Intuitive Density Functional Theory-Based Energy Decomposition Analysis for Protein-Ligand Interactions

M. J. S. Phipps,[†] T. Fox,[‡] C. S. Tautermann,[‡] and C.-K. Skylaris[*,†]

School of Chemistry, University of Southampton, Highfield, Southampton, SO17 1BJ, UK, and

Department of Medicinal Chemistry, Boehringer Ingelheim Pharma GmbH & Co. KG, 88397

Biberach, Germany

E-mail: C.Skylaris@soton.ac.uk

Abstract

First principles quantum mechanical calculations with methods such as density functional theory (DFT) allow the accurate calculation of interaction energies between molecules. These interaction energies can be dissected into chemically relevant components such as electrostatics, polarisation and charge transfer using energy decomposition analysis (EDA) approaches. Typically EDA has been used to study interactions between small molecules, however it has great potential to be applied to large biomolecular assemblies such as protein-protein and protein-ligand interactions. We present an application of EDA calculations to the study of ligands that bind to the thrombin protein, using the ONETEP program for linear-scaling DFT calculations. Our approach goes beyond simply providing the components of the interaction energy; we are also able to provide visual representations of the changes in density that happen

---

[*]To whom correspondence should be addressed
[†]School of Chemistry, University of Southampton, Highfield, Southampton, SO17 1BJ, UK
[‡]Department of Medicinal Chemistry, Boehringer Ingelheim Pharma GmbH & Co. KG, 88397 Biberach, Germany

as a result of polarisation and charge transfer, thus pinpointing the functional groups between the ligand and protein that participate in each kind of interaction. We also demonstrate with this approach that we can focus into studying parts (fragments) of ligands. The method is relatively insensitive to the protocol that is used to prepare the structures, and the results obtained are therefore robust. This is an application to a real protein drug target of a whole new capability where accurate DFT calculations can produce both energetic and visual descriptors of interactions. These descriptors can be used to provide insights for tailoring interactions, as needed for example in drug design.

1   Introduction

The ability to understand the driving forces of molecular interactions is a key goal towards solving many important problems of chemistry, for example in the development of new materials such as catalysts[1–3] and polymers,[4–8] and the fine tuning of drug-biomolecule interactions.[9–11] Determination of these driving forces is not trivial, with many different effects combining to produce an overall interaction energy. Energy decomposition analysis (EDA) refers to the number of approaches[12] that have been developed to aid our understanding of these forces. These methods separate the interaction energy (ΔE) into a series of chemically intuitive energy terms. Many such methods have been proposed in the literature,[13–24] typically partitioning ΔE into electrostatics, exchange, correlation, Pauli-repulsions, polarisation, charge transfer and dispersion terms, and combinations of these.

Many popular variational-based EDA methods have been developed from the pioneering Kitaura-Morokuma (KM) scheme.[13,25,26] An example of one such scheme is the absolutely localised molecular orbital (ALMO) EDA scheme,[20,27,28] closely related to the block-localised wavefunction (BLW) EDA,[29–31] that seeks to decompose the interaction energy into frozen density, polarisation and charge transfer energy terms. More recently, alternative treatments for polarisation have been proposed such as the EDA of Wu[22,32] and the approach of Horn and Head-Gordon.[33] We have published a review[12] comparing the different EDA approaches within the scope of drug

design which may be of interest to the inquisitive reader.

Despite the extensive progress in the development of EDA methods, there have been so far few attempts to use them in drug design applications. Examples of such work include the application of the pair interaction EDA[34] (PIEDA) by Heifetz et al. to study the binding of inhibitors to interleukin-2 inducible T-cell kinase,[35] the orexin-2 receptor[36] and other G-protein-coupled receptors,[37] and the application of the KM and natural[16,17] EDA approaches to study the binding of Cl- ions in the Cl-/H+ exchange transporter protein by Church et al.[38] Here, polarisation and charge transfer were found to play important roles in these binding interactions. Recently, Fedorov and Kitaura[39] demonstrated a subsystem analysis of binding in the 20-residue Trp-cage protein. In this work ligand binding in solution was analysed using their PIEDA scheme with the polarisable continuum model of solvation. EDA methods have the potential to provide valuable information with regard to the chemical nature of the interactions in order to guide pharmaceutical chemists to modify chemical functional groups in the ligands accordingly.

EDA has the exciting potential to guide the pharmaceutical modification of drug-protein interactions in systems many thousands of atoms in size. We have recently presented in literature[21] an EDA method developed in the ONETEP DFT package capable of such large-scale analyses. Towards this goal we have sought to study with EDA the thrombin protein which contains an active site consisting of three well-shaped pockets: S1, S2, and S3/4.[40] This a serine protease that plays a central role in the blood coagulation cascade, in which insoluble fibrin is formed by the action of thrombin on fibrinogen.[40,41] The delay of coagulation by direct inhibition of thrombin is therefore a key target for the treatment of thrombosis, for example by interaction of the drug molecules ximelagatran or dabigatran.[42]

In this paper, we investigate the interactions of thrombin in two cases. These two cases have separate objectives, but share an overall aim of exploring the usefulness of large-scale EDA with ONETEP for drug design. In the first case, we apply the ONETEP EDA to small, structurally comparable fragments interacting with the thrombin protein S1 pocket. Here, our aim is to demonstrate the practical use of our EDA for rationalising the energetics observed and identifying the key

functional groups involved in drug-ligand binding.

In the second case, we consider the interactions of larger thrombin drug binders partitioned into fragments. Our motivation is to assess the performance of the EDA for investigating the energetic contributions of specific functional motifs to binding and also investigate the level at which ligand fragmentation can have additive interactions.

2 Theory

We have implemented within the ONETEP[43] package a combination of the ALMO EDA approach[20] and the frozen density component analysis based upon the LMO EDA approach.[19] This development has been described in detail in earlier published literature.[21] In this section, we provide a brief overview of this EDA method and a trivial extension for solvation that is used within the studies of this work.

2.1 The ONETEP EDA

The original ONETEP EDA approach[21] decomposes the (gas phase) interaction energy ΔE_{vac} into electrostatic (ΔE_{ES}), exchange (ΔE_{EX}), Pauli-repulsion (ΔE_{REP}), correlation (ΔE_{CORR}), polarisation (ΔE_{POL}) and charge transfer (ΔE_{CT}) energy components as

$$\Delta E_{\text{vac}} = \Delta E_{\text{ES}} + \Delta E_{\text{EX}} + \Delta E_{\text{REP}} + \Delta E_{\text{CORR}} + \Delta E_{\text{POL}} + \Delta E_{\text{CT}} \quad . \tag{1}$$

The frozen density component (ΔE_{FRZ}) that represents the interaction of the frozen, unperturbed charge densities of the monomers is the total of the electrostatic, exchange, Pauli-repulsion, and correlation terms, i.e.

$$\Delta E_{\text{FRZ}} = \Delta E_{\text{ES}} + \Delta E_{\text{EX}} + \Delta E_{\text{REP}} + \Delta E_{\text{CORR}} \quad . \tag{2}$$

By substitution of Eq. 2 into Eq. 1, the decomposition may be expressed in compact form as

$$\Delta E_{\text{vac}} = \Delta E_{\text{FRZ}} + \Delta E_{\text{POL}} + \Delta E_{\text{CT}} \tag{3}$$

similar to that of the original ALMO EDA scheme.[20]

2.2 Solvation

The EDA scheme described above is trivially extended to account for solvation effects during binding. This component is included as a correction to ΔE_{vac}, and its inclusion does not affect the original energy component definitions described in Eq. 1.

The solvent interaction correction to ΔE_{vac}, ΔE_{SOLV}, is given by

$$\Delta E_{\text{SOLV}} = \Delta E_{\text{sol}} - \Delta E_{\text{vac}} \tag{4}$$

where E_{sol} is the total energy calculated using the ONETEP solvation model[44] which is a self-consistent model based on direct solution of the non-homogeneous Poisson equation. The calculation of the solvation contribution is shown diagrammatically in Figure 1.

The overall solvation corrected interaction energy decomposition is given by

$$\Delta E = \Delta E_{\text{ES}} + \Delta E_{\text{EX}} + \Delta E_{\text{REP}} + \Delta E_{\text{CORR}} + \Delta E_{\text{POL}} + \Delta E_{\text{CT}} + \Delta E_{\text{SOLV}} \quad . \tag{5}$$

We should note that the definition of the interaction energy in Eq. 5 describes a portion of the enthalpy (ΔH) component of the Gibbs' free energy of binding (ΔG). Some of the remaining enthalpic portion may be included as additional energy components to the EDA. For example, in ligand-protein binding the interaction energy may be augmented to include the dispersion energy of the ligand leaving the solvent using an empirical-based approach. Additionally, strain effects may be taken into account by relaxation of the molecules in isolation. A comparison of the binding enthalpies (calculated by inclusion of these effects) for the ligands studied in this work with
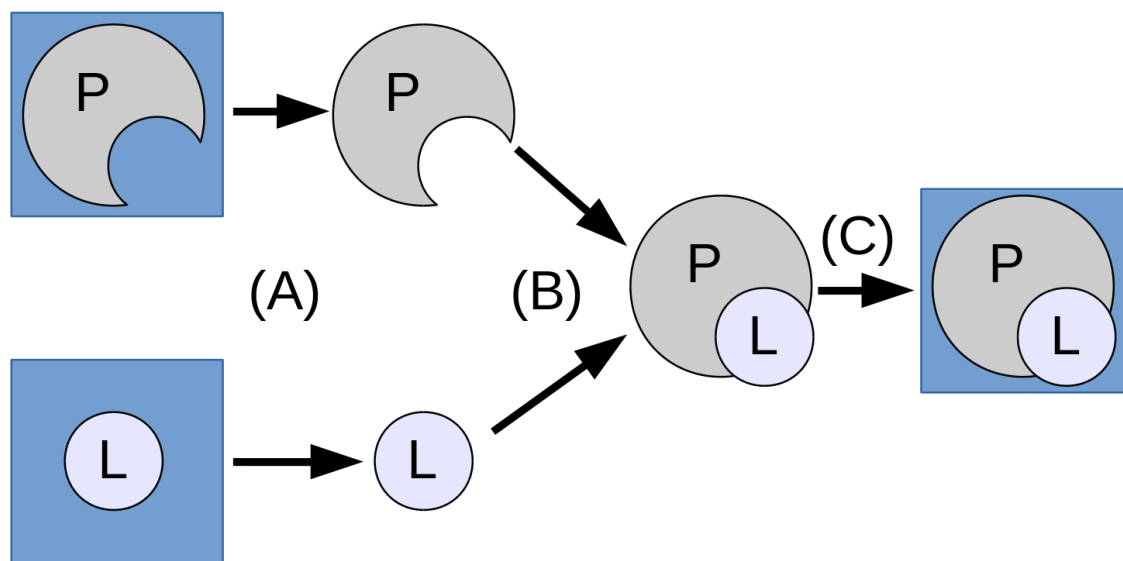
Figure 1: Calculation of the solvation contribution to ΔE for a model protein - ligand system. Energies of the protein (P) and ligand (L) calculated in solvent are represented using a blue box. The total (gas phase) interaction energy described in the original EDA (Eq. 1) is given by the energy change of the process (B). The correction for solvation, ΔE_{SOLV}, is given by the total energy change of the process $[(A) + (C)]$.

experimental thermodynamic parameters is provided in the Supporting Information (Supporting Information Available).

2.3 Electron density differences

Whilst the EDA provides energetic quantities useful for understanding the overall driving forces of binding, these do not provide any insight into the spatial regions of importance to binding. Notably, the energies provided by the EDA do not provide information of the key functional groups that contribute to binding through effects such as polarisation and charge transfer. This information would be highly valuable from a drug-design perspective, as this provides insight into regions of ligands relevant for structural modification.

As part of the ONETEP EDA approach, electron density difference (EDD) plots are computed that provide ready access to this valuable information. These plots that indicate the significant

electron density redistributions that occur during polarisation and charge transfer are calculated as

$$n_{\text{EDD(POL)}}(\mathbf{r}) = n_{\text{POL}}(\mathbf{r}) - n_{\text{FRZ}}(\mathbf{r}) \tag{6a}$$

$$n_{\text{EDD(CT)}}(\mathbf{r}) = n_0(\mathbf{r}) - n_{\text{POL}}(\mathbf{r}) \tag{6b}$$

where $n_{\text{FRZ}}(\mathbf{r})$, $n_{\text{POL}}(\mathbf{r})$, and $n_0(\mathbf{r})$ are the frozen,[45] polarised, and fully relaxed electron densities respectively. These can be understood by considering the densities used in these computations: the polarisation EDD plot shows the effect of polarising the frozen monomer densities, and the charge transfer EDD plot shows the effect of allowing full electron delocalisation having already polarised the individual monomer charge densities.

EDD plots of geometrically similar ligands may be subtracted[46] from one another to investigate the chemistry of their differing functional groups. In this work, these plots are referred to as ΔEDD plots and are calculated as

$$n_{\Delta\text{EDD(POL)}}(\mathbf{r}) = \text{abs}\left(n_{\text{EDD(POL,X1)}}(\mathbf{r})\right) - \text{abs}\left(n_{\text{EDD(POL,X2)}}(\mathbf{r})\right) \tag{7a}$$

$$n_{\Delta\text{EDD(CT)}}(\mathbf{r}) = \text{abs}\left(n_{\text{EDD(CT,X1)}}(\mathbf{r})\right) - \text{abs}\left(n_{\text{EDD(CT,X2)}}(\mathbf{r})\right) \tag{7b}$$

where X1 and X2 refer to the two ligands being analysed using ΔEDD plots, with positive iso-surface values representing greater interaction for X1 and negative isosurface values representing greater interaction for X2. Absolute values are used to allow identification of the ligand origins of the ΔEDD plot interactions especially in regions where the EDD values overlap.

3 Calculation details

The ability to observe differences in structural features of ligands is of key importance within this work. For this reason, thrombin fragments and ligands were taken from the literature of Rühmann et al.[40] and Baum et al.[47] based on their high suitability for structural comparison (i.e. the ligands are structurally different, but not so dissimilar as to prevent comparison). The PDB codes of the

structures selected for the S1 pocket analysis were 5AHG,[40] 5AFY,[40] 4UD9,[40] 5AF9,[40] 4UEH,[40] and 4UE7,[40] and the PDB codes of the thrombin-ligand complex structures selected for fragmentation studies were 4UDW,[40] 2ZGX,[47] and 5AFZ[40] (downloaded from the Protein Data Bank[48]). In this work, the S1 pocket binders are referred to as L1 to L6 respectively and the thrombin ligands are referred to as L7 to L9. A figure of these ligands superposed in the protein cavity is shown in Figure 2.

The following protocol was used to prepare the protein structures for EDA. The missing γ-loop (residues Trp147A, Thr147B, Ala147C, Asn147D, Val147E, Gly147F, and Lys147G) of each of the initial thrombin-inhibitor crystal structures was built using MOE homology modelling suite. Each structure was then capped using $COCH_3$ and $NHCH_3$ groups, and protonated using the Protonate3d software.[49] The structures were energy minimised using the MMFF94x forcefield with the generalised Born implicit solvent (GBIS) model using tether restraints with a force constant of $(3/2)kT/\sigma^2$ ($\sigma = 0.5$Å) for all atoms during the minimisation. Additional calculations have been performed demonstrating the stability of the ONETEP EDA with respect to geometry variations resulting when using different geometry optimisation procedures (Supporting Information Available). The hirudin and all waters (excluding the structural water molecules involved in the Asp189 salt bridge interaction in the L5, L6, L8, and L9 systems) were subsequently removed from the systems. These structures were then truncated to remove residues beyond 15Å of the ligand, and further truncated to ensure consensus of the protein residues across all of the systems. Further capping of the cleaved backbone of the structures was performed using $COCH_3$ and $NHCH_3$ groups to leave protein model structures 2740 atoms in size.

The ligands L7, L8, and L9 were further partitioned into fragments that interact with particular regions of the thrombin active site. Specifically, we have selected fragments with index F1 to interact with the S1 pocket, F2 with the S2 pocket, and F3 to interact with the S3/4 pocket. Positions of the ligand fragmentations are shown in Figure 3. To partition the ligands into fragments, the selected bonds were cleaved and capped with hydrogens.

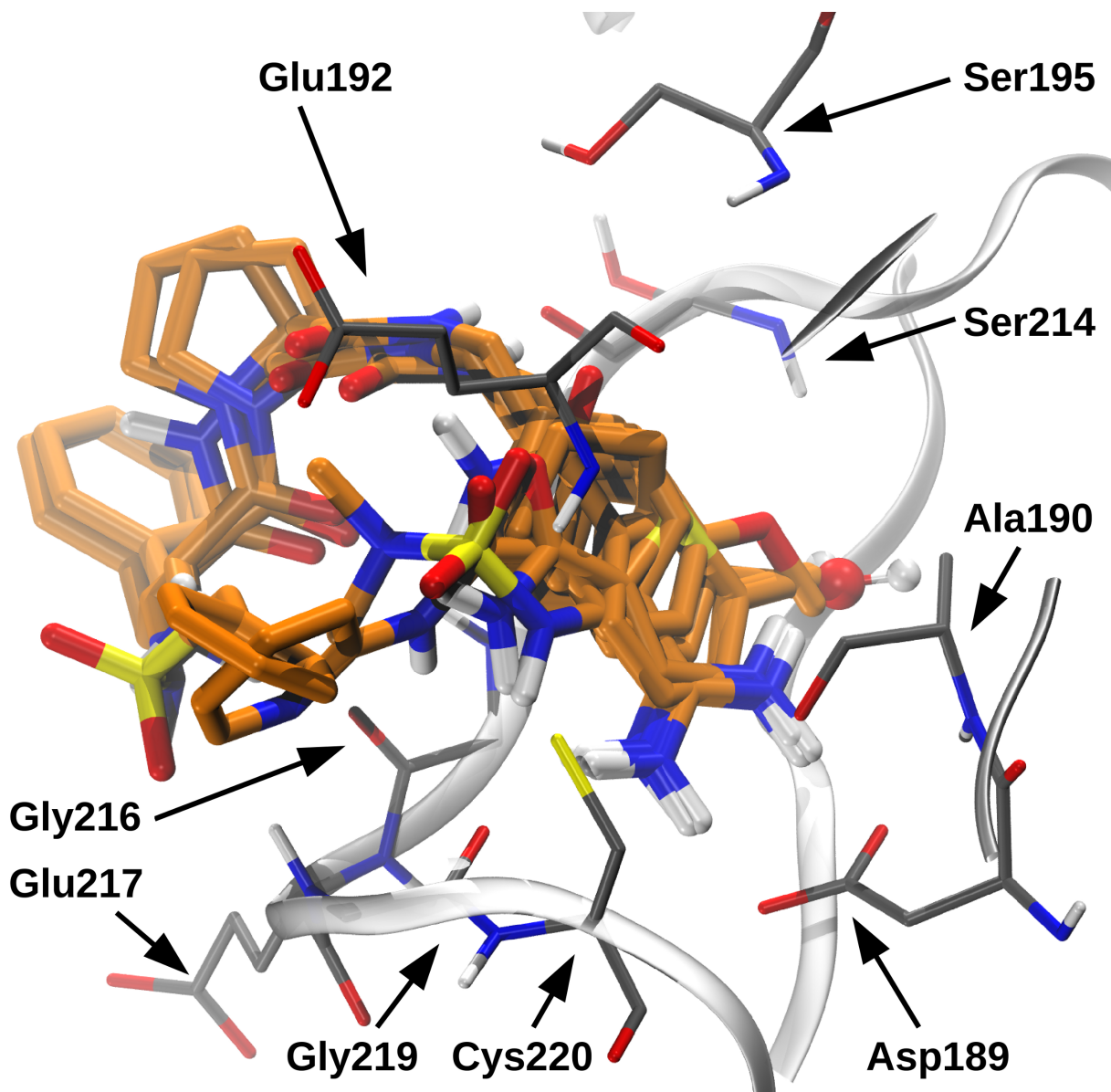ONETEP EDA was performed on the resulting thrombin-fragment and thrombin-ligand sys-

Figure 2: Superposition of the geometry optimised fragments and ligands within the thrombin pocket. The protein and water geometry shown is the optimised L8 system with non-contributing distant amino acids hidden for clarity.

tems using a 800 eV psinc basis set cutoff energy and 8.0 Bohr NGWF radii at the PBE level of theory. In this EDA calculation, the structural waters of L5, L6, L8, and L9 were treated as belonging to the thrombin monomer. The solvent correction energies (ΔE_{SOLV}) were calculated at the same level of theory including the ONETEP solvation model.[44] This was performed using a fixed cavity, with a unitless β parameter of 1.3, density threshold ρ_0 of 0.00035 e$^-$/$a_0{}^3$, and solvent surface tension parameter γ of 0.07415 N/m. In this work, the energies in implicit solvent were calculated using open boundary conditions while the vacuum energies and EDA components were calculated using periodic boundary conditions for reasons of computational performance. Thus the final interaction energies in solvent are in open boundary conditions and the implicit solvation energies include the correction from periodic to open boundary conditions.

To quantify the error introduced through fragmentation of the ligands, an 'additivity' error is defined for each energy component. This quantity is expressed as

$$\Delta E_{\text{err},Y}(X) = \Delta E_Y(X) - \sum_{A \in X}^{N_{\text{frag}}} \Delta E_Y(A) \tag{8}$$

where $\Delta E_Y(X)$ is the value of the EDA energy component Y for the ligand X - protein interaction, $\Delta E_Y(A)$ is the value of the EDA energy component Y for the fragment A - protein interaction, and where N_{frag} is the number of fragments A constituting the ligand X.

4   Analysis: S1 pocket ligands

In this section, we analyse the binding of the fragment ligands (L1 to L6) with the thrombin S1 pocket, with specific attention towards use of EDA as a tool for drug design. The structures of these ligands are shown in Figure 4, and plots of the EDA components are displayed in Figure 7 and Figure 8. Energy component values are also provided in the Supporting Information (Supporting Information Available). EDD and ΔEDD plots for the polarisation and charge transfer interactions are displayed in Figure 5 and Figure 6 respectively. Typically, fragment based drug design seeks to modify the structural motifs of ligands in a systematic fashion to improve the characteristics of
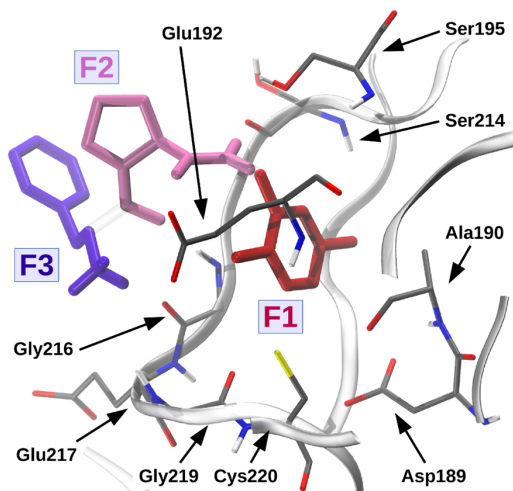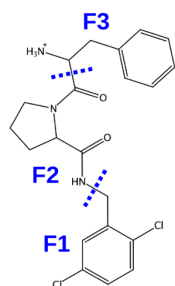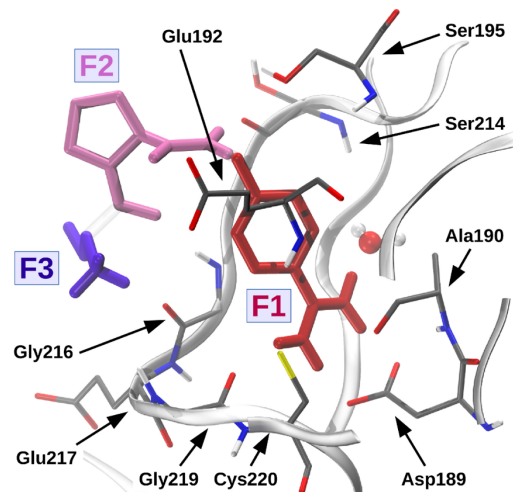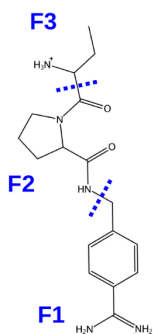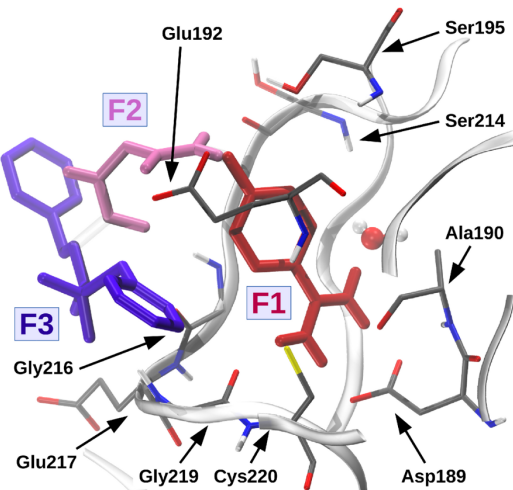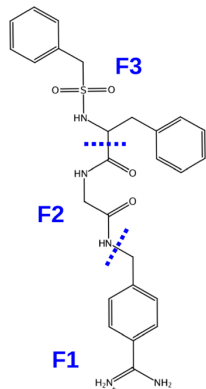
Figure 3: The ligand fragmentation scheme for the L7, L8, and L9 ligands. The ligand PDB codes are given in parentheses. The bond partitions of the ligands are shown by the dashed blue lines, with each fragment produced capped with a hydrogen atom.

a given drug. In the following discussion, we analyse the ligands as pairs in order to demonstrate how EDA can be used to pinpoint the changes in the energy components caused by modifying functional groups, and hence guide fragment based drug design applications.

4.1   Ligands L1 and L2

Here we consider the interactions of the structurally comparable L1 and L2 ligands which both share a similar chlorobenzene functional group. These ligands show notable overlap when superposed within the binding pocket and therefore present strong candidates for comparison using EDA.
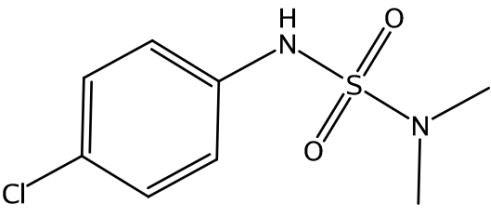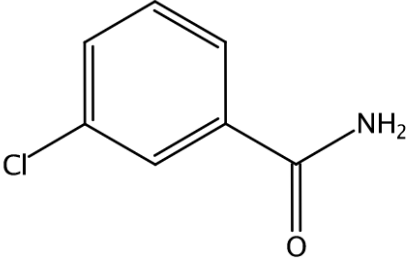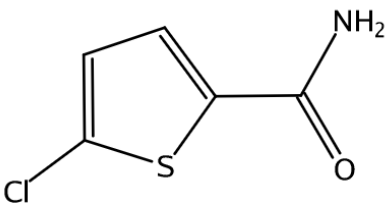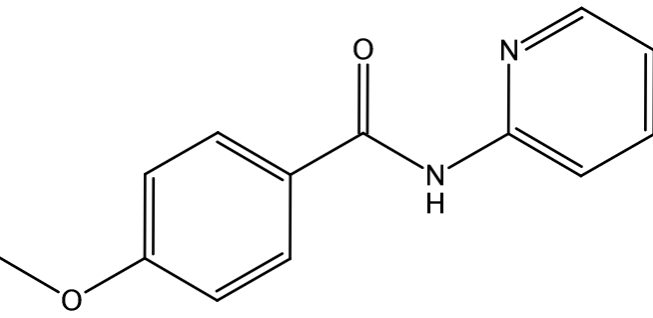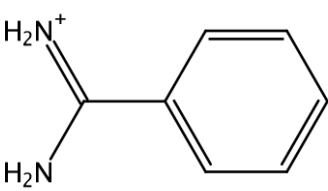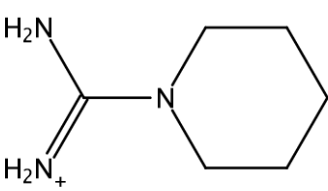
Although these ligands are conformationally similar, differences are observed in the Ser195 and Glu192 protein residue positions as seen in the EDD and ΔEDD plots of Figure 5(a)-(c) and Figure 6(a)-(c). In these plots, we see that the para- and meta-substitution of these ligands results in different interactions with alternate sides of the binding pocket. Specifically, the para-substituted L1 sulfonamide interacts primarily with the Gly219 residue, whereas the meta-substituted L2 amide is observed to interact with the Ser195 residue alone.

We begin by considering the frozen interactions of the L1 and L2 ligands with the thrombin protein. The magnitudes of the individual frozen density interaction components are much larger in the L1 system than the L2 system. Notably, Pauli repulsions are 13.0 kcal/mol more repulsive in the L1 interaction. This is due to the greater size of the L1 ligand than the L2 ligand, despite the close interaction (1.7Å separation) of the L2 ligand with the Ser195 hydrogen. Overall, the stronger interactions seen for L1 balance to give an overall frozen density component of 2.1 kcal/mol which is comparable to that seen in L2 (2.8 kcal/mol).

The EDD and ΔEDD plots connect the EDA energetics with the structural features of ligand-protein interactions. Importantly, these allow us to see the ligand functional groups that are most responsible for the interactions. In the EDD plots, we see polarisation and charge transfer interactions of the L1 sulfamoyl group with Gly219, and the L2 amide group with Ser195. In the ΔEDD plot, we observe clear cancellation of the minor Asp189 residue charge transfer interactions

Figure 4: Small fragment ligands of the thrombin S1 pocket for EDA.

Ligand	PDB code	Chemical structure
L1	5AHG	
L2	5AFY	
L3	4UD9	
L4	5AF9	
L5	4UEH	
L6	4UE7	

Figure 5: EDD and ΔEDD plots of polarisation for the S1 pocket ligand interactions. Green EDD isosurfaces are used to represent electron gain and magenta isosurfaces are used to represent electron loss. The ΔEDD plots are calculated by subtraction of the EDD plot values (see Eq. 7), with green isosurfaces used to represent contributions from the right-hand ligand and magenta isosurfaces used to represent contributions from the left-hand ligand. Isosurface contour levels are displayed at 0.040 electrons per cubic Angstrom for all systems excluding ligands labelled with asterisks (0.025 electrons per cubic Angstrom). Non-contributing distant amino acids have been hidden for clarity.
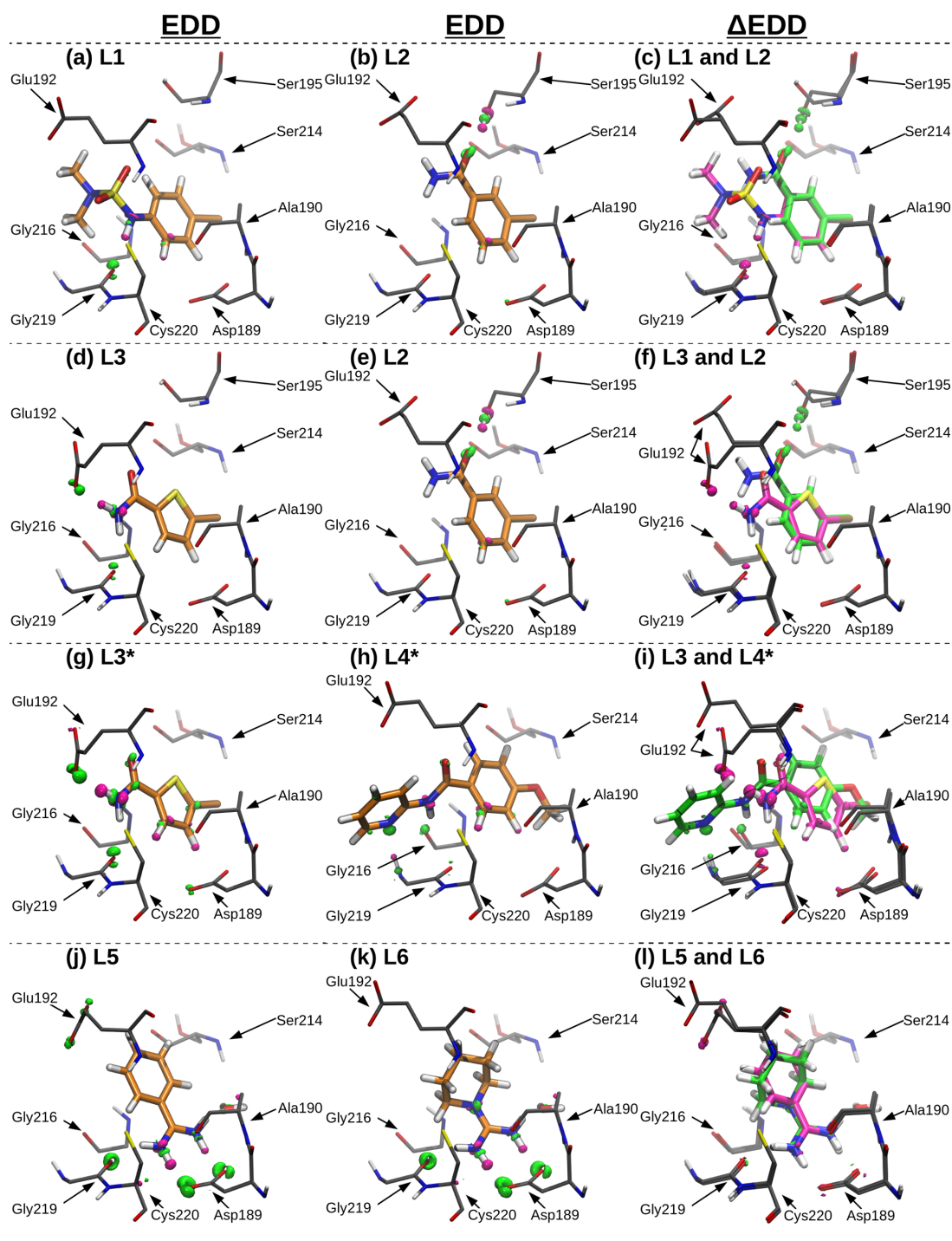
14
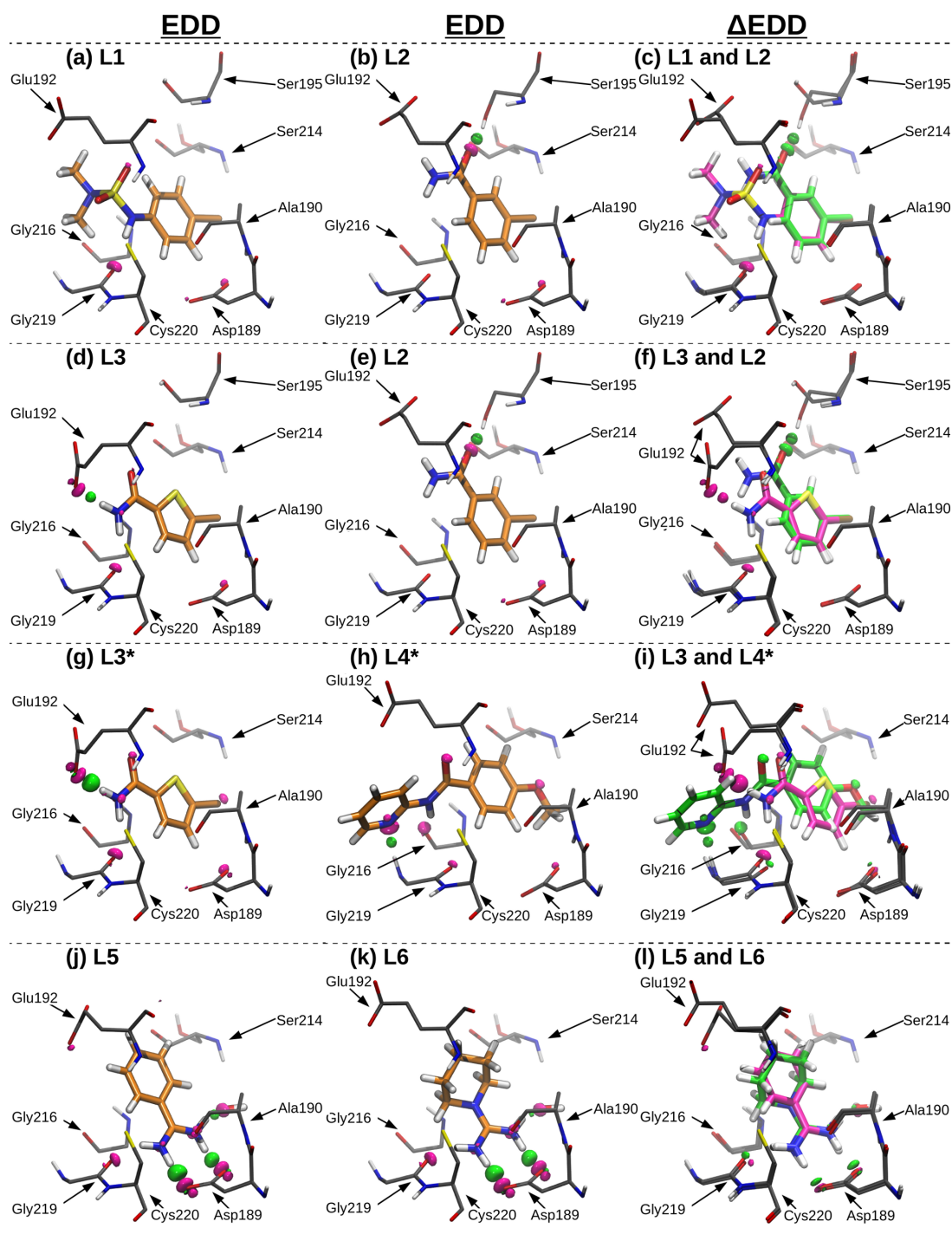
Figure 6: EDD and ΔEDD plots of charge transfer for the S1 pocket ligand interactions. Green EDD isosurfaces are used to represent electron gain and magenta isosurfaces are used to represent electron loss. The ΔEDD plots are calculated by subtraction of the EDD plot values (see Eq. 7), with green isosurfaces used to represent contributions from the right-hand ligand and magenta isosurfaces used to represent contributions from the left-hand ligand. Isosurface contour levels are displayed at 0.040 electrons per cubic Angstrom for all systems excluding ligands labelled with asterisks (0.025 electrons per cubic Angstrom). Non-contributing distant amino acids have been hidden for clarity.

with the two ligands, as we would expect considering their identical chlorobenzene functional groups. Here, the usefulness of ΔEDD plots that remove common interactions and leave only unique contributions is shown. The energetics of polarisation (-6.1 kcal/mol and -5.3 kcal/mol) and charge transfer (-11.4 kcal/mol and -11.1 kcal/mol) are comparable for the L1 and L2 ligands, with L1 displaying slightly stronger interactions. Again, the 0.8 kcal/mol greater polarisation in L1 is explained by the larger size of the ligand: this greater size increases the opportunity for stabilisation through intramolecular charge redistributions during binding. Charge transfer is only 0.3 kcal/mol greater in L1, indicating high energetic similarity of the L2 ligand charge transfer interaction with the Ser195 residue and the L1 ligand with the Gly219 and Ser214 residues.

4.2   Ligands L3 and L2

As in the case of the L1 and L2 ligand pair, the L2 and L3 ligands are structurally comparable and differ only by simple modification of their aromatic ring constituent as seen in Figure 4. Using basic chemical intuition alone, the dipolar thiophene ring of L3 might be expected to enhance electrostatic, polarisation and charge transfer effects above that of the symmetrical benzene ring of L2. While this simple picture is complicated by the presence of chlorine and amide substituents, overall this prediction holds true, with the interaction components of L3 all greater than L2. For example, electrostatics (-33.1 kcal/mol), polarisation (-6.5 kcal/mol), and charge transfer (-14.0 kcal/mol) are greater in L3 than in L2 (-30.3 kcal/mol, -5.3 kcal/mol, and -11.1 kcal/mol respectively).

However, these ligands' interactions differ significantly through the hydrogen bonds that they form with the protein cavity. The L3 ligand forms three observable hydrogen bonds with the Gly219 and Glu192 residues, whereas L2 forms an observable hydrogen bond with Ser195. These interactions are shown clearly in the EDD and ΔEDD plots of Figure 5(d)-(f) and Figure 6(d)-(f). More subtly, the benzene of L2 favours a conformation with two hydrogens interacting with the Asp189 residue, whereas the L3 conformation primarily involves one hydrogen interacting with the Asp189 residue and another with the Cys220 backbone. Overall therefore, these structurally

16
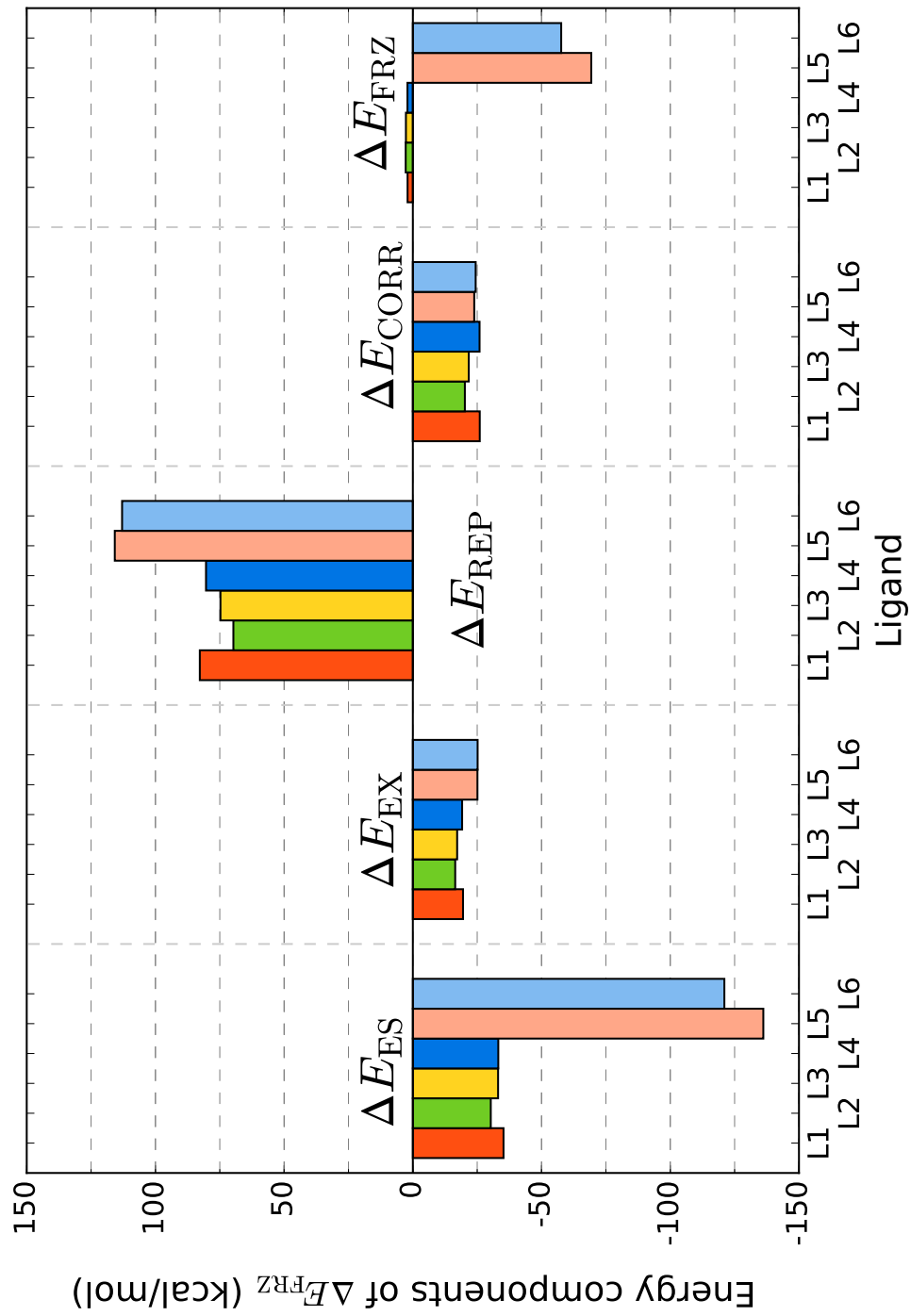
Figure 7: Frozen density analysis of the thrombin S1 pocket complexes calculated at the PBE/800 eV level of theory. The frozen density component (ΔE_{FRZ}) is the sum of the associated left hand bars for each given ligand (see Eq. 2).
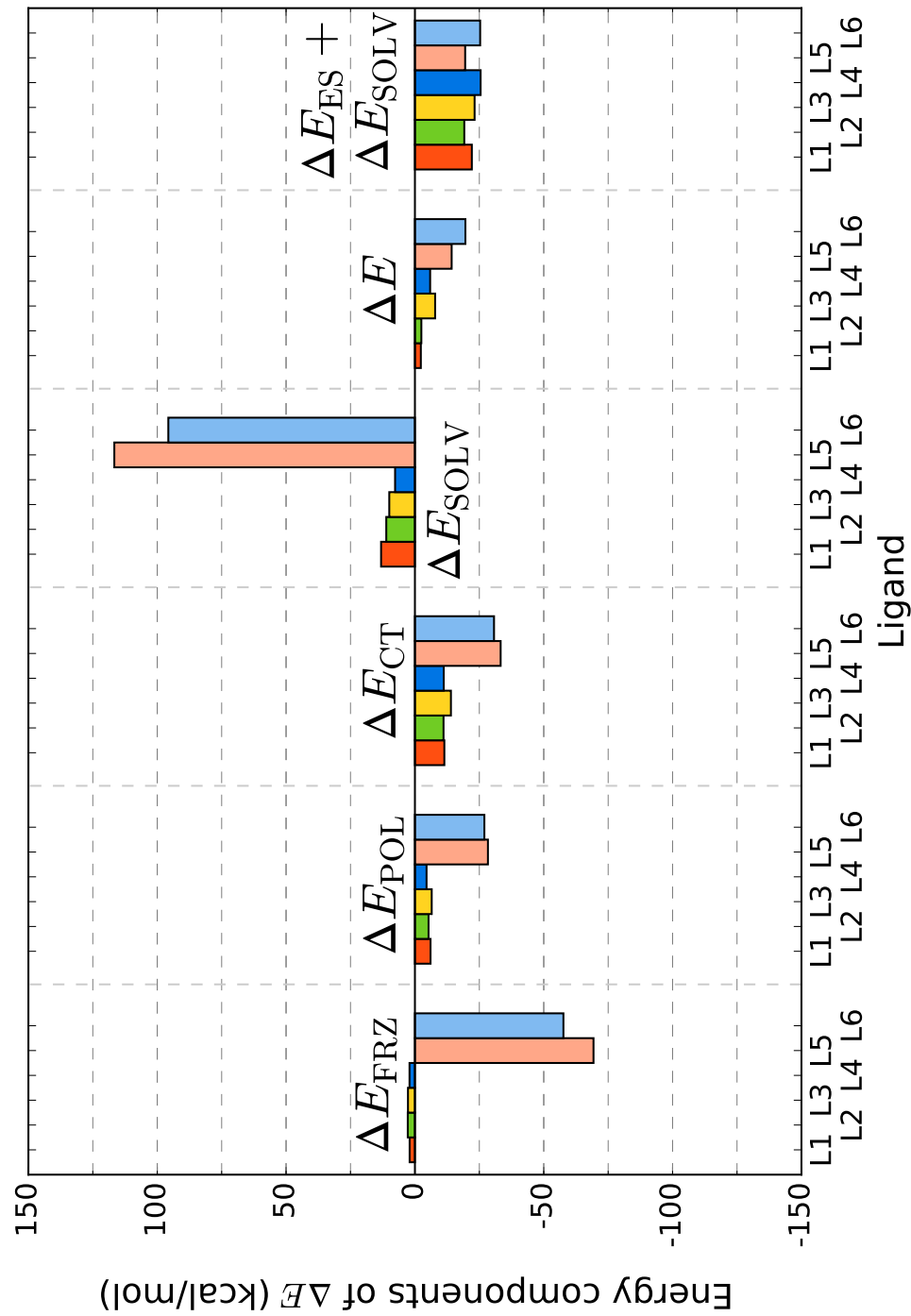
Figure 8: EDA components of the thrombin S1 pocket complexes calculated at the PBE/800 eV level of theory. The total interaction energy (ΔE) is the sum of the associated left hand bars for each given ligand. The separate sum of electrostatics and solvation is included to demonstrate the competition of these effects.

similar binders display very dissimilar interactions with the protein pocket: the difference in the EDA components cannot be convincingly interpreted by considering only the aromatic ring substituent.

The EDD and ΔEDD plots computed offer clear explanation for the driving forces of polarisation and charge transfer, however it is also important to consider the frozen density component of the L3 and L2 interactions. Overall, the frozen density components differ by only 0.1 kcal/mol. At first glance, this may appear counter-intuitive considering the very different interactions so far discussed of these ligands. Here, the ability of the EDA to decompose the frozen density interaction into its individual parts is very useful, allowing the chemist to quantify the cancellation of effects that results in this almost neutral interaction. The greater electrostatics, exchange and correlation in L3 (differing by -5.1 kcal/mol) almost exactly cancel with the greater Pauli repulsions cost (differing by 5.0 kcal/mol) compared with L2. These differences in the individual contributions are reconciled through a number of considerations. Similar to the arguments concerning the polarisation and charge transfer contributions presented above, electrostatics in L3 (-33.1 kcal/mol) are more stabilising than in L2 (-30.3 kcal/mol) due the presence of the dipole moment in the thiophene ring of L3 and due the shorter range interactions of this ligand with the protein. Additionally, the thiophene ring has a slightly greater electron count than benzene which we expect to be the reason for enhanced exchange (by -0.8 kcal/mol) and correlation (by -1.5 kcal/mol) effects in the L3 interaction above L2. Finally, whilst the L2 ligand is larger than the L3 ligand and interacts closely with the Ser195 residue, the L3 ligand is observed to interact strongly with both the Glu192 and Gly219 residues, resulting in overall greater Pauli repulsions for L3 (by 5.0 kcal/mol).

4.3   Ligands L3 and L4

We now consider the interactions of the L3 and L4 ligands. Despite the much smaller size of the L3 ligand compared to the L4 ligand, we observe similar electrostatics of L3 (-33.1 kcal/mol) and L4 (-33.2 kcal/mol). All the other contributions to the frozen density component remain larger in magnitude for the L4 ligand, as expected considering the larger size of this ligand. To understand

this observation, we begin by noting that the L3 and L4 ligands and their protein conformations overlap when superposed yet the Glu192 residue positions differ significantly. This is shown in Figure 5(g)-(i) and Figure 6(g)-(i), where we observe the Glu192 residue actively participating in binding to the L3 ligand. This interaction with the charged Glu192 residue provides additional stability to the L3 ligand through electrostatics, increasing the total electrostatics up to the level of the larger L4 ligand.

The frozen density contribution is slightly more repulsive for L3 (2.7 kcal/mol) than L4 (2.1 kcal/mol). This greater repulsion is due to the presence of additional short range stabilisations for the L3 ligand that reduce the intermolecular separation and increase Pauli repulsions. Specifically, polarisation (-6.5 kcal/mol) and charge transfer (-14.0 kcal/mol) interactions are also seen with the negatively charged Glu192 residue (in addition to the electrostatic interaction discussed above). These interactions favour paying the Pauli repulsions cost (74.7 kcal/mol) of an overall closer intermolecular separation, which in turn results in a more repulsive frozen density component for L3.

The L3 ligand displays stronger polarisation and charge transfer components than L4. However, solvation effects are more repulsive for L3 (10.0 kcal/mol) than L4 (7.7 kcal/mol), competing with polarisation and charge transfer effects. The 2.3 kcal/mol greater destabilisation of the L3 ligand through solvation effects is not as significant as the 4.8 kcal/mol greater stabilisation provided through polarisation and charge transfer and results in an overall stronger interaction energy for the L3 ligand.

4.4   Ligands L5 and L6

Here, we consider the interactions of the two charged ligands of our study, L5 and L6. We begin our discussion by considering the important overall differences between these two ligands' interactions and those of the remaining uncharged ligands. The charged ligands show the largest energy component magnitudes for all but correlation. This observation is explained by considering the structures of the ligands in addition to their charges. The L5 and L6 ligands are smaller

than the remaining uncharged ligands, and therefore we expect the correlation and exchange contributions (which depend on electron count) of L5 and L6 to be weakened by this. On the other hand, these two ligands are positively charged and so favour closer interaction with the protein pocket. A particularly strong interaction of the amidine group of the ligands with the negatively charged Asp189 residue is observed, as shown in Figure 5(j)-(l) and Figure 6(j)-(l). This closer interaction significantly increases Pauli repulsions (115.8 kcal/mol for L5 and 113.0 kcal/mol for L6) above even the large L4 ligand (80.4 kcal/mol). At the same time, this interaction enhances electrostatics (-136.2 kcal/mol for L5 and -121.0 kcal/mol for L6) as we would expect from this interaction of positively and negatively charged functional groups. Exchange is enhanced by this closer interaction such that L5 and L6 are the ligands that are most stabilised through this energy component (-25.1 kcal/mol and -25.2 kcal/mol respectively). However, correlation (-23.8 kcal/mol and -24.4 kcal/mol) is not enhanced enough to be above that seen in the L1 and L4 systems (-26.0 kcal/mol and -25.9 kcal/mol). This may be reconciled as the L1 and L4 ligands contain delocalised electron functional groups with high electron counts which are well suited to correlation interactions with the protein.

An important component of the total interaction energy is the solvation energy ΔE_{SOLV} which accounts for solvation effects omitted in the gas phase calculation of the EDA components (see Eq. 4). This serves to normalise the interaction energies to physically relevant quantities seen in drug-protein binding. Without the solvent interaction correction (i.e. when in gas phase), the L6 and L5 display interaction energies greater than 100 kcal/mol. This arises from the positive charges of these ligands, with electrostatics dominating the protein cavity interaction as discussed above. Necessarily, this stabilisation is also associated with an energy penalty to leave the solvent environment to bind with the protein. This desolvation energy is captured within the ΔE_{SOLV} component, which is 116.7 kcal/mol for L5 and 95.7 kcal/mol for L6. These values are far above those for the neutral ligands (7.7 kcal/mol to 13.1 kcal/mol), as we would expect from the stronger bonds the L5 and L6 would make with the solvent water molecules. Accounting for the effect of solvation brings the interaction energies of L5 (-14.3 kcal/mol) and L6 (-19.6 kcal/mol) into comparable

range of the neutral ligands (-2.3 kcal/mol to -5.9 kcal/mol). These charged ligands interact more strongly than the neutral binders even when including this solvation correction, indicating that payment of a high solvation penalty is acceptable through rewarding the ligands with a more favourable overall enthalpic stability.

Having discussed the important differences between the charged and uncharged ligand interactions, we now consider the interaction energy driving forces of the L5 and L6 ligands. As these ligands are structurally very similar, the differences in the energy components are relatively simple to reconcile. Electrostatics, polarisation and charge transfer are greater in L5 than in L6 by 15.1 kcal/mol, 1.4 kcal/mol and 2.6 kcal/mol respectively. This is likely a direct result of the additional interaction of L5 with the charged Glu192 residue as seen in Figure 5(j) and Figure 6(j). Due to this additional interaction, Pauli repulsions are greater in L5 (115.8 kcal/mol) than in L6 (113.0 kcal/mol), despite the greater number of protons present in L6 which would be expected to increase steric pressure. Correlation is slightly greater in L6 (-24.4 kcal/mol) than in L5 (-23.8 kcal/mol), which is explained by the greater electron count of the L6 ligand.

It is also important to note that whilst the EDA results account for the enthalpic stability (ΔH) of binding, these values do not account for entropic effects (ΔS) involved in quantifying the Gibbs' free energy of binding ($\Delta G = \Delta H - T\Delta S$). For example, the additional rigidity of the L5 ligand compared to L6 may result in the L5 ligand displaying an overall more favourable free energy of binding than L6, despite the greater enthalphic stability in L6. This insight is supported quantitatively by the thermodynamic parameters determined by Rühmann and coworkers.[40]

5   Analysis: fragmented ligands

In this section, we analyse the binding of the L7, L8, and L9 ligands and their fragments with thrombin (shown in Figure 3). Plots of the EDA components are displayed in Figure 9 and Figure 10. Energy component values are also provided in the Supporting Information (Supporting Information Available).

5.1 Frozen density component

We begin our discussion by noting the important role of electrostatics, and in particular those arising from the fragment charges, as a factor in determining the overall ΔE. We expect strong electrostatics in the L8 and L9 ligands through interaction of their amidine functional group with the Asp189 amino acid residue. Importantly, by partitioning the ligands into functional group regions, we are able to quantify the contribution of this amidine-Asp189 interaction. For these two ligands, the amidine functional group forms part of the F1 fragment. We observe total electrostatic energy components of -208.3 kcal/mol (L8) and -169.7 kcal/mol (L9) for the full ligand, of which the F1 fragments contribute -129.2 kcal/mol and -135.7 kcal/mol respectively (Figure 9). From this, we are able to conclude that this salt bridge interaction contributes well over half of the electrostatics in these systems.

Electrostatics are also shown to be important in the charged F3 fragments of L7 and L8. We expect that strong stabilising contributions arise through interaction of the charged amine group with the Gly216 oxygen as shown in Figure 11(a)-(c) and Figure 12(a)-(c). The electrostatic stability provided by these fragments is -71.2 kcal/mol (L7) and -59.3 kcal (L8), with greater stability seen in the L7 ligand due to the presence of benzene in this fragment (see Figure 3). The benzene group of L7 may enhance electrostatics above L8 through inductive effects stabilising the amine charge. Understanding effects such as this is important to developing knowledge of the subtle forces that enhance drug-protein binding, demonstrating the high value of the EDA.

Whilst the ability to extract functional group energetics from the EDA is a powerful ability, it is important to consider the error that may be introduced by the bond partitioning process. Notably, we observe a 6.5 kcal/mol difference in electrostatics for the F1 fragments of the L8 and L9 systems despite these fragments sharing identical chemical structures. Errors in the electrostatic component due to fragmentation (calculated using Eq. 8) are < 1 kcal/mol and therefore bond partitioning itself is unlikely to be the origin of this difference. For this reason we conclude that this difference most likely arises from variances in the structural geometries. This observation demonstrates that care must be taken drawing conclusions from EDA results using a single structure, and it is expected

Figure 9: Frozen density analysis of the fragmented L7, L8, and L9 ligand interactions calculated at the PBE/800 eV level of theory. F1, F2, and F3 refer to the fragments that compose the full ligand. For each energy component, the fragmentation error equals the sum of the fragment contributions subtracted from the contribution of the full ligand (Eq. 8). The frozen density component (ΔE_{FRZ}) is the sum of the associated left hand bars (see Eq. 2).
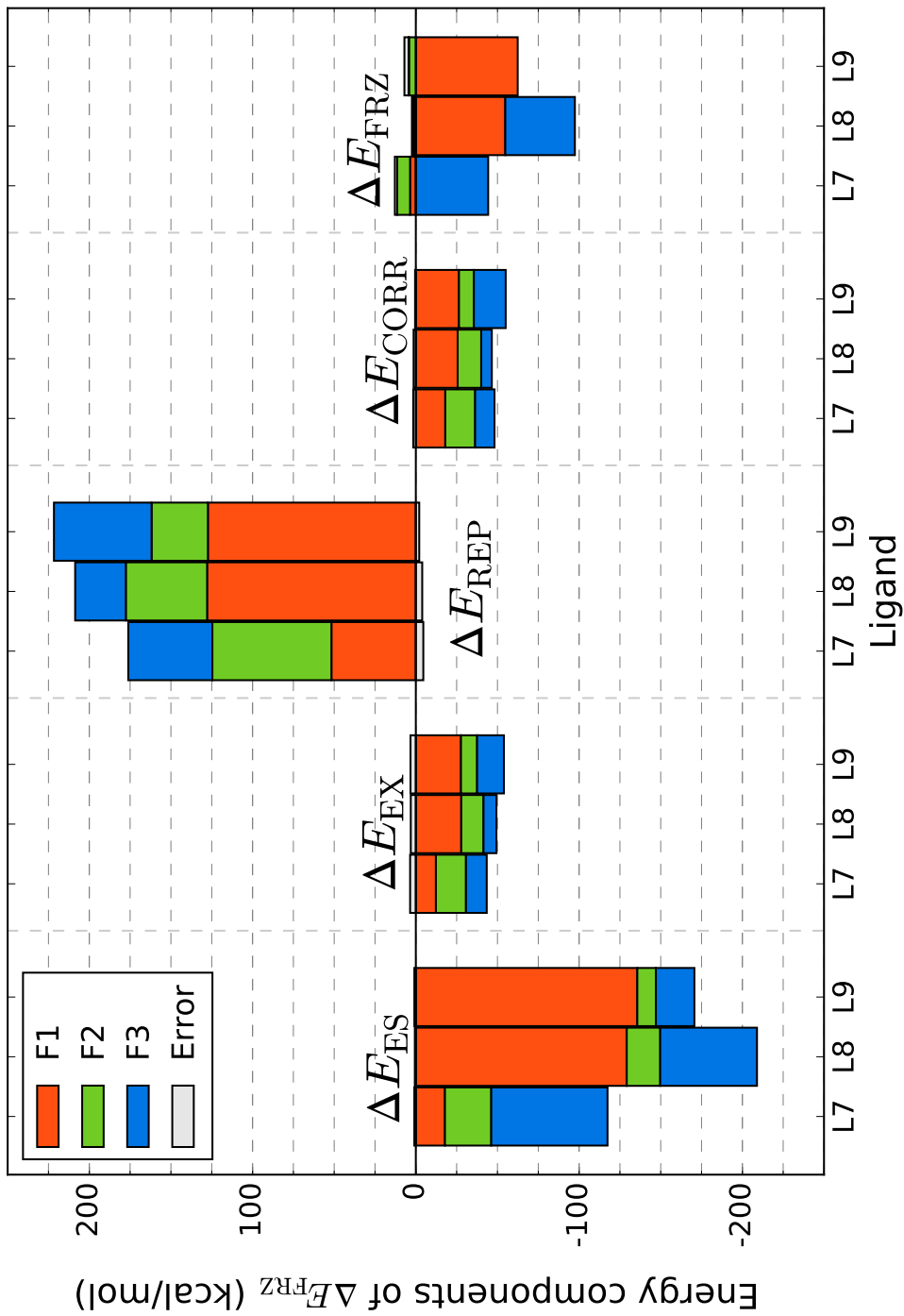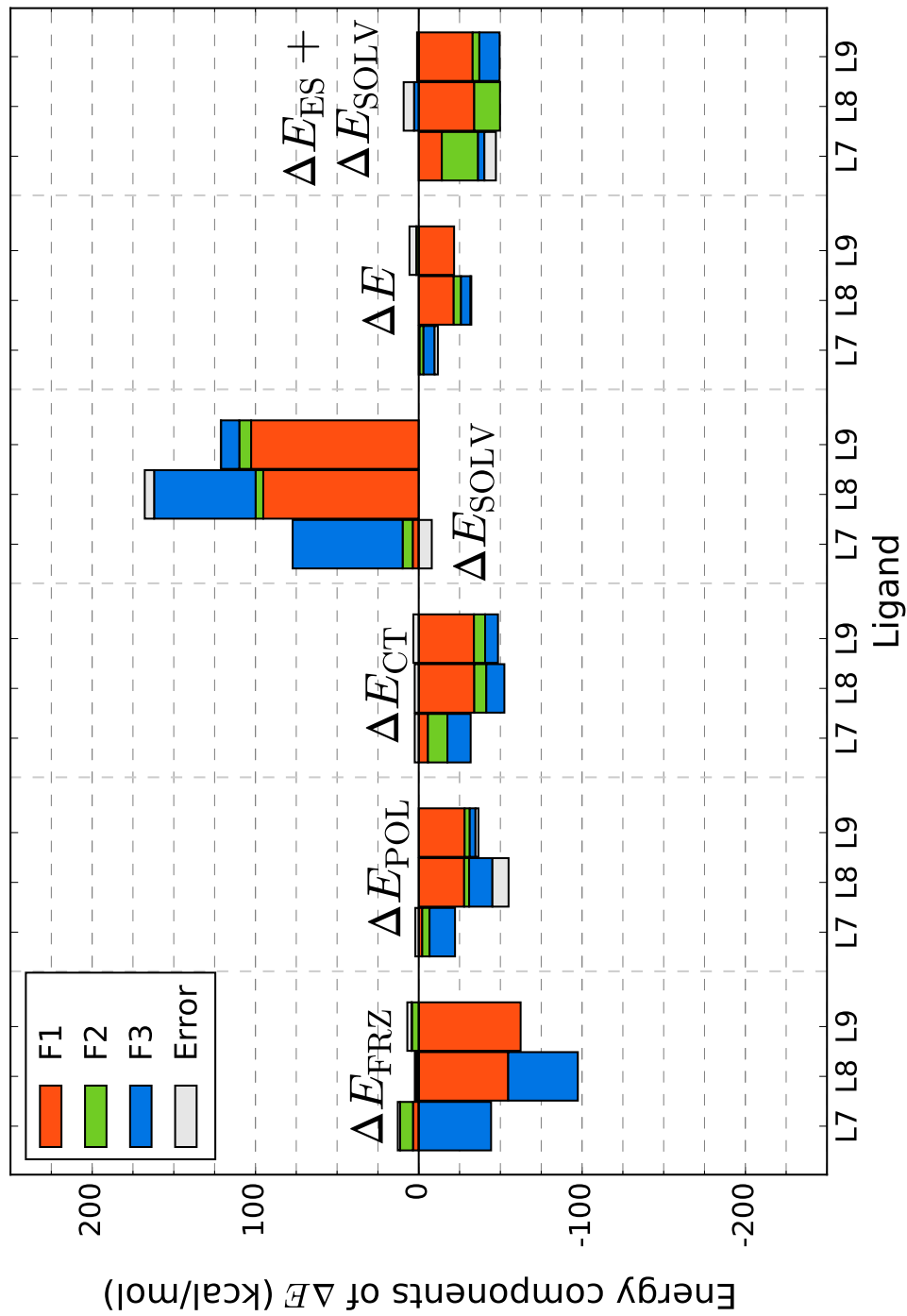
Figure 10: EDA components of the fragmented L7, L8, and L9 ligand interactions calculated at the PBE/800 eV level of theory. F1, F2, and F3 refer to the fragments that compose the full ligand. For each energy component, the fragmentation error equals the sum of the fragment contributions subtracted from the contribution of the full ligand (Eq. 8). The total interaction energy (ΔE) is the sum of the separate sum of electrostatics and solvation is included to demonstrate the competition of these effects.

Figure 11: EDD and ΔEDD plots of polarisation for the ligand interactions. Green EDD isosurfaces are used to represent electron gain and magenta isosurfaces are used to represent electron loss. The ΔEDD plots are calculated by subtraction of the EDD plot values (see Eq. 7), with green isosurfaces used to represent contributions from the right-hand ligand and magenta isosurfaces used to represent contributions from the left-hand ligand. All isosurface contour levels are displayed at 0.040 electrons per cubic Angstrom. Non-contributing distant amino acids have been hidden for clarity.
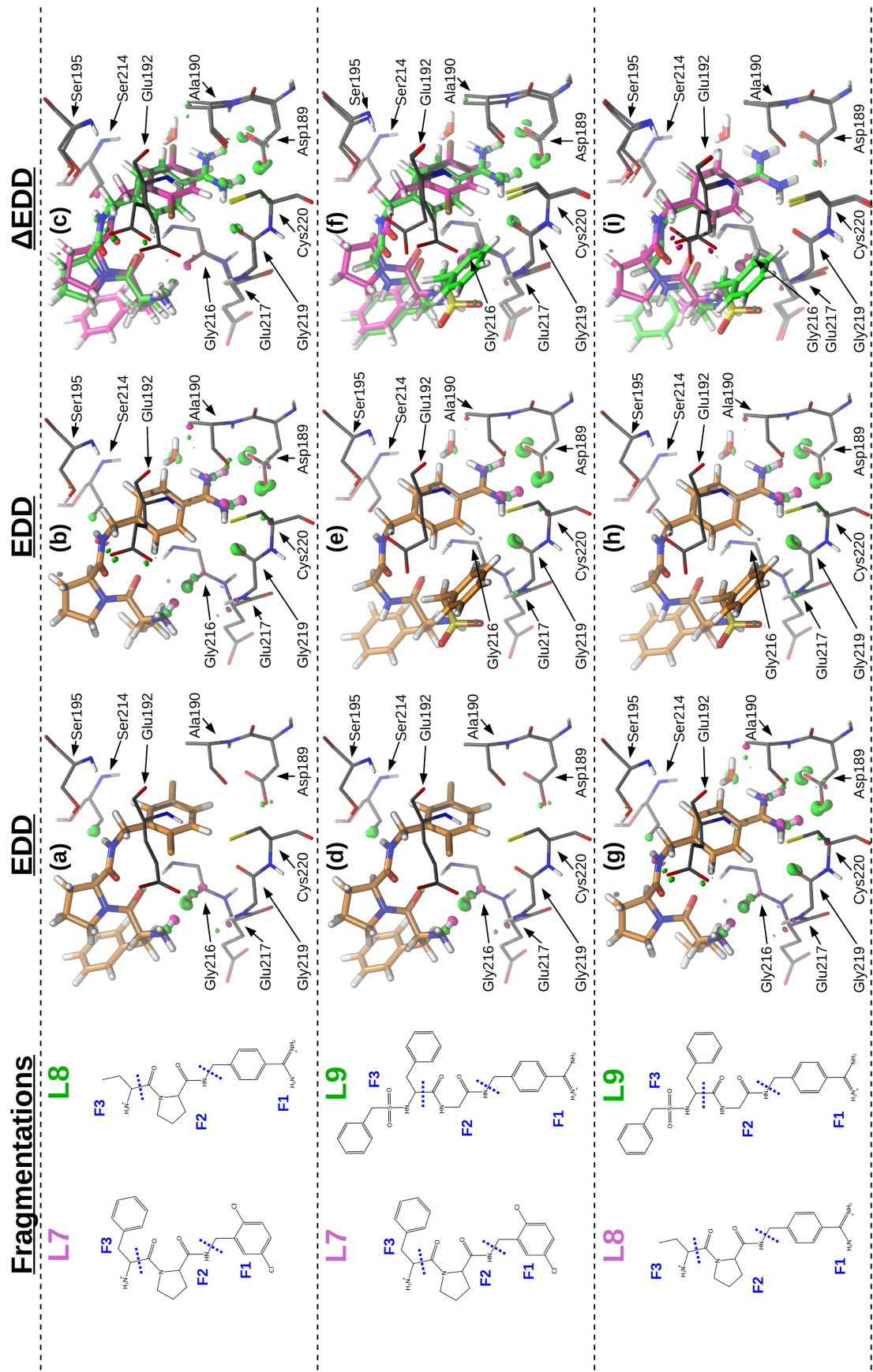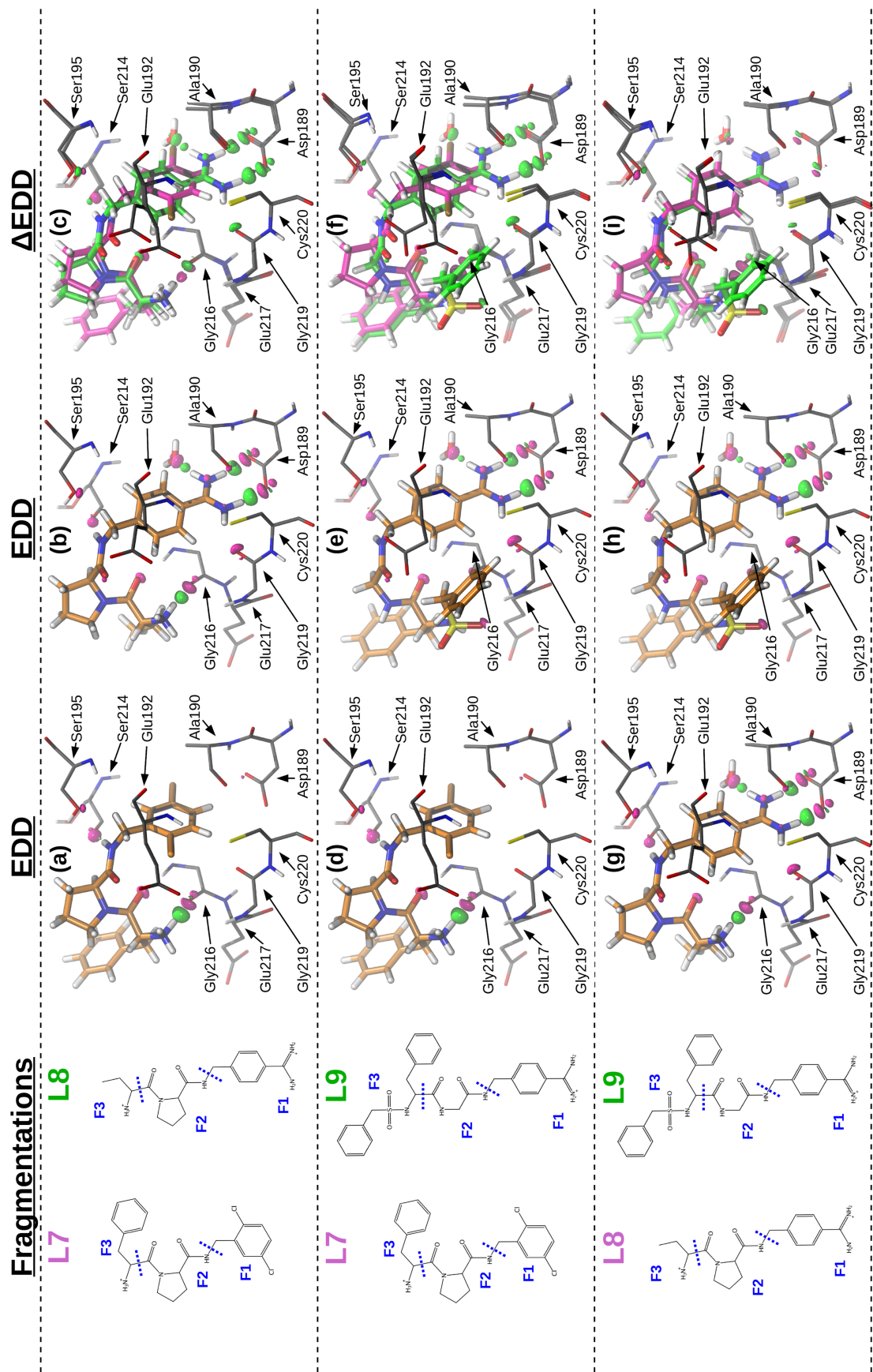
Figure 12: EDD and ΔEDD plots of charge transfer for the ligand interactions. Green EDD isosurfaces are used to represent electron gain and magenta isosurfaces are used to represent electron loss. The ΔEDD plots are calculated by subtraction of the EDD plot values (see Eq. 7), with green isosurfaces used to represent contributions from the right-hand ligand and magenta isosurfaces used to represent contributions from the left-hand ligand. All isosurface contour levels are displayed at 0.040 electrons per cubic Angstrom. Non-contributing distant amino acids have been hidden for clarity.

that improvements in the accuracy of EDA results would be made by computing averages of EDA component values using snapshots taken from molecular dynamics simulations.

Overall, the fragment charge is observed to play an important role in determining the frozen density component. For L7 and L9, each containing one charged fragment, the frozen density component is predominantly stabilised by its charged fragment (-44.3 kcal/mol by F3 of L7, and -62.4 kcal/mol by F1 of L9). For L8, containing two charged fragments, the frozen density component is stabilised by similar amounts through both of its charged fragments (-54.8 kcal/mol by F1 and -42.6 kcal/mol by F3). The L8 and L9 F1 fragments show strong electrostatic stabilisation through salt bridge formation with the Asp189 residue. A high Pauli repulsions penalty (approximately 75.9 kcal/mol) is paid by L8 and L9 to form the salt bridge, highlighting the importance of sterics when suggesting modifications for drug design.

The fragment exchange and correlation contributions to ΔE are largely determined by the fragment sizes and proximity to the protein. For example, the F3 fragment of the L8 system is shown to contribute -7.9 kcal/mol and -6.5 kcal/mol stability to ΔE through exchange and correlation effects respectively. As we would expect, in the larger and more closely interacting F3 fragment of L7 exchange and correlation are also greater (-12.7 kcal/mol and -11.9 kcal/mol respectively). Overall, the full ligand contributions of exchange and correlation (up to -50.7 kcal/mol and -54.6 kcal/mol respectively) are comparable to the contributions of polarisation (up to -55.0 kcal/mol) and charge transfer (up to -50.0 kcal/mol). These effects therefore contribute important driving forces to binding. The fragmentation errors associated with exchange are consistently positive (3.1 kcal/mol to 3.5 kcal/mol), indicating that the fragment contributions slightly overestimate the stabilisation provided through exchange. This error is moderate when considering that this represents a 6.5-8.8% exchange contribution unattributable to any particular fragment.

An important aspect of the EDA is its ability to highlight potential drug modification regions that would induce stronger protein interaction. For example, in the F1 fragment interaction of L8 and L9 discussed earlier the benzene provides stability to the salt bridge interaction through resonance effects involving the amidine group. Here, the EDA might be used to evaluate placement of

a functional group such as a halide to the benzene ring to both stabilise this interaction further and offer enhancement of electrostatic effects. Additionally, the EDA may support modification of the ligand F2 fragments due to their consistently repulsive frozen density component. Specifically, the electrostatic interactions of the F2 fragments are observed to be relatively small (-11.4 kcal/mol to -28.4 kcal/mol) when compared to the electrostatic interactions of the remaining fragments. Substitutions of these fragments tailored towards enhancing electrostatics may therefore be desired to attempt to increase the overall ligand binding enthalpy.

5.2   Polarisation and charge transfer

Here we consider the polarisation and charge transfer EDA components for the fragmented systems through their energy values and the EDD and ΔEDD plots shown in Figure 11 and Figure 12.

The L8 and L9 ligands are observed to largely contribute to polarisation and charge transfer through their F1 fragments. These fragments show a similar total polarisation and charge transfer ($\Delta E_{\mathrm{POL}} + \Delta E_{\mathrm{CT}}$) contribution of -61.8 kcal/mol and -61.9 kcal/mol respectively. This strong interaction is due to the salt bridge interaction with Asp189, with contributions also observed from Gly219 and Cys220 and the water molecule in this pocket. As shown in Figure 11(g)-(i), charge is seen to gather at the Asp189 oxygens by attraction of the positively charged amidine group of the ligand. Simultaneously, charge density is repelled away from the amidine group hydrogen atoms in the direction of the benzene ring. Two hydrogen bonds are subsequently formed by charge transfer from the Asp189 residue to the amidine group, as shown in Figure 12(g)-(i).

The EDD and ΔEDD plots show a number of other protein and ligand regions involved in key interactions. The Ser214 residue is shown to be involved in interactions with the F2 fragments of every ligand of this study. For example, in the L7 system, charge is seen to gather on the oxygen of Ser214 through polarisation and form a bond with a nitrogen of the ligand prolinamide group. The overall interaction of the protein with the L7 F2 fragment is moderately strong and likely stabilised greatly by this specific interaction: polarisation is -4.5 kcal/mol and charge transfer is -11.9 kcal/mol. This interaction is shown for L7 in Figure 11(a) and Figure 12(a).

In the L7 and L8 F3 fragments, stabilising contributions are also seen to arise through interaction of the ligands' charged amine groups with the Gly216 oxygen that result in formation of a $^{\delta+}$NH···OC bond (Figure 11(a)-(c) and Figure 12(a)-(c)). In the case of L9, the F3 fragment has a lesser interaction with the Gly216 oxygen atom due to the fact that this fragment is uncharged. However, this fragment is more extended than the F3 fragments of L7 and L8 and as a result is able to interact with a larger area of the cavity surface. The overall weaker stabilisation provided by these interactions is reflected by a > 10 kcal/mol decrease in polarisation compared to the F3 fragments of L7 and L8, and a > 3 kcal/mol decrease in charge transfer.

Notably, the fragmentation error for the polarisation component of the L8 system is observed to be relatively high. This error is −9.9 kcal/mol, and therefore 18% of the polarisation is not attributable to any particular fragment. This is possibly the result of the fragmentation scheme we have adopted for this system, as the partitioned bonds will be unable to reproduce the polarisation effects seen in the full ligand. Related to this idea, it is possible that giving the F3 fragment of L8 a +1 integer charge is physically unjustifiable. This may be the case as this fragment is relatively small and may stabilise its electron deficiency by withdrawing charge from its neighbouring fragment (F2). This is evidenced by population analysis of the full ligand (performed at the same level of theory as the EDA) in which the fragment is observed to have a net charge of +0.91e (Mulliken population) and +0.89e (natural population), different to a +1 integer charge. This observation shows that care must be taken when partitioning bonds, especially in the case of small, charged fragments which should ideally be avoided.

5.3   Solvation

As we may expect for the ligands studied, solvation effects are seen to make substantial contributions to the binding energy of the ligands. By decomposing the solvation correction into different contributing regions of the ligand, we are able to gain insight into the effects of functional groups on this energy component. As discussed earlier, for charged fragments and ligands the electrostatic component is negatively correlated with the solvation energy. This is due to fact that elec-

trostatically driven binding to the protein is associated with an electrostatic energy cost of solvent bond breaking to leave the solvent environment. For example, solvent effects for the charged F1 fragments of L8 (95.2 kcal/mol) and L9 (102.7 kcal/mol) mediate the strongly stabilising electrostatic interactions of these fragments with the protein (-129.2 kcal/mol and -135.7 kcal/mol respectively). Including the solvation effects in our description of the interaction energy ΔE, we see that overall the interactions are dominated by the salt bridge formation. This interaction involves the amidine groups contained in the F1 fragments of L8 ($\Delta E = -21.4$ kcal/mol) and L9 ($\Delta E = -21.6$ kcal/mol).

6   Conclusions

In this work, we have presented an application of the ONETEP energy decomposition analysis (EDA) to a series of binders interacting with the thrombin protein. Investigations of biomolecular interactions using EDA in the literature have often been limited to smaller model system studies and without seeking to visually analyse the specific structural regions involved in the interactions.[38,50,51] Here, we have shown the potential for EDA to aid in drug design through use of the energy components in combination with electron density difference (EDD) and ΔEDD plots for visualising the interactions. We note that the visualisation of electron redistributions through such plots is not directly possible using molecular mechanics-based approaches. This work applying EDA to protein-ligand systems greater than 2740 atoms in size therefore exemplifies the strength of large-scale quantum approaches for drug design.[52,53]

We have used the EDA to highlight a number of important interactions during thrombin binding. Specifically, we considered the interactions of a series of small S1 pocket fragments and larger drug-sized ligands with the binding cavity. A partitioning protocol for the ligands was introduced that enabled their interactions to be decomposed into their structural origins. This protocol was a simple yet robust procedure which we showed to maintain a high degree of chemical accuracy.

Our analysis evidenced the importance of the salt bridge interaction with the charged amidine

group of the ligands, and highlighted the Asp189, Glu192, Ser195, Gly219 and Gly216 amino acid residues as playing important roles in thrombin binding through polarisation and charge transfer effects. The conformation of the highly flexible Glu192 was shown to be a significant factor for differences in the interactions of structurally comparable ligands such as the L5 and L6 ligand pair and the L8 and L9 ligand pair.

This work demonstrates the application of a whole new capability to a real protein drug target where accurate DFT calculations can produce both energetic and visual descriptors of interactions. These descriptors can be used to provide insights for tailoring interactions, as needed for example in drug design.

Acknowledgement

Supporting Information Available

- SI.1 Stability of the ONETEP EDA approach with respect to geometric coordinate variations.

- SI.2 Comparison with experimental thermodynamic parameters.

- SI.3 PBE/800 eV ONETEP EDA values: the thrombin S1 pocket small binders (Table S1), the fragmented ligands (Table S2).

- SI.4 ONETEP EDA input files.

This material is available free of charge via the Internet at `http://pubs.acs.org`.

References

(1) Ewing, C. S.; Veser, G.; McCarthy, J. J.; Lambrecht, D. S.; Johnson, J. K. Predicting catalyst-support interactions between metal nanoparticles and amorphous silica supports. *Surf. Sci.* **2016**, *652*, 278–285.

(2) Wheeler, S. E.; Seguin, T. J.; Guan, Y.; Doney, A. C. Noncovalent Interactions in Organocatalysis and the Prospect of Computational Catalyst Design. *Acc. Chem. Res.* **2016**, *49*, 1061–1069.

(3) Cavani, F.; Trifirò, F. Selective oxidation of light alkanes: interaction between the catalyst and the gas phase on different classes of catalytic materials. *Catal. Today* **1999**, *51*, 561–580.

(4) Sweetnam, S.; Vandewal, K.; Cho, E.; Risko, C.; Coropceanu, V.; Salleo, A.; Brédas, J.-L.; McGehee, M. D. Characterizing the polymer: fullerene intermolecular interactions. *Chem. Mater.* **2016**, *28*, 1446–1452.

(5) Zhang, J.; Xu, Y.; Cui, L.; Fu, A.; Yang, W.; Barrow, C.; Liu, J. Mechanical properties of graphene films enhanced by homo-telechelic functionalized polymer fillers via $\pi - \pi$ stacking interactions. *Composites Part A* **2015**, *71*, 1–8.

(6) Yang, X.; Yu, H.; Wang, L.; Tong, R.; Akram, M.; Chen, Y.; Zhai, X. Self-healing polymer materials constructed by macrocycle-based host-guest interactions. *Soft Matter* **2015**, *11*, 1242–1252.

(7) Schwartz, B. J. CONJUGATED POLYMERS AS MOLECULAR MATERIALS: How Chain Conformation and Film Morphology Influence Energy Transfer and Interchain Interactions. *Annu. Rev. Phys. Chem.* **2003**, *54*, 141–172.

(8) Kwon, M. S.; Lee, D.; Seo, S.; Jung, J.; Kim, J. Tailoring Intermolecular Interactions for Efficient Room-Temperature Phosphorescence from Purely Organic Materials in Amorphous Polymer Matrices. *Angew. Chem. Int. Ed.* **2014**, *53*, 11177–11181.

(9) Delle Piane, M.; Corno, M.; Ugliengo, P. Does Dispersion Dominate over H-Bonds in Drug-Surface Interactions? The Case of Silica-Based Materials As Excipients and Drug-Delivery Agents. *J. Chem. Theory Comp.* **2013**, *9*, 2404–2415.

(10) Farmanzadeh, D.; Najafi, M. Theoretical study of anticancer properties of indolyl-oxazole drugs and their interactions with DNA base pairs in gas phase and solvent. *Struct. Chem.* **2015**, *26*, 831–844.

(11) Wang, Y.-T.; Chen, Y.-C. Insights from QM/MM Modeling the 3D Structure of the 2009 H1N1 Influenza A Virus Neuraminidase and Its Binding Interactions with Antiviral Drugs. *Mol. Inf.* **2014**, *33*, 240–249.

(12) Phipps, M. J. S.; Fox, T.; Tautermann, C. S.; Skylaris, C.-K. Energy decomposition analysis approaches and their evaluation on prototypical protein-drug interaction patterns. *Chem. Soc. Rev.* **2015**, *44*, 3177–3211.

(13) Morokuma, K. Molecular Orbital Studies of Hydrogen Bonds. III. C=O$\cdots$H-O Hydrogen Bond in H2CO$\cdots$H2O and H2CO$\cdots$2H2O. *J. Chem. Phys.* **1971**, *55*, 1236–1244.

(14) Stevens, W. J.; Fink, W. H. Frozen fragment reduced variational space analysis of hydrogen bonding interactions. Application to the water dimer. *Chem. Phys. Lett.* **1987**, *139*, 15–22.

(15) Bagus, P. S.; Hermann, K.; Bauschlicher, C. W. A new analysis of charge transfer and polarization for ligand-metal bonding: Model studies of Al4CO and Al4NH3. *J. Chem. Phys.* **1984**, *80*, 4378–4386.

(16) Glendening, E. D.; Streitwieser, A. Natural energy decomposition analysis: An energy partitioning procedure for molecular interactions with application to weak hydrogen bonding, strong ionic, and moderate donor-acceptor interactions. *J. Chem. Phys.* **1994**, *100*, 2900–2909.

(17) Glendening, E. D. Natural Energy Decomposition Analysis:âĂĽ Extension to Density Functional Methods and Analysis of Cooperative Effects in Water Clusters. *J. Phys. Chem. A.* **2005**, *109*, 11936–11940.

(18) Foster, J. P.; Weinhold, F. Natural hybrid orbitals. *J. Am. Chem. Soc.* **1980**, *102*, 7211–7218.

(19) Su, P.; Li, H. Energy decomposition analysis of covalent bonds and intermolecular interactions. *J. Chem. Phys.* **2009**, *131*, 014102.

(20) Khaliullin, R. Z.; Cobar, E. A.; Lochan, R. C.; Bell, A. T.; Head-Gordon, M. Unravelling the Origin of Intermolecular Interactions Using Absolutely Localized Molecular Orbitals. *J. Phys. Chem. A* **2007**, *111*, 8753–8765.

(21) Phipps, M. J.; Fox, T.; Tautermann, C. S.; Skylaris, C.-K. Energy decomposition analysis based on absolutely localised molecular orbitals for large-scale density functional theory calculations in drug design. *J. Chem. Theory Comp.* **2016**, *12*, 3135–3148.

(22) Wu, Q.; Ayers, P. W.; Zhang, Y. Density-based energy decomposition analysis for intermolecular interactions with variationally determined intermediate state energies. *J. Chem. Phys.* **2009**, *131*, 164112.

(23) Jeziorski, B.; Moszynski, R.; Ratkiewicz, A.; Rybak, S.; Szalewicz, K.; Williams, H. L. SAPT: A program for many-body symmetry-adapted perturbation theory calculations of intermolecular interaction energies. *METECC–94* **1993**, *B*, 79–129.

(24) Jeziorski, B.; Moszynski, R.; Szalewicz, K. Perturbation Theory Approach to Intermolecular Potential Energy Surfaces of van der Waals Complexes. *Chem. Rev.* **1994**, *94*, 1887–1930.

(25) Kitaura, K.; Morokuma, K. A new energy decomposition scheme for molecular interactions within the Hartree-Fock approximation. *Int. J. Quantum Chem.* **1976**, *10*, 325–340.

(26) Morokuma, K. Why do molecules interact? The origin of electron donor-acceptor complexes, hydrogen bonding and proton affinity. *Acc. Chem. Res.* **1977**, *10*, 294–300.

(27) Khaliullin, R. Z.; Head-Gordon, M.; Bell, A. T. An efficient self-consistent field method for large systems of weakly interacting components. *J. Chem. Phys.* **2006**, *124*, 204105.

(28) Khaliullin, R. Z.; Bell, A. T.; Head-Gordon, M. Analysis of charge transfer effects in molecular complexes based on absolutely localized molecular orbitals. *J. Chem. Phys.* **2008**, *128*, 184112.

(29) Mo, Y.; Gao, J.; Peyerimhoff, S. D. Energy decomposition analysis of intermolecular interactions using a block-localized wave function approach. *J. Chem. Phys.* **2000**, *112*, 5530–5538.

(30) Mo, Y.; Song, L.; Lin, Y. Block-Localized Wavefunction (BLW) Method at the Density Functional Theory (DFT) Level. *J. Phys. Chem. A* **2007**, *111*, 8291–8301.

(31) Mo, Y.; Bao, P.; Gao, J. Energy decomposition analysis based on a block-localized wavefunction and multistate density functional theory. *Phys. Chem. Chem. Phys.* **2011**, *13*, 6760–6775.

(32) Wu, Q. Variational nature of the frozen density energy in density-based energy decomposition analysis and its application to torsional potentials. *J. Chem. Phys.* **2014**, *140*, 244109.

(33) Horn, P. R.; Head-Gordon, M. Polarization contributions to intermolecular interactions revisited with fragment electric-field response functions. *J. Chem. Phys.* **2015**, *143*, 114111.

(34) Fedorov, D. G.; Kitaura, K. Pair interaction energy decomposition analysis. *J. Comput. Chem.* **2007**, *28*, 222–237.

(35) Heifetz, A.; Trani, G.; Aldeghi, M.; MacKinnon, C. H.; McEwan, P. A.; Brookfield, F. A.; Chudyk, E. I.; Bodkin, M.; Pei, Z.; Burch, J. D.; Ortwine, D. F. Fragment Molecular Orbital Method Applied to Lead Optimization of Novel Interleukin-2 Inducible T-Cell Kinase (ITK) Inhibitors. *J. Med. Chem.* **2016**, *59*, 4352–4363.

(36) Heifetz, A.; Aldeghi, M.; Chudyk, E. I.; Fedorov, D. G.; Bodkin, M. J.; Biggin, P. C. Using the fragment molecular orbital method to investigate agonist–orexin-2 receptor interactions. *Biochem. Soc. Trans.* **2016**, *44*, 574–581.

(37) Heifetz, A.; Chudyk, E. I.; Gleave, L.; Aldeghi, M.; Cherezov, V.; Fedorov, D. G.; Biggin, P. C.; Bodkin, M. J. The Fragment Molecular Orbital Method Reveals New Insight into the Chemical Nature of GPCR-Ligand Interactions. *J. Chem. Inf. Model.* **2016**, *56*, 159–172.

(38) Church, J.; Pezeshki, S.; Davis, C.; Lin, H. Charge Transfer and Polarization for Chloride Ions Bound in ClC Transport Proteins: Natural Bond Orbital and Energy Decomposition Analyses. *J. Phys. Chem. B* **2013**, *117*, 16029–16043.

(39) Fedorov, D. G.; Kitaura, K. Subsystem Analysis for the Fragment Molecular Orbital Method and Its Application to Protein-Ligand Binding in Solution. *J. Phys. Chem. A* **2016**, *120*, 2218–2231.

(40) Rühmann, E.; Betz, M.; Heine, A.; Klebe, G. Fragment Binding Can Be Either More Enthalpy-Driven or Entropy-Driven: Crystal Structures and Residual Hydration Patterns Suggest Why. *J. Med. Chem.* **2015**, *58*, 6960–6971.

(41) Stubbs, M. T.; Bode, W. A player of many parts: The spotlight falls on thrombin's structure. *Thromb. Res.* **1993**, *69*, 1–58.

(42) Lomas, O.; Argyle, R.; Prendergast, B. Will direct thrombin inhibition change the boundaries of oral anti-coagulation? *QJM* **2010**, *103*, 429–434.

(43) Skylaris, C.-K.; Haynes, P. D.; Mostofi, A. A.; Payne, M. C. Introducing ONETEP: Linear-scaling density functional simulations on parallel computers. *J. Chem. Phys.* **2005**, *122*, 084119.

(44) Dziedzic, J.; Helal, H. H.; Skylaris, C.-K.; Mostofi, A. A.; Payne, M. C. Minimal parameter implicit solvent model for ab initio electronic-structure calculations. *Europhys. Lett.* **2011**, *95*, 43001.

(45) The frozen density refers to the charge density constructed using the idempotent density kernel that ensures orthogonality of the molecular orbitals between the monomers.

(46) In practice, the electron densities of a given protein-ligand pair are prepared for ΔEDD plot computation by tricubic interpolation to superpose their shared protein atoms.

(47) Baum, B.; Muley, L.; Heine, A.; Smolinski, M.; Hangauer, D.; Klebe, G. Think Twice: Understanding the High Potency of Bis(phenyl)methane Inhibitors of Thrombin. *J. Mol. Biol.* **2009**, *391*, 552–564.

(48) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235–242.

(49) Labute, P. Protonate3D: Assignment of ionization states and hydrogen coordinates to macromolecular structures. *Proteins* **2009**, *75*, 187–205.

(50) Esrafili, M. D.; Behzadi, H. Investigation into the nature of interactions in aspirin–water clusters including SAPT, AIM and NBO theories. *Mol. Simul.* **2013**, *39*, 629–639.

(51) Thellamurege, N.; Hirao, H. Water Complexes of Cytochrome P450: Insights from Energy Decomposition Analysis. *Molecules* **2013**, *18*, 6782–6791.

(52) Merz, K. M. Using Quantum Mechanical Approaches to Study Biological Systems. *Acc. Chem. Res.* **2014**, *47*, 2804–2811.

(53) Raha, K.; Peters, M. B.; Wang, B.; Yu, N.; Wollacott, A. M.; Westerhoff, L. M.; Merz Jr., K. M. The role of quantum mechanics in structure-based drug design. *Drug Discovery Today* **2007**, *12*, 725–731.

Graphical TOC Entry