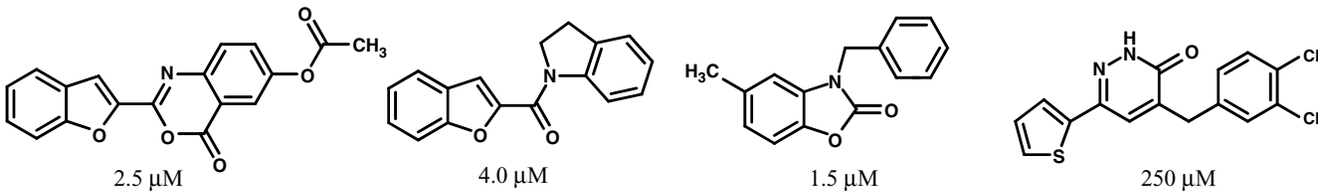

Figure 1.6. Structures and IC_{50} values for some inhibitors of MIF-CD74 binding discovered by virtual screening.

improvements with GLIDE 4.0 and the XP scoring, use of the large ZINC library, the relatively small binding site and consequently small number of rotatable bonds for potential inhibitors, and the human filtering.

ADME ANALYSES

As indicated in Figure 1.1, as one pursues leads it is important to be aware of potential pharmacological liabilities. The significance of this issue became increasingly apparent in the 1990s because of high failure rates for compounds in clinical trials that could be ascribed to ADME and toxicity problems.[31] This led to the introduction of Lipinski's rules and recognition that compounds developed in the post-HTS era frequently tended to be too large and hydrophobic, which is accompanied by solubility and bioavailability deficiencies.[32] In this atmosphere, more effort was placed on quantitative prediction of molecular properties beyond log $P_{o/w}$ using statistical procedures such as regression analyses and neural networks, which were trained on experimental data.[33,34] The typical regression equation is a linear one, Equation (1.1), where the sum is over molecular descriptors i that have values c_i for the given structure and the coefficients a_i are determined to minimize the error with the experimental data:

$$property = \sum_i a_i c_i + a_0. \qquad (1.1)$$

In Figure 1.1, the choice for ADME analyses is QIKPROP, which was among the earliest programs to predict a substantial array of pharmacologically relevant properties. Version 1.0, which was released in March 2000, provided predictions for intrinsic aqueous solubility, Caco-2 cell permeability, and hexadecane/gas, octanol/gas, water/gas, and octanol/water partition coefficients. The required input for QIKPROP is a three-dimensional structure of an organic molecule, and it mostly uses linear regression equations with molecular descriptors such as surface areas and hydrogen-bond donor and acceptor counts. By version 3.0 from 2006, the output covered eighteen quantities, including log BB for brain/blood partitioning, log K_{hsa} for serum albumin binding, hERG K^+ channel blockage, primary metabolites, and overall percentage human oral absorption.[15] The prediction of primary metabolites is based on literature precedents and recognition of corresponding substructures; for example, methyl ethers and tolyl methyl groups are typically metabolized to the alcohols. Execution time with QIKPROP is negligible because the most time-consuming computation is for the molecule's surface area. Average root-mean-square (rms) errors for most quantities are about 0.6 log unit, as in Figure 1.7.

To gauge acceptable ranges of predicted properties, QIKPROP 3.0 was used to process approximately 1,700 known neutral oral drugs,[13] which were compiled by Proudfoot.[35] For submission to QIKPROP, the original two-dimensional structures were converted to three-dimensional structures and energy-minimized with BOSS using the OPLS/CM1A force field.[10,16] Some key results from the analyses are summarized as histograms in Figures 1.8 and 1.9. Consistent with the log $P_{o/w}$ limit of 5 in Lipinski's rules,[32] 91% of oral drugs are found to have

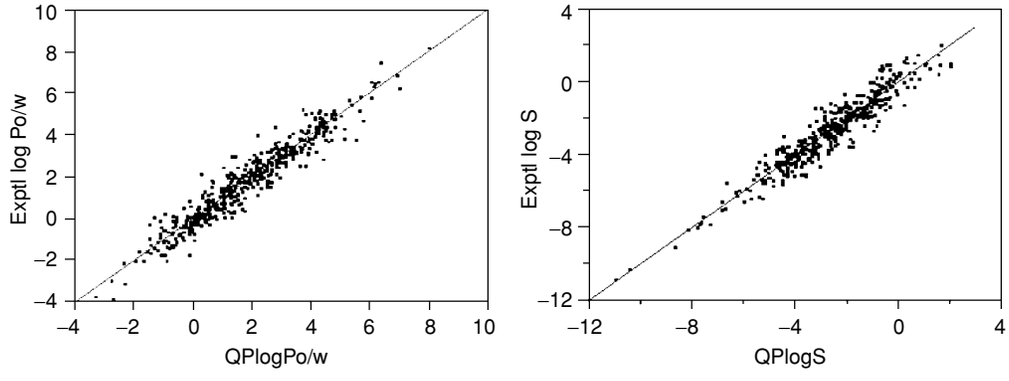

Figure 1.7. Experimental data and QikProp 3.0 results for 400–500 octanol/water partition coefficients (left) and aqueous solubilities (right). S is aqueous solubility in moles per liter. Correlation coefficients r^2 are 0.92 and 0.90 and the rms errors are 0.54 and 0.63 log unit, respectively.