Journal of Computer-Aided Molecular Design **17:** 383–397, 2003.

Binding of α-hydroxy-β-amino acid inhibitors to methionine aminopeptidase. The performance of two types of scoring functions

Anne Techau Jørgensen[‡], Morten Dahl Sørensen[§], Fredrik Björkling[§] & Tommy Liljefors[‡,*]
[‡] *Department of Medicinal Chemistry, The Danish University of Pharmaceutical Sciences, Universitetsparken 2, DK-2100 Copenhagen, Denmark;* [§] *Medicinal Chemistry Research, LEO Pharma, Industriparken 55, DK-2750 Ballerup, Denmark*

Received 11 December 2002; accepted in revised form 6 May 2003

Key words: scoring functions, CScore, Merck Molecular Force Field (MMFF), methionine aminopeptidase, angiogenesis, α-hydroxy-β-amino acid inhibitors, GRID, conformational energies, solvation

Summary

The binding mode of a recently described set of α-hydroxy-β-amino acid inhibitors of methionine aminopeptidase type 2 is suggested in the present work. The binding mode is supported by analysis of published structures of transition state analogues co-crystallised with *E. coli* methionine aminopeptidase and by a comparison of molecular interaction fields calculated using GRID for *E. coli* and human methionine aminopeptidase. Based on the suggested binding mode two types of scoring functions have been evaluated and compared. These are the commercially available consensus score, CScore, and scoring of the ligands employing energies calculated using the Merck Molecular Force Field (MMFF). Enriched subsets of ligands were obtained when using CScore, but these scores could not be used to assess the relative affinities of individual compounds. Although still not sufficiently accurate for reliable predictive purposes, an improved correlation was obtained between structure and affinity using a combined force field energy including contributions from solvation and conformational energy penalty for binding.

Abbreviations: MetAP–2: methionine aminopeptidase type 2; eMetAP-1: *E. coli* methionine aminopeptidase type 1; hMetAP-2: human methionine aminopeptidase type 2; MMFF: Merck Molecular Force Field.

Introduction

Angiogenesis, the outgrowth of new blood vessels, is an important process associated with the growth of tumours. Angiogenesis leads to an increased supply of nutrients and oxygen, and thus facilitates tumour growth as well as the incidence of metastasis [1, 2]. For this reason, there is currently an extensive search for anti-angiogenic drugs that are expected to inhibit the cancer process.

It has been shown that inhibition of the enzyme methionine aminopeptidase type 2 (MetAP–2) may play an important role in the angiogenic process by selectively blocking endothelial cell proliferation [3–7]. Furthermore, the potent angiogenesis inhibitor fumagillin irreversibly inhibits MetAP-2 by covalent binding to the active site [8]. These results suggest that MetAP-2 may be an interesting target for the treatment of undesired angiogenesis.

Methionine aminopeptidase occurs in two major isoforms in eukaryotes: type 1b and type 2b. By contrast, only one isoform of methionine aminopeptidase is found in bacteria and archaea. An additional N-terminal domain is the most significant difference between the eukaryote forms and the bacterial and archaeic forms, and the type 2 isoforms are distinguished by a single insert that forms a subdomain close to the active site. All isoforms share the same overall fold and function [9]. Experimental structures of three of the isoforms have been published: Human MetAP-2 (hMetAP-2), *E. Coli* MetAP-1 (eMetAP-1),

∗Author to whom correspondence should be addressed. E-mail: tl@dfh.dk

and *P. furiosus* MetAP-2 [10-12]. The active site of hMetAP-2 is located at the centre of a central β sheet surrounded by α-helices and shorter β sheets and includes a metal binding site that is strictly conserved among the methionine aminopeptidases [9, 10].

A series of α-hydroxy-β-amino acid compounds have been described, which exhibit affinities for MetAP-2 in the high nanomolar or low micromolar ranges. In contrast to fumagillin, these compounds are reversible inhibitors [13]. No experimental structure has been published for any of these inhibitors in complex with hMetAP-2, but a ligand exhibiting an α-hydroxy-β-amino acid moiety has been co-crystallised with eMetAP-1 [11]. In the present work, a binding mode to hMetAP-2 is suggested for the α-hydroxy-β-amino acid inhibitors based on the X-ray structure of this eMetAP-1 complex. The suggested binding mode is supported by comparison of molecular interaction fields calculated using the program GRID [14, 15] for eMetAP-1 and hMetAP-2.

Scoring functions are used to rank ligands, for instance when ligands are docked to proteins [16-18]. Previous studies suggest that a combination of scoring functions, i.e. generation of a consensus score, performs better than the use of a single one [19, 20]. We have compared consensus scoring with scoring employing force field energies, since carefully parameterised force field energy functions are expected to provide good estimates of the protein-ligand interactions and consequently form a sound basis for a scoring function. Moreover, docking followed by energy minimisation and reranking using a force field method have been shown to improve the predictive power compared to the scoring included in the docking procedure [21].

Based on these considerations we have evaluated and compared the performance of two types of scoring functions using a set of α-hydroxy-β-amino acid inhibitors (Figure 1). The commercially available consensus score, CScore implemented in Sybyl [22] is compared to scoring of the ligands employing energies calculated using the Merck Molecular Force Field (MMFF94s) [23–27] implemented in MacroModel [28]. Scoring functions are most often tested for their performance in virtual screening. Here the performance is tested on a data set of 34 inhibitors to allow thorough analysis of the data.

The evaluation and comparison are based on correlations between scores and experimental affinities [13] and on the fraction of nanomolar inhibitors identified in subsets selected using the scoring functions. Both parameters illustrate how well the scoring functions distinguish between the stronger and the weaker binders within a data set displaying a moderate spread of binding affinities. The ligands were scored in conformations generated by conformational analyses of the ligands within the active site of hMetAP-2.

Computational methods

GRID analyses

Molecular interaction fields of selected probes were calculated for the X-ray structures of hMetAP-2 (pdb code 1BOA [10]) and eMetAP-1 (pdb code 3MAT [11]) using GRID v. 16 [29] or 18 [30], respectively. The methyl probe was used to explore the steric characteristics of the site and the O, O1, N1, N2: and N3+ probes were used to explore the polar ones.

The protein structures were extracted from the Research Collaboratory for Structural Bioinformatics (RCSB) protein database [31]. A grid spacing of 0.25 Å was used in the GRID calculations for hMetAP-2, and the following adjustments were made to the X-ray structure. The incomplete residues were renamed to glycines with the deletion of any extra atoms. When two positions were listed in the pdb-file, position A was selected for the cysteines C228, C290 and C448. To avoid disturbance from the incomplete backbone, the nitrogen of E154 was treated as an amide nitrogen. The GRID calculations for eMetAP-1 were performed with a grid spacing of 0.33 Å. The incomplete residues in this structure were renamed to alanines and any extra atoms were deleted. Hydrogens were added to both X-ray structures using the GRIN module of GRID, and the cobalt ions were treated as belonging to the protein. Default values were used for all directives, except for the grid spacing.

Force field and energy minimisation routines

All energy minimisations and conformational analyses were performed using the Merck Molecular Force Field (MMFF94s) [23–27] as implemented in MacroModel v. 6.5 [28]. The MMFF94s force field was chosen because it has shown to be among the best at reproducing experimental data [23–27, 32]. Because no cobalt parameters are included in the MMFF94s force field, the cobalts of the metal-binding site were changed to zincs. Previous quantum chemical calculations for a model system of the methionine aminopeptidase metal-binding site have shown that

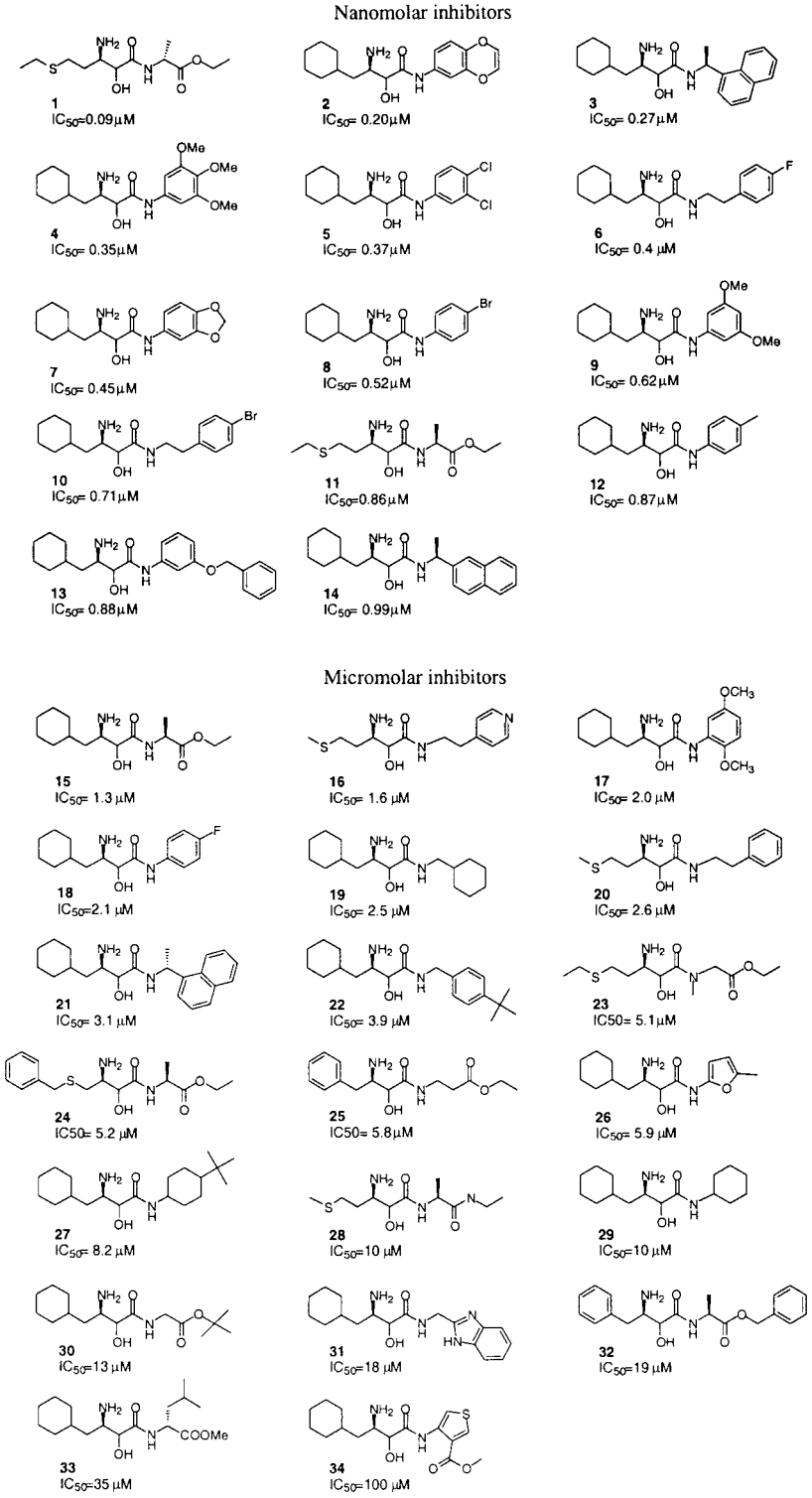


Figure 1. The α-hydroxy-β-amino acid inhibitors included in this study together with their affinities.

cobalts and zincs produce very similar site geometries [33].

Minimisations including the protein were performed with the Polak-Ribier Conjugate Gradient (PRCG) routine, whereas minimisations of small molecule ligands were performed with the truncated Newton-Raphson conjugate gradient algorithm (TNCG) algorithm.

The protein structure

The conformational analyses and scorings were performed using the published X-ray structure of hMetAP-2 pdb [31] code 1BOA [10] after addition of hydrogens and optimisation of hydrogen and water positions.

Prior to geometry optimisation, hydrogens were added to the X-ray structure using Insight II v. 98.0 [34] with pH set to 7.4. The following waters: 527, 528, 556, 586, 625, 643 and 662 were deleted due to possible overlap with ligands. Histidines were treated as Nτ-tautomers. Exceptions were H231 and H331, which were treated as Nπ-tautomers to obtain the optimal interactions with the ligands. The disulphide bridges were completed by addition of the necessary bonds. All residues lacking atoms were completed. The terminal residues were modelled as charged.

A two-step procedure was employed when optimising waters and hydrogens. Firstly, the positions of all hydrogens were optimised. All other atoms were frozen at their X-ray positions. Secondly, the positions of all water oxygens were relaxed. The water oxygens were tethered to their X-ray positions by harmonic flat-bottomed Cartesian constraints. All hydrogens remained free and all other atoms were frozen. The Cartesian constraints were centred at the respective atomic X-ray positions, the flat-bottomed half-radius was set to 0.3 Å and the force-constant applied beyond the half-radius was set to 400 kJ/mol$*$Å^2.

Selection of ligands

A subset of 34 compounds was selected among the published α-hydroxy-β-amino acid inhibitors [13] that covered the full span of affinities. The affinities of the selected compounds range from 0.09 μM to 100 μM (Figure 1). Subset selection was further based on the following criteria: The compounds had to be reasonably rigid, neutral and of well-defined stereochemisty. Compounds containing nitro-groups were also rejected due to possible force field problems.

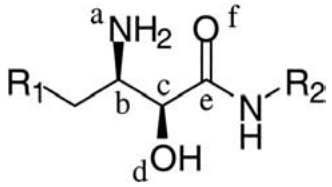
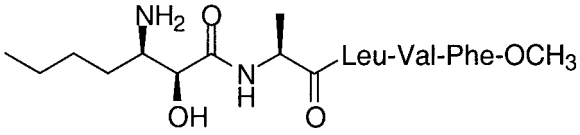


Figure 2. Illustration of the α-hydroxy-β-amino acid moiety. The torsions constrained in the conformational searches were: a-b-c-d and d-c-e-f.

Binding mode of the α-hydroxy-β-amino acid-moiety

The α-hydroxy-β-amino acid inhibitors were assumed to bind to hMetAP-2 in a binding mode analogous to that of compound **35** to eMetAP-1 (pdb code 3MAT) [11]. Analyses indicated that this binding mode was transferable to hMetAP-2 (see Results and discussion). Thus, this experimentally determined binding mode was preferred over binding modes generated by other means as for instance automated docking. The experimental structures of eMetAP-1 including bound **35** and hMetAP-2 were superimposed using the two metal ions of the active site and the five residues coordinating to them (hMetAP-2 D251, D262, H331, E364 and E459). The α-hydroxy-β-amino acid ligands were added to the hMetAP-2 protein structure by superimposing the heteroatoms of the amino acid portion onto the corresponding heteroatoms of the eMetAP-1 inhibitor **35**.



35

Conformational analysis of ligands within the active site

The conformation of the α-hydroxy-β-amino acid-moiety of **1**–**34** was assumed to be similar to that of **35**, as described above. Thus, conformational analyses were only performed for the side chains R1 and R2 (Figure 2).

The low-energy conformational spaces of the inhibitors were searched employing a two step search procedure. (1) Side chain R1 (Figure 2) occupies a narrow subsite in the protein and has a very limited conformational space. For the efficiency of the procedure and because of the limited conformational space it was decided to predetermine the conformation of

the R1 side chain and then keep it fixed during the conformational analysis of the R2 side chain. Five different R1 side chains are included in the set of inhibitors represented by compounds **11**, **15**, **24**, **25** and **28** (Figure 1). Conformational analyses of the R1 side chains of these compounds were performed within the active site of hMetAP-2. The conformation exhibiting the lowest complex energy was selected for each compound. Cluster analyses verified that the selected conformation was well represented among the low energy conformations. (2) The conformational space of the R2 side chain of each ligand was analysed, while the R1 side chain exhibiting the selected conformation was just energy minimised and not included in the conformational search.

Low energy conformations were sought using the systematic pseudo Monte Carlo search method (SPMC) [35]. During conformational analyses, the conformations were discharged without further energy minimisation if the initial total energy were more than 40 kJ/mol above the current global minimum, and they were saved if the energy after minimisation were less than 20 kJ/mol above the global minimum. Testing for identical conformations involved comparing the positions of all heavy atoms of the ligand. The positions of the amino and hydroxy hydrogens influenced the calculated internal energy. Consequently, these hydrogens were positioned prior to the conformational analyses in order to avoid internal hydrogen bonds and to simulate the co-ordination to the metal ions. All calculations were performed employing a dielectric constant of 1.

The conformational space was searched within a fully frozen site of hMetAP-2; i.e. all atoms in residues and waters within 16 Å of the ligand were frozen. Residues outside the 16 Å sphere were not included in the calculations. All atoms of the ligands were free to move with a few torsional constraints. The torsions of the amino acid moiety were constrained to $\pm$ 15 degrees with respect to their corresponding dihedral angles as determined in the conformational analysis of the relevant R1 side chain (torsions a-b-c-d and d-c-e-f in Figure 2). The force-constant applied was 9999 kJ/mol$*$Å^2. Side chain esters and amides were constrained to their lowest-energy torsional angles. No positional constraints were applied to the ligand atoms because test calculations showed insignificant differences in the positions of ligands generated with and without constraints (results not shown).

Conformational analysis of free ligands in aqueous solution

Low energy conformations of the neutral form of the ligands were sought using the Monte Carlo multiple minimum method (MCMM) [36]. During conformational analyses, the conformations were discharged without further energy minimisation if the initial total energy were more than 40 kJ/mol above the current global minimum, and they were saved if the energy after minimisation were less than 20 kJ/mol above the global minimum. The search was continued until the lowest energy conformations were found multiple times. The test for identical conformations included all heavy atoms as well as the hydroxyl and amino hydrogens of the amino acid moiety. Conformational analysis was performed including aqueous solvation *via* the generalised Born/solvent accessible surface area (GB/SA) continuum dielectric solvation model [37] as implemented in MacroModel. The following non-default cut-off values for non-bonded interactions were employed: 30 Å for charge–charge interactions, and 30 Å for van der Waals interactions. The conformations generated by the conformational analysis were further minimised using TNCG until all minimisations had converged according to the default convergence criteria. Duplicate structures resulting from the energy minimisation were deleted by the multi-conformer minimisation option provided by MacroModel.

Consensus score – CScore

The ligands were ranked using CScore as implemented in Sybyl v. 6.8 [22, 38]. CScore includes five scoring functions in this implementation: F_score (the FlexX scoring function [39]); G_score (based on the Gold scoring [40]); D_score (based on Dock scoring function [41]); PMF_score (based on the work of Muegge and Martin [42]); and ChemScore (based on the work of Eldridge *et al.* [43]). For each scoring function the better scored ligands are assigned a 1, while the rest is assigned a zero. CScore is the sum of the assignments over the scoring functions, and CScore will thus be an integer between 0 and 5. The protein coordinates obtained as described above were imported into Sybyl. MMFF94 charges were added, and atom and bond types for all histidines were corrected.

Two parallel analyses using CScore were performed based on two different sets of ligand conformations obtained from the conformational analyses of the ligands within the site as described above. In the

first CScore analysis, $CScore_{energy}$, the ligand conformation was obtained from the complex exhibiting the lowest complex energy. In the second analysis, the ligands were scored in conformations selected by using CScore. This procedure is denoted $CScore_{selected}$. The conformation used was the one that exhibited the highest CScore among all conformations generated in the conformational analysis of the ligand within the site. The consensus score used to select the conformation of the ligand was based on the default scoring functions as well as on the complex energy ($E_{complex}$) calculated during the conformational analyses. If more than one conformation achieved the same highest consensus score, the conformation with the lowest $E_{complex}$ was used in the scoring of the ligands.

Force field based scoring

The binding of a ligand to a protein includes desolvation, formation of intermolecular interactions with the protein and most often conformational changes of the ligand and the protein. This complex set of events can be represented by a sum of discrete components obtained from a thermodynamic cycle [44]. Accordingly, the overall free energy of the binding process, ΔG, can be calculated as a combination of these discrete components (Equation 1).

$$\begin{aligned}\Delta G = {} & \Delta G_{interact} + \Delta G_{solv}(complex) \\ & - \Delta G_{solv}(ligand) - \Delta G_{solv}(protein) \\ & + \Delta G_{conf}(ligand) + \Delta G_{conf}(protein)\end{aligned} \quad (1)$$

Based on this equation the α-hydroxy-β-amino acid inhibitors were scored using three combined energies related to the overall free energy of binding:

1) the interaction free energy, $\Delta G_{interact}$
2) the first combined energy, ΔG_1, taken as the combination of $\Delta G_{interact}$ and the conformational energy penalty of the ligand, ΔG_{conf}. $\Delta G_1 = \Delta G_{interact} + \Delta G_{conf}$
3) the second combined energy, ΔG_2, taken as the combination of ΔG_1 and the solvation free energy of the ligand, ΔG_{solv}. $\Delta G_2 = \Delta G_1 - \Delta G_{solv}$

The interaction free energy, $\Delta G_{interact}$, gives the strength of the interaction between ligand and protein, and was estimated by subtracting the calculated internal energy of the ligand in its calculated bioactive conformation, $E_{int,bound}$, from the energy of the complex, $E_{complex}$ (Equation 2). The steric energy of the protein is not included in $E_{complex}$, because the protein was frozen and the ligand was defined as a substructure during the conformational search. We assume that the entropic contribution to $\Delta G_{interact}$ is similar for the compounds studied. Since were are only interested in differences in $\Delta G_{interact}$ we have for simplicity omitted this assumed constant term in Equation 2.

$$\Delta G_{interact} \approx E_{complex} - E_{int,bound} \quad (2)$$

The conformational energy penalty of the ligand, ΔG_{conf}, estimates the conformational energy required for a ligand to adopt the bioactive conformation, and was calculated by subtracting the internal energy of the calculated global energy minimum conformation of the ligand in solution, $E_{int,free}$, from the internal energy of the calculated bioactive conformation of the ligand, $E_{int,bound}$ (Equation 3, [44]). The internal energy of the ligand in solution, $E_{int,free}$, was calculated by subtracting ΔG_{solv} from the total energy of the calculated global energy minimum conformation of the ligand in solution, $E_{tot,free}$.

$$\Delta G_{conf} \approx E_{int,bound} - E_{int,free} \quad (3)$$

The solvation free energy, ΔG_{solv}, was calculated for the neutral form of the ligand, and the protonation energy was not taken into account for any of the ligands. Moreover, the effect of protonation on the conformational ensemble of the ligands is neglected as well. The calculations were performed on the neutral form of the ligands, because the compounds bind in this form and because force field calculated energies of charged and neutral species are incomparable.

The complex energy, $E_{complex}$, used was that of the complex exhibiting the lowest total energy in the conformational analysis of the ligand within the site as described above. The energy of the ligand in the calculated bioactive conformation, $E_{int,bound}$, was calculated *in vacuo* for the ligand extracted from the complex used to calculate $E_{complex}$. Accordingly, $E_{complex}$ and $E_{int,bound}$ were calculated based on the complex from which the conformation used in the $CScore_{energy}$ was derived as well. The total energy, $E_{tot,free}$, and the solvation energy, ΔG_{solv}, were calculated for the ligand conformation exhibiting the lowest total energy in the conformational analysis of the free ligand in aqueous solution as described above. In these calculations, as well as in the calculation of the conformational energy penalty, ΔG_{conf}, one single

conformation, the global energy minimum conformation, was assumed to represent the whole aqueous ensemble, and consequently the conformational entropy effect was not taken into account.

Results and discussion

Suggested binding mode for α-hydroxy-β-amino acid inhibitors

As mentioned in the introduction, no experimentally determined binding mode of α-hydroxy-β-amino acid inhibitors to hMetAP-2 has been reported. However, an X-ray structure has been published of eMetAP-1 co-crystallised with compound **35** [11]. This ligand binds to the active site of eMetAP-1 via co-ordination between its α-hydroxy-β-amino acid moiety and the cobalt ions of the site.

Lowther *et al.* have determined experimental structures of eMetAP-1 co-crystallised with a number of phosphorus-based transition state analogues [45]. The analogues bind to the active site via co-ordination between the metal ions and the α-amino phosphonate, α-amino phosphinate or α-amino carboxylate groups of the ligands. Only small differences in co-ordination distances of the ligands were identified (Figure 3). This indicates that the co-ordination pattern is consistent and it is most likely valid for α-hydroxy-β-amino acid inhibitors in general.

Transfer of the binding mode of the α-hydroxy-β-amino acid moiety of **35** from eMetAP-1 to hMetAP-2 is supported by comparisons of the molecular interaction fields calculated using GRID for the two proteins. The interactions elucidated by the hydrophobic methyl probe as well as by the more specific polar probes O1 and O illustrate a significant similarity between the active sites of the two proteins (Figure 4). The calculated molecular interaction fields indicate that the energetically favourable interaction sites near the metal ions are very similar for eMetAP-2 and hMetAP-2, while the regions occupied by the side chains of **35** are more deviating. It should be noted that the GRID force field is not thoroughly tested for ligand-metal interactions [46]. However, since the favourable interaction sites are defined by the surrounding protein residues as well as by the metal ions, the analyses are still considered to be valid. Consequently, it is assumed that the α-hydroxy-β-amino acid moiety of the inhibitors bind to hMetAP-2 in such a way that the metal co-ordination distances are similar to those determined for **35** when co-crystallised with eMetAP-1. The positions of the side chains were then determined by conformational analyses within the site as described in the Computational methods section.

Analysis of the calculated bioactive conformations of compounds **1**–**34** revealed that their R1 side chains are positioned in a pocket defined by F219, H231, I338, H382, M384, A414 and Y444 (Figure 5). The R2 side chains protrude into the aqueous environment while mainly exhibiting hydrophobic interactions with the protein, most significantly with L328, N329, H331, H339, T343 and L447 (Figure 5).

CScore – Correlation to experimental affinities

CScore analyses ($CScore_{selected}$ and $CScore_{energy}$) were performed for compounds **1**–**34** based on ligand conformations obtained by conformational analyses within the active site. In both analyses, the correlation with experimental affinities was generally poor for the five included scoring functions as well as for CScore (Figure 6 and Figure 7). Compounds including an alkylthioether in the R1 side chain (**1**, **11**, **16**, **20**, **23** and **28**) were assigned lower scores than would be expected based on their experimental affinities by F_score, ChemScore and in some cases PMF_score. A possible explanation for the systematic error is difficulties estimating the hydrophobic interactions of the sulphur atom. When considering the underestimation of the alkylthioethers and defining compounds **31** and **34** as outliers a weak correlation is found between F_score and the experimental affinities.

When CScore was used to select the conformation of the ligand ($CScore_{selected}$), another conformation was selected for 18 of the ligands than the one obtained from the complex exhibiting the lowest complex energy. Thus, 16 of the ligands were scored in the same conformation in the two analyses, $CScore_{energy}$ and $CScore_{selected,}$ whereas a conformation with a higher $E_{complex}$ was selected for 18 of the ligands. This illustrates a discrepancy between the ranking of conformations based on CScore and that based on the energy of the complex as calculated by MMFF94s.

CScore – Number of nanomolar inhibitors in ligand subsets

Based on the poor correlation between the calculated scores and the experimental affinities, the scoring would not be expected to distinguish stronger binders from weaker binders. Nonetheless, 5 nanomolar inhibitors were among the 8 compounds scored to 4

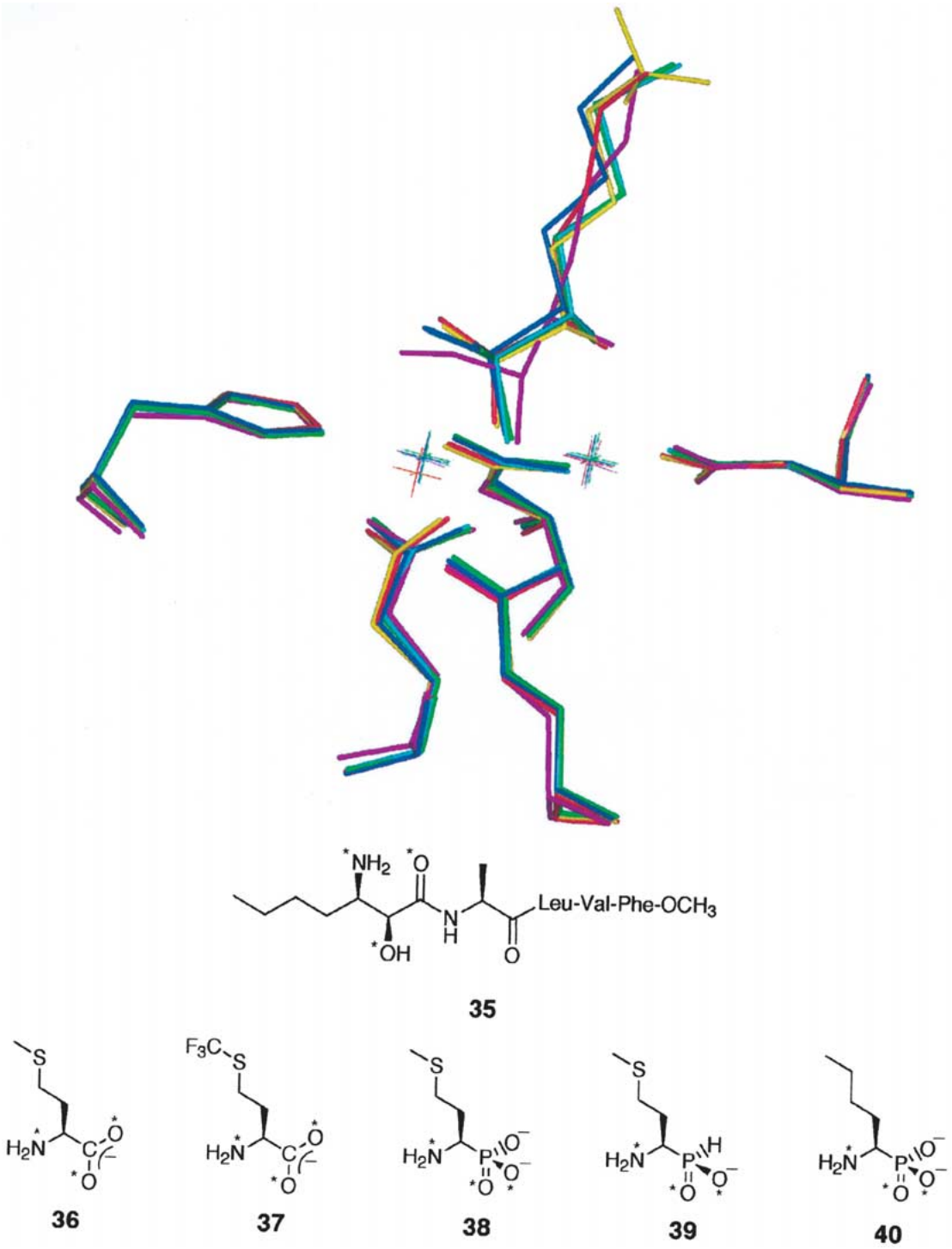


Figure 3. Superposition of the metal ions and various ligands illustrating the consistency of the co-ordination patterns. Compound **35** co-crystallised with eMetAP is shown in magenta. The X-ray structures of the phosphorus-based transition state analogues are shown in the following colours: **36** (pdb-entry 1C21) in red; **37** (pdb-entry 1C22) in yellow; **38** (pdb-entry 1C23) in green; **39** (pdb-entry 1C24) in cyan; and **40** (pdb-entry 1C27) in blue. All structures have been determined by Lowther *et al.* [45] Co-ordinating atoms are indicated with asterisks in the compound structures.

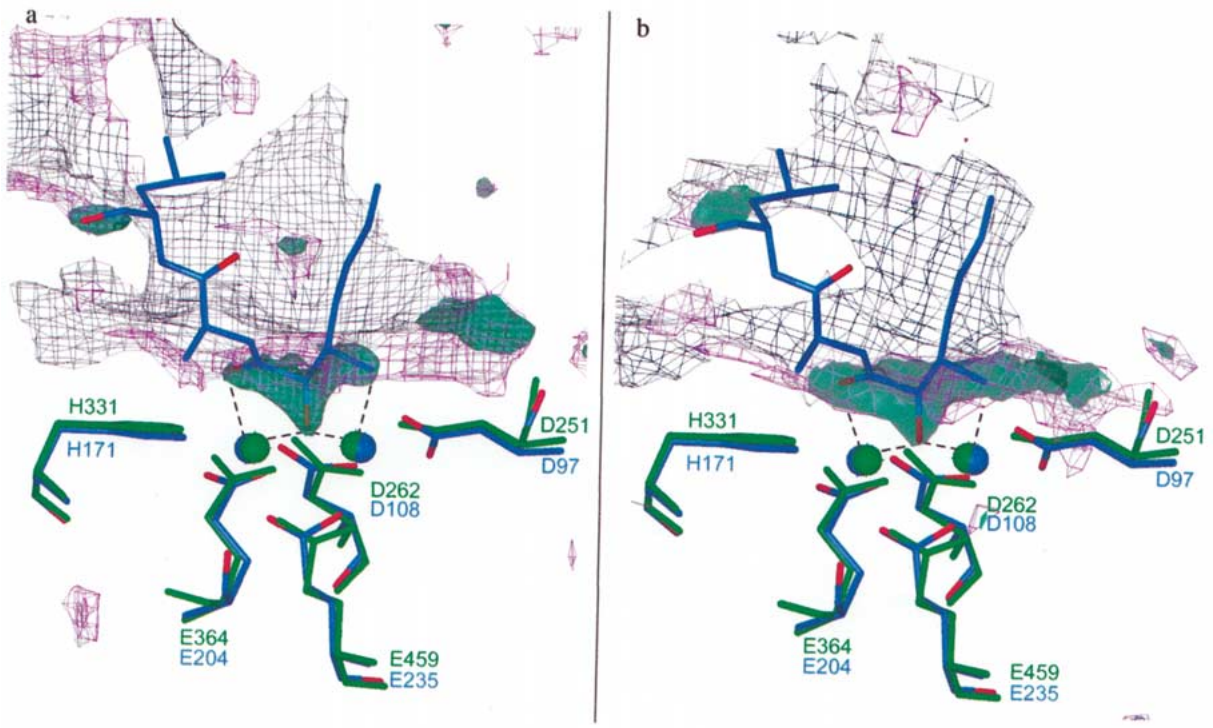


Figure 4. Comparison of the molecular interaction fields calculated for hMetAP-2 and eMetAP-1. The −2 kcal/mol isocontour plot of the methyl GRID probe is shown in grey, the −7 kcal/mol isocontour of the hydroxyl probe (O1) is shown as magenta grid and that of the carbonyl probe (O) is shown in transparent green. A superposition of the five metal-coordinating residues together with the two cobalt ions of eMetAP-1 (blue) and hMetAP-2 (green) is included as well as the experimentally determined part of ligand **35** from the X-ray structure of eMetAP-1. **a** The molecular interaction fields calculated for hMetAP-2. **b** The molecular interaction fields calculated for eMetAP-1.

or better by $CScore_{selected}$, whereas 8 nanomolar inhibitors were identified among the 18 compounds exhibiting a CScore of at least 4 scored by $CScore_{energy}$ (Figure 8). This should be compared to the number of nanomolar inhibitors, 14, included in the full data set of 34 compounds. These results indicate that the choice of ligand conformation is important for the performance of CScore. In the present case, the best performance was obtained when selecting ligand conformations using the CScore procedure.

Force field based scoring of ligands – Correlation to experimental affinities

The ligands were scored using three energies $\Delta G_{interact}$, ΔG_1 and ΔG_2 obtained using MMFF94s-GB/SA calculations as described in the Computational methods section. As shown in Figure 9, ΔG_2 displays the strongest correlation with the experimental affinities of the three energies calculated. This indicates that the conformational energy penalty and especially the solvation energies are important terms in a force field based scoring function. The fact that the correlation improves by addition of the conformational energy penalty and the solvation energy is conceptually appealing, when considering the physical process of ligand binding that includes adoption of the bioactive conformation by the ligand as well as desolvation of the ligand (and the protein).

Force field based scoring of ligands – Analysis of outliers

Three outliers, compounds **11**, **16** and **23**, were identified in the correlation between ΔG_2 and the experimental affinities (Figure 9). The energies and conformations calculated for the three inhibitors were analysed in more detail to try to explain their unexpectedly low ΔG_2. The complex of **16** with the protein was compared to that of the closely related **20**. The only difference between these two compounds is the pyridine ring of **16** that is replaced by a phenyl ring in compound **20**. The affinities are 2–3 μM for both compounds and it is thus interesting that the differ-

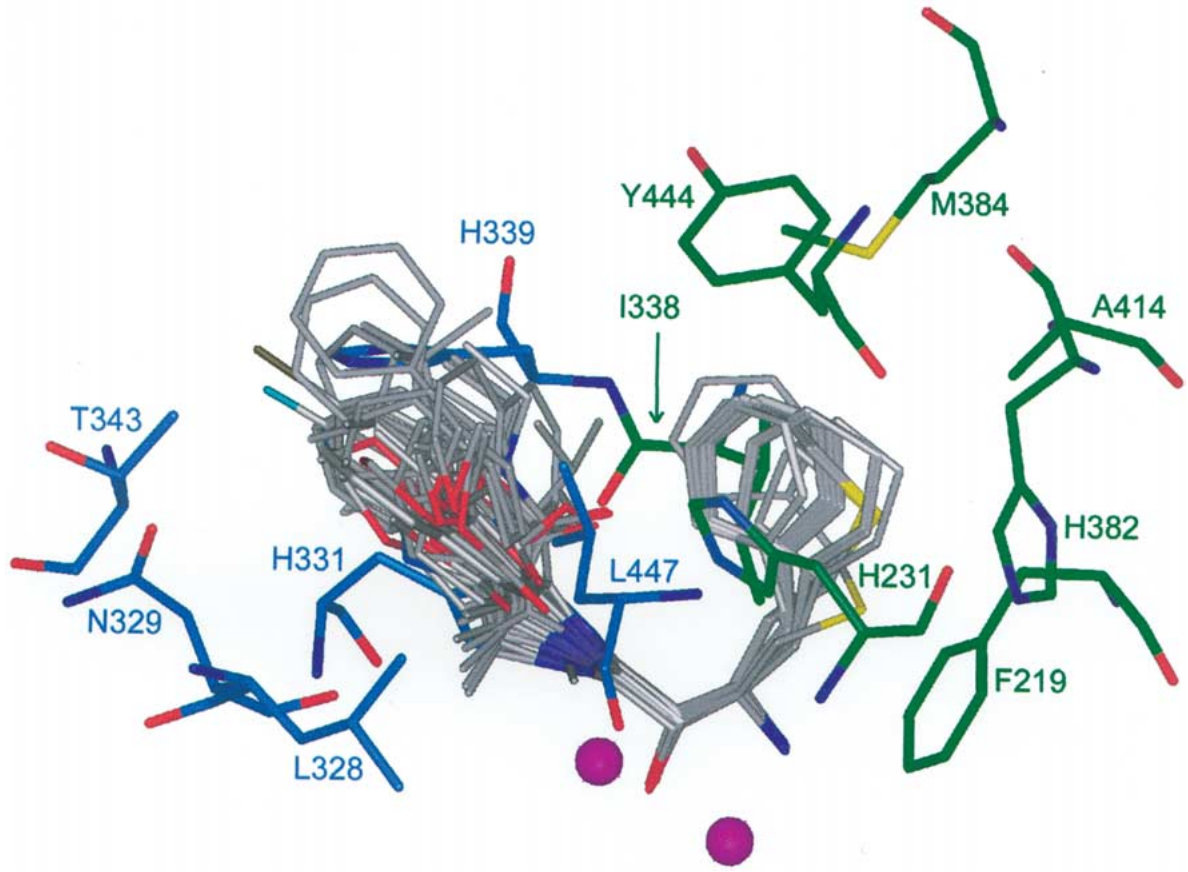


Figure 5. Illustration of the binding mode of the scored ligands in their calculated bioactive conformations. The carbons of residues interacting with the R1 side chains are shown in green and the carbons of residues interacting with the R2 side chains are shown in blue. The metal ions are shown as magenta spheres.

ence in ΔG_2 is as high as 73 kJ/mol. Analysis of the energies showed that the majority of the energy difference is of electrostatic origin, and analysis of the complexes revealed that the pyridine nitrogen of **16** accepts a hydrogen bond from the residue H339. However, the geometry of this hydrogen bond is non-optimal; the N-H-N angle is 131 degrees and the N-N distance is 3.17 Å. Additional calculations were performed on model systems to evaluate the contribution from this hydrogen bond. The model systems consisted of an imidazole ring and either a pyridine or a benzene ring in positions according to those of **16** or **20** and H339 in the ligand-protein complex. The calculations indicate that the difference in interaction energy between a pyridine and a benzene ring interacting with an imidazole ring is in the same range as the solvation energy difference. The difference in solvation energy ($\Delta\Delta G_{solv}$) for the two compounds is −25.2 kJ/mol when calculated using MMFF94s and GB/SA. This is in the same range as the difference in interaction energy ($\Delta\Delta G_{interact}$) that is −25.9 kJ/mol when calculated using MMFF94s and −16.2 kJ/mol when calculated as single point energies at the B3LYP/6-31G∗+//B3LYP/6-31G∗+ level. Thus, it is most likely that the affinity-gain obtained from the hydrogen bond is counteracted by the increased desolvation term for **16**. The difference in the calculated interaction energies for **16** and **20** is therefore probably not due to the presence of a hydrogen bond between **16** and H339. Most likely, it is due to long-range electrostatic interactions differences resulting from the different assignment of partial charges by the force field to the pyridinyl and phenyl groups of **16** and **20**, respectively.

A closer investigation of compounds **11** and **23** revealed no significant or systematic differences in the positions of the ligands or their energy contributions, and thus no explanation was found for their low scor-

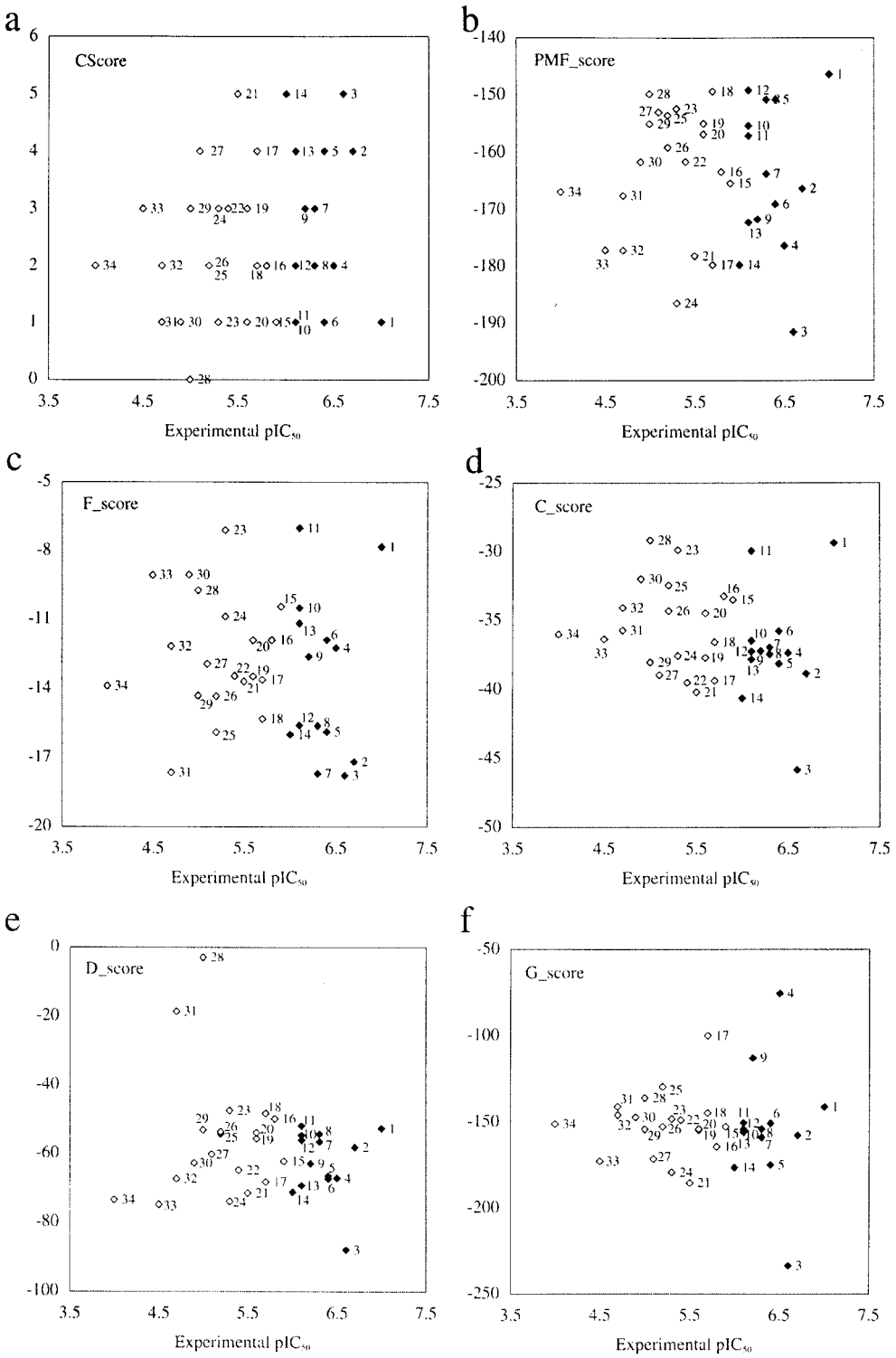


Figure 6. The scoring values calculated for each ligand in the $CScore_{selected}$ analysis plotted against the experimental affinities (pIC_{50}). Nanomolar inhibitors are represented by filled diamonds and micromolar inhibitors are represented by unfilled diamonds. **a**. CScore values. **b**. PMF_score values. **c**. F_score values. **d**. ChemScore values. **e**. D_score values. **f.** G_score values.

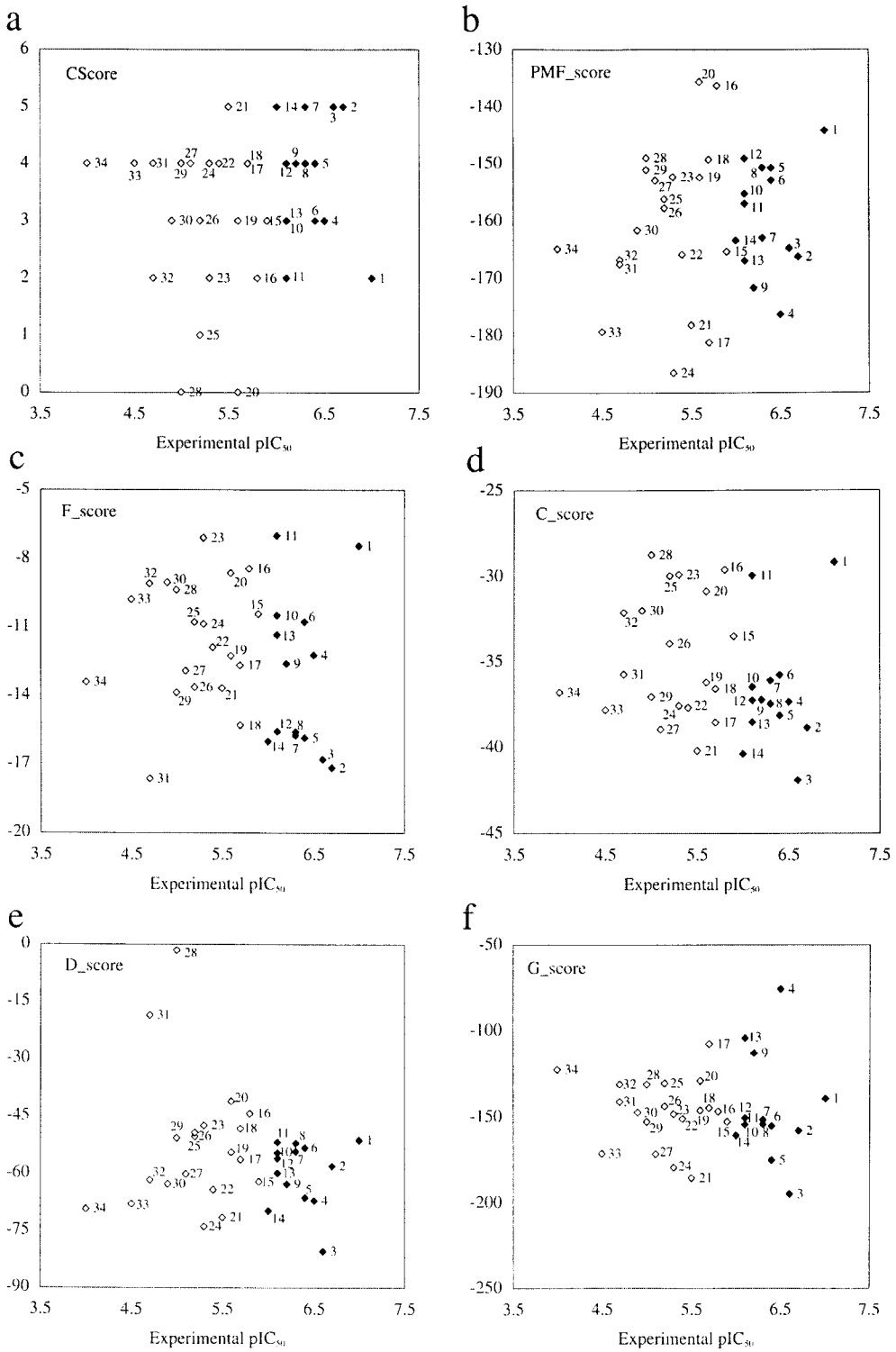


Figure 7. The scoring values calculated for each ligand in the $CScore_{energy}$ analysis plotted against the experimental affinities (pIC_{50}). Nanomolar inhibitors are represented by filled diamonds and micromolar inhibitors are represented by unfilled diamonds. **a**. CScore values. **b**. PMF_score values. **c**. F_score values. **d**. ChemScore values. **e**. D_score values. **f.** G_score values.

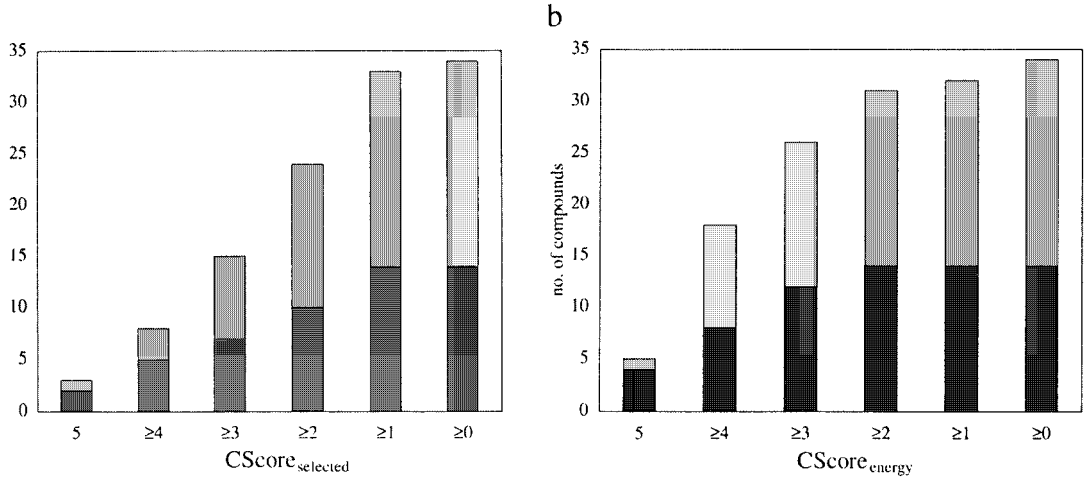


Figure 8. Fraction of nanomolar inhibitors in ligand subsets obtained when ranking the ligands using CScore. The dark grey portion of the bar represents the nanomolar inhibitors and the light grey portion represents the micromolar inhibitors. **a**. The result of $CScore_{selected}$ **b**. The result of $CScore_{energy}$.

ing energies. Both compounds possess a thioether as their R1 side chain and were scored as weaker binders by CScore. However, compounds **11** and **23** were scored as *stronger* binders by the force field based scoring and three of the six thioether compounds (**1**, **11**, **16**, **20**, **23** and **28**) were not outliers. These facts contradict a general trend.

Force field based scoring of ligands – Number of nanomolar inhibitors in ligand subsets

When the 7 compounds exhibiting the most negative ΔG_2 energies were selected, 4 of these were found to be nanomolar inhibitors, and the 14 compounds displaying the most negative ΔG_2 energies included 8 (57%) nanomolar inhibitors (Figure 10). This is a slight increase compared to the percentage of nanomolar inhibitors in the data set (41%).

Conclusions

In this evaluation, both CScore and the MMFF94s-based scoring function showed to be useful in the scoring of the compounds. Slightly enriched subsets of ligands were obtained when using CScore, but the scores could not be used to assess the relative affinities of individual compounds. When force field based scoring includes the conformational energy penalty and solvation energy it leads to a better correlation between predicted and experimental affinities. This is conceptually appealing because it is consistent with the physical process of ligand binding. Nonetheless, outliers were identified and the correlation must be improved if the method is to be used for prediction of binding affinities.

Acknowledgements

This work was financially supported by Graduate School of Drug Research and LEO Pharma, which is gratefully acknowledged.

References

1. Folkman, J., Nat. Med., 1 (1995) 27.
2. Folkman, J., N. Eng. J. Med., 333 (1995) 1757.
3. Griffith, E.C., Su, Z., Turk, B.E., Chen, S., Chang, Y.-H., Wu, Z., Biemann, K. and Liu, J.O., Chem. Biol., 4 (1997) 461.
4. Turk, B.E., Griffith, E.C., Wolf, S., Biemann, K., Chang, Y.H. and Liu, J.O., Chem. Biol., 6 (1999) 823.
5. Wang, J.Y., Lou, P.P. and Henkin, J., J. Cell. Biochem., 77 (2000) 465.
6. Kusaka, M., Sudo, K., Matsutani, E., Kozai, Y., Marui, S., Fujita, T., Ingber, D. and Folkman, J., Br. J. Cancer, 69 (1994) 212.
7. Ingber, D., Fujita, T., Kishimoto, S., Sudo, K., Kanamaru, T., Brem, H. and Folkman, J., Nature, 348 (1990) 555.
8. Griffith, E.C., Su, Z., Niwayama, S., Ramsay, C.A., Chang, Y.-H. and Liu, J.O., Proc. Natl. Acad. Sci. USA, 95 (1998) 15183.
9. Bradshaw, R.A., Brickey, W.W. and Walker, K.W., Trends Biochem. Sci., 23 (1998) 263.
10. Liu, S., Widom, J., Kemp, C.W., Crews, C.M. and Clardy, J., Science, 282 (1998) 1324.
11. Lowther, W.T., Orville, A.M., Madden, D.T., Lim, S., Rich, D.H. and Matthews, B.W., Biochemistry, 38 (1999) 7678.

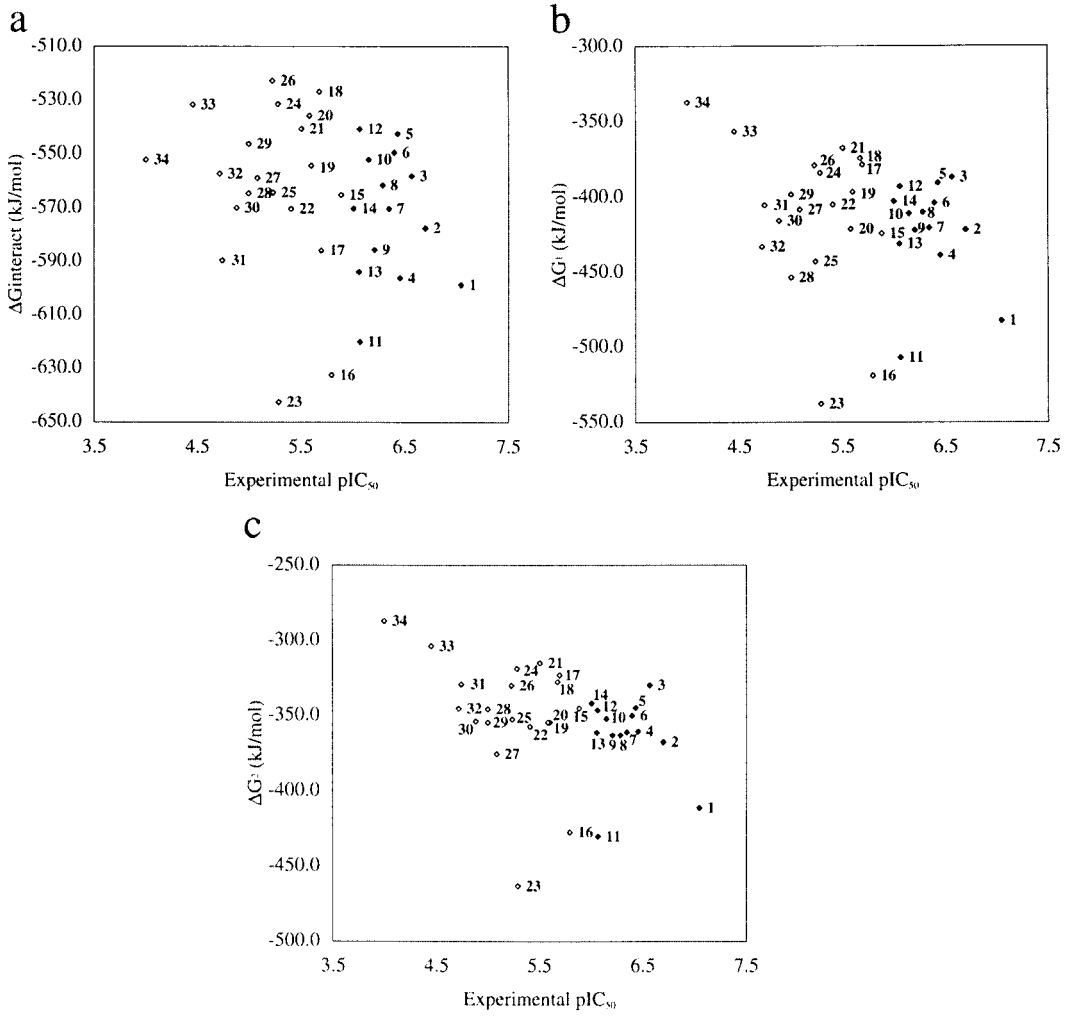


Figure 9. The calculated energies plotted against the experimental affinities, pIC_{50}. Nanomolar inhibitors are represented by filled squares and micromolar inhibitors by unfilled squares. **a**. The interaction energy, $\Delta G_{interact}$. **b.** The combination, ΔG_1, of $\Delta G_{interact}$ and ΔG_{conf}. **c.** The combination, ΔG_2, of $\Delta G_{interact}$, ΔG_{conf} and ΔG_{solv}.

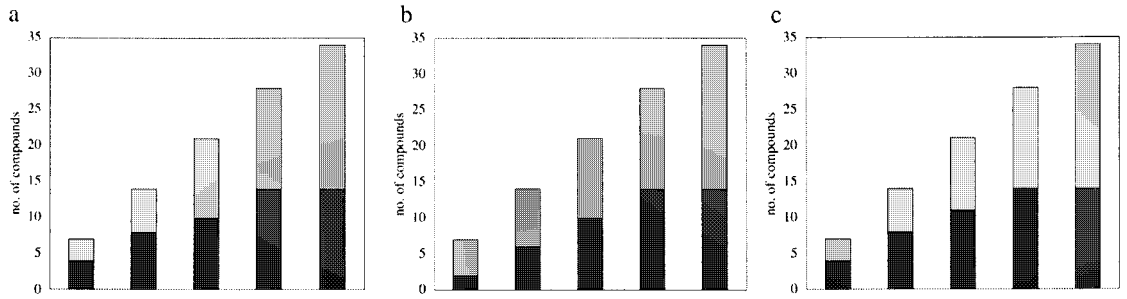


Figure 10. Fraction of nanomolar inhibitors in ligand subsets obtained when ranking the ligands using on the force field based scoring. The dark grey portion of the bar is the number of nanomolar inhibitors and the light grey portion is the number micromolar inhibitors. **a**, **b** and **c**. The distribution obtained when selecting groups of 7 compounds (the last group only includes 6 compounds) based on the interaction energy, $\Delta G_{interact}$, ΔG_1 and ΔG_2, respectively.

12. Tahirov, T.H., Oki, H., Tsukihara, T., Ogasahara, K., Yutani, K., Ogata, K., Izu, Y., Tsunasawa, S. and Kato, I., J. Mol. Biol., 284 (1998) 101.
13. Abbott Laboratories. Substituted beta-amino acid inhibitors of methionine aminopeptidase-2. Int. Pat. Appl. WO 99/57098, 1999.
14. Goodford, P.J., J. Med. Chem., 28 (1985) 849.
15. GRID v. 20, Molecular Discovery Ltd., 4 Chandos Street, London, England, 2001-2002.
16. Böhm, H.-J. and Stahl, M. In K. B. Lipkowitz and D. B. Boyd (ed.), Reviews in Computational Chemistry, Volume 18. John Wiley and Sons, Inc., Hoboken, New Jersey, 2002, pp. 41.
17. Gohlke, H. and Klebe, G., Cur. Opin. Struct. Biol., 11 (2001) 231.
18. Muegge, I. and Rarey, M. In K. B. Lipkowitz and D. B. Boyd (ed.), Reviews in Computational Chemistry, Volume 17. Wiley-VCH, John Wiley and Sons, Inc., New York, 2001, pp. 1.
19. Bissantz, C., Folkers, G. and Rognan, D., J. Med. Chem., 43 (2000) 4759.
20. Charifson, P.S., Corkery, J.J., Murcko, M.A. and Walters, W.P., J. Med. Chem., 42 (1999) 5100.
21. Hoffmann, D., Kramer, B., Washio, T., Steinmetzer, T., Rarey, M. and Lengauer, T., J. Med. Chem., 42 (1999) 4422.
22. Sybyl v. 6.8, Tripos Inc., 1699 South Hanley Rd, St. Louis, Missouri, USA,
23. Halgren, T.A., J. Comput. Chem., 17 (1996) 490.
24. Halgren, T.A., J. Comput. Chem., 17 (1996) 520.
25. Halgren, T.A., J. Comput. Chem., 17 (1996) 553.
26. Halgren, T.A. and Nachbar, R.B., J. Comput. Chem., 17 (1996) 587.
27. Halgren, T.A., J. Comput. Chem., 17 (1996) 616.
28. Mohamadi, F., Richards, N.G.J., Guida, W.C., Liskamp, R., Lipton, M., Caufield, C., Chang, G., Hendrickson, T. and Still, W.C., J. Comput. Chem., 11 (1990) 440.
29. GRID v. 16, Molecular Discovery Ltd., West Way House, Elms Parade, Oxford, England, 1996-1999.
30. GRID v. 18, Molecular Discovery Ltd., West Way House, Elms Parade, Oxford, England, 1999-2000.
31. Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N. and Bourne, P.E., Nucl. Acids. Res., 28 (2000) 235.
32. Gundertofte, K., Liljefors, T., Norrby, P.-O. and Pettersson, I., J. Comput. Chem., 17 (1996) 429.
33. Jørgensen, A.T., Norrby, P.-O. and Liljefors, T., J. Comput-Aided Mol. Des., 16 (2002) 167.
34. Insight II v. 98.0 Molecular Modeling System, Molecular Simulations, Cambridge, U.K., 1998.
35. Goodman, J.M. and Still, W.C., J. Comput. Chem., 12 (1991) 1110.
36. Chang, G., Wayne, C.G. and Still, W.C., J. Am. Chem. Soc., 111 (1989) 4379.
37. Still, W.C., Tempczyk, A., Hawley, R.C. and Hendrickson, T., J. Am. Chem. Soc., 112 (1990) 6127.
38. Clark, R.D., Strizhev, A., Leonard, J.M., Blake, J.F. and Matthew, J.B., J. Mol. Graphics. Mod., 20 (2002) 281.
39. Rarey, M., Kramer, B., Lengauer, T. and Klebe, G., J. Mol. Biol., 261 (1996) 470.
40. Jones, G., Willett, P., Glen, R.C., Leach, A.R. and Taylor, R., J. Mol. Biol., 267 (1997) 727.
41. Meng, E.C., Shoichet, B.K. and Kuntz, I.D., J. Comput. Chem., 13 (1992) 505.
42. Muegge, I. and Martin, Y.C., J. Med. Chem., 42 (1999) 791.
43. Eldridge, M.D., Murray, C.W., Auton, T.R., Paolini, G.V. and Mee, R.P., J. Comput.-Aided Mol. Des., 11 (1997) 425.
44. Boström, J., Norrby, P.-O. and Liljefors, T., J. Comput.-Aided Mol. Des., 12 (1998) 383.
45. Lowther, W.T., Zhang, Y., Sampson, P.B., Honek, J.F. and Matthews, B.W., Biochemistry, 38 (1999) 14810.
46. Goodford, P. J., Molecular Discovery Ltd., England, personal communication, 2001.