



**MISSION
INNOVATION**

accelerating the clean energy revolution

POA MATERIALI AVANZATI PER L'ENERGIA

**PROGETTO IEMAP - Piattaforma Italiana Accelerata per i Materiali per
l'Energia**

Implementazione della piattaforma
computazionale per il calcolo automatico di
migliaia di materiali con applicazioni nel
fotovoltaico

Ivan Infante, Liberato Manna

D1.19, Implementazione della piattaforma computazionale per il calcolo automatico di migliaia di materiali con applicazioni nel fotovoltaico

Ivan Infante, Liberato Manna

Febbraio 2023

Report MISSION INNOVATION

Ministero dell'Ambiente e della Sicurezza Energetica - ENEA

Mission Innovation 2021-2024 - I annualità

Progetto: Piattaforma accelerata per i Materiali per l'Energia

Work package: IEMAP: Italian Energy Materials Acceleration Platform

Linea di attività 1.16: Predisposizione codici e workflow per l'esplorazione di nuovi materiali adatti al PV

Responsabile del Progetto: Massimo Celino, ENEA

Responsabile della LA: Manna, Liberato, IIT

Indice

1	Introduzione	4
2	Un database di leganti... ..	5
3	... e di loro proprietà chimico-fisiche	6
3.1	Interazioni nanocristallo-legante	6
3.2	Interazioni steriche legante-legante	7
3.3	Interazioni legante-solvente	8
4	Conclusione	8
5	Abbreviazioni.....	9
6	Riferimenti Bibliografici	9

1 Introduzione

I nanocristalli (NC) semiconduttori colloidali, anche noti come “quantum dots”, sono emersi come materiali all'avanguardia nella moderna ricerca sui semiconduttori, soprattutto per la possibilità di controllare il loro bandgap e il loro colore di emissione in base alle loro dimensioni e per la loro facile (e a basso costo) processabilità in soluzione [1]. Queste caratteristiche conferiscono ai NC semiconduttori il potenziale di rivoluzionare molti settori della tecnologia, che vanno dai displays elettroluminescenti e a cristalli liquidi [2,3], lasers [4] e imaging infrarosso [5] alle applicazioni fotovoltaiche [6,7].

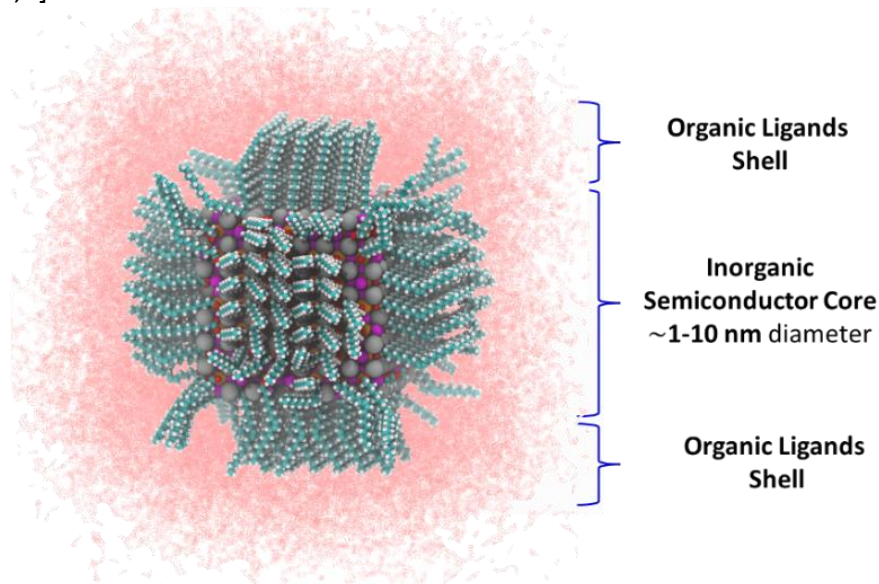


Figura 1: Rappresentazione di un nanocristallo colloidale di perovskite ricoperto da leganti oleati e sospeso in ottadecene. Il codice colore utilizzato per gli atomi è Cs, grigio; Pb, arancione; Br, magenta; O, rosso; C, ciano; H, bianco; il solvente è rappresentato da punti rosa.

I NC di semiconduttori colloidali possono essere descritti come materiali ibridi inorganici-organici composti da un “core” di materiale semiconduttore di dimensioni nanometriche ricoperto da una “shell” esterna di molecole organiche, chiamate anche leganti (**Figura 1**) [8]. Il “core” inorganico è la parte elettronicamente attiva che determina le caratteristiche optoelettroniche del materiale complessivo, mentre il ruolo principale della “shell” di leganti è quello di stabilizzare il NC in diversi tipi di solventi organici, un aspetto chiave per la sua processabilità in soluzione. Una seconda -e ancora più importante- funzione dei leganti organici consiste nella capacità, ad alte coperture superficiali, di passivare i difetti superficiali, migliorando così le prestazioni optoelettroniche del NC. Questi due aspetti possono essere formalizzati in un unico quadro teorico (**Figura 2**), in cui i leganti ideali dovrebbero favorire lo stato intermedio spingendo contemporaneamente l'equilibrio tra “binding” e distacco dei leganti verso la massimizzazione della copertura superficiale e l'equilibrio tra precipitazione e dissoluzione dei NC verso un'ottimizzata stabilità colloidale.

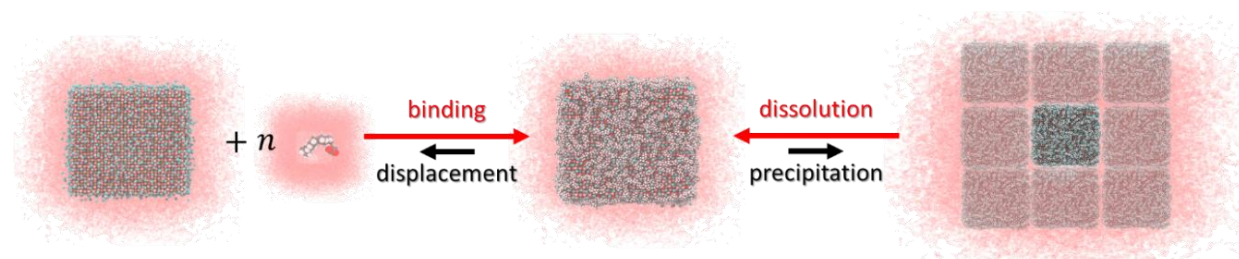


Figura 2: Meccanismo di “binding”/distacco dei leganti e precipitazione/dissoluzione dei NC.

Sebbene possiamo aspettarci che per ogni composizione di NC e per ogni solvente possa esistere un legante (o una combinazione di leganti) perfetto(i), la loro ricerca tra l’enorme numero di possibili molecole organiche che possono essere ancorate alla superficie di un NC rappresenta una formidabile sfida. Una fonte di ispirazione per affrontare questo problema proviene dal campo del “drug discovery”, dove la ricerca dei migliori farmaci si basa attualmente su modelli di “machine learning” (ML) addestrati *ad hoc* su dati provenienti da calcoli “high-throughput” di struttura elettronica, un approccio che permette di mantenere la massima precisione, ottimizzando al contempo le risorse.

Il successo del trasferimento di queste metodologie nel campo dei nanocristalli colloidali si basa sulla ricerca e il calcolo di descrittori specifici, cioè proprietà chimico-fisiche intrinseche dei leganti, che definiscano al meglio la chimica della superficie del NC tenendo conto delle varie interazioni dei leganti: (i) con il “core” del NC; (ii) tra loro sulla superficie del NC; e (iii) con il solvente. La principale sfida di questa attività è di calcolare tali descrittori in modo rapido ed efficiente senza perdere precisione, al fine di produrre un dataset di proprietà chimico-fisiche di circa diecimila strutture in un tempo pratico di settimane/mesi. Questo insieme di dati è destinato a essere utilizzato come set di addestramento per istruire algoritmi di ML, permettendo così di effettuare procedure di “screening” di grandi database alla ricerca di candidati leganti con le caratteristiche desiderate senza doverle calcolare esplicitamente.

2 Un database di leganti...

Si è scelto il database “Compounds” di PubChem, un database pubblico che elenca sostanze chimiche disponibili, come punto di partenza per questa attività. Innanzitutto, sono state selezionate solo molecole contenenti un singolo gruppo funzionale, scelto tra i noti gruppi di ancoraggio dei leganti (acido carbossilico, acido fosfonico, acido solfonico, tiolo, ammina primaria, ammina quaternaria, ...) tramite il dedicato workflow di Flamingo. Questo filtro ha permesso di ridurre il set di dati dalle $\sim 160 \cdot 10^9$ specie incluse nel database iniziale a ~ 37 milioni di molecole candidato che possono effettivamente comportarsi come leganti (**Figura 3**).

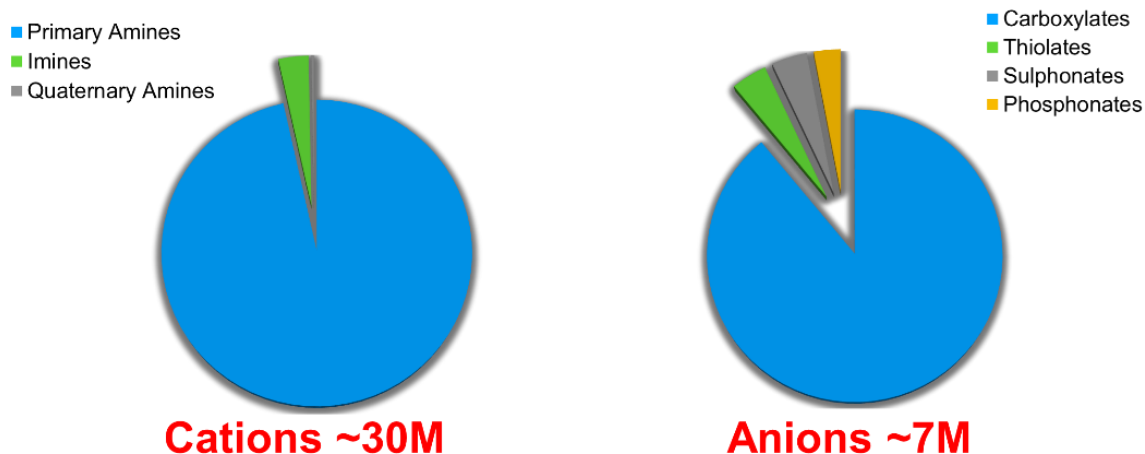


Figura 3: Distribuzione delle ~37 milioni di molecole candidato in base al loro gruppo funzionale.

Questa prima fase di selezione è stata gestita da Flamingo, una libreria basata sul linguaggio di programmazione Python che esegue un processo di scansione "smiles to filter", escludendo (o includendo) un insieme di molecole in base a loro caratteristiche strutturali come i gruppi funzionali o a loro proprietà derivate come l'indice di sintetizzabilità.

3 ... e di loro proprietà chimico-fisiche

Questa attività prevedeva inoltre di fornire, per ogni legante del database, alcune proprietà chimiche e fisiche specifiche che tengono conto delle interazioni con il NC, tra leganti, e con il solvente e altre proprietà rilevanti come punto di ebollizione e di fusione, ecc.

In questa seconda fase, per velocizzare e sincronizzare tutti i calcoli richiesti e raccogliere in modo efficiente tutti i dati prodotti sono stati sfruttati (e estesi) Compound Attachment Tools (CAT) nano-CAT and data-CAT, tre pacchetti software correlati scritti in linguaggio di programmazione Python.

3.1 Interazioni nanocristallo-legante

Per tenere in considerazione le interazioni nanocristallo-legante, è stato effettuato un benchmark della "binding energy" calcolata a livello di density functional theory (DFT) tramite il dedicato workflow di CAT. In particolare, si è stimata la "binding energy" tra due tipi di "core" (CsPbBr_3 e CdSe) e una serie rappresentativa di leganti che presentano: (i) i diversi gruppi di ancoraggio selezionati e (ii) catene alchiliche con diversi effetti elettronici. L'analisi di tali "binding energy" ha rivelato che le interazioni nanocristallo-legante sono principalmente determinate dal tipo di gruppo di ancoraggio già selezionato e possono quindi essere ritenute, in prima approssimazione, indipendenti dalla catena alchilica (**Figura 4**).

Il calcolo delle "binding energy" con il workflow di CAT è stato preceduto da un'ottimizzazione di geometria del solo "core" del NC, del solo legante, e del complesso NC-legante a livello di teoria DFT/PBE con il set di base DZVP [9,10] utilizzando CP2k [11].

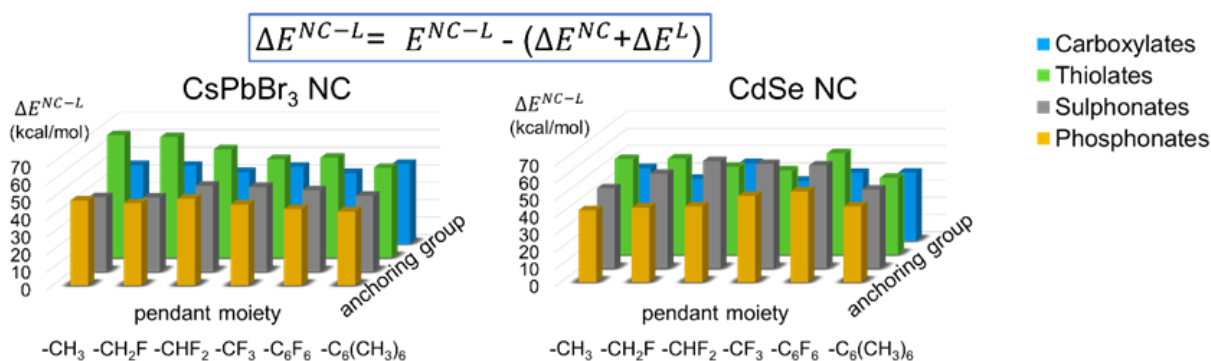


Figura 4: Benchmark della “binding energy” calcolata a livello DFT, tra due tipi di “core” di nanocrystallo CsPbBr₃ (sinistra) e CdSe (destra) e una serie rappresentativa di leganti che presentano i gruppi di ancoraggio selezionati e catene alchiliche con diversi effetti elettronici. Le interazioni nanocrystallo-legante si rivelano essere principalmente determinate dal tipo di gruppo di ancoraggio selezionato.

3.2 Interazioni steriche legante-legante

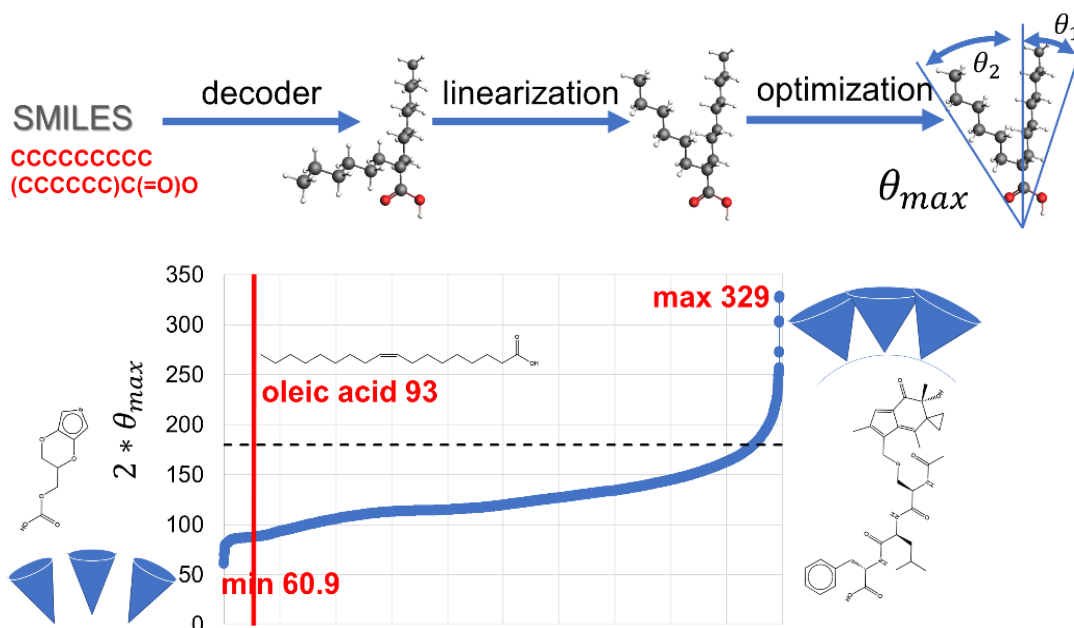


Figura 5: (sopra) Workflow di nano-CAT per il calcolo dell'angolo di cono, che prevede una preliminare linearizzazione o “biased conformational search” implementata in CAT e ottimizzazione della geometria. (sotto) Distribuzione dell’angolo di cono in una decina di migliaia di molecole. L’angolo di cono dell’acido oleico, il legante più comunemente utilizzato, è evidenziato in rosso e potrebbe servire di riferimento per un successivo filtraggio del database.

Per stimare le interazioni steriche tra leganti sulla superficie del nanocrystallo, è stato usato l'angolo di cono, un concetto ispirato alla chimica organo-metallica. Poiché per ottenere conformazioni molecolari realistiche si sono dimostrate necessarie procedure preliminari di linearizzazione e di ottimizzazione della geometria, l'angolo di cono è stato calcolato solo per circa 10 000 leganti, estratti in modo “random” dal database di riferimento utilizzando il dedicato workflow di nano-CAT (Figura 5). Questa serie di dati è destinata a essere utilizzata come set di addestramento/test per istruire algoritmi di ML, consentendo in ultima battuta di prevedere l'angolo di cono dell’intero set

di molecole candidato senza dovere effettuare esplicitamente le (lunghe) procedure preliminari richieste.

Il calcolo dell'angolo di cono con il workflow di nano-CAT è stato preceduto da una linearizzazione o "biased conformational search" implementata in CAT e da un'ottimizzazione della geometria a livello di teoria DFTB con il set di parametri GFN1-xTB [12] utilizzando Amsterdam Density Functional (ADF) [13].

3.3 Interazioni legante-solvente

Infine, l'energia di solvatazione e i coefficienti di attività in vari solventi organici come anche il loro logP sono stati selezionati come descrittori delle interazioni legante-solvente. Queste proprietà termodinamiche macroscopiche sono state ottenute sfruttando Fast-sigma, un programma di stima delle proprietà necessarie ai calcoli COMSO-RS (CONductor-like Screening MOdel for Realistic Solvents). Con l'ausilio del dedicato workflow di nano-CAT, è così stato possibile ottenere rapidamente le proprietà di solvatazione dell'intero database di molecole (Figura 6) come anche altre loro proprietà rilevanti come punto di ebollizione e di fusione, ecc. E' tuttavia da notare una perdita di precisione per i descrittori degli acidi fosfonici e solfonici, due gruppi funzionali che non rientrano nel training set del programma di previsione delle proprietà di solvatazione impiegato.

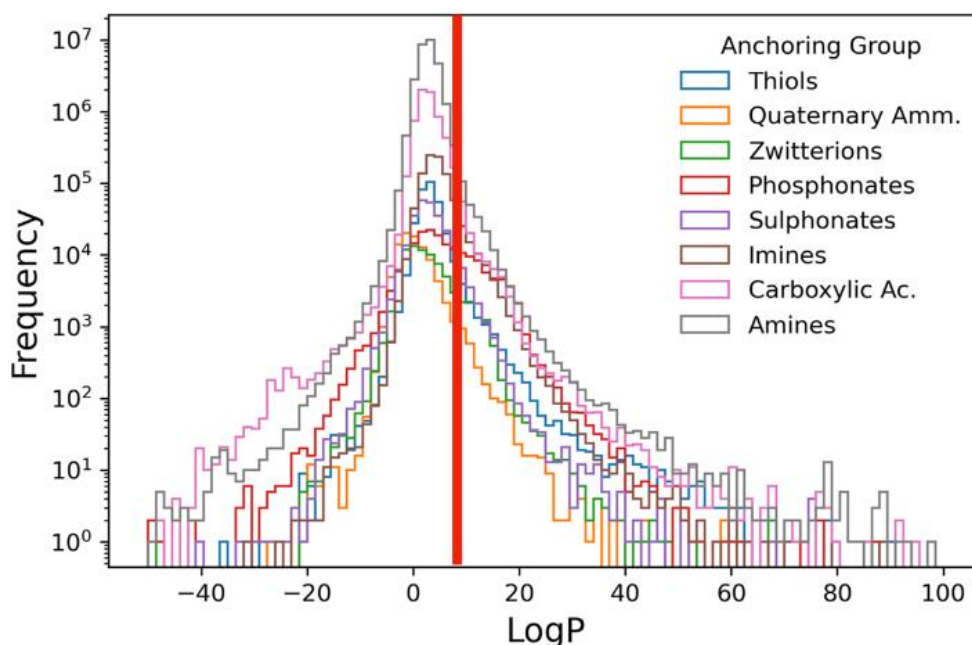


Figura 6: Distribuzione del LogP per le molecole del database in base al loro gruppo funzionale. Il valore di logP dell'acido oleico, il legante più comunemente utilizzato, è evidenziato in rosso e potrebbe servire di riferimento per un successivo filtraggio del database.

La stima delle proprietà sopradescritte con il workflow di nano-CAT si è basata sui risultati ottenuti con il programma di Fast Sigma implementato in ADF [13].

4 Conclusione

I nanocristalli colloidali possono essere descritti come materiali ibridi che combinano un "core" di un materiale inorganico semiconduttore e una "shell" di molecole organiche alla superficie, i

cosiddetti leganti. Un ruolo chiave svolto dai leganti in tutti questi sistemi è la stabilizzazione colloidale del “core” inorganico per prevenirne la dissoluzione nei tipici solventi organici. L'idoneità di un determinato legante dipende da diversi fattori, come l'entità della forza di legame legante-nanocristallo (“binding energy”), l'impacchettamento inter-legante alla superficie e le interazioni legante-solvente. In pratica, ci si aspetta che ogni tipo di nanocristallo abbia un suo legante migliore: trovare tale legante è tuttavia laborioso da un punto di vista sperimentale, se si considerano i costi di manodopera e materiali coinvolti e il numero di prove da eseguire. Questa LA propone uno screening computazionale come un'efficiente alternativa per esplorare lo spazio ampiamente inesplorato dei leganti potenzialmente interessanti. In questa attività, introduciamo per la prima volta un database pubblico di molecole organiche, derivato da un filtraggio avanzato del database PubChem, con l'obiettivo di costruire un sottoinsieme di candidati leganti potenzialmente adatti a passivare la superficie di qualsiasi “core” di materiale inorganico. Per ogni legante presente nel database, forniamo inoltre proprietà chimiche e fisiche rilevanti e proprietà più specifiche che tengono conto delle interazioni con altri leganti e con il solvente. Queste proprietà dei leganti sono indipendenti dal materiale e possono quindi essere utilizzate come strumento di orientamento per la ricerca sui materiali inorganici rivestiti di leganti in generale.

5 Abbreviazioni

ADF = Amsterdam Density Functional
CAT = Compound Attachment Tools
COMSO-RS = CONductor-like Screening MOdel for Realistic Solvents
DFT = density functional theory
DFTB = Density Functional based Tight Binding
DZVP = double zeta for valence electrons plus polarization function
ML = machine learning
NC = nanocristalli
PBE = Perdew–Burke–Ernzerhof

6 Riferimenti Bibliografici

- [1] Talapin, D. V., Lee, J.-S., Kovalenko, M. V. & Shevchenko, E. V. Prospects of Colloidal Nanocrystals for Electronic and Optoelectronic Applications. *Chem. Rev.* **110**, 389–458 (2010).
- [2] Bourzac, K. Quantum dots go on display. *Nature* **493**, 283–283 (2013).
- [3] Manders, J. R., Bera, D., Qian, L. & Holloway, P. H. Quantum Dots for Displays and Solid State Lighting. *Materials for Solid State Lighting and Displays* 31–90 (John Wiley & Sons, Ltd, 2016).
- [4] Wise, F. Colloidal nanocrystals: Virtues of defects. *Nat. Mater.* **17**, 8–9 (2017).
- [5] H. Sargent, E. Infrared Quantum Dots. *Adv. Mater.* **17**, 515–522 (2005).
- [6] Swarnkar, A. *et al.* Quantum dot-induced phase stabilization of CsPbI₃ perovskite for high-efficiency photovoltaics. *Science (80-.)*. **354**, 92–95 (2016).
- [7] Carey, G. H. *et al.* Colloidal Quantum Dot Solar Cells. *Chem. Rev.* **115**, 12732–12763 (2015).
- [8] Kovalenko, M. V *et al.* Prospects of Nanoscience with Nanocrystals. *ACS Nano* **9**, 1012–1057 (2015).

- [9] Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **18**, 3865–3868 (1996).
- [10] VandeVondele, J.; Hutter, J. Gaussian Basis Sets for Accurate Calculations on Molecular Systems in Gas and Condensed Phases. *J. Chem. Phys.* **11** (2007).
- [11] Hutter, J. *et al.* CP2K: An Electronic Structure and Molecular Dynamics Software Package - Quickstep: Efficient and Accurate Electronic Structure Calculations. *J. Chem. Phys.* **19** (2020).
- Klamt, A. The COSMO and COSMO-RS solvation models. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **8**, e1338 (2018).
- [12] Grimme, S., Bannwarth, C. & Shushkov, P. A Robust and Accurate Tight-Binding Quantum Chemical Method, *J. Chem. Theory Comput.* **13**, 1989–2009 (2017)
- [13] te Velde, G. *et al.* Chemistry with ADF. *J. Comput. Chem.* **22**, 931–967 (2001).