UPTEC X 02 040
OCT 2002
ISSN 1401-2138

EVA STJERNSCHANTZ

Evaluation of docking in ICM as a tool in structure-based drug design

Master's degree project

Molecular Biotechnology Programme
Uppsala University School of Engineering

UPTEC X 02 040 | **Date of issue 2002-10-04**

Author

Eva Stjernschantz

Title (English)

Evaluation of docking in ICM as a tool in structure-based drug design

Title (Swedish)

Abstract

The ICM docking algorithm was studied and evaluated as a tool in structure-based drug design. ICM VLS, v. 2.8 Molsoft L.L.C, was evaluated in terms of docking accuracy and scoring function performance. Different approaches to improve docking results were studied and generated conformations were analysed. The discriminative power of the implemented scoring functions i.e. the ability of distinguishing a small number of active compounds from a large compound database was studied. Existing methods of processing virtual screening results were applied and compared with other filtering method approaches. Consensus scoring and Bayesian classification based on the two scoring functions implemented in ICM were shown to be superior in discriminating active from inactive compounds compared to ranking based on single scoring function results.

Keywords

Docking, virtual screening, scoring functions, consensus scoring, Bayesian classification, conformational analysis

Supervisors

Micael Jacobsson
Department of Structural Chemistry, Biovitrum AB

Examiner

Dr Anders Karlén
Department of Organic Pharmaceutical Chemistry, Uppsala University

Project name	Sponsors
Language **English**	Security
ISSN 1401-2138	Classification
Supplementary bibliographical information	Pages **67**

Biology Education Centre Biomedical Center Husargatan 3 Uppsala
Box 592 S-75124 Uppsala Tel +46 (0)18 4710000 Fax +46 (0)18 555217

Evaluation of docking in ICM as a tool in structure-based drug design

Eva Stjernschantz

Sammanfattning

Förståelsen för hur en läkemedelsmolekyl interagerar med en receptor är ett viktigt steg i utvecklandet av ett läkemedel. Kunskapen om hur molekylerna interagerar, t ex var på receptorn läkemedelsmolekylen binder, den lokala miljön runt bindningsstället, typer av interaktioner och bindningsstyrka, är användbart för att förbättra eller hitta helt nya potentiella läkemedelsmolekyler. Som ett första steg för att hitta nya läkemedelsmolekyler screenas ofta en stor mängd olika molekyler kemiskt, vilket är en tidsödande och dyr process.

Dockning är ett sätt att modellera dessa molekylinteraktioner. Dockning kan användas som ett hjälpmedel för att förbättra potentiella läkemedelsmolekylers egenskaper eller som ett steg innan screening. Genom att utföra en ”virtuell screen” innan den kemiska screenen, dvs docka/modellera bindning av ett stort antal, t ex 100 000, molekyler till en utvald receptor, kan idealt ett stort antal av dessa molekyler filtreras bort. Den faktiska kemiska screenen kan därmed effektiviseras och inriktas på mer ”rimliga” molekyler.

Här har användningen av dockning, både för optimering och upptäckt av nya potentiella läkemedelsmolekyler, utvärderats. Pålitligheten hos dockning i olika sammanhang har studerats och olika metoder för att förbättra dockningsresultat har undersökts. För användning inom virtuell screen har olika statistiska metoder för att filtrera bort dåligt bindande molekyler studerats.

Examensarbete 20 p, Molekylär bioteknikprogrammet
Uppsala universitet Oktober 2002

1 Introduction

Existing docking programs are subjected to two main requirements for effective use in structure-based drug design. The docking part of the programs should model the interactions and energy of the receptor-ligand system accurately presenting the lowest-energy conformation as a reliable docking result. The scoring part of the programs should be able to discriminate between binders and non-binders and give an estimation of the binding affinity of a compound [1].

Docking as a tool in structure-based drug design has two main applications, docking in the lead optimization phase and virtual screening in the lead discovery phase. Docking in lead optimization is used to predict and evaluate interactions between a limited amount of similar ligands and a target protein. This requires docking programs simulating interactions correctly and preferably scoring functions giving an accurate indication of binding affinity [1].

Experimental high-throughput screening (HTS) of compound libraries is a highly expensive and complex technology. With the rapidly increasing amount of three-dimensional structures a theoretical alternative to HTS could be advantageous. High-throughput docking of virtual compound libraries, virtual screening, can be used as a filtering method requiring fewer compounds to be screened and improving hit rates in subsequent HTS. Using docking and scoring in lead discovery thus requires docking programs that accurately identify binders from large sets of diverse molecules. Well performing scoring functions able to discriminate between binders and non-binders as well as time-effectiveness of the docking program are requirements [1].

The presently existing docking programs display weaknesses in a variety of areas. The requirement of time-effectiveness combined with the importance of the ability to accurately model systems displaying diverse properties complicate the procedure of developing reliable docking programs.

Evaluation of the performance of the ICM docking software in terms of docking and scoring is presented below.

2 Theory

2.1 ICM docking algorithm

The ICM docking algorithm is a global energy optimization procedure based on Monte Carlo minimization, exploring conformational space of the ligand in the vicinity of protein. The protein is held rigid and represented by pre-calculated grid potential maps, describing the interaction energy between the protein and a probe at each grid point. Five grid potentials are calculated, with a grid spacing of 0.5 Å; van der Waals grid potential for a hydrogen probe, van der Waals grid potential for a non-hydrogen probe, hydrogen bond grid potential, electrostatic grid potential and hydrophobic grid potential [2, 3].

The ICM docking algorithm is based on global optimization of the energy function describing the intra-molecular ligand energy and the total interaction energy of the ligand-receptor complex [2].

2.1.1 Global optimization

The ICM docking algorithm is a Monte Carlo energy minimization consisting of the following [4]:

1. Random conformational and orientational change of the ligand. The ligand is positioned as well as rotated around its center of gravity by pseudo-Brownian random movements. Torsion angles are changed in a random manner, one at a time, with amplitude 180°.
2. Local energy minimization of analytically differentiable terms using the conjugate gradient minimization method. The surface-based solvation energy is non-differentiable and will not be included.
3. Calculation of non-differentiable terms, e.g. solvation energy, and addition of the terms to the local energy minimum obtained above.
4. Application of the Metropolis selection criterion to decide whether to discard or accept the suggested conformation. A calculated energy lower than previously obtained ones will be accepted and used as the new energy criterion. An energy higher than the energy criterion will be accepted with probability:

$$P_{acc} = e^{-(E_{new} - E_{crit})/kT} \quad (1)$$

k = Boltzmann's constant

T = Temperature of simulation (K)

The global optimization is performed from multiple starting points, the number of starting points depending on the size of the ligand [3].

An alternative approach to adding the non-differentiable energy terms calculated for the conformation obtained after local minimization is excluding local minimization from the procedure. It has, however, been shown that full local energy minimization after each random step improves the conformational search results significantly [3]. The approximation of adding the non-differentiable energy to the local energy minimum should not affect the results significantly if the added energy terms have "slow" derivatives i.e. are not greatly affected by small conformational changes [3].

2.1.2 Internal variables

The global energy optimization involves an extensive search of conformational space, including ligand conformation as well as orientation. Both the high dimensionality of the system and multiplicity of the energy hyper surface complicate the search procedure [4]. To reduce the dimensionality of the system ICM uses internal variables to describe the molecules [3]. The dimensionality can be reduced by converting Cartesian coordinates (x,y,z) describing the position of each atom independently to internal coordinates, dependent of each other. The internal coordinates are bond length, bond angle and torsion angle. As bond lengths and bond angles in most cases can be described as rigid the only remaining variable parameter will be torsion angle. The internal variables are constructed as a directed tree-like graph superimposed on the atoms of the system, including a number of virtual atoms. Each atom position is described by three geometrical parameters with respect to the preceding parts of the tree; bond length, bond angle and torsion or phase angle for main and side branches respectively [3]. Ring closures are considered fixed rigid bodies, implicating that the ICM docking program does not sample ligand ring conformations.

2.1.3 Saving low-energy conformations

ICM saves a specified number of low-energy conformations in a stack during global optimization [5]. Conformational suggestions during Monte Carlo minimization are compared energetically and geometrically with other conformations in the stack and appropriate substitutions of stack conformations are made to present only low-energy conformations. Incorrect docking results can in some cases be explained by inability of distinguishing

between stack conformations to present the correct, lowest energy conformation as the docking result. The energy differences between consecutive stack conformations as well as the number of visits to each local minimum, representing stack conformations, give an indication of sufficiency of sampling [5].

ICM offers the possibility of flexible receptor calculations of stack conformations subsequent to docking. The receptor is modeled in full-atom representation and receptor side-chains are relaxed taking the interactions with the docked ligand into consideration. The ligand is kept only partially flexible, with constrained heavy atom positions [5]. The receptor side-chain relaxation is performed applying a biased probability Monte Carlo conformational search [6].

2.2 Energy function

The energy function globally optimized in the ICM docking procedure includes two intra-molecular energy contributions describing ligand conformation, namely van der Waals interaction energy and torsion energy. The inter-molecular energy contributions included in the energy function are inter-molecular van der Waals energy, hydrogen bond energy, electrostatic energy and hydrophobic interaction energy. Equation 2 describes the total energy of the system calculated for each conformational suggestion. The first term includes the intra-molecular energy contributions, whereas the remaining terms describe the interaction energies of the complex [2, 3].

$$E = E_{ff} + E_{vW} + E_{hb} + E_{el} + E_{hp} \qquad (2)$$

2.2.1 Van der Waals interactions

The intra-molecular van der Waals interaction energy is described by the 6-12 potential (3):

$$E_{vW}(R_{ij}) = -\frac{A_{ij}}{R_{ij}^6} + \frac{B_{ij}}{R_{ij}^{12}} \qquad (3)$$

R_{ij} = Distance between atoms *i* and *j*

A and B for different atom pairs are determined by the MMFF force field used in the calculations [2].

The interaction consists of a rather weak attractive term, decreasing rapidly with distance, and a repulsive term, which at short inter-atomic distances rises steeply towards infinity. Being more difficult to estimate, the constant describing the repulsive term produces "background noise" when calculating binding energies. The inter-molecular van der Waals potential is therefore slightly modified introducing a maximum energy value thus truncating the repulsive part of the interaction, which results in a smoother potential (4). Negative energy values dominated by the attractive term remain indifferent to the modification, whereas positive values for short inter-atomic distances result in a potential asymptotically approaching the maximum energy [2].

$$E_{vW} = \begin{cases} E_{vW}^{0} & E_{vW}^{0} \leq 0 \\ \dfrac{E_{vW}^{0} E_{\max}}{E_{vW}^{0} + E_{\max}} & E_{vW}^{0} > 0 \end{cases} \qquad (4)$$

On the basis of the number of interactions between the molecules being approximately the same before and after interaction, the interactions with solvent being substituted by interactions within the complex, an alternative approach would be to exclude the inter-molecular van der Waals interactions from the energy function completely. Not considering inter-molecular van der Waals interactions would however result in loss of information about the molecular interface quality [2].

2.2.2 Hydrogen bonds

Hydrogen bonds form a significant part of complex formation and hydrogen bond networks play an important role for the affinity of as well as the specificity of ligands. The hydrogen bond energy is calculated according to Equation 5 [2].

$$\mathrm{E}_{\mathrm{hb}} = \mathrm{E}_{\mathrm{hb}}^{0} e^{-(t - r_{hb})^2 / {d_{hb}}^2} \qquad (5)$$

d_{hb} = Radius of interaction sphere set to 1.4 Å

r_{hb} = Radius vector of the interaction center 1.7 Å from the atom

$\mathrm{E}_{\mathrm{hb}}^{0}$ = Maximum interaction energy set to 2.5 kcal / mol

Hydrogen bonds display high directionality. The lone pairs of the acceptor atoms occupy the sp-orbitals, resulting in an uneven charge distribution around these atoms. Most force fields

do not take this uneven charge distribution into account but use spherical atom-centric potentials, which can cause inaccuracies in modeling of hydrogen bond interactions during docking calculations [3]. The radius of the interaction sphere in Equation 5 constrains the interaction to a certain direction, but allows a deviation of 30-40°. The interaction vector is placed along the axis of a covalent bond of the polar hydrogen atom and depending on hybridization of the acceptor atom in appropriate directions [3].

2.2.3 Electrostatic energy

The electrostatic energy contribution is described by Coulomb's law (6), using a distance-dependant dielectric constant $\varepsilon = \varepsilon_0 r$,

$$E_{el} = k\frac{q_1 q_2}{eR_{12}} \tag{6}$$

R_{12} = Inter-atomic distance.

The distance dependant dielectric constant offers an approximation of the screening of charges by solvent molecules, which is not included in the calculations [2]. Without any type of simulated solvent screening the calculations would not generate reasonable results, but strongly interacting charges would result in collapsed molecules. Including a static layer of water molecules in the simulation would not be more accurate, due to the dynamic nature of the solvent. A molecular dynamics simulation would have to be performed to generate the thermodynamic ensemble of solvent states, which in most cases is too rigorous an approach. The distance dependant dielectric constant partly simulates the shielding of charges by taking into account the weakened electrostatic interactions with distance. The distance dependant dielectric constant is however a rather large approximation, not considering the interactions between solvent molecules and molecular charges [2].

Alternative approaches to describing the electrostatic interactions are modeling the solvent as a continuous medium with high dielectric constant and solving the Poisson equation, either analytically or numerically [2]. The electrostatic energy is calculated as the energy of point charges in a medium of low dielectric constant, representing the protein, surrounded by a medium with high dielectric constant, representing the solvent. Approximating the boundary,

i.e. the molecular surface, to a spherical shape, makes it possible to solve the Poisson equation (7) analytically [2].

$$-\nabla[e(r)\nabla f(r)] = r(r) \tag{7}$$

$r(r)$ = Charge density

$f(r)$ = Electric potential

Not approximating the boundary surface to a spherical shape, a numerical solution can be obtained using the boundary element method [3]. Solving the Poisson Equation for two regions with diverse dielectric constants is equal to solving the Equation for a region with constant permittivity, i.e. applying Coulomb's law, if additional charges are distributed on the boundary between the two regions. The charge distribution on the dielectric boundary surface is calculated, making it possible to calculate the electrostatic potential at any point by applying Coulomb's law, the permittivity now being constant [3]. ICM uses a distance dependant dielectric constant for electrostatic energy calculations during the Monte Carlo simulation and re-evaluates the electrostatic free energy by solving the Poisson Equation numerically, applying the rapid boundary element solvation electrostatics algorithm, for score calculations [7]. The generalized Born approximation is an alternative approach to describing electrostatic interactions [5].

2.2.4 Hydrophobic interactions

Most ligands being amphiphilic, the hydrophobic effect forms an important part of complex formation.

The hydrophobic potential is described by Equation 8 [2].

$$E_{hp} = E_{hp}^{0} e^{-d_{surf}^2 / d_w^2} \tag{8}$$

d_{surf} = Distance to closest point of the hydrophobic surface

d_w = Effective radius of hydrophobic interaction

The hydrophobic potential is approximated to be proportional to the buried hydrophobic surface in the complex. The solvent accessible surface is obtained applying a modified form of the Shrake and Rupley algorithm [2]. The surface of a non-hydrogen atom is depicted by a set of points uniformly spread on a sphere. The number of dots not occluded by neighboring atom surfaces represents the accessible surface of the atom. The solvent accessible surface is used for hydrophobic interaction calculations as well as for application of the boundary element method for electrostatic interaction calculations described above [3].

2.3 Scoring functions

As a result of docking the lowest energy conformation is stored and scoring functions are applied to the result. Two scoring functions based on two diverse approaches of evaluating ligand-receptor interactions are implemented in the ICM docking software. The ICM Score is an empirical scoring function based on calculation of physiochemical properties of the receptor-ligand complex [5]. The PMF (Potential of Mean Force) score is a statistical knowledge based scoring function based on structural information of known protein-ligand complexes [5].

2.3.1 ICM score

The ICM Score is calculated as the weighted sum of scores describing the energy terms evaluated during docking simulations [5]. *Grid score* is the score calculated as the sum of internal torsion and van der Waals interaction energy for the obtained ligand conformation and the interaction energy with the five grid potential maps subtracted with the energy for the free ligand in aqueous solution. The ligand is subsequently combined with a full atom receptor model for electrostatic calculations. As a result the *hydrogen bond score* is calculated. The *electrostatic score* is calculated as the difference between electrostatic interaction energy evaluated applying the boundary surface method and the electrostatic interaction energy of the free ligand in solution. The *surface score* is calculated as the hydrophobic interaction energy and the *Hp score* as the difference between the hydrophobic interaction energy for the bound ligand conformation and the free ligand conformation in solution. The ICM Score is the result of the weighted sum of the different scores described above, where the coefficients have been determined by fitting the derived scoring function to activities of a training set of protein-ligand complexes [5].

2.3.2 PMF score

The PMF Score is a knowledge-based scoring function solely based on observed distance distributions of specific atom types in protein-ligand complexes extracted from the Brookhaven Protein Data Bank, PDB. The structural information from PDB is converted to Helmholtz free interaction energy for different protein-ligand atom pairs. The score is calculated from the atom pair distances of the docked receptor-ligand complex [8].

The protein-ligand free energy (PMF) between protein atom type *i* and ligand atom type *j* is described in Equation 9 [8].

$$A_{ij}(r) = -kT \ln\left(f^{j}_{Vol\_corr}(r) \frac{r^{ij}_{seg}(r)}{r^{ij}_{bulk}} \right) \quad (9)$$

k = Boltzmann's constant

T = Absolute temperature

r = Atom pair distance

$f^{j}_{Vol\_corr}(r)$ = Ligand volume correction factor

$r^{ij}_{seg}(r)$ = Number density of pairs *ij* occurring in radius range *seg* in structural database

r^{ij}_{bulk} = Distribution of i and j when no interaction occurs

$r^{ij}_{seg}(r) / r^{ij}_{bulk}$ describes the radial distribution function of protein atom type *i* paired with ligand atom type *j* in a structural database.

The PMF Score is defined as the sum over all interactions between atoms in the complex (10).

$$\sum_{kl}^{r<r^{ij}_{cut-off}} A_{ij}(r) \quad (10)$$

kl = All protein-ligand atom pairs in the database

$r^{ij}_{cut-off}$ = Cut-off radius of atom type pair *ij*

2.4 Consensus scoring

The requirement of scoring functions is to be able to predict the binding affinity of a compound, or at least give an estimate of the "tightness of binding", making it possible to distinguish between potential binders and non-binders in a large set of compounds. Ranking molecules according to score and determining threshold score values or threshold "rank positions" for defining binders has been the main approach for analyzing docking results of a large set of compounds from a virtual screen [9, 10, 11]. However, the docked structures very seldom obtain "crystallographically correct conformations", resulting in large inaccuracies in prediction of affinities for all presently existing scoring functions. Several studies have shown that different scoring functions display different strengths and weaknesses in describing ligand-receptor interactions. A natural extension to using one scoring function alone for discriminating binders from non-binders is applying a number of scoring functions to the docking results and only considering compounds receiving high ranks with two or more scoring functions. Several studies of this approach, i. e. consensus scoring, show significant improvement of hit rates compared to one-dimensional scoring [9, 10, 11].

Two scoring functions are implemented in ICM and an additional five scoring functions can be applied to docking results using the CScore module in SYBYL [12] comprising Gscore, Fscore, Dscore, Chemscore and PMFscore.

2.5 Bayesian classification

An alternative approach to consensus scoring for discriminating between active and inactive compounds is construction of simple classifiers and applying these on scoring results. Compounds are classified into two classes, binders and non-binders. A Bayesian classifier can be constructed using scoring results of non-binders as well as a training set of known binders. The test set can subsequently be classified accordingly. The simple Bayesian classifier approximates the score values of the two classes to be normally distributed. The docked compounds are classified according to Equation 11, derived from Bayes theorem [13].

$$\begin{aligned} &\log P_{act} - 0.5 * \log\left|C_{act}\right| + \left[S - M_{act}\right]' * inv\left[C_{act}\right] * \left[S - M_{act}\right] > \\ &\log P_{inact} - 0.5 * \log\left|C_{inact}\right| + \left[S - M_{inact}\right]' * inv\left[C_{inact}\right] * \left[S - M_{inact}\right] \end{aligned} \qquad (11)$$

P_{act} and P_{inact} are the a priori probabilities for a compound being active and inactive respectively, C_{act} and C_{inact} are the covariance matrices of the scores of the first ("active") and second ("inactive") set of compounds respectively, M_{act} and M_{inact} are the mean value

matrices of the scores of the first and second set respectively, and S is the score value matrix. When the inequality is satisfied the compound is classified active [13].

2.6 Analysis of conformations generated in ICM

To study the quality of conformations generated in ICM a conformational search can be performed, resulting in the global energy minimum conformation and a number of low-energy conformations of a compound [14]. The conformational search procedure is a systematic pseudo Monte Carlo search, modifying torsion angles of the molecule [15]. Local energy minimization is performed and conformations similar in energy or RMS deviation to previously found structures are discarded. The search procedure exploring conformational space is exhaustive, terminating when each low-energy conformation has been found a specified number of times [15]. The obtained global energy minimum can be used for comparison with the docked conformation generated in ICM. To study the reliability of the method of comparison, "correct" conformations obtained from crystal structures can be subjected to the same energy calculations and compared to the obtained global energy minimum conformation. *Boström et al* [16] have shown that approximately 70 % of ligands from crystal structures of ligand-receptor complexes are low-energy conformations and display conformational energies $\leq$ 3.0 kcal/mol above the global energy minimum of the ligand.

3 Calculations

3.1 Docking evaluation calculations

8 protein structures with a total of 71 ligands were used for the docking evaluation, as displayed in Table 1. The coordinates of the complex structures were obtained from the Brookhaven Protein Data Bank, PDB [17].

Characteristic	Protein	PDB accession number
Nuclear receptor	Estrogen receptor, ER	1ere
		1err
		1qku
		3erd
		3ert
	Retinoic acid receptor γ, RARγ	1exa
		1exx
		1fcx
		1fcy
		1fcz
		2lbd
		3lbd
		4lbd
Matrix metallo proteases	MMP-1, collagenase	1hfc
		2tcl
		966c
	MMP-3, stromelysin	1b3d
		1b8y
		1biw
		1bqo
		1c3i
		1caq
		1d5j
		1d7x
		1d8f
		1d8m
		1g05
		1g49
		1g4k
		1hfs
		1hy7
		1sln
		1usn
		2usn
Serine proteases	Thrombin	1etr
		1tes
		1ett
		1uvt
		1uvu
	Trypsin	1aq7
		1az8
		1bty
		1c5q
		1f0t
		1f0u
		1k1i

		1k1j
		1k1l
		1k1m
		1k1n
		1k1o
		1k1p
		1mts
		1qb6
		1zzz
Carbonic anhydrase	Carbonic anhydrase II	1a42
		1bn1
		1bn4
		1bnm
		1bnq
		1bnt
		1bnu
		1bnv
		1bnw
		1i8z
		1i90
		1i91
MAP kinase	P38	1b16
		1b17
		1bmk
		1di9

Table 1. PDB accession numbers for complex structures used in the docking evaluation.

Correct ligand coordinates were obtained from the receptor-ligand complex crystal structures and used for RMS deviation calculations and qualitative analysis of the docking results. Ligand structures were compiled originating from the corresponding crystal structures and using ReliBase [18]. To randomize the conformations of the ligands to be docked 2D structures were generated and new three-dimensional ligand structures were generated using Corina [19]. In an attempt to improve docking results diverse ligand representations were generated and docked. Manual modifications of ligand structures, involving atom type, enantiomeric form, specific protonation/deprotonation, were all made using SYBYL [12]. The PDB coordinates of protein structures and occasionally included solvent molecules were converted to ICM internal coordinates using either the "Convert" or the newer "ConvertObject" macro in ICM [5]. Two different ionization scripts were used to ionize ligands according to diverse ionization rules. Generation of ligands with alternative ring conformations was performed with a script, generating a maximum of ten different ring conformations per ligand. A set of ligands was docked to one or several complex structures of the same protein.

Conformational analysis, i.e. conformational search and energy calculations of docked and correct ligand structures, was performed using MacroModel [15] in Maestro [14]. The conformational search method used was SUMM, using the MMFFs [37] force field and water as solvent. The local minimization method used was PRCG and the energy window for saving low-energy conformations was set to 15 kJ/mol. Multiple energy minimization was performed, discarding conformations similar in energy or RMS deviation to any previously obtained low-energy conformations. The correct and docked ligand conformations were subjected to a constrained local energy minimization, with a distance constraint of 0.3 Å and a force constant of 500 kJ/(mol $Å^2$) for heavy atoms. The low-energy conformational limit was set to 15 kJ/mol above the obtained global energy minimum.

3.2 Scoring function evaluation calculations

1000 randomly chosen compounds from the MDDR database [20] were ionized and subsequently docked to stromelysin (1hy7), thrombin (1kts) and estrogen receptor (1ere) structures (PDB accession numbers [17]). The set of MDDR compounds was selected to maximize diversity and was considered inactive in the evaluation of the discriminative power of the scoring functions. Docking against these structures was also performed for 51 stromelysin inhibitors [21], 24 thrombin inhibitors [22] and 30 ER agonists [23] with known activity. In order to evaluate the accuracy of different score/rank thresholds obtained from docking results of the stromelysin inhibitor training set, an additional 20 stromelysin inhibitors, with known activity were docked to the same MMP-3 structure. The inhibitors of this test set all display diverse P and P' groups compared to the training set inhibitors.

At least two consecutive runs were performed to investigate the reproducibility of the docking results both for inactive and active compounds. Based on the scoring results of the stromelysin inhibitor training set 154 MDDR compounds were extracted from the set of 1000 random inactive compounds, with ICM- and PMF scores better than the lowest scoring active compound. (ICM score < -15.0 and PMF score< -37.5) These compounds as well as the 51 stromelysin inhibitors were re-docked to the same structure, using the same parameters. The test set was in a similar way re-docked to the same structure. A threshold Δ score value describing significant differences in binding mode was obtained qualitatively and used as a filtering method discarding compounds with high Δ score. 102 MDDR compounds were extracted from the set of 1000 inactive compounds docked to the ER structure, according to

ICM score< -27.0 and PMF score < -50.0 and re-docked to the estrogen receptor structure. The set of ER agonists were re-docked to the same structure.

To study the effect of induced fit the stromelysin inhibitor test set was docked in two consecutive runs to two additional stromelysin structures, 1caq [17] and 1biw [17]. Multiple ring conformations were also generated and different ionization rules were applied to the stromelysin test set. The set was also docked to the same structures using the generalized Born approximation as an alternative description of the solvation electrostatic term [5].

4 Qualitative and quantitative evaluation of the ICM docking program

4.1 Introduction

The ICM docking program was evaluated studying several ligand-receptor complexes to investigate when good reliable docking results can be expected as well as in what cases docking might fail. The proteins display different active site properties, involving size, type of interactions, variability, metal ions etc. The structures used for the evaluation also display different qualities/resolutions. Docking was performed against eight different proteins with a number of diverse ligands to each protein, obtained from crystal structures of ligand-receptor complexes. Root mean square deviation values between the docked and correct ligand structures were calculated and formed receptor-ligand interactions were studied qualitatively, to give a quantitative as well as qualitative analysis of the docking results. Different approaches to improve the docking results were tried, involving both ligand and protein representation.

One of the major expected problems with current docking tools is the induced fit phenomenon, i.e. the receptor flexibility during interaction with the ligand. For practical purposes, ie. time-effectiveness and computer power, present docking algorithms only take into account ligand flexibility and omits the flexibility displayed by the receptor [1]. For docking experiments in general and applied to virtual ligand screening in particular, induced fit can cause a large number of failed docking calculations. To study the impact of induced fit on docking results, docking of a set of ligands obtained from complex crystal structures was performed to one as well as several different protein structures. Different modes of protein representation were also investigated, involving orientation of polar hydrogens as well as including solvent molecules in the calculations.

As an alternative approach to solving the problems caused by induced fit, flexible receptor calculations on saved stack conformations can be performed in ICM [5], which was also evaluated.

Simulated electrostatic interactions will be dependent on the partial charges assigned by the force field, and thus dependant on atom types used, as well as ionization state of the receptor residues and ligand. Dielectric fit, i.e. changes of the ionization state due to local environment, will complicate the simulation of electrostatic interactions additionally [24]. The impact of ligand ionization was investigated docking ligands with diverse ionization states. To automate the docking procedure general ionization rules are required and a new ionization script was evaluated.

The ICM docking algorithm samples only torsional and orientational changes of the ligand, leaving bond angles and lengths fixed [3]. Different enantiomers of a compound will not be sampled and manual generation of enantiomers is needed. Ring closures are considered to be fixed, rigid bodies in the ICM internal coordinate system, and are not sampled during Monte Carlo minimization [3]. The impact on docking results of enantiomeric discrimination as well as generation of different ring conformations prior to docking was studied.

To investigate whether the docked ligand conformations generated by ICM are realistic, low-energy conformations a conformational search was performed for a majority of the ligands and energy calculations for the docked as well as the correct structures were made. For cases displaying high-energy conformations of the docked ligands, generated low-energy conformations were further used in rigid docking calculations, i.e. docking a rigid ligand to a rigid receptor.

The docking calculation, being a global energy optimization based on Monte Carlo minimization, is a probabilistic procedure and therefore requires a minimum number of simulation steps to give reproducible/sTable results [5]. The reproducibility of docking results and thus reliability of the docking program was investigated applying different numbers of Monte Carlo steps in the docking calculations.

General conclusions about the docking programs ability to model different types of interactions, the reproducibility of results/stability of the docking program and cases where docking is unsuccessful using the current docking tool in ICM are presented below. Practical aspects of docking in ICM are also presented.

4.2 Docking evaluation results

A detailed description of the different ligand-protein complexes used in the evaluation and their docking results is given below, followed by general conclusions about docking in ICM.

4.2.1 Estrogen receptor, ER

The estrogen receptor (ER) is a nuclear transcription factor induced by binding of estradiol [25]. Estradiol binds in a hydrophobic region, separated from the hydrophilic external environment. When estradiol is bound alpha helix twelve sits as a lid over the ligand-binding cavity. Glu353, Arg 394 and a water molecule form a hydrogen bond network anchoring ligands to the ligand-binding pocket. His524, on the opposite side of the inner parts of the pocket offers additional ligand anchoring through hydrogen bonding. The remaining part of the ligand-binding domain is mainly hydrophobic in nature. Antagonists show binding similarities to estradiol. The side chain of the antagonist is however bulky and cannot be contained within the binding cavity, and thus displaces alpha helix 12, from the "lid position" over the cavity. The resulting conformational change of the protein structure is the basis for inhibition [25].

Both agonists and antagonists were docked to ER. The ligand structures are displayed in appendix 1.

4.2.1.1 Docking results

Due to the major differences in protein structure of agonist and antagonist bound ER, the ligands need to be docked to different crystal structures. Docking agonist and antagonist ligands to their respective structures results in very well modeled interaction and low RMS deviations to the correct ligand structures. The smaller agonist ligands can be docked successfully to the larger antagonist-binding cavity. RMS deviation values are displayed in Table 2.

4.2.2 Retinoic acid receptor gamma, RARg

Retinoic acid receptors, RAR α, β and γ, belong to the nuclear receptor family, regulating transcription of target genes in a ligand inducible manner [26]. The ligand-binding cavity of RARγ is characterized by a hydrogen bond network at the inner end of the pocket, formed by

an Arg residue, a Ser residue a water molecule and a Leu residue. The ligand-binding pocket is elsewhere mainly constituted of hydrophobic residues. Enantiomer discrimination has been observed for RARγ agonists, one of the enantiomers lacking biological activity. Crystal structures of complexes with the inactive ligand have still been solved showing no differences in conformation of the receptor when bound to the two different enantiomers. This indicates that the lack of activity is not due to induction of conformational changes of the receptor as is the case for ER, where transactivation is inhibited through displacement of α helix twelve, but high-energy ligand conformation and unfavorable contacts with the receptor [26].

The inactive ligand, 1exx, docked to the active 1exa structure is displayed in Figure 1. The ligands docked to the RARγ structure are displayed in appendix 1.

4.2.2.1 Docking results

Important factors for obtaining good docking results docking to RARγ are listed below.

Including solvent molecules and ionization of ligands**.** Without including the water molecule, involved in the hydrogen bond network, or ionizing the ligands, docking results are generally poor. The water molecule could be important for anchoring the ligands, in particular for the retinoic acid ligands, which are elsewhere hydrophobic. The most important reason for not obtaining the correct binding mode is, however, most likely the lack of ionization of the ligands. For formation of electrostatic interactions it seems necessary to deprotonate carboxylic acids before docking. When including the water molecule involved in the hydrogen-bonding network and ionizing the ligands correct binding mode is obtained for all ligands with generally low RMS deviation values from the correct structures. The docking results for the 1exa ligand are shown in Figure 2.

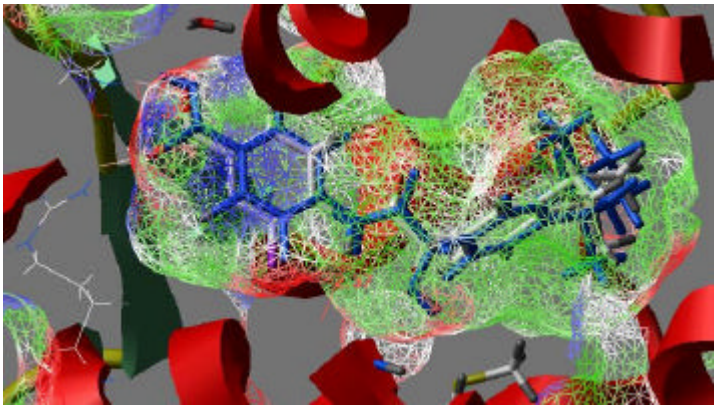


Figure 1. Ligand of complex crystal structure 1exx docked to structure 1exa. Correct ligand conformation colored in blue. The ligand-binding pocket is displayed. Green=Hydrophobic. Red=Hydrogen bond acceptor. Blue=Hydrogen bond donor.

The docking program does not give the expected lower score values for the inactive enantiomer, 1exx, as compared to 1exa, thus failing to illustrate the lack of activity of the enantiomer.

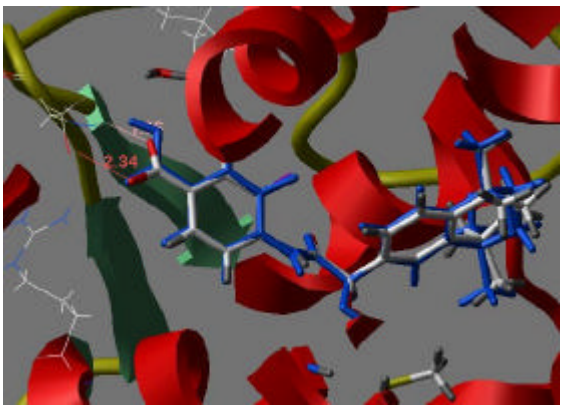


Figure 2. Results of docking the synthetic agonist, ligand of 1exa RARγ complex structure to the 1exa crystal structure. The correct ligand conformation is colored in blue.

The RMS deviation values between the correct ligand structures and the corresponding docked ligands are summarized in Table 2.

4.2.3 Carbonic anhydrase II

Carbonic anhydrase is a plasma membrane associated enzyme involved in catalyzing the zinc-dependant hydration of carbon dioxide [27]. The active site is a 15 Å deep, predominantly hydrophobic cavity, open towards bulk solvent. A catalytic zinc ion is positioned at the bottom of the cavity. The zinc ion is coordinated by three histidine residues and, in uninhibited form, a hydroxide ion.

The ionized sulphone amide nitrogen displaces the hydroxyl group and coordinates the zinc ion as well as participates in a hydrogen bond network with protein residues. One of the oxygen atoms of the sulphone amide is involved in hydrogen bond formation, whilst the other oxygen atom is neither involved in hydrogen bonding nor zinc coordination. The oxygens of the second sulphone amide form polar interactions with protein residues and a buried solvent molecule and the aliphatic tail form van der Waals contacts with hydrophobic residues on each side of the active site cleft [27].

Figure 3 shows the “1i90” carbonic anhydrase inhibitor superimposed on the active site of the 1bn1 protein structure.

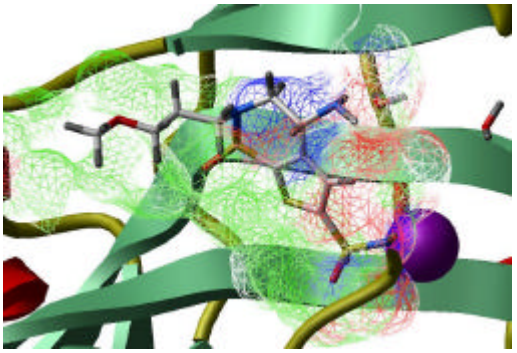


Figure 3. The inhibitor of the 1i90 complex structure superimposed on the 1bn1 protein structure. The active site cavity is displayed, green representing hydrophobic residues, blue hydrogen bond donors and red hydrogen bond acceptors. The catalytic zinc ion is colored violet. Two water molecules are included.

4.2.3.1 Docking results

The results of docking the inhibitor set to carbonic anhydrase II indicate some measures to be taken prior to docking as well as conditions where the ICM docking program fails to model the system correctly.

Ionization. Docking the sulphone amide based inhibitors without deprotonating the amide results in incorrectly predicted binding modes, with failed modeling of the of the zinc ion coordination. The nitrogen must be negatively charged to orient the ligand in the cavity and model the coordination of zinc ion correctly. A number of the docked ionized ligands still fail to coordinate the zinc ion with the sulphone amide nitrogen, but coordinate the ion with one or both of the sulphone oxygen atoms. As in the correct structure the amide nitrogen forms a hydrogen bond to Thr199, but with the backbone amide to which one of the oxygens normally forms a hydrogen bond with, as opposed to the side chain hydroxyl group. The sulphone amide moiety is tilted sideways thus exposing the oxygens to the zinc ion. Figure 4 shows two docked inhibitors displaying different modes of zinc ion coordination. A possible explanation for the occasionally incorrectly modeled zinc ion coordination could be that the difference in calculated energy between the nitrogen and the oxygens coordinating the zinc ion, due to the similar hydrogen bond formation, could be very small. As a result of the conformational sampling being limited the correct conformation with respect to coordination of the zinc ion might not be suggested in combination with correct conformation regarding other parts of the molecule.

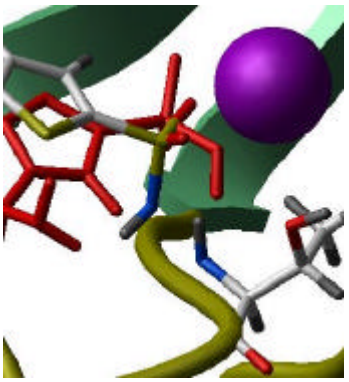


Figure 4. The zinc ion is correctly coordinated by the sulphone amide nitrogen H (red) forming a hydrogen bond with the hydroxyl group O of Thr199. The incorrect zinc ion coordination mode by both of the sulphone oxygens with the nitrogen N forming a hydrogen bond with the Thr199 backbone amide H is depicted in gray.

Including solvent molecules. Ionization of ligands results in correctly oriented ligands with correctly modeled zinc coordination for some inhibitors, however, in contrast to straight/relaxed conformations, as displayed in Figure 3, the docked ionized molecules adopt strained, bent conformations in the large hydrophobic cavity. Figure 5 shows the ionized 1bn1 ligand docked to the 1bn1 protein structure.

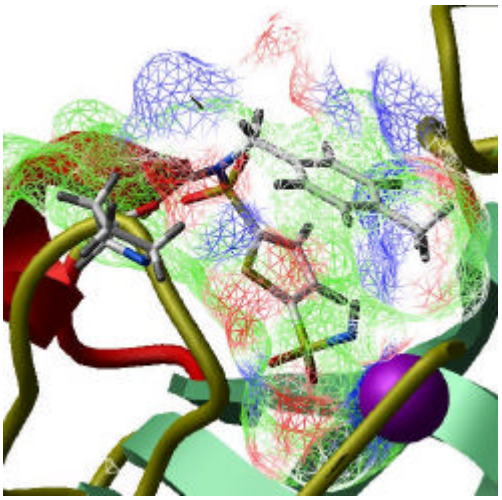


Figure 5. The ionized 1bn1 ligand docked to the 1bn1 protein structure. The zinc ion is correctly coordinated but the ligand adopts a bent conformation. The second sulphone amide nitrogen H interacts with the backbone carbonyl of a proline residue.

By including two water molecules in the docking calculations the volume where the aliphatic end of the inhibitor molecules enters in the strained conformation is excluded, as depicted in Figure 3, forcing the molecules to adopt unbent conformations. The obtained strained ligand conformations seem to be a result of the mainly hydrophobic environment surrounding the second sulphone amide in the relaxed structure. In the bent conformation the second sulphone amide nitrogen can form hydrogen bonds with the protein, as displayed in Figure 5. Analysis of the bent ligand conformations indicates that they are not high-energy conformations, giving yet another explanation for the occurrence of the bent conformations and indicating that ICM

does not fail in generating/distinguishing reasonable low-energy ligand conformations. Including the "correct" fixed solvent molecule meant to interact with the oxygen of the second sulphone amide results in straight ligand conformations. However, the position of this water molecule differs somewhat between the complexes and the hydrogen bond formation results in slightly wrong positioned ligands in a number of cases. Docking ionized ligands including two water molecules as a steric hinder, results in generally well-predicted binding modes, the ligands displaying reasonable interactions with the receptor.

van der Waals interactions. Possibly due to the lack of polar interactions stabilizing the ligands, aliphatic tails, are slightly tilted. The weak van der Waals interactions may result in a flexibility of the aliphatic tail, the correct structure representing one of several possible modes. Another explanation could be underestimation of van der Waals interactions during docking calculations.

Hydrogen bond formation vs. electrostatic interactions. Docking results display examples where incorrect binding mode is presented due to wrongly formed hydrogen bonds substituting correct electrostatic interactions. This could be a result of overestimation of hydrogen bond energy.

Instability of docking results. Re-docking the same ligands to the same protein structure gives rise to differences in docking results as well as score, between consecutive docking calculations. As a probable effect of insufficient sampling during Monte Carlo minimization, the number of Monte Carlo steps can be increased to stabilize the docking results. This is shown to be sufficient for a major part of the inhibitor set, however, results show that for the 1i91 and 1bnq ligands the docking results are insTable even after increasing the Monte Carlo steps by a factor of ten. The simulation time is thereby increased from approximately 2 minutes to 15 minutes per ligand, which is unacceptably high for practical purposes.

RMS deviation values between the docked and the correct structures, obtained from complex crystal structures, are summarized in Table 2.

4.2.4 Matrix metallo proteases

Matrix metallo proteases are zinc dependent endoproteases involved in the degradation of components of the connective tissue in the extra cellular matrix [28, 29,30, 31]. The active site of the MMPs is located in a large cleft and consists of a number of specificity pockets of variable size. The specificity pockets are designated S1', S2', S3', S1, S2, S3, where the prime corresponds to binding of the carboxyl end of the substrate. Most inhibitors of MMPs bind prominently in the prime sub-sites. The S1' site differs significantly between MMP-1

and MMP-3. MMP-1 has a small and hydrophilic S1' pocket, whereas MMP-3's S1' site is a large and open pocket with mainly hydrophobic residues. MMPs display a catalytic zinc ion, coordinated by three histidine residues and a water molecule substituted by the ligand in inhibitor complexes [28, 29, 30, 31].

4.2.4.1 MMP-1

For a detailed description of the ligands, see appendix 1. The hydroxamate nitrogen interacts with the carbonyl oxygen of an Ala residue, stabilizing the hydroxamate position. The P1' group, to be positioned in the S1' specificity pocket, is varied from small un-branched carbon compounds to the large bulky biphenyl group of ligand 966c. In most cases collagenase-1 displays a small S1' specificity pocket that cannot accommodate large P1' groups. To accommodate the biphenyl Arg214 at the back/bottom of the S1' sub-site is relocated when interacting with the ligand, i.e. induced fit, creating a larger more open pocket [28].

Figure 6 displays the 966c ligand structure superimposed on the 1hfc protein structure and the alternative position of Arg214 in the 966c protein structure.

4.2.4.2 MMP-1 docking results

The peptide-based inhibitors obtain correct binding mode with RMS deviation values from their correct structures of about 1Å, when docked to 1hfc. Due to steric hindrance, as a result of induced fit, the sulphone based hydroxamic acid ligand, 966c, is relocated in the active site.

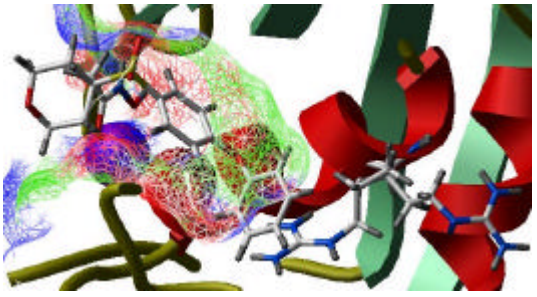


Figure 6. The correct 966c ligand structure superimposed on the active site of 1hfc. The biphenyl ether clashes with Arg214. The 1hfc position of Arg214 (left) as well as its position in the 966c structure (right) is displayed, illustrating the induced fit.

Docking to the 966c structure results in wrong location of the 1hfc ligand in the large S1' specificity pocket. A possible explanation is induced fit. The specificity pocket being much larger than necessary for the ligand, results in more possible binding modes and formation of "correct" interactions is complicated due to altered shape and distances. The 966c ligand is

located correctly in the S1'pocket but displays some differences from the correct structure in the S sub-sites and does not coordinate the zinc ion correctly. Ionizing the hydroxamic acid, i.e. assigning a negative charge to it, should assists in placing the hydroxamate in position to coordinate the catalytic zinc ion. Ionization of the hydroxamic acid group positions the hydroxamate in the vicinity of the zinc ion but does, however, not improve the overall docking results.

4.2.4.3 MMP-3

Docked ligand structures are displayed in appendix 1.

4.2.4.4 MMP-3 docking results

Docking of ligands displaying strongly diverse binding modes. A major part of the docked ligands display correct coordination of the catalytic zinc ion. Exceptions are the 1usn and 2usn ligands, coordinating the zinc ion with the exocyclic sulfur of thiadiazole. Explanations for this is the much diverse binding mode displayed by the ligands, compared to other known stromelysin inhibitors, as well as the charge distribution of the zinc coordinating group. Figure 7 displays the diverse binding modes of the 1usn and 1d8m ligands in the stromelysin active site.

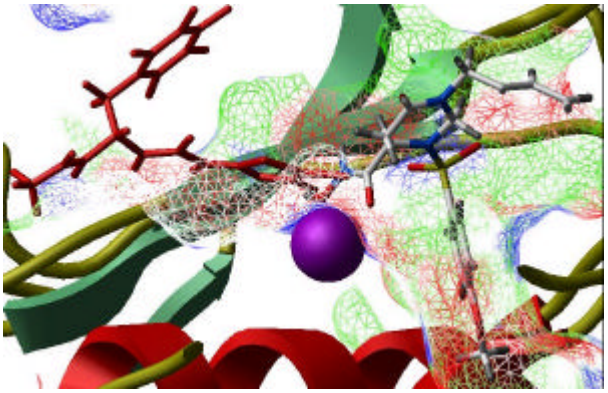


Figure 7. Diverse binding modes of the 1usn (red) and 1d8m ligands in the MMP-3 active site. The 1usn ligand interacts primarily with residues in the S sub-sites in contrast to other known stromelysin inhibitors, represented by 1d8m, binding in the S' pockets. 1d8m coordinates the catalytic zinc ion with its hydroxamic acid moiety. The aromatic ring inserts into the S1' specificity pocket.

In contrast to other ligands, these inhibitors bind in the more open non-prime side of the active site, the S sub-sites. The S1' pocket being empty adopts a closed conformation shifting Tyr223 at the back of the S1' pocket more than 7 Å inwards affecting the backbone position, and creating an entirely different active site conformation [30]. Important interactions for the non-prime site binding ligands occur in the S3 sub-site where the P3 group of the ligands, a

pentafluorophenyl and a phenyl group respectively, forms hydrophobic interactions with the side chain of Tyr155 and display π-π stacking with the Tyr side chain. In addition, the group interacts with a conserved buried water molecule [30]. Since π-π stacking is not simulated within the docking program, no grid map describing this type of interaction being present, the water molecule must be included to anchor the 1usn ligand to the S3 sub-site. To successfully dock the 1usn ligands, the docking must further be performed to a receptor structure with correct conformation, preventing the ligands to enter into the prime sub-sites of the receptor. The π-π stacking is substituted by an electrostatic interaction, with the rings halfway overlapping. Figure 8 displays the docked 1usn ligand. To obtain correct coordination of the zinc ion the thiadiazole group must be correctly modeled with the exocyclic sulfur partially negatively charged. It should be noted that including water molecules for specific interactions can result in incorrectly formed interactions due to the positioning of the polar hydrogens as well as differences in solvent molecule positions between complexes, as was shown in docking to carbonic anhydrase.

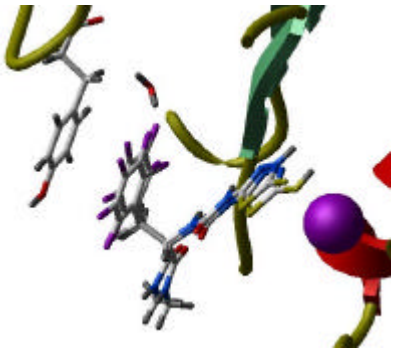


Figure 8. The 1usn ligand docked to the 1usn receptor structure. One water molecule is included. The pentafluorophenyl of the docked molecule is slightly tilted/turned, interacting electrostatically with the Tyr in contrast to the correct structure displaying π-π stacking with the aromatic ring.

Induced fit. A majority of the remaining non-peptide inhibitors obtain reasonable binding modes when docked to the hydroxamate type inhibitor, 1b3d. Some inhibitors, displaying large P and P1' groups, do not dock correctly to the 1b3d structure. Docking to the 1hy7 crystal structure improves docking results for some ligands displaying large P1'and P groups, e.g. 1g05, 1b8y, 1hy7 [31].

Ring conformations. The docking procedure does not involve sampling of diverse ring conformations, making it necessary to generate and subsequently dock molecules with multiple ring conformations for some ligands. A number of ligands, e.g. 1d8f, 1g49, with the hydroxamic acid group directly bound to a ring structure, do not obtain correct binding orientation with wrong ring conformation, due to inability to coordinate the zinc ion. It should be mentioned that overall good docking results can be achieved even when docking molecules with wrong ring conformation for some molecules.

Docking large flexible molecules. The peptide derived inhibitors all display wrong binding modes whether docked to the 1b3d structure or the 1hy7 structure. They fail to coordinate the zinc ion with their respective moieties, carboxylic acid and thiol. The peptide derived ligands are large, branched molecules, and the reason for the inability to dock them correctly could be the induced fit phenomenon. Other explanations could be insufficient sampling due to the large amount of variable torsion angles, or the large amount of interaction possibilities, possibly creating different complexes with similar energy.
Ionization. The carboxylic acids must be ionized to obtain correct docking results, displaying correct coordination mode of the catalytic zinc ion. Ionizing hydroxamic acids, however, does not improve docking results. Docking results indicate that thiol groups should not be deprotonated, in this particular environment, but the results are somewhat dubious since other difficulties obtaining correct binding mode for the thiol containing ligands are observed. Protonating amines do not seem to improve docking results for the tested stromelysin ligands, which is unexpected considering the charge assigned to the molecule as well as the altered hydrogen bond formation possibilities.

When performing docking calculations in consecutive runs some inconsistency of docking results was observed, a more detailed analysis is presented in the scoring function section.

RMS deviation between the correct and docked structures is displayed in Table 2.

4.2.5 Serine proteases

Trypsin and thrombin are serine proteases involved in the blood coagulation cascade. The proteins share many features and display similar interactions with their respective ligands. [32, 33, 34, 24]

The P1 moieties, binding in the S1 specificity pocket, of the docked inhibitor set are positively charged amidino- guadinino or 4amino pyridine groups forming hydrogen bonds and ionic interaction with Asp189. The P2 group inserting into the small hydrophobic S2 pocket consists of substituted aliphatic rings in chair conformation. The P3 group is the equivalent of the main chain of the peptide substrates of thrombin. The P4 residue, normally inserted into the aryl-binding pocket, constitutes of large aromatic groups displaying favorable perpendicular orientation to the side chain of Trp215. The inhibitors are generally large flexible molecules, the set of the docked inhibitors containing both branched and un-branched ligands. Some of the docked ligands display enantiomeric discrimination and ring conformation constraints. [32, 33, 34, 24]

For a more detailed description of the docked ligands, see appendix 1. Figure 9 shows the structure of 1etr, MQPA, in complex with thrombin.

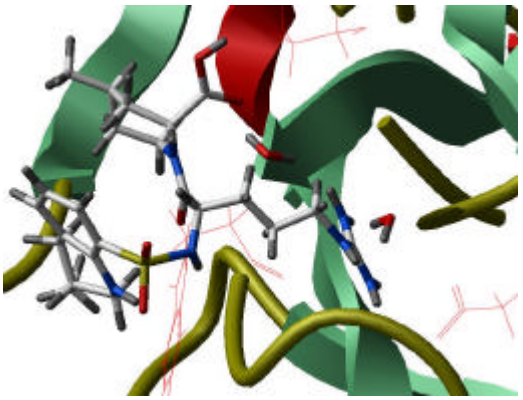


Figure 9. The 1etr thrombin complex structure. Two water molecules are included. The guanidino group interacts with Asp 189 in the S1 sub-site. The axially positioned hydroxyl of the piperidine ring in the S2 pocket interacts with Ser195. The naphtyl moiety in the aryl binding site is positioned perpendicular to Trp215.

For the majority of guanidinyl ligands hydrogen/ionic bonds with Asp189 are only formed with one of the terminal nitrogens of the guanidino group, resulting in an uneven charge distribution. Diverse from the guanidinyl inhibitors both the amidino nitrogens of a major part of the benzamidinyl inhibitors interact with the two Asp189 oxygens, displaying an even charge distribution. The un-branched 1uvt ligand represents a different type of thrombin inhibitor, with a positively charged 4-aminopyridine moiety inserted into the S1 pocket. The polar 4-amino group, positioned in a mainly hydrophobic environment gives rise to a conformational change of the protein structure to facilitate interaction with Glu192. Enantiomer type plays an important role for some of the ligands, e.g. the thrombin inhibitors of 1ets and 1ett, displaying binding with only one of the stereoisomers. Carboxylate groups on substituted aliphatic rings binding in the S2 pocket form hydrogen bonds to Ser195 in thrombin as well as trypsin. To achieve this the carboxylate group must be oriented in an axial direction as displayed in Figure 9 [32, 33, 34, 24].

4.2.5.1 Docking results

The serine protease docking results indicate some measures to be taken prior to docking as well as conditions where the ICM docking program fails to model the systems correctly.

Enantiomer discrimination. Molecules, in particular large, bulky ligands, with several enantiomeric forms must be constructed in correct enantiomeric conformation prior to docking. Steric hindrance or interactions impossible to form due to steric hindrance will result

in wrong binding modes for non-binding enantiomeric form of the ligands. Representative examples are the thrombin inhibitors of structures 1ets and 1ett.

Charge distribution. Depending on atom type different partial charges will be assigned to the molecule. ICM assigns evenly distributed partial charges when modeling the amidino group with one double and one single bonded nitrogen. The assigned partial charges are equal to modeling the amidino group with two single bonded nitrogens and a carbo cat ion. Modeling of unevenly distributed partial charges is thus harder to accomplish. The effect on docking of the charge distribution is unclear, due to observed inconsistency of results when docking with the two models of the amidino group mentioned above, which should be equal. The explanation can be the size and flexibility of the molecules, i. e. insufficient sampling. Figure 10 displays the 1ets ligand docking results, with the amidino group represented by a carbo cat ion and two single bonded nitrogens.

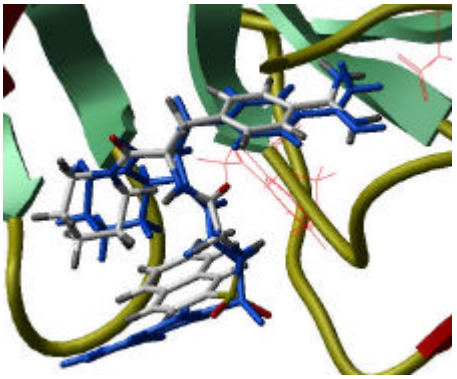


Figure 10. The docked 1ets ligand with the terminal charge represented by a carbo cation. The correct structure is colored blue. The P1, P2 and P3 groups are positioned correctly. The large naphtyl group is turned ~90 °, however still showing favorable perpendicular orientation towards the Trp215 side chain.

Including fixed solvent molecules. Water molecules involved in hydrogen bond networks between the receptor and the ligand must often be included in the calculations. Docking of the 1etr ligand without including the two water molecules in the calculations results in incorrect binding mode. The second terminal guanidine nitrogen points out of the pocket towards bulk solvent. Including the two fixed water molecules results in a correctly bent guanidino group. The hydrogen bond formation between the water molecule and the second terminal guanidino nitrogen could be important for positioning the P1 group correctly into the S1 specificity pocket, as illustrated in Figure 11. The thrombin inhibitor 1uvu displays better docking results including one fixed solvent molecule, indicating the difficulties when docking a set of different compounds to one protein.

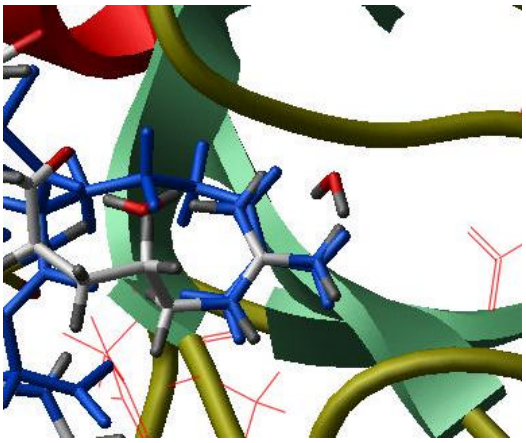


Figure 11. Two examples of docking results in the S1 sub-site of 1etr including and excluding two water molecules in the calculations. Incorrectly docked ligand is depicted in blue. The incorrect structure is positioned perpendicular to the correct one with one of the terminal nitrogens towards solvent in contrast to inwards interacting with one of the water molecules.

Docking ligands with correct /or several alternative ring conformation/s. Rings containing one or several heteroatoms or substituents, perhaps involved in specific interactions with the receptor, must be docked in correct or several alternative ring conformations. Docking of the 1etr ligand with the piperidine in chair conformation but with substituents in wrong orientations results in wrong overall binding mode and large deviations from the correct structure. In particular, the interaction between the carboxylate substituent and Ser195 in 1etr cannot be formed with wrong ring conformation. Docking with correct ring conformation results in an overall correctly docked molecule, even without anchoring the guanidinyl with the water molecules described above, see Figure 12 for details.

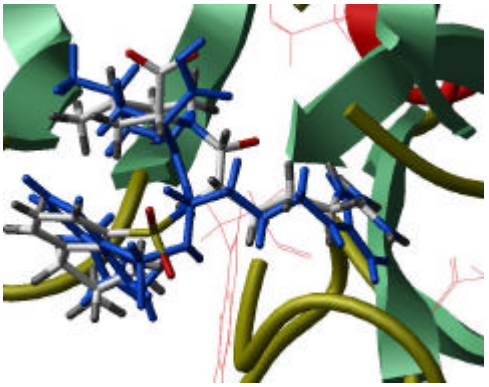


Figure 12. Correctly docked ligand 1etr to the 1etr receptor structure. Correct structure colored in blue. With the carboxylate of the piperidine ring placed in an axial position putting the group in an energetically and sterically favorable position an overall correct structure is obtained even without including the water molecules in the calculations.

Ionization. Correctly modeled ionization state of molecules is important for obtaining reasonable docking results. Dielectric fit, i.e. change of protonation state upon complex formation, as illustrated by the carboxyl acid H of the piperidine ring forming a hydrogen bond with the Ser195 hydroxyl group O in the S2 sub-site, complicates “preparation” of the ligands to be docked [24]. Modeling the ligands correctly with an automatic ionization

procedure is in these cases impossible. As is seen with the docked 1etr ligand the impact of the failed hydrogen bond formation can be minor, probably due to favorable electrostatic environment.

Induced fit. Docking the 1uvt ligand to the 1etr crystal structure results in completely failed docking. Due to the rotation of the peptide plane mentioned above, the hydrogen bond between Glu192 and the 4-amino group cannot be formed and since no other polar interactions exist outside the S1 pocket, anchoring the ligand to Glu192 is essential for obtaining correct binding mode.

RMS deviation values for the docked thrombin and trypsin ligands with respect to the correct structures are shown in Table 2.

4.2.6 MAP kinase, P38

The mitogen-activated protein (MAP) kinase P38 is an important signaling molecule, involved in production of cytokines during inflammation. The inhibitors, displayed in appendix 1, bind in the ATP-binding site. The crystal structures are low-resolution structures [35].

4.2.6.1 Docking results

The docking of kinase inhibitors to P38 displays generally poor docking results, with incorrect ligand orientation and wrongly formed or neglected hydrogen bonds. The poor quality of the structures as well as the rather large amount of non-polar interactions could be explanations for the docking failure. The ligands are not particularly flexible; insufficient sampling should thus not be an issue in the case of docking to P38.

Protein (PDB accession number)	Ligand (PDB accession number)	RMSD (Å)	Protein (PDB accession number)	Ligand (PDB accession number)	RMSD (Å)
ER agonist (1ere)	1qku	0.43	Carbonic anhydrase II (1bn1)	1i90	0.69
	1ere	0.54	(1i91)	1i91	2.00
	3erd	1.61		1bnv	1.75
ER antagonist (1err)	3ert	1.85		1bnt	1.11
	1err			1bnu	0.97
RARγ (1exa)	1fcy	0.49		1bnq	0.78
	1fcz	0.47		1a42	
	1exx	0.46		1bnm	1.80
	4lbd	1.15		1bn1	1.60
	2lbd	0.74		1i8z	3.47
	1fcx	0.49		1bnw	1.46
	1exa	0.33	(1i91)	1bn4	2.45
	3lbd	0.76	Thrombin (1etr)	1etr	1.10
MMP-1 (1hfc)	2tcl	1.10		1ett	1.29
	1hfc	1.18		1ets	2.89
MMP-1 (966c)	966c	2.35	1uvt	1uvt	1.24
MMP-3 (1b3d)	1g4k	2.54		1uvu	2.58
	1hfs	1.98	Trypsin (1k1m)	1k1m	1.70
	1d8m	1.56		1mts	1.84
	1d7x	2.27		1k1p	0.71
	1b3d	1.38		1f0u	2.93
	1d8f	2.77		1qb6	
1hy7	1hy7	2.39		1k1l	1.59
	1bqo	1.69		1f0t	
1b3d	1g05	2.77		1k1j	2.47
	1biw	1.23		1az8	1.22
	1d5j	3.55		1k1n	2.57
1hy7	1b8y	2.18		1bty	1.93
1usn	2usn	6.26		1k1i	3.00
1hy7	1g49	1.15		1zzz	5.40
1usn	1usn	0.55		1k1o	1.40

Table 2. RMS deviations between correct ligand structures, obtained from crystal structures, and docked ligands. Protein structures to which specific ligands have been docked are marked in parenthesis.

Ligand	Global minimum (kJ/mol)	Energy Correct (kJ/mol)	Energy ICM (kJ/mol)	global min (kJ/mol)	min (kJ/mol)	Correct (kJ/mol)
1b3d	184.3621	210.854	228.387	26.4919	44.0249	17.533
1b8y	-244.041	-223.423	-234.533	20.618	9.508	-11.11
1biw	58.63832	80.469	118.837	21.83068	60.19868	38.368
1bn1	-548.7178	-533.256	-539.482	15.4618	9.2358	-6.226
1bn4	-549.255	-537.661	-526.602	11.594	22.653	11.059
1bnm	-184.8193	-175.12	-172.386	9.6993	12.4333	2.734
1bnq	-264.4059	-230.406	-210.768	33.9999	53.6379	19.638
1bnt	-184	-164.88	-166.359	19.12	17.641	-1.479
1bnu	-379.2401	-365.581	-373.742	13.6591	5.4981	-8.161
1bnv	-175.3906	-160.93	-162.173	14.4606	13.2176	-1.243
1bnw	-638.4528	-618.171	-617.003	20.2818	21.4498	1.168
1c3i	322.4784	346.264	393.167	23.7856	70.6886	46.903
1caq	-181.1566	-146.247	-161.571	34.9096	19.5856	-15.324
1d5j	168.2579	236.186	189.98	67.9281	21.7221	-46.206
1d8f	15.85995	33.842	61.534	17.98205	45.67405	27.692
1d8m	-394.3281	-331.646	-340.045	62.6821	54.2831	-8.399
1ere	197.9	198.443	198.461	0.543	0.561	0.018
1err	252.8399	281.773	280.041	28.9331	27.2011	-1.732
1ett	-251.9153	-236.027	-227.999	15.8883	23.9163	8.028
1exa	125.3051	152.679	152.629	27.3739	27.3239	-0.05
1exx	125.2603	143.982	141.287	18.7217	16.0267	-2.695
1fcx	88.08007	94.541	93.777	6.46093	5.69693	-0.764
1fcy	164.0578	165.467	166.582	1.4092	2.5242	1.115
1fcz	89.00036	103.591	102.604	14.59064	13.60364	-0.987
1g05	-385.8404	-351.192	-337.94	34.6484	47.9004	13.252
1g49	-323.3552	-259.874	-266.295	63.4812	57.0602	-6.421
1g4k	-305.1378	-296.803	-290.164	8.3348	14.9738	6.639
1hfc	-1.119897	18.203	28.724	19.322897	29.843897	10.521
1hfs	-252.1404	-230.729	-216.307	21.4114	35.8334	14.422
1hy7	-190.3644	-160.697	-148.181	29.6674	42.1834	12.516
1i90	-341.0965	-332.305	-334.84	8.7915	6.2565	-2.535
1i91	24.52798	101.059	86.355	76.53102	61.82702	-14.704
1qku	197.9	198.451	198.4	0.551	0.5	-0.051
1sln	-520.676	-482.332	-465.502	38.344	55.174	16.83
1usn	-468.191	-452.608	-449.869	15.583	18.322	2.739
1uvt	-18.34403	-2.171	-4.361	16.17303	13.98303	-2.19
1uvu	-130.2101	-123.896	-128.696	6.3141	1.5141	-4.8
2lbd	-207.0278	-201.48	-202.457	5.5478	4.5708	-0.977
2usn	-492.9697	-461.71	-469.512	31.2597	23.4577	-7.802
3erd	64.93	66.879	71.103	1.949	6.173	4.224
3ert	327.8429	344.39	345.284	16.5471	17.4411	0.894
3lbd	-206.5296	-197.408	-200.198	9.1216	6.3316	-2.79
4lbd	125.2792	155.748	151.317	30.4688	26.0378	-4.431
966c	174.6399	187.978	183.549	13.3381	8.9091	-4.429
1bqo	-326.64	-293.689	-256.381	32.951	70.259	37.308

Table 3. Conformational energies of the correct ligand structures, obtained from crystal structures, the docked ligand structures and the global minimum of the free ligand in solution. Differences in energy between correct and docked structures respectively and the global minimum as well as between the correct and docked structures are also shown.

4.3 Analysis of conformations generated in ICM

The results of the conformational searches and the constrained energy minimizations of the docked ligand and crystal structure conformations are presented in Table 3. The threshold energy value for low-energy conformations is set to 15kJ/mol above the global energy minimum. A major part of the correct as well as the docked structures display energy values higher than 15 kJ/mol above the global energy minimum, i.e. are high-energy conformations according to the pre-set definition. This is in contrast to *Boström et al* [16] stating that the conformational energy penalties for ligand binding in 70 % of the crystal structure complexes studied ≤ 3.0 kcal/mol, which corresponds to 12.6 kJ/mol. The same parameters for constrained relaxation of the docked and correct ligands structures have been applied. A few cases of high-energy ligand conformations obtained from crystal structures could be expected, due to structure solvation methods and low-resolution structures. The reason for the deviating results, with a large part of the ligand conformations obtained from crystal structures displaying high-energy values is, however, unclear. The docked conformations similar in energy to the obtained global energy minimum are mostly small, rather rigid molecules, e.g. ER and RARγ ligands. To be noted is that the inactive enantiomer of the RARγ ligand, 1exx, displays a lower energy than the active enantiomer, 1exa, in contrast to previously stated [26]. Both these observations indicate that the method for conformational energy comparisons used here is inaccurate. The energy difference between the crystal structure and docked ligand conformations is, however, in a majority of cases lower than 15 kJ/mol and in a number of cases small, indicating that conformations generated in ICM should not be entirely unreasonable.

4.4 Rigid ligand docking

In an attempt to improve docking results, docking with a set of low-energy conformations, obtained from the conformational search was studied. During the docking calculation the ligand is kept rigid, i.e. only rotational and translational sampling is done. The results indicate that slight improvements of binding mode is obtained, displaying better RMS deviation values from the correct structures. However, no major improvements of docking results can be expected. Other factors, as induced fit or ionization state of the ligands, seem to play more important roles for the general outcome of docking calculations. Docking a set of low-energy conformations rigidly can be used to study the details of a hypothetical interaction between a specific compound and a protein, but the usability of the method for general purposes seems limited. This observation is consistent with several studies showing the improved

performance of flexible ligand docking algorithms as compared to docking programs applying rigid ligand docking [1].

4.5 Flexible receptor calculations on stack conformations

The ICM docking program includes a method for relaxation of receptor side-chains for low-energy stack conformations generated during Monte Carlo minimization [5]. The method was tested for a number of docked ligands. As a full-atom receptor representation is required, the procedure is time-consuming and the improvement of the results seems to be limited. During the relaxation heavy ligand atom movements are constrained, i.e. no major difference in docking results can be expected.

4.6 Conclusions and discussion

The qualitative and quantitative evaluation of the ICM docking tool has revealed cases where the docking algorithm performs well as well as complex docking cases that cannot be solved with the present docking program. The usability and the reliability of the ICM docking program will therefore strongly depend on the aim of the docking calculation being performed as well as the properties of the system being studied. A number of preparative steps, regarding protein and ligand representations in the docking calculations, useful for improving docking results are also suggested.

4.6.1 Considerations regarding protein representations

4.6.1.1 Converting crystal structures to ICM objects

Crystal structures obtained from the Protein Data Bank [17] must be converted to ICM objects, i.e. Cartesian coordinates are converted to internal coordinates. Crystal structures, in most cases not containing any information about hydrogen positions, will be assigned hydrogen atoms where appropriate in the conversion process. Due to the directionality of hydrogen bond interactions, randomly assigned hydrogen orientations could result in a number of wrongly formed or neglected interactions. The accepted deviation for hydrogen bonds is set to 30-40° allowing for some inaccuracy of the hydrogen atom orientation. However, a better way of assigning hydrogen atoms to crystal structures is to consider favorable interactions for polar hydrogen atoms when orienting these in the protein structure. One negative effect is the inability of the docking program to simulate the flexibility of the

protein when interacting with the ligand. Favorably oriented hydrogen atoms in the unbound protein structure may in reality be re-oriented when interacting with the ligand, which is not simulated during docking calculations. The ICM command "Convert" assigns hydrogens randomly to crystal structures when converting them to ICM objects. In contrast, "ConvertObject" orients polar hydrogen atoms favorably [5]. In addition, the "ConvertObject" command considers both tautomeric forms of histidine residues in the protein structure and assigns the most favorable one, which can be advantageous, e.g. for docking against stromelysin. Although the effect of the "ConvertObject" as opposed to the "Convert" macro is difficult to estimate, the latter being a random process, the overall impression is that it should be more advantageous to consider polar hydrogen orientations in a protein structure than not to consider them at all.

4.6.1.2 Including water molecules in docking calculations

Water molecules in crystal structures are in most cases not included in docking calculations. The reason for this is the inaccuracy of positioning water molecules in fixed places when in reality the bulk solvent is a dynamic environment, which must be modeled in a molecular dynamics simulation. In some cases including fixed water molecule positions can be justified. A water molecule involved in a hydrogen bond network with the protein or with the protein and ligand in complex has a fairly fixed position and can be essential for anchoring the ligand in the binding pocket. It should be noted that the orientation of the included water molecules, and thus the occurrence of specific interactions, most probably will be random and the position of the water molecule can differ somewhat between complexes. There is a risk that including this type of solvent molecule can result in incorrectly formed interactions and thus be more disadvantageous than advantageous. The water molecules can still make small changes to the receptor environment and perhaps most importantly occlude certain volumes for the ligand to occupy. To improve docking results additionally it can be justifiable to include certain "unspecific" water molecules in the active site region serving as obstacles, sterically hindering the ligand to enter certain areas. Even if these solvent molecules are not involved in specific interactions with the ligand, it can be thermodynamically explained by the fact that energy must be provided to remove these solvent molecules when binding into the active site. As can be seen from the crystal structure of the complex the ligand does not enter these areas because the energy required to remove the solvent molecules is too high. This justifies the approximation of including these "non-fixed/unspecific" solvent molecules. Since the solvent cannot be simulated with the docking program some type of approximation is needed in certain cases, e.g. docking to carbonic anhydrase II, and including solvent molecules serving as obstacles is one such approximation. Another possibility of/an

additional explanation for the ligand not entering these areas can of course be bad fit of the ligand there, which the docking program fails to model. This type of improvement of docking results does require detailed information about the receptor and several receptor-ligand complexes. In reality, this type of improvement could only be made for docking studies of one protein with a very limited amount of similar ligands, where information about similar complexes is available. In a virtual screen, screening a large amount of compounds for new potential binders, it would be less useful, in particular if searching for new binding modes. To be able to model the solvent molecules correctly a molecular dynamics simulation would be necessary, which is both time consuming and computationally expensive, and therefore impossible for application in a virtual screen.

4.6.1.3 Induced fit

A number of examples of the importance of docking to the correct, or at least very similar model, of the complex structure have been encountered. Since specific interactions play an important role for obtaining correct binding mode, a slight displacement of a key residue in the protein can cause inability to dock correctly. The induced fit problem ranges from large steric effects, e.g. the agonist/antagonist ER receptor structures involving the displacement of helix 12, to smaller steric effects as for the MMP-3 inhibitors 1usn and 2usn. The induced fit problem also involves more subtle structural changes with only slight displacements of key residues, as shown by docking the smaller MMP-1 ligand 1hfc to the larger binding cavity of the 966c structure. Docking against a model of a rigid receptor represented by grid maps makes it impossible to simulate receptor flexibility and docking against several diverse structures of the same protein is advisable. Relaxing the structure after docking to obtain more favorable interactions/binding modes is shown to have little effect on docking results and is very time-consuming. Docking against several different crystal structures of the protein is probably the best way of simulating the protein flexibility, and can be done realistically if docking to a well-studied protein and investigating specific interactions between a limited amount of similar ligands and the protein. For a virtual screen one structure must be chosen, preferably a complex structure binding large ligands, to avoid steric hindrance. The fact that specific interactions might be impossible to simulate due to the differences in structure still remains. More possible binding modes, possibly with similar energy, can be found, and insufficient sampling can result in incorrectly modeled binding modes. The docking program evaluation also indicates that docking must be performed to crystal structures of proteins. Considering the impact of the induced fit phenomenon between crystal structures of different protein-ligand complexes, it should be virtually impossible to obtain reliable docking results when docking against homology models, displaying large inaccuracies. Similar problems are

encountered when docking to poor quality structures. Ideally, the protein would be modeled with flexible side chains in the docking calculations. Due to the computational expenses and requirement of high time-effectiveness this is, however, not possible with the present algorithms and computer power.

4.6.2 Considerations regarding ligand representation

4.6.2.1 Ligand ionization state

One observation made throughout the evaluation is the impact of the protonation state of the ligands to be docked. Carboxylic acids, in particular if interacting with positively charged groups, must be ionized to dock correctly, e.g. docking to stromelysin. The ionization state is, however, strongly dependant on the local environment and the pK_as will in reality change during interaction between the ligand and the receptor, i.e. dielectric fit. The dielectric fit can cause problems due to the ionization state of the ligand being determined prior to docking and the importance of having general ionization rules. It should be kept in mind, if docking specific compounds to a specific receptor, that the ionization state of groups can change due to the local environment, and can thus be altered if docking problems are encountered. For virtual screens a general ionization script based on information about the ligand only, must be used and the dielectric fit cannot be simulated. Unexpectedly, the ionization of amines and hydroxamic acids does not seem to have great impact on overall docking results in the studied systems, considering the introduction of an extra charge, which should form electrostatic interactions. In the case of tertiary amines, the group obtaining ability to act as a hydrogen bond acceptor as opposed to donor should alter the interaction with the receptor.

4.6.2.2 Enantiomer discrimination

Sampling only torsional and translational changes of the ligand, no enantiomer sampling is performed during the ICM docking calculations. The occurrence of enantiomer discrimination, due to steric or energetic factors, makes it necessary to consider generation of all enantiomeric forms of compounds to use for docking. Cases where different enantiomers of a compound both display binding but show differences in activity are more complicated to discriminate between. If the reason for the differences in activity is somewhat unfavorable interactions with the protein the differences should be mirrored in score values. If the activity differences are due to energetically unfavorable ligand conformations, a conformational search should be able to mark the high-energy conformation. As an example, the ICM

docking program fails in discriminating between the two RARγ ligands, one active and one inactive enantiomeric form, by score. When evaluating the energy of the docked ligand conformation, the inactive ligand seems to have a lower energy than the active one, in contrast with the expected, thus failing to discriminate between the inactive and active enantiomer. This indicates that the force fields display weaknesses.

4.6.2.3 Charge distribution

The choice of ligand atom types, determining the partial charges assigned to atoms, is shown to be of importance. Problems occur in particular for resonance structures, displaying delocalized electrons. The docking results show examples of dependency of the local environment in the complex for charge distributions in resonance structures, e.g. amidino groups in serine protease inhibitors. Some inhibitors show an expected even charge distribution, whereas interactions with the protein result in an uneven charge distribution for the same group in other inhibitors. Assigning of atom types and thus partial charges should be considered carefully in particular for systems with delocalized electrons.

4.6.2.4 Sampling of ring conformations

The ICM docking program does not sample different ring conformations during the Monte Carlo simulation [3]. The evaluation shows the importance of docking with correct ring conformation. The reasons for this is mainly due to steric factors, hindering the ligand to adopt correct conformation and orientation in the active site, but also inability to form correct interactions with wrong ring conformation. Docking with wrong ring conformation can result in overall good docking results, interactions with high directionality might however not be formed. Due to the fact that some compounds completely fail to dock correctly with wrong ring conformation, different ring conformations should be generated and docked separately to obtain reliable docking results. If studying a specific protein's interactions with a limited amount of similar ligands, the results of docking different ring conformations can be evaluated manually. To be used in a virtual screen the wrong ring conformations must be filtered out subsequent to docking, otherwise the amount of generated results to be studied becomes unrealistically large. The evaluation indicates that although the correct ring conformation does not always obtain the highest score in the set of the same docked ligand with different ring conformations, it usually displays a high ranking in the set.

4.6.3 Ligand – receptor interactions

A number of general observations considering different types of interactions have been observed throughout the evaluation. Electrostatic interactions involving perpendicularity of aromatic rings can result in slightly tilted or dispostioned rings. The docking program representing the aromatic rings by rings of negative and positive partial charges respectively, should favour perpendicularity of aromatic rings, even though the energy cost for a slightly tilted ring might not be too great. π-π interactions between aromatic rings, only possible to describe in quantum mechanical terms, are not modeled by the docking program. In contrast to overlapping parallel aromatic ring systems halfway overlapping rings are observed, i.e. π-π interactions are substituted by pure electrostatic interactions. Few polar interactions seems to result in a general "floppiness" of the part of the molecule interacting mainly hydrophobically with the protein. The reproducibility of results seems to be poorer and the deviations from the correct conformation larger for these parts of the molecules. The unspecific weak non-polar interactions can account for the observation. The energetic cost for a tilted hydrophobic tail of a molecule will be low or a number of diverse binding modes will result in similar overall energy. Insufficient sampling can contribute to the deviations. Conformations with specific "tail" conformations in combination with overall geometry and position might not be suggested in some docking calculations. In reality a hydrophobic tail of a molecule might not have a fixed position, as represented in the crystal structure, but could show flexibility and switch somewhat between conformations. Correlating docking results with electron density of crystal structures would give an indication as to whether some of the "docking failures" correspond to flexible parts of the molecules. Some results indicate that the hydrogen bonding energy contribution might be slightly overestimated, as exemplified by docking to carbonic anhydrase.

Can the ICM docking program be used as a tool in structure-based drug design? ICM docking results are generally accepTable, when correct measures have been taken prior to docking. The system must be well studied and docking to several different complex structure of the protein is advisable. Accuracy and consistency of ICM docking results will depend on the properties of the system being studied. Large hydrophobic binding cavities are more difficult to dock against, e.g. Carbonic anhydrase, than smaller pockets displaying specific interactions with the ligands, as RARγ or ER. Highly flexible ligands are more difficult to handle. Induced fit and inconsistency of docking results are the largest drawbacks of the docking program.

5 Virtual screening. Evaluation of scoring functions implemented in ICM

5.1 Introduction

A condition for a successful virtual screen in the lead discovery phase is well performing scoring functions, both able to discriminate binders from non-binders and preferably also rank binding compounds according to activity [9]. In order to evaluate the two scoring functions implemented in the ICM docking program, sets of compounds with known activity as well as randomly chosen inactive compounds were docked and their ranking and scores studied. Different approaches for determining score or rank thresholds discriminating between active and inactive compounds were tested. The robustness of the method as well as the reproducibility of the docking results, both considering binding mode and score, was studied. The scoring functions ability to rank active compounds according to activity was also evaluated.

The most straightforward way of analyzing the docking results of a large set of compounds, e.g. 100 000, containing no previously known active compounds, is by ranking the score values and scanning through a top percentage, usually 0.5-1 %, of the docking results, searching for potential hits [36]. Alternatively, an empirical score threshold can be determined, above which most active compounds of different systems can be found, typically an ICM score of –32 [5].

The PMF (Potential Mean Force) [8] scoring function is a knowledge-based scoring function implemented in ICM, which can be combined with the ICM score in a consensus scoring approach. The distribution of ICM- and PMF score for sets of compounds can be used to guide the decision of threshold score values. In an attempt to improve the hit rate Bayesian classifiers [13] were also constructed from docking results of the inactive compounds and the sets of inhibitors with known activities.

Due to difficulties in reproducing docking results, both in binding mode and score values, observed during evaluation of the docking tool, consecutive runs were performed for the inactive compounds as well as for some of the inhibitor sets. The instability of docking and scoring is probably an effect of insufficient sampling during Monte Carlo simulation [3], resulting in lower probability of finding the global energy minimum for some systems. The instability could possibly be related to the quality of the minimum. A non-smooth surface near the global energy minimum will make the minimum more difficult to reach if the number of Monte Carlo steps is finite. This could possibly be a result of docking of non-binders. If consecutive runs with different types of active compounds would show good reproducibility

whereas the results of docking of inactive compounds would be less sTable, the difference in score/binding mode could be used as an indication of non-binding. The observation could possibly be used as an additional filtering method when trying to discriminate active from inactive compounds and was evaluated.

Another explanation for difficulties reproducing docking results could be the induced fit phenomenon, i.e. the receptor flexibility during interaction with the ligand. Induced fit, an important factor for the outcome of docking different ligands to the same protein structure, will probably have an effect both on the amount of false negatives as well as the docking programs ability to reproduce results. To study the impact of induced fit on score values, compounds with known activity obtaining unusually low scores and showing poor reproducibility were docked against different structures of the same protein.

To determine whether the differences in docking results of consecutive runs for some compounds is related to the Monte Carlo and/or the local minimization step, the number of visits to different suggested energy minima during the simulation was studied. The number of visits to a particular energy minimum corresponds to the number of times a certain conformation, or a conformation very similar to that, has been suggested [5]. A conformation corresponding to an energy minimum visited a large number of times, can indicate that the random step meant to transport the conformation a large distance away from its current position in conformational space might be insufficient or not occur often enough during the simulation. Instead of covering a representative part of conformational space the simulation can "get stuck" in a local energy minimum. If the active compounds with insTable docking results display few visits to the corresponding energy minimum the sampling might be insufficient, resulting in difficulties reaching the same minimum in consecutive runs.

To optimize the set of ligand conformations multiple conformations of flexible rings were generated for active compounds obtaining low scores. Different ionization rules were also applied and the ligands were re-docked to the same structures to study the importance of protonation and ring conformation for score values and reproducibility.

The ICM score is the weighted sum of different energy terms calculated during docking [5]. The different contributions to the total energy for low scoring active compounds were studied in an attempt to explain which part of the energy function that fails to model the system accurately. For active compounds with large differences in score between two consecutive runs the Δ energy terms were calculated to determine whether one term consistently differs between runs.

5.2 Scoring function evaluation results

5.2.1 MMP-3

5.2.1.1 MMP-3 scoring results

Figure 13 shows the results of docking 51 stromelysin inhibitors and 1000 random, "inactive", compounds to the stromelysin structure 1hy7. In this particular case the active compounds form a rather distinct cluster, although not entirely separated from the inactive compounds. Choosing score threshold values ICM score ≤ -23.0 and PMF score ≤ -50.0, the consensus scoring approach results in identification of 38 of 51 active compounds and 22 of 1000 inactive compounds classified active. This corresponds to finding 74.5% of the actives when scanning through 5.8% of the docked compounds obtaining 36.1% false positives. The results are summarized in Table 4.

Set	C_{act}	C_{in}	$\%C_{act}$	$\%C_{in}$	$\%C_{tot}$	EF
Stromelysin						
Training set						
1	38	22	74.5	36.1	5.8	13
2	49	58	96.1	54.2	10.2	9
3	29	7	56.9	19.4	3.4	17
4	45	33	88.2	42.3	7.4	12
5	45	181	88	80.2	21.6	4
Test set						
4	7	33	36.8	82.5	3.9	9
6	13	67	68.4	83.8	7.8	8
ER						
7	21	7	70	25.0	2.7	26
4	28	45	93.3	61.6	6.1	13
Thrombin						
4	15	16	62.5	51.6	3.0	21

Table 4. Classification results based on scores of active and inactive compounds docked to stromelysin, ER and thrombin. C_{act} and C_{in} are number of active and inactive compounds obtained in the scanned set of compounds, $C_{tot.}$ Total number of inactive compounds is 1000. Total numbers of active compounds are 51 and 19 MMP-3 inhibitors, training set and test set respectively, 30 ER agonists and 24 thrombin inhibitors. EF is the enrichment factor calculated as the percentage of actives identified divided by the percentage of the total set where these actives are found, i. e. C_{tot}.

Classification methods:

1 ICMScore ≤ -23.0 and PMFScore ≤ -50.0

2 Bayesian classifier: $P_{act} = P_{inact} = 0.5$

3 Bayesian classifier: $P_{act} = 0.055$, $P_{inact} = 0.945$

4 Bayesian classifier: P_{act} = 0.2, P_{inact} =0.8

5 Ranking based on ICMScore

6 Bayesian classifier: P_{act} = 0.2, P_{inact} =0.8 and ICMScore ≤ -23.0 and PMFScore ≤ -100.0

7 ICMScore ≤ -35.0 and PMFScore ≤ -55.0

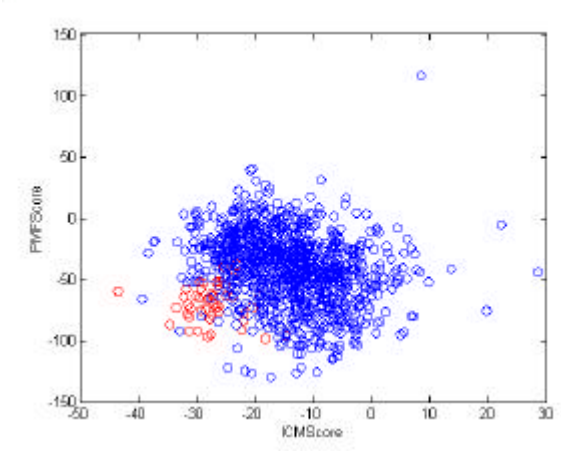


Figure 13. Scoring results of docking 51 active stromelysin inhibitors and 1000 inactive MDDR compounds to the 1hy7 MMP-3 crystal structure. Red=active compounds, Blue=MDDR compounds.

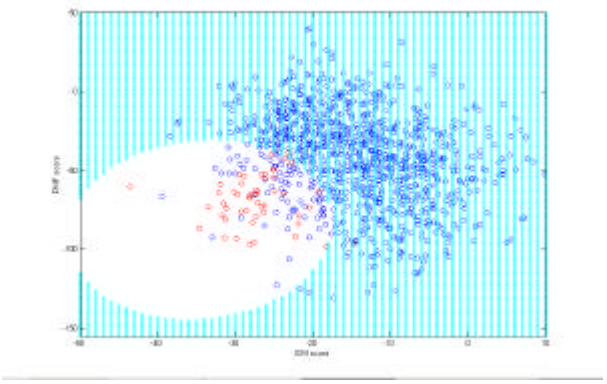


Figure 14. Bayesian classifier applied to scoring results of docking 51 active and 1000 inactive compounds to MMP-3. The a priori probabilities are set to 0.5 for inactive and active. The white area represents the area where active compounds are found.

The Bayesian classifier, constructed to improve the hit rate, is displayed in Figure 14 where the white ellipsoid area represents the area where active compounds can be found. Considering the ellipsoid shape of the area, fewer false negative hits can be expected. The a priori probability of each class, active and inactive, is here set to 0.5, which is not optimal, considering the large amount of inactive compounds compared to the number of active ones. This classifier identifies 96.1% of the actives when scanning through 10.2 % of the docked compounds obtaining 54.2% false positives. By optimizing the probability term, the area identifying "active" compounds can be narrowed down, resulting in fewer false positives combined with the loss of some true positives. An expected hit rate when screening a large set of compounds is 0.5 % [5]. In this particular case with 51 known actives the a priori probability of the active class would be 5.5%. A classifier based on this probability identifies

56.9% of the actives, when scanning through 3.4% of the docked compounds, obtaining 19.4% false positives. When no active compounds are previously known the a priori probability for identifying actives within the set is 0.5%. However, the score does not describe the interactions well enough, and is not reliable enough, to merely scan the expected 0.5 % of the results. Constructing the classifier with an a priori probability of 20% to identify actives might be more realistic, compensating for the instability and inaccuracy of the docking program and its scoring function.

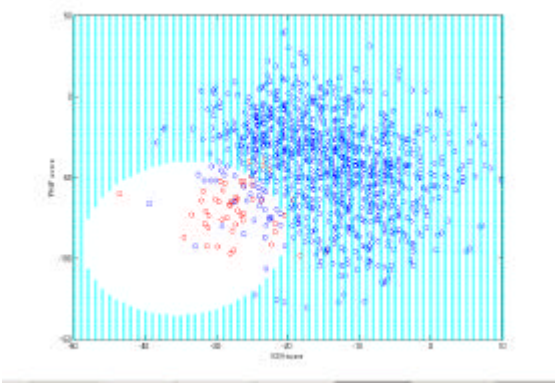


Figure 15. Bayesian classification applied to scoring results of docking 51 active 1000 inactive compounds to MMP-3. The a priori probabilities are set to 0.8 and 0.2 for inactive and active respectively. The white are represents the area where active compounds are found.

Figure 15 shows the Bayesian classifier constructed of results of docking to stromelysin with the a priori probabilities set to 0.8 and 0.2, inactive and active class respectively. This classifier identifies 45 of 51 active compounds and classifies 33 of 1000 MDDR compounds as active, corresponding to finding 88.2% of the actives when scanning 7.4% of the docking results obtaining 42.3% false positives.

To identify 88% of the actives merely using the ICM score to rank compounds, 21.6% of the docking results would have to be scanned containing 80.2% false positives, as displayed in Figure 16.

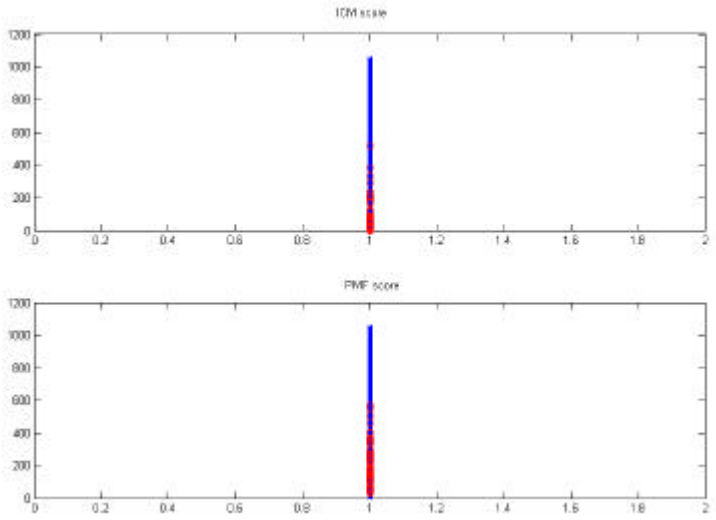


Figure 16. Ranking of compounds by ICM and PMF score. Blue = inactive compounds. Red = active compounds.

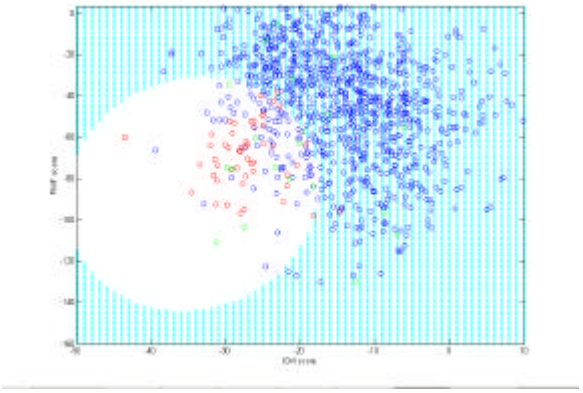


Figure 17. Bayesian classifier constructed based on docking results for 51 active biphenyl stromelysin inhibitors and 1000 inactive MDDR compounds, red and blue respectively. The white area represents the are where active compounds are found. 20 active test set inhibitors, green, are classified accordingly.

Figure 17 shows the docking results of a test set of 20 stromelysin inhibitors of known activity to the stromelysin structure 1hy7. The Bayesian classifier was constructed based on the docking results of the stromelysin inhibitor training set to the same structure. The probabilities are set to 0.8 and 0.2 for inactive and active class respectively. The classifier identifies 7 of 19 active compounds (1 of the 20 compounds is considered to be inactive) and classifies 33 of 1000 MDDR compounds as active. This corresponds to finding 36.8% of the actives when scanning 3.9% of the docking results, obtaining 82.5 % false positives. The very large amount of unidentified active compounds could be due to the induced fit phenomenon, i.e. the receptor flexibility when interacting with the ligand. For some of the false negatives only one of the scoring functions seems to fail in identifying the active compounds. Including compounds with ICM score < -23.0 and PMF score< -100.0 results in identification of 68.4%

actives when scanning 7.8% of the docking results obtaining 83.8% false positives. The results are summarized in Table 4.

5.2.1.2 Reproducibility : Re-docking compounds to the stromelysin structure

Figure 18 and 19 show the histograms of ΔICM score and ΔPMF score values for the MDDR compounds (blue, 154 compounds) and the stromelysin inhibitor training set (red, 51 compounds).

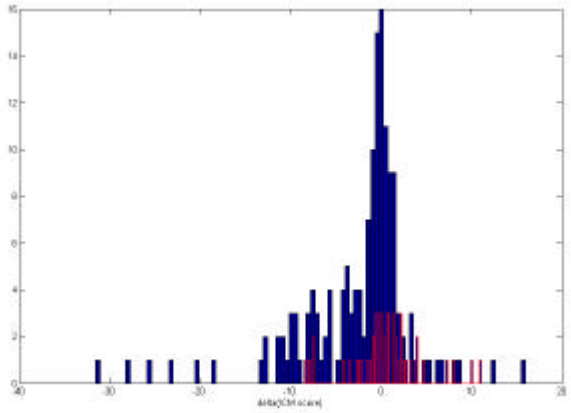


Figure 18. Histogram of difference in ICM score of two consecutive runs for 51 active (red) and 154 inactive (blue) compounds docked to MMP-3.

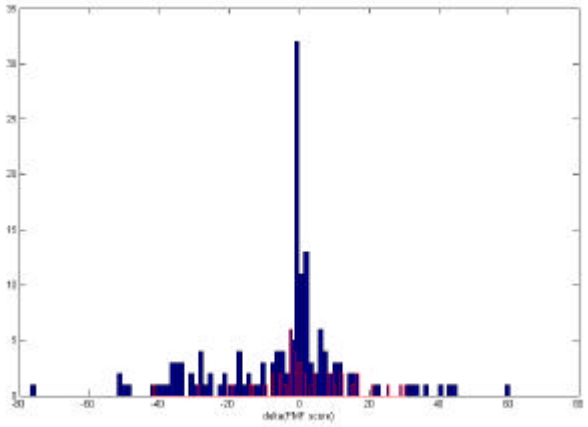


Figure 19. Histogram of difference in PMF score of two consecutive runs for 51 active (red) and 154 inactive (blue) compounds docked to MMP-3.

The average ΔICM score and Δ PMF score values of the inhibitor training set are slightly lower than the corresponding average values of the inactive compounds; 2.68 compared to 4.16 in Δ ICM score. The corresponding values for the PMF score are 9.03 compared to

13.03. The variance of the Δ PMF scores of the training set seems similar to that of the inactive compounds. The Δ ICM score values of the inactive set of compounds seem to vary more than the corresponding values of the inhibitor training set.

Re-docking of the inhibitor test set shows that the reproducibility of score values and binding modes for these compounds is poor. The average Δ ICM score and ΔPMF score are 3.63 (5.13) and 8.77(12.38) (additional run) and the delta score values vary over a large range.

The correlation between the difference in both ICM score and PMF score and a difference in binding mode was studied qualitatively. Approximate Δ score values that correlate with significant difference in binding mode were determined to ΔICM score ≥ 4 and ΔPMF score ≥10. Restricting classification of compounds to the active class to compounds classifying as active in two consecutive runs and applying the Δ score threshold values as an additional filter results in the hit rates displayed in Table 5. The Bayesian classifier was constructed with the probability parameters set to 0.5.

Filtering method	Docked compound	# classified active	Total number of compounds	# classified active run 1 and run2
1	MDDR (run 1)	58	1000	44
	MDDR (run 2)	45	154	
	Training set (run 1)	49	51	46
	Training set (run 2)	47	51	
	Test set (run 1)	10	20	6
	Test set (run 2)	6	20	
2	MDDR (run 1)	42	1000	31
	MDDR (run2)	37	154	
	Training set (run 1)	45	51	42
	Training set(run 2)	42	51	
	Test set (run 1)	7	20	5
	Test set (run 2)	5	20	
3	MDDR (run 1)	67	1000	60
	MDDR (run 2)	64	154	
	Training set (run 1)	50	51	48
	Training set (run 2)	49	51	
	Test set (run 1)	14	20	8
	Test set (run 2)	12	20	
4	MDDR (run 1)	46	1000	42
	MDDR (run 2)	49	154	
	Training set (run 1)	46	51	44
	Training set (run 2)	44	51	
	Test set (run 1)	10	20	7
	Test set (run 2)	11	20	

Table 5. Results of applying different filtering methods to scoring results of docking 51 and 20 inhibitors from training and test set respectively and 1000 MDDR compounds to MMP-3. Two consecutive runs were performed, the second run with only the best scoring MDDR compounds (154/1000).

1 Bayesian classification, 51 inhibitors used as training set, $P_{act}=P_{inact}=0.5$

2 Bayesian classification, 51 inhibitors used as training set, $P_{act}=P_{inact}=0.5$ and discarding compounds with ΔICM score ≥ 4 and ΔPMF score ≥ 10

3 Bayesian classification, 51 inhibitors used as training set, $P_{act}=P_{inact}=0.5$ and including compounds with ICM score < -23.0 or PMF score <- 100.0

4 Bayesian classification, 51 inhibitors used as training set, $P_{act}=P_{inact}=0.5$ and including compounds with ICM score< -23.0 or PMF score< -100.0 and discarding compounds with ΔICM score ≥ 4 and ΔPMF score ≥ 10

Based on the results of docking to the stromelysin structure, the best approach to identifying maximum number of active compounds of different types is applying consensus scoring and classification methods and including compounds displaying good ICM- or PMF score. This results in identification of 94.1% of the active training set compounds and 42.1% of the test set compounds. 10.8% of the docking results would have to be scanned containing 5.6% false positives. Discarding compounds with high Δ score does not improve the hit rate; the loss of

active compounds is significant. The fact that these compounds probably obtain different binding modes when docked several times still remains.

Docking to different stromelysin structures, in an attempt to overcome the effect of induced fit does not improve the hit rate. Single compounds can however obtain improved score values when docked to a different structure. The overall reproducibility does not seem to change.

Generation of multiple ring conformations and application of different ionization rules do not change the overall hit rate. Compounds with specific ring conformations do obtain improved scores, but the reproducibility seems poor.

Using the generalized Born approximation as an alternative representation of the electrostatic interaction energy does not improve the score values for the active compounds or the reproducibility of the docking and scoring results.

Studying the different energy terms contributing to the ICM score, the only terms that differ significantly when the Δ score for two consecutive runs is high are the electrostatic energy and van der Waals energy terms. Average values are about 3 compared to less than 0.5 for the other Δ energy terms. The van der Waals and electrostatic contributions to the score do not seem to be generally higher for compounds with large differences in scores. The electrostatic score does, however, always contribute with a positive component, i.e. has a negative effect on the score. The difference of the electrostatic energy term between runs seems to have greater impact on the ICM score than on the PMF score, which is expected. Similar results are obtained when comparing energy terms of consecutive docking calculations of the inhibitor training set to stromelysin.

Studying the number of visits to different energy minima in the conformational stack shows that a large number of visits to the best scoring conformation in the stack corresponds to well reproducible docking results, as expected. This is probably an effect of large, smooth energy minima. Few visits to the best scoring conformation and many visits to other conformations in the stack correspond to instability in reproducing docking results, a probable effect of insufficient sampling. Large energy differences between consecutive conformations in the stack in combination with approximately the same amount of visits to each conformation seem to be a sign of poor reproducibility. This indicates that increasing the number of Monte Carlo steps will not affect the reproducibility for these compounds.

5.2.2 Estrogen receptor (ER)

5.2.2.1 ER scoring results

Figure 20 shows the docking results (ICM-, PMF score) for 1000 inactive compounds and 30 ER agonists with known activity docked to the estrogen receptor structure, 1ere. A Bayesian classifier, with probability parameters set to 0.8 and 0.2, inactive and active respectively, identifies 28 of 30 active compounds and classifies 45 of 1000 MDDR compounds as active. This corresponds to finding 93.3% of the actives scanning through 6.1% of the docking results obtaining 61.6% false positives. The results are summarized in Table 4. Applying rigid score thresholds results in reduction of the amount of false positives combined with the loss of a number of true positives, binders. Applying the score thresholds ICM score ≤ -35.0 and PMF score ≤ -55.0 results in the identification of 70.0% of the active compounds when scanning through 2.7% of the docking results obtaining only 25.0% false positives.

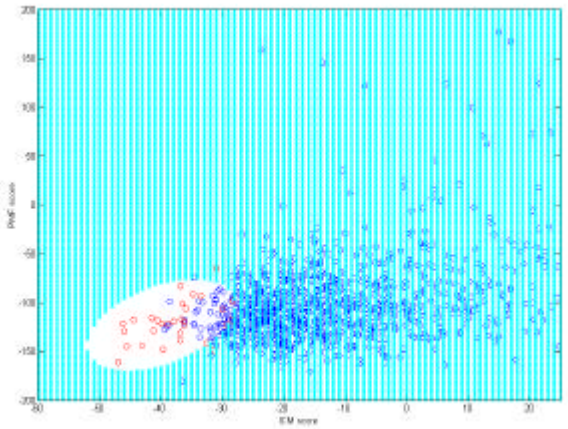


Figure 20. Bayesian classification applied to scoring results of docking 30 ER agonists (red) and 1000 inactive MDDR compounds (blue) to the ER structure 1ere. The white area represents the area where active compounds are found. The a priori probabilities are set to 0.8 and 0.2 for inactive and active respectively.

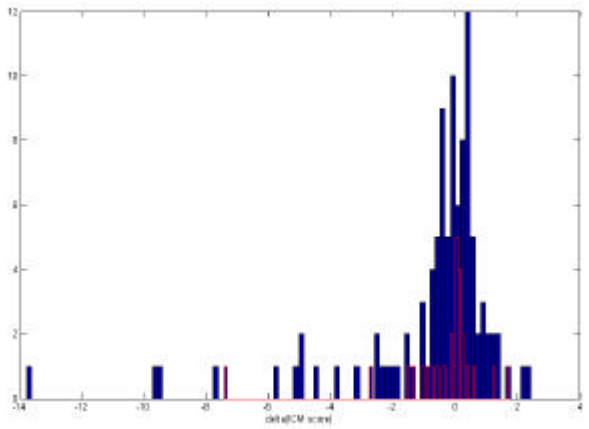


Figure 21. Histogram of difference in ICM score of two consecutive runs for 30 active (red) and 102 inactive (blue) compounds docked to ER.

5.2.2.2 Reproducibility : re-docking compounds to the ER structure

The differences in score values between consecutive runs display the same properties for docking to the estrogen receptor as for docking to stromelysin. The histograms of the difference in ICM and PMF score between two consecutive runs are displayed in Figure 21 and 22 respectively. The average value of Δ ICM score is somewhat higher for the inactive compounds than for the ER agonists. The Δ PMF score average value is approximately the same for inactive and active compounds. The average values for Δ ICM score and Δ PMF score are 1.32, 4.18 and 0.81, 4.16 for MDDR compounds and ER agonists respectively. Generally the difference in score between consecutive runs, for active as well as inactive compounds, seems to be lower for the estrogen receptor than for stromelysin. The estrogen receptor is most probably a more simple system to model. The ER active site is smaller, resulting in fewer possible binding modes and a more "sTable" system, than for the open active site cleft of stromelysin.

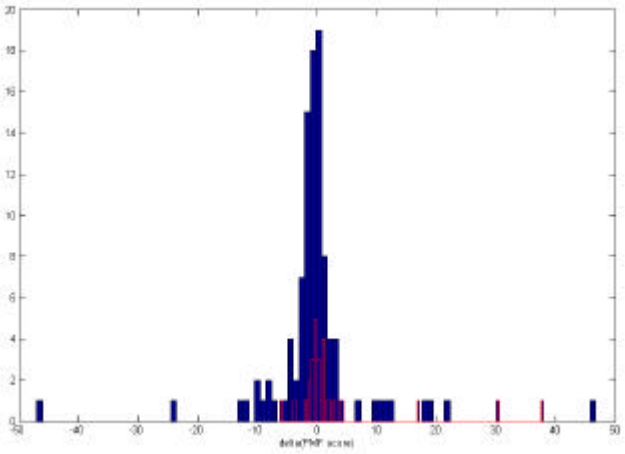


Figure 22. Histogram of difference in PMF score of two consecutive runs for 30 active (red) and 102 inactive (blue) compounds docked to ER.

The difference in energy terms contributing to the score for consecutive docking calculations of ER agonists are shown in Table 6.

Δ ICM	Δ Hbond	Δ Grid	Δ Ege	Δ Surf	Δ El	Δ Hp	Δ PMF
0.60	0.35	0.02	0.06	0.03	0.48	0.02	1.27
1.04	0.14	0.89	0.01	0.04	0.75	0.04	0.81
0.58	0.13	0.32	0.03	0.01	0.84	0.01	1.31
1.39	1.46	5.04	1.17	0.00	4.12	0.04	30.26
0.74	0.30	0.07	0.03	0.00	0.04	0.06	1.02
0.13	0.15	0.48	0.07	0.02	0.03	0.04	0.66
0.05	0.06	0.18	0.07	0.00	0.01	0.01	0.35
0.20	0.06	0.26	0.11	0.02	0.14	0.01	0.77
0.36	0.01	0.42	0.03	0.00	0.18	0.04	0.06
0.07	0.01	0.37	0.04	0.15	0.59	0.03	1.21
0.17	0.04	0.08	0.03	0.02	0.25	0.06	0.63
0.20	0.01	0.18	0.10	0.02	0.61	0.10	1.85
1.74	0.24	0.57	0.92	0.00	2.68	0.03	37.99
0.42	0.05	0.08	0.02	0.02	0.80	0.09	1.44
2.69	1.15	1.08	0.09	0.01	2.06	0.09	6.19
0.01	0.00	0.51	0.03	0.02	0.83	0.02	3.93
0.12	0.29	1.91	0.59	0.01	3.72	0.11	2.95
0.87	0.05	0.89	0.02	0.17	0.42	0.01	0.14
0.15	0.11	0.39	0.03	0.00	0.36	0.04	4.13
0.12	0.05	0.00	0.03	0.01	0.33	0.05	1.86
1.31	0.08	1.59	0.03	0.08	0.20	0.01	0.10
1.56	0.07	1.70	0.01	0.00	0.57	0.08	0.34
1.25	0.02	1.17	0.05	0.00	0.20	0.01	1.09
0.07	0.03	0.18	0.12	0.01	0.03	0.02	1.38
0.00	0.06	0.05	0.03	0.15	0.40	0.06	0.37
0.41	0.02	0.29	0.06	0.00	0.92	0.05	0.41
0.12	0.83	2.14	0.25	0.15	0.53	0.12	3.72
7.44	5.11	3.86	0.66	0.00	2.44	0.00	17.05
0.05	0.03	0.01	0.05	0.02	0.24	0.05	0.31
0.33	0.03	0.00	0.03	0.02	0.63	0.03	1.11
0.81	0.37	0.82	0.16	0.03	0.85	0.05	4.16

Table 6. Absolute differences in score of two consecutive runs for docking a set of 30 ER agonists to the ER structure, 1ere.

The only energy terms differing significantly are the van der Waals and electrostatic terms. No difference in contribution from these terms to the score for compounds with poor reproducibility can be detected. As observed for docking to stromelysin the electrostatic energy contribution is positive, i.e. has a negative effect on the score.

5.2.3 Thrombin

5.2.3.1 Thrombin scoring results

Figure 23 shows the docking results (ICM-,PMF score) for 1000 inactive compounds and 24 thrombin inhibitors with known activity docked to the thrombin structure 1kts. Due to the large variance in ICM score for the active compounds the area defining the active class becomes hyperbolic. Discarding active compounds with positive ICM score, the classifier is able to identify 15 of 24 active compounds and classifies 16 of 1000 MDDR compounds active, with the active probability parameter set to 0.2. This corresponds to finding 62.5% of the actives when scanning through 3.0% of the docking results with 51.6% false positives.

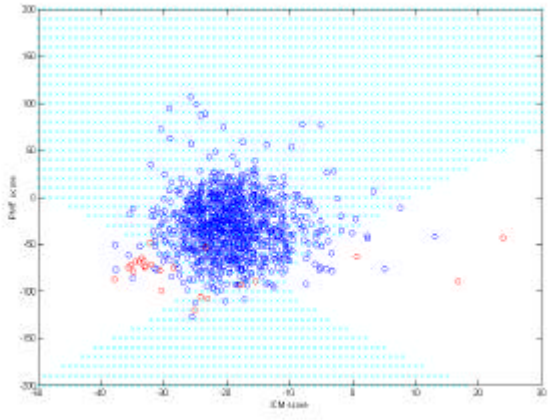


Figure 23. Bayesian classification applied to scoring results of docking 24 active (red) and 1000 inactive (blue) compounds to the thrombin structure, 1kts. The white area represents the area where active compounds are found.

The PMF score for the thrombin inhibitors is much more consistent and probably describes the system more accurately than the ICM score. The ICM score fails to discriminate between active and inactive compounds and displays instability with scores varying between –40 and 25 for active compounds from the same SAR set. A rigid score threshold would yield either a large amount of false positives or even fewer identified active compounds, which in the case of the Bayesian classifier already is at a minimum.

5.2.4 Score-activity relationship

The ICM score does not display any correlation with inhibitor activity. The PMF score – activity correlation factor is ~0.2. Results for docking ER agonists to the estrogen receptor are shown in Figure 24.

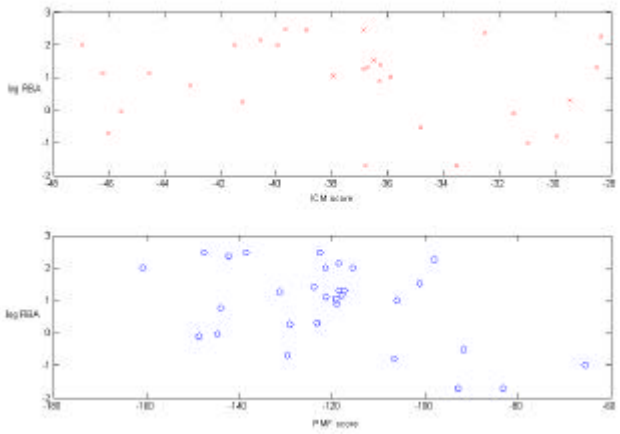


Figure 24. ICM and PMF score – activity relationship. The PMFScore – activity correlation factor is ~0.2. The ICM score – activity correlation factor is even less significant, ~0.03.

5.3 Conclusions and discussion

Using the ICM score combined with the PMF score implemented in ICM in a consensus scoring approach, whether by deciding on rigid thresholds or classifying compounds according to the Bayesian classification, reduces the number of false positives drastically. The amount of docking results to be scanned will thereby decrease considerably. The use of rigid score thresholds generally seems to reduce the amount of false positive hits more than the Bayesian classification method does, but at the cost of an increased amount of false negatives. The choice of classification method must therefore be dependent on the amount of false negatives that is considered accepTable. When applying optimal rigid score thresholds, that reduce the number of known false positives to a minimum, the thresholds vary between different systems, in particular the ICMScore (from -23 to -35).

If sets of active compounds are known prior to a virtual screen either a score threshold determination or a Bayesian classification can be performed and the results from the virtual screen classified accordingly. A condition for this is high reproducibility of score values for active compounds and relationship between scores of different types of active compounds when docked to the same structure. If no previously known active compounds exist the thresholds must be estimated from apparent distribution of compounds only, which will reduce the hit rate. However, the results indicate that applying consensus scoring on apparent distributions will produce a better hit rate than screening of a top percentage of ICM- or PMF score ranked compounds.

The scoring results display some cases where only one of the scoring functions seems to be able to distinguish an active compound; information that is lost when applying classifiers or a consensus scoring approach to the scoring results. The loss of these actives must be weighted against the increased amount of false positives obtained if altering the score thresholds. Application of a combination of classification methods and score thresholds could result in more actives being found, but requires a thoroughly studied system, and is probably unrealistic. Discarding compounds displaying poor reproducibility is shown to be inefficient.

There are two main applications of determining which compounds from a large set used in a virtual screen to extract for further studies. Extracting a small amount, usually not more than 1000 compounds, for further analysis to obtain more information about the system and set of ligands is an important application. The binding mode and interactions are analyzed qualitatively, discarding compounds with strained conformations or bad interactions with the receptor. The remaining ligand conformations can be used for construction of a SAR hypothesis or deciding upon discrete compounds to buy or synthesize for further studies. A quite different application is performing a virtual screen prior to HTS to discard a large subset of compounds where screening will be unnecessary. An alternative approach is screening a virtual library to decide which compounds to synthesize for subsequent screening. Neither approaches involve further qualitative studies of the docking results, but application of a simple filter based on score. The objectives of the two different applications are obtaining more knowledge of a system versus enrichment of hits in a screen. Depending on application quite diverse ways to determine which compounds to extract should be applied. The amount of true positive hits should be maximized for the first application, e.g. for constructing a SAR hypothesis, whereas losing a number of true positives can be considered accepTable for the second application, e.g. enrichment of hits in screening. A suiTable approach for the first case would therefore be applying the Bayesian classifier to the docking results, which includes a maximum number of binders. For the second application conservative rigid score thresholds could be considered to keep the number of inactive compounds at a minimum, but consequently also loosing a number of true actives in the process.

The poor reproducibility of docking results of some proteins is a major drawback for the interpretation and reliability of ICM generated virtual screening results. Increasing the number of steps in the Monte Carlo simulation will improve the overall consistency of docking results, but at a cost of time-effectiveness and might therefore not be a realistic approach. Studying the factors contributing to the scoring reproducibility also show that in some cases increasing the amount of Monte Carlo steps in the simulation will not be sufficient for obtaining well reproducible results. Some systems will probably have a higher level of

reproducibility than others, as is shown by the comparison of ER and MMP-3. The nature of the active site, i.e. size, accessibility and flexibility, are important factors. Before performing a virtual screen the system has to be closely studied and preferably docking to several different structures of the protein tried.

Determining the simplicity of the system prior to a virtual screen is important both considering the reliability of the docking results and the effects of induced fit, two of the major problems in using docking and scoring as tools in structure-based drug design.

Acknowledgements

I would like to thank my supervisor Micael Jacobsson for his support and guidance and the people at the Department of Structural Chemistry, Biovitrum, for making this an interesting and positive experience.

6 References

1 Abagyan R. and Totrov M., *Curr. Opin. Chem. Biol.*, **5**, 375, (2001).

2 Totrov M., Abagyan R., *Drug receptor thermodynamics: Introduction and applications, vol 1.* Edited by Raffa RB. New York: John Wiley & Sons, Ltd; 603, (2001).

3 Abagyan R., Totrov M. and Kuznetsov D., *J. Comp. Chem*, **15**, 488, (1994).

4 Totrov M., Abagyan R., *Proteins: Structure, Function and Genetics*, **Suppl.1**, 215, (1997).

5 Abagyan R., ICM manual v .2.8, Molsoft LLC, (2002).

6 Abagyan R., and Totrov M., *J. Mol. Biol.*, **235**, 983, (1994).

7 Totrov M., Abagyan R., *Biopolymers*, 60(2), 124, (2001).

8 Muegge I. And Martin Y. C., *J. Med. Chem.*, **42**, 791, (1999).

9 Stahl M. and Rarey M., *J. Med. Chem.*, **44**, 1035, (2001).

10 Charifson P. S., Corkerey J. J., Murcko M. A. and Walters W. P., *J. Med. Chem.*, **42**, 5100, (1999).

11 Wang R. and Wang S., *J. Chem. Inf. Comput. Sci.*, **41**, 1422, (2001).

12 SYBYL v. 6.8, Tripos Inc., (2002).

13 Duda, R. O. And Hart, P. E., *Pattern classification and scene analysis*, NY: Wiley, (1973).

14 Maestro v. 5.0, Schrödinger LLC, (2002).

15 MacroModel v. 8.0, Schrödinger LLC, (2002).

16 Boström J., Norrby P-O. and Liljefors T., *J. Comput.-Aided Mol. Design*, **12**, 383, (1998).

17 Bernstein F.C., Koetzle T.F., Williams G.J.B., *et al, J. Mol. Biol.*, **112**, 535, (1977).

18 ReliBase http://relibase.ebi.ac.uk/, 3 April 2002

19 CORINA v. 2.4, Molecular Networks GmbH

20 MDDR, MDL Information System Inc.

21 Muegge I. and Podlogar B. L., *Quant. Struct.-Act. Relat.*, **20**, 215, (2001).

22 Hauel N. H., Nar H., Priepke H., Ries U., Stassen J-M. and Wienen W., *J. Med. Chem.*, **45**, 1757, (2002).

23 Sippl W., *J. Comput.-Aided Mol. Design*, **14**, 559, (2000).

24 Dullweber F., Stubbs M. T., Musil D., Sturzebecher J. and Klebe G., *J. Mol. Biol*, **313**, 593, (2001).

25 Brzozowski A. M., Pike A. C. W., Dauter Z., Hubbard R. E., Bonn T., Engström O., Öhman L., Greene G.L., Gustafgsson J-Å. and Carlquist M., *Nature*, **389**, 753, (1997).

26 Klaholz B. P., Mitschler A., Belema M., Zusi C. and Moras D., *Proc. Natl. Acad. Sci*., **97**, 6322, (2000).

27 Stams T., Chen Y., Boriack-Sjodin P. A., Hurt J. D., Liao J., May J. A., Dean T., Laipis P., Silverman D. N. and Christianson D. W., *Protein Science*, **7**, 556, (1998).

28 Lovejoy B., Welch A. R., Carr S., Luong C., Broka C., Hendricks R. T., Campbell J. A., Walker K. A. M., Martin R., Van Wart H. and Browner M. F., *Nat. Struct. Biol.*, **6**, 217, (1999).

29 Chen L., Rydel T. J., Gu F., Dunaway C. M., Pikul S., McDow Dunham K. and Barnett B. L., *J. Mol. Biol*, **293**, 545, (1999).

30 Pavlovsky A. G., Williams M. G., Ye Q-Z., Ortwine D. F., Purchase II C. F., White A. D., Dhanaraj V., Roth B. D., Johnson L. L., Hupe D., Humblet C. and Blundell T. L., *Protein Science*, **8**, 1455, (1999).

31 Finzel B. C., Baldwin E.T., Bryant JR G. L., Hess G. F., Wilks J. W., Trepod C. M., Mott J. E., Marshall V. P., Petzold G. L., Poorman R. A., O'Sullivan T. J., Schostarez H. J. and Mitchell M. A., *Protein Science*, **7**, 2118, (1998).

32 Brandstetter H., Turk D., Hoeffken H. W., Grosse D., Stuzebecher J., Martin P.D., Edwards B. F. P. and Bode W., *J. Mol. Biol.*, **226**, 1085, (1992).

33 Engh R. A., Brandstetter H., Sucher G., Eichinger A., Baumann U., Bode W., Huber R., Poll T., Rudolph R. and von der Saal W., *Structure*, **4**, 1353, (1996).

34 Maignan S., Guilloteau J-P., Pouzieux S., Choi-Sledeski Y. M., Becker M. R., Klein S. I., Ewing W. R., Pauls H. W., Spada A. P. and Mikol V., *J. Med. Chem*., **43**, 3226, (2000).

35 Wang Z., Canagarajah B. J., Boehm J. C., Kassisá S., Cobb M. H., Young P. R., Abdel-Meguid S., Adams J. L. and Goldsmith E., *Structure*, **6**, 1117, (1998).

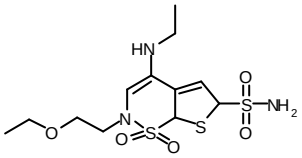
36 Muegge I., Martin Y. C., Hajduk P. J. And Fesik S. W., *J. Med. Chem*., **42**, 2498, (1999).

37 Halgren T. A, *J. Comp. Chem*., **20**, 730, (1999).

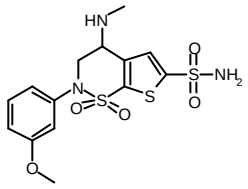
7 Appendix 1

(PDB accession numbers)

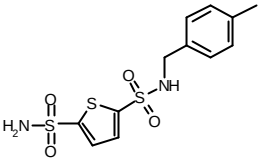
Carbonic anhydrase II



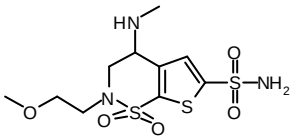
1A42



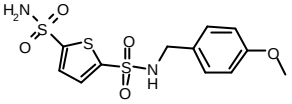
1BNM



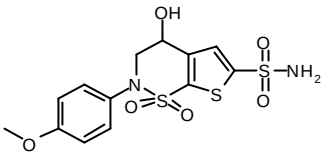
1BN1



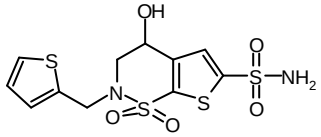
1BNQ



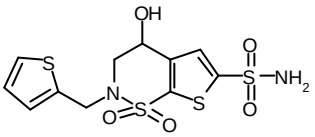
1BN4



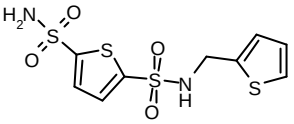
1BNT



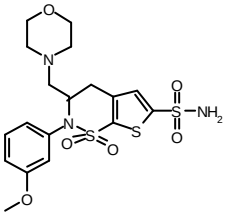
1BNV



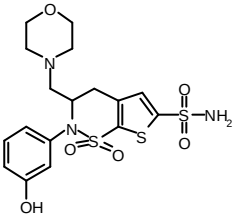
1BNU



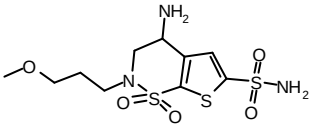
1BNW



1I8Z

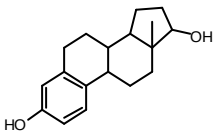


1I91

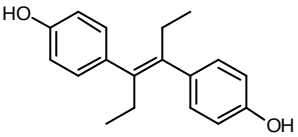


1I90

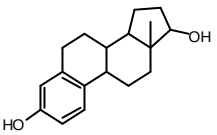
ER



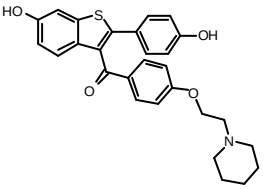
1ERE



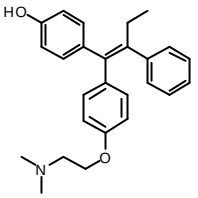
3ERD



1QKU

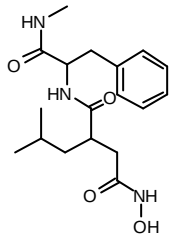


1ERR

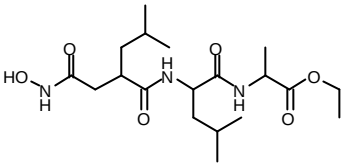


3ERT

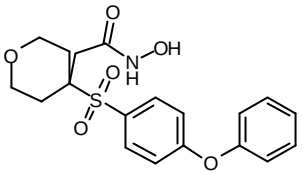
MMP-1



1HFC

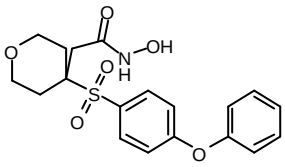


2TCL

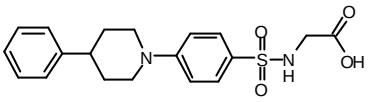


966C

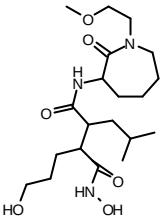
MMP-3



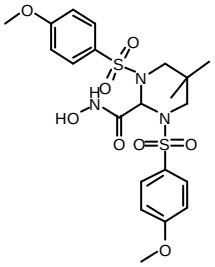
1B3D



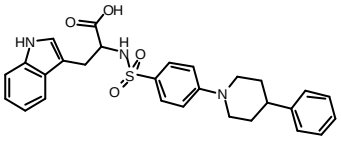
1B8Y



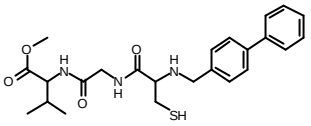
1BIW



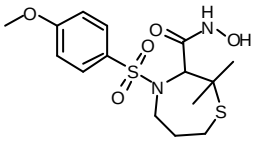
1BQO



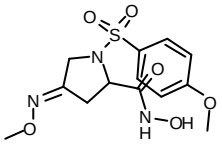
1CAQ



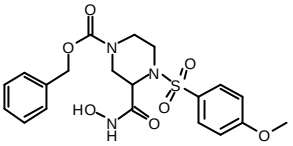
1C3I



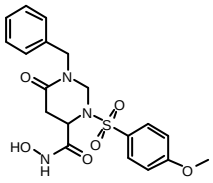
1D5J



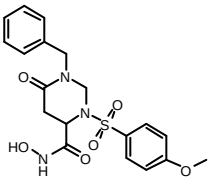
1D7X



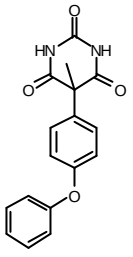
1D8F



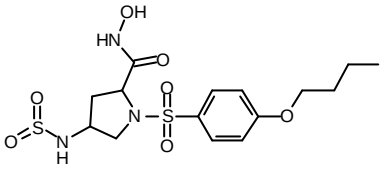
1D8M



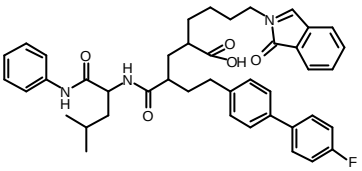
1G05



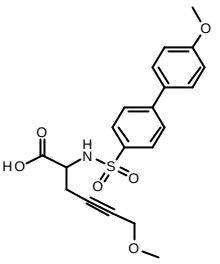
1G4K



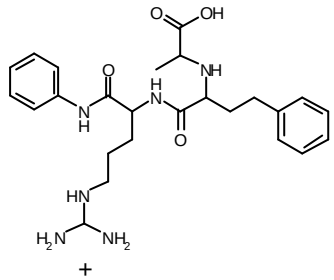
1G49



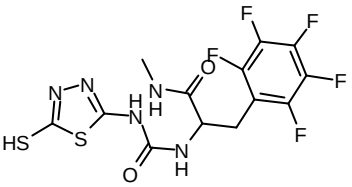
1HFS



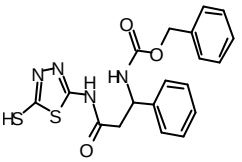
1HY7



1SLN

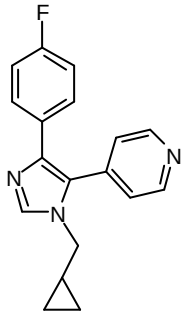


1USN

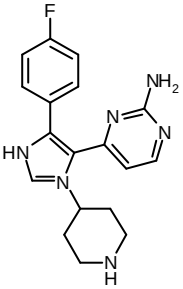


2USN

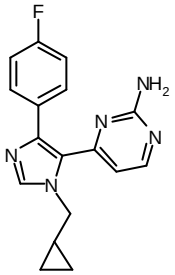
P38



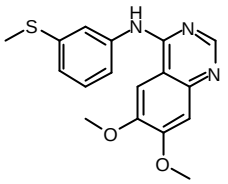
1BL6



1BL7

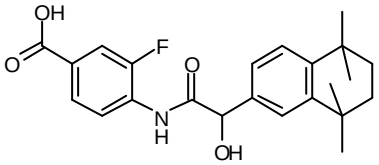


1BMK

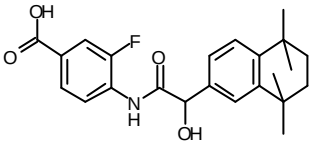


1DI9

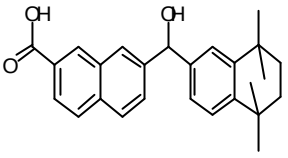
RARγ



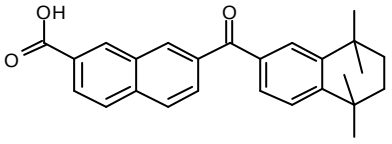
1EXA



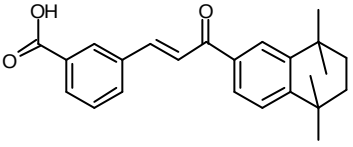
1EXX



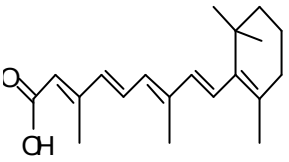
1FCX



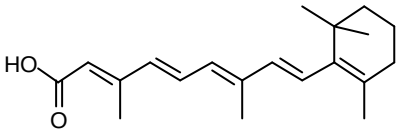
1FCY



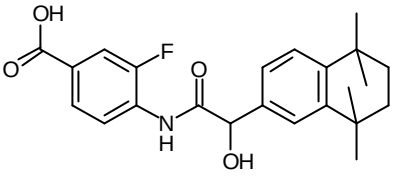
1FCZ



2LBD

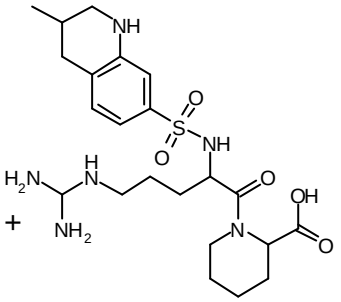


3LBD

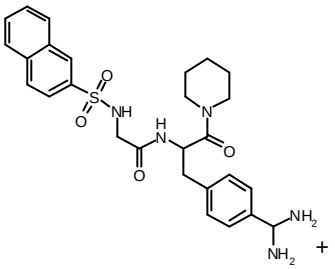


4LBD

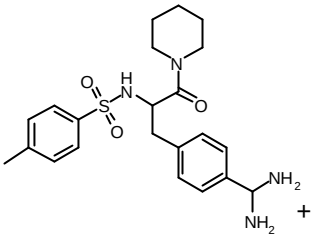
Thrombin



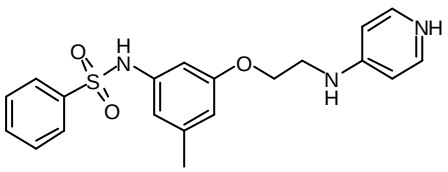
1ETR



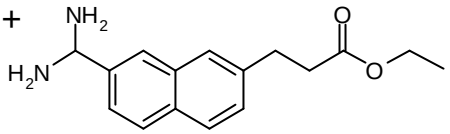
1ETS



1ETT

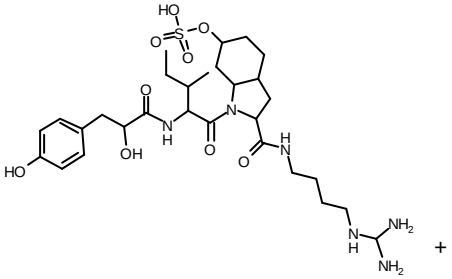


1UVT

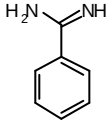


1UVU

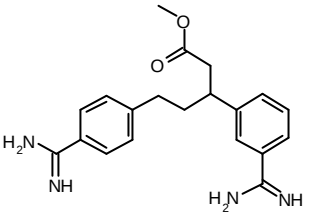
Trypsin



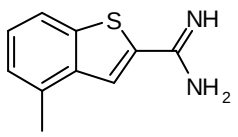
1AQ7



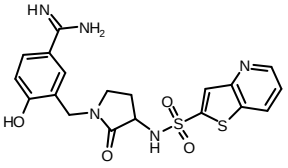
1BTY



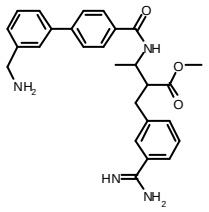
1AZ8



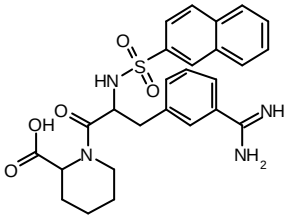
1C5Q



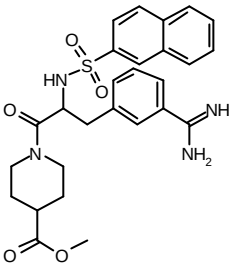
1F0T



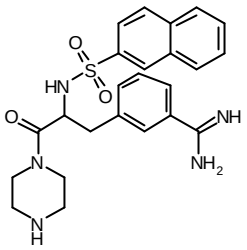
1F0U



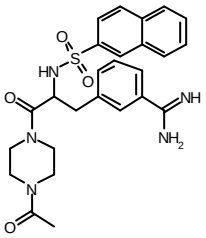
1K1I



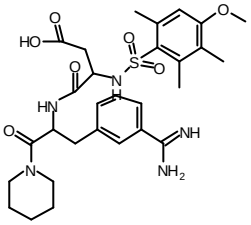
1K1J



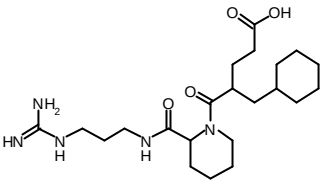
1K1L



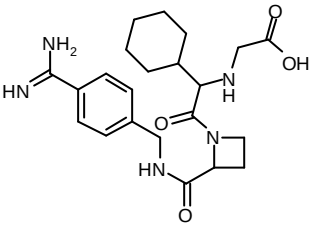
1K1M



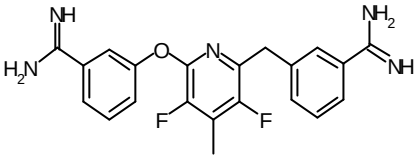
1K1N



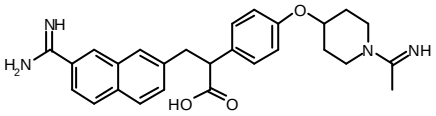
1K1O



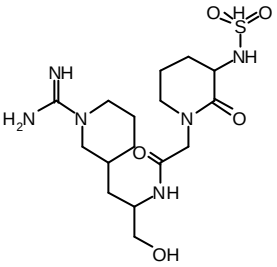
1K1P



1QB6



1MTS



1ZZZ