MOLECULAR MODELLING EXERCISES

EXERCISE 20.7 Asunaprevir

INTRODUCTION

Asunaprevir (Fig. 1) is a new antiviral drug for the treatment of hepatatis C (section 20.10 in the textbook). It was designed to inhibit a non-structural viral protein called NS3/4A protease. This protein is an enzyme that catalyses the hydrolysis of peptide bonds. The binding interactions of the drug with its binding site are predicted to involve a number of key hydrogen bonds as shown in figure 2. In this exercise, we will download the crystal structure of the enzyme/drug complex and identify the interactions highlighted in figure 2.

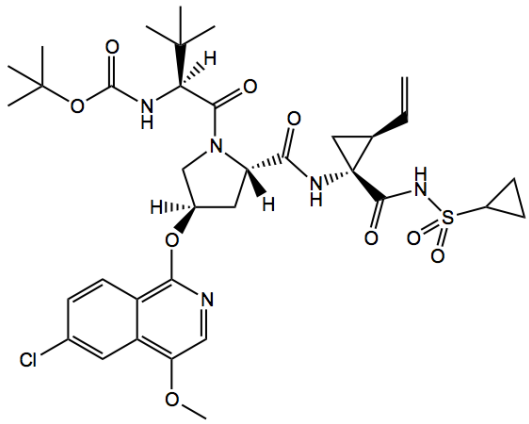


Figure 1 Asunaprevir.

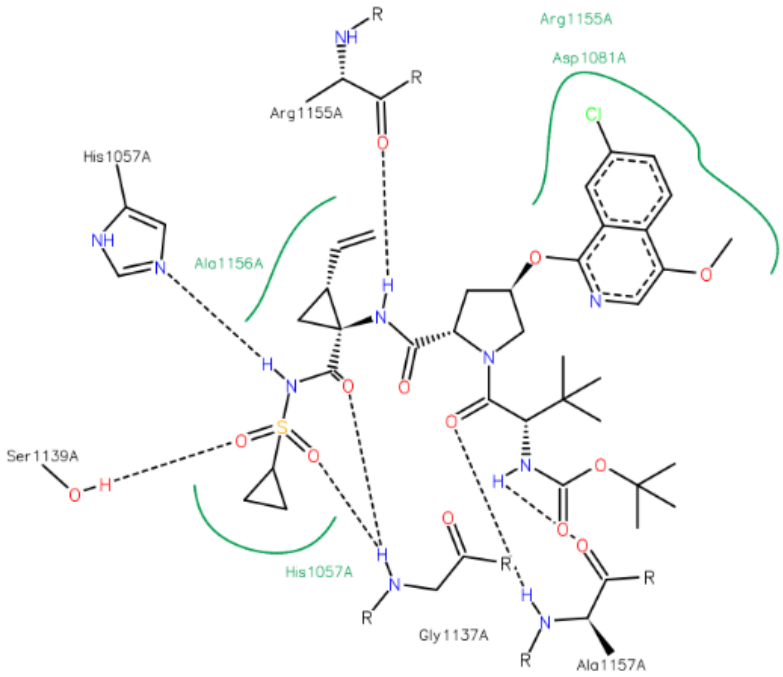


Figure 2 Binding interactions of asunaprevir in the active site of NS3/4A protease – a PoseView diagram available from the protein data bank (pdb 4wf8).

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. The ChemDraw file for asunaprevir is available in the ChemDraw folder.

*Create an energy-minimised structure of asunaprevir and identify the number of rotatable bonds, the number of HBAs, the number of HBDs, the molecular weight and the LogP value.
*Download the crystal structure of the NS3/4A protease enzyme with bound asunaprevir (pdb 4WF8).
*Examine the features that are present in the enzyme, and the location of the binding site and the inhibitor.
*Identify the active conformation of asunaprevir and compare it with the energy-minimised conformation created in part A.
*Create a model of the binding site and identify key binding interactions between asunaprevir and amino acid residues.
*Identify the atoms acting as HBDs and HBAs, and measure the lengths of hydrogen bonds
*Identify the amino acid residues involved.

PROCEDURES

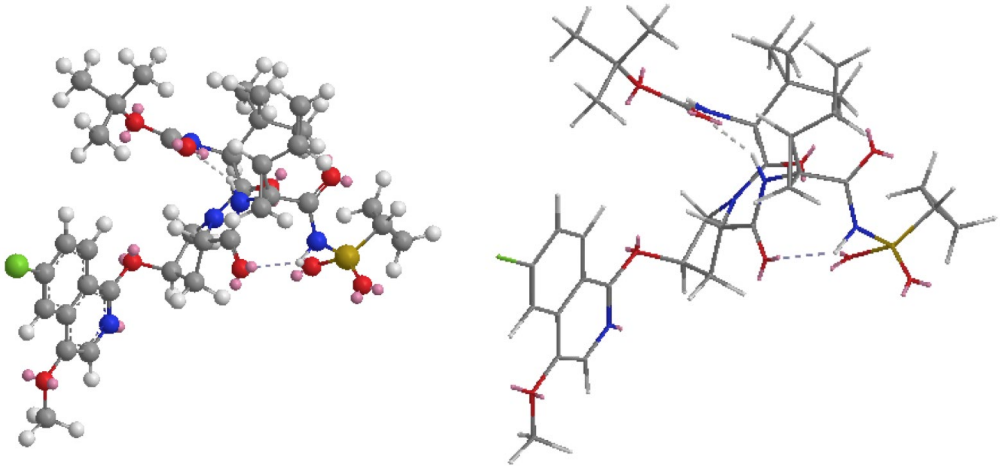
There are various approaches that you can use to tackle these molecular modelling exercises. The following procedures illustrate how you might tackle this particular exercise, but they are not meant to be prescriptive. Note also that the results obtained may vary depending on the computer and the version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

PART A Asunaprevir

1. Create the energy-minimised 3D model of asunaprevir.

Open* **ChemBio3D or **Chem3D**.
*From the **File** menu, *choose* **Open**, then *select* the ChemDraw file for asunaprevir.
Click* **Open.
*Energy minimise the structure 

. You may get something like the following conformation (Fig. 3). The steric energy is given in the bottom window as 202 kcal/mol.



A B

Figure 3 Energy-minimised structure of asunaprevir as A) a ball and stick model and B) a stick model.

2. Identify the number of rotatable bonds in asunaprevir.

In medicinal chemistry, rotatable bonds are those that would result in significant changes in conformation. Therefore, rotatable bonds to substituents such as CH_3, OH and NH_2 are not included in the total. The number of rotatable bonds in the structure can be identified as follows.

*From the **Calculations** menu, *choose* **Compute Properties**.
Expand* **Molecular Topology and *select* **Num Rotatable Bonds,** then *click* **OK**.
The number of rotatable bonds is given as 18 in the bottom window.

However, ChemBio3D/Chem3D overestimates the number of rotatable bonds by including bonds that have restricted rotation such as amide bonds. Figure 5 shows 11 rotatable bonds coloured red that would result in significant changes in conformation. The blue-coloured bonds are other bonds that have been identified by ChemBio3D/Chem3D as rotatable. However, the bonds where nitrogen is linked to a carbonyl group or sulphone group are not freely rotatable, while the bonds to *tertiary*-butyl groups would not result in significantly different conformations when they rotate. This illustrates that software property predictions are not necessarily accurate. Nevertheless, there are a large number of rotatable bonds in asunaprevir, and so there is no guarantee that the energy-minimised conformation obtained represents a global minimum, let alone the active conformation.

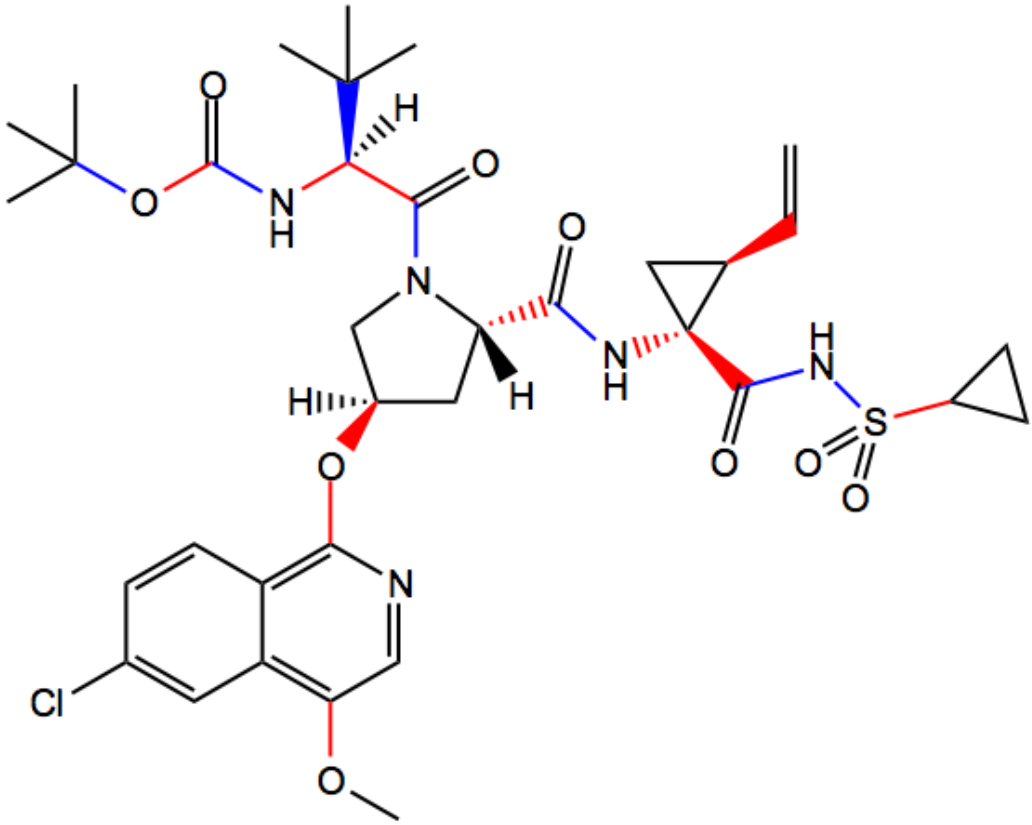


Figure 5 Rotatable bonds that would result in significantly different conformations are shown in red. Additional rotatable bonds identified by ChemBio3D/Chem3D are coloured blue.

3. Identify the number of hydrogen bond acceptors and hydrogen bond donors in the structure.

ChemBio3D/Chem3D can calculate the various factors covered by Lipinski's rules (section 11.3 in the textbook) – molecular weight, log*P*, the number of hydrogen bond donors, and the number of hydrogen bond acceptors. This is carried out as follows.

*From the **Calculations** menu, *choose* **Compute Properties**.
Expand* the entry for **ChemPropPro.
Click* on **Lipinski's Rule, then *click* **OK.**
The molecular weight, number of HBAs, number of HBDs, number of rotatable bonds and log*P* values are shown in the bottom window as 747.27, 9, 3, 18, and 6.312 respectively. The HBAs and HBDs in the structure are shown in figure 6. Note that nitrogen atoms next to carbonyl groups are not counted as HBAs because their lone pair of electrons interacts with the carbonyl group.

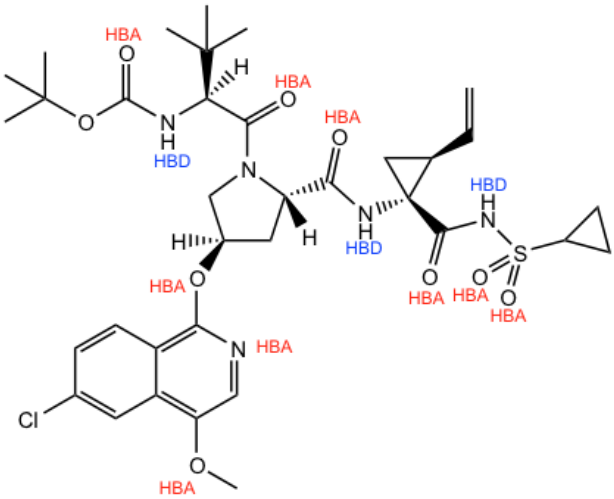
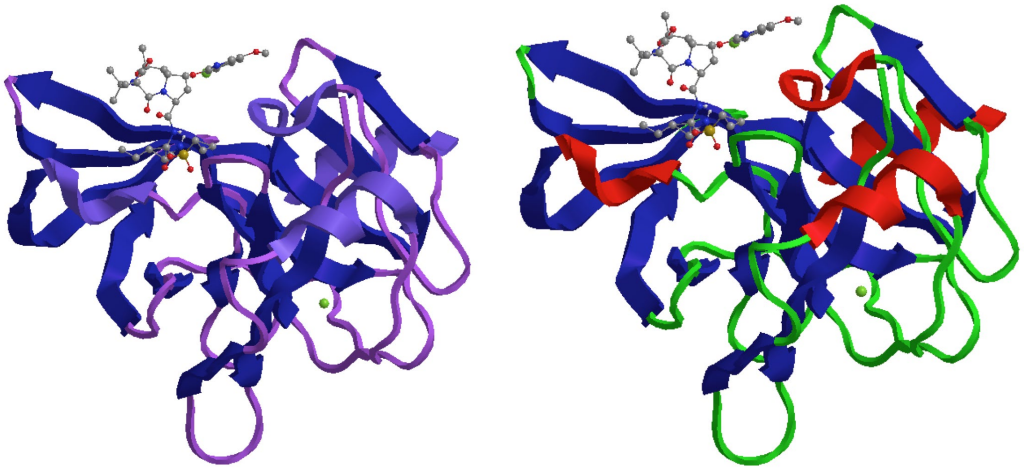


Figure 6 HBAs and HBDs in the structure.

PART B Features of the asunaprevir - NS3/4A protease crystal structure.
A study of the asunaprevir-NS3/4A protease crystal structure allows you to identify the location of the binding site of asunaprevir, and to identify the active conformation of the ligand.

1. Download the crystal structure of the NS3/4A protease with bound asunaprevir
You will need to be connected to the internet in order to do this.
*From the **Online** menu, *choose* **Find Structure from PDB ID** (PDB stands for the protein data bank). A text box opens up.
Enter* the PDB code (4WF8**) into the text box.
Click* on **Get File.
The protein will now appear as a ribbon diagram, with any ligands present shown in the ball and stick mode (Fig. 7). The ribbon format represents the polypeptide backbone and allows you to identify regions of secondary structure. You will see evidence of alpha helices (coils) coloured light purple, beta-pleated sheets (broad arrows) coloured dark blue, and connecting strands coloured pink.



A B

Figure 7 Protein-ligand complex (pdb code 4WF8).

The distinction between the alpha-helices and the beta sheets is not that clear in figure 7A since the colours are so similar, and so you might wish to recolour the structure. This can be done as follows.
*From the **File** menu, *choose* **Model Settings**.
Click* on the tab for **Colors and Fonts.
*Under the section on **Model Colors**, *modify* the colours used for the Alpha helix, Beta sheet and Coil. *Choose* red for alpha helices, blue for sheets and green for coils.
Click* on **Apply to see the effect in the main window. If satisfied, *click* **OK.**
The resulting structure (Fig.7B) shows the 4 helices in the structure more clearly.

2.Investigate the main features of the crystal structure.

Click on the **Model Explorer** tab at the left hand margin of the window. This opens up the Model Explorer Table to show you that there is 1 protein chain present (Chain A) (Fig. 8). *Click* on it and you will find that the whole structure is highlighted in yellow in the main window (Fig. 9).

Figure 8 Model Explorer table.

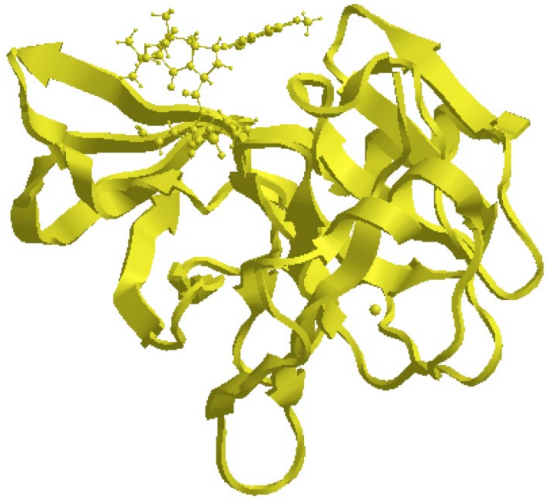


Figure 9 Identification of Chain A.

Note that the ligands are also selected when selecting chain A. This is because the ligands are included under the heading of Chain A.
*Go into the Model Explorer table and *click* on the + sign to the left of Chain A. This expands the entry to show what features are included (Fig. 10). One Fragment representing amino acids 989-1180 is revealed along with entries for three ligands.

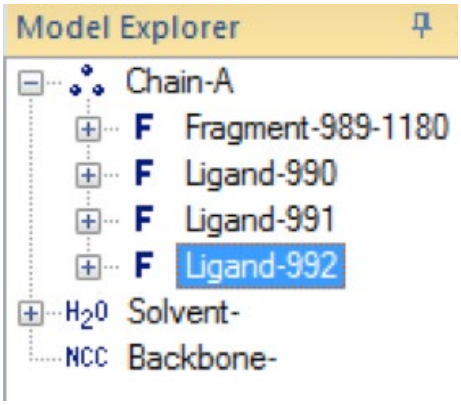


Figure 10 Model Explorer – Chain A before and after expansion.

Click* on **Fragment-989-1180 and you will find that the protein ribbon structure is highlighted (Fig. 11).
Click* on **Ligand-992 and you will find the main ligand (asunaprevir) is highlighted (Fig. 12).

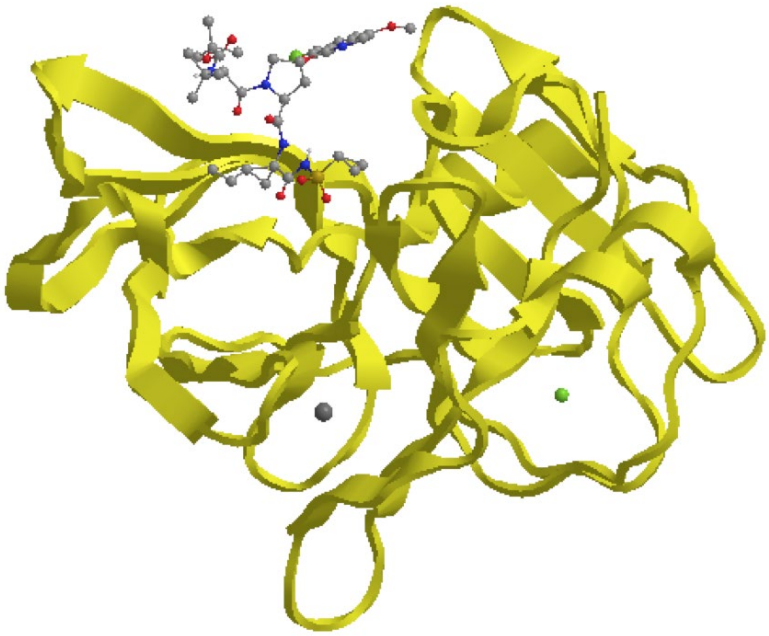


Figure 11 Protein chain highlighted.

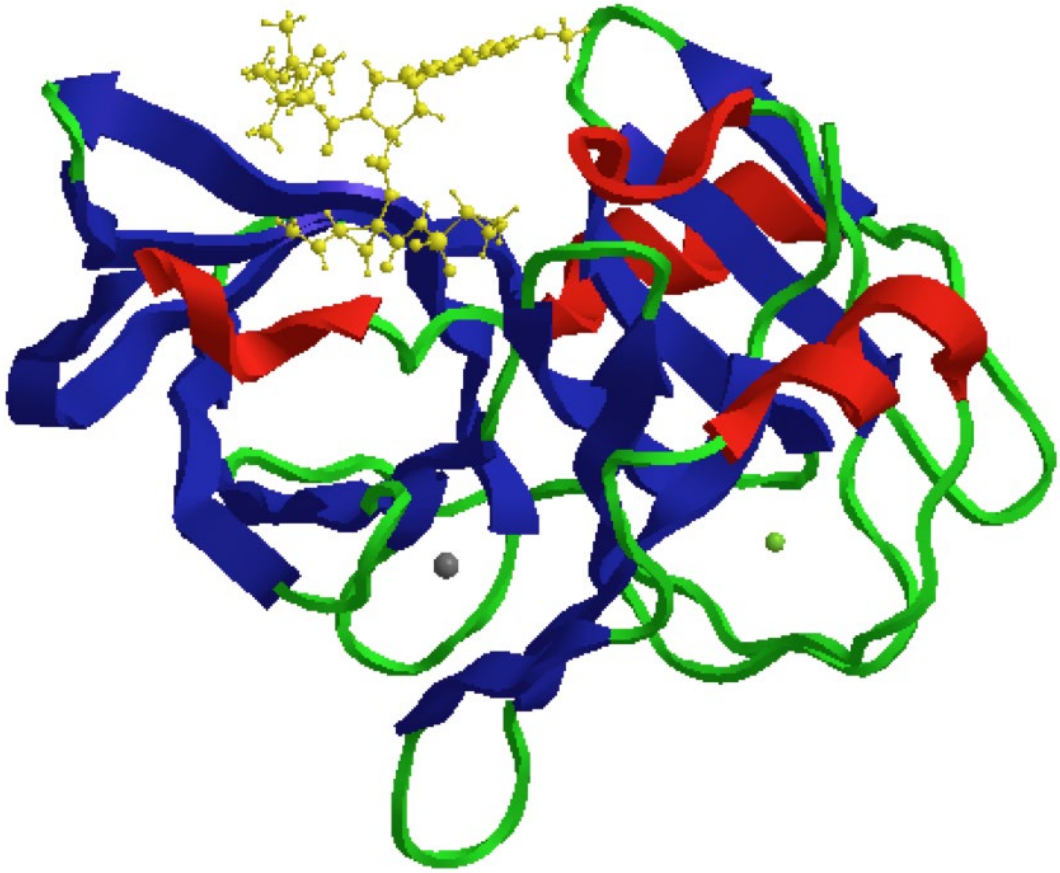


Figure 12 Ligand highlighted in the crystal structure.

3. Identify the active conformation of asunaprevir.

*_Use_ the **translate** tool to centre your view on asunaprevir bound to chain A.
*Use the mouse scroll to *zoom* in on the ligand and the binding site (Fig. 13).
This demonstrates that the binding site is quite shallow and that much of the bound ligand lies close to the surface of the protein.

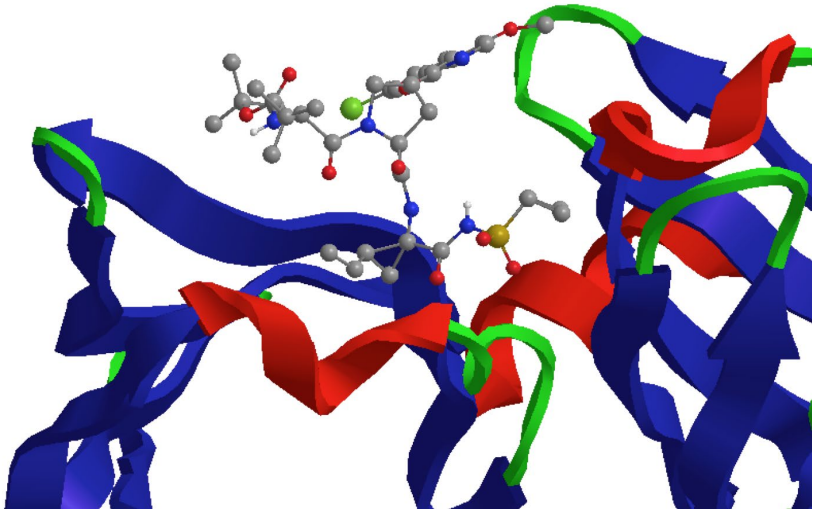
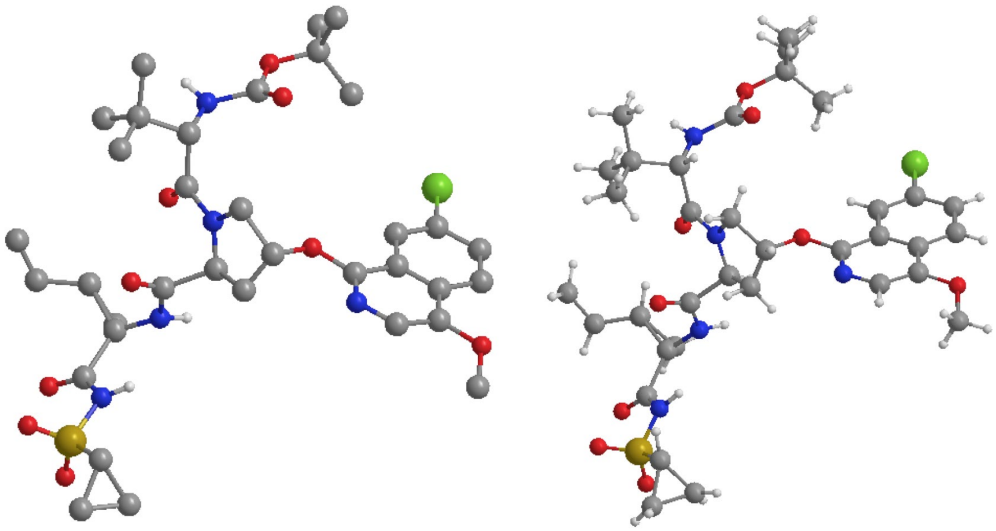


Figure 13 Asunaprevir bound to the binding site.

*In the **Explorer Menu**, *select* **Ligand-992**. The ligand becomes highlighted in the main window.
*From the **Edit** menu, *choose* **Copy**.
*From the **File** menu, *choose* **New** to open a new window.
*From the **Edit** menu, *choose* **Paste** to paste the active conformation of asunaprevir into the new window (Fig. 14).
*From the **View** menu, *choose* **Model Display**, then **Show Hydrogen Atoms**. *Choose* **Show Polar**. This corresponds to the image shown in figure 14A.
*From the **View** menu, *choose* **Model Display**, then **Show Hydrogen Atoms**. *Choose* **Show All**. This corresponds to the image shown in figure 14B.
*From the **View** menu, *choose* **Model Display**, then **Show Hydrogen Atoms**. Choose **Hide**. This corresponds to the image shown in figure 15B.

Note: Hydrogen atoms are included in the structure of the copied ligand. This is not always the case.



A B

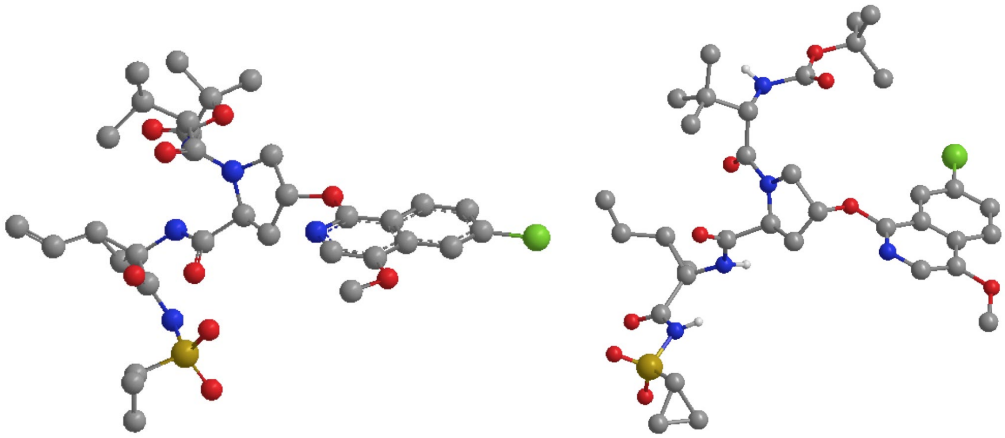
Figure 14 The active conformation with A) polar hydrogen atoms visible and B) with all hydrogens visible .

4. Compare the active conformation of asunaprevir with the energy-minimised conformataion created in Part A.

*_Copy_ and _paste_ both conformations into a new window.

*_Translate_ and _rotate_ the structures such that they are side by side in similar orientations (Fig. 15). One structure can be moved without moving the other if it is first _selected_, then _rotated / translated_ with the shift key depressed.

Both conformations are relatively 'open' and extended. However, they are not identical. The conformations in figure 15 were manipulated such that the central 5-membered ring was orientated in the same way. However, the substituents from this ring are clearly orientated differently for the two conformations.



A **B**

Figure 15 Comparison of A) energy-minimised structure with B) active conformation.

PART C Creation of a model binding site
To assess binding interactions, it is best to extract the ligand with the closest amino acid residues and paste the selection into a separate window.

1. Select asunaprevir along with the closest amino acid residues.
*Return to the window containing the crystal structure.
Choose* the **select tool .
**Zoom* into the structure of asunaprevir.
Double click* on any of asunaprevir's atoms to select the whole molecule. Alternatively, *click* on the entry for asunaprevir in the **Model Explorer table (**Ligand-992**).
**Position* the mouse cursor over the selected feature and *right click* the mouse to reveal a menu.
Choose* **Select, then *choose* **Select groups within Distance of Selection**. *Select* **4 Angstrom.** The ligand and the selected amino acids are now highlighted (Fig. 16).

2. Copy and paste the selection into a new window.
From the **Edit menu, *choose* **Copy.**
*From the **File** menu, *choose* **New** to open a new window.
*From the **Edit** menu, *choose* **Paste** to paste the selection into the new window (Fig. 17).

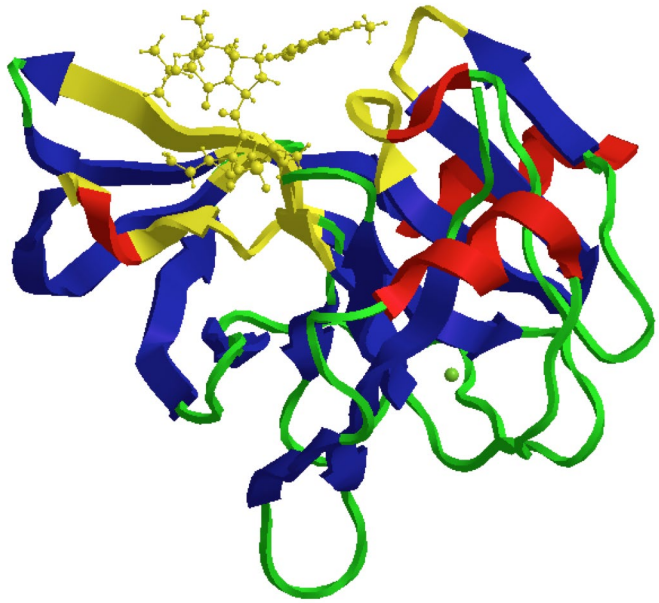


Figure 16 Selection of asunaprevir and amino acid residues within 4Å.

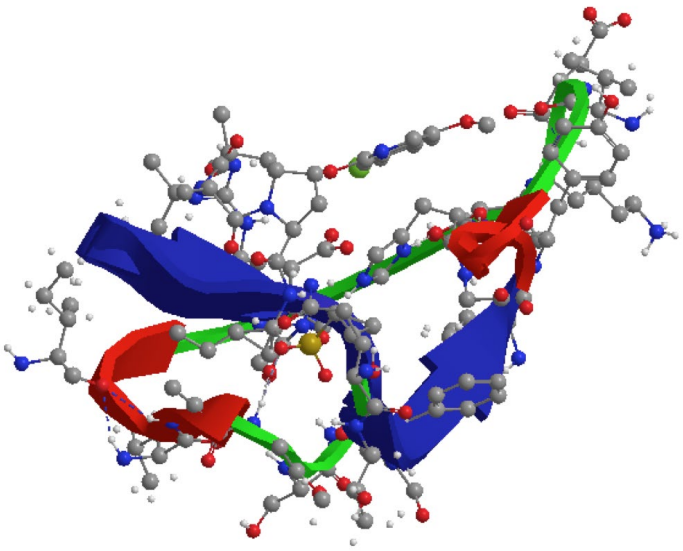


Figure 17 Initial pasted selection of asunaprevir and surrounding amino acid residues.

3. Modify the model binding site to distinguish asunaprevir from the amino acid residues.
The selection can be tidied up as follows.
*From the **View** menu *choose* **Model Display**, then **Display Mode.** *Choose* **Sticks**.
*To alter the ligand back to ball and stick, *select* the ligand by finding it in the **Model Explorer** Menu. Alternatively, *double click* on any of the atoms belonging to the ligand.
*Make sure that the mouse is placed over the selection and r*ight click* the mouse to produce a menu.
Choose* **Display Mode, then **Ball and Stick.**
*From the **View** menu, *choose* **Model Display**, then **Show Hydrogen Atoms**.
Select **Show Polar.**
*From the **View** menu, *choose* **Hydrogen Bonds**, then *select* **Show Intermolecular**.
*In the **Model Explorer** table, *click* on **Solvent**, then *right click* the mouse to reveal a menu. *Choose* **Cut.** This removes water molecules and also allows the structure to be rotated and translated more easily.

The resulting model should now distinguish between the ligand and the surrounding amino acid residues (Fig. 18). Note that hydrogen atoms are visible on heteroatoms and intermolecular hydrogen bonds are indicated by dashed lines.

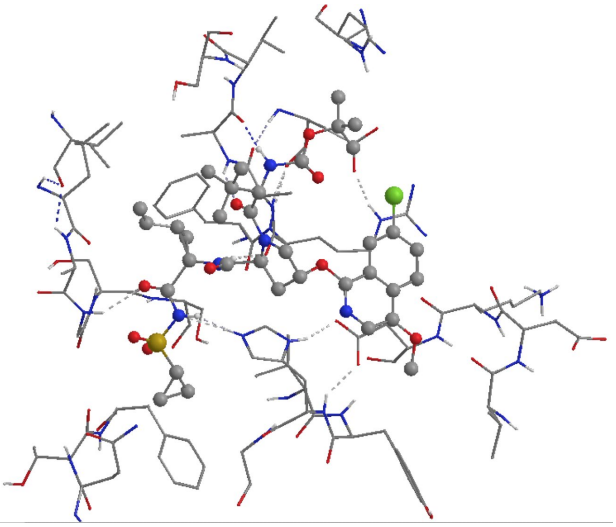


Figure 18 Asunaprevir in the binding site of the protease enzyme.

4. Identify hydrogen bonds involving the acyl sulphonamide region of asunaprevir.
It is useful to use perspective when identifying binding interactions.
*From the **View** menu, *choose* **Model Display**, then **Perspective**.
**Zoom* into the relevant region of the binding site to identify the binding interactions.
*With the shift key depressed, *click* on the HBD and the HBA of a relevant H-bond.
* From the **Structure** menu, *choose* **Measurements**, then **Display Distance Measurement**.

The following image (Fig. 19) shows one hydrogen bond measuring 2.2Å from the carbonyl oxygen (HBA) of the acyl sulphonamide group, as well as two hydrogen bonds measuring 2.2Å and 2.3Å from the NH (HBA and HBD) moiety of the same group.

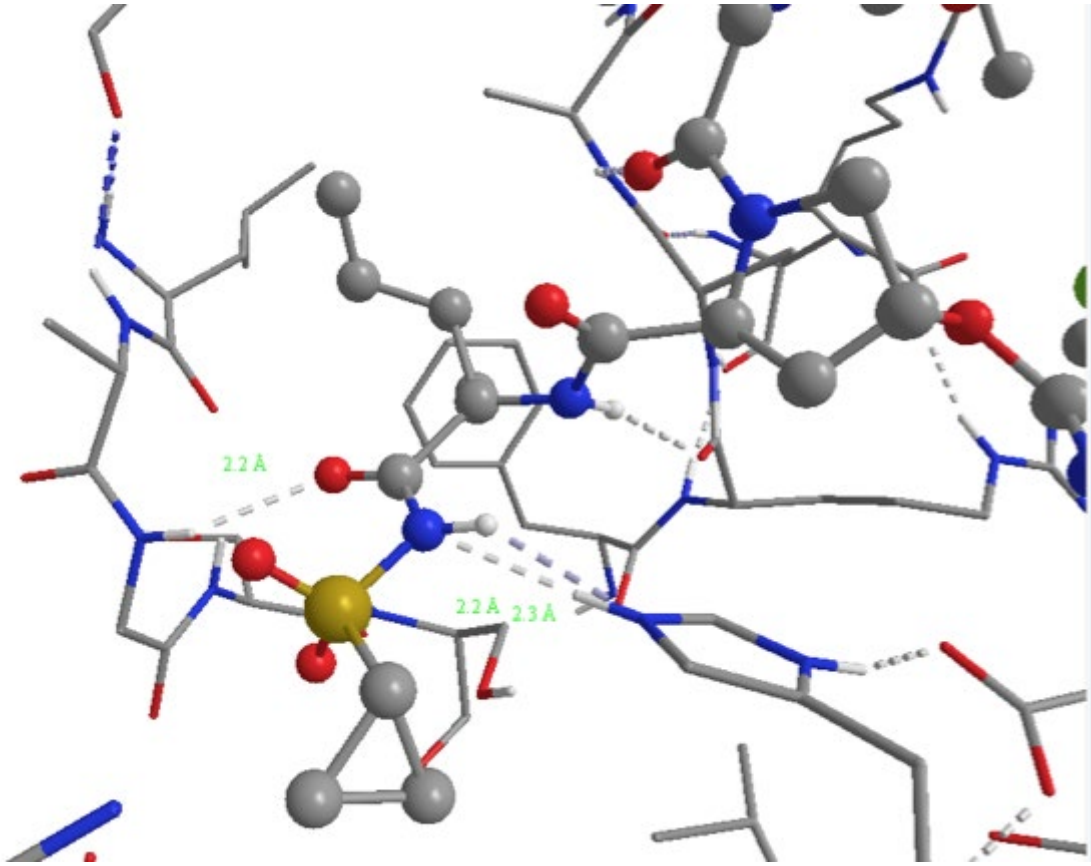


Figure 19 Hydrogen bonds identified for the sulphonamide carbonyl region.

The amino acid residues can be identified by displaying residue labels.
*From the **View** menu *choose* **Model Display** and **Show Residue Labels** (Fig. 20).

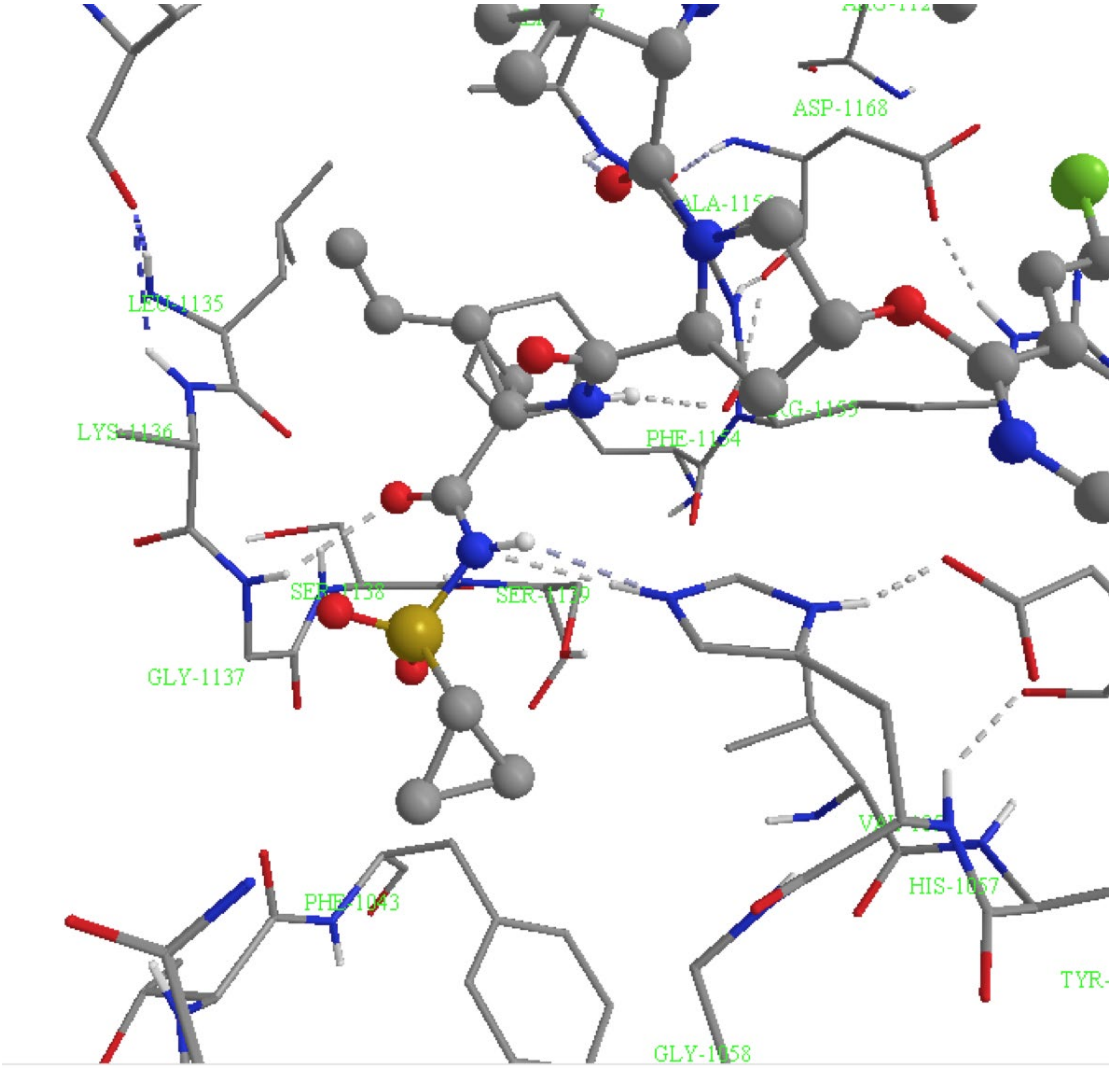


Figure 20 Identification of amino acid residues.

The two hydrogen bonds from the acyl sulphonamide NH involve the imidazole ring of a histidine residue (His-1057). The hydrogen bond from the carbonyl oxygen involves the NH proton of glycine-1137. This is part of the peptide bond between Gly-1137 and Lys-1136.
Note that the Pose View identifies interactions from the sulphonamide oxygen atoms to the side chain hydroxyl group of Ser-1139 and the peptide NH proton of Gly-1137. The peptide bond concerned is between Gly-1137 and Lys-1136 (Fig. 22). The lengths of these H-bonds can be measured (Figs. 21).

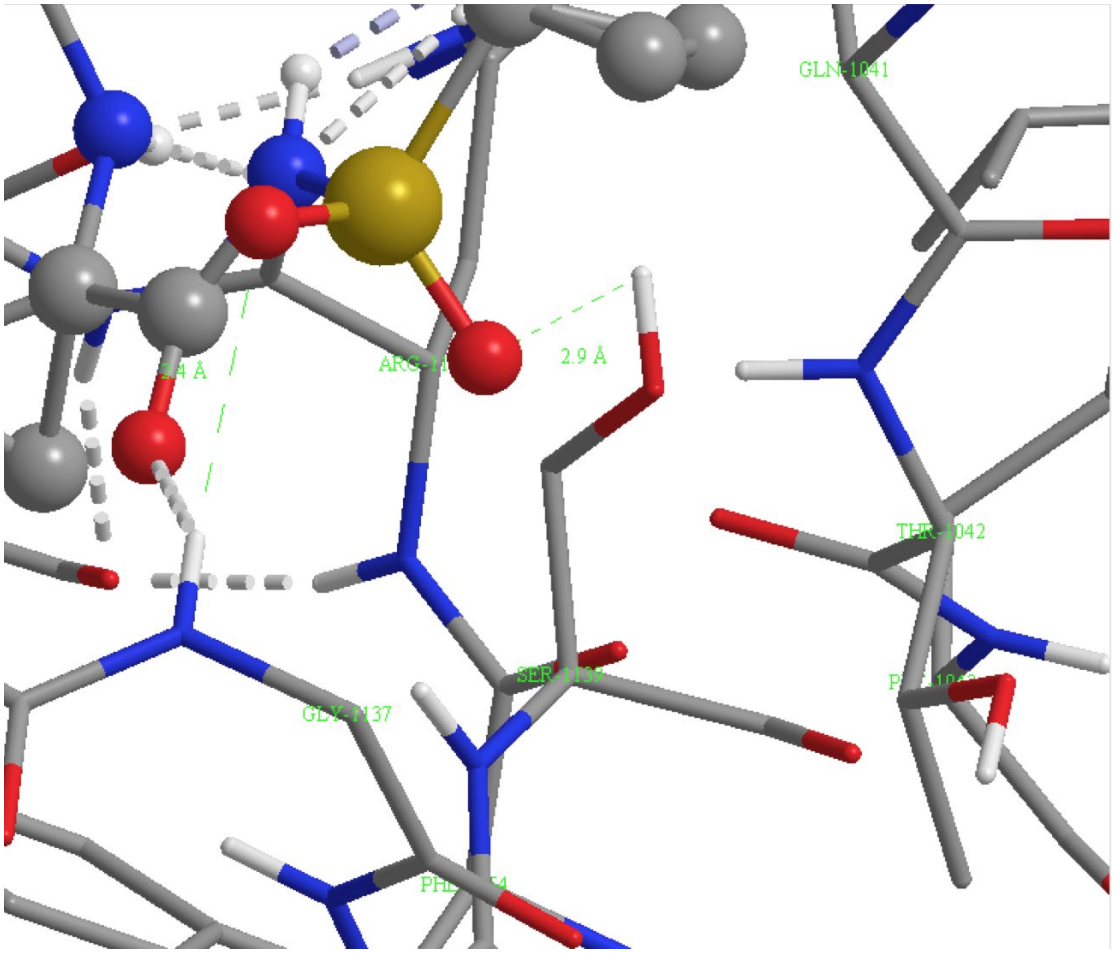


Figure 21 Possible interactions with Ser-1139 (2.9Å) and Gly-1137 (2.4Å).

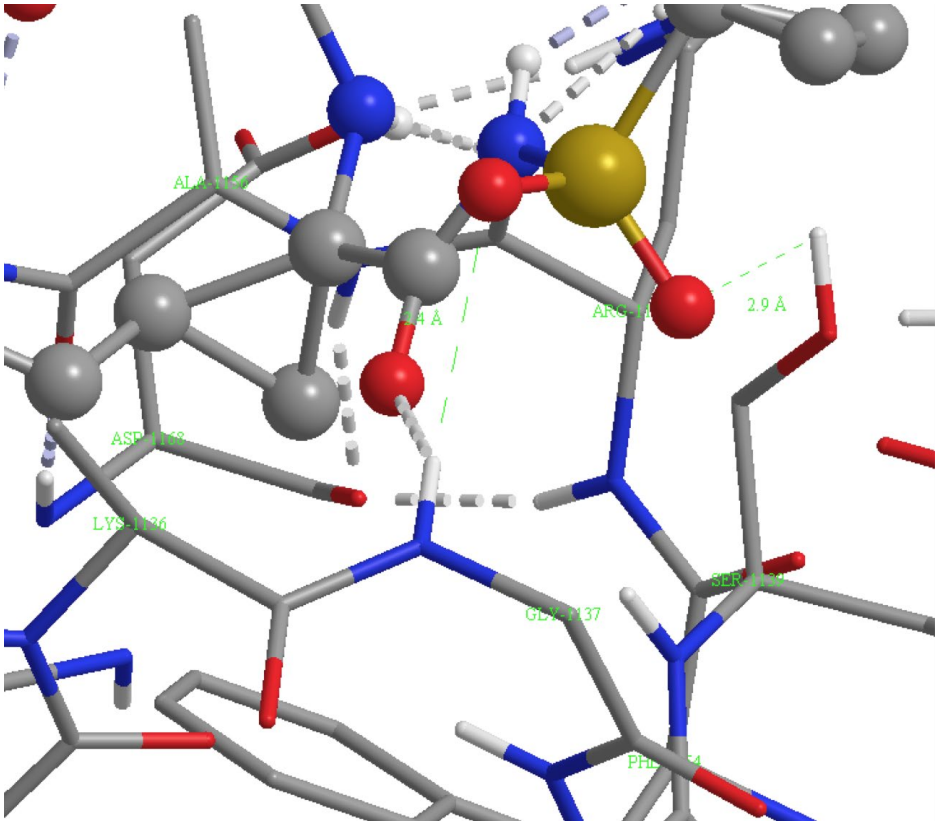


Figure 22 Identification of Ser-1139, Gly-1137 and Lys-1136.

5. Identify the interaction involving the amide group of asunaprevir

An interaction measuring 2.1Å is observed between the secondary amide NH proton acting as an HBD with the peptide carbonyl oxygen of Arg-1155 (Figs. 22 and 23). The carbonyl oxygen of the tertiary amide is interacting with the peptide NH proton of Ala1157 via a H-bond measuring 2.1Å in length (Figs. 23 and 24). The peptide bonds involved in these interactions are between Arg-1155 and Ala1156, and between Ala-1157 and Ala-1156.

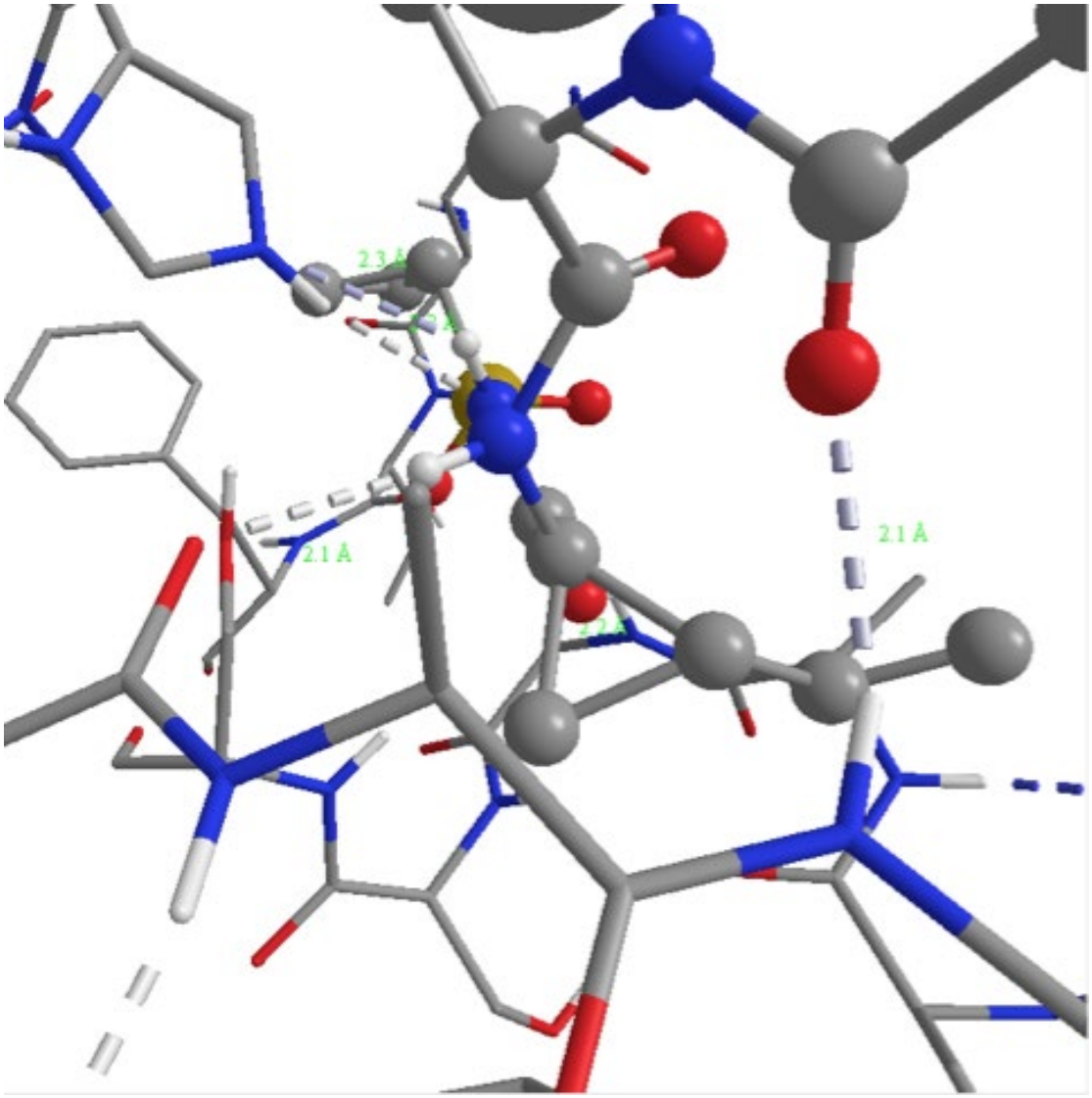


Figure 23 Hydrogen bond interactions involving the amide groups.

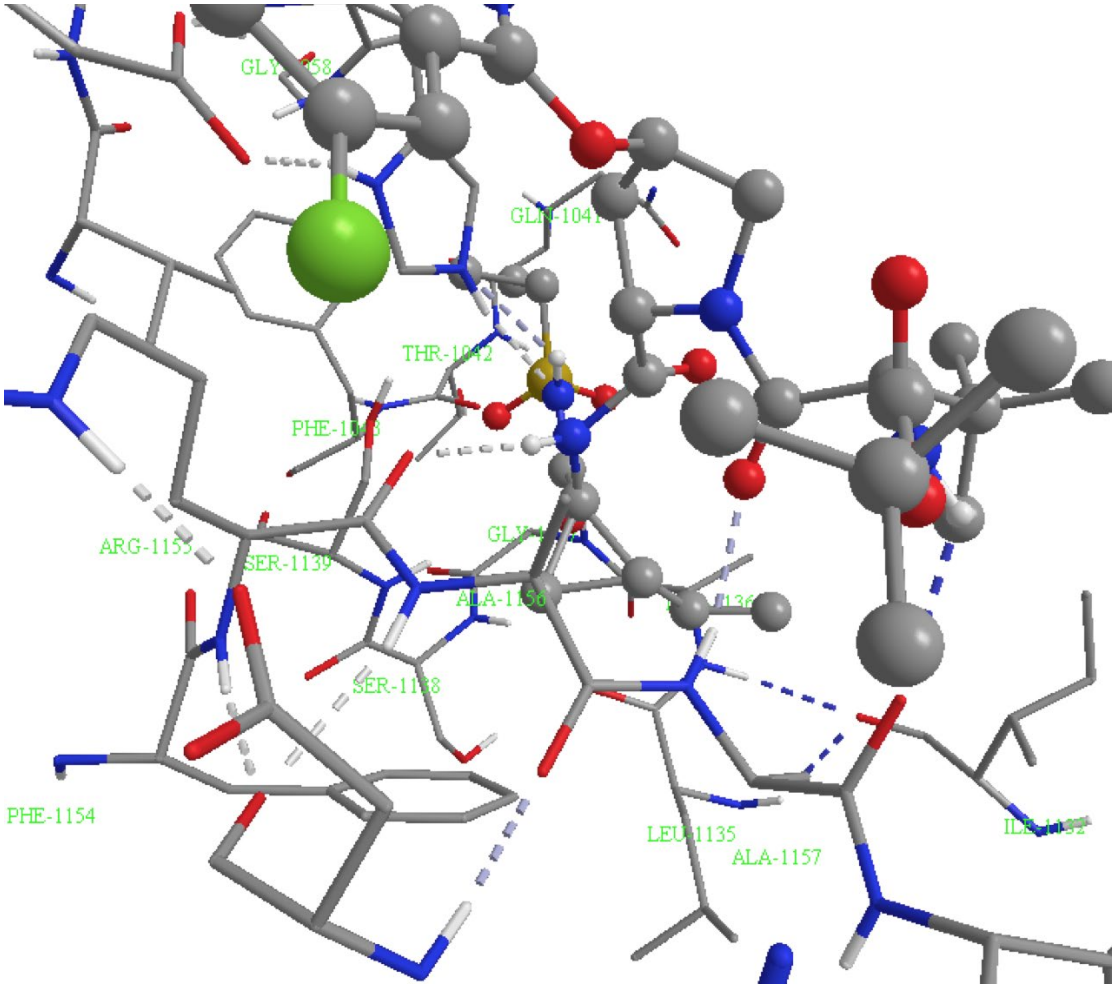


Figure 24 Interactions with Arg1155 and Ala1157

7. Identify the H-bond interaction involving the urethane group
This interaction measures 2.2Å and involves the NH proton of the urethane group acting as an HBD. It is interacting with the peptide carbonyl oxygen of Ala-1157 (Fig. 25 & 26). The peptide bond is between Ala-1157 and Val-1158.

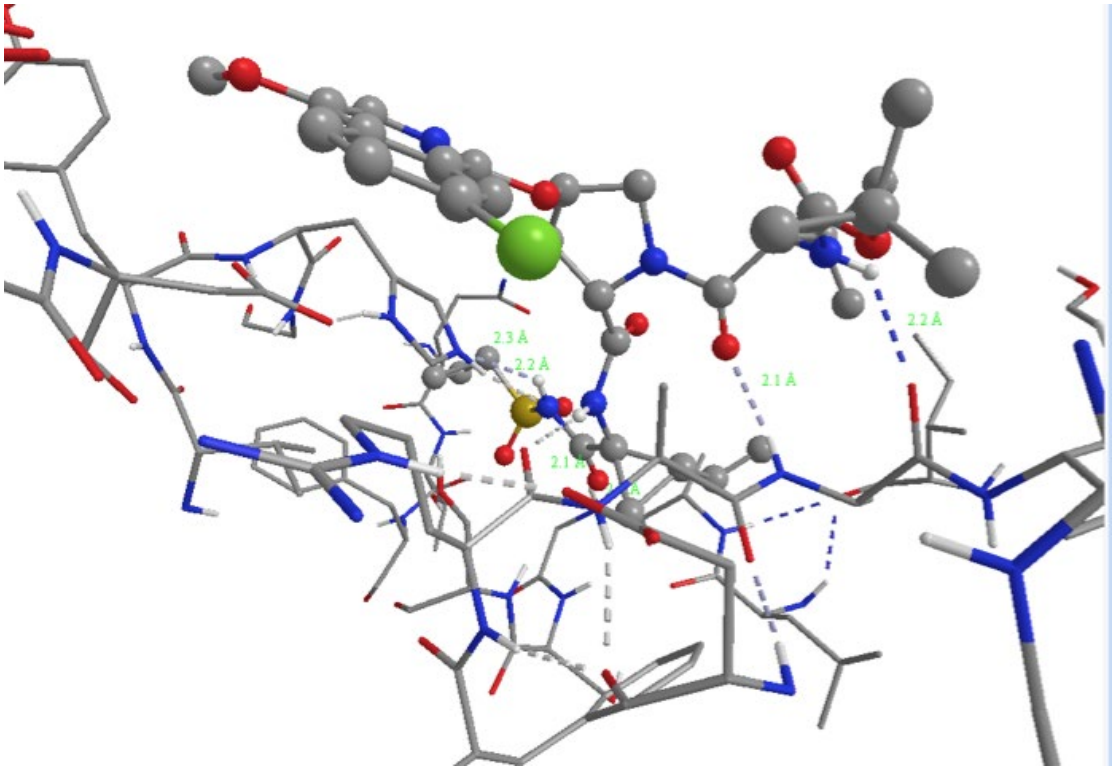


Figure 25 Interaction involving the urethane group (on the right).

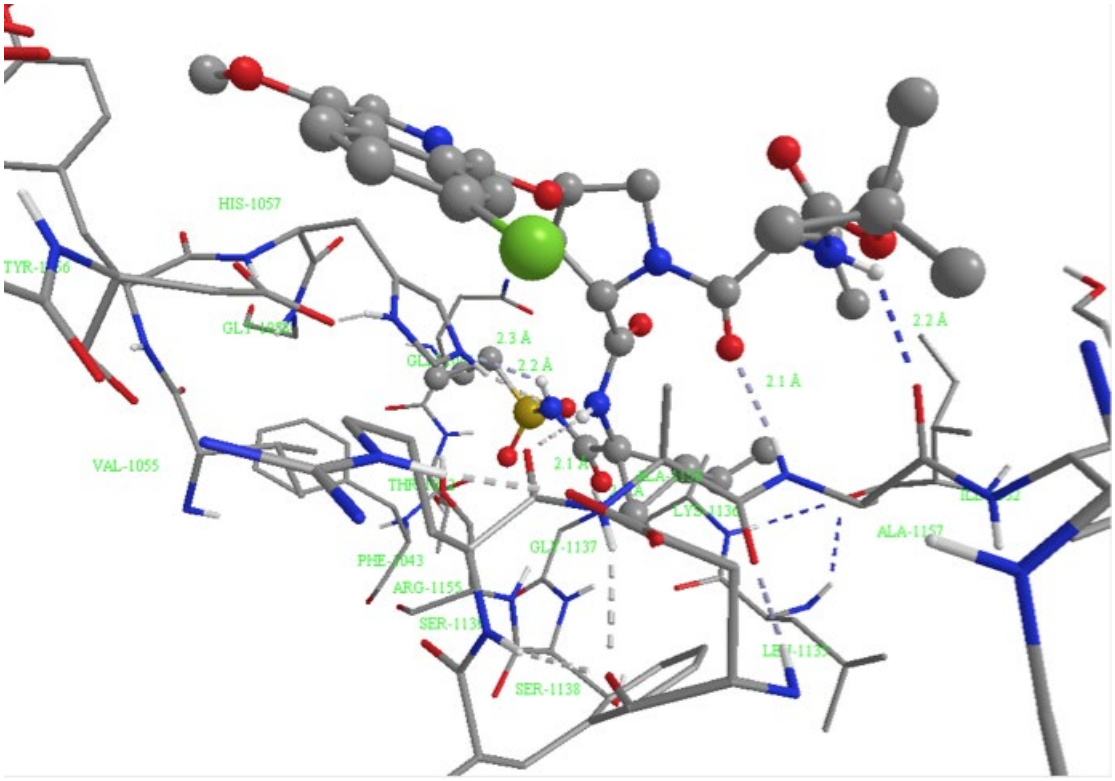


Figure 26 Interaction involving the peptide bond between Ala-1157 and Val 1158