Zhang calculated the ligand/DNA/RNA interaction at high levels of theory.[88] Although further validation is needed for evaluating the ability of such a method to calculate binding free energies, it clearly has potential.

Fukuzawa et al. have used another approach – ab intio fragment molecular orbital (FMO) – to calculate the interaction energy of ligands that bind to the human estrogen receptor.[89] Although the agreement between the calculated and observed binding affinity is modest, they have examined the feasibility of modeling the receptor using only a few of the residues surrounding the ligand. They found no significant difference in the computed interaction energy between the complete receptor and the pruned receptor that had residues surrounding only the ligand. This hints toward a strategy to reduce the time taken for such calculations even further. However, a more thorough validation study is still needed.[2]

### Interaction energy decomposition with QM and QM/MM

Experimental measures of binding affinity give very little insight into the relationship of the binding pose of an active inhibitor and its interaction with the receptor. Such insights can be very useful for the process of going from a lead to a drug. Computational methods, in general, provide access to the decomposition of the interaction energy between the ligand and the receptor. However, with the application of QM to RBDD, these insights are more grounded theoretically and can often be validated by experiments. These insights can be utilized in design cycles comprising prediction and testing for increasing the potency of submicromolar leads in drug discovery.

Both QM/MM and linear-scaling QM methods have been used to dissect the interaction of a ligand with its receptor. Hensen et al. used MD and QM/MM to dissect the interaction of inhibitors bound to the HIV-1 protease.[64] They demonstrated that a 4-hydroxy-dihydropyrone substructure of the most potent inhibitor, tripnavir, made favorable interactions with the catalytic aspartates and isoluecine residues of the HIV-1 protease. He et al. have used the linear-scaling DM-MFCC approach to dissect the interaction between the HIV-1 reverse transcriptase (RT) and its drug-resistant mutants with the inhibitor nevirapine. The authors calculate a QM interaction spectrum that sheds light on crucial aspects of resistance to RT.[90]

Raha et al., using linear-scaling QM and a pairwise energy decomposition (PWD) scheme, dissected the interaction of a series of fluorine-substituted ligands [N-(4-sulfamylbenzoyl)benzylamine or SBB] with human carbonic anhydrase.[91] They divided the enzyme and inhibitors into subsystems and calculated the exchange energy that consisted of the off-diagonal elements of the density matrix and the one-electron matrix elements between subsystems:

$$E_{AB} = \sum_{\mu}^{A}\sum_{\nu}^{B} P_{\mu\nu}^{AB}\left(2H_{\mu\nu}^{AB} - \frac{1}{2}\sum_{\lambda}^{B}\sum_{\sigma}^{A} P_{\lambda\sigma}^{BA}\left(\mu^{A}\sigma^{A}\,|\,\lambda^{B}\nu^{B}\right)\right).$$

Here, $A$ and $B$ are residue subsystems, and $P$ and $H$ are the density matrix and the one-electron matrix, respectively. Using this PWD scheme, the authors investigated the effect of substitution of fluorines on the distal aromatic rings of SBB inhibitors and on its interaction with human carbonic anhydrase. The authors probed at the relationship of various pairwise interactions with the free energy of binding of the inhibitors. It was found that the substitution of fluorine at the distal group did not directly affect the free energy of binding. Rather, it geometrically influenced the strongest interaction between the sulfonamide group of the inhibitor and the Thr199 residue of the protein. This strong interaction, which was chemically identical in each of the inhibitors, was directly correlated with the binding affinity of the ligand. Such insights can be valuable in designing new and potent inhibitors. The PWD scheme was also incorporated into the comparative binding energy analysis (COMBINE)[92] methodology of Ortiz and coworkers to create SE-COMBINE by Peters and Merz.[93] This method elucidated the most important interactions between trypsin and a series of trypsin inhibitors. The multivariate statistical tools, principal component analysis and partial least squares (PLS), were used to mine the interactions between the receptor residues and the ligand fragments to generate QSAR models. The authors introduced so-called IMMs (intermolecular interaction maps), which enable the researcher to graphically view where a candidate drug could be modified or optimized.

## LIGAND-BASED DRUG DESIGN

One of the oldest tools used in rational drug design is QSAR. QSAR models are derived for a set of compounds with dependent variables (activity values, e.g., $K_i$, $IC_{50}$), and a set of calculated molecular properties or independent variables called descriptors. Each compound in the data set is assumed to be in its active conformation. Models are generated using techniques such as multiple linear regression (MLR), principal component regression (PCR), partial least squares regression (PLSR), and computer neural networks (CNNs) to name a few. Ligand-based methods can be further divided into two categories, 3D-QSAR and field-based methods. Both will be touched on under "3D-QSAR with QM Descriptors."

## 3D-QSAR WITH QM DESCRIPTORS

The descriptors used in 3D-QSAR are usually divided into three categories: (1) electronic [e.g., highest occupied molecular orbital (HOMO) and lowest unoccupied molecular orbital (LUMO) energies], (2) topological (e.g., connectivity indices), and (3) geometric (e.g., moment of inertia). The models in all cases are often created using multivariate statistical tools due to the large number and high degree of collinearity of descriptors. An excellent review by Karelson, Lobanov, and Katritzky provides details of QM-based