

MOLECULAR MODELLING EXERCISES

EXERCISE 20.2

SAQUINAVIR

INTRODUCTION

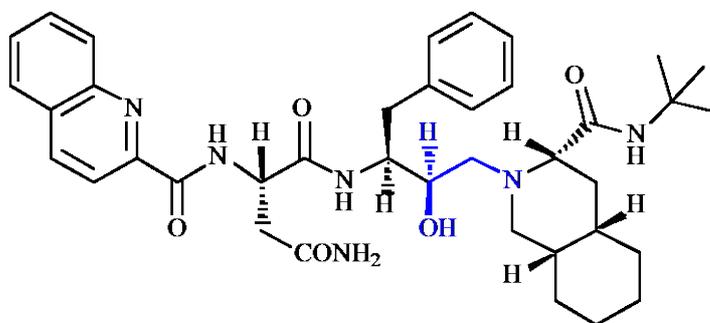


Figure 1 Saquinavir

Saquinavir was the first HIV protease inhibitor (PI) to reach the market and was the benchmark structure for all subsequent PIs. The lead compound used in the development of saquinavir was a peptide structure, and various peptidomimetic strategies were employed to diminish peptide character and increase oral bioavailability (section 20.7.4.3 of the textbook). Despite this, saquinavir is a relatively large and flexible molecule with a rather low oral bioavailability. Because of the number of rotatable bonds, there are a large number of possible conformations and it is impossible to predict what the active conformation might be. The only way to identify the active conformation is to crystallise the target protein with saquinavir bound to the active site.

In this exercise, you will identify the active conformation by extracting saquinavir from its binding site. You will also identify binding interactions between saquinavir and various amino acid residues.

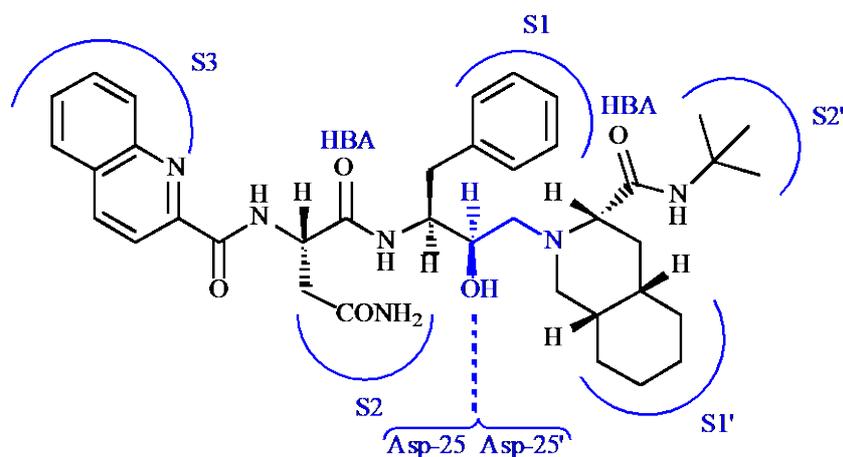


Figure 2 Interaction of saquinavir with Asp-25, Asp-25', and five binding subsites.

INSTRUCTIONS

It is suggested that you attempt to carry out the following instructions yourself before following the more detailed **Procedures** that follow. You may find the file entitled **Common Operations for ChemBio3D/Chem3D** a useful guide on how to carry out various operations. ChemDraw files for structures can be found in the ChemDraw folder.

PART A

*Create an energy-minimised structure of saquinavir and identify the number of rotatable bonds present.

PART B

*Download the crystal structure of the protease enzyme with saquinavir bound to the active site (pdb 1hxb).

*Identify the various components of the crystal structure, as well as the active site.

*Locate the position of the catalytic amino acid residues (Asp-25 & Asp-25') and describe their position relative to saquinavir.

*Extract the structure of saquinavir and compare its conformation with the energy-minimised conformation obtained in Part A.

PART C

*Create a model binding site that will allow you to identify the amino acid residues surrounding saquinavir.

PART D

*Create a model binding site that will allow you to identify hydrogen bond interactions between saquinavir and surrounding amino acid residues.

PART E

Identify hydrogen bonding interactions between saquinavir and the following amino acid residues (Asp-25, Asp-25', Asp-30, Asp-29, Gly-48, Gly-27, Ile-50, & Ile-50').

PROCEDURES

There are various approaches that you can use to tackle these molecular modelling exercises. The following procedures illustrate how you might tackle this particular exercise, but they are not meant to be prescriptive. Note also that the results obtained may vary depending on the computer and the version of ChemBio3D or Chem3D used. For example, the specific conformations obtained from energy minimisation may differ, as may quantitative results such as steric energies.

ChemDraw files for structures can be found in the ChemDraw folder.

PART A

1. Open the ChemDraw file of saquinavir into ChemBio3D or Chem3D and energy minimise the structure.

*Open **ChemBio3D** or **Chem3D**.

*From the **File** menu choose **Open**, then find the ChemDraw file for saquinavir. *Select* it, then *click* on **Open**.

*Energy minimise the structure .

The steric energy is given in the bottom window and was -0.71 kcal/mol for the conformation obtained.

*From the **View** menu, choose **Model Display**, then *click* on **Show Lone Pairs**. Choose **Hide**. The structure is shown in figure 3A.

*From the **View** menu, choose **Model Display**, then *click* on **Show Hydrogen Atoms**. Choose **Hide**. The skeleton of the structure is shown more clearly without the hydrogen atoms (Fig. 3B).

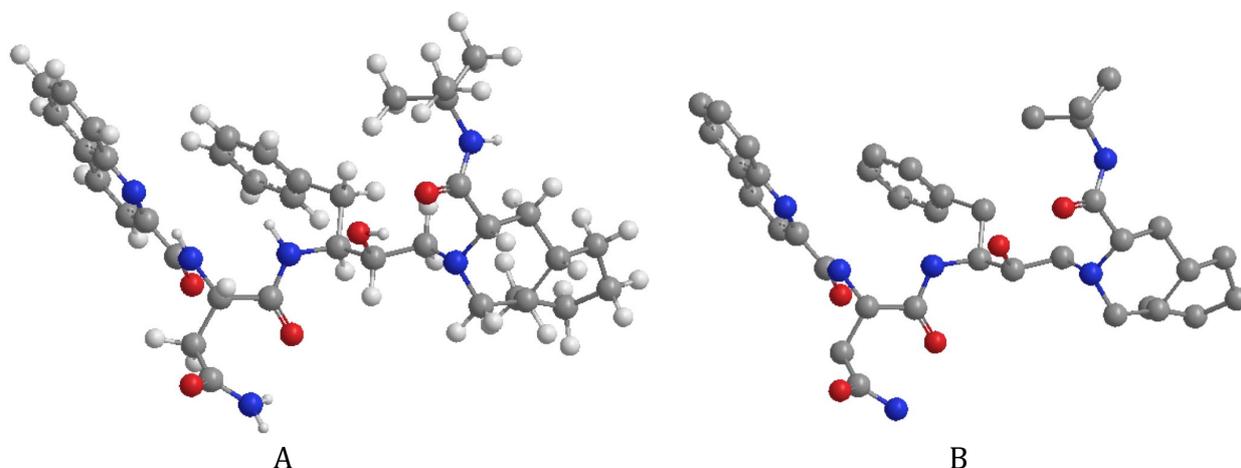


Figure 3 Energy-minimised structure of saquinavir with A) hydrogen atoms present and B) with hydrogen atoms hidden.

2. Identify the number of rotatable bonds in the structure.

*From the **Calculations** menu choose **Compute Properties**.

**Expand Molecular Topology* and *select Num Rotatable Bonds*. *Click OK*.

The number of rotatable bonds is revealed in the bottom window as 16. However, 3 amide bonds are included in this total. These are not truly rotatable bonds and a more accurate figure would be 13. These are the rotatable bonds that would be responsible for significant changes in conformation (Fig. 4).

With so many rotatable bonds, it would be a huge task to identify the most stable conformation. Even if we did identify the most stable conformation, we cannot assume that it

is the active conformation. Therefore, it makes more sense to identify the active conformation by studying the crystal structure of the HIV-protease enzyme with saquinavir bound to the active site.

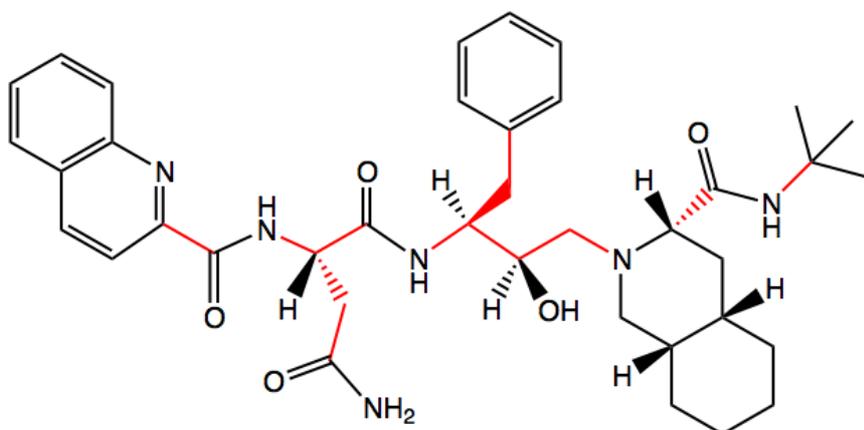


Figure 4 Rotatable bonds in saquinavir.

PART B

1. Download the crystal structure of the HIV-protease enzyme with bound saquinavir.

The crystal structure of HIV-protease with bound saquinavir can be retrieved from the protein data bank. You will need to be connected to the internet in order to do this.

* From the **Online** menu, choose **Find Structure from PDB ID**. Wait for the correct window to show up, then *enter* the PDB code for the protease enzyme with bound saquinavir (**1HXB**).

* If you wish to save the file, *click* on **Save as** and give the file a name and a location, then *click* **Save**. However, this step is not essential.

Click* on **Get File and the crystal structure will appear in the main window (Fig. 5).

The protein is represented as a ribbon structure with light purple sections corresponding to alpha-helices, and dark blue sections corresponding to beta sheets. The pink regions are connecting regions (or coils) that do not have a secondary structure. The saquinavir ligand is shown as a ball and stick model. Note that the protein is relatively small, which means that saquinavir is clearly visible in the central active site.

Note: An oddity with this crystal structure is the fact that two overlapping molecules of saquinavir are shown. This does not mean that two molecules are actually bound to the active site, because the molecules will clash. One possible explanation for this result is that the crystals are made up of protein molecules with saquinavir bound in two different orientations.

2. View the crystal structure from different aspects.

The protease enzyme is a symmetrical structure that is made up of two protein subunits with a hole through the middle that acts as the active site. The active site is occupied by saquinavir shown as a ball and stick model, and there are protein 'flaps' at the top that help to seal the ligand within the active site.

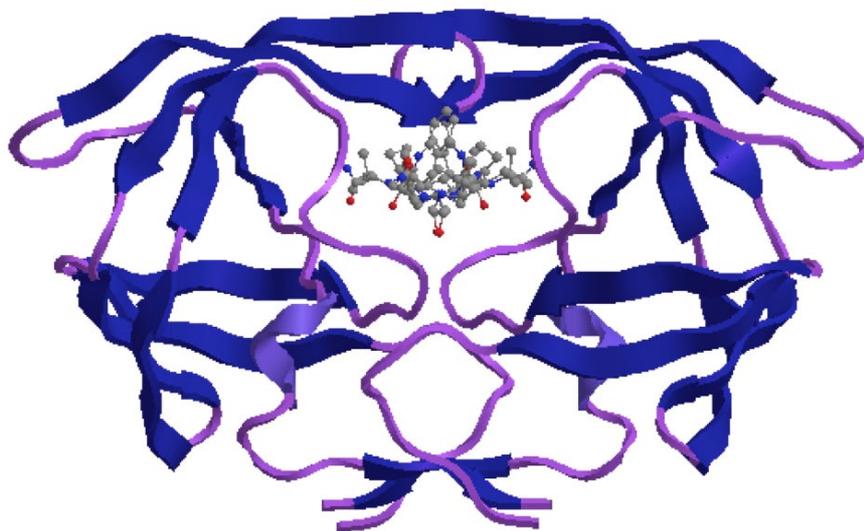


Figure 5 Crystal structure of saquinavir bound to the HIV-protease enzyme.

*Choose the **rotate** tool and *rotate* the structure to view it from different angles (Fig. 6).

In the viewpoints shown, the structure has been rotated 90° in the y-axis and 90° in the x-axis, showing that the inhibitor essentially traverses the width of the protein.

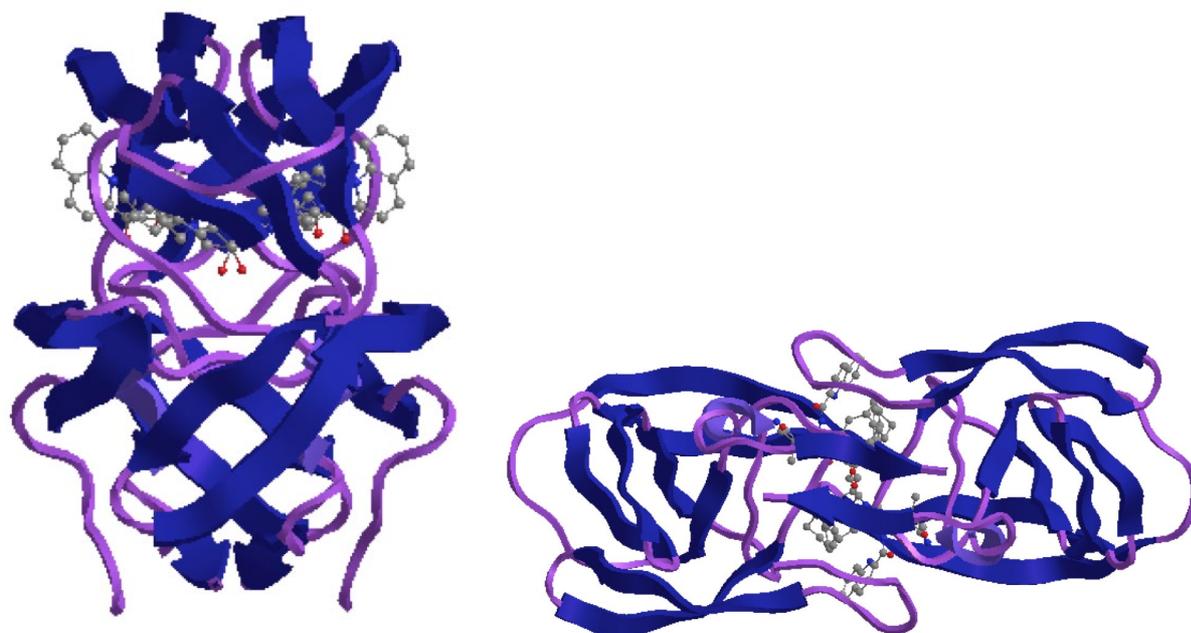


Figure 6 The protease enzyme from different viewpoints.

*Return to the original viewpoint, then *rotate* the structure 180° in the y-axis. You will find that the protein structure looks identical, indicating the symmetry of the structure (Fig. 7).

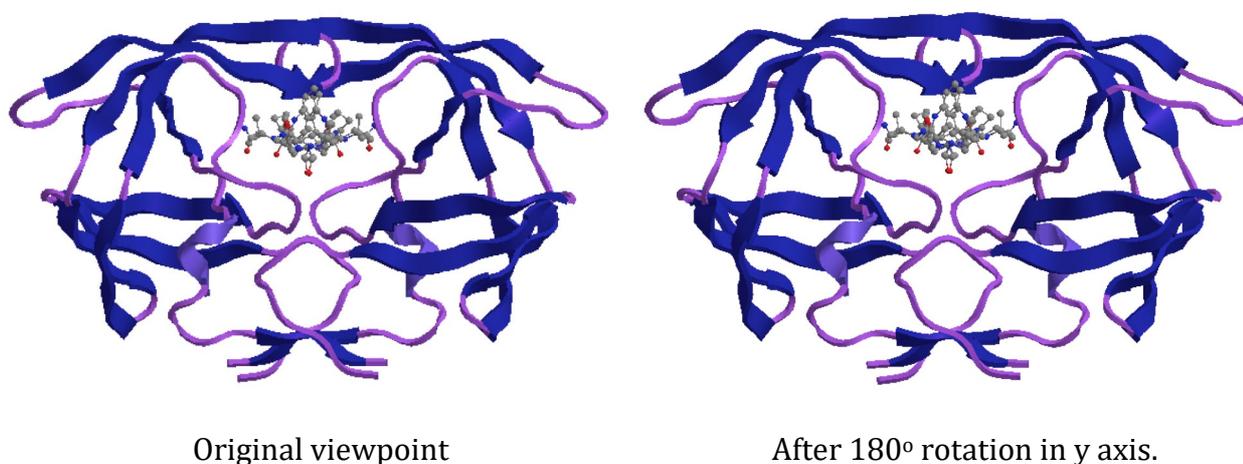


Figure 7 Rotation of the crystal structure by 180° round the y-axis.

3. Identify the main components of the crystal structure.

The Model Explorer table is useful in identifying the different components of the structure. There should be a tab for this to the left of the main window. *Clicking* on the tab will open up the Model Explorer table (Fig. 8). Failing that, you can access the table from the **View** menu, by *choosing* **Model Explorer**. There are four entries present - Chain B, Chain A, Solvent and Backbone.



A

Figure 8 The Model Explorer table.

**Click on Chain B.*

The protein subunit corresponding to chain B becomes highlighted in yellow in the main window (Fig. 9A).

**Click on Chain A.*

The protein subunit corresponding to chain A is now highlighted. Note that saquinavir also becomes highlighted.

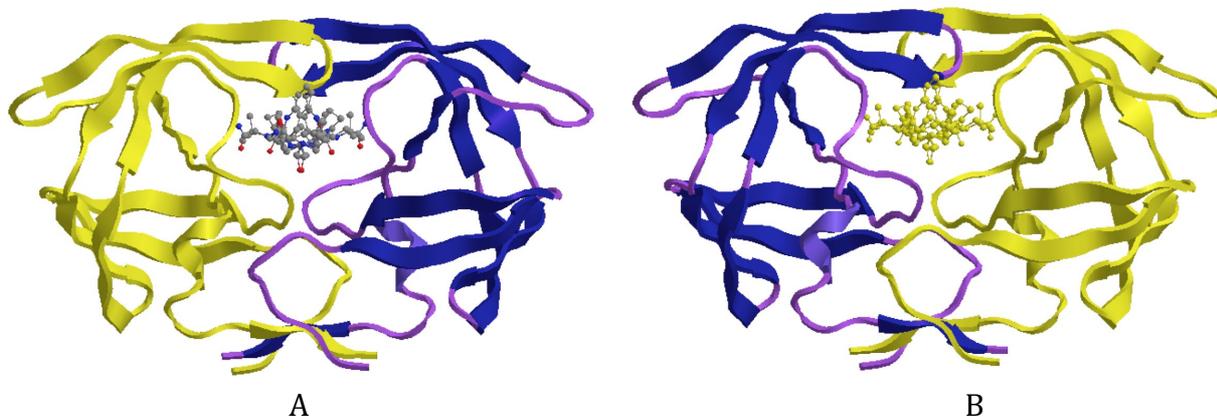


Figure 9 A) Chain B is highlighted

B) Chain A is highlighted along with saquinavir.

**Click on the + sign to the left of Chain B.*

This expands the entry for chain B to show all the constituent amino acids (Fig.10).

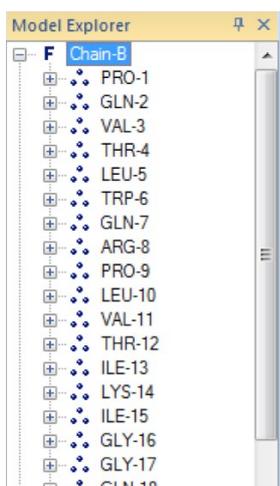


Figure 10 Expansion of Chain B in the Model Explorer table to reveal constituent amino acids.

**Click on Asp-25*

Scroll down the list of amino acids and *click* on Asp-25. This is one of the catalytic aspartate residues within the active site. Its position is now highlighted in yellow in the main window (Fig. 11).

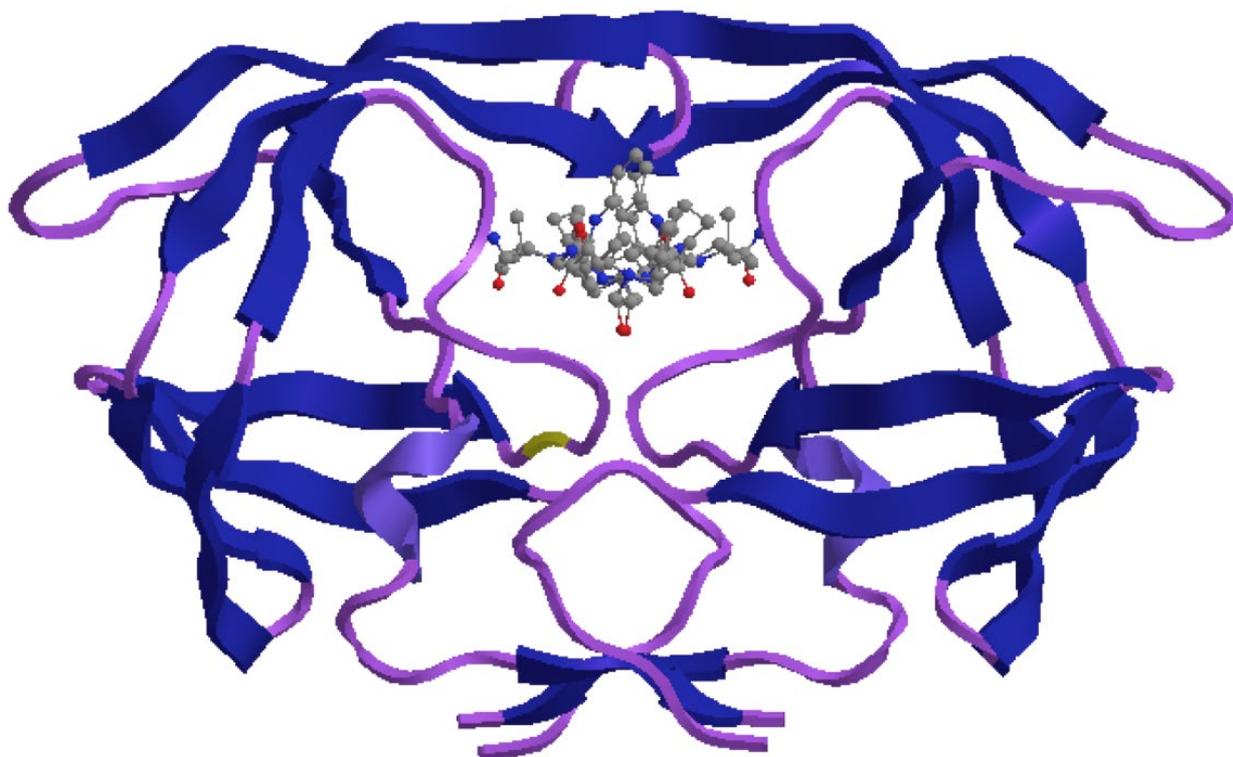


Figure 11 Position of Asp-25 in the HIV-protease enzyme.

*Click on the minus sign to the left of Chain B in the **Model Explorer** table to shrink the list of amino acids.

*Click on the + sign to the left of Chain A.

This expands the entry for Chain A to show two main entries entitled **Fragment 1-99** and **Ligand -2**.

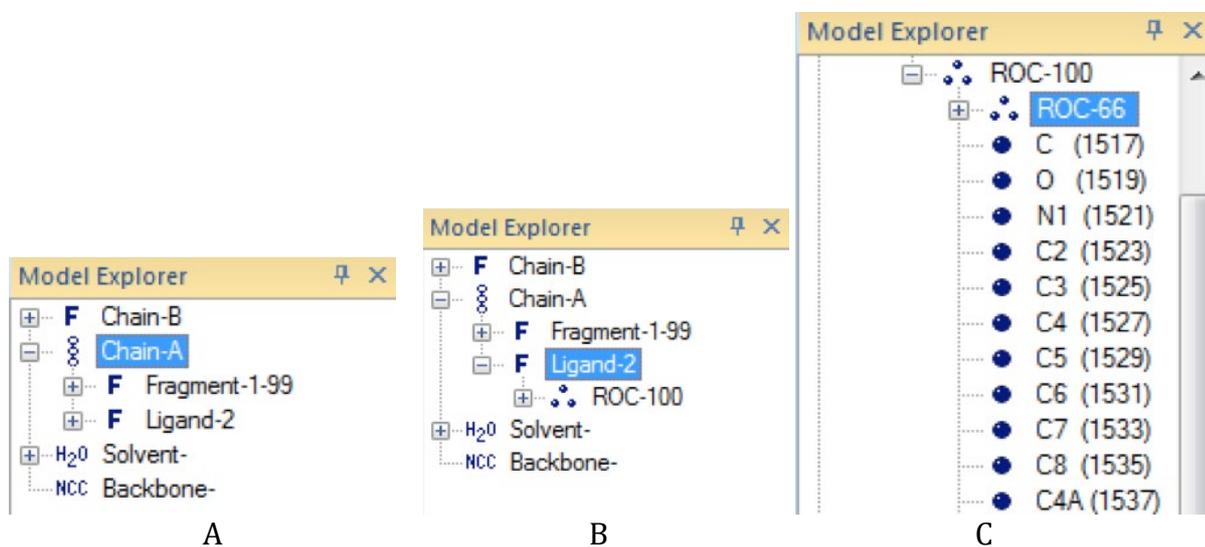


Figure 12 A) The Model Explorer table with Chain A expanded. B) Expansion of Ligand-2 C) Expansion of ROC-100.

* Click on **Fragment 1-99** to highlight the protein chain. Saquinavir is not highlighted as it is not part of the protein chain (Fig. 13A).

*Click on **Ligand-2** to highlight saquinavir (Fig. 13B).

*Expand the entry for **Ligand-2** by clicking on the + sign to the left of the entry to reveal **ROC-100** (Fig. 12B).

*Expand the entry for ROC-100 to reveal a series of amino acids, as well another ligand identified as **ROC-66** (Fig. 12C).

*Click on the entry for ROC-66

This appears to highlight half of the ligand (Fig. 14). In fact, there are two molecules of saquinavir present in the crystal structure (ROC-100 and ROC-66). This is a quirk of the experimental measurement. There cannot be two molecules of saquinavir present within the binding site at any one time as they would be occupying the same region of space. The two molecules represent saquinavir binding to the symmetrical binding site in both possible orientations. In order to study the active conformation of saquinavir and its binding interactions, we will remove ROC-66.

*Make sure that ROC-66 is still selected. From the **Edit** menu choose **Clear**.

Only one molecule of saquinavir is now shown in the active site (Fig. 15).

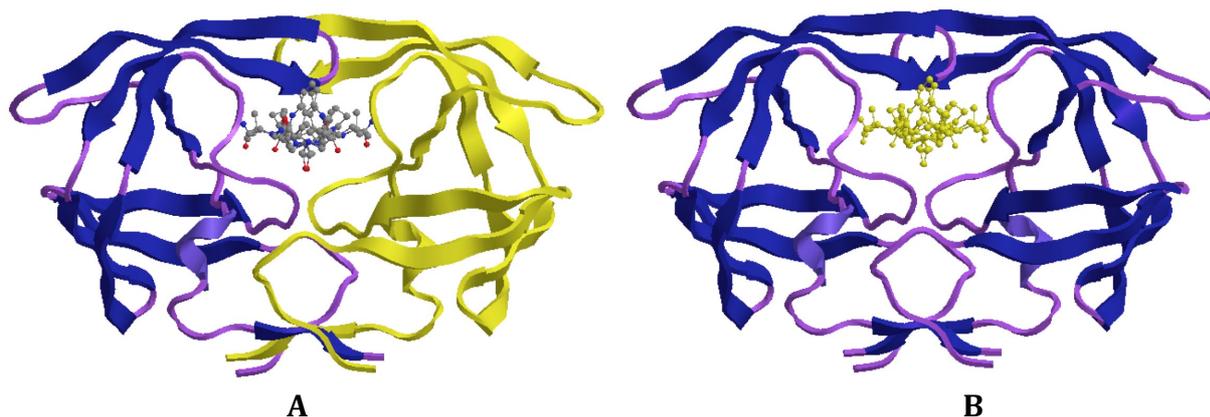


Figure 13 A) Fragment 1-99 highlighted. B) Ligand-2 (saquinavir) highlighted.

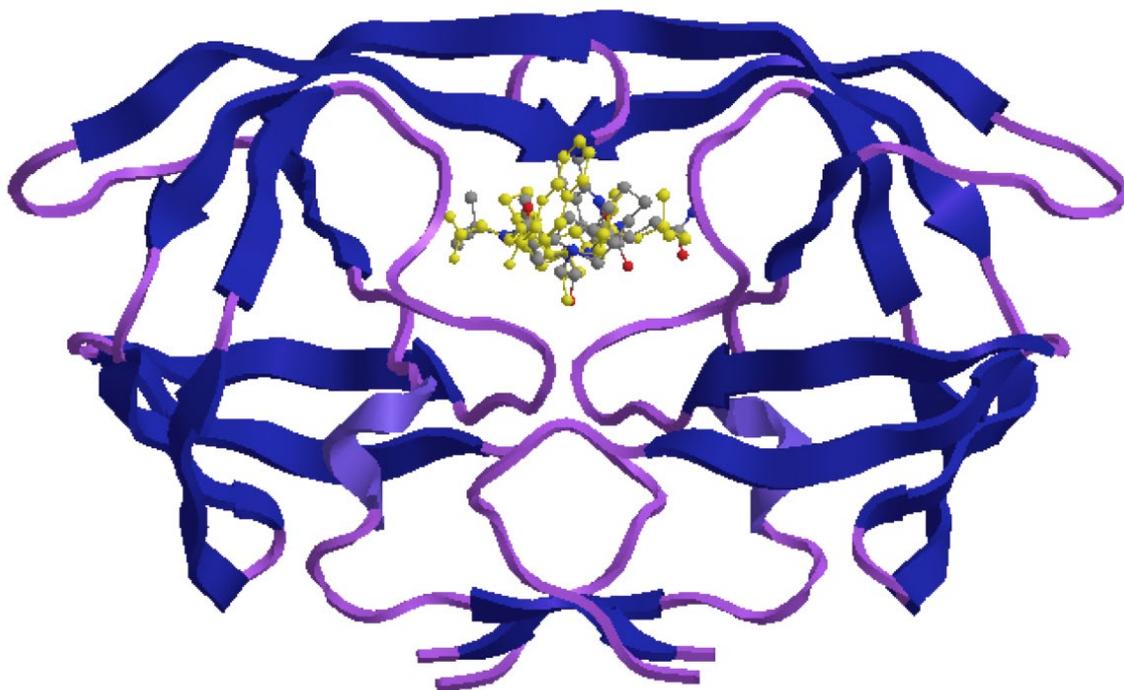


Figure 14 Crystal structure with ROC-66 highlighted.

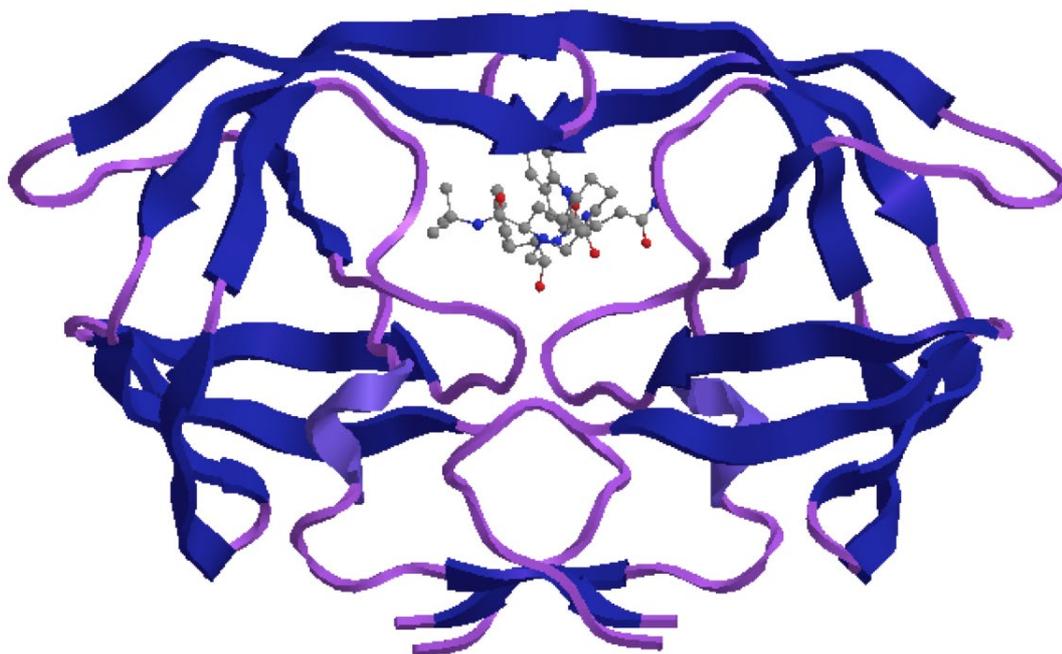


Figure 15 Crystal structure with ROC-66 removed.

4. Identify amino acid residues in the active site

*Use the **zoom** tool or the scroll control on your mouse to *zoom* into the binding pocket (Fig. 16).

*From the **View** menu, *choose Model Display*, then **Show Residue Labels**.

Shrink* the entry for **Ligand-2 and *expand* the entries for **Chain B** and for **Fragment 1-99**.

**Click* in the main window with the arrow tool to deselect any highlighted regions.

*Return to the Model Explorer table. With the Control key depressed, *click* on the Asp-25 residue that is present in Chain B, and the Asp-25 residue that is present in Fragment 1-99.

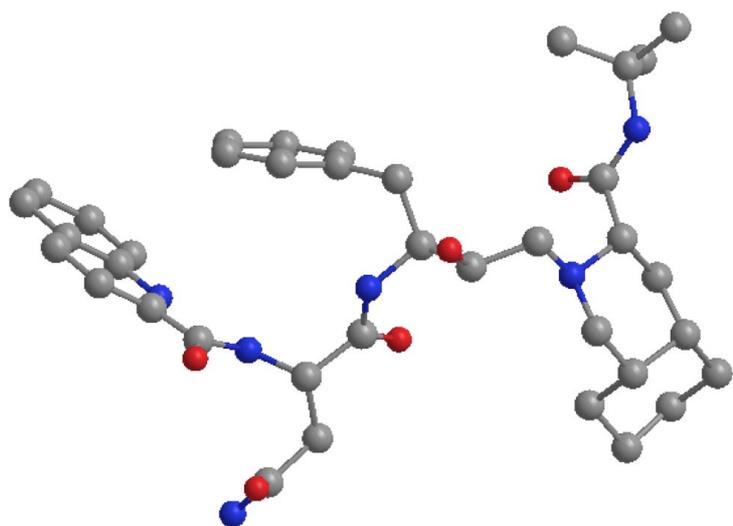


Figure 17 Active conformation for saquinavir.

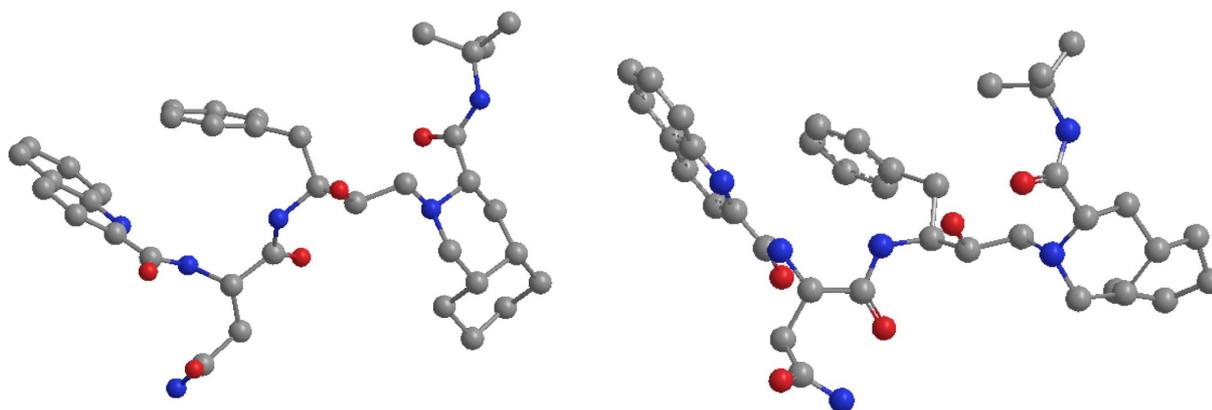


Figure 18 Comparison of A) the active conformation of saquinavir with B) the energy-minimised conformation of saquinavir from Part A.

Note that it is not valid to measure the steric energy of the active conformation because the positions of the atoms in saquinavir are obtained by experiment and so there is an experimental error involved in their position. This means, in turn, that bond lengths, bond angles and torsion angles are not exact. Moreover, the hybridisation states of the atoms are not properly defined, which means that it is not possible to carry out an energy minimisation.

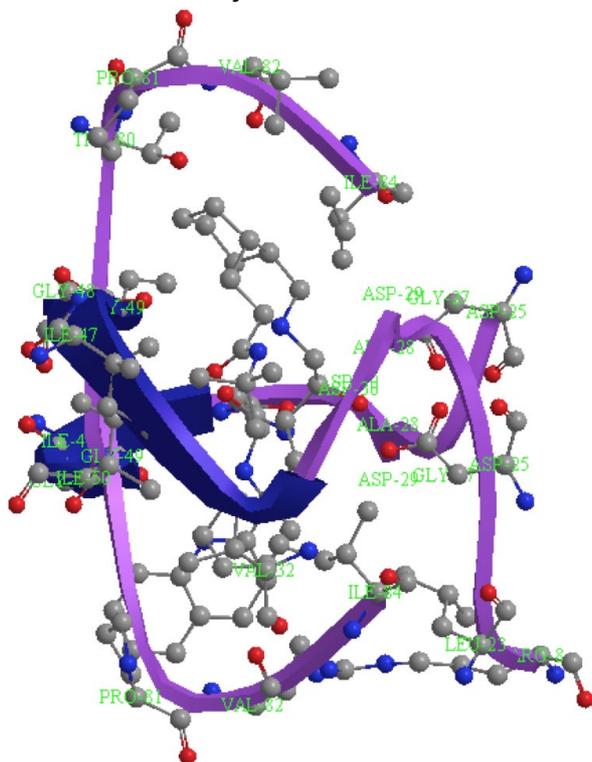


Figure 20 Initial paste of saquinavir and the closest amino acid residues in the binding site.

The following operations are carried out.

*From the **View** menu, choose **Model Display**, then **Display Mode. Select Sticks**.

*To alter saquinavir back to a ball and stick model, *select* the ligand by finding it in the **Model Explorer** table (**Ligand 2-2**). Alternatively, *double click* on any of the atoms belonging to the ligand.

*Make sure that the mouse is placed over the selection and *right click* the mouse to produce a menu. Choose **Display Mode**, then **Ball and stick**.

Select Solvent* in the **Model Explorer table and *right click* the mouse to reveal a menu.

**Choose Cut*.

A ball and sphere model for the ligand will now be visible, surrounded by labelled amino acid residues in stick format (Fig. 21). However, no hydrogen atoms or hydrogen bonds are visible in the model binding site.

PART D

The model binding site obtained in part C identifies the amino acids that are close to saquinavir, but fails to reveal the hydrogen atoms or the hydrogen bonding interactions. However, it is possible to create the binding site in a different way such that hydrogen bonding interactions *can* be observed. Admittedly, it is a bit 'klunky' and is not totally accurate.

1. Select saquinavir and the closest amino acid residues in the active site.

*Go back to the window containing the full crystal structure. Saquinavir and the closest amino acids should still be highlighted.

*Choose the **rotate** or **translate** tool, then *click* in the main window. It's important not to use the Select tool, or the selection will be lost.

*From the **Edit** menu, choose **Copy As**, then *select ChemDraw Structure*.

*Open a new window and *paste* the selection into the new window (Fig. 22).

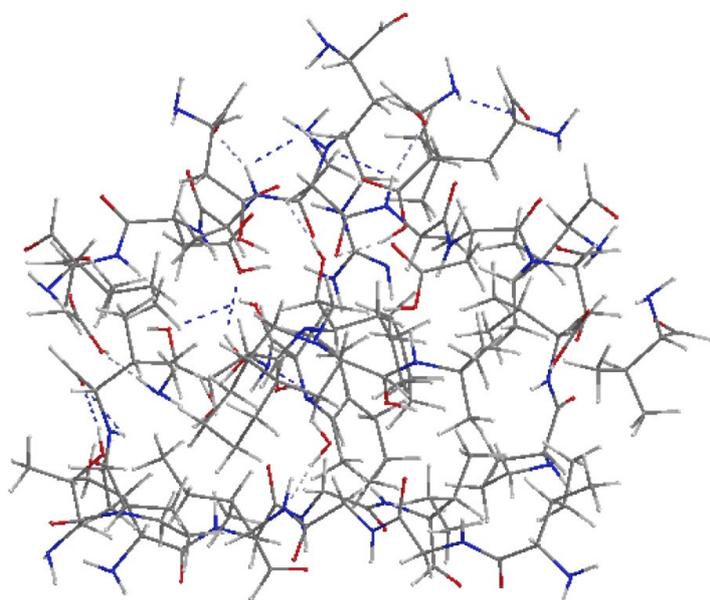


Figure 22 The structure of saquinavir and closest amino acid residues obtained from the initial paste.

The resulting image (Fig. 22) needs to be tidied up. If the structure is not already in the stick format, carry out the following.

*From the **View** menu choose **Model Display**, then **Display Mode. Select Sticks**.

The following operations can then be carried out.

* *Select* saquinavir from the **Model Explorer** table (**Fragment-1**). Alternatively, *double click* on saquinavir in the main window.

* Keep the mouse over the selected entry and *right click* the mouse.

* From the menu that appears, choose **Display Mode**, then **Ball and stick**.

Unfortunately, it is not possible to label the amino acid residues in this extract. However, it is possible to show hydrogen atoms and hydrogen bonds.

*From the **View** menu, choose **Model Display. Select Show Hydrogen Atoms**, then **Show Polar**.

*From the **View** menu, choose **Model Display. Select Hydrogen Bonds**, then *select Show Intermolecular*.

*Select **Solvent** in the **Model Explorer** table and *right click* on the mouse to reveal a menu. Choose **Visibility**, then *select* **Hide group**.

The resulting structure now shows hydrogen atoms on heteroatoms such as oxygen and nitrogen. Hydrogen bonds are also shown as dashed lines (Fig. 23).

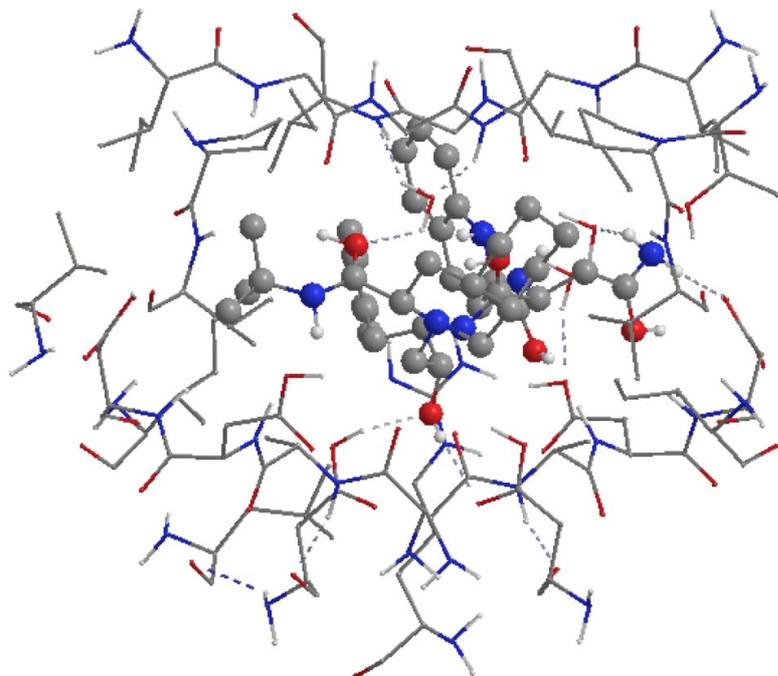


Figure 23 Saquinavir in the active site. Note that not all the hydrogen atoms shown are accurate.

Before identifying possible H-bond interactions, it is necessary to make some corrections to the ligand since the process of adding hydrogen atoms assumes that all the atoms in the structure are sp^3 hybridised. This means that some atoms (such as carbonyl oxygen atoms) have a hydrogen atom present which should not be there. Therefore, these hydrogen atoms have to be hidden. There are five 'wayward' hydrogen atoms to be considered, one on the nitrogen of the heteroaromatic ring and four on the carbonyl oxygen atoms (Fig. 24A). These can be hidden as follows.

*Choose the **Select** tool and *select* the atom or atoms concerned.

**Hold* the mouse cursor over any of the selected atoms and *right click* to open a menu.

* Choose **Visibility**, then **Hide**.

The corrected ligand should now look like figure 24B and the binding site should now look like figure 25.

Note: Although the structure has been corrected to hide hydrogen atoms that should not be present, the orientation of some hydrogen atoms may be faulty. For example the NH protons of amide bonds are not orientated properly since they should be in the plane of the amide group. Instead they are orientated as if the nitrogen atom is sp^3 hybridised.

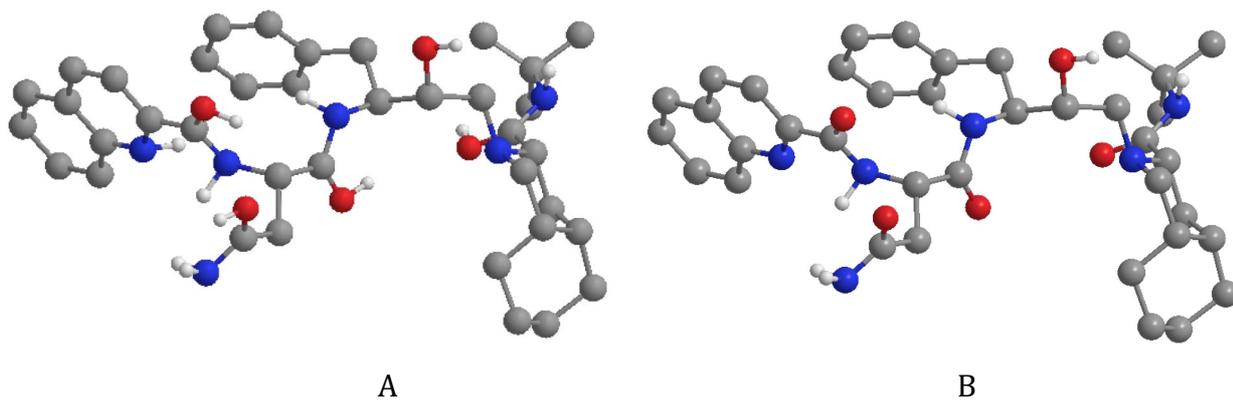


Figure 24 Saquinavir. A) Faulty structure and B) corrected structure.

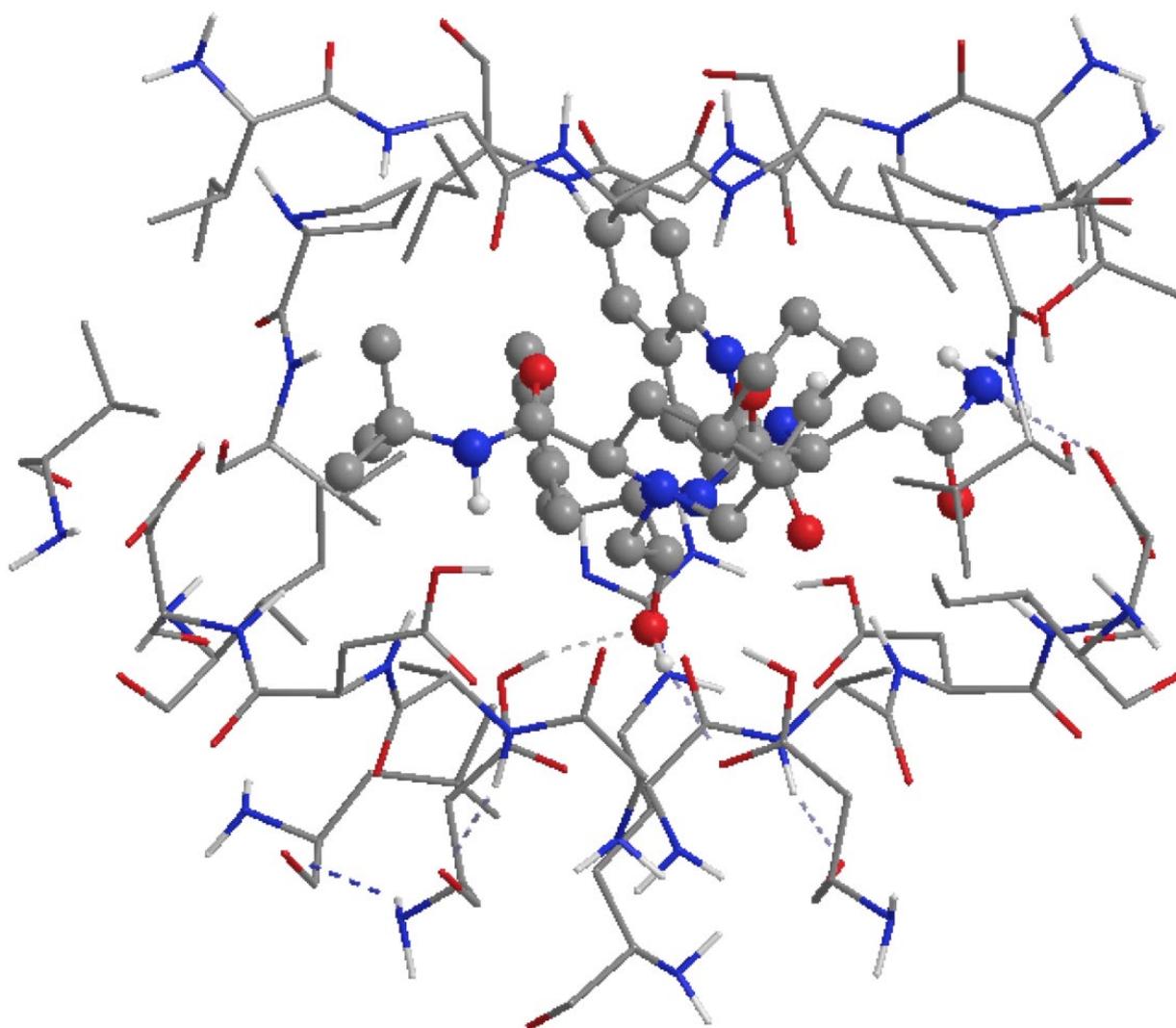


Figure 25 Corrected structure of the binding site showing polar hydrogen atoms and H-bonds.

PART E

The model binding sites obtained in parts C and D can now be used to identify specific binding interactions and the amino acid residues involved. It helps to view these structures in perspective in order to appreciate depth of field.

*From the **View** menu, choose **Model Display**, then **Perspective**.

1. Identify the binding interactions of saquinavir with Asp-25 and Asp-25'.

**Zoom* into the relevant region of the active site. Identify the hydrogen bonds indicated by dashed lines.

*Choose the *Select* tool .

*With the shift key depressed, *click* on the HBD and the HBA involved in the H-bond.

*From the **Structure** menu, choose **Measurements**, then **Display Distance Measurement**.

Two H-bonds are visible between the aspartate residues and the alcohol group of saquinavir. The alcohol acts as an HBD to one of the aspartate residues, and as an HBA to the other. The lengths of the H-bonds are 2.1Å and 1.9Å respectively (Figs. 26 & 27).

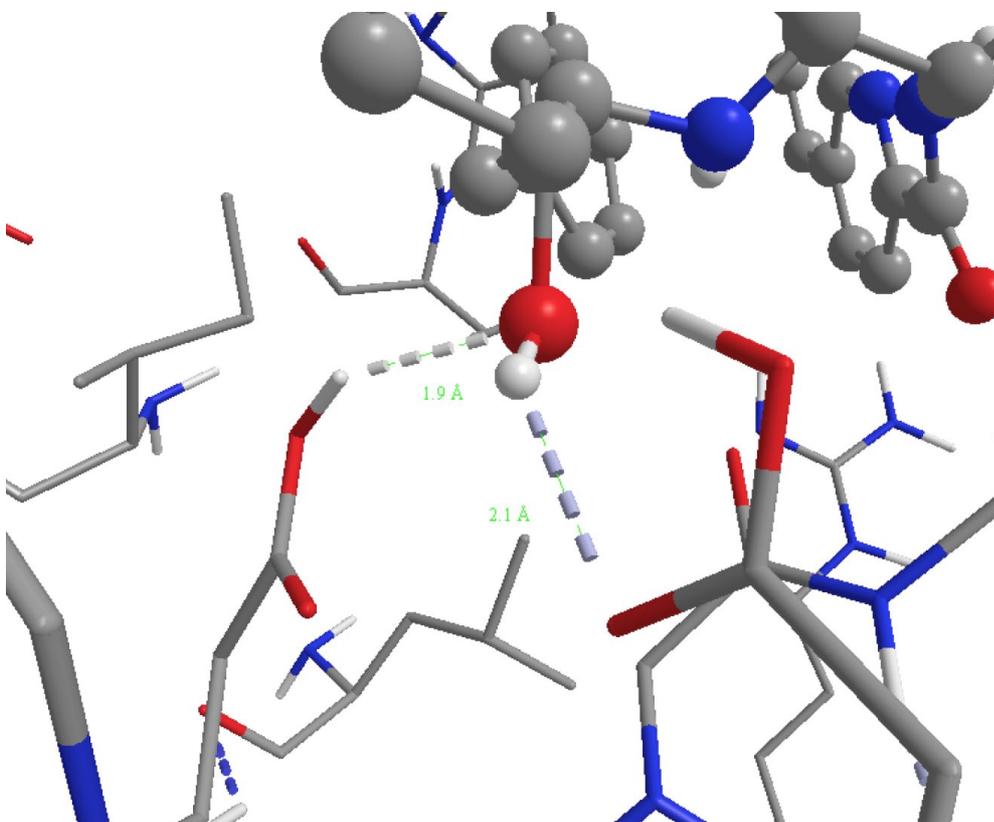


Figure 26 Interactions with Asp-25 and Asp25' (1.9Å and 2.1Å).

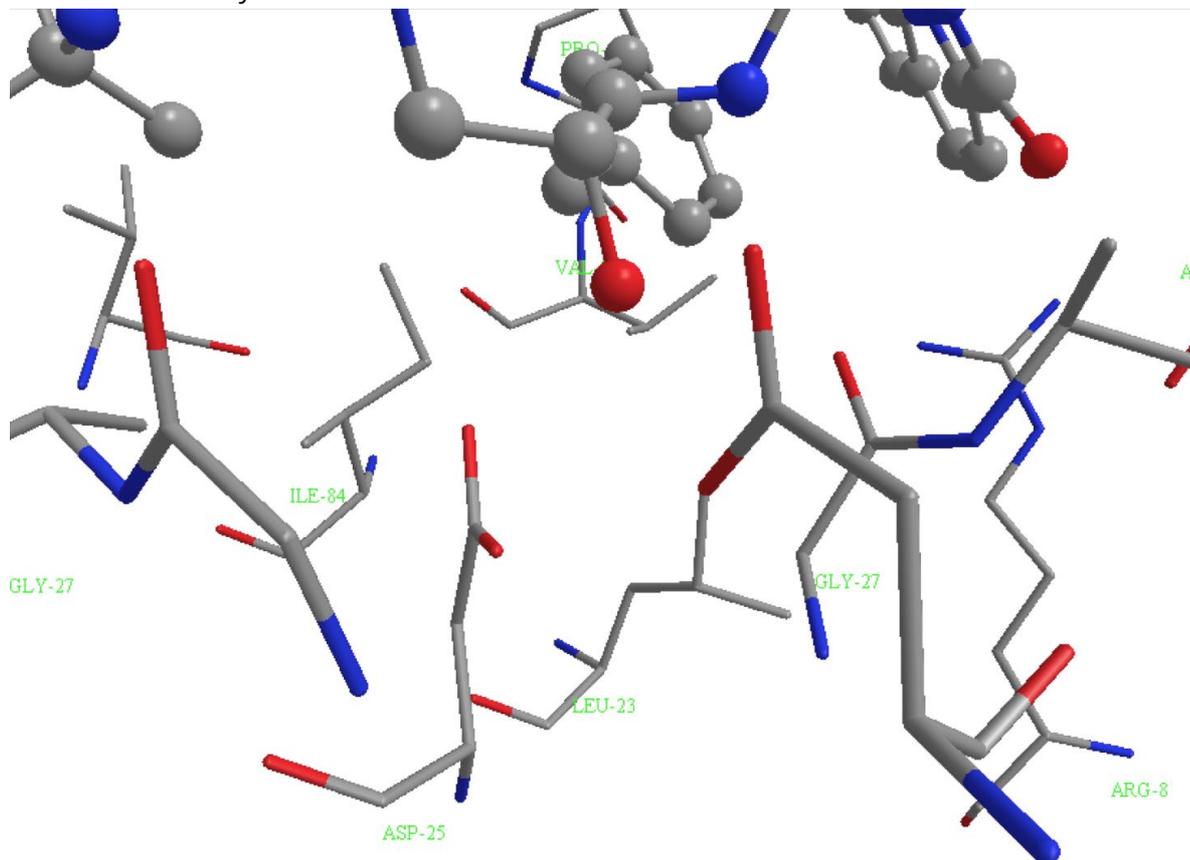


Figure 27 Identification of Asp-25 and Asp25' (compare figure 26).

2. Identify binding interactions with Asp-30.

The primary amide group of saquinavir forms a hydrogen bond measuring 2.4Å with the carboxylate group of Asp-30 (Figs. 28 & 29). The amide NH₂ acts as a HBD, and the carboxylate group acts as an HBA.

In addition, there is a hydrogen bond involving the carbonyl oxygen of the primary amide, which has not been picked up by Chem3D. This is to the NH proton of Asp-30 and measures 2.2Å. The NH is part of the peptide bond between Asp-30 and Asp-29.

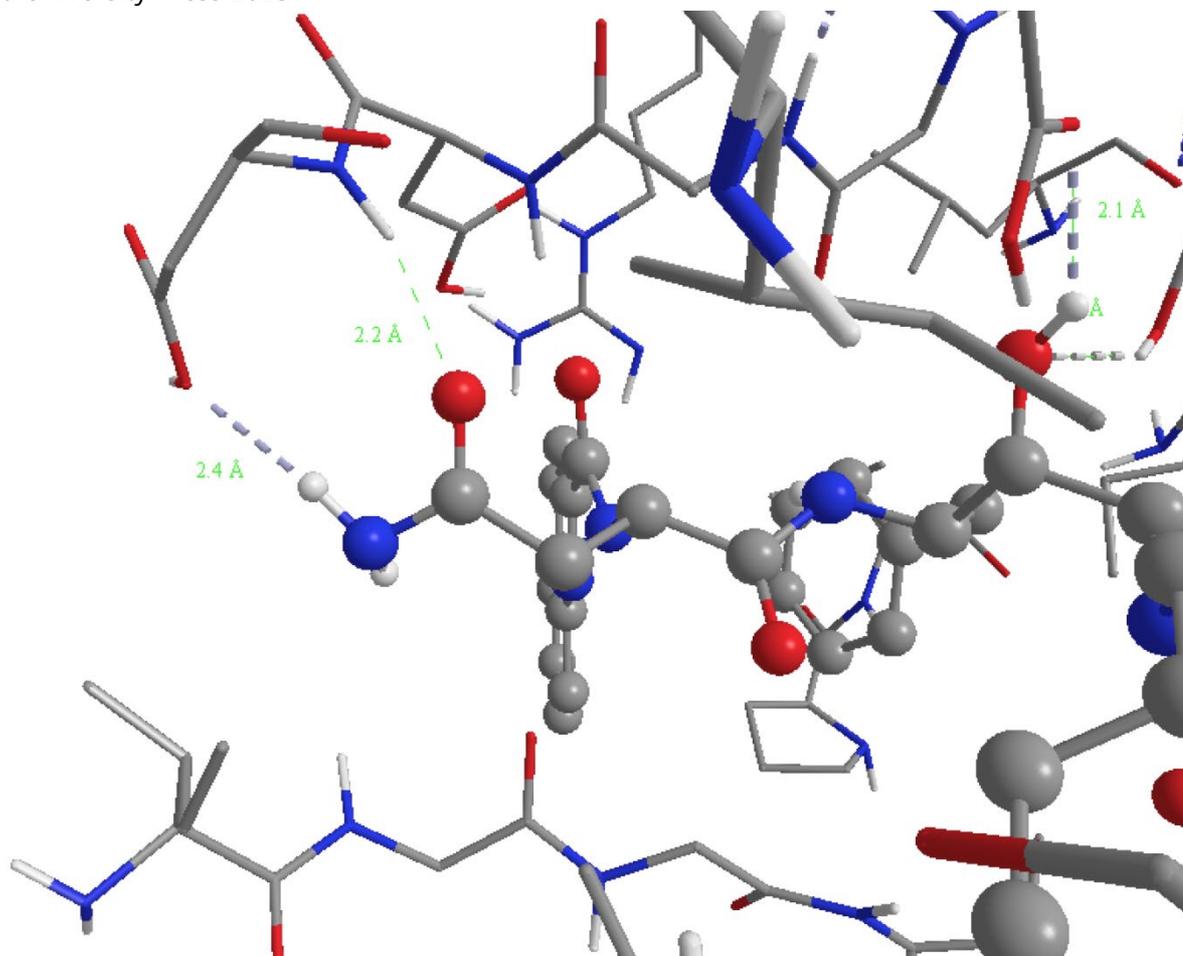


Figure 28 Hydrogen bond between aspartate (Asp-30) and the primary amide NH₂ of saquinavir. The peptide NH group of Asp-30 is also within H-bonding range of the carbonyl oxygen.

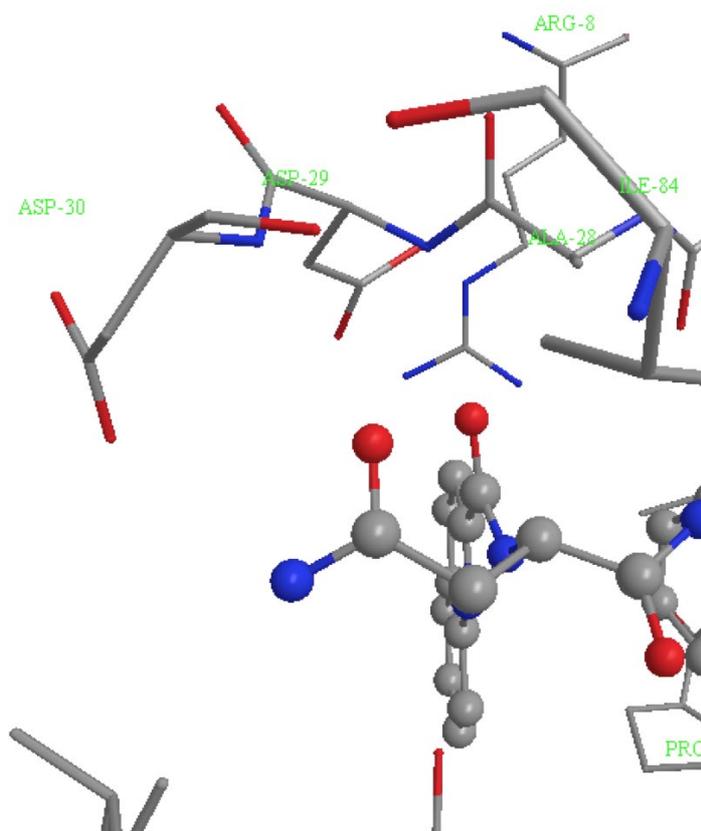


Figure 29 Identification of amino acid residue Asp-30 (compare Fig 28).

3. Identify hydrogen bond interactions with Asp-29 and Gly-48.

Another hydrogen bond that has not been identified by Chem3D is between Asp-29 and the amide group attached to the quinoline ring system of saquinavir (Figs. 30 and 31). This measures 1.9Å and involves the carbonyl oxygen of the amide group acting as a HBA. The NH proton of Asp-29 is the HBD and is part of the peptide bond between Asp-29 and Ala-28. The NH proton of saquinavir's amide group is also capable of forming a hydrogen bond measuring 2.5Å to the carbonyl oxygen Gly-48. This is part of the peptide bond between Gly-48 and Gly-49. The length of this H-bond is likely to be shorter than this, since the orientation of the NH proton is faulty (see note 1).

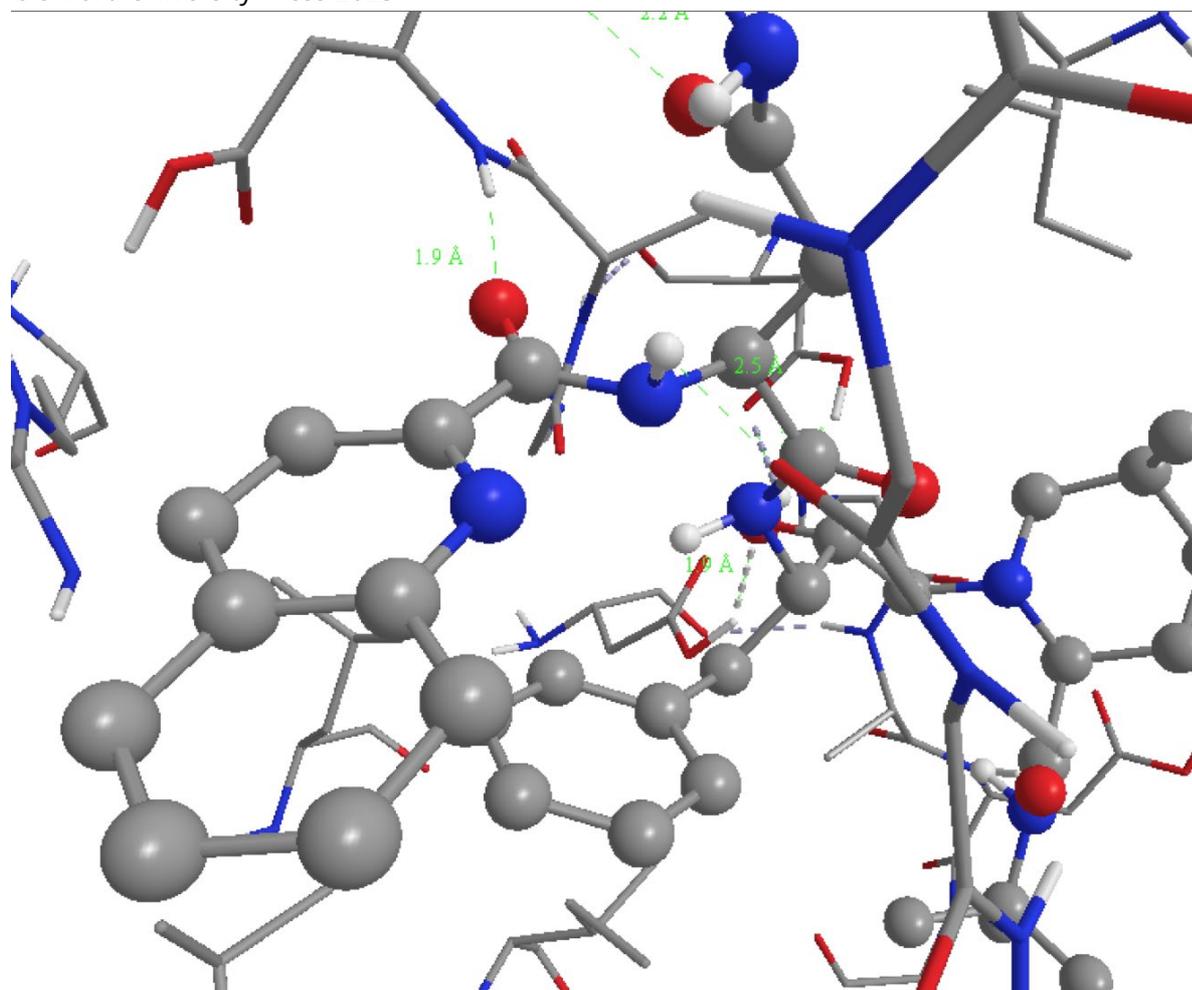


Figure 30 Hydrogen bonds involving the amide group of saquinavir with Asp-29 and Gly-48.

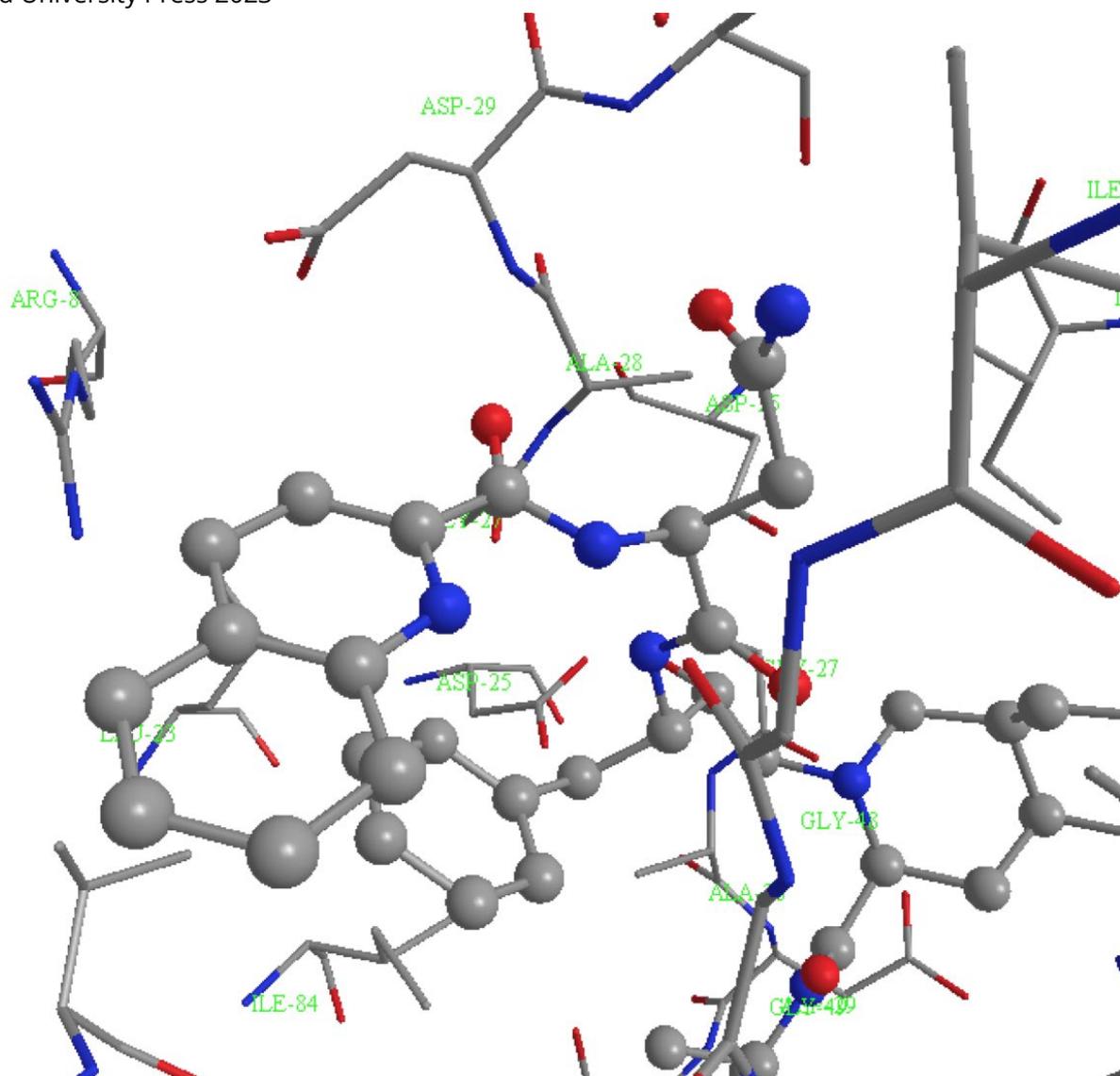


Figure 31 Identification of amino acid residues Asp-29 and Gly-48 (compare Fig. 30).

4. Show the H-bond interaction between saquinavir and Gly-27

There is meant to be a hydrogen bond interaction involving another amide group in saquinavir where the NH proton acts as a HBD (Figs. 32 & 33). The hydrogen bond should be to the carbonyl oxygen of Gly-27, where the carbonyl group is part of the peptide bond between Gly-27 and Ala-28. This measures 3\AA , which is too long for an effective hydrogen bond, but it has to be appreciated that the N-H proton in the model binding site is wrongly orientated (see note 1). If it was correctly orientated it would be placed much closer to the carbonyl oxygen of Gly-27 allowing a more effective interaction.

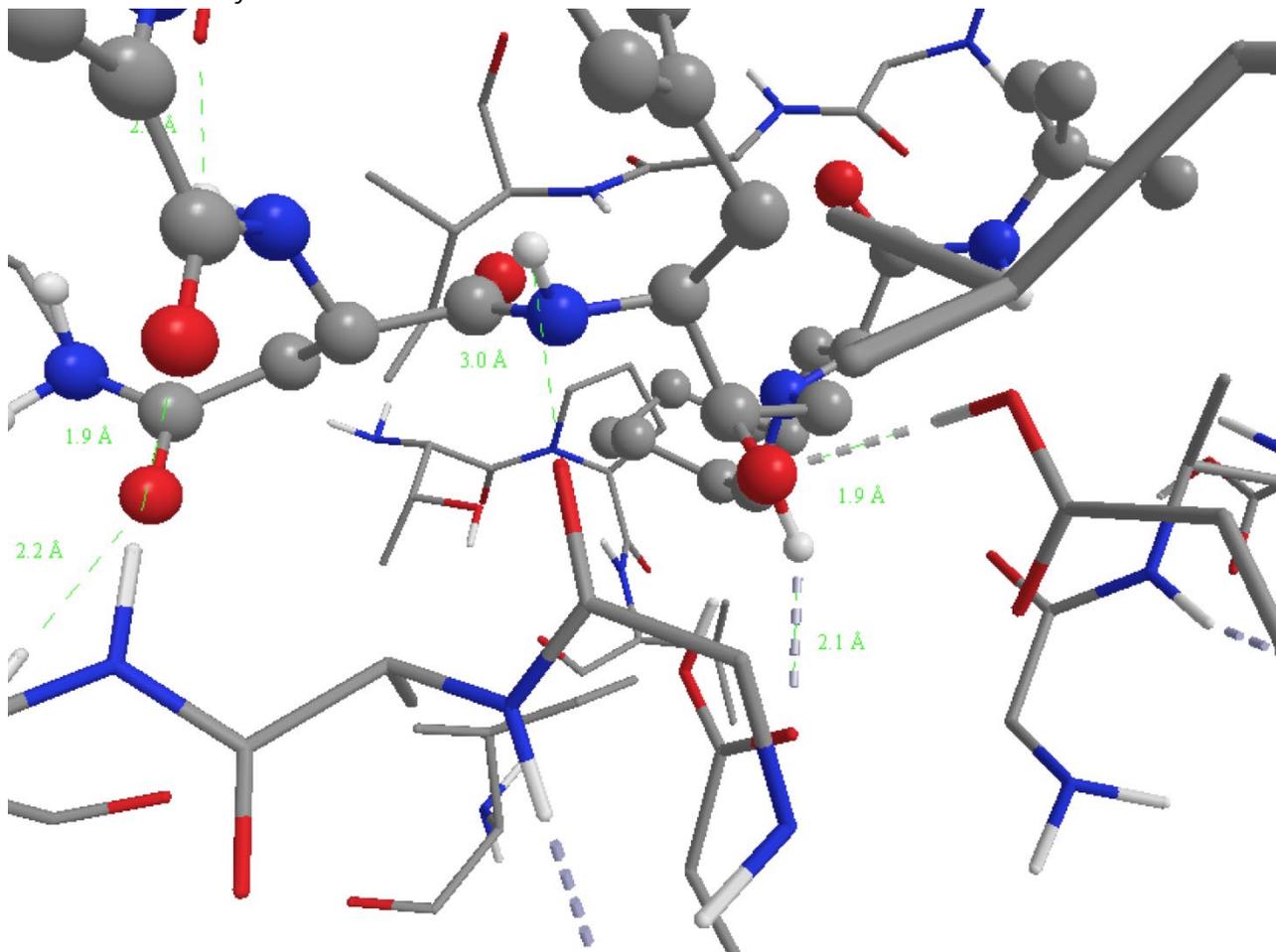


Figure 32 Identification of a hydrogen bond between saquinavir and Gly-27.

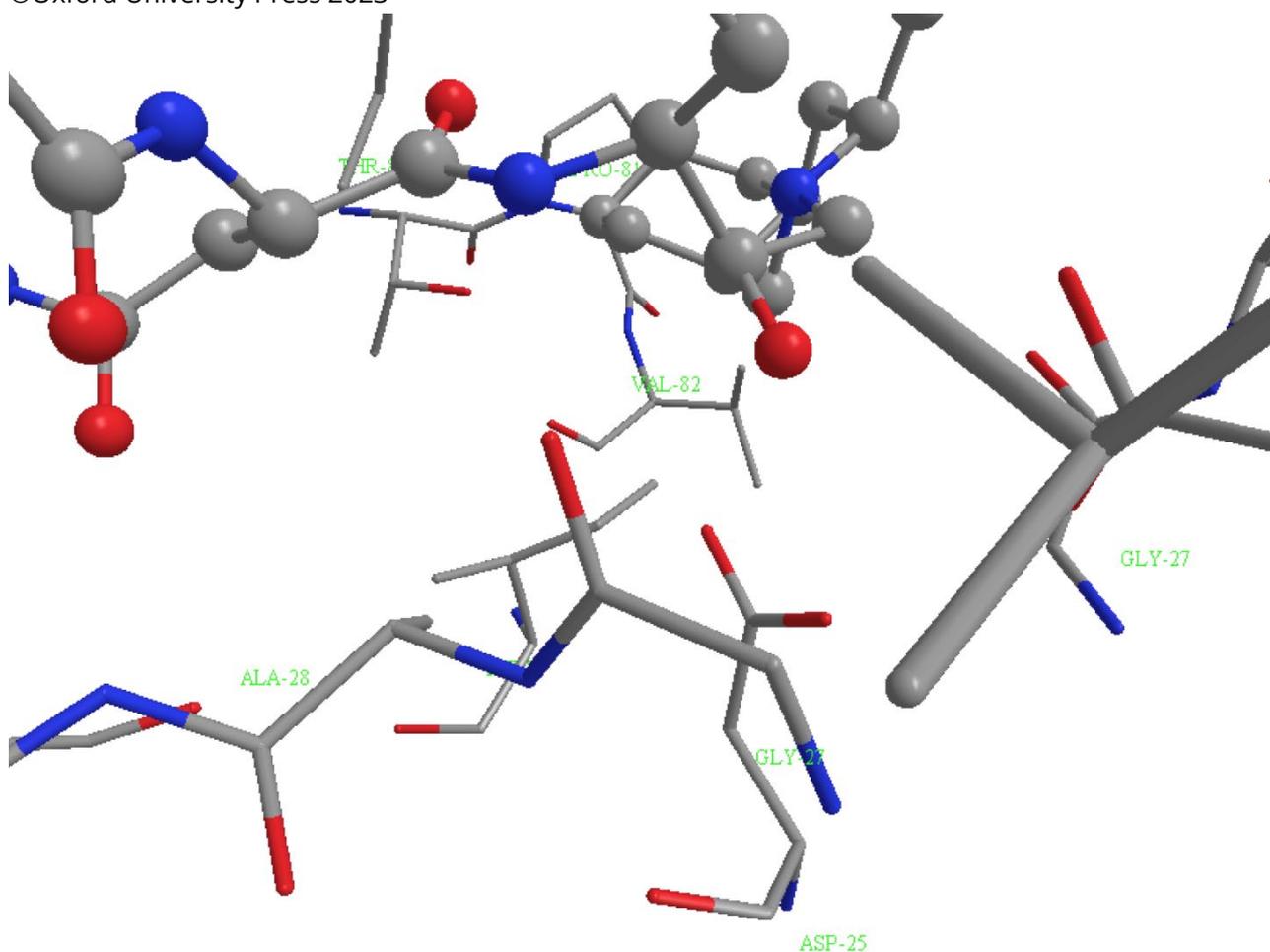


Figure 33 Identification of amino acid residue Gly-27 (compare Fig. 32).

5. Show the H-bond interactions between saquinavir and the amino acid residues Ile-50 & Ile-50'.

Hydrogen bond interactions are possible between saquinavir and the amino acid residues Ile-50 & Ile-50' in the flap region of the enzyme. This involves a bridging water molecule. In order to observe this, it will be necessary to reveal the water molecules that were previously hidden.

* **Select Solvent** in the **Model Explorer** table and *right click* the mouse to show a menu.

* **Choose Visibility** then **select Show Group**.

* **Zoom** into the region of the binding site occupied by the terminal amide group of saquinavir containing the *tertiary* butyl substituent (Figs. 34 and 35).

A hydrogen bond interaction measuring 2.2\AA is revealed between the amide carbonyl oxygen and the bridging water molecule. The water molecule then forms a hydrogen bond to the peptide bond between Gly-49 and Ile-50 (Ile-50NH) in one protein subunit, and a second hydrogen bond to the peptide bond between Gly-49' and Ile-50' (Ile-50'NH) in the other protein subunit. These measure 2.4\AA and 2.0\AA .

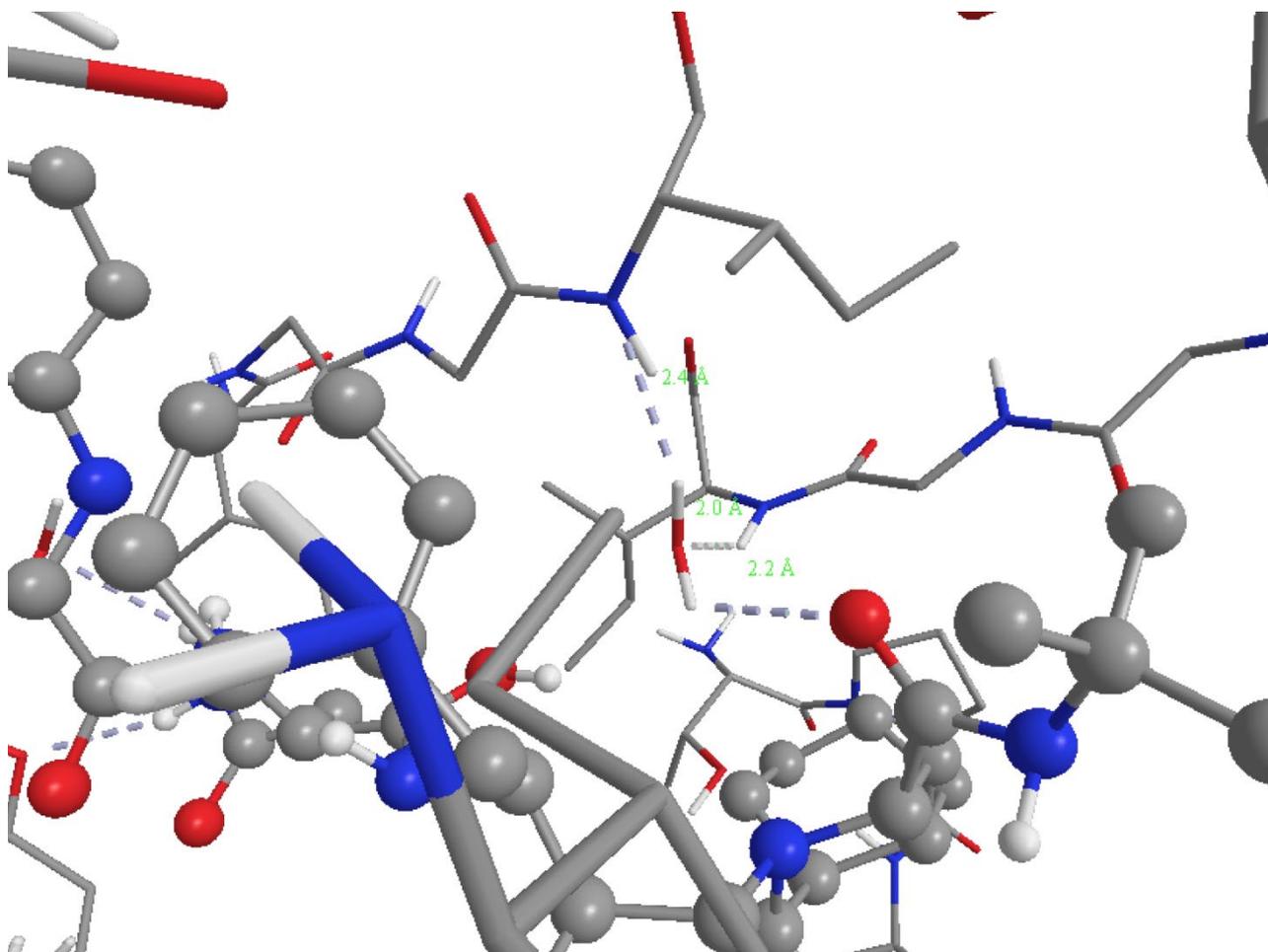


Figure 34 Water acting as a hydrogen-bonding bridge between saquinavir and the flap region (Ile-50 and Ile-50').

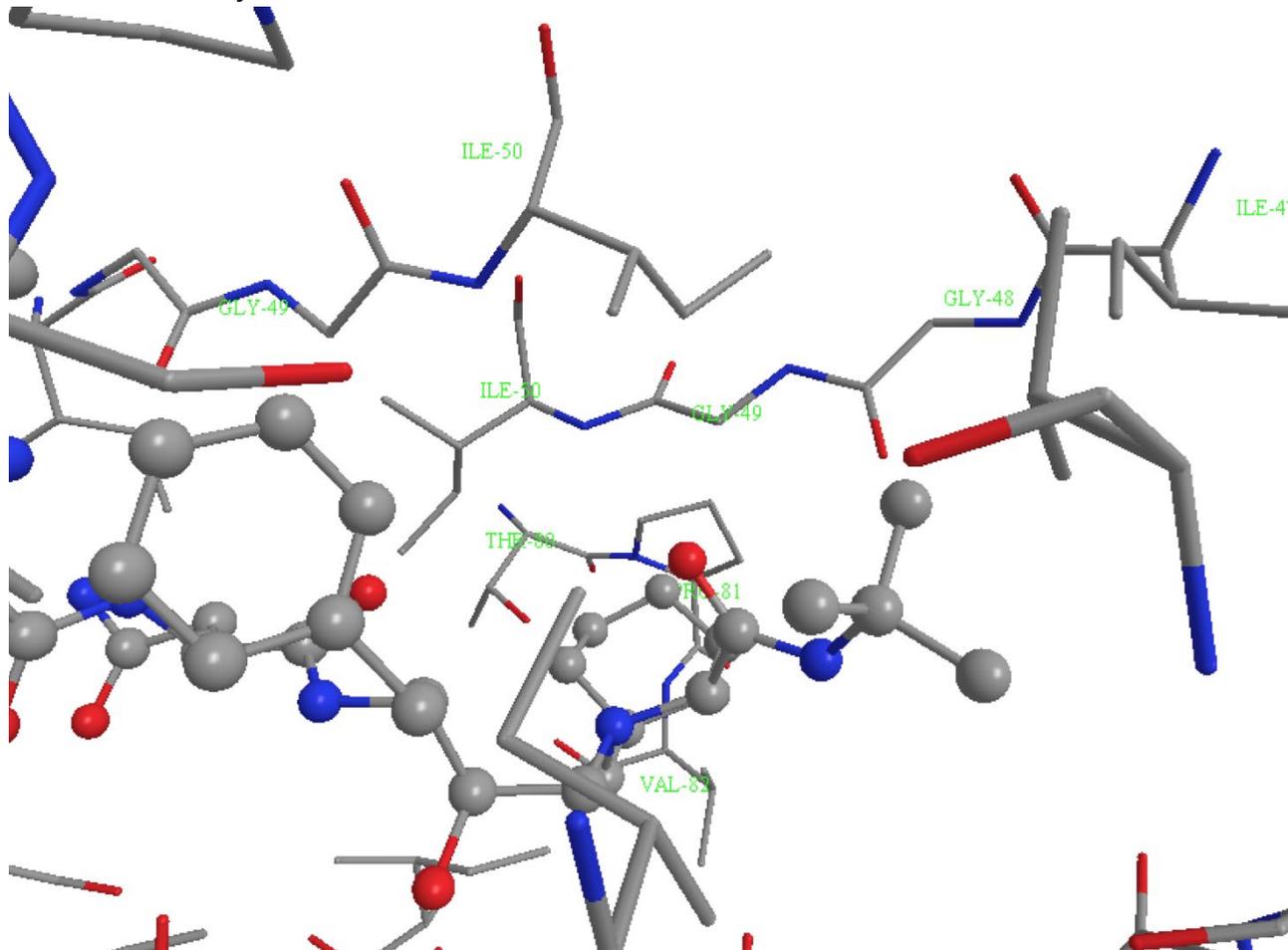


Figure 35 Identification of relevant amino acid residues (compare Fig. 34).