

Extending quantum-mechanical benchmark accuracy to biological ligand-pocket interactions

Mirela Puleva^{1,2}, Leonardo Medrano Sandonas^{1,3,*}, Balázs D. Lőrincz^{4,5,6}, Jorge Charry^{1,7}, David M. Rogers⁸, Péter R. Nagy^{4,5,6,*}, and Alexandre Tkatchenko^{1,2,*}

¹Department of Physics and Materials Science, University of Luxembourg, L-1511 Luxembourg City, Luxembourg

²Institute for Advanced Studies, University of Luxembourg, Campus Belval, L-4365 Esch-sur-Alzette, Luxembourg

³Institute for Materials Science and Max Bergmann Center of Biomaterials, TUD Dresden University of Technology, 01062 Dresden, Germany

⁴Department of Physical, Chemistry and Materials Science, Faculty of Chemical Technology and Biotechnology, Budapest University of Technology and Economics, H-1111 Budapest, Hungary

⁵HUN-REN-BME Quantum Chemistry Research Group, H-1111 Budapest, Hungary

⁶MTA-BME Lendület Quantum Chemistry Research Group, H-1111 Budapest, Hungary

⁷Luxembourg Researchers Hub asbl, 223 rue de Luxembourg, L-4222, Esch-sur-Alzette, Luxembourg

⁸National Center for Computational Sciences, Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA

*Corresponding authors: Leonardo Medrano Sandonas (leonardo.medrano@tu-dresden.de), Péter R. Nagy (nagy.peter@vbk.bme.hu), Alexandre Tkatchenko (alexandre.tkatchenko@uni.lu)

ABSTRACT

Predicting the binding affinity of ligands to protein pockets is key in the drug design pipeline. The flexibility of ligand-pocket motifs arises from a range of attractive and repulsive electronic interactions during binding. Accurately accounting for all interactions requires robust quantum-mechanical (QM) benchmarks, which are scarce for ligand-pocket systems. Additionally, disagreement between “gold standard” Coupled Cluster (CC) and Quantum Monte Carlo (QMC) methods casts doubt on many benchmarks for larger non-covalent systems. We introduce the “QUantum Interacting Dimer” (QUID) benchmark framework containing 170 non-covalent (non-)equilibrium systems modeling chemically and structurally diverse ligand-pocket motifs. Symmetry-adapted perturbation theory shows that QUID broadly covers non-covalent binding motifs and energetic contributions. Robust binding energies are obtained using complementary CC and QMC methods, achieving agreement of 0.5 kcal/mol. The benchmark data analysis reveals that several dispersion-inclusive density functional approximations provide accurate energy predictions, though their atomic van der Waals forces differ in magnitude and orientation. Contrarily, semiempirical methods and empirical force fields require improvements in capturing non-covalent interactions (NCIs) for out-of-equilibrium geometries. The wide span of NCIs, highly accurate interaction energies, and analysis of molecular properties take QUID beyond the “gold standard” for QM benchmarks of ligand-protein systems.

Accurate computational modeling of physicochemical phenomena in protein-ligand systems is vital for accelerating the early stages of the drug development pipeline^{1–4}. Reliable and “clean room” experimental measurements of the binding affinity can be costly due to multiple factors, *e.g.* dissolved electrolytes, solvent concentration, and target protein misfolding⁵, which makes robust computational methods crucial for improving efficiency and gaining detailed mechanistic understanding into the ligand-protein systems⁶. Understanding and consequently controlling non-covalent interactions (NCIs) in a targeted way can aid the compound design process to achieve optimal target selection⁷. Hence, an in-depth description of NCIs – dominant interactions determining structural configuration and ligand-pocket binding mechanism – is indispensable for binding affinity simulations. Accurate calculations are indeed critically important as even errors of 1 kcal/mol can lead to erroneous conclusions about relative binding affinities⁸.

Molecular mechanics (MM) force fields (FF), along with docking and free-energy methods, have been widely used due to their affordable computational cost for estimating structural and thermodynamical properties of complex (bio)molecular systems^{9,10}. Despite significant progress in the development of accurate and polarizable MMFFs^{11,12}, most of them treat ubiquitous non-covalent polarization and dispersion interactions using effective pairwise approximations, often resulting in inaccuracies or lack of transferability between different chemical subspaces, both being prerequisites for accurate novel drug predictions^{13,14}. Meanwhile, a broad range of quantum-mechanical (QM) methods have become available with different trade-offs between accuracy and size spanning from the less expensive but more approximate semiempirical (SE)^{15,16} approaches, passing through the density functional theory (DFT)^{17–25} ones to the “gold standard” Coupled Cluster (CC)²⁶ and Quantum

Monte Carlo (QMC)^{27–29} methods. However, achieving a sufficiently reliable and reproducible QM description of NCIs remains computationally prohibitive for realistic ligand-pocket systems, preventing further development of accurate and efficient free-energy simulation methods as well as enhanced mechanistic models for ligand-protein design.

Aiming to close this gap and improve our understanding of ligand-pocket binding, here we develop the “Quantum Interacting Dimer” (QUID) benchmarking framework. At this first stage, QUID contains 170 chemically diverse large molecular dimers (42 equilibrium and 128 non-equilibrium) of up to 64 atoms, including the H, N, C, O, F, P, S, and Cl chemical elements, encompassing most atom types of interest for drug discovery purposes. The selection of ligand-pocket motifs is accomplished through exhaustive exploration of different binding sites of nine large flexible chain-like drug molecules from the Aquamarine dataset³⁰ systematically probed with a small monomer, once with benzene (C₆H₆) and once with imidazole (C₃H₄N₂). Eight non-equilibrium conformations are also generated per each for a selection of 16 equilibrium dimers, sampling along the non-covalent bond dissociation direction. Given the structural and chemical diversity of the resulting conformations, a single dimer can exhibit multiple types of steric effects and NCIs simultaneously, including, but not limited to, polarization, π – π stacking, hydrogen and halogen bonds.

The design process of QUID was inspired by several landmark QM datasets. Most of them are focused on DFT-based physicochemical properties of single molecules up to a few hundred atoms^{30–34}. Only a limited number investigate NCIs, via the interaction energies (E_{int}), in molecular systems at the benchmark *ab initio* level of CCSD(T)/CBS. Among them, one can find the well-established S22^{35,36} and S66(x8)^{37,38} datasets; L7 with a few specific larger systems³⁹, as well as the newer NENCI⁴⁰, DES370K⁴¹, and SAPT10K⁴² ones with improved chemical diversity. Specifically for modeling ligand-pocket interactions, a recent dataset Splinter⁴³ has been developed – in it two distinct small molecules represent common fragments in proteins and small-molecule ligands. While Splinter features charged monomers and good chemical diversity, its compounds are all of similar size, up to ≈ 40 atoms, thus offering limited venue for reproducing size-dependent NCIs or geometric arrangements typical of ligands in a pocket. Furthermore, large QM datasets of energies and atomic forces have been generated for non-covalent systems by combining structural data from MMFF-based simulations with DFT methods^{44–47} to develop more robust machine-learned FFs for (bio)molecular simulations.

Our QUID framework aims to redefine the state-of-the-art in benchmarking NCIs in complex molecular systems. First, we define a “platinum standard” for ligand-pocket interaction energies (not to be confused with the platinum standard as used in the quantum chemistry community). It is obtained by establishing a tight agreement between two completely different “gold standard” methods for solving the Schrödinger equation: LNO-CCSD(T) and FN-DMC, thereby largely reducing the uncertainty in highest-level QM calculations. Second, the analysis of interaction components allows us to describe a wide range of NCIs of relevance to ligand-protein systems. Third, a comprehensive analysis of approximate empirical, semiempirical, and first-principles calculations allows us to pinpoint improvements required in each of these methodologies to move towards trustworthy free-energy simulation methods. We suggest that only a comprehensive combination of such benchmark analyses enables an unbiased understanding of NCIs in realistic molecular complexes as illustrated in this work for model ligand-pocket systems.

Results

Quantum-mechanical exploration of binding interactions. Modeling the NCIs of a ligand in a protein pocket is essential for determining the structural arrangements of natural enzyme substrates and drug candidates. This is why the QUID model systems, which represent the most frequent ligand-pocket interaction types, were created to investigate the impact of adequately describing NCIs on binding features and the influence of structural binding conformations on electronic properties. Each QUID dimer comprises a large monomer as a host and a small monomer representing a ligand motif. To achieve a proxy model representation of the interactions on the pocket-ligand surface, the large monomers are chosen from chemically diverse drug-like molecules of ~ 50 atoms, with flexible chain-like geometry allowing for folding and multiple accessible binding sites (aromatic rings) (see Fig. 1). In doing so, nine molecules (including C, N, O, H, F, P, S, and Cl atoms) that met our criteria were extracted from the Aquamarine dataset³⁰. Two small monomers were selected to represent the ligand interactions: benzene, the quintessential aromatic compound present in the phenylalanine side-chain, and imidazole, present in histidine, a more reactive and also a commonly used drug motif⁵⁰. The resulting complexes represent the three most frequent interaction types appearing on the pocket-ligand surface, that is aliphatic-aromatic, H-bonding, and π -stacking, which are found in more than 100,000 interactions within PDB structures⁵¹. The QUID dimers are comprised of monomers interacting in one or more of the aforementioned ways, with many presenting non-covalent effects of mixed character, e.g. combining π -stacking and H-bonding. In each initial dimer conformation, the aromatic ring of the small monomer was aligned with that of the binding site at a distance of 3.55 ± 0.05 Å (similar to S66 dimers³⁷), and the dimer was then optimized at the PBE0+MBD level of theory.

Post-optimization, we obtained 42 QUID equilibrium dimers that were split into three categories based on the structural shape of the corresponding large monomer: ‘Linear’, in which the original chain-like geometry is mainly retained; ‘Semi-Folded’, in which parts of the large monomer are bent while other sections remain linear, and ‘Folded’, in which the big

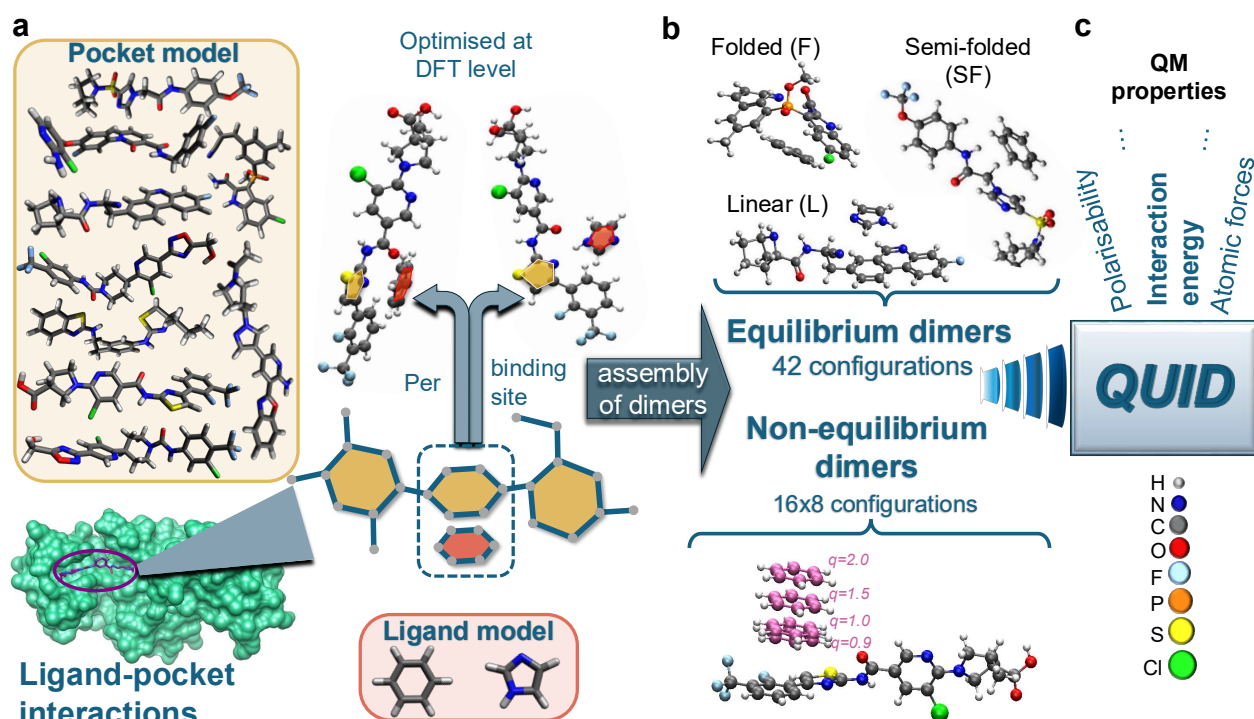


Figure 1. Schematic representation of the generation of the QUANTUM Interacting Dimer (QUID) dataset. In panel **a**, the molecules forming a QUID structure modeling a protein pocket and those modeling are presented at the top and right side of a protein-ligand complex⁴⁸ visualized with ChimeraX⁴⁹. Each dimer is composed of one of nine big monomers containing multiple potential binding sites and a small monomer binding to one of them. The resultant dimer arrangement is optimized at Density Functional Theory (DFT) level (PBE0+MBD). The resulting conformer geometries are shown in panel **b**, categorized as Linear, Semi-Folded, and Folded. For 16 equilibrium dimers, eight non-equilibrium conformations are designed along the dissociation of the non-covalent bond as illustrated by an example. q is a multiplicative dimensionless factor in the range of 0.9 to 2, which denotes the ratio of the inter-monomer distance to that of the equilibrium dimer. In panel **c**, a graphic summary of the QUID dataset with its chemical composition and some available QM (quantum mechanical) molecular properties are shown.

monomer encapsulates the smaller one. Thus, a variety of pockets with different packing densities is modeled by the QUID dimers, e.g. the folded F1I3 mimicking a more crowded binding pocket⁵², or the linear L2B1 representing a toy model of a more open surface pocket⁵³. This classification is shown in terms of the radius of gyration and chemical diversity in Fig. S1 of the Supplementary Information (SI). As a result, a wide range of interaction energies E_{int} between the monomers is produced, ranging from -24.3 to -5.5 kcal/mol at PBE0+MBD level, with imidazole usually resulting in stronger non-covalent bonds (see Fig. S2 of the SI). The prediction of the E_{int} values will be investigated in “Assessing the performance of DFT, semiempirical, and empirical methods” section with diverse computational methods.

Next, a representative selection of 16 dimers is used to construct non-equilibrium conformations along the dissociation pathway of the non-covalent bond (along π - π or H-bond vector), modeling snapshots of a ligand binding to a pocket. These conformations were generated at eight distances, characterized by a multiplicative dimensionless factor q , defined as the ratio of the inter-monomer distance to that of the equilibrium dimer. The chosen values of q are 0.90, 0.95, 1.00, 1.05, 1.10, 1.25, 1.50, 1.75, and 2.00, where $q = 1.00$ denotes the equilibrium dimer. The structure of these non-equilibrium dimers was also optimized at PBE0+MBD level with the heavy atoms of the small monomer and the respective binding site kept frozen. The resulting systems demonstrate the varied E_{int} spectrum for different pocket types via the structure categories in equilibrium and along the dissociation paths (see Fig. S2 of the SI). The generation protocol for the 42 equilibrium and 128 (16x8) non-equilibrium dimers is schematically outlined in Fig. 1 and detailed in the “Methods” section. These model systems represent a significant step forward in accurately investigating ligand-pocket interactions, characterized by robustly optimized molecular dimers that exhibit chemical diversity, larger size, and complex binding conformations.

Analysis of non-covalent interaction components. A detailed characterization of the physical interactions in ligand-pocket systems is needed to both aid understanding and ensuring diversity in the coverage of interactions. To that end, a decomposition

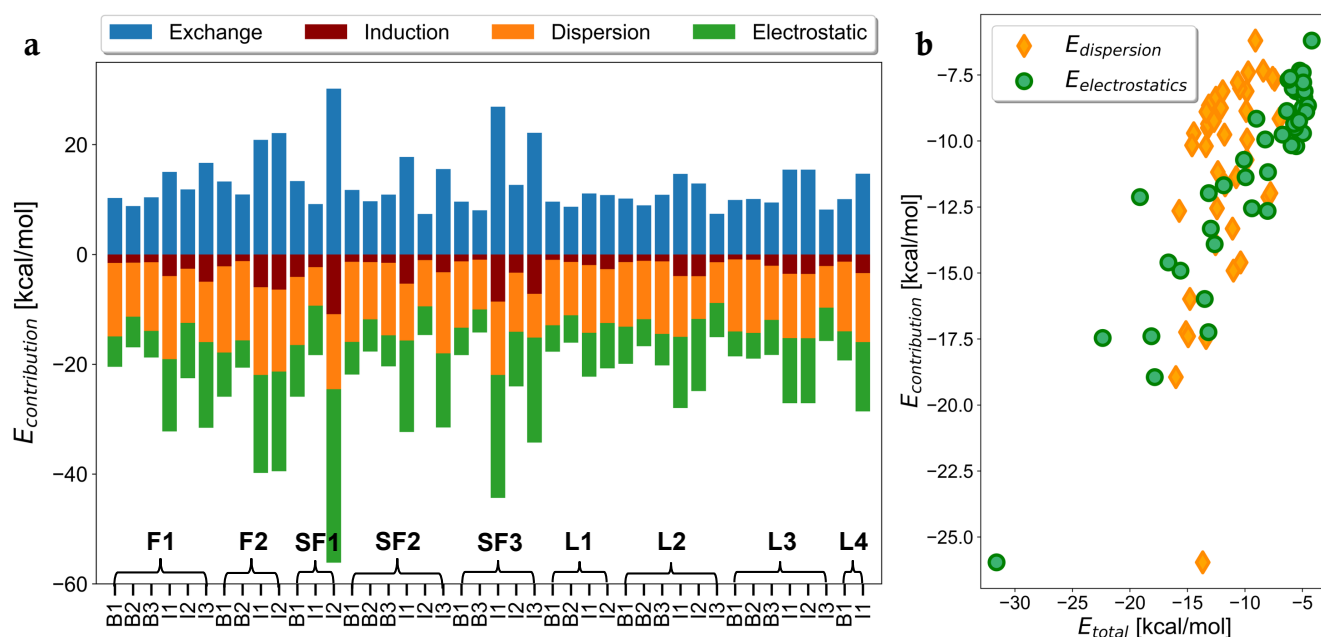


Figure 2. Symmetry-Adapted Perturbation Theory (SAPT) energy decomposition analysis. **a** The Exchange, Induction, Dispersion, and Electrostatic contributions to the interaction energies E_{int} of all 42 equilibrium QUANTUM Interacting Dimers (QUID) are depicted for the results obtained with SAPT level sSAPT0/jaDZ⁵⁴. The dimer names on the x axis can be read as the bracket first half followed by the tick label, *e.g.* F1 followed by B1 for the F1B1 dimer first on the x axis. **b** Scatter plot of the Dispersion and Electrostatics contributions compared to the total interaction energies is shown for all 42 equilibrium QUID dimers.

of E_{int} of QUID systems with SAPT analysis was performed, specifically with the sSAPT0 version due to its good balance between accuracy and computational cost⁵⁴, which is a slightly modified recipe that involves a rescaling of $E_{\text{exch-disp}}^{(20)}$ and $E_{\text{exch-ind}}^{(20)}$ terms based on an empirically adjusted proportion between $E_{\text{exch}}^{(10)}$ and $E_{\text{exch}}^{(10)}(S^2)$. The sSAPT0 E_{int} predictions for the equilibrium dimers were found qualitatively consistent with those computed at the LNO-CCSD(T) level (MAE of 0.85 kcal/mol). The largest discrepancies are found for dimers with imidazole as the small monomer and with both π - π stacking and H-bonding contributions to the non-covalent interaction. The most pronounced discrepancy occurs for the folded F1I1 dimer, with a value of 1.97 kcal/mol. Case-by-case results for all equilibrium QUID dimers are shown in Fig. S3 of the SI. The SAPT analysis provides insight into the energy components of the NCIs, namely induction, dispersion, electrostatic, and exchange contributions, which elucidate the balance of intermolecular interactions. This measure of dispersion and electrostatic components has been used before to gain insight into specific protein-ligand interactions^{42,55} and will provide a solid basis for interpreting the different predictions of E_{int} from different QM methods.

The variety in the spread of the sSAPT0 components for the QUID equilibrium dimers is shown in a stacked histogram in Fig. 2a. This showcases the diversity as a result of the different chemical environments and NCI types (see the associated NCI-plots⁵⁶ in Fig. S4 of the SI). Notably, the dispersion and electrostatic terms are strongly represented, while the induction contribution is the smallest, about 15-20% of the total value of E_{int} . This indicates a supplemental role of the polarization effects on one monomer as a result of permanent dipoles of the other, with the imidazole-based dimers consistently having larger induction components compared to their benzene counterparts when settled in the same environment. For 9 dimers, electrostatics are the dominant term (see Fig. S5 in SI), being particularly strong for the SF1I2 dimer. Such strong electrostatics are rare in the QUID systems as seen in Fig. 2b, where single outliers are seen for electrostatic contributions higher than -14 kcal/mol, while there is a skew towards smaller values in the range -8.0 to -4.0 kcal/mol. All 9 electrostatic-dominant cases involve imidazole as the small monomer. This is consistent with the presence of the two N atoms and their lone pair orbitals, capable of forming H-bonds and dipole-dipole interactions. For SF1I2, also a sulphonyl group is found near the binding site (see Fig. S4 in SI), and the amino H atom in imidazole is interacting with the SO₂ functional group, possibly forming an H-bond, in addition to the second H-bond between the imidazole and an amino group in the large monomer. Dispersion is the dominant component for the other 33 equilibrium dimers, as expected given the choice of binding sites (pre-optimization) on accessible aromatic rings arranged for π - π stacking. The spreads of the electrostatic and dispersion components are contrasted

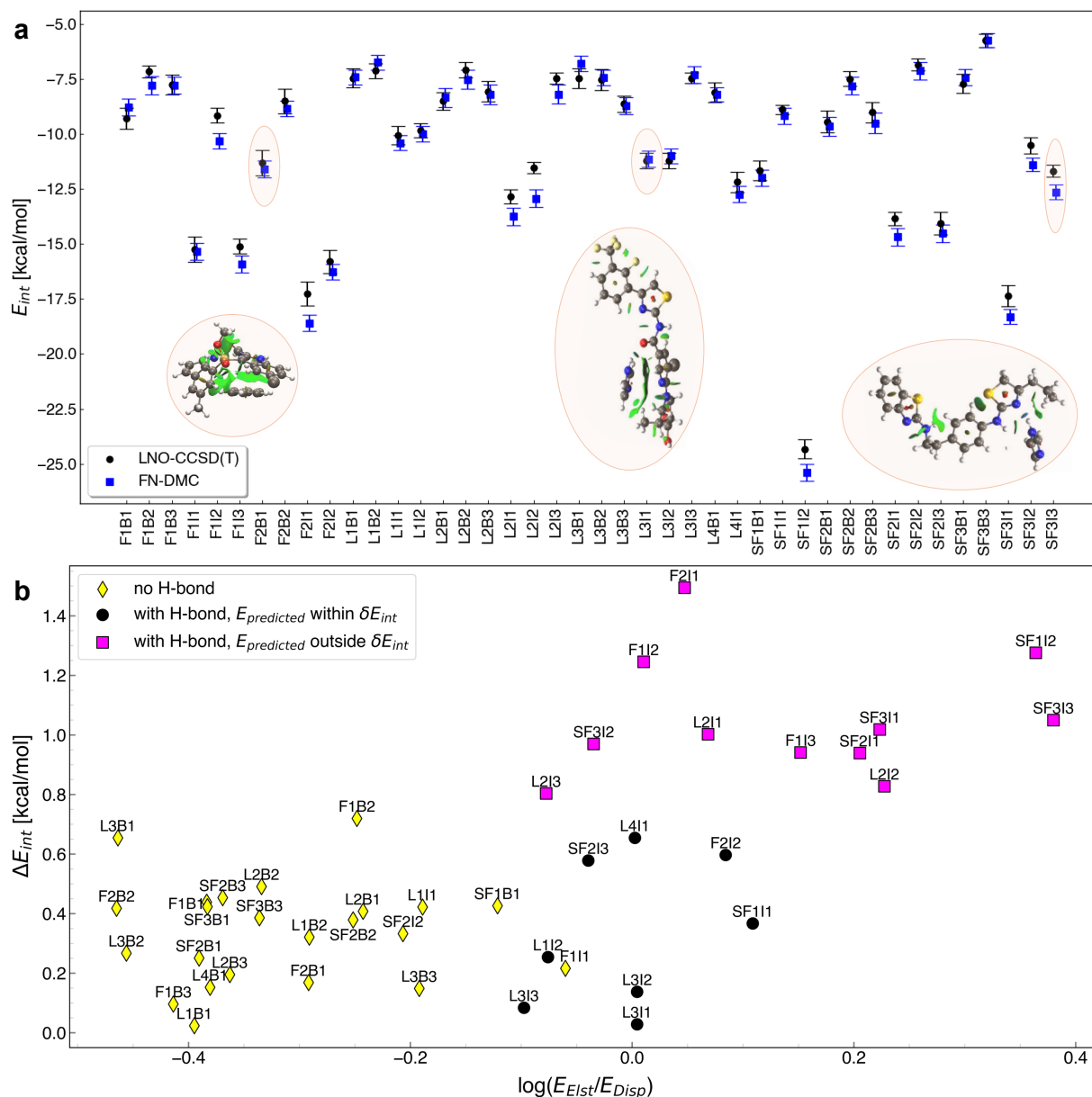


Figure 3. Towards “platinum standard” in E_{int} by benchmarking “gold standard” methods LNO-CCSD(T) and FN-DMC. **a** Comparison of the interaction energies (E_{int}) computed using Fixed-Node Diffusion Monte Carlo, FN-DMC (0.015 or 0.025 time step) and Local Natural Orbitals - Coupled Cluster with Singles, Doubles, and perturbative triplets, LNO-CCSD(T) (extrapolated to Complete Basis Set (CBS) and Local Approximation-Free (LAF) limit) for the 42 equilibrium Quantum Interacting Dimer (QUID) dataset. For FN-DMC, error bars represent estimated one- σ statistical error for 4×10^8 configurations. For LNO-CCSD(T), the error bars correspond to the estimated uncertainty from the best CBS and LAF extrapolations. Three cases are highlighted: L3I1 for which the methods are in perfect agreement, F2B1 for which the methods agree within their uncertainty estimates, and SF3I3 as the one case for which they are in slight disagreement. The NCI plots illustrating the non-covalent interactions in those molecular dimers are also shown⁵⁶. **b** Scatter plot of the absolute differences in the prediction of the interaction energies between LNO-CCSD(T) and FN-DMC, ΔE_{int} , versus the log of the ratio between the electrostatic (Elst) and dispersion (Disp) Symmetry Adapted Perturbation Theory (SAPT) components from sSAPT0⁵⁴. The equilibrium QUID dimers are divided into three subsets: yellow (no H-bond in non-covalent interaction) and black (H-bond in non-covalent interaction) symbols indicate cases where the E_{int} predictions agree within their uncertainty estimates. Pink symbols denote dimers for which the predictions do not agree within uncertainty estimates, all of these feature an H-bond between the monomers.

against the E_{int} spread in Fig. 2. E_{int} ranges from -8 to -20 kcal/mol, clearly larger than π - π stacked or single H-bonds in small dimers, which usually contribute between -2 to -8 kcal/mol to the interaction energy³⁷. Furthermore, the dispersion component around -10 kcal/mol is larger than that obtained in a benzene dimer around -5 kcal/mol⁵⁷. Some dimers can also depict mixed NCI character, *e.g.* both an H-bond and π - π stacking in F1I3, or both sandwich π - π stacking and T-shaped interaction of aromatic groups simultaneously in F1I1. Therefore, we can conclude that the ligand binding is enhanced due to the collective interactions with the pocket with SAPT characterizing the specifics of those interactions and also revealing the complexity of the physical interactions of the QUID systems.

Benchmarking the quantum-mechanical benchmarks towards “platinum standard”. Reliable models for pocket-ligand systems rest on robust methods performing consistently and accurately in such systems. A prerequisite towards understanding and estimating the performance of current methods is the existence of reliable data, which can be challenging when results in literature obtained at “gold standard” level of computation have been found to disagree^{58,59}, and comprehensive and computationally expensive studies are needed to explore sources of the discrepancy⁶⁰. Hence, to establish a thoroughly dependable reference for the interaction of ligands with a protein pocket, E_{int} for the QUID proxy systems has been obtained and compared for the two gold standard methods LNO-CCSD(T)^{61–64} and Fixed-Node Diffusion Monte Carlo (FN-DMC)^{27–29} to produce a “platinum standard”. Within our methodology, both approaches were employed with particular care to achieve convergence, see the “Methods” section for more details. The results were compared for the equilibrium set of QUID dimers and found in agreement within the uncertainty estimates of the two reference quantum methods in 31 of the 42 cases (*i.e.* 74%) as seen in Fig. 3. The MAE between the two methods is 0.47 kcal/mol compared to 0.38 kcal/mol mean absolute value of the uncertainty estimate for both FN-DMC and LNO-CCSD(T), respectively. The benchmark *ab initio* methods are in good agreement for the QUID systems, a result previously found unattainable for larger non-covalent systems with dispersion-dominated interactions, *e.g.* in the L7 dataset^{58,65}. In QUID, let us take as an example one case (SF3I3) out of the studied ones, where the LNO-CCSD(T) prediction with its uncertainty estimate lies outside one-sigma agreement with the FN-DMC result but remains within two sigmas. This still means statistical consistency considering 68% and 95% assigned to one and two sigma intervals, respectively.

Analysis of the discrepancy patterns between FN-DMC and LNO-CCSD(T) was performed by assessing the character of the non-covalent bonding from sSAPT0/jaDZ. The results were found to be consistent with a recent study performed on the S66 dimers dataset⁵⁹ - in both cases the results indicate that the dominant electrostatic component of the interaction energy correlates with the disagreements between the gold benchmark methods. Details are given in Fig. 3b, presenting a plot of the difference in the interaction energy predictions versus a log of the ratio of the electrostatic and dispersion sSAPT0 components. For the QUID systems, all the cases of disagreement involve a H-bond between the monomers, although some dimers with H-bonds in the interaction with higher dispersion contribution were found to be in agreement. These results are in line with the 0.9 kcal/mol deviation found by Shi et al.⁵⁹ for the acetic acid dimer. From this perspective, the QUID systems differ from the supramolecular complexes with extended π - π interactions, where some considerable disagreements between CCSD(T) and FN-DMC were uncovered.⁵⁸ As noted in Ref. 58 and recent studies,^{59,66–68} the FN approximation, time-step discretization, pseudopotential, post-CCSD(T) terms, basis set extrapolation used in CCSD(T), and other high-order effects could be notable for extended π - π interactions. However, beyond CCSD(T) corrections are deemed to be very small for our purposes in H-bonded dimers.^{66,68} Hence, based on these studies and our comparison between “gold standard” methods, we take LNO-CCSD(T) as a practical and reliable reference for E_{int} of ligand-pocket NCIs in the complex QUID dimers. LNO-CCSD(T) results were subsequently obtained for all 42 equilibrium dimers and the full dissociation curves of a representative selection of 6 dimers (details in “Methods” section).

Assessing the performance of DFT, semiempirical, and empirical methods. Given the “platinum standard” E_{int} reference, we next conduct a comprehensive and reliable examination of its prediction and reliability obtained from different approximate computational methods for capturing NCIs in QUID equilibrium systems. This is done with the goal of identifying approximate methods that can be used in eventually building a trustworthy pipeline for calculating binding affinities.

With the aim of providing a systematic investigation of QM and MM approximations, we include a wide selection of methods. First, we study a variety of DFT functionals (*e.g.*, global, range-separated, and double hybrids) with dispersion interactions selected from previous benchmark studies^{78,79}, namely PBE0+MBD, PBE0+D4, ω B97X-V, ω B97X-D3, ω B97M-V, PBE+MBD, PBE0+XDM, PBE-QIDH-D3, CAM-B3LYP-XDM, B3LYP-D3, M06-2X, PBE0+MBD-NL, and PBE0+TS. Second, among the SE methods we study the third-order Density Functional Tight Binding DFTB3+MBD¹⁵ and GFN2-xTB¹⁶. From the available empirical classical FFs, we included GAFF2 (computed with AMBER)¹¹ and CHARMM-CGenFF¹² (computed with OpenMM)⁸⁰. The results are presented in Table S1 of the SI and as a spread of E_{int} predictions obtained with these methods with respect to the LNO-CCSD(T) reference values, ΔE_{int} , shown in Fig. 4a. They are presented in ascending order of MAE, whose values can be found in the first column of Fig. 4b (full results in Fig. S7 of the SI). The performance must be analyzed in the context of the intrinsic uncertainty estimate of the LNO-CCSD(T) method on the QUID dataset (mean value 0.39 kcal/mol). From this perspective, PBE0+MBD, ω B97X-V, and PBE0+D4 are within the uncertainty of LNO-CCSD(T),

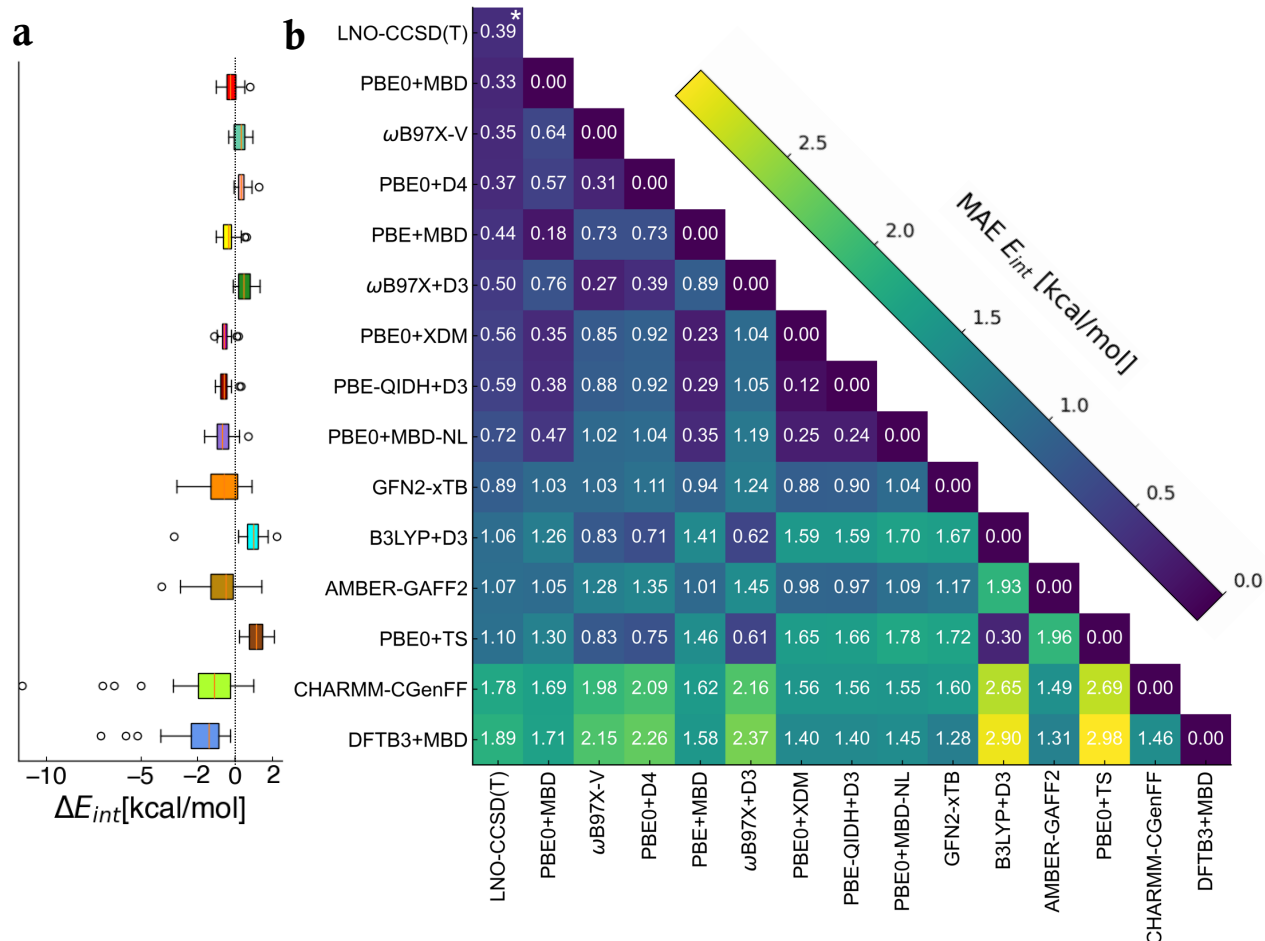


Figure 4. Comparison of interaction energy predictions using high level and approximate methods. Specifically, Density Functional Theory (DFT), semiempirical, and empirical methods are compared to Local Natural Orbitals - Coupled Cluster with Singles, Doubles, and perturbative triplets (LNO-CCSD(T)^{61–64}) and to each other. **a** Distributions of interaction energy E_{int} predictions w.r.t. LNO-CCSD(T), ΔE_{int} , showed via box plots, for a selection of computational methods - DFT methods: PBE0¹⁷+Many-Body Dispersion (MBD)^{18,19}, ω B97X-V⁶⁹, PBE0¹⁷+D4^{23,24}, PBE⁷⁰+MBD^{18,19}, ω B97X²¹+D3^{71,72}, PBE0¹⁷+eXchange-hole Dipole Moment (XDM)⁷³, PBE-QIDH⁷⁴+D3^{71,72}, PBE0¹⁷+MBD-NL²⁰, B3LYP^{75,76}+D3^{71,72}, PBE0¹⁷+TS²⁵; semiempirical methods: DFTB3¹⁵+MBD^{18,19}, GFN2-xTB⁷⁷; and classical force fields: AMBER-GAFF2¹¹ and CHARMM-CGenFF¹². The negative ΔE_{int} values signify underbinding, while the positive ones overbinding. Each boxplot shows the median of the error distribution with a vertical median, a box covering the interquartile range from the 25th to the 75th percentile is shown in colour, with whiskers spanning horizontally 1.5 times the interquartile range and data points outside that range plotted individually as outliers. **b** A heatmap of mean absolute error (MAE) values of predicted E_{int} w.r.t LNO-CCSD(T) for the 42 QUID equilibrium dimers in the first column, and the MAE of all methods w.r.t each other in subsequent columns. The computational methods to predict E_{int} were the same methods as in panel **a**. *For the LNO-CCSD(T) method, the value shown with asterisk is the mean absolute of the uncertainty estimates for E_{int} .

although a case-by-case analysis reveals deviations beyond the uncertainty of the benchmark data.

Overall, it is reassuring that all recent DFT approximations yield rather accurate results. Even if some of them underestimate (PBE0+XDM, M06-2X, PBE0+MBD-NL, PBE-QIDH-D3) or overestimate (B3LYP-D3, PBE0+TS) the reference interaction energies on average, the spread of deviations is rather narrow for *all* DFT methods (with B3LYP-D3 being the only exception in the def2-QZVPPD basis set). On the other hand, both empirical and semiempirical methods show a tendency to underbind, producing larger spreads and exhibiting large outliers. The most prominent outliers are found for the methods CHARMM-CGenFF and DFTB3+MBD, with errors ranging from -12.5 kcal/mol to -5 kcal/mol (details in Fig. S6 of the SI) and -7.5 kcal/mol to -4.5 kcal/mol, correspondingly. Examining the DFTB3+MBD outliers—SF112, SF311, and SF313 dimers—reveals that for SF112 the strongest interactions are driven by electrostatics, including contributions from a sulfonyl group at the

binding site, while the SF3Ix dimers exhibit reactive thiazole groups. It is noticeable that the error distributions of GFN2-xTB (a SE method) and AMBER-GAFF2 (an empirical FF) are quite similar, although GFN2-xTB has a slightly lower average deviation from LNO-CCSD(T). Since GFN2-xTB was partially fitted to CCSD(T) data while AMBER-GAFF was not, this is not surprising.

For AMBER-GAFF2, dimers with high electrostatic contributions result in larger errors (see Fig. S5 of the SI) pointing to a limitation of fixed partial charges. Contrarily, for GFN2-xTB, the higher errors appear to be associated with the local chemical environment. For example, the presence of P (in all 2Ix dimers), S (in all SF3Ix dimers) or Cl atoms (in both Folded F1I1, F1I3 and Linear structures L1I2, L2I2) as well as H-bonds (in L3I3) affects the bonding of imidazole ligands, presenting greater challenges for the method.

To evaluate the level and areas of agreement between methods, we have also computed the MAE values for E_{int} of QUID equilibrium dimers relative to each other (see Fig. 4b and Fig. S6 of the SI). The functionals ω B97X-D3 and ω B97X-V show excellent agreement with each other (MAE = 0.25 kcal/mol) despite the distinct incorporation of dispersion terms, D3 and non-local correlation VV10, in ω B97X-D3 and ω B97X-V, respectively^{21,69}. This indicates that the critical similarity between ω B97X-V and ω B97X-D3, which sets them apart from the Minnesota functional M06-2X or PBE0+MBD, is the range-separation treatment of the DFT functional⁸¹. Particularly important for QUID systems appears to be the long-range handling of the electron-electron interactions, as the short-range ones differ for the GGA and meta-GGA functionals⁸¹. In the same vein, we consider the related PBE0+MBD and PBE0+MBD-NL methods (MAE of 0.47 kcal/mol) - we notice that the MBD-NL method increases the deviation compared to MBD in almost all cases except for the dimer with two S atoms and imidazole ligand, SF3I1-3 (see Fig. S6 of the SI). We note that the MBD-NL functional was designed to achieve broad applicability to inorganic solids and molecular systems, while the original MBD (or MBD@rsSCS) method was developed for molecular systems. Interestingly, while the PBE0+MBD and the PBE+MBD functionals produce comparable results, with MAEs of 0.33 kcal/mol and 0.44 kcal/mol respectively, the double-hybrid PBE-QIDH-D3 functional, based on the same PBE method, achieves an MAE of 0.59 kcal/mol (2.63 kcal/mol without the D3 correction). This may be a result of the shortcomings found in MP2 contributions for large flexible π - π dispersion-dominated molecular systems⁸² and the role of van der Waals parameterizations for double hybrid functionals⁸³.

Let us now investigate more in-depth the performance of the three best performing methods PBE0+MBD, ω B97X-V, and PBE0+D4, all of which obtain E_{int} within the LNO-CCSD(T) uncertainty estimate (0.39 kcal/mol) at 0.33 kcal/mol, 0.35 kcal/mol, and 0.37 kcal/mol, respectively. As the chemical environment and energetic balance in the NCIs proved to be a more distinguishing factor for the method than the structure types, we focus on a consideration of dispersion versus electrostatics contributions to E_{int} . Overall, the MAE value of the 14 electrostatics-dominated dimers for PBE0+MBD is 0.26 kcal/mol, notably better than the PBE0+D4 results with an MAE of 0.56 kcal/mol. On the other hand, for the 28 dispersion-dominated dimers, PBE0+D4 yields 0.27 kcal/mol, while PBE0+MBD obtains a close MAE of 0.35 kcal/mol. This suggests that systems with stronger electrostatic interactions pose a greater challenge for the D4 dispersion correction. This could stem from the different underlying mechanisms of the two approaches for modeling long-range correlation effects^{18,19,24,84,85}. ω B97X-V achieves correspondingly 0.25 kcal/mol, outperforming for the electrostatics-dominated dimers but has higher MAE of 0.40 kcal/mol for the dispersion-dominated ones.

In summary, empirical and semiempirical methods have demonstrated variable performance for ligand-pocket model systems in QUID, yielding a MAE of about 1 kcal/mol or higher and exhibiting a tendency to underbind. In contrast, among the many DFT methods examined, PBE0+MBD, ω B97X-V, and PBE0+D4 proved most effective in capturing the complex QM effects contributing to E_{int} calculations, while PBE0+XDM also showed excellent performance as a pairwise dispersion method. These findings enhance our understanding of the applicability and limitations of the various investigated computational methods. However, the choice of method for simulating a ligand binding to a protein pocket should also account for non-equilibrium conformations.

Non-covalent bond dissociation pathways: non-equilibrium dimers. A key factor in modeling the dynamics of ligand-pocket systems is the capability of a physical model to investigate systems out of equilibrium accurately. To that end, we have considered six representative dissociation curves (*i.e.* F2B1, F2I1, L2B3, L2I3, SF2B2, and SF2I2) and conducted an in-depth analysis of the performance of selected computational methods: PBE0+MBD, PBE0+D4, PBE0+XDM, GFN2-xTB, DFTB3+MBD, and AMBER-GAFF2. The choice of DFT functional was based on the long-range effects in the elongated non-covalent bond regime as seen in Table S2 of the SI and allowed for direct comparison within the same functional PBE0. In Fig. 5a, we present the averaged results (over six dimers) for the 'Delta metric' Δ that measures the agreement between the dissociation curves of a given computational method and the LNO-CCSD(T) reference (see Fig. S8 of the SI). Indeed, it is evident that SE and classical FF methods, which are the tools of choice for biomolecular modeling, perform notably worse than DFT methods. The Δ values per dimer are listed in Table S3 of the SI, with the best performing equilibrium methods PBE0+MBD and PBE0+D4 achieving smaller Δ values. These findings are confirmed by analyzing the average error of E_{int} w.r.t. LNO-CCSD(T) at each q , see Fig. 5b (corresponding six individual plots are available in Fig. S9-S14 of

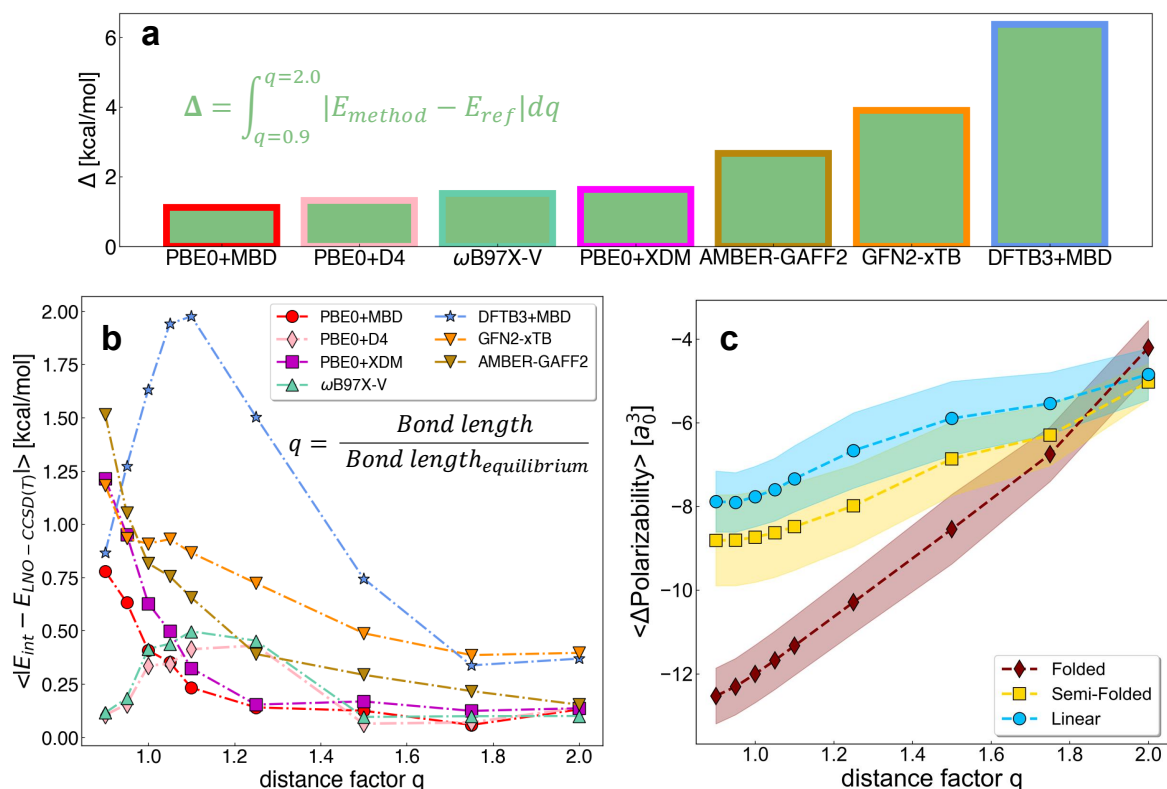


Figure 5. Interaction energy and molecular polarizability along non-covalent bond dissociations. **a** The delta metric (Δ , see formula on the plot) results are shown for four Density Functional Theory methods: PBE0¹⁷ including Many-Body Dispersion (MBD)^{18,19}, D4^{23,24}, and eXchange-hole Dipole Moment (XDM)⁷³, ω B97X-V⁶⁹; two semiempirical methods: DFTB3¹⁵+MBD^{18,19}, GFN2-xTB⁷⁷; and a classical force field method: AMBER-GAFF2¹¹. **b** Average of the absolute difference of predicted interaction energy with the Local Natural Orbitals - Coupled Cluster with Singles, Doubles, and perturbative triplets (LNO-CCSD(T)) reference along the dissociation of the non-covalent bond of a selection of dimers. The average is calculated at each multiplicative distance factor q (ranging from 0.9 to 2.0), defined as the ratio between the bond length and the equilibrium non-covalent bond length for the corresponding dimer. The average is shown for six selected molecular dimers: F2B1, F2I1, SF2B2, SF2I2, L2B3, and L2I3, using the same methods and corresponding color highlights as in panel **a**. **c** Average of the difference in the molecular polarizability of the dimer and the sum of isotropic polarizabilities of its corresponding monomers at each distance factor q . The results are shown as an average (points) with a standard error (background) over all 128 non-equilibrium dimers in the QUAntum Interacting Dimer (QUID) dataset, split by structural type in Linear, Semi-Folded, and Folded.

the SI). Notably, the performance of each method shows a strong dependence on the intermolecular distance. To elucidate the results, the dissociation curve profiles for all methods are presented individually in Fig. S9-S14 of the SI. Interestingly, unlike DFT methods, AMBER-GAFF2 either underestimates or overestimates E_{int} , depending on the dimer configuration. The discrepancies are more pronounced in configurations where dispersion components dominate the NCI, and for those dimers particularly at distances with factor $q < 1.0$, where dispersion interactions are stronger. On the other hand, both SE methods predominantly underestimate E_{int} and fail to accurately capture the position of the minimum on the dissociation curve or its overall shape. This behavior changes only for GFN2-xTB at $q < 1.0$, where it overestimates E_{int} in most cases. Hence, to assess the efficiency of the methods in different interaction regimes, two ranges have been defined: 'compressed' for $q \leq 1.0$ and 'elongated' for $q > 1.0$.

Concerning DFT methods, the best performance across the dissociation curves is consistently displayed by PBE0+MBD (underestimation) and PBE0+D4 (overestimation), with errors remaining within the uncertainty estimates of LNO-CCSD(T). The MAE values for E_{int} in the 'compressed' and 'elongated' regimes obtained using all methods are provided in Table S2 of the SI. As expected from previous results, PBE0+D4, ω B97X-V, ω B97X-D3, PBE0+D4, and PBE0+MBD yield the best results in the 'compressed' regime, while PBE0+MBD, PBE+MBD, PBE0+XDM, PBE0+D4, ω B97M-V, and ω B97X-V perform best in the 'elongated' regime, indicating the consistently good performance of the PBE0 and ω B97X functionals.

The SF2B2 and SF2I2 dimers proved to be the most challenging among those examined, likely due to the interaction of a 5-membered oxadiazole ring (C_2N_2O) via π - π stacking with the small monomer. The presence of two N and one O atoms in the aromatic ring contributes to an increase in the dipole moment and polarizability of the monomer, thereby enhancing both electrostatic and dispersion interactions. Interestingly, F2B1 and F2I1 are the easiest dimers to predict among the examined methods, as the molecular environment contributing to NCI is located within a few Å of the molecule.

While the analysis of the dissociation curves has confirmed the performance of the methods for computing E_{int} (*vide supra*), it has also revealed that the accuracy of SE methods and classical FF strongly depends on the distance range and the dimer configuration. This is a critical result, as both approaches are widely used to investigate intermolecular interactions in biomolecular simulations, raising questions about the reliability of the results obtained in molecular dynamics simulations carried out with empirical methods.

Quantum-mechanical property space of QUID systems. To enhance our understanding of the effects of dimer configuration and intermolecular distances on the properties of pocket-ligand systems, several global and local physicochemical properties, in addition to E_{int} , were computed for all equilibrium and non-equilibrium QUID dimers at PBE0+MBD level of theory (see “Methods” section for more details). A full list of properties, similar to those in the Aquamarine³⁰ dataset of large monomers, is provided in Table S5 of the SI. Further, quantities describing the Hirshfeld partitioning, *i.e.* Hirshfeld volumes, ratios, charges, (scalar) dipole moments provides information for the electron response of an atom-in-molecule environment. Overall the molecular properties can also be useful as ML descriptors⁸⁶. Here, we first focus on the molecular polarizability, α , as an additional measure of the NCIs.

To that end, analogous to E_{int} , the difference in α between each dimer and the sum of its corresponding large and small monomers, $\Delta\alpha$, was calculated for all 128 non-equilibrium conformations. The average $\Delta\alpha$ values for each structure type as a function of the distance factor q are plotted in Fig. 5c). Overall, the three structure types exhibit an almost linear behavior, with slight deviations near the equilibrium distance for the Linear and Semi-Folded structures. According to a recent concept of chemical bonding based on α , proposed by D. Hait and M. Head-Gordon⁸⁷, this linear behavior suggests no significant modification in the covalent bond arrangements along the dissociation curve. The variation can thus be attributed to the self-consistent screening effect between the monomers. Indeed, in both Linear and Semi-Folded structures, the small monomer affects fewer atoms of the large monomer. In contrast, in Folded structures, the small monomer remains within 5 Å of a significant number of atoms of the large monomer, substantially influencing the electrostatics and dispersion in the pocket site, resulting in a steeper curve. Moreover, no correlation between α with E_{int} and μ emerges from the exploration QUID dimers (see Fig. S15 of the SI). On the one hand, this could suggest flexibility for rational ligand design as observed for small molecules of up to 7 heavy atoms⁸⁶. At this stage, we can ascertain the interplay between electrostatic and dispersion interactions in a structurally and chemically complex local environment in the QUID pocket-ligand proxies requires exact QM models beyond the capture of a single key global property in the high-dimensional chemical compound space.

Another relevant property for understanding NCIs in QUID dimers is the atomic forces, which are widely used to parameterize MLFFs for (bio)molecular systems. Indeed, the molecular conformational sampling at a given temperature strongly depends on the accuracy of the chosen computational method in adequately describing the forces acting on the atoms in the molecule. Accordingly, we have analyzed the atomic force distributions using a selection of DFT methods: the best performing ones PBE0+MBD and PBE0+D4, and the well-performing pairwise PBE0+XDM method. Since these methods share the DFT functional and our primary interest lies in NCIs, the focus will be on the vdW components of the atomic forces. We take the MBD method as the reference for comparison because geometry optimizations have been carried out with the PBE0+MBD level of theory. In addition, MBD is the only method that includes non-local screening effects in the polarizability, which were shown to be important in Fig. 5c).

The differences in the vdW atomic forces are examined in terms of their magnitude and direction, treated as two distinct yet interrelated driving factors in MD simulations. The results are presented in Fig. 6a as a radial plot illustrating the difference in forces of D4 and XDM with respect to MBD. In this plot, the angular part represents the arccosine between the dot product of two force vectors, while the radius is the difference in force magnitudes, scaled by the MBD force. This analysis focuses on the differences observed per atom type, with the results for N, O, Cl, and S atoms shown in Fig. 6a (see other atom types in Fig. S16 of the SI).

The overall force analysis is provided by the plot in Fig. 6b, which shows that MBD yields the smallest forces on average. The vdW force magnitudes are generally not much higher than 1 kcal/mol/Å; this is expected given that equilibrium geometries are involved. Higher vdW forces would be expected for non-equilibrium geometries. Further, largest discrepancies in force magnitude were found in D4 compared to XDM. This is corroborated by the distribution of the atomic force magnitudes for all atoms. The most significant outliers in force magnitude discrepancies are associated with D4 and Fig. 6a demonstrates that those outliers of up to 3-4 times the magnitude of the MBD force are found on C atoms in descending order for L2I3, L2B3, L3I1, and L3I2 (Fig. S16 of the SI). The FFAST software⁸⁸ allows for visualization of the discrepancies between the vdW atomic forces to confirm that the C atoms were found at the binding site, an example of such a visualization can be seen in

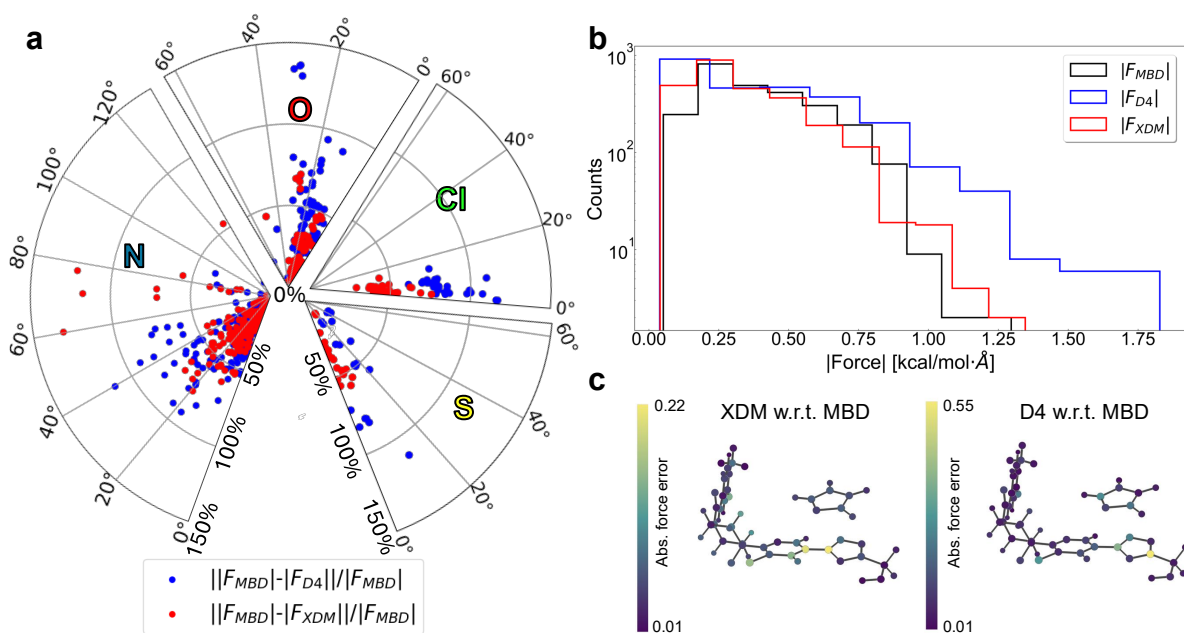


Figure 6. Atomic van der Waals (vdW) force differences between Many-Body Dispersion (MBD)^{18,19}, D4^{23,24}, and eXchange-hole Dipole Moment (XDM) methods⁷³. **a** A polar plot, where the radius is the difference in magnitude of the vdW forces (given as a percentage) and the angle is the arccosine between the dot product of the force vectors. Four such polar plots depict the distributions of the forces acting on all the N, O, Cl, and S atoms, respectively, in all 42 equilibrium QUantum Interacting Dimer (QUID) structures. **b** Overall distribution of the magnitudes of the atomic vdW forces on all atoms for all 42 equilibrium QUID dimers using the MBD, D4, and XDM methods. **c** An example of the deviation in atomic van der Waals (vdW) forces between different methods is visualized for the SF2I2 dimer (produced by the FFAST software⁸⁸). The range is represented by colors varying from purple to yellow in a viridis color bar.

Fig. 6c) depicting SF2I2. The difference in force magnitudes is visible not only on the atoms of the binding site as expected but also in proximity to it showing the differing impact of the ligand interactions in the different methods. For the SF2I2, there is a notable difference between the comparisons of XDM and D4 w.r.t. MBD (the MBD depiction is available in Fig. S17 of the SI).

This also holds true in general for the vdW forces on the Cl atoms in the QUID dimers. As seen on Fig. 6a XDM is in better agreement with MBD than D4 for Cl atoms, and the same is true for the other halogen element F, as well as the few P atoms (Fig. S16 of the SI). Hence, the presence of more electronegative atoms can hint at a systematic difference between the different vdW atomic forces. In that vein, a particular outlier is seen for the S element, with a 113% gap between the magnitude of the MBD and D4 forces for the SF1I1 dimer, in the sulfonyl group of the binding site of the imidazole ligand (Fig. S4). Interestingly, there is a split in the vdW force directions between ‘heavier’ atoms in the QUID dimers, *i.e.* O, F, Cl, S, and P and the ‘lighter’ ones *i.e.* H, C, and N. The ‘lighter’ atoms represented more in organic molecules demonstrate a larger spread of angle difference between the forces up to 120°-180°. By construction, for the pairwise D4 and XDM methods, the vdW force is a simple vector sum where all force vectors are aligned along the vector connecting pairs of atoms. In contrast, many-body effects that are treated to infinite order in MBD can thus substantially alter the force directions, and this difference is much more pronounced than for force magnitudes. This could have a visible effect on MD trajectories, although these implications remain to be assessed in a future study.

The observed difference in force directions has potential repercussions for MLFF methods, where molecular data is routinely optimized at one level of theory, method or functional or dispersion correction, and computed at a different one to serve as input or reference for energies and forces. Unfortunately, while for E_{int} we have highly accurate LNO-CCSD(T) reference data and even the ability to achieve a “platinum standard” confidence by comparing with FN-DMC, obtaining forces at benchmark *ab initio* level is prohibitively expensive. Current research in developing gradients for LNO-CCSD(T)⁸⁹ and FN-DMC⁹⁰ could facilitate a reference benchmark for the vdW forces in the future and provide a clearer picture for the accuracy of the DFT methods for MD of ligands binding to a pocket.

Discussion

The "Quantum Interacting Dimer" (QUID) benchmarking framework was developed here to redefine the state of the art in QM-based modeling of ligand-pocket motifs. Currently, QUID contains 170 structurally and chemically diverse large molecular dimers (42 equilibrium and 128 non-equilibrium) of up to 64 atoms, including the H, N, C, O, F, P, S, and Cl chemical elements, encompassing most atom types of interest for drug discovery purposes. This diversity enables a single dimer to exhibit multiple types of steric effects and NCIs simultaneously, including, but not limited to, π - π stacking, hydrogen, and halogen bonds. Accordingly, we conducted a series of analyses that provided valuable insights into inter- and intramolecular interactions of these model protein-ligand systems from a QM perspective.

Indeed, the energy decomposition of the interaction energy E_{int} , as obtained through SAPT analysis, revealed that ligand-pocket interactions are predominantly governed by dispersion and electrostatics—types of interactions often inadequately represented by MM methods. Moreover, we defined a "platinum standard" of accuracy for E_{int} of ligand-pocket interactions by contrasting the results with the "gold standard" methods such as LNO-CCSD(T) and FN-DMC. Notably, the previously reported disagreement between LNO-CCSD(T) and FN-DMC for large non-covalent systems⁵⁸ was not observed at such scale for the QUID dimers, where the disagreements are driven by predominantly electrostatic interactions as in a recent study on the S66 dataset⁵⁹, and the overall discrepancy is small, approximately 0.5 kcal/mol. Those identified dimers can serve as model systems for future exhaustive analysis in both QM methods. Thus, our findings demonstrated that among all studied MM and QM approaches, DFT methods such as PBE0+MBD, ω B97X-V, and PBE0+D4 achieve excellent agreement with the costly LNO-CCSD(T) method in the determination of E_{int} for equilibrium and non-equilibrium dimers. Additionally, we have identified certain limitations in widely used semiempirical (*e.g.*, GFN2-xTB and DFTB3+MBD) and MM methods (*e.g.*, AMBER-GAFF2 and CHARMM-CGenFF) for investigating complex non-covalent motifs, which raises questions about their reliability in binding affinity simulations. These intriguing results highlight the relevance of determining the appropriate level of theory to accurately characterize protein-ligand systems, particularly in the development of extensive QM datasets utilized in physical method benchmarking and ML-based investigations.

Furthermore, QUID provides access to a diverse set of extensive and intensive (global and local) QM properties beyond E_{int} (at PBE0+MBD level), enabling the electronic characterization of chemical environments within ligand-pocket motifs—a limitation of current benchmark frameworks, which primarily focus on structural and energetic features. The structural dependence of the polarizability change in dimers as a function of the monomer separation, as well as the lack of correlation between global electronic properties and E_{int} , offers a perspective for understanding the NCIs in protein-ligand systems. These insights can rationalize the design of drug-like molecules targeting specific pocket sites with a desired set of QM properties⁸⁶. The observed discrepancies in the van der Waals (vdW) component of the atomic forces using MBD, D4, and XDM methods also show the importance of investigating additional properties beyond the traditional E_{int} in non-covalent complexes. An inaccurate description of vdW forces can strongly impact the reaction pathway and the resulting binding pose when simulating the interaction of ligands with protein pockets. Hence, QUID has the potential to revolutionize standard procedures in approaching the modeling of ligand-pocket interactions in physical and ML-based frameworks by providing global and local electronic property data for large molecular dimers and their building blocks, which are critical for a faithful incorporation of long-range effects^{46,47}.

In summary, the QUID benchmark framework presents a rigorously designed approach for accurately analyzing ligand interactions at various protein binding sites, facilitating the development of robust QM datasets for reliable predictions of binding affinities and structural conformations. While the insights gained from this work highlight the importance of an appropriate QM description for inter- and intramolecular properties of ligand-pocket motifs, we acknowledge that further efforts should incorporate more flexible and charged pocket structures, as well as solvation effects^{30,91}. The sampled chemical space should ultimately encompass full pocket-ligand molecular systems, similar to those in MM datasets (*e.g.* PL-REX⁹² and QR50⁹³), to enhance the reliability of the findings. We expect this work to pave the way for a more informed use and refinement of physical and chemical models for simulating ligand-pocket interactions, offering particular value in fine-tuning MLFFs and ML-augmented semiempirical models, which are increasingly integrated into screening pipelines for drug discovery.

Methods

Generation procedure of QUID systems. A schematic representation of the procedure used to generate the 170 (equilibrium and non-equilibrium) QUID dimers is outlined in Fig. 1. Each QUID dimer consists of one of nine chemically diverse large monomers (of ~ 50 atoms) selected from the Aquamarine dataset of drug-like molecules³⁰, paired with a small monomer, either benzene C_6H_6 or imidazole $\text{C}_3\text{H}_4\text{N}_2$. Each large monomer features at least 2 sterically accessible aromatic rings, which serve as binding sites. The larger monomer initially has chain-like geometry, modeling a protein pocket, while the small molecule represents the ligand. The chemical composition of QUID systems reflects this purpose: each conformer contains not only C, N, O, and H atoms but also at least one of the elements F, P, S, or Cl for eight out of the nine large monomers. These elements

are incorporated into the following functional groups: substituted five- or six-membered aromatic rings, aliphatic heterocycles, ketones, ethers, hydroxyls, amines, haloalkanes, and sulfonyl. Moreover, the choice of benzene and imidazole molecules enables a comparison of the effects of small aromatic rings and amphoteric compounds (imidazole can be both H-bond donor and acceptor), which is a key characteristic for many compounds of biomedical significance such as amino-acid histidine and anti-inflammatory drugs⁵⁰.

In the initial dimer conformation, prior to structural optimization, benzene and imidazole occupy the same position relative to a given binding site, specifically at 3.55 ± 0.05 Å parallel to the site. A similar distance was used in the generation procedure of S66x8 dataset³⁸. Here, the aromatic ring of the small monomer was aligned with that of the binding site, ensuring that the distance between the corresponding pair of heavy atoms on the two rings remained within a margin of 0.05 Å. Each dimer was then optimized using non-empirical hybrid density functional theory (DFT) with many-body dispersion approach (range-separated self-consistent screening, MBD@rsSCS), namely PBE0+MBD, in conjunction with tightly-converged numeric atom-centered orbitals (“tight settings”), as implemented in the FHI-aims⁹⁴ software, version 221103. We classify the resulting structures into three categories based on their shape: ‘Linear’, where the original chain-like geometry is mainly preserved; ‘Semi-Folded’, where parts of the large monomer bend while other sections remain linear; and ‘Folded’, where the large monomer encases the smaller ones. Following this classification, the first letter of the dimer names corresponds to the structural category: ‘F’ for Folded, ‘SF’ for Semi-Folded, and ‘L’ for Linear. The small monomer is indicated by the letter ‘B’ for benzene or ‘I’ for imidazole, while the number that follows indicates the binding site (in QUID dimers as found viable post-optimization). The chemical diversity of the equilibrium QUID dimers (represented by the counts of non-H elements) is plotted in Fig. S1 against the radius of gyration, illustrating the spread of the three structural categories. The software Avogadro⁹⁵ (version 1.2.0) and Visual Molecular Dynamics (VMD)⁹⁶ (version 1.9.4a57-arm64-Rev12) were used for visualization of the molecules.

Furthermore, we acknowledge that modeling ligand-pocket systems would be incomplete without computing non-equilibrium dimer conformations, which are crucial for understanding the binding of drugs to protein pockets. To this end, we generated eight optimized out-of-equilibrium structures along the reaction pathway of the non-covalent bond for 16 representative molecular dimers (F2B1, F2I1, F2B2, F2I2, SF2B1, SF2I1, SF2B2, SF2I2, SF2B3, SF2I3, L2B1, L2I1, L2B2, L2I2, L2B3, L2I3). These non-equilibrium conformations were constructed with inter-monomer distances ranging from 3.2 Å to 7.1 Å, defined by a dimensionless multiplicative factor q , which represents the ratio of the inter-monomer distance in a given conformation to that in its equilibrium state. The chosen values of q are 0.90, 0.95, 1.00, 1.05, 1.10, 1.25, 1.50, 1.75, and 2.00, as shown in Fig. 1. For the π - π interactions between aromatic rings, the dissociation vector was defined as the distance between the plane of the aromatic ring on the large monomer and the center of mass of the small monomer. In the case of H-bonds, dissociation was measured between the proton acceptor and donor atoms. All the non-equilibrium geometries were also optimized at PBE0+MBD level with “tight” settings using FHI-Aims (version 221103). Unlike the equilibrium dimers, the heavy atoms of the binding site and small monomer were kept frozen in their respective positions during optimization. An example of the dissociation geometries for one of the non-equilibrium dimers is shown in Fig. 1.

Property calculation. The interaction energies E_{int} of QUID dimers were calculated using the supramolecular approach,

$$E_{\text{int}} = E_{\text{dimer}} - (E_{L_{\text{monomer}}} + E_{S_{\text{monomer}}}). \quad (1)$$

Counterpoise corrections were applied to PBE0+MBD, PBE-QIDH-D3, and CCSD(T) single-point calculations. The basis set superposition error was negligibly small (under 1.5%) for DFT and ca. 4% on the average for CCSD(T) when extrapolated to the complete basis set (CBS) limit (see results in Fig. S18 of the SI). To investigate the level of agreement among QM methods for calculating E_{int} of QUID dimers, we have considered a selection of well-performing hybrid, double hybrid, range-separated hybrids DFT functionals, including M06-2X²², ω B97X-D3²¹, ω B97M-V⁹⁷, ω B97X-V⁶⁹, PBE+MBD⁷⁰, PBE-QIDH-D3^{71,72,74}, B3LYP-D3^{71,72,75,76}, CAM-B3LYP-XDM⁹⁸ and BH&HLYP-XDM⁹⁹. Additionally, the PBE0 functional was combined with multiple two-body or many-body corrections: MBD^{18,19} (range-separated self-consistent screening (MBD@rsSCS) approach), MBD-NL²⁰ (Non-Local), XDM⁷³ (eXchange Dipole Moment), TS (Tkatchenko-Scheffler)-vdW²⁵, D4^{23,24}, ω B97X-D3²¹, ω B97M-V⁹⁷. These calculations were performed using either the FHI-Aims (version 221103) software¹⁰⁰ with “tight” settings, the Psi4 software^{101,102} (version 1.9.1) with the quadruple-zeta def2-QZVPPD basis set or the QChem software¹⁰³ (version 6.1) with the quadruple-zeta def2-QZVPPD basis set except in the CAM-B3LYP and BH&HLYP cases using the XDM dispersion, where the aug-cc-pVTZ basis set was used in accordance with previous studies⁷⁹ and the available parametrization from literature¹⁰⁴. The PBE-QIDH-D3 implementation made use of the counterpoise correction as needed for the double hybrid functional method using MP2 correlation in the calculation⁷⁴. All DFT calculations on Gaussian basis set employed the resolution of identity RI technique to accelerate the calculation of electron repulsion integrals. Notably, for the XDM method, a specific parametrization for the PBE0 functional was applied and then computed using FHI-aims software on “tight” settings ($a_1=0.4710$ and $a_2=2.3857$)¹⁰⁵. SAPT energy decomposition calculations were carried out at the sSAPT0/jaDZ level⁵⁴ employing the Psi4 software¹⁰¹ (version 1.9.1). At the semiempirical level, E_{int} was calculated via single-point

calculations using Density Functional Tight Binding¹⁵ DFTB3+MBD with DFTB+ software¹⁰⁶ (version 23.1) and GFN2-xTB with the xTB software⁷⁷ (version 6.7.0).

Regarding MM methods, the results for AMBER¹¹ were obtained using Openbabel¹⁰⁷ for molecular format conversion. The parametrization with AmberTools (version 23.6) and GAFF2¹¹ required manual assignment and adjustment of bonds for more complex cases, such as ring interactions, as well as modification of the self-consistent loop limits for the F2I1 dimer. The CHARMM-CGenFF¹² calculations were conducted using OpenMM⁸⁰ (version 8.1.2) following a CGenFF2¹⁰⁸ parametrization. For these calculations, manual inclusion of the dihedral angles for the flexible side chains, such as the 'C-C-N' type, was necessary. An example is the L4B1 dimer, which was assumed to exhibit relatively low flexibility due to the nature of its bonds and chemical environment. A list of the E_{int} storage details is available in Table S4 of the SI.

Additionally, the optimized structures of equilibrium and non-equilibrium QUID dimers were also utilized for more accurate QM single-point calculations using PBE0+MBD level of theory to compute other physicochemical properties (as detailed in Table S5 of the SI). For these calculations, we have used the FHI-aims code¹⁰⁰ together with "tight" settings for basis functions and integration grids. Energies were converged to 10^{-6} eV and the accuracy of the forces was set to 10^{-4} eV/Å. The convergence criteria used during self-consistent field (SCF) optimizations were 10^{-3} eV for the sum of eigenvalues and 10^{-6} electrons/Å³ for the charge density. The MBD energies and MBD atomic forces were here computed using the range-separated self-consistent screening (rsSCS) approach¹⁹, while the atomic C_6 coefficients, isotropic atomic polarizabilities, molecular C_6 coefficients and molecular polarizabilities (both isotropic and tensor) were obtained *via* the SCS approach¹⁸. Here, we have also computed van der Waals forces using D4 and XDM methods. Hirshfeld ratios correspond to the Hirshfeld volumes divided by the free atom volumes. The TS dispersion energy refers to the pairwise Tkatchenko-Scheffler (TS) dispersion energy in conjunction with the PBE0 functional²⁵. The vdW radii were also obtained using the SCS approach *via* $R_{\text{vdW}} = (\alpha^{\text{SCS}}/\alpha^{\text{TS}})^{1/3} R_{\text{vdW}}^{\text{TS}}$, where α^{TS} and $R_{\text{vdW}}^{\text{TS}}$ are the atomic polarizability and vdW radius computed according to the TS scheme, respectively. Atomization energies were obtained by subtracting the atomic PBE0 energies from the PBE0 total energy of each molecular conformation.

LNO-CCSD(T) reference interaction energies. The large CCSD(T) computations well beyond the limits of conventional implementations were performed with the highly optimized local natural orbital (LNO) CCSD(T)⁶¹⁻⁶⁴ method in the MRCC^{109,110} (version 2024) quantum chemistry suite. First, a detailed basis set and LNO approximation convergence study is performed to determine the most efficient LNO and basis set settings that provide high accuracy for the interactions relevant to QUID. To that end, the interaction energies were tested for three representative dimers, namely SF2I2 (Fig. 7), F2B1, and L2B3 (Fig. S19 of the SI), at $1\times$, $0.9\times$, and $2.0\times$ equilibrium distances. Here, we used the systematically improving series of aug-cc-pV(X+d)Z basis sets with $X = \text{D, T and Q}$ ¹¹¹ as well as Normal (N), Tight (T), and very Tight (vT) LNO thresholds. The convergence toward the complete basis set (CBS) limit was accelerated for the HF¹¹² and correlation¹¹³ energies *via* CBS extrapolation [CBS(X,X+1)]. To accelerate the convergence toward conventional CCSD(T), that is the local approximation free (LAF) limit, the LAF extrapolations^{63,64} N-T employs the Normal and Tight and the T-vT use the Tight and very Tight LNO settings.

Our best estimate of CCSD(T)/CBS for the above nine dimers is given by a composite ($\bar{E}$) scheme benefiting from CBS(T,Q) and LAF T-vT extrapolations as well as counterpoise corrections¹¹⁴:

$$E_{\text{CCSD(T)}}^{\text{CBS}} \approx \bar{E}_{\text{T-vT LNO-CCSD(T)}}^{\text{CBS(T,Q)}} \approx E_{\text{T-vT LNO-CCSD(T)}}^{\text{CBS(D,T)}} - E_{\text{Tight LNO-CCSD(T)}}^{\text{CBS(D,T)}} + E_{\text{Tight LNO-CCSD(T)}}^{\text{CBS(T,Q)}} \quad (2)$$

Moreover, an uncertainty estimate can be assigned to the LNO approximation as the difference between the usually monotonically converging steps, that is, $\pm 0.5 |E_{\text{veryTight}}^{\text{XZ}} - E_{\text{Tight}}^{\text{XZ}}|$ for T-vT.

Inspecting all nine cases (Fig. 7 and Fig. S19), we find fast convergence with the LNO settings reaching ca. ± 0.1 kcal/mol uncertainties at the T-vT LAF extrapolated level. Importantly, the N-T LAF extrapolated results are very close to the T-vT extrapolated ones, thus the most costly very Tight settings can be spared when targeting the entire dataset. Regarding the basis set convergence, for all nine dimers, we find a good agreement between the counterpoise corrected and uncorrected results by reaching the CBS(T,Q) level. An additional indication of excellent convergence is the close match between counterpoise corrected aug-cc-pV(Q+d)Z as well as CBS(D,T) results, also suggesting that the most expensive aug-cc-pV(Q+d)Z computations are not needed for the entire dataset. Furthermore, the difference of the best-converged basis set levels, e.g., between counterpoise corrected CBS(D,T) and CBS(T,Q), can be used as an uncertainty estimate for the basis set convergence. Then, our conservative formula for the total uncertainty estimate w.r.t. CCSD(T)/CBS adds together the above LNO and basis set uncertainty estimates, which indicates about ± 0.2 kcal/mol uncertainty at the $\bar{E}_{\text{T-vT LNO-CCSD(T)}}^{\text{CBS(T,Q)}}$ level [Eq. (2)].

On the basis of the above analysis, a more affordable composite energy expression can be recommended:

$$\bar{E}_{\text{N-T LNO-CCSD(T)}}^{\text{CBS(D,T)}} \approx E_{\text{N-T LNO-CCSD(T)}}^{\text{DZ}} - E_{\text{Normal LNO-CCSD(T)}}^{\text{DZ}} + E_{\text{Normal LNO-CCSD(T)}}^{\text{CBS(D,T)}}, \quad (3)$$

which will be used to compute reference E_{int} values for QUID dimers. In accordance with the above analysis of the nine convergence plots, the more efficient $\bar{E}_{\text{N-T LNO-CCSD(T)}}^{\text{CBS(D,T)}}$ results are in excellent agreement with the tightly converged $\bar{E}_{\text{T-vT LNO-CCSD(T)}}^{\text{CBS(T,Q)}}$ ones, their difference is mostly below 0.1 kcal/mol, and only up to 0.2–0.3 kcal/mol for the worst cases of L2B3 (equilibrium) and SF2I2 (0.9× equilibrium). The corresponding uncertainty estimate for $\bar{E}_{\text{N-T LNO-CCSD(T)}}^{\text{CBS(D,T)}}$ is the sum of $0.5|E_{\text{Tight}}^{\text{DZ}} - E_{\text{Normal}}^{\text{DZ}}|$ and $0.33|E_{\text{Normal}}^{\text{CBS(D,T)}} - E_{\text{Normal}}^{\text{TZ}}|$, resulting on the average about ± 0.4 kcal/mol, in agreement with the observations for the nine dimers where $\bar{E}_{\text{T-vT LNO-CCSD(T)}}^{\text{CBS(T,Q)}}$ is available.

The benefit of $\bar{E}_{\text{N-T LNO-CCSD(T)}}^{\text{CBS(D,T)}}$ over $\bar{E}_{\text{T-vT LNO-CCSD(T)}}^{\text{CBS(T,Q)}}$ is about an order of magnitude computer time and 4-fold memory demand reduction. Namely, the three kinds of LNO-CCSD(T) interaction energy computations needed for Eq. (3) can be completed in ca. 30–70 hours for each dimer on 16 cores and at most 20 GB memory, making it ideal for high-throughput computations even on clusters with short wall-time limits. For the sake of completeness, the frozen-core approximation and conventional auxiliary basis sets^{115,116} were employed for all LNO-CCSD(T) interaction energies.

FN-DMC Interaction energy calculation. We performed Fixed-Node Diffusion Monte Carlo (FN-DMC) calculations^{27–29} for QUID dimers corroborating the LNO-CCSD(T) E_{int} with another accurate QM method. Our FN-DMC wave function ansatz was built using a single Slater determinant in addition to a Jastrow factor with one-body, two-body, and 3/4-body terms to account for cusps conditions at the nuclei, fermionic pair correlations, and product of pair correlations at the field of the nuclei, respectively. See references^{117,118} for further details about the electronic wave function ansatz. The molecular orbitals of the Slater determinant were taken from a previous DFT calculation employing ORCA code¹¹⁹ (version 5.0.4) with the Local-density approximation (LDA) exchange-correlation functional, a cc-pVTZ basis set for all atoms, and the ccECP pseudopotential¹²⁰.

All variational parameters of the Jastrow factor were variationally optimized at the Variational Monte Carlo (VMC) level with the stochastic reconfiguration optimization method¹²¹, while molecular orbital coefficients, basis set contraction coefficients, and exponents were kept fixed from the initial DFT calculation. The optimized VMC wave functions were taken as guiding functions in the FN-DMC calculation, in which we also employed the ccECP pseudopotentials to approximate core electrons for each atom, integrated with the determinant localization approximation (DLA)¹²². Consequently, the FN-DMC calculations were computed at two time-steps of 0.050 and 0.025 (a.u.), using 12800 walkers divided into 300 blocks, each 100 steps long, for a total of 4×10^8 sampled configurations. FN-DMC statistical error bars represent one- σ standard error of the mean estimated using binning technique (reblocking) to avoid autocorrelation. For some systems, it was required to run an additional third calculation with a time-step of 0.015 a.u. to get statistical agreement within 1- σ in the observed binding energies. In Fig. 7 we displayed the time-step convergence against the CCSD(T) reference values. Both VMC wave function optimization and FN-DMC calculations were performed with the QMeCha code^{117,118,123} (version Dec22, 2024, commit b296fc0).

Data availability

The QUID dataset generated in this study is openly available in the GitHub repository [Repo-QUID](#) with corresponding Zenodo DOI: [10.5281/zenodo.15880531](#).

Code availability

The MRCC^{109,110} code and source are open-access for academic use. The QMC QMeCha package is available in a private Github repository¹²³, and can be requested for academic use under a CC BY-NC-ND license by contacting Dr. Matteo Barborini at matteo.barborini@uni.lu. Source data are provided with this paper.

Acknowledgements

MP acknowledges with gratitude financial support from the Institute for Advanced Studies (IAS) Luxembourg for the PhD “Young Academics” program. AT was funded by the Luxembourg Research Fund (FNR Core Grant MBD-in-BMD/18093472). The authors thank Matteo Barborini and Matteo Gori for their support and discussions, Sergio Suárez Dou for his support in biomolecular force fields, and Gregory Cordeiro Fonseca for the support in the use of his FFAST software. The financial support from the ERC Starting Grant No. 101076972, “aCCuracy”, the ERC Advanced Grant No. 101054629 “FITMOL”, the National Research, Development, and Innovation Office (NKFIH, Grant No. FK142489), the János Bolyai Research Scholarship of the Hungarian Academy of Sciences, and the computing time granted by the Hungarian Governmental Information-Technology Development Agency on the Komondor and the LEONARDO supercomputers are gratefully acknowledged. Some of the calculations presented in this paper were carried out using the HPC facilities of the University of Luxembourg¹²⁴ (see hpc.uni.lu). The QMC simulations were performed on the Luxembourg national supercomputer MeluXina. The authors

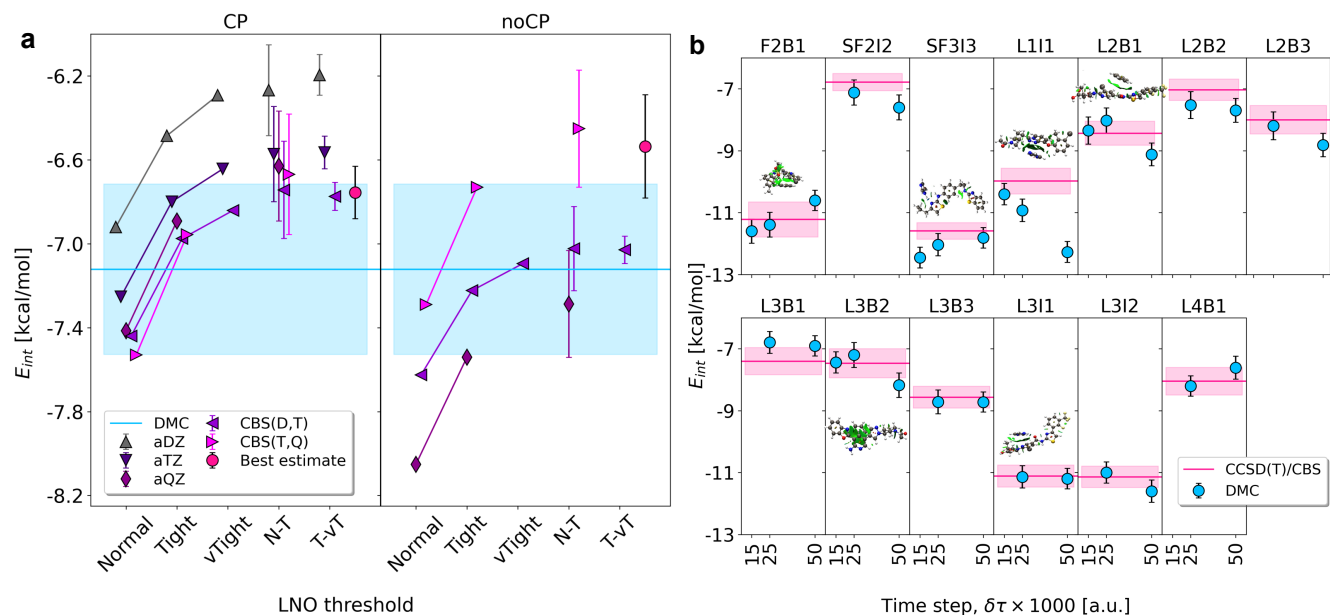


Figure 7. Mutual agreement for interaction energies from benchmark *ab initio* methods LNO-CCSD(T) and FN-DMC. **a** Local Natural Orbitals - Coupled Cluster with Singles, Doubles, and perturbative triplets (LNO-CCSD(T)) interaction energy E_{int} convergence analysis with respect to the LNO thresholds (x axis) and aug-cc-pV($X+d$)Z (aXZ) basis set choices with (CP) and without (noCP) CounterPoise corrections for the SF2I2 dimer at equilibrium distance, including the best estimate interaction energy corresponding to Eq. (2); the horizontal line indicates the Fixed-Node Diffusion Monte Carlo (FN-DMC) interaction energy ($0.025 \delta\tau$) with its statistical error in a box. For FN-DMC, error bars represent estimated one- σ statistical error for 4×10^8 configurations. For LNO-CCSD(T), the error bars correspond to the estimated uncertainty in the different extrapolations to the Complete Basis Set (CBS) and Local Approximation-Free (LAF) limits. **b** FN-DMC time-step convergence plots for a selection of 13 dimers. The reference LNO-CCSD(T) values are added for comparison, represented as horizontal lines with their uncertainty estimates shown as boxes.

gratefully acknowledge the LuxProvide teams for their expert support. An award of computer time was provided by the Innovative and Novel Computational Impact on Theory and Experiment (INCITE) program. This research used resources of the Argonne Leadership Computing Facility, which is a DOE Office of Science User Facility supported under Contract DE-AC02-06CH11357.

Author contributions

M.P. generated the QUID dimers, performed DFT optimization and calculations and semiempirical, empirical, and SAPT calculations, organized and analyzed the results, and drafted the original manuscript. L.M.S. contributed to the conceptualization, manuscript writing, conducted DFT optimizations and property calculations, participated in the analysis, and served in a supervisory role. B.L. carried out the LNO-CCSD(T) calculations. J.C. performed the FN-DMC calculations and supported the SAPT calculations. D.M.R. contributed to the study design, and facilitated the procurement of computational resources through the INCITE project. P.R.N. conducted the LNO-CCSD(T) calculations, procured HPC resources, and served in a supervisory role. A.T. contributed to supervision, conceptualization, analysis, manuscript writing, and the procurement of computational resources for HPC MeluXina. All authors discussed the results and contributed to the review and editing of the article.

Competing interests

The authors declare no conflict of interest.

References

1. Lingè, D. *et al.* PLBD: protein–ligand binding database of thermodynamic and kinetic intrinsic parameters. *Database* **2023**, baad040 (2023).
2. Ryde, U. & Söderhjelm, P. Ligand-binding affinity estimates supported by quantum-mechanical methods. *Chem. Rev.* **116**, 5520–5566 (2016).
3. Mucs, D. & Bryce, R. A. The application of quantum mechanics in structure-based drug design. *Expert. Opin. on Drug Discov.* **8**, 263–276 (2013).
4. Sohraby, F. & Nunes-Alves, A. Advances in computational methods for ligand binding kinetics. *Trends Biochem. Sci.* **48**, 437–449 (2023).
5. Jarmoskaite, I., AlSadhan, I., Vaidyanathan, P. P. & Herschlag, D. How to measure and evaluate binding affinities. *Elife* **9**, e57264 (2020).
6. Ryde, U. & Soderhjelm, P. Ligand-binding affinity estimates supported by quantum-mechanical methods. *Chem. Rev.* **116**, 5520–5566 (2016).
7. Davis, H. J. & Phipps, R. J. Harnessing non-covalent interactions to exert control over regioselectivity and site-selectivity in catalytic reactions. *Chem. Sci.* **8**, 864–877 (2017).
8. Ross, G. A. *et al.* The maximal and current accuracy of rigorous protein-ligand binding free energy calculations. *Commun. Chem.* **6**, 222 (2023).
9. Abel, R. A critical review of validation, blind testing, and real-world use of alchemical protein-ligand binding free energy calculations. *Curr. Top. Medicinal Chem.* **17**, 2577–2585 (2017).
10. Chodera, J. D. *et al.* Alchemical free energy methods for drug discovery: progress and challenges. *Curr. Opin. Struct. Biol.* **21**, 150–160 (2011).
11. Case, D. *et al.* AMBER 2016 (2016).
12. Vanommeslaeghe, K. *et al.* CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J. Comput. Chem.* **31**, 671–690 (2010).
13. Dauber-Osguthorpe, P. & Hagler, A. T. Biomolecular force fields: where have we been, where are we now, where do we need to go and how do we get there? *J. Comput. Mol. Des.* **33**, 133–203 (2019).
14. Stöhr, M. & Tkatchenko, A. Quantum mechanics of proteins in explicit water: The role of plasmon-like solute-solvent interactions. *Sci. Adv.* **5**, eaax0024 (2019).
15. Gaus, M., Cui, Q. & Elstner, M. DFTB3: Extension of the self-consistent-charge density-functional tight-binding method (SCC-DFTB). *J. Chem. Theory Comput.* **7**, 931–948 (2011).

16. Bannwarth, C., Ehlert, S. & Grimme, S. GFN2-xTB—An accurate and broadly parametrized self-consistent tight-binding quantum chemical method with multipole electrostatics and density-dependent dispersion contributions. *J. Chem. Theory Comput.* **15**, 1652–1671 (2019).
17. Adamo, C. & Barone, V. Toward reliable density functional methods without adjustable parameters: The PBE0 model. *J. Chem. Phys.* **110**, 6158–6170 (1999).
18. Tkatchenko, A., DiStasio Jr, R. A., Car, R. & Scheffler, M. Accurate and efficient method for many-body van der Waals interactions. *Phys. Rev. Lett.* **108**, 236402 (2012).
19. Ambrosetti, A., Reilly, A. M., DiStasio, R. A. & Tkatchenko, A. Long-range correlation energy calculated from coupled atomic response functions. *J. Chem. Phys.* **140**, 18A508 (2014).
20. Hermann, J. & Tkatchenko, A. Density functional model for van der Waals interactions: Unifying many-body atomic approaches with nonlocal functionals. *Phys. Rev. Lett.* **124**, 146401 (2020).
21. Lin, Y.-S., Li, G.-D., Mao, S.-P. & Chai, J.-D. Long-range corrected hybrid density functionals with improved dispersion corrections. *J. Chem. Theory Comput.* **9**, 263–272 (2013).
22. Zhao, Y. & Truhlar, D. G. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals. *Theor. Chem. Accounts* **120**, 215–241 (2008).
23. Caldeweyher, E., Bannwarth, C. & Grimme, S. Extension of the D3 dispersion coefficient model. *J. Chem. Phys.* **147**, 034112 (2017).
24. Caldeweyher, E. *et al.* A generally applicable atomic-charge dependent London dispersion correction. *J. Chem. Phys.* **150**, 154122 (2019).
25. Tkatchenko, A. & Scheffler, M. Accurate molecular van der Waals interactions from ground-state electron density and free-atom reference data. *Phys. Rev. Lett.* **102**, 073005 (2009).
26. Bartlett, R. J. & Musiał, M. Coupled-cluster theory in quantum chemistry. *Rev. Mod. Phys.* **79**, 291 (2007).
27. Foulkes, W. M. C., Mitas, L., Rajagopal, R. J. N. & Rajagopal, G. Quantum Monte Carlo simulations of solids. *Rev. Mod. Phys.* **73**, 33–83 (2001).
28. Dubecký, M., Mitas, L. & Jurečka, P. Noncovalent Interactions by Quantum Monte Carlo. *Chem. Rev.* **116**, 5188–5215 (2016).
29. Becca, F. & Sorella, S. *Quantum Monte Carlo approaches for correlated systems* (Cambridge University Press, 2017).
30. Sandonas, L. R. M. *et al.* Dataset for quantum-mechanical exploration of conformers and solvent effects in large drug-like molecules. *Sci. Data* **11**, 742 (2024).
31. Hoja, J. *et al.* QM7-X, a comprehensive dataset of quantum-mechanical properties spanning the chemical space of small organic molecules. *Sci. data* **8**, 43 (2021).
32. Smith, J. S., Isayev, O. & Roitberg, A. E. ANI-1, A data set of 20 million calculated off-equilibrium conformations for organic molecules. *Sci. data* **4**, 170193 (2017).
33. Isert, C., Atz, K., Jiménez-Luna, J. & Schneider, G. QMugs, quantum mechanical properties of drug-like molecules. *Sci. Data* **9**, 273 (2022).
34. Chmiela, S. *et al.* Accurate global machine learning force fields for molecules with hundreds of atoms. *Sci. Adv.* **9**, eadf0873 (2023).
35. Jurečka, P., Šponer, J., Černý, J. & Hobza, P. Benchmark database of accurate (MP2 and CCSD (T) complete basis set limit) interaction energies of small model complexes, DNA base pairs, and amino acid pairs. *Phys. Chem. Chem. Phys.* **8**, 1985–1993 (2006).
36. Hobza, P. Calculations on noncovalent interactions and databases of benchmark interaction energies. *Accounts Chem. Res.* **45**, 663–672 (2012).
37. Rezáč, J., Riley, K. E. & Hobza, P. S66: A well-balanced database of benchmark interaction energies relevant to biomolecular structures. *J. Chem. Theory Comput.* **7**, 2427–2438 (2011).
38. Rezáč, J., Riley, K. E. & Hobza, P. Extensions of the S66 data set: More accurate interaction energies and angular-displaced nonequilibrium geometries. *J. Chem. Theory Comput.* **7**, 3466–3470 (2011).

39. Sedlak, R. *et al.* Accuracy of quantum chemical methods for large noncovalent complexes. *J. Chem. Theory Comput.* **9**, 3364–3374 (2013).
40. Sparrow, Z. M., Ernst, B. G., Joo, P. T., Lao, K. U. & DiStasio, R. A. NENCI-2021. I. A large benchmark database of non-equilibrium non-covalent interactions emphasizing close intermolecular contacts. *J. Chem. Phys.* **155**, 184303 (2021).
41. Donchev, A. G. *et al.* Quantum chemical benchmark databases of gold-standard dimer interaction energies. *Sci. Data* **8**, 55 (2021).
42. Villot, C. & Lao, K. U. Ab initio dispersion potentials based on physics-based functional forms with machine learning. *J. Chem. Phys.* **160**, 184103 (2024).
43. Spronk, S. A., Glick, Z. L., Metcalf, D. P., Sherrill, C. D. & Cheney, D. L. A quantum chemical interaction energy dataset for accurately modeling protein-ligand interactions. *Sci. Data* **10**, 619 (2023).
44. Eastman, P. *et al.* Spice, a dataset of drug-like molecules and peptides for training machine learning potentials. *Sci. Data* **10**, 11 (2023).
45. Wang, R., Fang, X., Lu, Y. & Wang, S. The PDBbind database: Collection of binding affinities for protein-ligand complexes with known three-dimensional structures. *J. Med. Chem.* **47**, 2977–2980 (2004).
46. Unke, O. T. *et al.* Biomolecular dynamics with machine-learned quantum-mechanical force fields trained on diverse chemical fragments. *Sci. Adv.* **10**, eadn4397 (2024).
47. Kabylda, A. *et al.* Molecular simulations with a pretrained neural network and universal pairwise force fields. *ChemRxiv* 10.26434 (2024).
48. Ackloo, S. *et al.* A target class ligandability evaluation of WD40 repeat-containing proteins. *J. Medicinal Chem.* **68**, 1092–1112 (2024).
49. Meng, E. C. *et al.* UCSF ChimeraX: Tools for structure building and analysis. *Protein Sci.* **32**, e4792 (2023).
50. Siwach, A. & Verma, P. K. Synthesis and therapeutic potential of imidazole containing compounds. *BMC Chem.* **15**, 12 (2021).
51. Ferreira de Freitas, R. & Schapira, M. A systematic analysis of atomic protein-ligand interactions in the PDB. *Med. Chem. Commun.* **8**, 1970 (2017).
52. Gao, M. & Skolnick, J. A comprehensive survey of small-molecule binding pockets in proteins. *PLoS computational biology* **9**, e1003302 (2013).
53. Johnson, D. K. & Karanicolas, J. Druggable protein interaction sites are more predisposed to surface pocket formation than the rest of the protein surface. *PLoS computational biology* **9**, e1002951 (2013).
54. Parker, T. M., Burns, L. A., Parrish, R. M., Ryno, A. G. & Sherrill, C. D. Levels of symmetry adapted perturbation theory (SAPT). I. Efficiency and performance for interaction energies. *J. Chem. Phys.* **140**, 094106 (2014).
55. Kumar, K. *et al.* Cation- π interactions in protein-ligand binding: theory and data-mining reveal different roles for lysine and arginine. *Chem. Sci.* **9**, 2655–2665 (2018).
56. Contreras-García, J. *et al.* NCIPLOT: A program for plotting noncovalent interaction regions. *J. Chem. Theory Comput.* **7**, 625–632 (2011).
57. Sinnokrot, M. O. & Sherrill, C. D. Highly accurate coupled cluster potential energy curves for the benzene dimer: Sandwich, T-shaped, and parallel-displaced configurations. *J. Phys. Chem. A* **108**, 10200–10207 (2004).
58. Al-Hamdani, Y. S. *et al.* Interactions between large molecules pose a puzzle for reference quantum mechanical methods. *Nat. Commun.* **12**, 3927 (2021).
59. Shi, B. X. *et al.* Systematic discrepancies between reference methods for noncovalent interactions within the S66 dataset. *The J. Chem. Phys.* **162** (2025).
60. Fishman, V., Lesiuk, M., Martin, J. M. L. & Boese, A. D. A new angle on benchmarking noncovalent interactions. *arXiv* 2410.12603 (2024).
61. Nagy, P. R. & Kállay, M. Optimization of the linear-scaling local natural orbital CCSD(T) method: Redundancy-free triples correction using Laplace transform. *J. Chem. Phys.* **146**, 214106 (2017).
62. Nagy, P. R., Samu, G. & Kállay, M. Optimization of the linear-scaling local natural orbital CCSD(T) method: Improved algorithm and benchmark applications. *J. Chem. Theory Comput.* **14**, 4193 (2018).

63. Nagy, P. R. & Kállay, M. Approaching the basis set limit of CCSD(T) energies for large molecules with local natural orbital coupled-cluster methods. *J. Chem. Theory Comput.* **15**, 5275 (2019).
64. Nagy, P. R. State-of-the-art local correlation methods enable affordable gold standard quantum chemistry for up to hundreds of atoms. *Chem. Sci.* **15**, 14556–14584 (2024).
65. Al-Hamdani, Y. S. & Tkatchenko, A. Understanding non-covalent interactions in larger molecular complexes from first principles. *J. Chem. Phys.* **150**, 010901 (2019).
66. Schäfer, T., Irmeler, A., Gallo, A. & Grüneis, A. Understanding discrepancies of wavefunction theories for large molecules. *arXiv e-prints* arXiv:2407.01442 (2024).
67. Fishman, V., Lesiuk, M., Martin, J. M. L. & Daniel Boese, A. Another angle on benchmarking noncovalent interactions. *J. Chem. Theory Comput.* **21**, 2311 (2025).
68. Semidalas, E., Boese, A. D. & Martin, J. M. Post-CCSD(T) corrections in the S66 noncovalent interactions benchmark. *Chem. Phys. Lett.* **863**, 141874 (2025).
69. Mardirossian, N. & Head-Gordon, M. ω B97X-V: A 10-parameter, range-separated hybrid, generalized gradient approximation density functional with nonlocal correlation, designed by a survival-of-the-fittest strategy. *Phys. Chem. Chem. Phys.* **16**, 9904–9924 (2014).
70. Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77**, 3865 (1996).
71. Grimme, S., Antony, J., Ehrlich, S. & Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (dft-d) for the 94 elements h-pu. *The J. chemical physics* **132** (2010).
72. Grimme, S., Ehrlich, S. & Goerigk, L. Effect of the damping function in dispersion corrected density functional theory. *J. computational chemistry* **32**, 1456–1465 (2011).
73. Becke, A. D. & Johnson, E. R. A density-functional model of the dispersion interaction. *J. Chem. Phys.* **123**, 154101 (2005).
74. Brémond, É., Sancho-García, J. C., Pérez-Jiménez, Á. J. & Adamo, C. Communication: Double-hybrid functionals from adiabatic-connection: The QIDH model. *The J. Chem. Phys.* **141** (2014).
75. Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *The J. Chem. Phys.* **98**, 5648–5652 (1993).
76. Lee, C., Yang, W. & Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* **37**, 785 (1988).
77. Bannwarth, C., Ehlert, S. & Grimme, S. GFN2-xTB—An accurate and broadly parametrized self-consistent tight-binding quantum chemical method with multipole electrostatics and density-dependent dispersion contributions. *J. Chem. Theory Comput.* **15**, 1652–1671 (2019).
78. Alfonso-Ramos, J. E., Adamo, C., Brémond, É. & Stuyver, T. Improving the reliability of, and confidence in, dft functional benchmarking through active learning. *J. Chem. Theory Comput.* **21**, 1752–1761 (2025).
79. Gray, M. & Herbert, J. M. Density functional theory for van der waals complexes: Size matters. In *Annual Reports in Computational Chemistry*, vol. 20, 1–61 (Elsevier, 2024).
80. Eastman, P. *et al.* OpenMM 8: molecular dynamics simulation with machine learning potentials. *The J. Phys. Chem. B* **128**, 109–116 (2023).
81. Mardirossian, N. & Head-Gordon, M. Thirty years of density functional theory in computational chemistry: an overview and extensive assessment of 200 density functionals. *Mol. Phys.* **115**, 2315–2372 (2017).
82. Sedlak, R. *et al.* Accuracy of quantum chemical methods for large noncovalent complexes. *J. Chem. Theory Comput.* **9**, 3364–3374 (2013).
83. Sancho-Garcia, J.-C., Brémond, É., Savarese, M., Pérez-Jiménez, A. & Adamo, C. Partnering dispersion corrections with modern parameter-free double-hybrid density functionals. *Phys. Chem. Chem. Phys.* **19**, 13481–13487 (2017).
84. Tang, K. & Toennies, J. P. An improved simple model for the van der Waals potential based on universal damping functions for the dispersion coefficients. *J. Chem. Phys.* **80**, 3726–3741 (1984).
85. Chai, J.-D. & Head-Gordon, M. Long-range corrected hybrid density functionals with damped atom–atom dispersion corrections. *Phys. Chem. Chem. Phys.* **10**, 6615–6620 (2008).

86. Medrano Sandonas, L. *et al.* “Freedom of design” in chemical compound space: towards rational in silico design of molecules with targeted quantum-mechanical properties. *Chem. Sci.* **14**, 10702–10717 (2023).
87. Hait, D. & Head-Gordon, M. When is a bond broken? The polarizability perspective. *Angewandte Chemie* **62**, e202312078 (2023).
88. Fonseca, G., Poltavsky, I. & Tkatchenko, A. Force field analysis software and tools (FFAST): Assessing machine learning force fields under the microscope. *J. Chem. Theory Comput.* **19**, 8706–8717 (2023).
89. Mester, D. *et al.* Overview of developments in the MRCC program system. *The J. Phys. Chem. A* **129**, 2086–2107 (2025).
90. Sliotman, E. *et al.* Accurate quantum Monte Carlo forces for machine-learned force fields: Ethanol as a benchmark. *J. Chem. Theory Comput.* **20**, 6020–6027 (2024).
91. Hilfiker, M., Medrano Sandonas, L., Klähn, M., Engkvist, O. & Tkatchenko, A. Leveraging quantum mechanical properties to predict solvent effects on large drug-like molecules. In Clevert, D.-A., Wand, M., Malinová, K., Schmidhuber, J. & Tetko, I. V. (eds.) *AI in Drug Discovery*, 47–57 (Springer Nature Switzerland, Cham, 2025).
92. Pecina, A., Fanfrlík, J., Lepšík, M. & Řezáč, J. SQM2.20: Semiempirical quantum-mechanical scoring function yields DFT-quality protein–ligand binding affinity predictions in minutes. *Nat. Commun.* **15**, 1127 (2024).
93. Yan, Z., Wei, D., Li, X. & Chung, L. W. Accelerating reliable multiscale quantum refinement of protein–drug systems enabled by machine learning. *Nat. Commun.* **15**, 4181 (2024).
94. Blum, V. *et al.* Ab initio molecular simulations with numeric atom-centered orbitals. *Comput. Phys. Comm.* **180**, 2175–2196 (2009).
95. Hanwell, M. D. *et al.* Avogadro: an advanced semantic chemical editor, visualization, and analysis platform. *J. Cheminformatics* **4**, 17 (2012).
96. Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38 (1996).
97. Mardirossian, N. & Head-Gordon, M. ω B97M-V: A combinatorially optimized, range-separated hybrid, meta-GGA density functional with VV10 nonlocal correlation. *J. Chem. Phys.* **144**, 214110 (2016).
98. Yanai, T., Tew, D. P. & Handy, N. C. A new hybrid exchange–correlation functional using the coulomb-attenuating method (cam-b3lyp). *Chem. Phys. Lett.* **393**, 51–57 (2004).
99. Becke, A. D. A new mixing of hartree-fock and local density-functional theories. *J. Chem. Phys.* **98**, 1372–1377 (1993).
100. Blum, V., Rossi, M., Kokott, S. & Scheffler, M. The FHI-aims code: All-electron, ab initio materials simulations towards the exascale. *arXiv* 2208.12335 (2022).
101. Turney, J. M. *et al.* Psi4: An open-source ab initio electronic structure program. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2**, 556–565 (2012).
102. Lehtola, S., Steigemann, C., Oliveira, M. J. & Marques, M. A. Recent developments in libxc—A comprehensive library of functionals for density functional theory. *SoftwareX* **7**, 1–5 (2018).
103. Shao, Y. *et al.* Advances in molecular quantum chemistry contained in the Q-Chem 4 program package. *Mol. Phys.* **113**, 184–215 (2015).
104. Otero-De-La-Roza, A. & Johnson, E. R. Non-covalent interactions and thermochemistry using XDM-corrected hybrid and range-separated hybrid density functionals. *J. Chem. Phys.* **138**, 204109 (2013).
105. Nickerson, C. J., Bryenton, K. R., Price, A. J. & Johnson, E. R. Comparison of density-functional theory dispersion corrections for the DES15K database. *J. Phys. Chem. A* **127**, 8712–8722 (2023).
106. Hourahine, B. *et al.* DFTB+, a software package for efficient approximate density functional theory based atomistic simulations. *J. Chem. Phys.* **152**, 124101 (2020).
107. O’Boyle, N. M. *et al.* Open Babel: An open chemical toolbox. *J. Cheminformatics* **3**, 1–14 (2011).
108. Vanommeslaeghe, K., Raman, E. P. & MacKerell Jr, A. D. Automation of the CHARMM general force field (CGenFF) II: assignment of bonded parameters and partial atomic charges. *J. Chem. Inf. Model.* **52**, 3155–3168 (2012).
109. Kállay, M. *et al.* The MRCC program system: Accurate quantum chemistry from water to proteins. *J. Comput. Phys.* **152**, 074107 (2020).
110. Kállay, M. *et al.* MRCC, a quantum chemical program suite. see <https://www.mrcc.hu/> (Accessed May 1, 2024).
111. Dunning Jr., T. H., Peterson, K. A. & Wilson, A. K. Gaussian basis sets for use in correlated molecular calculations. X. The atoms aluminum through argon revisited. *J. Chem. Phys.* **114**, 9244 (2001).

- 787 **112.** Karton, A. & Martin, J. M. L. Comment on: “Estimating the Hartree–Fock limit from finite basis set calculations”. *Theor.*
788 *Chim. Acta* **115**, 330 (2006).
- 789 **113.** Helgaker, T., Klopper, W., Koch, H. & Noga, J. Basis-set convergence of correlated calculations on water. *J. Chem. Phys.*
790 **106**, 9639 (1997).
- 791 **114.** Boys, S. F. & Bernardi, F. The calculation of small molecular interactions by the differences of separate total energies.
792 Some procedures with reduced errors. *Mol. Phys.* **19**, 553 (1970).
- 793 **115.** Weigend, F. Hartree–Fock exchange fitting basis sets for H to Rn. *J. Comput. Chem.* **29**, 167 (2008).
- 794 **116.** Weigend, F., Köhn, A. & Hättig, C. Efficient use of the correlation consistent basis sets in resolution of the identity MP2
795 calculations. *J. Chem. Phys.* **116**, 3175 (2002).
- 796 **117.** Charry Martinez, J. A., Barborini, M. & Tkatchenko, A. Correlated wave functions for electron–positron interactions in
797 atoms and molecules. *J. Chem. Theory Comput.* **18**, 2267–2280 (2022).
- 798 **118.** Ditte, M., Barborini, M., Medrano Sandonas, L. & Tkatchenko, A. Molecules in environments: Toward systematic
799 quantum embedding of electrons and Drude oscillators. *Phys. Rev. Lett.* **131**, 228001 (2023).
- 800 **119.** Neese, F., Wennmohs, F., Becker, U. & Riplinger, C. The ORCA quantum chemistry program package. *J. Chem. Phys.*
801 **152**, 224108 (2020).
- 802 **120.** Wang, G. *et al.* A new generation of effective core potentials from correlated calculations: 4s and 4p main group elements
803 and first row additions. *J. Chem. Phys.* **151**, 144110 (2019).
- 804 **121.** Sorella, S. Wave function optimization in the variational Monte Carlo method. *Phys. Rev. B - Condens. Matter Mater.*
805 *Phys.* **71**, 241103 (2005).
- 806 **122.** Zen, A., Brandenburg, J. G., Michaelides, A. & Alfè, D. A new scheme for fixed node diffusion quantum Monte Carlo
807 with pseudopotentials: Improving reproducibility and reducing the trial-wave-function bias. *J. Chem. Phys.* **151**, 134105
808 (2019).
- 809 **123.** Barborini, M. Quantum Mecha (QMeCha) package <https://github.com/qmecha> (private repository) (2024).
- 810 **124.** Varrette, S., Bouvry, P., Cartiaux, H. & Georgatos, F. Management of an academic HPC cluster: The UL experience. In
811 *Proc. of the 2014 Intl. Conf. on High Performance Computing & Simulation (HPCS 2014)*, 959–967 (IEEE, Bologna,
812 Italy, 2014).