81	-4.48	-5.65	-4.63	-3.10
6	btt-1	-1.77	-2.56	-4.52	-3.29	-3.04
7	btt-2	-2.10	-3.24	-4.50	-3.26	-2.78
8	tatb	-2.44	-4.27	-5.84	-4.18	-1.60
9	pur	-3.47	-6.42	-7.50	-6.55	-4.80
10	ade	-3.67	-7.01	-8.17	-7.16	-5.09
11	abdc	-2.91	-3.79	-7.50	-4.64	-3.52
12	ambdc	-3.20	-4.44	-8.20	-5.14	-3.04
13	diamino	-2.56	-3.48	-5.04	-4.18	-2.70
14	pei-1	-2.38	-5.04	-6.82	-5.63	-3.20
15	pei-2	-1.95	-4.61	-5.91	-5.11	-2.74
16	hmta	-2.30	-4.48	-6.15	-5.50	-3.42
17	azpy	-1.95	-3.41	-5.19	-4.02	-3.64
18	abtc	-2.31	-3.92	-5.98	-4.24	-3.62
19	cnIm	-2.23	-2.83	-3.73	-3.27	-2.86
20	nIm	-2.21	-1.79	-3.97	-3.06	-2.88
21	nbIm	-2.44	-2.24	-4.71	-3.61	-2.80
22	cbIm-1	-2.07	-3.22	-4.28	-3.42	-2.64
23	cbIm-2	-1.10	-1.83	-2.78	-1.88	-1.21
24	dhbf	-2.15	-3.25	-3.85	-3.70	-2.46
25	pycf3-1	-1.76	-1.71	-3.08	-0.88	-1.06
26	pycf3-2	-1.45	-2.22	-2.99	-2.02	-2.04
27	zn4o-1	-4.15	-4.34	-15.62	-7.59	-3.84
28	zn4o-2	-4.38	-2.91	-10.74	-10.90	-4.61

[a] Reference structures and energies from Ref. 513.

Table A7.13.: Association energies of rare gases with cucurbituril hosts (RG-CBx)[a] computed with different semiempirical methods. The values are given in kcal mol^{-1}.

#	system	GFN-xTB	PM6-D3H4X	HF-3c	ref.
1	He · CB_5	-1.21	-0.52	0.12	-1.10
2	Ne · CB_5	-1.41	-2.58	1.51	-1.06
3	Ar · CB_5	-5.35	2.22	-1.47	-5.94
4	Kr · CB_5	-6.88	55.02	-9.41	-6.77
5	Xe · CB_5	-8.23	6.77	-61.10	-8.39
6	He · CB_6	-0.64	-0.71	-2.43	-0.64
7	Ne · CB_6	-0.97	-2.66	-0.90	-0.54
8	Ar · CB_6	-3.46	4.73	-3.63	-3.11
9	Kr · CB_6	-4.77	36.20	-8.44	-4.49
10	Xe · CB_6	-6.78	-2.33	-34.65	-7.08

[a] Reference structures at the PBEh-3c level, while reference energies are obtained at the LDLPNO-CCSD(T)/def2-QZVPP (counterpoise and deformation energy corrected).

Table A7.14.: Center of mass (CMA) distances of 22 non-covalently interacting systems (S22)[a] computed with different semiempirical methods. All atom position root mean square deviations (RMSDs) are given in parentheses. The values are given in pm.

	GFN-xTB	PM6-D3H4X	MSINDO-D3(BJ)H+	OM2-D3(BJ)	DFTB3-D3(BJ)	HF-3c	ref.
1	306.6 (12.0)	351.0 (13.2)	366.0 (20.8)	329.5 (3.2)	330.0 (4.7)	331.4 (4.0)	324.6
2	288.6 (5.3)	291.1 (4.3)	263.8 (20.9)	274.6 (15.2)	288.8 (6.4)	290.1 (8.0)	291.1
3	296.1 (3.1)	311.1 (7.3)	269.8 (33.5)	288.0 (8.0)	302.7 (5.4)	302.7 (7.6)	301.1
4	330.8 (3.6)	330.2 (7.8)	311.1 (14.3)	311.4 (8.6)	326.3 (3.3)	325.0 (1.7)	325.8
5	609.1 (2.9)	612.6 (6.5)	604.2 (11.8)	590.9 (8.4)	611.3 (4.2)	602.7 (2.9)	606.0
6	510.8 (14.8)	536.8 (22.9)	513.1 (17.5)	508.5 (24.0)	526.7 (17.1)	512.6 (25.5)	517.7
7	590.2 (11.5)	618.8 (11.8)	597.3 (29.2)	589.2 (8.1)	613.5 (7.5)	599.1 (4.4)	602.6
8	372.5 (0.5)	352.1 (10.0)	372.0 (1.6)	350.9 (10.4)	340.8 (15.5)	360.1 (5.8)	371.7
9	409.4 (18.8)	389.8 (9.1)	407.6 (18.0)	359.9 (6.0)	362.1 (5.0)	362.4 (4.8)	371.8
10	376.8 (2.4)	364.3 (3.7)	356.0 (7.2)	353.8 (8.2)	355.1 (7.6)	370.3 (0.8)	371.6
11	353.3 (23.6)	377.0 (0.9)	314.3 (33.5)	423.6 (45.6)	347.2 (17.5)	367.7 (8.9)	376.4
12	332.6 (32.6)	337.1 (36.9)	311.3 (18.7)	307.2 (24.7)	328.5 (11.4)	349.2 (2.7)	347.9
13	299.3 (17.0)	325.1 (12.7)	345.6 (78.5)	278.9 (22.1)	306.3 (24.8)	318.4 (14.4)	317.6
14	335.6 (26.1)	408.3 (67.9)	297.4 (44.8)	400.1 (76.2)	334.7 (12.4)	351.1 (1.5)	349.8
15	307.6 (25.5)	327.1 (35.2)	288.6 (36.4)	291.0 (36.6)	301.6 (22.9)	320.0 (11.7)	316.7
16	429.4 (6.3)	426.7 (7.8)	419.6 (11.2)	376.3 (32.1)	413.1 (14.3)	442.2 (0.0)	442.2
17	316.6 (16.8)	301.9 (25.1)	288.9 (28.3)	277.8 (32.5)	291.9 (27.1)	311.3 (22.1)	333.7
18	304.3 (41.0)	311.0 (39.5)	307.8 (40.9)	290.6 (44.8)	308.5 (34.4)	328.6 (23.5)	353.9
19	416.3 (27.4)	399.4 (4.5)	310.5 (49.8)	379.4 (28.4)	387.0 (1.0)	385.8 (1.6)	389.2
20	502.3 (5.8)	491.4 (1.0)	485.8 (2.8)	479.8 (5.6)	490.2 (0.7)	490.0 (0.7)	490.8
21	478.7 (43.3)	490.5 (12.2)	489.7 (23.3)	459.5 (41.3)	480.3 (16.7)	485.5 (7.8)	488.7
22	512.7 (44.3)	493.7 (13.1)	308.5 (135.2)	498.2 (58.4)	488.3 (6.7)	527.9 (45.5)	494.0

[a] Reference structures taken from Ref. 492.

1: ammonia dimer, 2: water dimer, 3: formic acid dimer, 4: formamide dimer, 5: uracil dimer, 6: 2-pyridoxine · 2-aminopyridine, 7: adenine · thymine, 8: methane dimer, 9: ethene dimer, 10: benzene · methane, 11: benzene dimer, 12: pyracine dimer, 13: uracil dimer, 14: indole · benzene, 15: adenine · thymine (stack), 16: ethene · ethine, 17: benzene · water, 18: benzene · ammonia, 19: benzene · cyanide, 20: benzene dimer, 21: indole · benzene (T-shape), 22: phenol dimer.

Table A7.15.: Heavy atom position root mean square deviations (RMSDs) for structures of small peptides containing an aromatic side chain (P26)[a] The values are given in pm.

#	sytem	GFN-xTB	PM6-D3H4X	MSINDO-D3(BJ)H+	OM2-D3(BJ)	DFTB3-D3(BJ)	HF-3c
1	FGG_215	23.1	53.8	93.5	12.1	11.0	17.5
2	FGG_252	9.5	10.3	47.6	15.5	9.6	7.1
3	FGG_357	42.1	27.5	24.7	16.0	32.3	30.5
4	FGG_412	4.6	12.9	31.7	31.0	11.5	20.1
5	FGG_470	49.6	61.3	75.8	85.5	6.4	45.1
6	FGG_55	8.9	9.5	18.6	17.0	7.0	8.9
7	FGG_99	9.1	15.6	26.7	47.9	18.5	5.5
8	GFA_01	4.7	42.3	48.0	49.5	13.0	14.5
9	GFA_02	23.0	46.9	75.8	40.7	13.4	8.9
10	GFA_08	2.7	22.1	35.4	18.1	12.2	6.3
11	GFA_16	20.7	23.8	24.5	17.0	16.2	14.9
12	GGF_01	33.6	75.9	73.9	36.9	15.7	35.8
13	GGF_05	9.2	44.1	50.6	22.4	13.9	16.5
14	GGF_12	8.8	44.7	51.0	20.1	13.3	15.2
15	GGF_14	3.3	20.2	43.5	7.6	10.9	8.7
16	WG_03	3.4	47.0	104.9	34.1	4.2	11.8
17	WG_06	2.0	13.1	84.4	14.3	7.2	12.6
18	WG_10	7.0	25.9	34.6	35.7	9.9	17.8
19	WG_14	8.2	19.8	29.6	29.0	8.2	16.6
20	WG_15	10.0	12.4	30.7	24.9	9.0	9.9
21	WGG_03	3.8	28.1	43.8	30.8	6.4	30.8
22	WGG_04	2.2	10.9	47.6	20.5	10.1	12.2
23	WGG_05	4.9	12.1	37.9	17.1	5.2	12.6
24	WGG_06	5.7	41.4	26.8	21.7	6.0	53.5
25	WGG_08	4.0	29.2	41.9	31.9	8.1	34.2
26	WGG_11	13.6	33.0	34.1	36.4	18.1	40.8

[a] Reference structures taken from Ref. 514.

Table A7.16.: Equilibrium center-of-mass distances between non-covalently bound systems obtained from the S66x8 set[a]. The values are given in pm and obtained from a cubic spline interpolation based on the interaction energies computed on the 8 structures along the potential energy curve for each complex.

#	sytem	GFN-xTB	PM6-D3H4X	MSINDO-D3(BJ)H+	OM2-D3(BJ)	DFTB3-D3(BJ)	HF-3c	ref.
1	H_2O · H_2O	293.9	293.9	278.3	278.3	287.9	289.4	293.9
2	H_2O · MeOH	312.9	311.5	294.2	296.6	306.7	309.6	309.6
3	H_2O · $MeNH_2$	321.1	351.2	319.1	319.1	336.4	333.0	334.4
4	H_2O · peptide	384.9	385.3	370.1	370.1	378.4	384.4	384.9
5	MeOH · MeOH	354.6	356.9	336.8	340.0	350.0	348.2	350.0
6	MeOH · $MeNH_2$	325.0	355.0	319.1	323.0	341.7	332.4	335.8
7	MeOH · peptide	420.3	423.4	405.7	406.1	417.2	418.5	420.7
8	MeOH · H_2O	328.6	333.9	312.3	312.3	327.7	322.4	328.6
9	$MeNH_2$ · MeOH	356.7	342.0	351.0	336.9	345.4	345.9	354.4
10	$MeNH_2$ · $MeNH_2$	345.7	364.9	372.2	350.8	349.7	344.0	348.0
11	$MeNH_2$ · peptide	362.0	365.8	358.2	349.6	358.8	363.1	366.9
12	$MeNH_2$ · H_2O	295.2	323.9	287.4	296.1	311.6	304.4	303.4
13	peptide · MeOH	396.8	389.3	380.9	377.9	390.3	383.9	388.3
14	peptide · $MeNH_2$	382.5	402.9	406.0	382.5	392.7	380.0	388.7
15	peptide · peptide	470.0	465.6	459.8	451.2	466.1	461.8	468.1
16	peptide · H_2O	382.8	379.8	364.4	364.4	382.3	373.6	382.3
17	uracil · uracil (BP)	575.5	576.3	565.3	558.9	577.2	569.1	574.6
18	H_2O · pyridine	417.9	447.3	410.9	420.4	428.4	428.9	426.9
19	MeOH · pyridine	441.8	472.7	433.2	446.5	452.3	449.9	449.9
20	AcOH · AcOH	404.9	416.0	392.3	392.3	407.5	405.8	407.9
21	$AcNH_2$ · $AcNH_2$	433.5	433.0	420.0	415.8	431.1	427.4	432.5
22	AcOH · uracil	505.6	512.0	490.4	490.4	507.3	502.2	506.0
23	$AcNH_2$ · uracil	512.5	512.9	501.7	495.7	512.9	506.0	512.0
24	benzene · benzene (π-π)	366.6	373.0	337.9	349.1	353.7	371.1	387.6
25	pyridine · pyridine (π-π)	352.7	357.0	324.1	324.1	339.8	356.1	369.9
26	uracil · uracil (π-π)	300.8	325.2	289.6	289.6	302.9	314.8	314.8
27	benzene · pyridine (π-π)	360.1	365.6	330.6	337.1	347.2	364.7	379.2
28	benzene · uracil (π-π)	333.4	338.9	307.7	307.7	321.0	336.6	338.9
29	pyridine · uracil (π-π)	325.9	334.6	299.4	299.4	315.4	332.9	333.7
30	benzene · ethene	328.6	338.2	308.4	314.6	314.6	338.2	353.2
31	uracil · ethene	331.2	329.0	303.3	303.3	311.6	329.7	331.2
32	uracil · ethyne	327.7	334.5	298.1	298.1	308.8	326.2	326.2
33	pyridine · ethene	327.9	335.1	307.0	307.0	311.8	334.3	345.5
34	pentane · pentane	405.6	395.6	434.3	367.6	367.6	378.8	382.5
35	neopentane · pentane	468.1	462.5	505.4	431.7	433.4	447.0	452.8
36	neopentane · neopentane	520.1	513.7	560.2	498.2	498.2	509.6	525.6
37	cyclopentane · neopentane	480.0	468.9	526.2	441.4	444.9	459.1	465.7
38	cyclopentane · cyclopentane	439.5	437.9	482.9	405.0	407.8	419.4	421.7
39	benzene · cyclopentane	391.9	391.9	386.2	376.5	377.3	388.6	394.3
40	benzene · neopentane	441.4	433.5	432.7	423.9	427.1	440.6	448.5
41	uracil · pentane	351.9	353.3	346.7	337.9	337.9	351.1	353.3
42	uracil · cyclopentane	372.8	374.3	364.4	359.7	360.5	374.3	375.8
43	uracil · neopentane	431.4	432.2	422.0	414.1	419.6	433.7	434.5
44	ethene · pentane	400.2	390.5	409.7	362.6	361.0	368.8	375.6
45	ethyne · pentane	358.8	354.7	343.9	334.8	340.6	353.9	362.9
46	peptide · pentane	366.7	367.3	379.3	350.2	348.2	360.9	362.2
47	benzene · benzene (TS)	494.3	484.9	477.8	469.0	479.3	481.7	490.4
48	pyridine · pyridine (TS)	480.5	492.9	472.9	468.1	475.0	479.1	481.9
49	benzene · pyridine (TS)	493.3	482.4	476.9	466.6	477.7	479.2	487.1
50	benzene · ethyne (CH-π)	434.0	421.3	404.4	390.4	401.9	406.8	410.1
51	ethyne · ethyne (TS)	436.4	434.9	409.9	409.9	418.5	429.9	435.6
52	benzene · AcOH (OH-π)	398.7	418.0	394.8	394.2	404.5	416.1	417.3
53	benzene · $AcNH_2$ (NH-π)	464.8	476.1	475.1	457.7	467.4	475.6	476.1
54	benzene · H_2O (OH-π)	311.9	321.3	304.6	304.6	311.9	325.6	329.2
55	benzene · MeOH (OH-π)	330.2	342.1	327.6	319.6	330.9	342.1	342.1
56	benzene · $MeNH_2$ (N H-π)	354.7	354.0	354.0	341.9	345.5	354.7	358.2
57	benzene · peptide (N H-π)	400.9	401.7	401.7	388.4	393.1	398.6	404.0
58	pyridine · pyridine (N H-π)	582.8	600.3	603.0	578.0	586.7	577.0	585.7
59	ethyne · H_2O (CH-O)	413.2	392.5	379.7	379.7	386.9	389.1	399.2
60	ethyne · AcOH (OH-π)	383.9	409.6	376.6	376.6	388.5	398.7	396.4
61	pentane · AcOH	371.9	365.3	376.2	356.2	356.2	367.7	373.1
62	pentane · $AcNH_2$	353.4	353.4	363.8	338.1	338.1	351.4	358.6
63	benzene · AcOH	373.8	371.0	348.7	348.7	360.9	372.4	375.3
64	peptide · ethene	356.8	354.8	340.0	338.7	340.0	356.1	360.0
65	pyridine · ethyne	538.7	552.8	512.2	512.2	517.3	522.4	533.0
66	$MeNH_2$ · pyridine	368.9	386.6	373.6	371.9	370.1	372.5	372.5

[a] Reference structures taken from Ref. 494.

Table A7.17.: Association energies for 30 large non-covalent complexes consisting purely of main group elements (S30L)[a]. The association energies are computed with the composite method PBEh-3c[110] based on reoptimized geometries at the level of theory given in each column. The values are given in kcal mol^{-1}.

	GFN-xTB	PM6-D3H4X	DFTB3-D3(BJ)	PBEh-3c
1	-29.58	-29.38	-29.40	-31.43
2	-17.54	-18.95	-17.59	-19.86
3	-23.55	-14.94	-15.44	-26.42
4	-20.49	-18.65	-14.72	-22.41
5	-32.66	-29.21	-26.63	-34.11
6	-29.55	-32.62	-28.93	-32.31
7	-32.22	-31.87	-29.00	-32.41
8	-35.22	-38.56	-34.38	-39.23
9	-33.18	-32.83	-31.67	-33.29
10	-34.18	-33.45	-32.83	-35.02
11	-54.10	-26.62	-34.43	-39.01
12	-53.93	-30.67	-33.25	-39.75
13	-23.12	-21.12	-22.75	-24.52
14	-23.92	-23.19	-24.17	-26.38
15	-14.48	-4.70	-19.05	-20.54
16	-22.89	-25.54	-24.46	-25.12
17	-29.10	-24.59	-28.76	-31.90
18	-17.68	-17.52	-16.33	-22.24
19	-14.57	-5.74	-13.10	-16.51
20	-18.29	-8.56	-14.36	-21.74
21	-14.41	-14.75	-16.48	-25.97
22	-32.37	-37.39	-35.91	-35.37
23	-66.21	-64.10	-65.22	-67.61
24	-113.58	-97.60	-123.74	-125.21
25	-27.29	-27.08	-25.39	-28.31
26	-27.51	-26.19	-25.13	-27.84
27	-74.98	-64.10	-72.77	-81.22
28	-72.34	-69.32	-71.77	-79.81
29	-54.08	-52.95	-52.89	-55.04
30	-45.33	-45.99	-42.63	-45.70
MSD:	1.93	5.94	4.10	–
MAD:	3.88	6.14	4.14	–
SD:	5.37	6.73	2.89	–
MAX:	15.09	27.61	10.98	–

[a] Reference structures and numbering is taken from Ref. 139.

Table A7.18.: Relative conformer energies for different alkane conformers (ACONF)[a] computed with different semiempirical methods. The values are given in kcal mol^{-1}.

	GFN-xTB	PM6-D3H4X	MSINDO-D3(BJ)H+	OM2-D3(BJ)	DFTB3-D3(BJ)	HF-3c	ref.
1	0.328	0.878	1.056	0.739	0.317	0.109	0.598
2	0.298	0.659	1.152	0.621	0.277	0.093	0.614
3	0.743	1.709	2.386	1.508	0.474	-0.170	0.961
4	1.761	2.198	3.417	2.846	1.721	2.052	2.813
5	0.277	0.684	1.131	0.606	0.268	0.096	0.595
6	0.278	0.510	1.255	0.491	0.254	0.134	0.604
7	0.683	1.503	2.483	1.317	0.400	-0.183	0.934
8	0.554	1.305	2.121	1.177	0.542	0.142	1.178
9	0.618	1.435	2.474	1.312	0.563	0.283	1.302
10	1.019	2.501	3.706	2.075	0.533	-0.506	1.250
11	1.694	2.015	3.467	2.607	1.628	1.947	2.632
12	1.719	2.004	3.618	2.629	1.680	2.062	2.740
13	1.933	2.689	4.694	3.094	1.939	2.103	3.283
14	2.114	3.033	5.078	3.426	1.667	1.866	3.083
15	3.569	4.055	6.851	5.231	3.113	4.062	4.925

[a] Reference data taken from Ref. 516.

Table A7.19.: Relative conformer energies for cysteine (CYCONF)[a] computed with different semiempirical methods. The values are given in kcal mol^{-1}.

conformer	GFN-xTB	PM6-D3H4X	MSINDO-D3(BJ)H+	DFTB3-D3(BJ)	HF-3c	ref.
1	2.057	3.083	1.501	1.339	2.601	1.522
2	-0.625	-3.073	-4.651	-0.414	0.901	1.609
3	0.606	-0.941	-3.858	-0.921	1.229	1.948
4	-0.902	-3.520	-6.114	-0.582	1.518	1.795
5	2.889	2.805	3.477	2.380	1.940	2.098
6	-0.734	-3.353	-4.711	-0.635	1.021	1.933
7	1.868	0.919	-0.188	0.845	1.231	2.177
8	2.342	1.791	1.428	1.342	1.978	2.359
9	0.441	-2.975	-4.603	-0.066	3.107	2.562
10	1.030	-2.227	-5.364	0.733	1.831	2.674

[a] Reference data taken from Ref. 517.

Table A7.20.: Relative conformer energies for different melatonine conformers (MCONF)[a] computed with different semiempirical methods. The values are given in kcal mol^{-1}.

	GFN-xTB	PM6-D3H4X	MSINDO-D3(BJ)H+	OM2-D3(BJ)	DFTB3-D3(BJ)	HF-3c	ref.
1	0.653	1.366	2.419	1.310	1.457	1.535	0.593
2	1.526	2.108	1.048	1.611	1.557	2.037	1.873
3	1.275	2.515	2.067	2.011	2.008	1.962	1.402
4	0.287	1.913	-0.515	1.361	0.267	1.097	2.071
5	-0.046	0.558	2.009	0.756	-0.626	1.740	2.362
6	0.520	0.440	0.292	1.132	-0.173	1.854	2.678
7	0.477	2.455	-0.208	1.878	0.643	1.656	2.802
8	1.790	3.483	2.777	2.819	2.619	2.940	2.483
9	1.375	1.418	1.614	0.731	0.126	1.778	2.879
10	2.403	3.107	-2.549	0.776	4.681	4.317	4.134
11	2.468	3.410	-2.140	1.101	4.679	4.548	4.325
12	3.771	4.170	0.462	1.688	6.764	5.614	3.867
13	2.110	3.051	2.803	3.041	2.955	3.216	2.548
14	3.933	4.341	0.575	1.862	6.915	5.882	4.008
15	2.527	2.225	0.645	0.736	3.972	5.338	5.172
16	2.784	1.888	-1.872	0.458	4.607	5.281	5.111
17	3.182	4.561	4.314	3.972	3.937	4.434	3.468
18	1.699	2.143	2.186	1.338	0.602	2.604	3.839
19	2.055	2.824	3.424	2.484	1.633	2.961	4.081
20	2.412	3.710	3.310	3.602	3.377	3.931	3.412
21	3.766	2.982	-0.037	0.281	4.797	5.390	5.316
22	3.799	3.132	0.078	0.482	4.814	5.487	5.359
23	4.423	4.439	1.113	2.528	7.510	6.716	4.806
24	4.009	4.899	4.526	3.505	3.045	4.130	4.215
25	4.440	4.530	1.132	2.616	7.490	6.760	4.877
26	2.405	3.545	4.003	3.098	2.118	3.790	5.042
27	2.616	2.945	4.553	2.982	1.820	3.319	4.661
28	6.067	5.618	2.226	2.391	7.265	6.856	5.651
29	3.121	3.798	2.239	3.993	2.354	4.306	5.680
30	3.026	3.328	4.678	3.539	2.474	3.951	5.187
31	4.495	4.404	1.838	2.092	6.317	6.595	6.528
32	4.506	4.418	1.822	2.140	6.338	6.646	6.542
33	3.610	4.539	4.266	4.377	2.813	5.061	6.076
34	3.969	4.870	4.845	3.130	2.117	3.642	5.671
35	3.166	3.752	5.002	3.740	2.560	4.345	5.845
36	5.846	5.207	0.775	3.557	7.215	7.591	7.589
37	3.501	4.220	5.325	4.268	3.043	4.830	6.207
38	4.093	4.566	5.854	4.432	2.843	4.164	5.960
39	5.244	4.586	2.835	2.713	6.769	7.129	7.309
40	5.337	4.736	2.885	2.867	6.857	7.255	7.384
41	4.410	5.687	5.407	3.812	2.668	4.580	6.733
42	5.582	5.019	3.121	3.229	7.241	7.596	7.672
43	5.653	5.133	3.162	3.338	7.300	7.689	7.724
44	4.831	5.922	5.830	4.159	2.984	4.816	6.548
45	6.463	6.418	3.165	2.739	6.905	7.353	8.174
46	6.567	6.561	3.265	2.853	6.942	7.395	8.249
47	4.698	5.584	6.503	5.290	3.531	5.231	7.136
48	6.795	6.343	4.252	4.212	7.818	8.020	8.584
49	6.806	6.323	4.225	4.203	7.822	8.012	8.594
50	5.358	6.891	6.469	4.935	3.589	5.816	7.668
51	7.477	7.682	4.237	3.901	7.867	8.626	9.141

[a] Reference data taken from Ref. 519.

Table A7.21.: Relative conformer energies for conformers of a tripeptide (PCONF)[a] computed with different semiempirical methods. The values are given in kcal mol^{-1}.

	GFN-xTB	PM6-D3H4X	MSINDO-D3(BJ)H+	OM2-D3(BJ)	DFTB3-D3(BJ)	HF-3c	ref.
1	3.45	2.36	8.44	2.09	0.50	3.74	0.14
2	5.69	5.09	10.92	2.09	2.21	4.02	0.90
3	3.71	4.68	5.46	2.11	3.93	4.75	1.15
4	3.02	3.10	8.18	2.50	0.13	3.85	0.79
5	1.62	1.79	4.47	1.21	2.85	3.05	1.31
6	7.78	7.84	10.65	3.74	2.47	5.11	1.87
7	2.75	1.86	2.06	1.65	1.83	2.50	2.37
8	3.16	4.56	7.21	2.68	4.09	4.41	2.07
9	3.57	3.85	6.87	2.99	2.54	4.23	2.51
10	5.42	4.81	7.15	3.52	0.32	5.28	2.04

[a] Reference data taken from Ref. 518.

Table A7.22.: Relative conformer energies for conformers of long alkane chains (hairpin)[a] computed with different semiempirical methods. The values are given in kcal $mol^{-1}$for conformer energy of the folded (hairpin) relative to the linear *all-anti* conformer.

	GFN-xTB	PM6-D3H4X	MSINDO-D3(BJ)H+	OM2-D3(BJ)	DFTB3-D3(BJ)	HF-3c	ref.
$C_{14}H_{30}$	-0.17	0.71	7.03	0.24	-2.10	-2.46	0.04
$C_{16}H_{34}$	-0.94	-0.12	6.57	-0.74	-3.08	-3.58	-0.54
$C_{18}H_{38}$	-1.56	-0.89	6.53	-1.66	-4.20	-4.65	-1.59

[a] Reference data taken from Ref. 520.

Table A7.23.: Relative conformer energies for conformers of diuracilphosphate (UpU46)[a] computed with different semiempirical methods. The values are given in kcal mol^{-1}.

conformer	GFN-xTB	PM6-D3H4X	MSINDO-D3(BJ)H+	DFTB3-D3(BJ)	HF-3c	ref.
1	6.46	10.07	10.85	7.09	10.06	4.87
2	3.69	4.79	4.61	4.35	4.89	2.97
3	11.88	13.33	12.21	10.46	14.45	8.90
4	3.05	6.69	5.73	2.25	6.94	2.22
5	2.68	2.21	3.76	3.32	3.25	2.02
6	2.22	5.29	5.15	2.63	5.98	3.14
7	0.00	0.00	0.00	0.00	0.00	0.00
8	0.92	0.97	1.42	-1.86	2.02	0.57
9	2.36	5.17	2.80	2.67	2.89	3.32
10	6.96	10.65	11.30	6.50	13.38	7.26
11	5.00	5.31	6.49	4.24	7.65	3.96
12	11.98	13.46	13.33	11.61	16.44	11.13
13	5.60	9.07	9.21	5.17	9.13	4.82
14	13.85	17.73	15.80	13.23	18.86	14.41
15	3.38	7.82	9.76	4.15	9.22	5.15
16	4.15	4.98	4.02	2.78	5.55	5.48
17	7.83	10.24	9.27	5.82	9.77	6.84
18	2.21	4.01	5.16	1.29	3.91	3.90
19	4.72	6.83	8.18	4.78	8.95	6.43
20	4.89	8.83	9.55	4.11	9.63	5.42
21	3.53	5.32	4.71	3.60	4.84	6.70
22	7.50	10.78	10.34	6.97	9.03	5.60
23	12.31	14.87	14.43	11.42	15.49	10.42
24	5.10	7.80	6.16	4.17	7.55	6.09
25	10.16	9.66	12.12	5.55	11.19	8.48
26	9.84	12.27	9.88	8.08	11.63	10.24
27	4.97	11.24	10.22	6.45	11.21	5.68
28	7.03	11.38	11.52	8.17	12.29	9.44
29	10.62	11.29	9.47	7.71	11.06	8.03
30	12.44	15.37	7.60	10.71	12.30	11.08
31	4.62	9.77	10.13	4.81	8.64	5.84
32	10.94	14.26	15.32	10.26	14.61	13.98
33	6.80	11.07	9.80	5.35	8.40	4.92
34	7.18	9.58	8.62	5.79	10.76	6.74
35	3.32	6.84	6.66	3.14	5.87	3.25
36	9.10	7.69	14.60	8.22	10.49	14.66
37	10.00	8.81	9.28	8.28	10.52	8.04
38	3.96	6.75	7.30	3.31	7.94	5.99
39	4.48	6.20	6.58	4.75	6.81	6.73
40	5.72	7.94	6.82	4.44	9.58	6.34
41	7.35	8.00	8.55	5.96	9.40	6.82
42	15.55	16.76	15.50	13.48	17.30	12.32
43	19.46	20.05	18.85	16.81	20.97	18.20
44	15.09	19.00	18.68	14.68	19.30	14.80
45	13.72	13.55	12.22	12.02	15.78	12.48
46	15.16	19.75	17.39	16.01	19.22	16.61

[a] Reference data taken from Ref. 521.

Table A7.24.: Relative energies for isomerization reactions of large compounds (ISOL22)[a] containing main group elements. The values are given in kcal mol^{-1}.

	GFN-xTB[b]	PM6-D3H4X	MSINDO-D3(BJ)H+	OM2-D3(BJ)	DFTB3-D3(BJ)	ref.
1	54.45 (39.18)	34.76	60.27	38.23	34.93	40.59
2	27.41 (3.99)	11.69	25.00	10.01	7.55	11.68
3	56.31 (41.91)	37.93	69.70	34.71	36.78	34.94
4	43.89 (15.36)	22.91	6.61	8.51	28.90	25.89
5	25.29 (17.41)	13.26	23.98	20.46	16.11	18.79
6	-5.23 (15.88)	-2.08	-25.15	21.76	40.17	18.30
7	17.92 (18.48)	16.88	23.40	23.63	18.08	22.31
8	-0.55 (-0.14)	1.71	2.00	6.40	4.05	7.91
9	43.95 (56.40)	57.68	107.69	64.76	37.24	38.13
10	-2.99 (-3.23)	-2.32	4.84	2.48	-3.52	0.96
11	44.19 (45.10)	25.02	30.39	27.99	49.39	35.08
12	0.06 (-0.07)	10.40	4.05	8.45	6.00	5.20
13	29.16 (17.72)	26.52	64.40	10.42	25.28	3.87
14	27.10 (11.90)	20.73	29.38	17.58	26.34	22.59
15	25.86 (10.81)	6.92	11.39	11.11	4.78	11.07
16	12.17 (21.87)	10.18	24.74	31.48	28.39	26.08
17	10.74 (18.42)	26.29	32.91	0.21	19.61	17.02
18	2.40 (2.43)	4.43	4.07	2.72	4.09	4.52
19	16.43 (-0.38)	11.09	38.14	1.60	-5.54	13.38
20	-3.71 (-17.22)	-5.83	7.35	2.98	-15.69	0.48
21	24.00 (6.23)	24.45	38.44	1.19	7.05	27.61
22	17.43 (18.41)	16.25	9.70	1.06	0.00	16.05
MSD:	2.90 (-2.82)	-1.53	8.67	-2.49	-1.48	–
MAD:	9.77 (7.62)	6.97	16.19	7.24	8.01	–
SD:	12.03 (9.53)	9.67	24.00	10.73	11.09	–
MAX:	25.29 (21.38)	22.65	69.56	26.63	21.87	–

[a] Data and numbering taken from Ref. 523.

[b] Values in parentheses include atomization energy correction.

Table A7.25.: Relative energies for isomerization reactions of organic compounds (ISO34)[a]. The values are given in kcal mol^{-1}.

	GFN-xTB[b]	PM6-D3H4X	MSINDO-D3(BJ)H+	OM2-D3(BJ)	DFTB3-D3(BJ)	ref.
1	5.73 (5.80)	-11.19	-0.97	-0.27	4.56	1.62
2	21.89 (28.50)	12.20	27.95	22.47	31.75	21.88
3	-3.27 (3.22)	3.68	10.27	4.84	6.45	7.20
4	0.80 (0.75)	1.44	2.04	1.99	0.74	0.99
5	1.43 (1.17)	0.74	-0.65	1.69	0.86	0.93
6	3.42 (3.40)	2.37	2.86	3.65	3.50	2.62
7	-2.35 (7.99)	6.07	11.27	11.98	7.92	11.15
8	23.50 (21.93)	20.09	40.06	22.57	23.90	22.90
9	7.46 (7.35)	2.80	3.76	5.34	7.74	6.94
10	4.41 (3.59)	0.07	-3.22	7.84	0.66	3.58
11	3.09 (1.61)	-9.93	-20.44	11.03	-1.64	1.91
12	32.63 (42.75)	39.64	50.56	43.70	50.92	46.95
13	38.43 (37.07)	20.68	35.85	45.63	41.98	36.04
14	35.92 (45.41)	25.14	9.15	22.16	10.61	24.20
15	12.43 (12.54)	9.26	6.12	3.22	9.24	7.26
16	-7.85 (7.59)	5.94	15.14	22.77	17.70	10.81
17	33.22 (34.05)	22.74	21.65	15.80	41.29	26.98
18	5.20 (5.49)	5.54	20.34	12.36	9.13	11.16
19	3.77 (3.91)	1.67	0.32	0.20	4.22	4.60
20	13.88 (13.83)	7.35	10.44	7.38	26.77	20.23
21	0.94 (0.55)	-0.10	-1.79	-1.07	0.91	0.94
22	14.56 (6.00)	5.65	6.89	-2.14	3.50	3.23
23	8.49 (8.99)	1.84	5.26	3.92	7.49	5.26
24	3.06 (3.30)	15.23	19.91	10.85	11.84	12.52
25	9.20 (30.21)	26.86	36.37	30.37	31.98	26.49
26	10.54 (10.06)	15.82	29.66	21.66	18.31	18.16
27	64.48 (62.07)	77.58	63.26	58.42	66.40	64.17
28	12.71 (37.18)	20.79	33.35	41.44	29.54	31.22
29	22.74 (14.81)	15.02	35.79	19.26	14.04	11.90
30	16.12 (15.64)	14.33	15.79	8.69	10.36	9.50
31	22.57 (15.11)	16.18	29.81	17.41	18.04	14.05
32	18.76 (-3.29)	4.49	5.21	-11.17	3.58	7.10
33	-1.48 (-0.18)	14.09	21.32	0.54	10.45	5.62
34	10.08 (9.32)	0.89	0.22	5.28	11.79	7.26
MSD:	-1.20 (0.30)	-2.72	1.65	-0.52	1.45	–
MAD:	6.46 (4.12)	5.12	6.64	4.55	3.37	–
SD:	8.63 (5.87)	6.13	9.16	6.32	4.70	–
MAX:	18.66 (21.21)	15.36	23.89	18.27	14.31	–

[a] Data taken from Ref. 522.

[b] Values in parentheses include atomization energy correction.

Table A7.26.: HOMO-LUMO gaps (in eV) computed with GFN-xTB and PBE/def-TZVP for a set of compounds with valence-type HOMOs and LUMOs. The same set of molecules has been used as a crosscheck for excitation energies computed with sTDA-xTB[a].

		GFN-xTB	PBE/def-TZVP
1	$MePC_2H_4$	5.37	5.82
2	P_2H_4	5.13	5.31
3	CF_3COOH	5.20	5.23
4	Si_4H_8	2.43	4.78
5	saccharin	3.74	3.75
6	hexatriyne	3.90	3.77
7	ASS	3.77	3.87
8	purine	3.13	3.52
9	dithiacyclohexane	3.95	3.72
10	fluorisochinoline	3.26	3.51
11	bisthiophene	2.85	3.02
12	c-propenone	3.62	3.26
13	terpyridine	2.94	2.82
14	silabenzene	2.74	3.70
15	$B(C_6F_5)_3$	2.88	2.66
16	corannulene	2.92	3.14
17	S0904 dye	2.39	2.48
18	S2127 dye	2.16	2.24
19	acenaphthene	2.42	2.62
20	proflavine	2.06	2.28
21	S2408 dye	2.20	2.47
22	HCSOH	2.79	3.20
23	S2153 dye	2.11	2.23
24	CT5 system	1.34	1.52
25	S0491 dye	1.61	1.77
26	S2084 dye	1.22	1.54
27	thioindigo	0.92	0.90
	MSD:	−0.23	–
	MAD:	0.30	–
	SD:	0.49	–
	MAX:	2.34	–

[a] Geometries taken from Ref. 477.

Table A7.27.: Zero point vibrational energies E_{ZPVE} and free energy contributions G_{TRV} from the translational, rotational, and vibrational degrees of freedom. These are computed for the ROT34[77,504] and HMGB11[110] systems (G_{TRV} at $T = 298.15\,\mathrm{K}$ according to Ref. 42). Values are given in kcal mol^{-1}, relative deviations (in parentheses) in %.

	E_{ZPVE}		G_{TRV}	
system	GFN-xTB	PBEh-3c	GFN-xTB	PBEh-3c
		ROT34		
ethynyl-cyclohexane	110.9	110.2	90.8	90.3
isoamyl-acetate	125.4	125.0	101.9	102.2
diisopropyl-ketone	122.3	121.8	100.2	100.2
bicyclo[2.2.2]octadiene	97.4	97.5	78.8	79.0
triethylamine	126.9	126.2	105.7	105.1
vitamin C	90.9	93.1	67.7	70.0
serotonine	126.6	128.4	103.4	105.4
aspirin	95.4	97.1	71.8	73.6
cassyrane	196.4	195.4	171.1	170.6
Ac-Pro-NH2	144.8	144.3	122.3	122.2
lupinene	176.2	176.0	153.9	153.8
limonene	119.1	120.2	96.4 9	97.7
		LB12		
DIAD	470.1	465.7	441.8	437.6
FLP	369.4	375.7	317.7	325.2
DTFS	58.7	59.4	37.1	37.9
MESITRAN	183.7	183.3	160.2	160.0
S_8^{2+}	8.5	7.9	-14.7	-15.5
HAPPOD	103.8	107.9	72.9	78.7
KAMDOR	116.5	118.6	78.5	81.7
PP	161.4	161.8	136.4	137.2
BRCLNA	65.3	66.2	39.8	41.0
C_2Br_6	9.3	9.0	-16.4	-16.5
RESVAN	195.7	196.0	168.7	169.2
BHS	358.6	364.7	318.5	324.8
		HMGB11		
$As_2(CH_3)_4$	91.3	91.7	67.4	68.5
Br_2	0.5	0.5	-15.1	-15.1
Cl_2	0.8	0.7	-13.3	-13.4
Ge_2H_6	27.2	28.6	8.6	10.3
H_2S_2	11.1	11.4	-4.5	-4.2
H_2Se_2	9.7	10.0	-7.6	-7.3
$P_2(CH_3)_4$	93.3	93.7	71.2	72.1
$Pb_2(CH_3)_6$	132.8	132.8	102.7	102.7
$Sb_2(CH_3)_4$	89.2	89.9	63.7	65.0
$Sn_2(CH_3)_6$	133.4	133.4	104.8	104.4
$Te_2(CH_3)_2$	46.0	45.5	24.1	23.6
MSD (MRD):	-0.6 (0.0)	–	-0.9 (-1.5)	–
MAD (MURD):	1.2 (1.7)	–	1.4 (2.1)	–
SD (SRD):	1.9 (2.6)	–	2.1 (3.5)	–
MAX (MAXR):	6.2 (8.1)	–	7.5 (16.3)	–

Overlay of 10 larger transition metal complexes optimized by PM6-D3H4X and PBEh-3c

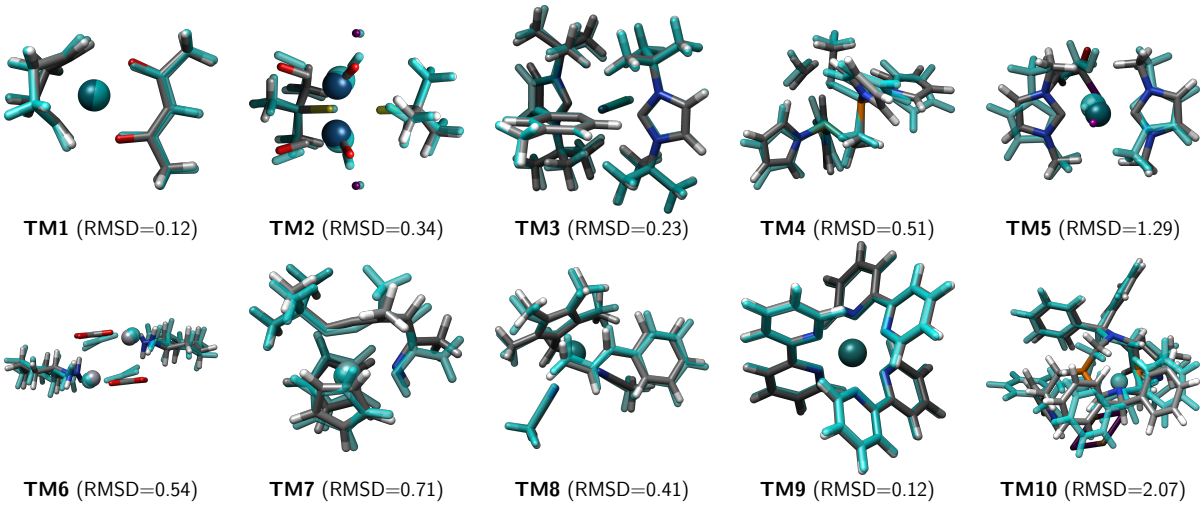


Figure A7.1.: Overlay of 10 larger transition metal complex[49] structures optimized with PBEh-3c (transparent cyan) and PM6-D3H4X (color coded atoms). The root-mean-square-deviation (RMSD) for an all-atom best fit is given in Å.

A8. Supporting Information to Part V

Appendix A8 contains:

- Computational Details for Figure 9.12

Computational Details

All ECD spectra are computed with sTDA-xTB and shifted by −0.45 eV. The minimum structure spectrum is computed on the PBE-D3-gCP/def2-SV(P) geometry used in Chapter 4 (see Appendix A3 for details). The MD simulation is run under NVT conditions for 10 ps (preceding equilibration run of 11 ps) with a timestep of 1 fs using the GFN-xTB approach, which is presented in Chapter 9. For the sTDA-xTB//MD spectrum, 100 equidistant snapshots are considered. All tight-binding calculations are performed with the `xtb` standalone program (version 4.8) in combination with the GBSA solvation model (H_2O parametrization).[490] The ECD intensities (velocity formalism) computed on the minimum geometry and on the MD snapshots are convoluted with Gaussians of 0.4 eV and 0.3 eV width at $1/e$ maximum, respectively. The latter are averaged and scaled by a factor of two. For the sTDA-xTB part, the VTB and XTB parameters of palladium were fitted and kindly provided by Jana Pisarek.

Acknowledgment

Without the support of several people, this thesis would not have been possible.

In particular, I wish to thank my supervisor Prof. Dr. Stefan Grimme for many things: the opportunity to work on interesting projects, the liberty to develop own ideas, and his scientific and financial support, which allowed me to present my work at many occasions.

I am grateful to Prof. Dr. Thomas Bredow for being my second referee, but also for all scientific and non-scientific interactions I had with him.

I thank Fabian Bohle, Dr. Jan Gerit Brandenburg, Markus Bursch, Eike Caldeweyher, Dr. Andreas Hansen, Jana Pisarek, and Jakob Seibert for proof-reading parts of this thesis. Additionally, I am grateful to my current and former colleagues Dr. Jan Gerit Brandenburg, Eike Caldeweyher, Dr. Andreas Hansen, Jana Pisarek, Jakob Seibert, and Dr. Philip Shushkov for the great teamwork during joint projects.

Next, I wish to thank all of my collaborators outside the Grimme group that I had the pleasure to work with. In the context of this thesis, I would like to express my gratitude to Prof. Dr. Arne Lützen and Dr. Christoph Gütz for the fruitful collaboration.

I am also grateful that I had the opportunity to meet and work with Prof. Dr. Tom Ziegler, Prof. Dr. Sergey A. Katsyuba, and Prof. Dr. Jan Kubelka during their visits in Bonn.

Throughout my studies, I had the pleasure to work in a welcoming and scientifically productive atmosphere. Hence, I want to thank all past and current members of the Grimme group for contributing to it in their own way. I often reminisce about the summer schools, which I attended together with my valued colleagues Christoph Bauer, Dr. Jan Gerit Brandenburg, and Dr. Rebecca Sure. Furthermore, I enjoyed sharing an office with Markus Bursch, Azaline Dunlap-Smith, Dr. Philip Shushkov, and Dr. Elena E. Zvereva. For being a great colleague, advisor, and friend during my entire studies, I want to thank Dr. Andreas Hansen.

Special thanks go to my girlfriend Sibel Ciftci for her ongoing support and patience. I also want to take this opportunity to thank my family: my sister Carolin, my brothers Matthias and Sebastian, and in particular my parents, for celebrating the moments of joy with me, but even more, for their continuous support during the stressful times. Without Sibel and my family, my studies would not have been possible in this way.