Journal of Computer-Aided Molecular Design, **16:** 479–499, 2002.
KLUWER/ESCOM

De novo ligand design with explicit water molecules: an application to bacterial neuraminidase

Ricardo L. Mancera*
De Novo Pharmaceuticals Ltd., Compass House, Vision Park, Chivers Way, Histon, Cambridge CB4 9ZR, United Kingdom.

Received 26 September 2001; Accepted in final form 21 August 2002

Key words: De novo drug design, structure generation, water molecules, ligand-protein interactions, X-ray crystal structures, hydration, neuraminidase.

Summary

Most computer-aided drug design methods ignore the presence of crystallographically-determined water molecules in the binding site of a target protein. In this paper, our *de novo* ligand design methods are applied to the X-ray crystal structure of bacterial neuraminidase in the presence of some selected water molecules. We have found that, for this particular protein, the complete removal of all bound water molecules leads to difficulties in generating any potential ligands if the unsatisfied hydrogen-bonding sitepoints left by removing these water molecules are to be satisfied by a ligand. As more of the crystallographically determined water molecules are allowed in the binding site, it becomes much easier to generate ligands in larger numbers and with wider chemical diversity. This example shows that, in some cases, bound water molecules can be more accessible for hydrogen bonding to an incoming ligand than the actual protein binding sitepoints associated with them. From the point of view of *de novo* ligand design, water molecules can thus act as versatile amphiprotic hydrogen-bonding sitepoints and reduce the conformational constraints of a particular binding site.

Introduction

Most *de novo* and structure-based computational approaches to drug design make use of the known molecular architecture of a binding site as revealed by X-ray crystallography [1-11]. The crystal structure of a free or complexed protein can be analysed to extract its most relevant features: the size and shape of the binding site, the distribution of hydrogen-bonding sitepoints and hotspots, the location of hydrophobic regions, electrostatic potential energy maps and the binding mode of any bound ligand. However, more often than not, any ordered water molecules that appear in the crystal structure are removed prior to any such analysis or the subsequent ligand design methodology applied. Only occasionally may one or two catalytic water molecules be retained for the design of enzyme inhibitors, based on further experimental structural or kinetic data. A detailed analysis of tightly bound water molecules is also important because they can be used to design new ligands acting to mimic their position and interactions in a binding site [12].

From a thermodynamic point of view, it is necessary to assess the effect of the presence of water molecules at or near the protein-ligand interface. It has been estimated that the inclusion of water molecules can increase the affinity of ligands for their protein receptors up to a limit of -9 kJ/mol in the entropy and of -16 kJ/mol in the enthalpy, at 298 K, for the transfer of a water molecule from the bulk solvent [13,14]. It then becomes clear that the inclusion of a single water molecule into the ligand-protein interface can potentially decrease the free energy of interaction up to -7.0 kJ/mol, leading to an increase in affinity of more than one order of magnitude [15]. Consequently, water molecules trapped in the binding site between the protein and its ligand can outweigh the negative entropic cost of their restricted freedom of movement

*E-mail: Ricardo.Mancera@denovopharma.com

by providing hydrogen-bonding interactions to both the protein and the ligand. These results constitute compelling evidence for the need to ensure that any strategy for drug design takes into account the possibility of using crystallographic water molecules to increase ligand binding to a receptor protein.

Several studies have reported that some water molecules occupy the same positions in crystal structures of the same enzyme under different conditions [16] and/or different ligands [17-19], or in a set of enzyme structures which are structurally related to each other [20-23]. There have also been attempts to predict ordered hydration sites by modular neural networks using protein sequence information [24]. Genetic algorithms have been used to predict polar ligand interactions mediated by conserved water molecules [25]. A cluster analysis of consensus water sites in thrombin and trypsin has shown the conservation of these sites between serine proteases and their contribution to ligand specificity [26]. Empirical relationships between water occupancy and accessibility and structural properties of the protein-solvent interface have also been studied by statistical analysis of crystal structures [27-29].

We have developed a set of empirical rules to determine which water molecules in a binding site are likely to be displaced by an incoming ligand and which water molecules should be retained for their use in a ligand design application. These rules take into account the hydrogen-bonding energy, the crystallographic B-factors, the number of contacts with the protein and the solvent-accessible surface of the water molecules (A.T. García-Sosa and R.L. Mancera, manuscript in preparation). Such water molecules contribute to the plasticity of a binding site, mainly by modifying its shape and by participating as hydrogen-bonding groups. Often these water molecules can be seen acting as hydrogen-bonding 'bridges' between the protein and a ligand in the crystal structure of their complexed form, establishing a complex network of hydrogen bonds that stabilises the protein-ligand interaction [30-34]. It would be desirable for the drug designer to be able to exploit those same features when designing new ligands, so that novel molecular scaffolds can bind effectively to a binding site not only through direct ligand-protein interactions but also through ligand-water interactions.

The general principle of including the treatment of hydration when dealing with ligand-protein interactions is by no means new [35]. There are many techniques currently used to deal with the energetic and entropic contributions of aqueous solvation. Implicit hydration methods treat the solvent as a dielectric continuum which can be polarised by the protein (and ligand) and thus produce a reaction field on such species [36]. These methods tend to be fast and easy to implement, while showing moderate to good efficacy in accounting for the solvation contributions to the free energy of binding of a ligand to its receptor [37]. Explicit hydration methods make use of free energy perturbation methods [38,39] and linear-interaction energy models [40,41] within molecular dynamics or Monte Carlo simulations. These methods include large numbers of explicit water molecules around a protein binding-site, tending to be slower and more involved in their implementation, but producing significantly better results than implicit hydration methods.

While such approaches can in principle provide a reasonably good description of the energetics of ligand-protein binding, there is still the need to be able to include explicit water molecules in the binding site of a protein and use them to modify the chemical environment where new ligands are to be designed. In fact, the use of explicit water molecules has already been used in ligand docking [42] and QSAR studies [43].

Our approach so far has relied on the information provided by crystal structures of free and complexed proteins to determine rules for classifying water molecules as displaceable or bound. Care must be taken when using such crystallographically-determined water molecules, since the reliability of the X-ray determination of water molecules decreases as the resolution decreases. The purpose of including a 'rational' selection of water molecules in a drug design application is to increase the plasticity of the binding site and, consequently, to provide with alternative strategies for the generation of novel molecular scaffolds.

In this paper we present the results of the use of such bound water molecules in the binding site of neuraminidase from Salmonella typhimurium with our ligand design methods. We show how, in this particular case, the presence of such water molecules has a profound effect on the strategies followed for *de novo* ligand design and the nature of the molecular scaffolds that can be generated.

Methods

The structure of neuraminidase (E.C. 3.2.1.18) from Salmonella typhimurium was extracted in its free (2SIL) and complexed form (2SIM) from the Protein Data Bank [44]. The complexed form illustrates the binding mode of a sialic acid-based inhibitor: 2-deoxy-2,3-dehydro-N-acetyl-neuraminic acid (Neu5Ac2en or DANA). The crystal structures of these forms of the protein have been resolved to a high resolution (1.6 Å) [45,46] and provide a good example of a highly hydrated protein of pharmaceutical interest.

Analysis of water molecules in the binding site

A statistical analysis had been previously performed on a test set of proteins which included the 2SIL/2SIM pair (A.T. García-Sosa and R.L. Mancera, manuscript in preparation) to determine which water molecules should be classified as bound (all the others being displaceable). Our method determines the statistical correlation between various structural properties of water molecules in an uncomplexed protein and the probability of observing the same water molecules at the same location in the complexed form of the protein.

This analysis showed that four water molecules (HOH 613, HOH 689, HOH 692 and HOH 694) were bound to the binding site. Two of them (HOH 613 and HOH 689) are buried deep in the binding cavity and are 'trapped' between the binding site and the ligand DANA. All four water molecules were included in the strategy for *de novo* ligand generation as hydrogen-bonding sitepoints, while simultaneously modifying the steric shape of the binding site.

Binding site analysis

The binding site of 2SIM, the complexed form of neuraminidase, was stripped of its ligand and water molecules. We analysed the binding site to identify all accessible hydrogen-bonding atoms that could form hydrogen bonds (sitepoints) [47] and any overlapping hydrogen-bonding regions (hotspots) [48]. Hotspots provide useful information regarding the extent to which one or more sitepoints can contribute to the hydrogen-bonding properties of a binding site [48]. The hotspots were then used to prioritise and/or subdivide sitepoints into neighbouring groups.

Separately, hydrogens were added to the protein structure assuming a pH of 7.4 and standard aminoacid pKa using the Biopolymer module in MSI Insight 2000. All water molecules were then removed from the system, except for the ones indicated above. The system was then ready to be used in the ligand generation stage.

De novo *ligand design*

De novo ligand design was carried out using the programme SKELGEN [49-51]. This in-house programme can generate molecular scaffolds in a binding site based on a Monte Carlo simulated annealing minimisation algorithm. An in-house library of the most common acyclic and cyclic fragments was compiled from the World Drug Index after filtering for toxic or synthetically undesirable fragments. Structures are built by connecting such molecular fragments. Given a random initial structure, it is modified in a succession of steps by using several types of 'move'. Structures can change by fragment additions, fragment removals, fragment mutations, molecular translations and rotations, and conformational changes. Thus, by applying these moves it is possible to remove previously incorporated fragments or replace them with different fragments, allowing the ligand to gradually adjust to the binding site constraints. The combinatorial problem of assembly of molecular structures in such a way is addressed by using simulated annealing optimisation to direct the molecular evolution of a scaffold. There are a number of requirements on the ligands that have to be satisfied before a structure is accepted, which include a favourable free energy of binding, a low internal energy, synthetic feasibility and receptor specificity (see below). All of these are expressed as penalties within a scoring function that vanish when satisfied. In particular, the issue of synthetic feasibility of the ligands is accounted for by using fragments from synthetically-accessible molecules and adding penalties for undesirable combinations of atoms when joining fragments. Other chemical rules allow the user to control the molecular size and diversity of the ligands generated. Full details of all the algorithms used as mentioned above can be found elsewhere [49-51].

Receptor specificity was considered by biasing ligand generation to explore important binding-site regions. These regions of interest are specified by combinations of hydrogen-bonding sitepoints, which were introduced as constraints (compulsory hydrogen bonds with the receptor), within the boundaries of an appropriately selected simulation box and a steric grid on the binding site. For the generation of any ligand,

an initial high temperature was used for the simulated annealing optimisation, which was cooled down by a factor of 0.9 between successive Markov chains up to a maximum of 40 chains, each 2000 steps long. A weighted score average of all the constraints imposed on the system (sitepoints, steric clashes and chemical rules, all usually carrying a unitary weight) was obtained during each Markov chain that attempted to generate a ligand, and only scores of 0.1 or less were acceptable for a ligand to be generated without severe steric clashes with the receptor and satisfying the hydrogen-bonding constraints imposed. At this point the annealing was terminated and the optimised molecular scaffold was outputted. Details of the scoring function and simulated annealing procedure can be found elsewhere [49].

Due to the restrictions imposed by using a fragment database and certain chemical rules, the chemical diversity of the ligands was not exhaustively explored. It should be borne in mind that our purpose was not to generate known or novel actual inhibitors, but to investigate the effect that the use of explicit water molecules can have on the above described ligand generation method. The actual suitability as potential inhibitors of the ligands generated is limited by the fragment database used, the chemical rules used to join fragments and the strategy for drug design followed.

Evaluation of de novo *ligands*

All *de novo* ligands generated by SKELGEN were post-processed to analyse their chemical nature, their binding modes and their hydrogen-bonding patterns with the protein.

A list of all putative hydrogen bonds made by the ligand with the protein was obtained following the same criteria as in the Ligand-Protein Contacts approach [52]. Only putative donor-acceptor hydrogen bonding pairs of 3.3 Å in length or less were considered. This distance threshold provided a sufficiently large list of potential hydrogen-bonding atom pairs to be used as constraints in the initial stages of the energy minimisation protocol (see below).

Ligands were given a Tanimoto index score based on the pairwise chemical similarity of their constituting fragments, and were subsequently clustered into families. Ligands in each family were manually filtered based on the presence of undesirable chemical or conformational features and, most importantly, on the basis of common binding modes. As a result of this procedure, many undesirable ligands were filtered out: this included duplicate structures, structures with unlikely conformations and chemically equivalent structures with similar binding modes. The energy minimisation protocol described below helped to filter out several structures whose binding modes were lost during the energy minimisation due to steric clashes, insufficient hydrogen bonding and/or internal conformational strain. The objective was to obtain a final list of ligands with both relatively unique chemical structures and binding modes.

Energy minimisation of ligand-protein interactions

The above selected ligands were all energy-minimised in the binding site of 2SIM using the Discover3 module in Insight2000 (Accelrys). The CFF forcefield [53] was used throughout. For some ligands, additional torsional or out-of-plane restraints were imposed to ensure planarity of aromatic or conjugated systems. The cell-multipole method [54,55] was used to treat non-bonded interactions, with a dielectric constant of 2.0 for the electrostatic interactions.

Although the protein receptor was kept entirely rigid in its original crystal conformation, hydrogen atoms in the binding site were allowed re-orientational freedom during the minimisation in order to optimise the hydrogen-bonding network between the ligand, the water molecules and the protein. For this purpose, all hydrogens in any aminoacid residue with at least one atom in contact (maximum distance of 3.5 Å) with any atom of the ligand were not kept fixed during the minimisation. The ligands were always allowed full flexibility during the minimisations, and water molecules had re-orientational freedom (through the optimisation of the position of their hydrogen atoms) but were kept fixed in their original positions.

The minimisations were carried out in a staged manner to try to retain the original binding mode. Consequently, this is a crucial stage in the post-processing of our *de novo* generated ligands. The potential energy landscape for a receptor-ligand pair can be quite complex, and subtle changes in the balance between electrostatic and steric interactions can easily shift a ligand from two neighbouring low energy binding modes. There are also four reasons why an energy minimisation is required for the ligands generated by SKELGEN. Firstly, this programme evaluates steric clashes in an approximate manner: any ligand atom crossing a pre-determined surface grid on the receptor is given a penalty according to the number of close contacts [49]. Secondly, the intramolecular steric

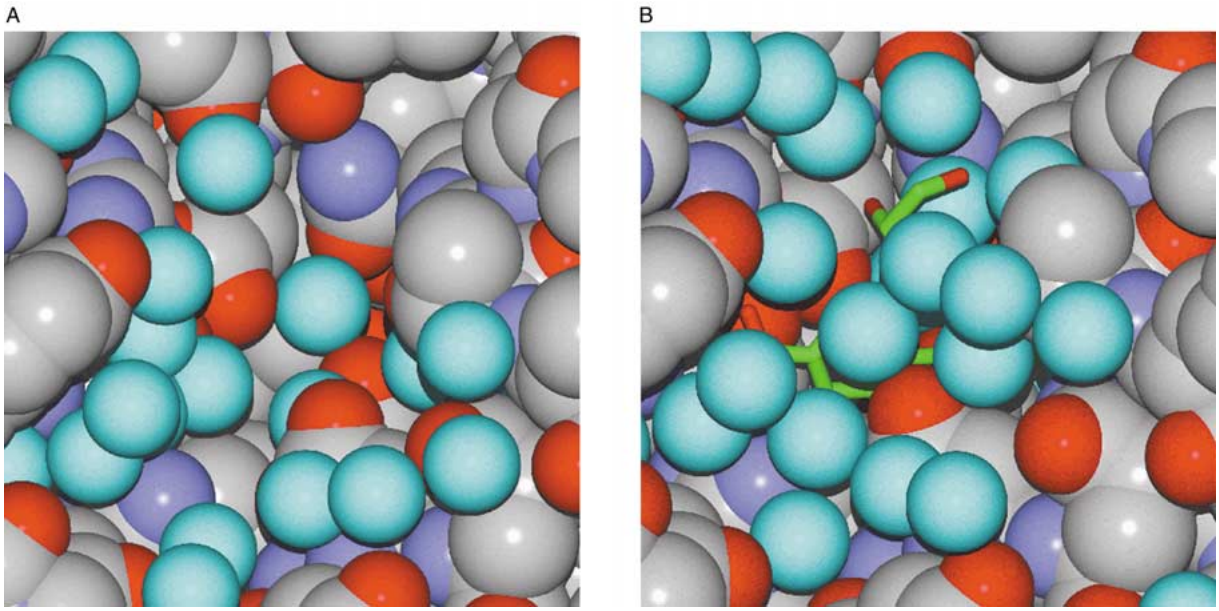

Figure 1. Comparison between the (a) free (2SIL) and (b) complexed (2SIM) forms of neuraminidase. The empty site is highly hydrated with many water molecules, some of which are displaced upon binding of the ligand.

term between non-bonded atoms uses a cut-off method between all non-hydrogn atoms, which is a crude approximation [49]. Thirdly, hydrogen-bonding constraints are satisfied by placing a ligand atom within a certain distance and at a certain angle with respect to a given sitepoint; consequently, there is no energy criterion applied to optimise such interactions. And fourthly, the use of a consistent forcefield for the protein and the ligand allows us to take into account in a more realistic manner hydrogen bonding, the intra- and inter-molecular van der Waals attractions and repulsions and the internal potential energy terms for the ligand. As a consequence, some of the binding modes generated by SKELGEN can potentially be lost during an energy minimisation to nearby low-energy states.

Initially, all putative hydrogen-bonded pairs (see above) with distances of 3.3 Å or less were given quadratic constraints with force constants of 100 kcal/mol. At the same time, van der Waals interactions were scaled down to 10% of their real value. This was done to allow for hydrogen re-orientation while preventing the ligand from experiencing high repulsion forces arising from any steric clashes present in the complex. An initial minimisation of 1000 steps using steepest descents [56] was performed. Van der Waals interactions were scaled up to 33% of their real value and a second minimisation of 1000 steps using steepest descents was carried out. At this point, the hydrogen bond quadratic constraints were modified to represent a weak constraint, being given smaller force constants of 10 kcal/mol. A third minimisation of 1000 steps using steepest descents was then carried out. At this point all hydrogen bond constraints were completely removed and a fourth minimisation of 1000 steps using conjugate gradients was performed. This was followed by all van der Waals interactions being scaled up to 67% of their real value. A fifth minimisation of 1000 steps using conjugate gradients [56] was then carried out. Finally, van der Waals interactions were scaled up to 100% of their real value. A sixth and final minimisation of 1000 steps using conjugate gradients was then performed.

It should be noted that all the previously described minimisations, except for the last two, were carried out for 1000 steps or up to a convergence gradient of 0.1 kcal/mol, whichever was reached first. In the last two minimisations, the gradient for convergence was reduced to 0.01 kcal/mol. Each full energy minimisation took an average of 4 minutes to complete on an SGI R10000 processor.

Results and discussion

The crystallographic studies of the bacterial neuraminidase in its complexed form with different inhibitors reveal that the active site is in a well-defined

and rigid conformation with very little flexibility [45,46]. Inhibitors have to acquire a certain degree of distorsion to occupy the space available and to allow their negatively charged carboxy and phosphonate groups (see below) to interact with the binding-site arginine residues [46]. These observations make this target well suited for our ligand design purposes, as molecules can be then designed against a structurally well-conserved binding site. It also adds further credibility to our use of explicit water molecules, as there are no expected conformational changes in the binding site upon ligand binding which could remove one or more of the water molecules. Furthermore, only a few sialic acid-based compounds and phosphonic acid analogues have been reported to be inhibitors of the bacterial neuraminidases and just three have actually been co-crystalised [46,57,58]. Consequently there is scope for generating novel molecular scaffolds in the active site of the enzyme that can act as tight binders, although that is not the purpose of this work.

If we compare the crystal structure of the native protein (2SIL) with the complexed form (2SIM), we can see that the main differences in solvent structure between the two forms are close to the binding site, as can be seen in Figure 1. In the native crystal form, water molecules fill the space occupied by the carboxylate oxygens and the O4 of DANA. There is also a water molecule interacting with Tyr 342 and one in the region of the N-acetyl group of DANA [46]. The complexed form is a co-crystal, obtained by soaking the protein crystal with a solution of the ligand (DANA) [46]. Upon binding of the ligand, some of the above-mentioned water molecules are expelled from the binding site, but eight more water molecules can be found in the complex, presumably through the existence of new ligand-solvent interactions. One water molecule is part of a hydrogen-bonding network around the carboxylate group of DANA, while the other seven water molecules occupy positions between the binding site residues and the glycerol chain of DANA, bridging the interactions between this ligand and the protein [46]. As explained earlier in the Methods section, only a few of the water molecules in the complex can be considered to be truly bound to the protein surface.

The binding mode observed for DANA is identical to that observed for the phosphonate analogs [46], and further comments are provided below in the description of the different regions of the binding site. However, it is clear that the arginine triad is responsible for anchoring strongly the carboxylic or phosphonic acid groups of these compounds, thus determining the main features of their binding mode. We provide an account of the main features of the binding mode of our designed ligands.

Binding site analysis

The first stage in a typical ligand design strategy involves the analysis of the binding site of a protein and the selection of appropriate sitepoints and, in this case, of those water molecules that are deemed to be bound to the binding site. This allows for the size and shape of the binding cavity to be defined; in the case of the water molecules, their gradual or total removal re-defines the binding site accordingly.

We obtained a list of sitepoints: hydrogen-bond donors, hydrogen-bond acceptors and hydrogen-bond amphiprotic groups (typically -OH groups) within 6.0 Å of any atom of the ligand [47]. We then generated hotspots or clusters of overlapping hydrogen-bonding regions (normally between two or three hydrogen-bonding groups) within the binding site and up to a distance of 5.0 Å from any atom of the ligand [48].

The results of these computations can be seen in Figure 2. There is a large spread of hydrogen-bond donors (in purple) lining the front of the cavity, made up by Arg 246, Arg 309, Arg 37 and Arg 56. There are a few more hydrogen-bond donors in deeper parts of the binding site, defined by Gln 63, Trp 128 and Gln 194. The hydrogen-bond acceptors (in yellow) are Asp 62, Asp 100 and Glu 231, and line the middle of the binding site, from front to back. The hydrogen-bonding amphiprotic groups (in white) are more randomly distributed in the binding site, and are made up of Tyr 342, Thr 127 and Ser 230.

Interestingly, there are several hotspot clusters in the binding site, some of which have ligand polar atoms in contact or in close proximity. Water molecules HOH 613, HOH 689 and HOH 692 are relatively close to three of these hotspot clusters. The wide spread of the various hotspot clusters shows that the binding site of 2SIM is a rather versatile cavity for hydrogen bonding.

As we shall describe below, the binding site was partitioned into several regions based on the position of the various sitepoints and their associated hotspots. Such partitioning follows the spirit of our ligand generation programme, SKELGEN, which allows for the separation of hydrogen-bonding constraints (the site-

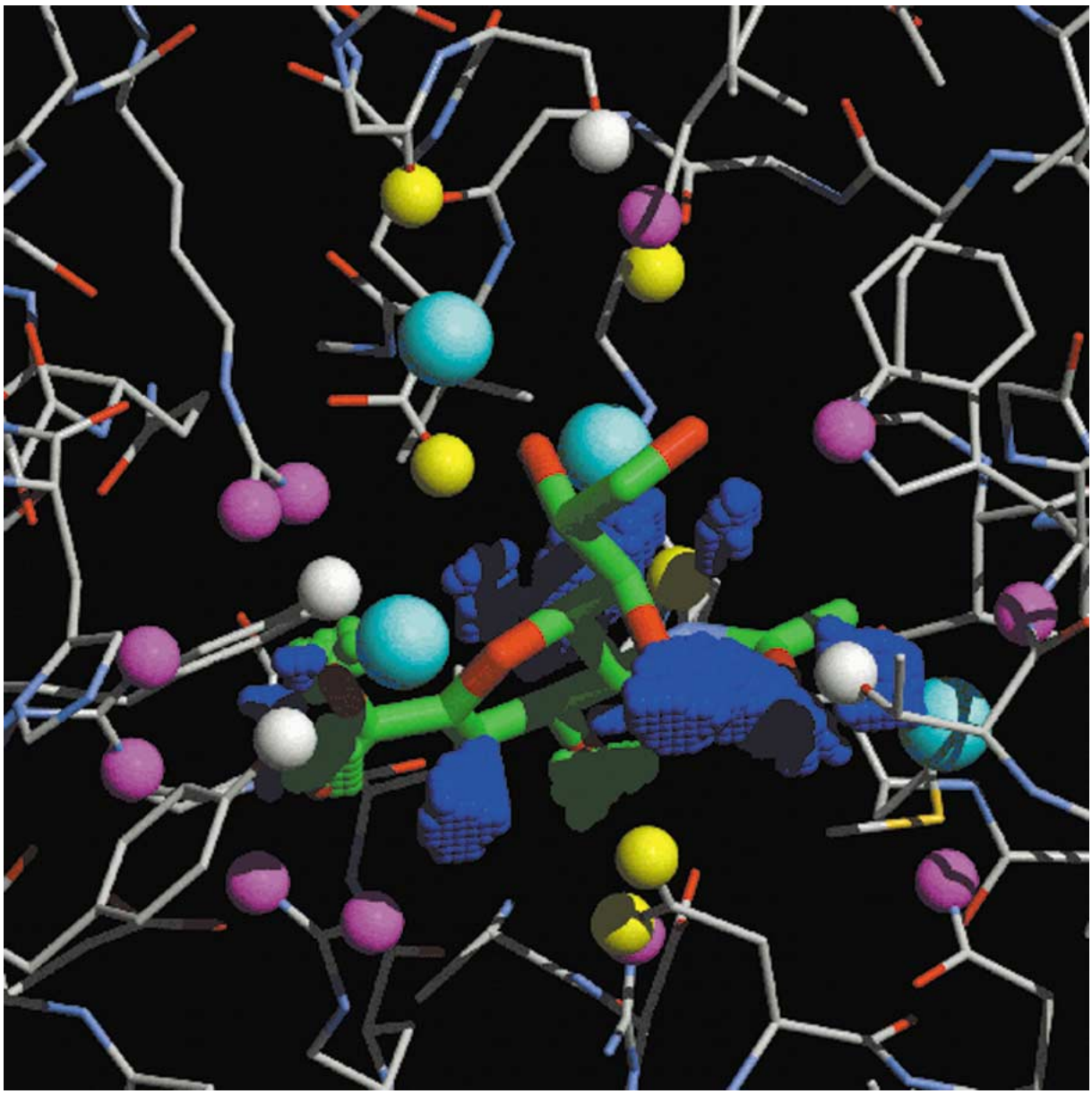

Figure 2. The site analysis of 2SIM reveals several kinds of sitepoints: hydrogen-bond donors (in purple), hydrogen-bond acceptors (in yellow) and hydrogen-bond amphiprotic groups (in white). There are two kinds of hotspot clusters: where two hydrogen-bonding groups overlap (in blue) and where three hydrogen-bonding groups overlap (in green).

points) into separate groups that have to be satisfied independently.

Ligand generation strategies

The second stage of the ligand design process involves defining a benchmark strategy for the generation of de novo ligands through the selection of sitepoints and their division into appropriate regions within the binding cavity. Such regions can be defined, for example, in terms of a catalytic region, a co-factor region and adjacent pockets and/or grooves. The regions can also be defined in a more intuitive manner by the clusters of hotspots, their occupation by ligand atoms (in a co-crystal), the position of water molecules and the presence of hydrophobic patches on the surface of the protein. As described below, we have taken a mixed

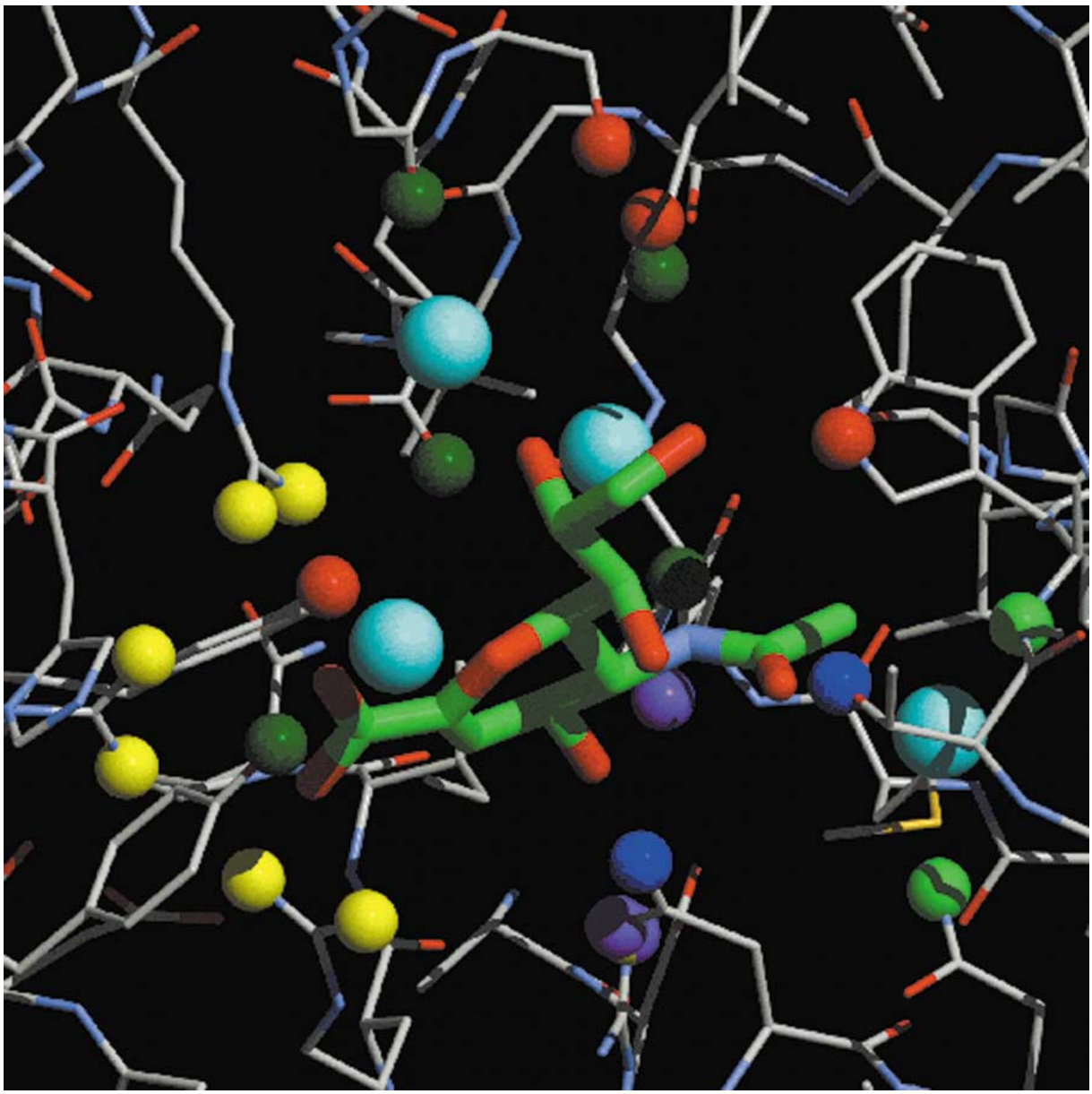

Figure 3. The binding site of 2SIM, with its bound ligand and the four selected water molecules, has its sitepoints divided into different regions. The sitepoints have been coloured according to the region they belong to: Region 1 in yellow, Region 2 in purple, Region 3 in blue, Region 4a in green, Region 5a in dark green and Region 6 in red.

approach to the way that the binding site of 2SIM can be partitioned.

Our benchmark strategy for de novo ligand generation subdivided the binding site of 2SIM into six regions, each one with its own set of hydrogen-bonding sitepoints. Figure 3 provides a colour-coded graphic illustration of the different regions described below.

(1) Region 1, with six sitepoints. This was chosen to be the catalytic arginine triad: NH1 and NH2 of Arg 37, NH1 and NH2 of Arg 246, and NH1 and NH2 of Arg 309. These arginines stabilise either the carboxylic acid group of natural sialic acid derivatives or the phosphonate group of analogues [45,46]. In the case of DANA, Arg37 and Arg309 form hydrogen

bonds to its O1A [46]. Only three sitepoints are to be satisfied.

(2) Region 2, with three sitepoints. This was chosen to be comprised of the sitepoints associated with a small spread of hotspots: OD1 of Asp 62, NH2 of Arg 56 and OD1 of Asp 100. These sitepoints are responsible in the complexed co-crystal structure (2SIM) for one weak and two strong hydrogen bonds to the O4 of DANA [45,46]. Only three sitepoints are to be satisfied simultaneously with Region 3 below.

(3) Region 3, with two sitepoints. This was chosen to encompass the sitepoints associated with a large spread of hotspots in the front of the binding site: OD2 of Asp 62 and OG1 of Thr 127. Only three sitepoints are to be satisfied simultaneously with Region 2 above.

(4) Region 4a, with two sitepoints. This reflects the sitepoints that become available when bound water molecule HOH 689 is removed, as well as an associated hotspot region: NE2 of Gln 63 and N of Trp 128. Both sitepoints are to be satisfied since they are very close to each other.

(5) Region 4b, with one sitepoint. This is bound water molecule HOH 689, which probably acts as a hydrogen-bond donor, since the associated sitepoints (see above) would donate hydrogen bonds to this water molecule. This sitepoint is to be satisfied.

(6) Region 5a, with five sitepoints. This reflects the sitepoints that become available when bound water molecules HOH 613, HOH 692 and HOH 694 are removed: OD2 of Asp100, OE1 of Gln 194, O of Gly 229, OE2 of Glu 231 and OH of Tyr 307. Asp 100 also forms a hydrogen bond to the N-acetyl nitrogen of DANA in the co-crystal complex [46]. Only one sitepoint is to be satisfied.

(7) Region 5b, with thee sitepoints. This comprises the bound water molecules HOH 613, HOH 692 and HOH 694, which were treated as donor sitepoints due to the number of associated sitepoints than can donate hydrogen bonds to them. Only one sitepoint is to be satisfied.

(8) Region 6, with four sitepoints. This includes the sitepoints associated with other small spreads of hotspots: NE1 of Trp 128, NE2 of Gln 194, OG of Ser 230 and OH of Tyr 342. Among these sitepoints one can find the hydrogen bond donor of Trp 128 responsible for one direct weak hydrogen bond to the O9 of DANA [45]. Incidentally, it has already been said that Ser 230 can only interact with a ligand through water molecules [46]. This region is rather spread out and is meant to include some extra sitepoints; as a consequence, only one sitepoint is to be satisfied.

Table 1. Differences between the ligand generation strategies

Run	Water molecules present	Ligand generation (%) success rate (%)
1	689, 613, 692, 694	25
2	689, 613	23
3	689	2
4		0

Table 1 describes the differences between each of the strategies used for ligand generation, as well as the rate of sucess for generating ligands prior to any selection or filtering of structures, as described earlier. The only difference between the various runs was the number of water molecules that were allowed in the binding site. Run 1 included all four water molecules. Run 2 included the deeper water molecules HOH 613 and HOH 689, but excluded water molecules HOH 692 and HOH 694. Run 3 only included water molecule HOH 689. Finally, Run 4 had no water molecules at all, being equivalent to a 'standard' ligand generation procedure without any water molecules and including all unsatisfied hydrogen bonding sitepoints. As can be seen from Table 1, regions 4a/b and 5a/b define those sitepoints which are available depending on whether the corresponding water molecules are present or not. Consequently, the removal of any water molecule meant the inclusion of its associated sitepoints, which resulted in more sitepoints available for a designed ligand.

The purpose of all the above strategies was to investigate the effect of gradually reducing the number of water molecules in the binding site on the *de novo* ligand generation process. Due to the nature of the binding site of 2SIM, the presence of water molecules adds accessible sitepoints on the surface of the binding cavity. The gradual removal of some of the water molecules and their substitution with surface sitepoints decreases the accessibility of the corresponding sitepoints (even though their numbers increase). We were naturally interested in observing the effect this would have on the success rate of ligand generation and, most importantly, on the chemical nature and diversity of the ligands generated.

Ligand structures and binding modes

We have observed that the orientational conformation of water HOH 689 is in general well conserved with all ligands. This is probably due to the translational constraints imposed and the strong hydrogen bond made by this water molecule with the backbone nitrogen atom of Trp 128.

Considering that only hydrogen atoms were allowed to have conformational flexibility during the energy minimisations, the only sidechain that showed a significant range of conformational flexibility was Thr 127. This was, of course, restricted to rotations of the -OH group. On the other hand, the terminal -OH group of Tyr 342 showed a strong preference for forming an intramolecular hydrogen bond with the terminal carboxylate of Glu 231. This is in contrast to the suggested intramolecular hydrogen bonds between Tyr 342 and Arg 246, and between Glu 231 and Arg 246, based on the co-crystal structure [46].

RUN1. SKELGEN's success rate for generating bound ligands was around 25%, meaning that from 100 attempted runs to generate a ligand, in only 25 of them a ligand with a sufficiently low score was obtained. This value of 25% constitutes our benchmark rate for de novo ligand design in the presence of all selected water molecules HOH 692, HOH 694, HOH 613 and HOH 689. The presence of all water molecules replaces several sitepoints (see Ligand generation strategies above) by hydrogen-bonding sitepoints that can be more accessible to an incoming ligand. The chemical structures of 19 selected unique ligands can be seen in Figure 4, while Figure 5 shows a two-dimensional representation of the binding modes of about half of these ligands after a full energy minimisation.

The first observation that can be made is that the main molecular anchor group interacting with the arginine catalytic triad is in most cases a tetrazole ring, both in its deprotonated form and in its 2-H tautomer. Molecules 1-1 to 1-16 have a tetrazole-based anchor group. This is most certainly due to the fact that Region 1, which contains the arginine triad, has several sitepoints out of which three had to be satisfied. This can easily be accomplished by a tetrazole ring but not by many other common rings. Another alternative was found with carboxylic acids (as expected from the structure of DANA), as seen in molecules 1-17 and 1-18 (molecule 1-10 uses the tetrazole ring as anchor group). Molecule 1-19 also has an alternative interesting anchor group: a lactone is able to form hydrogen bonds with all three arginines.

A kind of compound that was repeatedly generated contained a tetrazole ring linked to either a 2-pyridone or a 2,4-pyrimidinedione (uracil) ring, as represented by molecules 1-11, 1-12 and 1-13. All of them have the same binding mode, with their tetrazole ring forming hydrogen bonds with at least two of the three arginines in the catalytic triad. Only the deprotonated tetrazole-based compound (molecule 1-11) forms hydrogen bonds with all three arginines, as it lacks a hydrogen bond with Asp 62 (which both molecules 1-12 and 1-13 have) and is thus able to approach more closely all arginines. The only acceptable protonated form for the tetrazole ring was the 2-H tautomer, since the 3-H tautomer generated an electrostatic repulsion with the arginine triad and could not retain the same binding mode.

Another kind of compound that was generated on several occasions was made up of a tetrazole ring, a sulfonamide linker and either another tetrazole ring (as in molecule 1-14) or a different kind of heterocycle (as in molecule 1-15). Although in both compounds there is a strong hydrogen bond between their NH group and Asp 62, the binding modes are slightly different due to the different hydrogen-bonding nature of the second heterocycle, as can be seen in Figure 5. Furthermore, molecule 1-15 forms no hydrogen bonds with any water molecule in its minimised form, reflecting more dominant hydrogen bonding with Trp 128 and Arg 56. Molecule 1-14 cannot form a hydrogen bond with HOH 689 because it has a strong hydrogen bond between the second tetrazole ring and Asp 100.

Molecule 1-16 has an interesting binding mode. Despite the fact that it has a sulfonamide linker (as described before), due to the lack of a tetrazole ring the sulfono group becomes the anchor group, which forms several hydrogen bonds with the arginine triad, as seen in Figure 5.

When comparing the binding mode of DANA (and the phosphonate analogs) with those of the ligands generated here, it can be seen that the main feature that they all have in common, as expected, is a strong anchoring to the arginine triad. There also seems to be a tendency for some of the ligands to form a hydrogen bond with Asp 62, a similar feature of DANA, but hydrogen bonding is widespread on the binding site, reflecting the accessibility of various sitepoints.

Almost all ligands in their minimised conformations could form one or more hydrogen bonds with the water molecules. Water molecules HOH 689 and

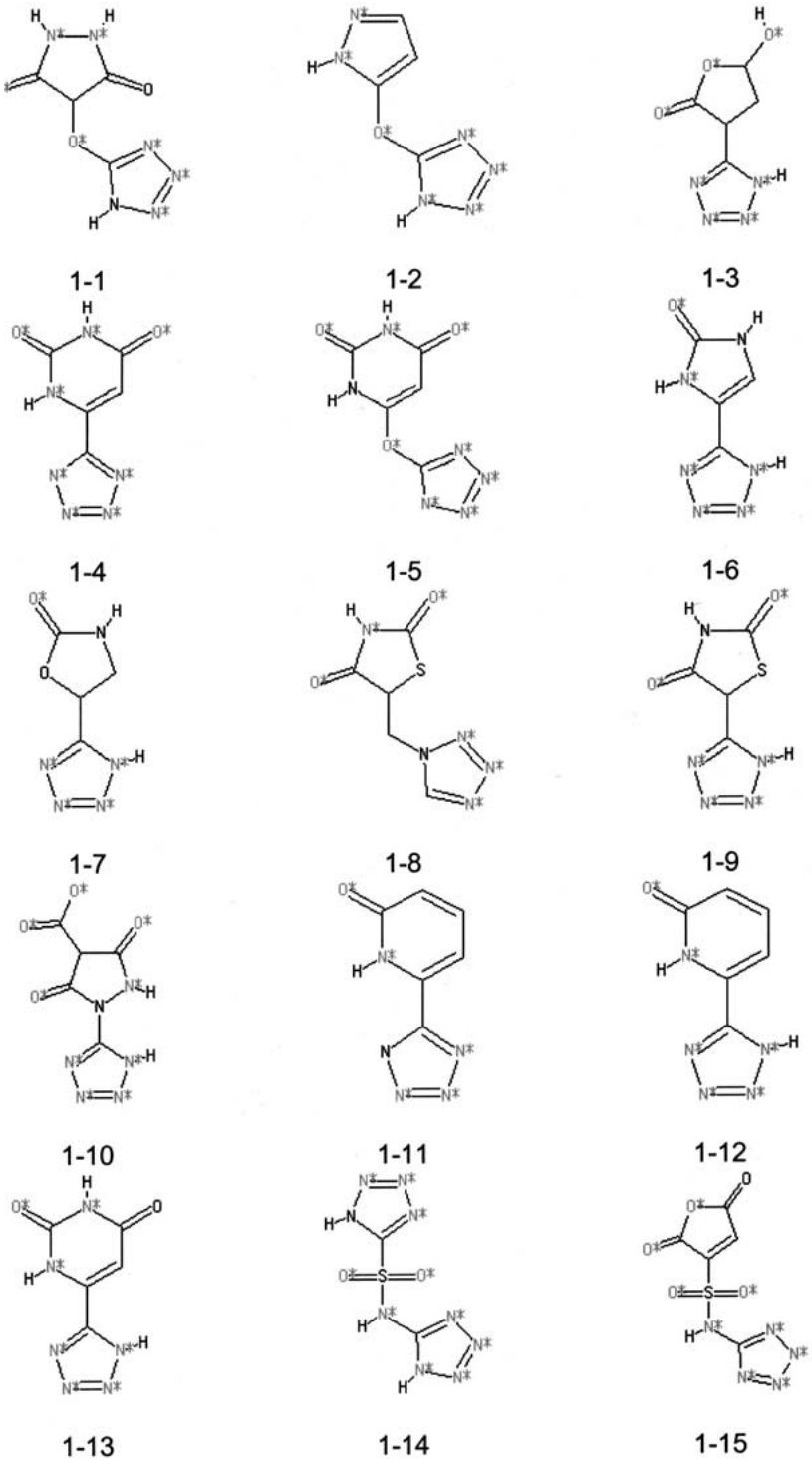

Figure 4. Chemical structures of all selected generated ligands from Run 1. Heteroatoms involved in putative hydrogen bonds are highlighted with a star (*).

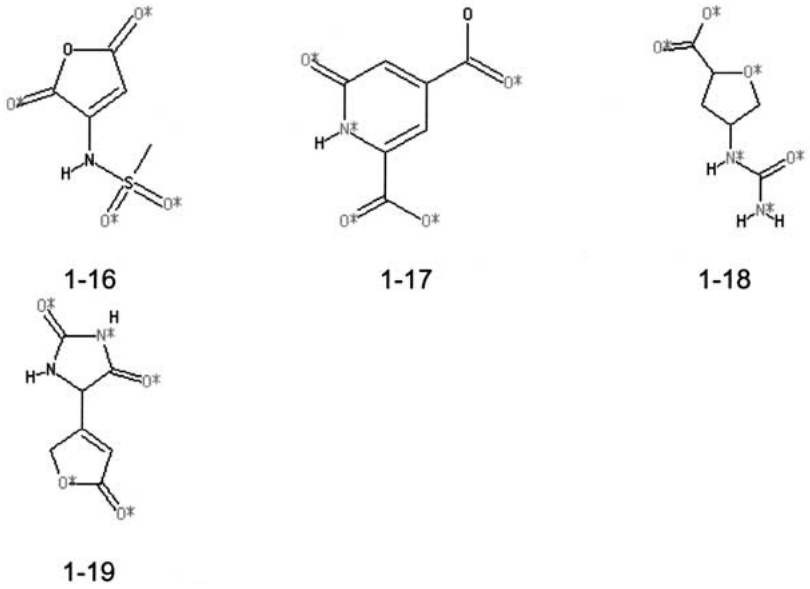

Figure 4. Continued.

HOH 692 were able to form hydrogens bonds with almost all the ligands. Water molecule HOH 613 was significantly less successful in forming a hydrogen bond with the ligands. Water molecule HOH 694 made a hydrogen bond with only one ligand (molecule 1-17). These results justified deleting the out-lying water molecules HOH 692 and HOH 694 for Run 2, as removing the latter would make little difference and removing the former would then leave water molecule HOH 689 as the only 'good' sitepoint.

RUN 2. SKELGEN's success rate for generating bound ligands was around 23%, which is only slightly worse than in the previous run. This means that removing water molecules HOH 692 and HOH 694 had indeed little effect on the ease with which ligands could be generated. This is most certainly due to the fact that the arginine triad and several other neighbouring sitepoints are close to these two water molecules. As a consequence, the removal of these water molecules has little effect on the overall accessibility of a ligand to the associated sitepoints. At the same time, the removal of these two water molecules modifies the shape of the binding site and can potentially remove some steric constraints for the position and orientation of ligands. The requirement that all ligands form at least one hydrogen bond with a water molecule means that the choice is now limited to water molecules HOH 613 and HOH 689.

The chemical structures of six selected unique ligands can be seen on Figure 6. Figure 7 shows a two-dimensional representation of the binding modes of all of these ligands after a full energy minimisation.

The main kind of compound that was now generated on several occasions was made up, as in Run 1, of a tetrazole ring, a sulfonamide linker and another tetrazole. On this occasion, however, there was a wider range of binding modes, due to the absence of steric clashes with water molecules HOH 692 and HOH 694. This explains why such compounds, while being all structurally related, were not generated in Run 1.

The first ligand to consider, molecule 2-1, shows a new binding mode produced by the strong interaction of the sulfono group with the arginine catalytic triad. The absence of water molecule HOH 692 allows the ligand to translate within the binding site without producing a steric clash. The second ligand to consider, molecule 2-2, is the same as molecule 1-14 generated in Run 1; however, the binding mode is different since now water molecule HOH 692 is absent and, consequently, the tetrazole ring is not pushed away. Other ligands to consider are molecules 2-3 and 2-4. Due to the absence of water molecule HOH 692, their tetrazole rings interacting with the arginine triad can adopt a slightly different position in the binding site, which now allows the sulfono group to form hydrogen bonds with Arg 56 and Tyr 342, while retaining the hydrogen bond with Asp 100.

Molecule 2-5 is identical to molecule 1-13 from Run 1, with a tetrazole ring attached to a uracil ring. Both molecules exhibit the same binding mode. Since

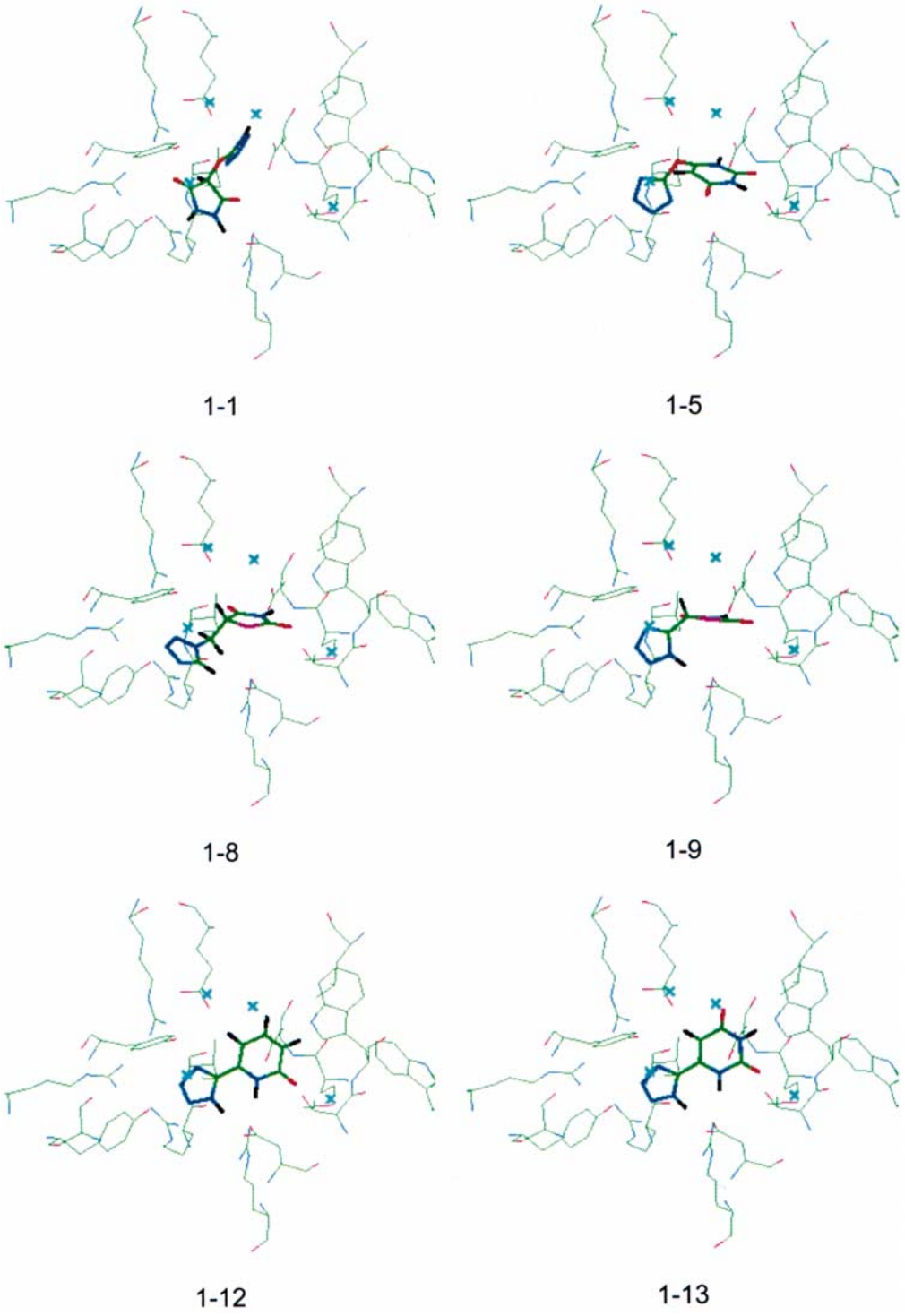

Figure 5. Representation of the binding modes of some of the ligands from Run 1 after a full energy minimisation. Water molecules are represented by crosses in cyan.

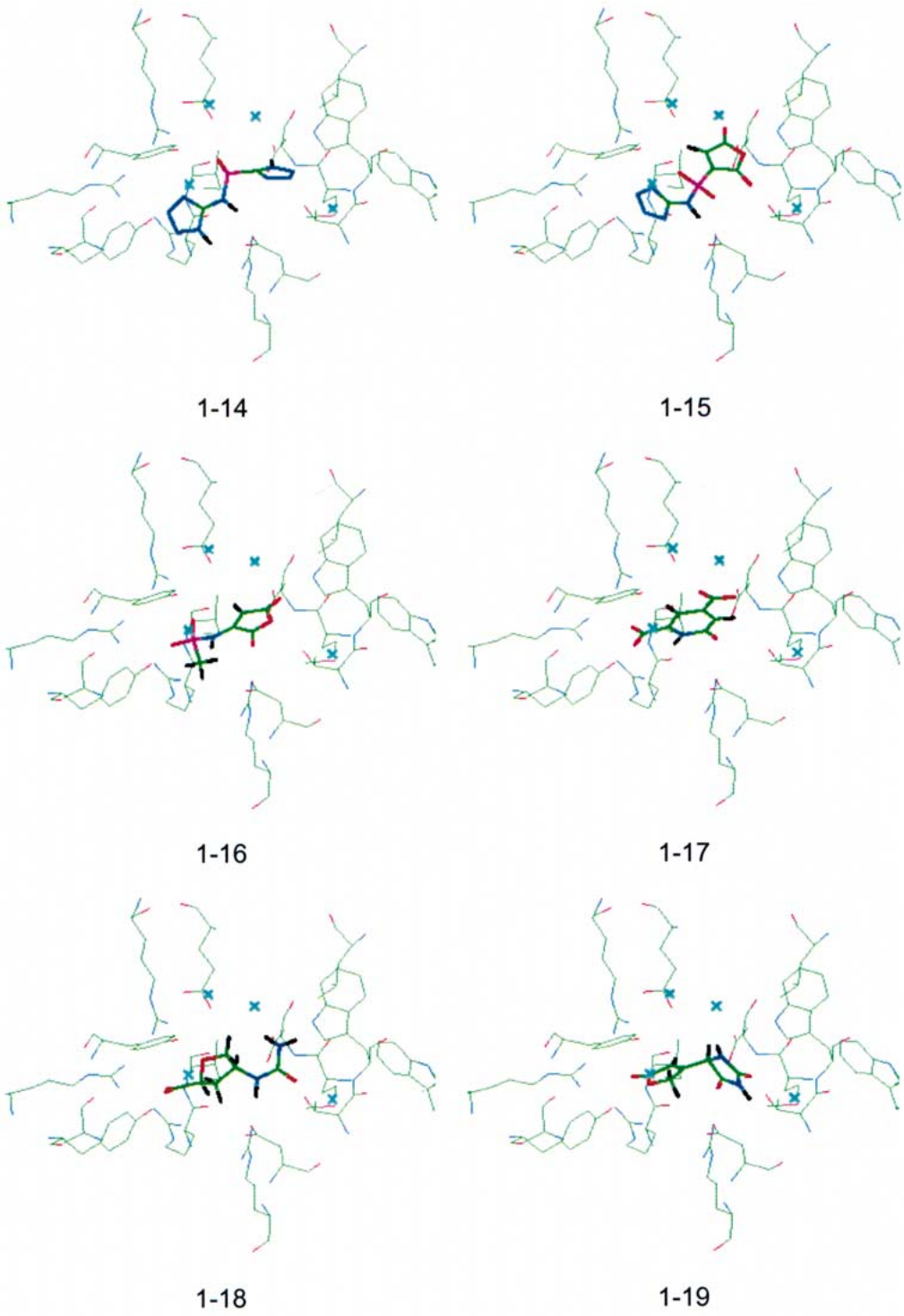

Figure 5. Continued.

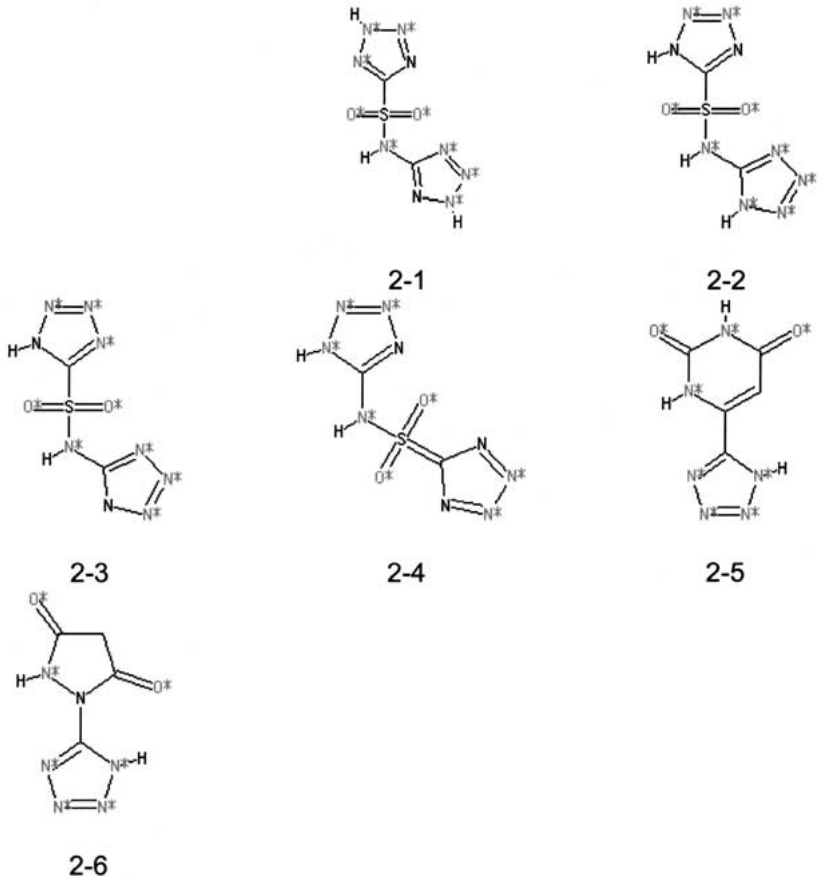

Figure 6. Chemical structures of all selected generated ligands from Run 2. Heteroatoms involved in putative hydrogen bonds are highlighted with a star (*).

the main ligand-water interaction occurs via a hydrogen bond with water molecule HOH 689, the removal of water molecule HOH 692 had no effect on the binding mode of the ligand.

Finally, molecule 2-6 has a very similar binding mode as that one previously observed for molecule 1-13 from Run 1, forming hydrogen bonds with the catalytic triad, with Asp 62 (two) and water molecule HOH 689. It is surprising that this molecule was not generated in Run1, as it would also be able to form strong hydrogen bonds not only with water molecule HOH 692 but also with water molecule HOH 694. This ligand may not have been generated in Run 1 because of insufficient sampling.

The above binding modes confirm that binding to the arginine triad is very important for binding, and seem to indicate that removing water molecules HOH 692 and HOH 694 allow for more flexibility in the way the preferred anchor group, tetrazole, can sit in the vicinity of the arginine triad. However, an expected increase in molecular diversity is not observed because the imposed constraints of forming a hydrogen bond with either water molecule HOH 689 or HOH 613 limit the nature of the ligands that can be generated. This naturally begs the question of whether having the two remaining water molecules is desirable at all. We comment further on this below. In any case, the generated ligands share some features with DANA and the phosphonate analogs: they can also form hydrogen bonds with Asp 62, Arg 56 and/or Asp 100. A picture of a consistent binding mode seems to be emerging, although it is not surprising since the anchor group is nearly always the same in the generated ligands.

Interestingly, most of the above mentioned ligands do not have any hydrogen bonds with any of the two water molecules present. Only molecules 2-5 and 2-6 form a hydrogen bond with water molecule HOH 689. This shows that the minimisation of the ligands in the binding site probably replaced weak hydrogen bonds

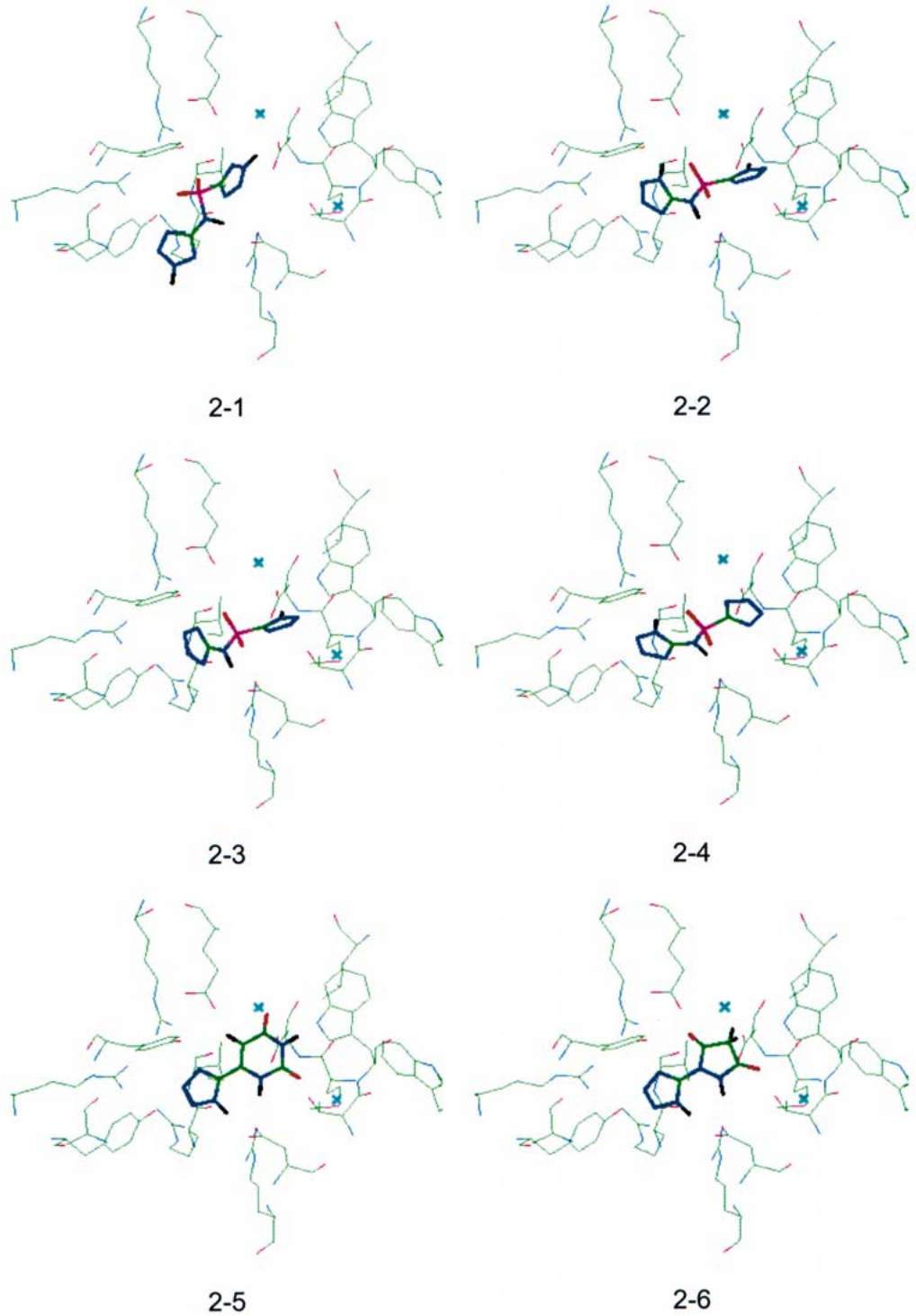

Figure 7. Representation of the binding modes of some of the ligands from Run 1 after a full energy minimisation. Water molecules HOH 613 and HOH 689 are represented by crosses in cyan.

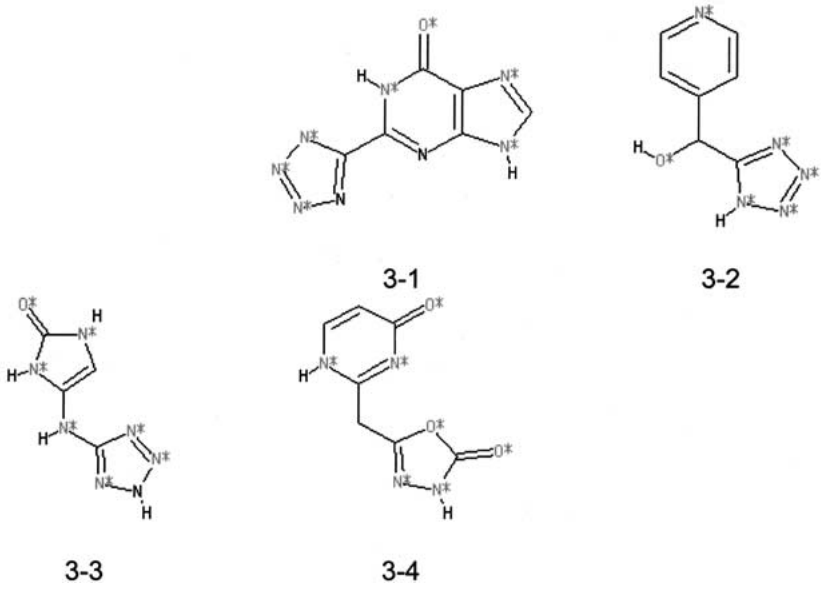

Figure 8. Chemical structures of all selected generated ligands from Run 3. Heteroatoms involved in putative hydrogen bonds are highlighted with a star (*).

with strong ones, resulting in the loss of ligand-water interactions. This was probably compounded by the fact that the presence of water molecules HOH 692 and HOH 694 in the previous Run 1 imposed steric constraints during the energy minimisation that prevented the ligands from moving away from their original binding modes. Although these results might seem contradictory, they provide a good example of what can happen when introducing water molecules in a drug design strategy. In this case, they certainly also point to the fact that the *theoretical* binding to this protein by a ligand is certainly possible without water molecules. This is by no means surprising, particularly because in our example the use of water molecules simply introduced additional constraints. There is no evidence that such water molecules are in fact a necessary condition for ligand binding, although removing them completely would lead to unsatisfied hydrogen-bonding sitepoints on the surface of the binding site, as described further on.

RUN 3. SKELGEN's success rate for generating bound ligands was around only 2%. The main obstacle for generating structures was achieving a sufficiently low value for the sitepoints term in the scoring function. This clearly shows that, with respect to Run 2 (which included also water molecule HOH 613), it was now much more difficult to generate a ligand which was forced to make a hydrogen bond with the remaining water molecule HOH 689.

The chemical structures of the selected four unique ligands can be seen on Figure 8, while Figure 9 shows a two-dimensional representation of the binding modes of these molecules after a full energy minimisation. Three molecules (3-1, 3-2 and 3-3) have once again a tetrazole ring as anchor group, which can be seen in each of its ionic forms (deprotonated and its two tautomers) making hydrogen bonds with the arginine triad. Molecule 3-1 also makes hydrogen bonds with Asp 100 and Thr 127. Molecule 3-2 also makes hydrogen bonds with Asp 62; however, this ligand has an unfavourable ring-to-ring conformation. Molecule 3-3 also makes hydrogen bonds with Asp 62 and Tyr 342. Finally, molecule 3-4 has an oxadiazolinone ring as anchor group, hydrogen bonded to the arginine triad; the rest of the ligand is hydrogen bonded to Asp 100 and Thr 127. All the molecules are able to form a hydrogen bond with water molecule HOH 689.

Figure 10 shows the four selected ligands in their minimised conformations in the binding site. If we look at the binding modes of these selected ligands, which were generated in the presence of water HOH 689 only (in cyan), we can see that the anchor groups (hydrogen-bonded to the arginine catalytic triad) show significant variability in the orientation of the ring plane, revealing the versatility of the arginine catalytic triad for making hydrogen bonds. It is thus reasonable to assume that the main geometric constraints imposed on the generated ligands

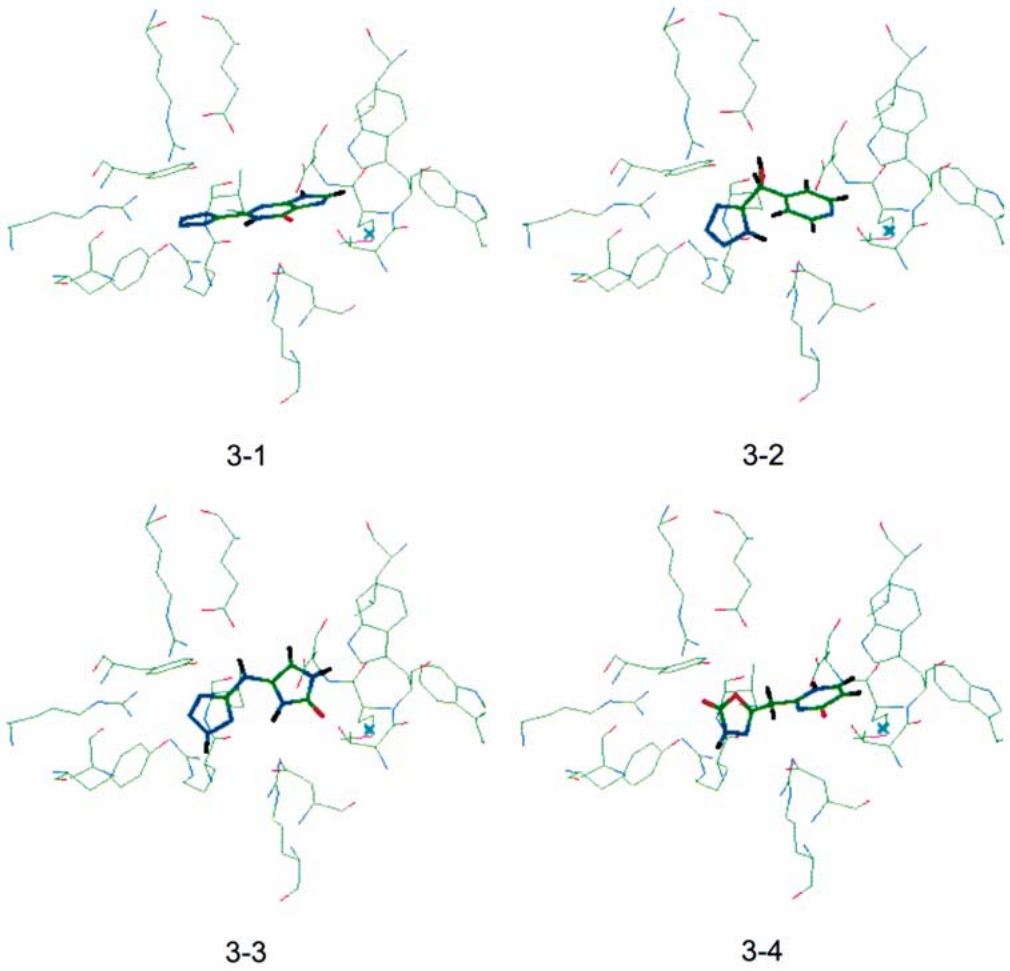

Figure 9. Representation of the binding modes of some of the ligands from Run 3 after a full energy minimisation. Water molecule HOH 689 is represented by a cross in cyan.

arise from having to satisfy a hydrogen bond to the distant water molecule HOH 689. The large variability in the orientation of the ring plane is also the result of the lack of steric constraints imposed by water molecules HOH 692, HOH 694 and HOH 613. The five-membered ring anchor groups can thus adopt more orientations with respect to the arginine triad. Other hydrogen bonds formed between the generated ligands and the binding site are similar to those observed with DANA, indicating that a similar binding mode is nearly always observed.

It should be stressed that these results are a mere example of what can happen when trying to satisfy hydrogen bonding to a bound water molecule that is not readily accessible. Although there might be a compelling reason for trying to design a ligand that will form a hydrogen bond with a strongly-bound water molecule, this might prove difficult on some occasions.

RUN4. SKELGEN was unable to generate any ligands at all. As in the previous run, the main obstacle for generating structures was achieving a sufficiently low value for the sitepoints term in the scoring function. This clearly shows that the sitepoints associated with bound water molecule HOH 689 (Region 4a, as described earlier) are not accessible in practice by an incoming ligand. Consequently, it would seem that the presence of this water molecule is ultimately required on the protein surface to ensure that a generated ligand can establish hydrogen-bonding interactions with that region of the binding site (which would otherwise have unsatisfied hydrogen bond groups). Furthermore, hydrogen-bonding interactions between a ligand and this water molecule would ensure that no hydration

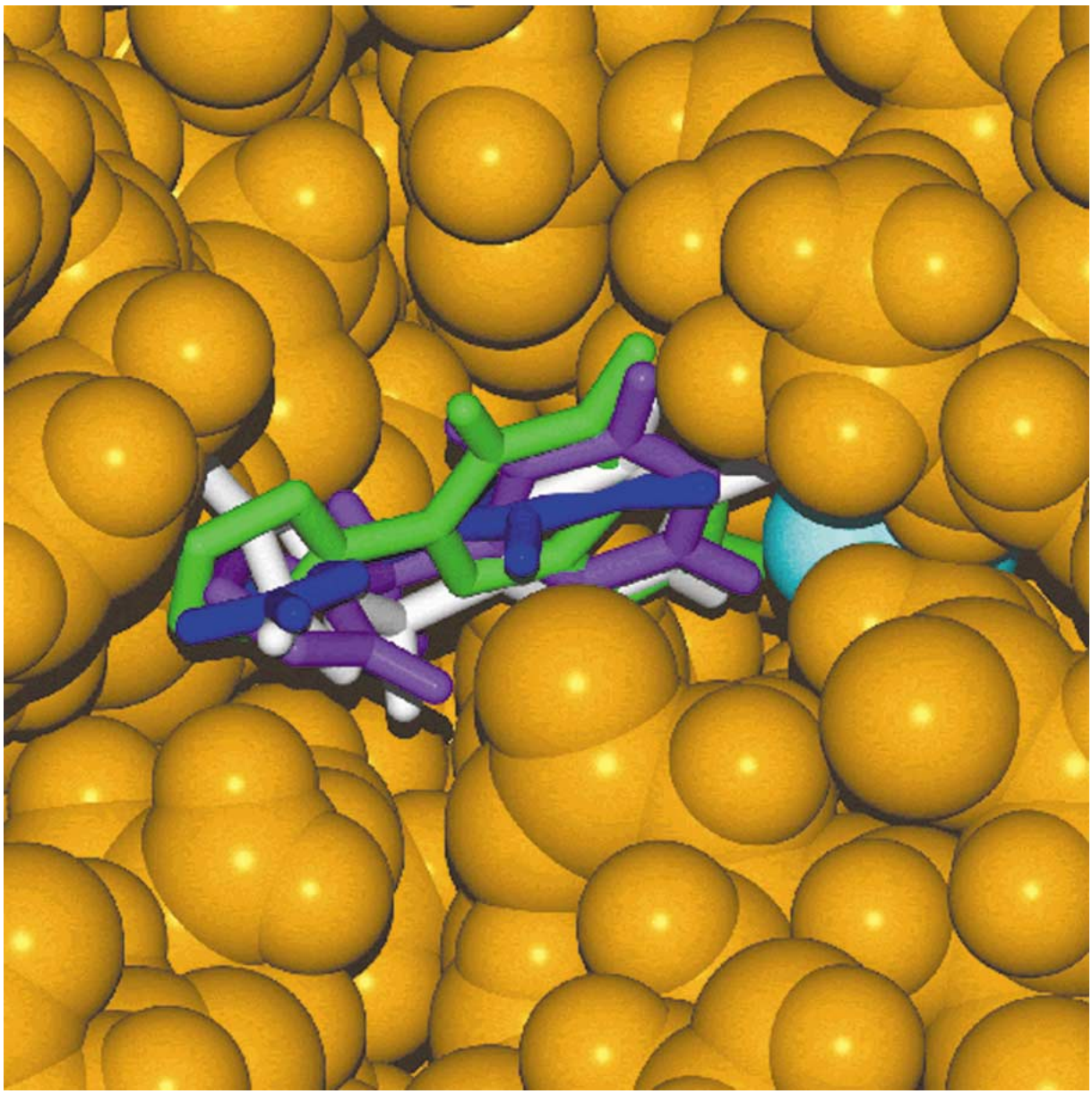

Figure 10. Molecular representation of the four ligands generated in Run 3, showing their conformational variability around the catalytic arginine triad. All the ligands are hydrogen-bonded to water molecule HOH 689 (highlighted in cyan).

penalties are incurred as a result of a lack of hydrogen bonding between HOH 689 and the bulk of the solvent. Our results serve to illustrate the fact that including sitepoints to replace an associated bound water molecule can be counter-productive in a drug design strategy if these sitepoints are not easily accessible.

Conclusions

We have examined the effect of including explicit crystallographically-determined water molecules in the binding site of neuraminidase on our ability to produce *de novo* designed ligands. As far as the generated ligands are concerned, a preferential anchor group was observed (tetrazole), which allowed for a large flexibility in the hydrogen bonding of this group

to the arginine triad and the adoption of alternative but related binding modes. In general it was observed that the general features of the binding mode of the co-crystalised carboxylic- and phosphonic-acid based ligands were retained by the generated ligands.

For this particular protein, our results suggest that two water molecules (HOH 613 and HOH 689) become crucial for generating new ligands. Their removal led to serious difficulties when applying our computer-aided *de novo* ligand design methods. This seems to occur because the associated protein sitepoints that they replace are not readily accessible to an incoming ligand. Our results for neuraminidase seem to suggest that, if water molecules are to be included, the gradual addition of more of the crystallographically determined water molecules (HOH 692 and HOH 694) in the binding site, allows for ligands not only to be generated in larger numbers, but also with wider chemical diversity. This is an indirect effect arising from one constraint imposed on all ligands: that a hydrogen bond be made with at least one water molecule. Consequently, when water molecules HOH 692 and HOH 694 were present, one of them could form a hydrogen bond with the ligand, allowing for more chemical structures to be accepted. Since hydration penalties can only be avoided by generating ligands which form hydrogen bonds with buried water molecules (HOH 613 and HOH 689), imposing hydrogen-bonding interactions with such water molecules in the absence of other water molecules reduced drastically the success rate of ligand generation.

The example that we have studied also shows that, in some cases, bound water molecules can be more accessible for hydrogen bonding to an incoming ligand than the actual protein binding sitepoints associated with them. This is enhanced by the fact that water molecules can act as versatile amphiprotic hydrogen-bonding groups. At the same time, we have observed that water molecules (or their removal) can modify the conformational constraints of a particular binding site. This allows for different ligands to bind or new conformations of the same ligands to become accessible within a binding site.

Bound water molecules, although intrinsically displaceable, should only be retained in a given structure-based or *de novo* drug design algorithm if they facilitate ligand generation. The example that we have presented here shows that in any protein-binding site it is important to determine the influence that water molecules present in the crystal structure can have on the generation of novel ligands. Once the number and position of water molecules that are bound to the binding site have been determined, and it has been decided to include them, any strategy for ligand generation should explore the effect of gradually removing some or all of them on the sucess rate of production of new molecular scaffolds.

Strongly-bound water molecules tend to satisfy hydrogen bonds with the protein surface and, in the absence of such water molecules, have to be satisfied by a bound ligand. Retaining such water molecules keeps those important hydrogen bonds to the protein, but in general will mean that further hydrogen bonds have now to be satisfied with these bound water molecules. This can be accomplished by other water molecules (either other bound water molecules or from the bulk solvent) or by a ligand. However, as was illustrated in this work, satisfying hydrogen bonds with bound water molecules can prove to be tricky if they are not immediately accessible.

All of these observations reveal the importance and pitfalls of the use of crystallographically-determined water molecules in any drug design strategy. Either through the provision of hydrogen-bonding sitepoints or by the direct modification of the accessible conformational space within a binding site, water molecules can play a vital role in the search for novel molecular scaffolds that can bind effectively to a protein receptor.

Acknowledgements

RLM is also a Research Fellow of Hughes Hall, Cambridge. Alfonso Garcia-Sosa is gratefully acknowledged for providing the list of selected water molecules on the 2SIL/2SIM protein pair.

References

1. Goodford, P.J., J. Med. Chem., 28 (1985) 849.
2. Kuntz, I.D., Science, 257 (1992) 1078.
3. Kuntz, I.D., Meng, E.C. and Shoichet, B.K., Acc. Chem. Res., 27 (1994) 117.
4. Lewis, R.M. and Leach, A.R., J. Comp-Aided Mol. Des., 8 (1994) 467.
5. Bohm, H.J., J. Comp-Aided Mol. Des., 6 (1992) 593.
6. Gillet, V.J., Johnson, A.P., Mata, P., Sik, S. and Williams, P., J. Comp-Aided Mol. Des., 7 (1993) 127.
7. Miranker, A. and Karplus, M., Proteins: Structure, Function and Genetics, 11 (1991) 29.
8. Moon, J.B. and Howe, W.J., Proteins: Structure, Function and Genetics, 11 (1991) 314.
9. Glen, R.C. and Payne, A.W.R., J. Comp-Aided Mol. Des., 9 (1995) 181.

10. Pearlman, D.A. and Murcko, M.A., J. Med. Chem. 39 (1996) 1651.
11. Liu, H., Duan, Z., Luo, Q. and Shi, Y., Proteins: Structure, Function and Genetics, 36 (1999) 462.
12. Mikol, V., Papageorgiou, C. and Borer, X., J. Med. Chem., 38 (1995) 3361.
13. Dunitz, J.D., Science, 264 (1994) 670.
14. Dunitz, J.D., Chemistry & Biology, 2 (1995) 709.
15. Ladbury, J.E., Chemistry & Biology, 3 (1996) 973.
16. Nakasako, M., J. Mol. Biol., 289 (1999) 547.
17. Faerman, C.H. and Karplus, P.A., Proteins: Structure, Function and Genetics, 23 (1995) 1.
18. Schwabe, J.W.R., Curr. Op. Struct. Biol., 7 (1997) 126.
19. Carrell, H.L., Glusker, J.P., Burger, V., Manfre, F., Tritsch, D., Biellmann, J-F., Proc. Natl. Acad. Sci. USA, 86 (1989) 4440.
20. Chung, E., Henriques, D., Renzoni, D., Zvelebil, M., Bradshaw, J.M., Waksman, G., Robinson, C.V. and Ladbury, J.E., Structure, 6 (1998) 1141.
21. Baker, E.L. and Hubbard, R.E., Prog. Biophys. Molec. Biol., 44 (1984) 97.
22. Loris, R., Langhorst, U., De Vos, S., Decanniere, K., Bouckaert, J., Maes, D., Transhue, T.R. and Stayaert, J., Proteins: Structure, Function and Genetics, 36 (1999) 117.
23. Loris, R., Stas, P.P. and Wyns, L., J. Biol. Chem., 269 (1994) 26722.
24. Ehrlich, L., Reckzo, M., Wade, R.C., Protein Eng., 11 (1998) 2054.
25. Raymer, M.L., Sanschagrin, P.C., Punch, W.F., Venkataram, S., Goodman, E.D. and Kuhn, L., J. Mol. Biol., 265 (1997) 445.
26. Sanschagrin, P.C. and Kuhn, L.A., Protein Science, 7 (1998) 445.
27. Carugo, O., Protein Eng., 12 (1999) 1021.
28. Carugo, O. and Bordo, D., Acta Crystallographica D, D55 (1999) 479.
29. Carugo, O. and Argos, P., Proteins: Structure, Function and Genetics, 31 (1998) 201.
30. 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 (1996) 1353.
31. Rejto, P.A. and Verkhivker, G.M., Proteins: Structure, Function and Genetics, 28 (1997) 313.
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. Palomer, A., Perez, J.J., Navea, S., Llorens, O., Pascual, J., Garcia, L. and Mauleon, D., J. Med. Chem., 43 (2000) 2280.
34. Poornima, C.S. and Dean, P.M., J. Comp-Aided Mol. Des., 9 (1995) 500.
35. Smith, P.E. and Pettitt, B.M., J. Phys. Chem., 98 (1994) 9700.
36. Honig, B. and Nicholls, A., Science, 268 (1995) 1144.
37. Gilson, M.K. and Honig, B., Proteins: Structure, Function and Genetics, 4 (1988) 7.
38. Straatsma, T.P., in Lipkwowitz, K.B. and Boyd, D.B. (Eds.), Reviews in Computational Chemistry, VCH Publishers Inc., U.S.A., 9 (1996) 81.
39. Lamb, M.L. and Jorgensen, W.L., Curr. Opinion Chem. Biol., 1 (1997) 449.
40. Aqvist, J., Medina, C., Samuelsson, J-E, Protein. Eng., 7 (1994) 385.
41. Hansson, T., Marelius, J., Aqvist, J., J. Comp-Aided Mol. Des., 12 (1998) 27.
42. Rarey, M., Kramer, B. and Lengauer, T., Proteins: Structure, Function and Genetics, 34 (1999) 17.
43. Pastor, M., Cruciani, G. and Watson, K.A., J. Med. Chem., 40 (1997) 4089.
44. Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N. and Bourne, P.E., Nucleic Acids Res., 28 (2000) 235.
45. Crennell, S.J., Garman, E.F., Laver, W.G., Vimr, E.R. and Taylor, G.L., Proc. Natl. Acad. Sci. USA, 90 (1993) 9852.
46. Crennell, S.J., Garman, E.F., Philippon, C., Vasella, A., Laver, W.G., Vimr, E.R. and Taylor, G.L., J. Mol. Biol., 259 (1996) 264.
47. Mills, J.E.J., Perkins, T.D.J. and Dean, P.M., J. Comp-Aided Mol. Des., 11 (1997) 229.
48. Mills, J.E.J. and Dean, P.M., Proceedings of the 12th. European Symposium on Quantitative Structure-Activity Relationships: Molecular Modelling and Prediction of Bioactivity, Plenum, New York, 1998.
49. Todorov, N.P. and Dean, P.M., J. Comp-Aided Mol. Des., 11 (1997) 175.
50. Todorov, N.P. and Dean, P.M., J. Comp-Aided Mol. Des., 12 (1998) 335.
51. Todorov, N.P. and Dean, P.M., in Frauenfelder, H., Hummer, G. and Garcia, R. (Eds.), Biological Physics, American Institute of Physics, U.S.A., 1999.
52. Sobolev, V., Wade, R.C., Vriend, G. and Edelman, M., Proteins: Structure, Function and Genetics, 25 (1996) 120.
53. Dinur, U. and Hagler, A.T., in Lipkwowitz, K.B. and Boyd, D.B. (Eds.), Reviews in Computational Chemistry, VCH Publishers Inc., U.S.A., 2 (1991).
54. Greengard, L. and Rokhlin, V.I., J. Comp. Phys., 73 (1987) 325.
55. Ding, H.Q., Karasawa, N. and Goddard, W.A., J. Chem. Phys., 97 (1992) 4309.
56. Leach, A.R., Molecular Modelling: Principles and Applications, Longman, U.S.A., 1996.
57. Angus, D.I. and von Itzstein, M., Carbohydrate Res., 274 (1995) 279.
58. Walliman, K. and Vasella, A., Helvet. Chim. Acta, 73 (1990) 1359.