Journal of Computer-Aided Molecular Design, **16:** 855–869, 2002.
KLUWER/ESCOM

Sensitivity of molecular docking to induced fit effects in influenza virus neuraminidase

Louise Birch[1], Christopher W. Murray[2,*], Michael J. Hartshorn[2], Ian J. Tickle[2] & Marcel L. Verdonk[2]
[1]*The University Chemical Laboratory, University of Cambridge, Lensfield Road, Cambridge, CB2 1EW, UK;*
[2]*Astex Technology Ltd, 250 Cambridge Science Park, Milton Road, Cambridge,CB4 0WE, UK*

MS received 11 Sept 2002; accepted in final form 23 Dec 2002

Key words: Ligand-protein docking, GOLD, induced fit effects, rigid protein approximation, influenza, neuraminidase

Summary

Many proteins undergo small side chain or even backbone movements on binding of different ligands into the same protein structure. This is known as induced fit and is potentially problematic for virtual screening of databases against protein targets. In this report we investigate the limits of the rigid protein approximation used by the docking program, GOLD, through cross-docking using protein structures of influenza neuraminidase. Neuraminidase is known to exhibit small but significant induced fit effects on ligand binding. Some neuraminidase crystal structures caused concern due to the bound ligand conformation and GOLD performed poorly on these complexes. A 'clean' set, which contained unique, unambiguous complexes, was defined. For this set, the lowest energy structure was correctly docked (i.e. RMSD < 1.5 Å away from the crystal reference structure) in 84% of proteins, and the most promiscuous protein (1mwe) was able to dock all 15 ligands accurately including those that normally required an induced fit movement. This is considerably better than the 70% success rate seen with GOLD against general validation sets. Inclusion of specific water molecules involved in water-mediated hydrogen bonds did not significantly improve the docking performance for ligands that formed water-mediated contacts but it did prevent docking of ligands that displaced these waters. Our data supports the use of a single protein structure for virtual screening with GOLD in some applications involving induced fit effects, although care must be taken to identify the protein structure that performs best against a wide variety of ligands. The performance of GOLD was significantly better than the GOLD implementation of ChemScore and the reasons for this are discussed. Overall, GOLD has shown itself to be an extremely good, robust docking program for this system.

Abbreviations: RMSD: root mean square deviation, CSD: Cambridge Structural Database, PDB: Protein Data Bank, GA: Genetic Algorithm, vdw: van der Waals.

Introduction

Recently there has been great interest in virtual screening as an efficient method of generating new leads for drug discovery [1, 2]. However, virtual screening is hampered by difficulties in obtaining the correct binding mode and ranking molecules and binding modes accurately. The binding mode is important because it is difficult to believe that correct ranking can be routinely obtained with incorrect binding modes.

There are many docking programs available so it is often difficult to judge which is best for a given protein target (for a recent review of docking methods see reference [3]). This has led to interest in assessing the performance against large test sets of protein-ligand complexes. However, such test sets are intrinsically

*Corresponding author (Present address: Astex Technology Ltd, 250 Cambridge Science Park, Cambridge CB4 0WE, UK. email: c.murray@astex-technology.com, Tel: 00 44 (1223) 226 228, Fax: 00 44 (1223) 226 201).

biased because if one docks a ligand into a protein that has moved to accommodate it, the problem of locating the binding mode in the pre-formed pocket has been made easier.

In most, if not all enzymes, conformational changes occur on ligand binding. This may only involve small side chain rotations to maximize interactions with the ligand [4], or the change may also be associated with small main chain movements. In extreme cases large loop movements or even domain shifts are induced on ligand binding. Cases of large induced fit are difficult to account for in any docking program, although some docking methods can address large domain motions by use of hinge bending [5]. A more realistic goal would be a method robust enough to deal with relatively small changes in the active site when an analogous ligand binds. This is especially important in virtual screening of novel compounds where, at best, there is a crystal structure of a different ligand bound to the protein and, at worst, no structure for a protein-ligand complex. The sensitivity of results to small structural changes is very important because virtual screening of compound databases against a single protein may often miss hits that require small structural modifications of the protein to bind well.

The ideal situation would be a program able to dock a ligand into a protein structure from a different complex easily and with reasonable accuracy. This has been referred to as 'cross-docking' [6]. There are a number of ways that the issue of local flexibility can be approached when applied to systems with only small induced-fit effects. Different docking methods and the ways in which they accommodate protein flexibility may be found elsewhere [7].

Mobility within the active site

The most rigorous methods for modeling protein flexibility are Monte Carlo (MC) or Molecular Dynamics (MD) simulations. These are the only simulations that can model movement of the backbone but they are extremely computationally expensive [8]. Simplifications have been investigated such as implicit solvation models [9] and using rotamer libraries to create likely conformations of side chains in the active site before minimization[10, 11], but to our knowledge no docking program that attempts to do this has been validated extensively.

Multiple binding site models

An alternative method, that takes the flexibility of the protein into account, is to consider a number of different possible conformations as an ensemble and either dock compounds against each structure individually or against an averaged protein. In the averaged protein method used by FlexE [12], areas of similarity are merged whereas dissimilar areas are treated as separate instances. This allows the different structures to be recombined to new protein conformations during the docking process in order to select the conformation that fits best to the placed ligand. In the averaged protein method used by DOCK [13], a weighted average protein structure is constructed prior to docking. These methods may also be extended to pharmacophore modelling [14].

Use of a rigid protein with a robust scoring method

Using scoring functions that are less sensitive to small changes in structure is potentially a fast way of taking into account minor conformational changes both of side chains and of small backbone shifts. This is the way in which GOLD [15] deals with small conformational changes, relying heavily on the rigid protein approximation. SLIDE [16], and its precursor, SPECITOPE [17] incorporate induced fit as a post optimization process by placing the ligand and then removing side chain clashes by rotating protein side chains and ligand dihedrals based on mean field theory [18].

Neuraminidase is a good test case for looking at the robustness of GOLD because docking results have previously proved quite sensitive to protein structure [19] and the induced fit effects observed, whilst small, are significant to binding. There are also a large number of crystal structures of neuraminidase complexes available from the RCSB Protein Data Bank (PDB) [20], which allows an extensive cross-docking study to be done.

Whilst a large number of crystal complexes is useful, the quality of the data in the crystal structures is also important. It is not likely that a program will be able to reproduce a docking mode if the electron density is not sufficiently clear to allow the accurate definition of the binding mode or torsions of the ligand, or if the modelled conformation of the ligand in the crystal structure is in a very unfavourable conformation. We have used objective criteria to separate the neuraminidase structures into those in which we have confidence in the reported binding modes and those

for which we have reason to believe some aspect of the binding mode may be less reliable.

In this report we investigate the limits of the rigid protein approximation used by GOLD through cross-docking experiments of 21 ligands into 31 protein structures, and also look at the effects of the quality of the X-ray structure on the performance of GOLD.

Neuraminidase

Influenza is the major cause of mortality and morbidity among patients with respiratory diseases. Neuraminidase is a surface glycoprotein that catalyses the cleavage of an α-ketosidic linkage that exists between a terminal sialic acid, N-acetyl neuraminic acid (Neu5Ac, Figure 1(a)), and an adjacent sugar residue. This action has a number of important effects that enable the spread of the virus within the respiratory tract [21].

The active site in neuraminidase is conserved across both A and B type influenza, despite the fact that there is up to 75% sequence variation. This makes it an attractive target for structure-based drug design [22].

The sialic acid binding site of neuraminidase consists of four pockets, which lock the sialic acid ligand into an unfavourable boat-like conformation, with a $K_i \sim 10^{-3}$M.

i. An acid binding site, where a triad of arginines (residues 118, 292 and 371) bind the carboxylic acid group.
ii. An acetamido-binding pocket that has a hydrophobic region (Trp178 and Ile222) to interact with the acyl group and Arg152, which forms a hydrogen bond with the amide oxygen. The amide proton forms water-mediated hydrogen bonds to Glu277 (not shown) and Glu227.
iii. A glycerol binding pocket that includes Glu276, which hydrogen bonds to the two terminal hydroxyls. The precise orientation of Glu276 varies between Type A and B but the interactions remain the same.
iv. Another pocket binds the ring hydroxyl, which forms hydrogen bonds via water to Glu119 and Glu227.

Figure 2 shows the crystal structure of transition state analogue DANA binding to the four pockets in the sialic acid-binding site. The earliest neuraminidase inhibitors were transition state analogues, such as Neu5Ac2en (DANA, Figure 1(b)) which had a binding affinity $\sim 10^3$ times greater than sialic acid [23]. However, the affinity was similar for all sialidases, including mammalian ones, and they showed no antiviral activity in animals.

Zanamivir [24], Figure 1(c), was the first neuraminidase inhibitor to reach clinical trials. It was based upon the deoxysialic acid derivatives but the guanidinium group displaces the water in the ring hydroxyl-binding pocket and leads to an $IC_{50} \sim 10^{-10}$ M in both A and B strains. However, due to its highly polar nature, the bioavailability is very poor and oral inhalation or intravenous injection is necessary.

An orally active alternative, oseltamivir Figure 1(d), was found and proved to be as potent as zanamivir [25]. Taken as the ethyl ester prodrug to allow oral administration, the drug is cleaved by esterases to release the active carboxylic acid form. The heterocycle was replaced by a more stable carbocycle and a hydrophobic group replaced the glycerol side chain. The interaction with the neuraminidase active site causes Glu276 in the glycerol-binding pocket to rotate and form a salt bridge with Arg224 [26]. This creates a hydrophobic region that allows interaction with the hydrophobic side chain.

Another inhibitor, peramivir [27, 28], utilises the same interactions as oseltamivir but on a cyclopentane scaffold, and was found to have a nanomolar IC_{50} for both A and B strains. Peramivir is currently undergoing clinical trials.

Methods

The experiments described here attempt to simulate the virtual screening of compounds using default parameters, as in the applications normally studied using GOLD. To avoid biasing the results unfairly towards the docked mode, a number of measures were taken to make the docking as objective as possible, although some assumptions were unavoidable.

Definition of active site

The active site determines the search space considered during the docking run and must be large enough to encompass all possible docked modes. All protein-ligand X-ray complexes were superimposed, and the active site was defined as all protein atoms within 6 Å (unless otherwise stated) of any ligand heavy atom. Other than confining the search to the active site, no additional constraints on the translation of the ligand were made.

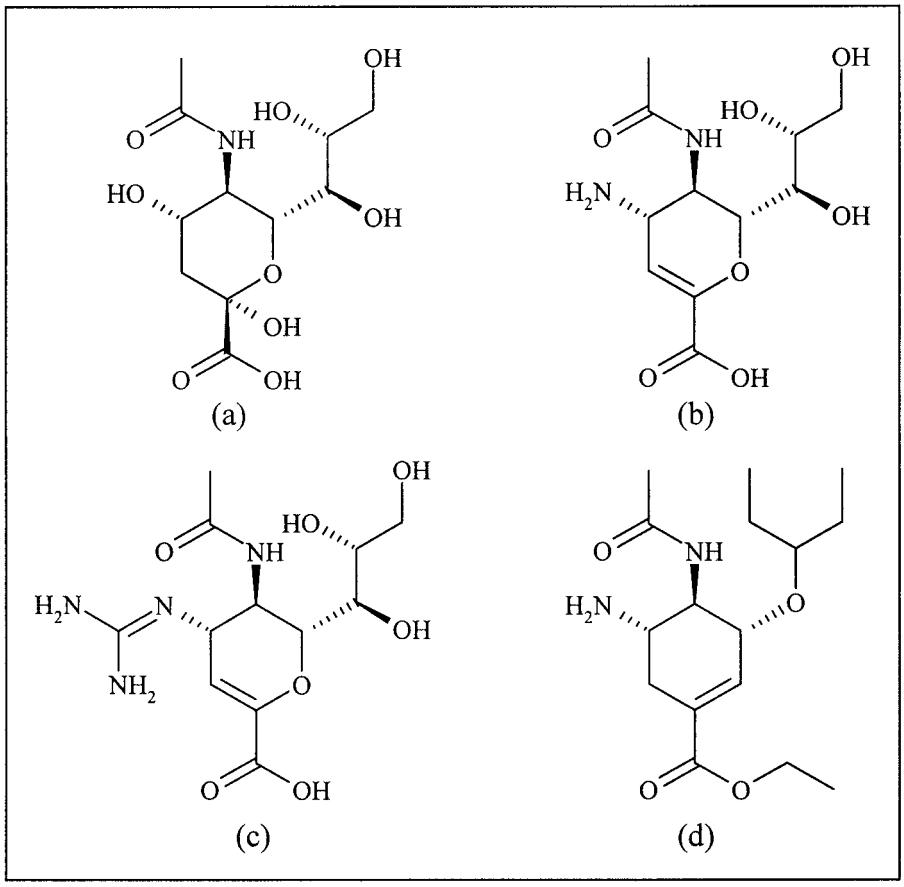


Figure 1. Chemical structures of some influenza neuraminidase inhibitors: (a) Sialic acid (N-acetylneuraminic acid (Neu5Ac)); (b) DANA (2-deoxy-2,3-dehydro-N-acetylneuraminic acid (Neu5Ac2en)); (c) Zanamivir (4-guanidio-Neu5Ac2en); (d) Oseltamivir (4-N-acetyl-5-amino-3-(1-ethylpropoxy)-1-cyclohexene-1-carboxylic acid methyl-ester prodrug).

Ligand flexibility

In order to avoid a biased starting conformation, ligand conformations were created automatically from SMILES strings (charged at pH 7) using the program CORINA [29]. As a result, any distortion of bond angles and distances in the active site conformation were not automatically included in the starting geometry used by GOLD.

All rotatable bonds were randomized at the beginning of a GOLD docking and treated as completely flexible during the docking runs. Bond lengths, angles and torsions associated with non-rotatable bonds were held fixed in GOLD whilst amide bonds were constrained to the *trans* form unless specifically allowed to flip to the *cis* form.

Flexibility of ring systems to explore different ring conformations was allowed using the 'Flip ring corners' option in GOLD (unless otherwise specified) [15]. This was necessary because some of the ligands based on sialic acid bind in a boat conformation.

Protein flexibility

Although the whole protein is generally treated as rigid in GOLD, partial flexibility is conferred by the rotation of protein hydroxyl protons and NH_3^+ groups, to allow optimization of hydrogen bonding to the ligand.

GOLD docking

For each independent genetic algorithm (GA) run, a maximum number of 100,000 operations were performed on a population of 5 islands of 100 individuals. Operator weights for crossover, mutation and migration were set to 95, 95 and 10 respectively. To allow poor van der Waals contacts and poor hydrogen bonds

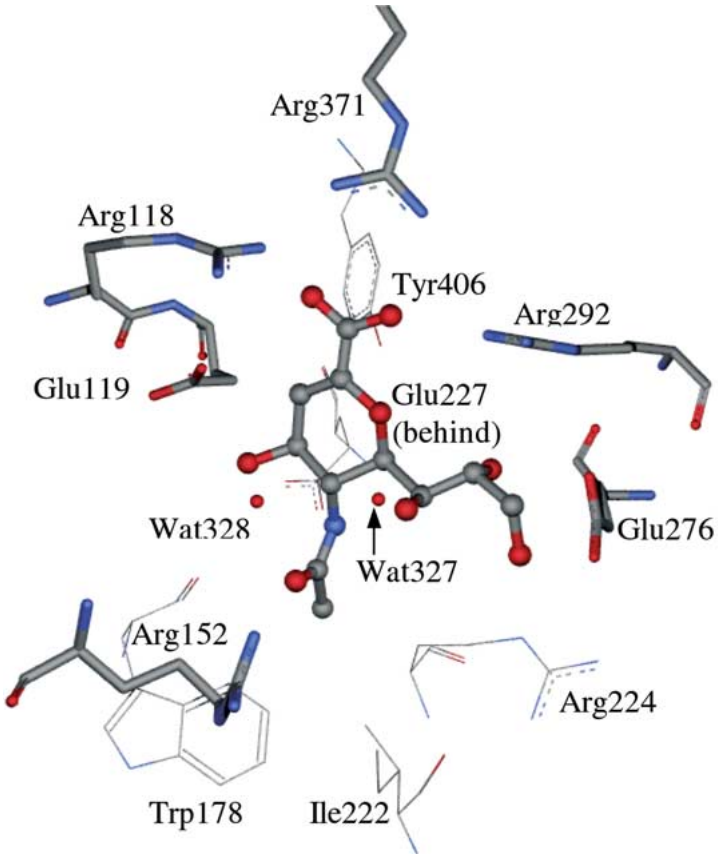


Figure 2. Neu5Ac2en (DANA, Figure 1(b), ball and stick) bound into influenza type A neuraminidase (PDB code: 1f8b). Residues making direct hydrogen bonds to the ligand are shown as sticks whilst those making water mediated hydrogen bonds or other important interactions are shown as lines.

to occur at the beginning of a GA run (in the expectation that they will evolve to better solutions), the initial external van der Waals (vdw) energies were cut-off at $7.5*k_{ij}$, where k_{ij} is the depth of the vdw well between two atoms i and j, and the maximum distance between a donor hydrogen and a fitting point was set to 4 Å. In order to speed up the calculations, the docking was terminated when the top three solutions were within 1.5 Å RMSD of each other *or* after 20 runs.

For the instance where virtual screening settings were used, a maximum of 10,000 operations were performed upon a single population of 100 individuals. Operator weights for crossover, mutation and migration were set to 100, 100 and 0 respectively. The initial, external vdw energy cut-off was reduced to $3*k_{ij}$ and the initial maximum distance between hydrogen to fitting point increased to 6 Å. No torsion distributions were used for the screening. Early termination was not allowed but only five runs were conducted per ligand.

ChemScore

Part of this work compares the GOLD scoring function to a version of the ChemScore function. The energy function used by ChemScore has a constant, ΔG_0, plus contact terms to estimate lipophilic and metal-ligand binding contributions, an explicit form for the hydrogen bonds (taking into account distances and angles between hydrogen bonding atoms) and a term that penalises flexibility, which is treated as constant for all conformations of the same ligand. The coefficients for each term were derived empirically using a regression based on a set of crystallographic complexes for which the binding constant is known. When used in docking the function is augmented by terms to penalize poor ligand torsions, clashes with the enzyme and internal clashes between ligand atoms. The full form of the empirical energy function used in this work is given elsewhere [30].

Selection and preparation of input data

Proteins

X-ray structures of 38 influenza virus neuraminidases complexed with small molecules, were retrieved from the RCSB Protein Data Bank [20]. The heavy atoms of these structures were superimposed. Of these complexes, seven (2qwb-h) were excluded because they were R292K mutants of other represented structures. A summary of the remaining 31 complexes may be found in Table 1. No other influenza neuraminidase complexes were knowingly excluded from this set.

The ligand and all crystallographic waters, including those mediating hydrogen bonds to the ligand, were removed from the remaining 31 structures and hydrogens were added using InsightII [43]. The protonation state of charge groups was set assuming pH 7. In the active site of some of the protein structures (1bji, 1f8c-e, 1iny, 1mwe, 1nnc, 1ivb and 1ivd) the amide oxygen and nitrogen of Asn294 appeared to be wrongly assigned. In these cases the sidechain amide was flipped, on the basis of the hydrogen bonds made to Arg292 and Ala246.

To look at the effect of specific waters on the dockings, the protein 1mwe was also prepared with water molecules 327 and/or 328 left in the active site. For these situations the hydrogens were added using InsightII and positioned manually to form the appropriate hydrogen bonds to the protein.

Ligands

The ligand structure containing only the heavy atoms was extracted from each superimposed complex using InsightII [43]. This is the reference ligand for each complex, to which RMSD values are calculated. In total, the complexes contained 21 different ligands (shown in Table 2). Also indicated in Table 2 are the ligands that: (i) contain a cyclohexyl ring in the boat conformation, (ii) displace specified waters from the active site, or (iii) induce movement of Glu276 to create a hydrophobic pocket.

The ligand structures were converted into SMILES strings, and charged as if at pH 7 (i.e. carboxylic acid and phosphate groups were negatively charged, whilst amino and guanidine groups were positively charged). Three-dimensional structures for the ligands were generated using CORINA [29]. CORINA makes amide bonds *trans* by default, however none of the ligands contain *cis* amide bonds.

Selection of a subset of complexes

The full set of 31 complexes was tested and the RMSD to the reference structures calculated. The binding modes for some of the ligands in the crystal structures gave some cause for concern and in general, as discussed below, it was observed that GOLD performed significantly worse for these ligands. It therefore seemed reasonable to focus on a smaller set of complexes for which we felt more confident about the binding modes. In many cases there were duplicates of the same ligand, and these were also removed. Complexes were not removed from the full set because GOLD performed badly on them but rather because they failed to meet particular objective criteria, outlined below.

Ligand-protein clash

Complexes 1ivd and 1ive were excluded from the clean set because clashes between the protein and the ligand were found in the crystal structure. A clash was defined as any heavy atom-heavy atom distance that was less than 80% of the minimal values observed in a study of similar atom types in the CSD [44, 45].

Strained amide torsion

Nine of the ligands contained an amide directly attached to a mono- or di-substituted phenyl ring, some of which were found to have unusually large amide torsions. In a search of the CSD [44], small molecule crystal structures containing a phenyl-N-acetyl moiety were located. Of the 129 molecules found, 59 were ortho-substituted and 7 were di-ortho-substituted but no molecules were found to have an amide C-N bond dihedral greater than 20 degrees. The mean torsion was very close to zero, as expected, and the standard deviation over the structures was 4.2°. The distribution indicates that there is an extremely large energy penalty associated with twisting the amide C-N bond.

Complexes in which the amide dihedral in the ligand was greater than 21.2° (mean + 5σ) were removed. It is questionable whether such energetically strained systems are really present in the protein complexes or whether they are some artifact of the refinement protocol. Using this criterion, the complexes 1b9s (ligand FDI), 1ing (ligand ST5), 1ivb (ligand ST1), 1ivc (ligand ST2) and 1ivd (ligand ST1) were excluded from the clean set.

Table 1. The 31 neuraminidase protein-ligand complexes used in the full set. Details of the ligands can be found in Table 2.

PDB code	sub-type	Strain	Ligand abbrevn.	Resolution. [Å]
1A4G [31]	–	B/BEIJING/1/87	ZAN	2.2
1A4Q [31]	–	B/BEIJING/1/87	DPC	1.9
1B9S [32]	–	B/LEE/40	FDI	2.5
1B9T [32]	–	B/LEE/40	RA1	2.4
1B9V [32]	–	B/LEE/40	RA2	2.35
1BJI [32]	N9	A/TERN/AUSTRALIA/G70C/75	DPC	2.0
1F8B [33]	N9	A/TERN/AUSTRALIA/G70C/75	DAN	1.8
1F8C [33]	N9	A/TERN/AUSTRALIA/G70C/75	4AM	1.7
1F8D [33]	N9	A/TERN/AUSTRALIA/G70C/75	9AM	1.4
1F8E [33]	N9	A/TERN/AUSTRALIA/G70C/75	49A	1.4
1INF [34]	–	B/LEE/40	ST4	2.4
1ING [35]	N2	A/TOKYO/3/67	ST5	2.4
1INH	N2	A/TOKYO/3/67	ST6	2.4
1INV [36]	–	B/LEE/40	EQP	2.4
1INW [36]	N2	A/TOKYO/3/67	AXP	2.4
1INX [36]	N2	A/TOKYO/3/67	EQP	2.4
1INY [36]	N9	A/TERN/AUSTRALIA/G70C/75	EQP	2.4
1IVB [37]	–	B/LEE/40	ST1	2.4
1IVC [37]	N2	A/TOKYO/3/67	ST2	2.4
1IVD [37]	N2	A/TOKYO/3/67	ST1	1.8
1IVE [37]	N2	A/TOKYO/3/67	ST3	2.4
1IVF [37]	N2	A/TOKYO/3/67	DAN	2.4
1MWE [38]	N9	A/TERN/AUSTRALIA/G70C/75	SIA	1.7
1NNB [39]	N9	A/TERN/AUSTRALIA/G70C/75	DAN	2.8
1NNC [40]	N9	A/TERN/AUSTRALIA/G70C/75	ZAN	1.8
1NSC [41]	–	B/BEIJING/1/87	SIA	1.7
1NSD [41]	–	B/BEIJING/1/87	DAN	1.8
2BAT [42]	N2	A/TOKYO/3/67	SIA	2.0
2QWI [26]	N9	A/TERN/AUSTRALIA/G70C/75	G20	2.0
2QWJ [26]	N9	A/TERN/AUSTRALIA/G70C/75	G28	2.0
2QWK [26]	N9	A/TERN/AUSTRALIA/G70C/75	G39	1.8

Ambiguous electron density

We have recreated the electron density of the complexes using the deposited structure factors of each protein (where available). The method used to construct these electron densities has been outlined in a previous paper [45].

In our hands, the electron density recreated from the 1inh file was not sufficiently well defined to unambiguously position the ligand (ST6). As no paper has ever been published to describe the solving of this structure, this complex (and hence its associated ligand) was excluded from the clean set.

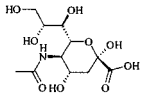
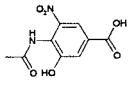
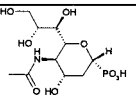
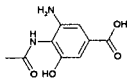
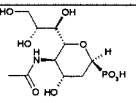
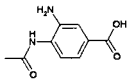
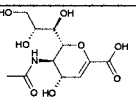
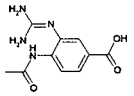
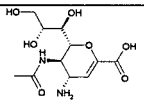
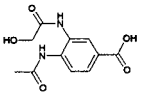
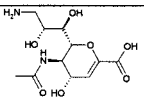
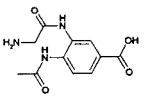
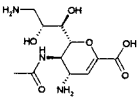
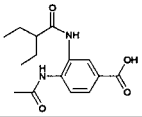
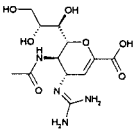
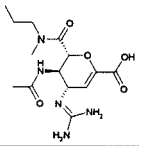
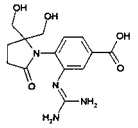
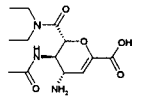
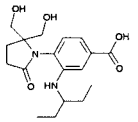
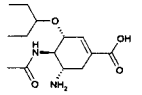
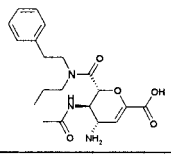
Duplicate ligands

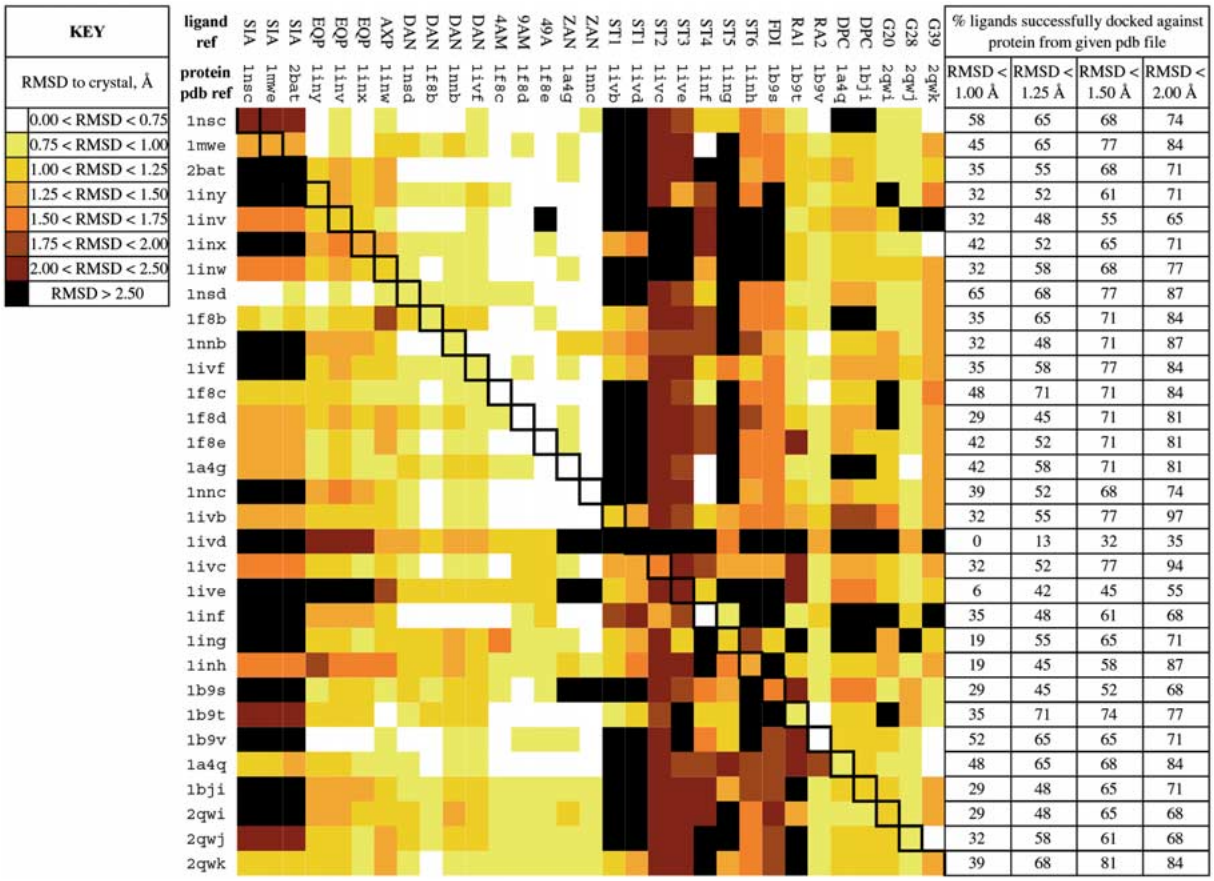
These adjustments resulted in a clean set with duplicates of some ligands (SIA, EQP, DAN, ZAN and DPC). The lowest resolution complexes were eliminated in each case and, if more than one protein structure remained, a complex was chosen on the basis of its electron density. The final set of clean, unique protein-ligand complexes contains 15 structures.

Results and Discussion

Initially the full set of 31 complexes was run using a 6 Å cavity radius and 100,000 operations (as de-

Table 2. The 21 ligands found in crystal complexes with influenza neuraminidase; [a]binds in boat conformation; [b]displaces water in ring hydroxyl-binding pocket; [c]displaces water in ring hydroxyl-binding and acetamido-binding pockets; [d]induces Glu276=Arg224 salt bridge, forming a hydrophobic pocket.

Ligand abbrevn.	Structure	Ligand abbrevn.	Structure
SIA[a] (sialic acid)		ST1	
EQP		ST2	
AXP[a]		ST3	
DAN (DANA)		ST4	
4AM		ST5[d]	
9AM		ST6	
49A		FDI	
ZAN[b] (Zanamivir)		G20[b,d]	
RA1[c]		G28[d]	
RA2[c]		G39[d] (Active form of Oseltamivir)	
DPC[d]			



% ligands successfully docked against protein from given pdb file			
RMSD < 1.00 Å	RMSD < 1.25 Å	RMSD < 1.50 Å	RMSD < 2.00 Å
58	65	68	74
45	65	77	84
35	55	68	71
32	52	61	71
32	48	55	65
42	52	65	71
32	58	68	77
65	68	77	87
35	65	71	84
32	48	71	87
35	58	77	84
48	71	71	84
29	45	71	81
42	52	71	81
42	58	71	81
39	52	68	74
32	55	77	97
0	13	32	35
32	52	77	94
6	42	45	55
35	48	61	68
19	55	65	71
19	45	58	87
29	45	52	68
35	71	74	77
52	65	65	71
48	65	68	84
29	48	65	71
29	48	65	68
32	58	61	68
39	68	81	84

Figure 3 'Full' set all-pairs docking experiment for 31 influenza neuraminidase complexes and ligands. The RMSD of the lowest energy solution, relative to the crystallographic position of the ligand in its native protein, is given in colour-coded form. At the right of the figure, the percentage of ligands correctly docked, at four different RMSD cut-offs relative to the reference ligand, is given for each protein structure.

tailed above) against all 21 ligands. The RMSD of the reference ligand to the top ranked (i.e. lowest energy) solution was calculated for each docking and is shown in Figure 3. GOLD found the crystal structure to within 1.5 Å RMSD in 66% of cases (Table 3).

The GOLD results showed variable performance against different protein conformations, with some proteins producing very good results at less than 1.5 Å RMSD to the reference ligand (e.g., 77% of ligands were correctly docked against the 1mwe protein structure) whilst others performed poorly (e.g. only 32% of ligands were correctly docked against 1ivd). This effect is not removed if longer protocols are used and appears to be due to the hypersensitivity of docking results to relatively small changes in protein geometry. This hypersensitivity is of concern in virtual screening applications where one often uses only a single protein structure for the docking runs. It would seem that some protein structures are more promiscuous, and hence more useful in virtual screening methodologies, than others.

The clean set with duplicates performed noticeably better than the full set but when the duplicates were removed there was no further change in performance (Table 3). This suggests that the duplicated structures were representative of the set.

The matrix of RMSDs for the clean set (colour-coded) can be seen in Figure 4. All proteins dock their native ligand to within 1.5 Å RMSD and ten to within 1.0 Å (third column of Table 4). The dockings of the lowest energy solution of each ligand into its native protein can be accessed via our web interface [46, 47].

The dockings are much more consistent than for the full set and most protein structures perform extremely well (at least 80% of ligands correctly predicted to within 1.5 Å of the reference structure in thirteen out of fifteen proteins). The hypersensitivity to small changes in protein structure is less pronounced

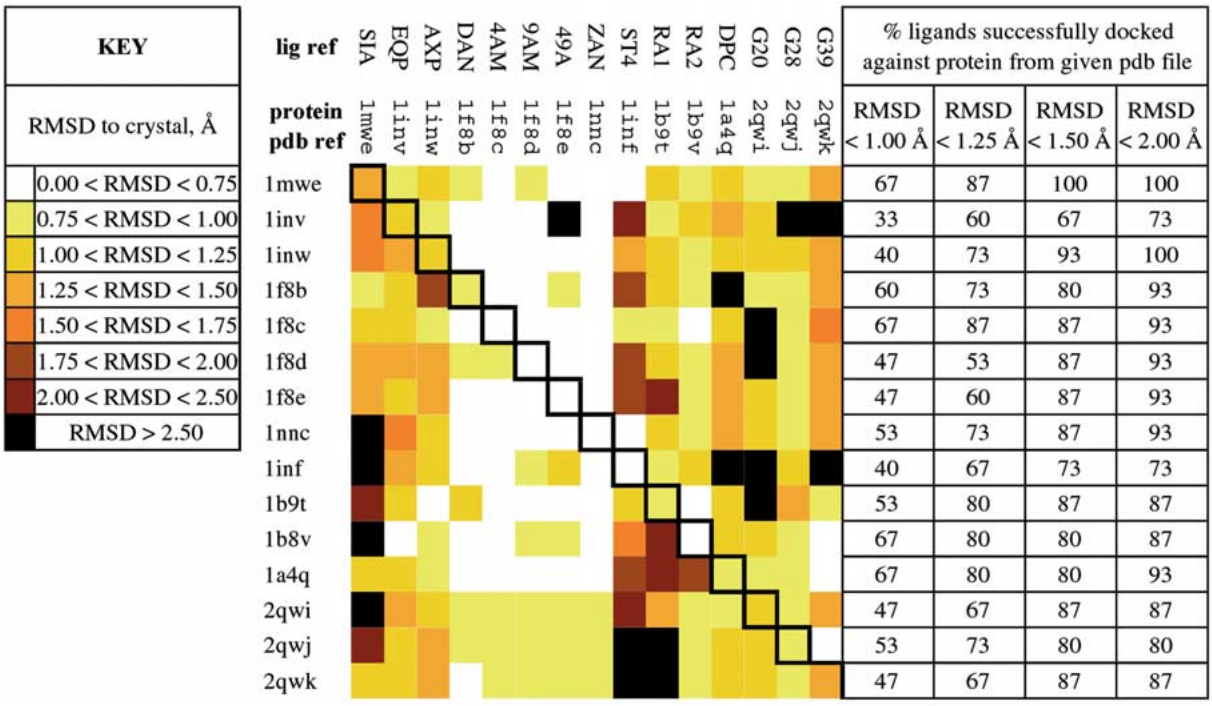


Figure 4 'Clean' set all-pairs docking experiment for 15 neuraminidase complexes and ligands. The RMSD of the lowest energy solution, relative to the crystallographic position of the ligand in its native protein, is given in colour-coded form. At the right of the figure, the percentage of ligands correctly docked, at four different RMSD cut-offs relative to the reference ligand, is given for each protein structure.

Table 3 Average percentage of ligands successfully docked (relative to their native crystal structures) into the protein structures of the full set and the reduced, 'clean' sub-sets. The percentages are given for four different RMSD cut-off definitions for successful docking.

RMSD cut-off, Å	full set	clean set but with duplicates	clean set
<2.00	76%	87%	89%
<1.50	66%	84%	84%
<1.25	54%	72%	72%
<1.00	35%	49%	52%

for the clean set. The most promiscuous protein in the clean set, 1mwe (the native ligand for which is sialic acid), correctly docks all 15 ligands to within 1.5 Å of the reference structures. This is more indicative of the kind of performance that could be expected in a well-constructed virtual screen. The lowest energy solution for each of the ligands docked into 1mwe can also be seen on the website [46].

Table 4 also illustrates the effect on docking performance of a particular induced fit effect. The ligands DPC, G20, G28 and G39, which require the formation of the hydrophobic pocket (the 'open' structure), were all able to dock within 1.5 Å RMSD of the reference ligand in five of the eleven complexes that did not have this pocket (the 'closed' structure). The remaining 'closed' complexes were able to dock at least one of the ligands correctly. Although, in general, the dockings to the open forms are slightly better than to the closed structures, some of the closed proteins (e.g. 1mwe, Table 4) were able to correctly dock DPC, G20, G28 and G39, despite the induced fit effects. We consider this 'robustness' to be desirable in virtual screening or docking applications because the correct conformation for the protein (or mediating water positions) may not be available. Any good docking program would be expected to perform better against the correct protein conformation but the observed robustness with GOLD implies that the dominant interactions necessary for good docking are sufficient to overcome unfavourable interactions caused by using the non-native protein conformation. One reason for this may be the soft clash term used by GOLD which may be of general utility when applied within the rigid protein approximation (although it cannot be

Table 4 Average RMSD of lowest energy docked ligand to reference ligand for proteins in the clean set. The proteins used are: the ligand's native protein (third column); all clean proteins (fourth column); 'closed' proteins, i.e. proteins without the hydrophobic pocket (fifth column); 'open' proteins, with a hydrophobic pocket (sixth column); 1mwe, the most promiscuous 'closed structure' (seventh column). In their native structures, the ligands DPC, G20, G28 and G39 all bind into the 'open' form of the protein.

		Average RMSD, Å, relative to reference ligand				
Native pdb ref	lig ref	Native protein	All 'clean' proteins	Closed only	Open only	1mwe (closed)
1mwe	SIA	1.28	2.36	2.34	2.41	1.28
1inv	EQP	1.18	1.17	1.16	1.20	0.87
1inw	AXP	1.01	1.14	1.13	1.19	1.24
1f8b	DAN	0.88	0.75	0.72	0.83	0.79
1f8c	4AM	0.74	0.71	0.67	0.82	0.65
1f8d	9AM	0.56	0.72	0.69	0.80	0.78
1f8e	49A	0.61	1.00	1.08	0.79	0.64
1nnc	ZAN	0.56	0.63	0.59	0.75	0.61
1inf	ST4	0.44	1.69	1.32	2.71	0.66
1b9t	RA1	0.99	1.62	1.28	2.57	1.05
1b9v	RA2	0.71	0.92	0.86	1.09	0.80
1a4q	DPC	0.95	1.57	1.77	1.02	1.24
2qwi	G20	1.15	2.07	2.41	1.13	0.84
2qwj	G28	0.86	1.18	1.29	0.90	0.86
2qwk	G39	1.35	1.56	1.79	0.95	1.37

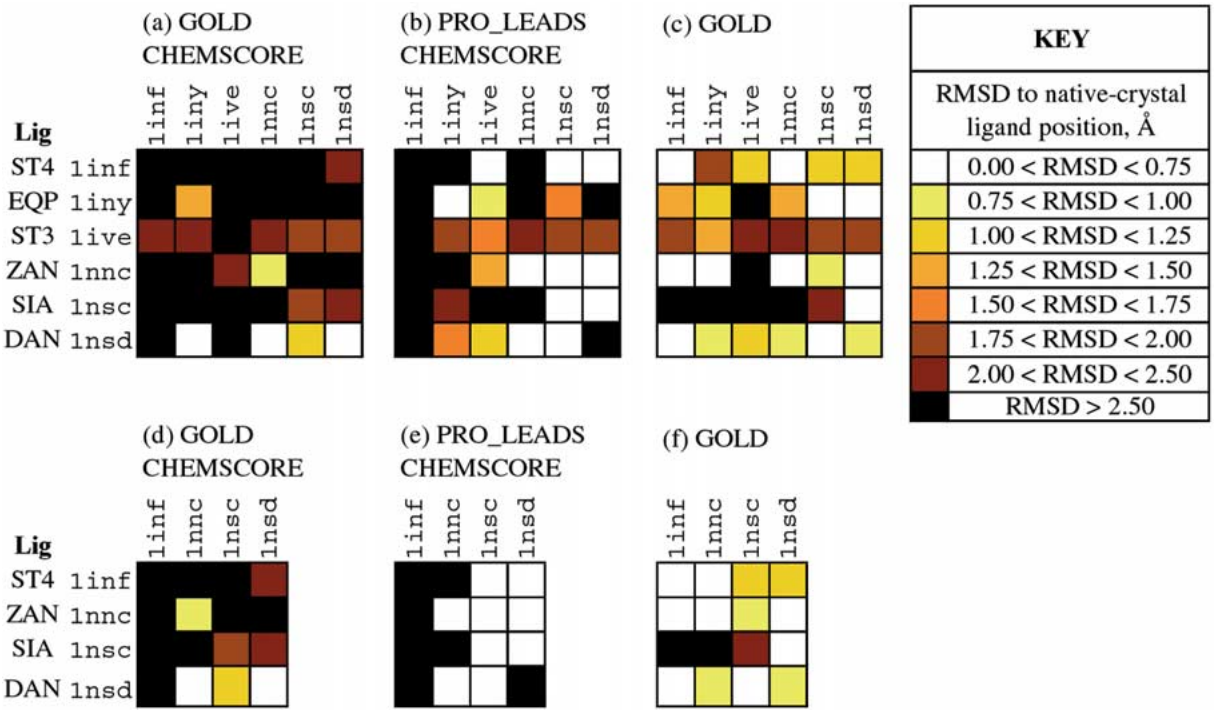


Figure 5 Docking results for the neuraminidase complexes used in a previous neuraminidase docking study [19] with (a) the current GOLD implementation of the ChemScore function, (b) that observed by PRO_LEADS, and (c) with the GOLD function. The proteins that were included in the clean set are shown in (d), (e) and (f), respectively.

Table 5 Average percentage of ligands successfully docked into the protein structures of the full set (top) and the clean set (bottom), for different cavity sizes, ring flipping option, GA parameters and scoring function.

	6 Å	4 Å	6 Å ring flip OFF		6 Å virtual screening settings	6 Å ChemScore
RMSD cut-off, Å	full	full	full	full (not boat)	full	full
<2.00	76%	77%	71%	75%	78%	35%
<1.50	66%	67%	59%	63%	67%	28%
<1.25	54%	53%	49%	47%	55%	23%
<1.00	35%	36%	33%	28%	37%	16%
RMSD cut-off, Å	clean	clean	clean	clean (not boat)	clean	clean
<2.00	89%	89%	76%	84%	83%	49%
<1.50	84%	85%	68%	76%	78%	44%
<1.25	72%	71%	59%	60%	67%	36%
<1.00	52%	51%	43%	45%	50%	22%

discounted that the soft clash term may be detrimental to successful docking for other proteins).

With decreasing sensitivity, however, comes the potential for lack of selectivity. Will a system that allows ligands to bind into the wrong conformation of a protein allow docking of compounds that do not bind? The problem of these 'false positives' in screening is becoming increasingly important as large databases of compounds are tested. However, the converse problem of 'false negatives', i.e. missing potentially novel 'hits', is perhaps more serious for lead discovery and virtual screening.

Reduced cavity size

Reducing the cavity size might be expected to improve the performance by restricting the conformational space and biasing the algorithm towards the correct solutions. Dockings with cavity radii around all reference ligands of 6 Å and 4 Å perform equally well on both the full set and the clean set, which indicates that GOLD is not significantly impaired by the larger search space and there is no bias (Table 5).

Reduced ring flexibility

With the rings fixed into the chair conformation, it is not possible to dock the ligands AXP and SIA, but the performance of the docking for the remaining ligands is not significantly better (see Table 5). For this reason we recommend running GOLD with ring flipping allowed because this permits exploration of other ring conformations without having a detrimental effect on those which bind in the chair position. It should also be noted that allowing amide bonds to flip into the *cis* form did not improve the overall dockings for the full set, or allow the ligands with strained amides to be docked correctly but neither did it adversely affect the dockings of the other ligands at an appreciable level (data not shown).

Number of operations

The fast screenings, as used in virtual screening (10,000 operations rather than 100,000) do not show a large decrease in the performance on the full set, although there is a slight decrease in performance on the clean set at lower RMSD cut-off values (Table 5). This is as expected but the effect is probably not large enough to lead to concerns of excessive false negatives in virtual screenings. The average CPU time using 100,000 operations is 7.2 minutes per molecule on a Pentium III 1 GHz processor. The virtual screening settings (10,000 operations) take 1.4 minutes per molecule.

Scoring function

A version of the ChemScore function was used as an alternative function in docking [48]. ChemScore does not perform well on the neuraminidase set, with

less than 50% of complexes docking to within 2 Å of the reference ligand. When compared to previous equivalent cross-dockings (with PRO_LEADS [19]) the observed RMSD results, whilst following the same trend, are noticeably worse than the values in the literature (Figure 5). The largest discrepancy between our ChemScore set and the PRO_LEADS set is the protein 1ive, for which PRO_LEADS achieved a much better success rate. This structure was removed from our clean set because of clashes between ligand and protein. Eliminating this, along with 1iny, which was removed from the clean set on the basis of its electron density, leaves a more relevant sub-set, Figure 5. Whilst the two ChemScore sets show a similar pattern, with 1inf performing very poorly, the GOLD implementation of ChemScore seems to give worse results. Whilst 1nsc and 1nsd were previously able to dock all four ligands to less than 0.75 Å, only dan (DANA) is now docked <1.5 Å. In contrast, the GOLD scoring function is able to dock st4, zan and dan into all four proteins to within 1.5 Å RMSD and 1nsd can dock all of the ligands at the same level.

We believe that this discrepancy between the ChemScore results is due to a protein relaxation procedure performed upon the complexes prior to docking with PRO_LEADS. This relaxation procedure regularizes the hydrogen bond networks and will make the docking easier by increasing the energy gap between native and non-native ligand geometries. It appears that for neuraminidase, the ChemScore function may be overly sensitive to small changes in protein structure.

The relatively poor performance of the ChemScore function is in contrast to its performance on the 300 structures in the extended GOLD validation set, where success rates with the ChemScore function are comparable to GOLD [48]. Neuraminidase is a system that is heavily based on electrostatic interactions and the GOLD scoring function may provide a better treatment of this type of recognition. It may be necessary to choose the most appropriate scoring function according to the precise nature of the binding site.

Water mediated interactions

For the most promiscuous protein complex, 1mwe, the two waters most significant for binding were included individually and together, to see how this altered the performance of GOLD.

From the full set there is no overall improvement on inclusion of one or more waters (Table 6). The ligands ST1, ST5 and ST6 were able to dock when one or more water was present (ST1 required water 328 specifically) but these ligands were originally removed from the clean set because doubts about their structures had been raised. For the clean set (Table 7), the docking was not significantly improved by inclusion of these waters but the inclusion of waters did prevent docking of ligands that displaced the water molecules.

Table 6 For the 21 ligands of the full set, percentage of ligands able to dock successfully into 1mwe with zero, one or two waters present in the binding site. Water 327 is involved in binding the N-acetyl group and water 328 in binding the C4 hydroxyl.

RMSD cut-off, Å	No water	328 and 327	328	327
<2.00	84%	81%	84%	84%
<1.50	77%	74%	74%	77%
<1.25	65%	55%	65%	58%
<1.00	45%	39%	29%	39%

Table 7 RMSD, Å, relative to the native crystal structure for ligand docking into 1mwe with neither, both, or a single water present. Percentage successfully docked is given for each case at four different cut-off values. ZAN and G20 displace water 328; RA1 and RA2 displace both water 328 and water 327.

Lig ref	pdb ref	No water	328 and 327	328 only	327 only
SIA	1mwe	1.28	1.33	1.22	1.32
EQP	1inv	0.87	1.15	1.19	1.17
AXP	1inw	1.24	0.72	1.05	0.82
DAN	1f8b	0.79	0.61	0.73	0.58
4AM	1f8c	0.65	0.66	0.66	0.66
9AM	1f8d	0.78	0.63	0.67	0.65
49A	1f8e	0.64	0.69	0.71	0.67
ZAN	1nnc	0.61	5.55	3.28	0.57
ST4	1inf	0.66	0.89	0.90	0.84
RA1	1b9t	1.05	2.13	2.48	2.50
RA2	1b9v	0.80	3.76	1.60	1.99
DPC	1a4q	1.24	1.35	1.04	1.07
G20	2qwi	0.84	5.50	5.37	1.15
G28	2qwj	0.86	0.87	0.84	0.87
G39	2qwk	1.37	1.34	1.37	1.39
RMSD < 2.00 Å		100%	73%	80%	93%
RMSD < 1.50 Å		100%	73%	73%	87%
RMSD < 1.25 Å		87%	53%	67%	73%
RMSD < 1.00 Å		67%	47%	40%	53%

GOLD does not appear to require the inclusion of relevant water molecules for good docking of ligands that make water mediated hydrogen bonds to neuraminidase. It appears that GOLD is robust enough to reproduce binding modes even when a favourable interaction with a water molecule is not included. This implies that the dominant interactions necessary for good docking are sufficient to overcome the omission of favourable interactions formed with mediating waters. This is useful because including water molecules may prevent the detection of novel compounds that displace 'conserved' waters. Examples where this might lead to false negatives in virtual screening are cyclic urea inhibitors in HIV-1 protease [49] or inhibitors with uncharged S1 binders in thrombin [50].

Conclusions

This is the first substantial study of a docking program on neuraminidase, and GOLD has shown itself to be an extremely good, robust docking program for this system. For example, after restriction to a clean test set of the most reliable neuraminidase structures, GOLD was able to dock all 15 ligands back into their own protein conformation (native docking) to within 1.5 Å. This is considerably better than the 70% success rate seen with GOLD against general validation sets. Additionally, it was possible to dock ligands into non-native proteins with a high success rate even for ligands that cause small but significant induced fit effects in the native complex. The top-ranked solution was within 1.5 Å heavy atom RMSD for all 15 ligands in the clean set, when docking was performed against the most promiscuous enzyme structure (PDB code: 1mwe). This is a success for the rigid protein approximation and it seems reasonable to suppose that virtual screening based on the most promiscuous enzyme structure would be a successful strategy. For this enzyme GOLD performs better than a previous smaller study with the program PRO_LEADS, and it has been shown that the GOLD scoring function is superior to the ChemScore function for neuraminidase. It has also been demonstrated that for neuraminidase, the inclusion of waters involved in water-mediated hydrogen bonds was not necessary to obtain the correct binding mode for these ligands.

References

1. Abagyan, R. and Totrov, M., Curr. Opin. Chem. Biol., 4 (2001) 375.
2. Walters, P.W., Stahl, M.T. and Murcko, M.A., Drug Disc. Today, 3 (1998) 160.
3. Taylor, R.D., Jewsbury, P.J., and Essex, J.W., J. Comput. Aided Mol. Des., 16 (2002) 151.
4. Najmanovich, R., Kutter, J., Sobolev, V., and Edelman, M., Proteins, 39, (2000) 261.
5. Sandak, B., Wolfson, H., and Nussinov, R., Proteins, 32 (1998) 159.
6. Kramer, B., Rarey, M., and Lengauer, T., Proteins, 37 (1999) 228.
7. Carlson, H.A., and McCammon, J.A., Mol. Pharmacol., 57 (2000) 213.
8. Bouzida, D., Rejto, P.A. Arthurs, S. Colmson, A.B., Freer, S.T., Gehlhaar, D.K., Larson, V., Luty, B.A., Rose, P.W., and Verkhivker, G.M., Int. J. Quantum Chem., 72 (1999) 73.
9. Luty, B.A., Wasserman, Z.R., Stouten, P.F.W., Hodge, C.N., Zacharias, M., and McCammon, J.A., J. Comput. Chem., 16 (1995) 713.
10. Leach, A., and Lemon, A., Proteins, 33 (1998) 227.
11. Schaffer, V., and Verkhivker, G.M., Proteins, 33 (1998) 74.
12. Claussen, H., Buning, C., Rarey, M., and Lengauer, T., J. Mol. Biol., 308 (2001) 377.
13. Knetgel, R.M.A., Kuntz, I.D., and Oshiro, C.M., J. Mol. Biol., 229, (1997) 424.
14. Carlson, H.A., Masukawa, K.M., and McCammon, J.A., J. Phys. Chem. A, 103 (1999) 10213.
15. Jones, G., Willett, P., Glen, R.C., Leach, A.R. and Taylor, R., J. Mol. Biol., 267 (1997) 727.
16. Schnecke, V., and Kuhn, L.A., Perspect. Drug Des. Discov., 20 (2000) 171.
17. Schnecke, V., Swanson, C.A., Getzoff, E.D., Tainer, J.A., and Kuhn, L.A., Proteins, 33 (1998) 74.
18. Koehl, P., Delarue, M., J. Mol. Biol., 239 (1994) 249.
19. Murray, C.W., Baxter, C.A., and Frenkel, A.D., J. Comput. Aided Mol. Des., 13 (1999) 547.
20. Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., and Bourne, P.E., Nucleic Acids Research, 28, (2000) 235.
21. Air, G.M., and Laver, W.G., Proteins, 6 (1989) 341.
22. Wade, R.C., Structure, 5 (1997) 1139.
23. Meindl, P., Bodo, G., Palese, P., Schulman, J., and Tuppy, H., Virology, 58 (1974) 457.
24. Von Itzstein, M., Wu, W.Y., and Kok, G.B., Nature, 363 (1993) 418.
25. Kim, C.U., Lew, W., Williams, M.A., Liu, H.T., Zhang, L.J., Swaminathan, S., Bischofberger, N., Chen, M.S., Mendel, D.B., Tai, C.Y., Laver., and Stevens, R.C., J. Am. Chem. Soc., 119 (1997) 681.
26. Varghese, J.N., Smith, P.W., Sollis, S.L., Blick, T.J., Sahasrabudhe, A., McKimm-Breschkin, and J.L., Colman, P.M., Structure, 6 (1998) 735.
27. Babu, Y.S., Chand, P., Bantia, S., Kotian, P.L., Dehghani, A., El-Kattan, Y., Lin, T.H., Hutchinson, T.L., Elliott, A.J., Parker, C.D., Ananth, S.L., Horn, L.L., Laver, G.W., and Montgomery, J.A., J. Med. Chem., 43 (2000) 3482.
28. Chand, P., Kotian, P.L., Dehghani, A., El-Kattan, Y., Lin, T.H., Hutchinson, T.L., Babu, Y.S., Bantia, S., Elliott, A.J., and Montgomery, J.A., J. Med. Chem., 44 (2001) 4379.

29. Gasteiger, J.; Sadowski, J.; Schuur, J.; Selzer, P.; Steinhauer, L., and Steinhauer, V., J. Chem. Inf. Comput. Sci., 36 (1996) 1030.
30. Baxter, C.A., Murray, C.W., Clark, D.E., Westhead, D.R., and Eldridge, M.D., Proteins: Struc. Funct. Genet., 33 (1998) 367.
31. Taylor, N.R., Cleasby, A., Singh, O., Skarzynski, T., Wonacott, A.J., Smith, P.W., Sollis, S.L., Howes, P.D., Cherry, P.C., Bethall, R., Colman, P., and Varghese, J., J. Med. Chem., 41 (1998) 798.
32. Finley, J.B., Atigadda, V.R., Duarte, F., Zhao, J.J., Brouillette, W.J., Air, G.M., and Luo, M., J. Mol. Biol., 293 (1999) 1107.
33. Smith, B.J., Colman, P.M.,Von Itzstein, M., Danylec, B., and Varghese, J.N., Protein Sci., 10 (2001) 689.
34. Sudbeck, E.A., Jedrzejas, M.J., Singh, S., Brouillette, W.J., Air, G.M., Laver, W.G., Babu, Y.S., Bantia, S., Chand, P., Chu, N., Montgomery, J.A., Walsh, D.A., and Luo M., J. Mol. Biol., 267 (1997) 584.
35. Jedrzejas, M.J., Singh, S., Brouillette, W.J., Air, G.M., and Luo, M., Proteins, 23 (1995) 264.
36. White, C.L., Janakiraman, M.N., Laver, W.G., Philippon, C., Vasella, A., Air, G.M., and Luo, M., J. Mol. Biol., 245 (1995) 623.
37. Jedrzejas, M.J., Singh, S., Brouillette, W.J., Laver, W.G., Air, G.M., and Luo, M., Biochemistry, 34 (1995) 3144.
38. Varghese, J.N., Colman, P.M., van Donkelaar, A., Blick, T.J., Sahasrabudhe, A., and McKimm-Breschkin, J.L., Proc. Natl. Acad. Sci. USA, 94 (1997) 11808.
39. Bossart-Whitaker, P., Carson, M., Babu, Y.S., Smith, C.D., Laver, W.G., and Air, G.M., J. Mol. Biol., 232 (1993) 1069.
40. Varghese, J.N., Epa, V.C., and Colman, P.M., Protein Sci., 4 (1995) 1081.
41. Burmeister, W.P., Henrissat, B., Bosso, C., Cusack, S., and Ruigrok, R.W., Structure, 1 (1993) 19.
42. Varghese, J.N., McKimm-Breschkin, J.L., Caldwell, J.B., Kortt, A.A., and Colman, P.M., Proteins, 14 (1992) 327.
43. Molecular Simulations Inc., San Diego, CA. InsightII, v2000, 2000.
44. Allen, F.H., Davies, J.E., Galloy, J.J., Johnson, O., Kennard, O., Macrae, C.F., Mitchell, E.M., Mitchell, G.F., Smith, J.M., and Watson, G.G., J. Chem. Inf. Comput. Sci., 31 (1991) 187.
45. Nissink, J.W.M., Murray, C.W., Hartshorn, M.J., Verdonk, M.L., Cole, J.C. and Taylor, R. Proteins, 49 (2002) 457.
46. www.astex-technology.com/AstexViewer and www.astex-technology.com/AstexViewer/neuraminidase
47. Hartshorn, M.J., J. Comput. Aided Mol. Des., 16 (2002) 871–881.
48. Verdonk, M.L, Cole, J.C., Hartshorn, M.J., Murray, C.W., and Taylor, R.D. *to be published.*
49. Lam, P.Y.S., Jadhav, P.K., Eyermann, C.J., Hodge, C.N., Ru, Y., Bacheler, L.T., Meek, O.M.J. and Rayner, M.M., Science, 263 (1994) 380.
50. Tucker, T.J., Brady, S.F., Lumma, W.C., Lewis, S.D., Gardell, S.J., Naylor-Olsen, A., Yan, Y., Sisko, J.T., Stauffer, K.J., Lucas, B.J., Lynch, J.J., Cook, J.J., Stranieri, M.T., Holahan, M.A., Lyle, E.A., Baskin, E.P., Chen, I.W., Dancheck, K.B., Krueger, J.A., Cooper, C.M., and Vacca, J.P., J. Med. Chem., 41 (1998) 3210.