

Structure-Based Design and QSAR using GRID Molecular Interaction Fields - Homework

Nicolás Rojo Esteban



1 3KRR

The structure with PDB code 3KRR corresponds to the kinase domain of Janus kinase 2 (JAK2) in complex with a molecule inhibitor bound at the ATP binding site. JAK2 is a member of the Janus kinase family and plays a central role in cytokine mediated signal transduction pathways, where it regulates cell proliferation, differentiation, and survival by activating STAT transcription factors.

Dysregulation of JAK2, most notably through the V617F mutation, leads to constitutive signaling independent of cytokine binding. This persistent activation results in uncontrolled proliferation of hematopoietic cells and is strongly associated with myeloproliferative neoplasms, particularly polycythemia vera, as well as essential thrombocythemia and primary myelofibrosis [1]. In polycythemia vera, excessive production of red blood cells increases blood viscosity, which can cause severe clinical complications such as thrombosis, hypertension, and increased risk of stroke. Due to its direct involvement in the molecular origin of these diseases, JAK2 has emerged as a key therapeutic target, and the development of potent and selective JAK2 inhibitors represents an important strategy for the treatment of myeloproliferative disorders.

The bound ligand in this structure is the quinoxaline-based inhibitor:

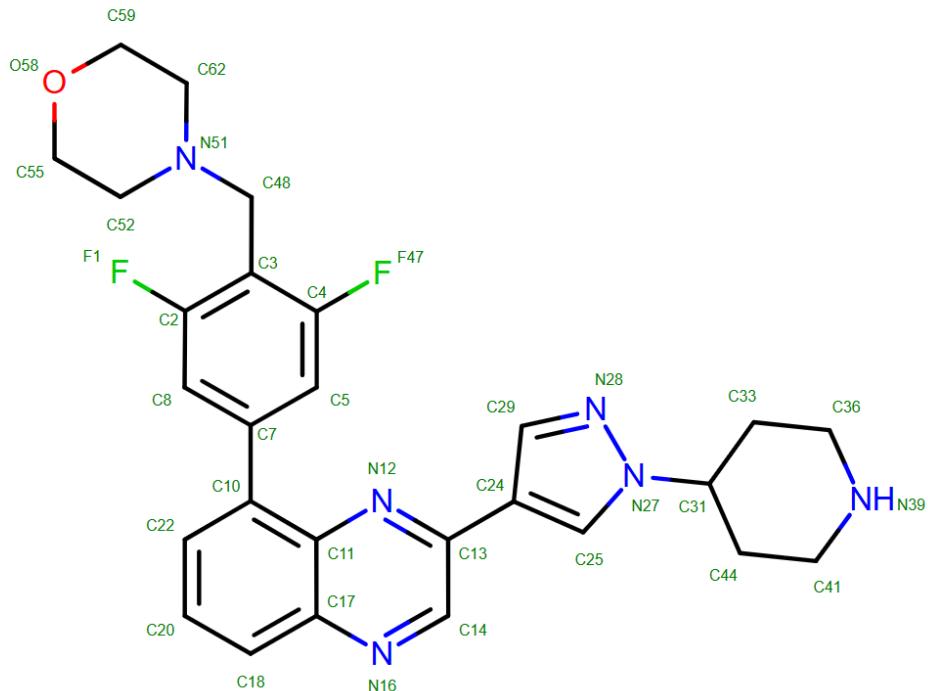


Figure 1: 2D representation of the ligands structure

8-[3,5-difluoro-4-(morpholin-4-ylmethyl)phenyl]-2-(1-piperidin-4-yl-1H-pyrazol-4-yl)quinoxaline

This compound corresponds to NVP-BSK805, a potent and selective ATP-competitive JAK2 inhibitor reported in the associated publication. The structure was determined by

X-ray diffraction at a resolution of 1.80 Å, providing sufficient structural detail for structure-based analysis using GRID and FLAP.

The paper entitled “*Potent and Selective Inhibition of Polycythemia by the Quinoxaline JAK2 Inhibitor NVP-BSK805*” describes the discovery and pharmacological characterization of NVP-BSK805 as a selective JAK2 inhibitor. Biochemical assays demonstrate that the compound acts through an ATP-competitive mechanism.

In cellular systems driven by constitutively active JAK2(V617F), NVP-BSK805 inhibits downstream signaling, notably reducing STAT5 phosphorylation. This leads to inhibition of cell proliferation and induction of apoptosis. In vivo, the compound shows good oral bioavailability and a long half-life, and it is effective in multiple disease-relevant models, including a Ba/F3 JAK2(V617F) leukemia model and erythropoietin-induced polycythemia models in rodents. Overall, the study highlights NVP-BSK805 as a well-balanced example of a potent, selective, and pharmacologically viable JAK2 inhibitor.

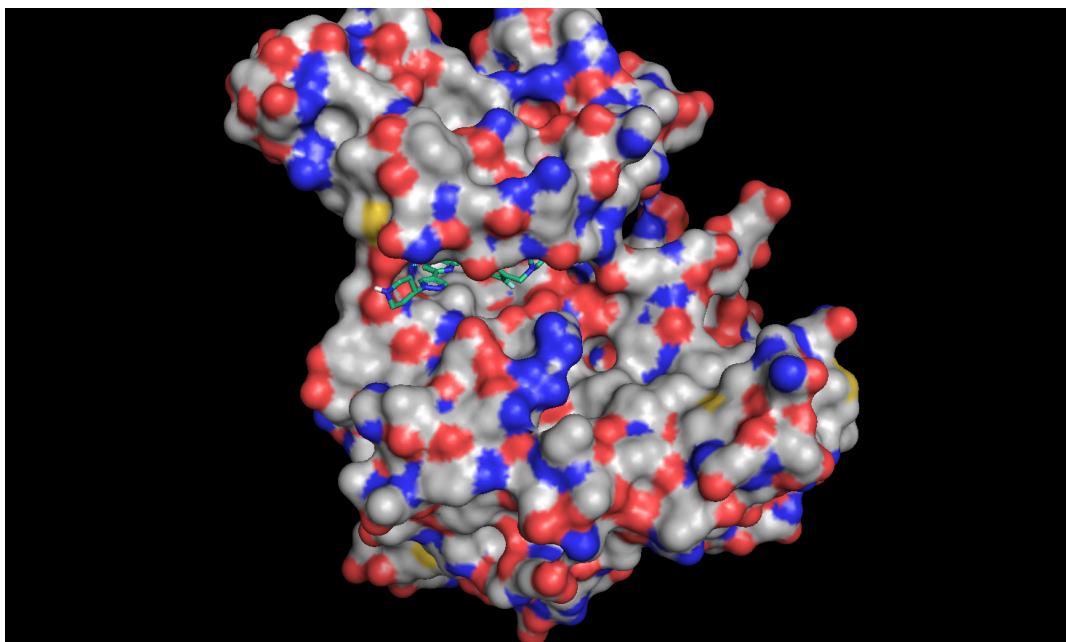


Figure 2: Overall surface representation of JAK2 (PDB: 3KRR). The bound inhibitor is shown within the ATP-binding pocket.

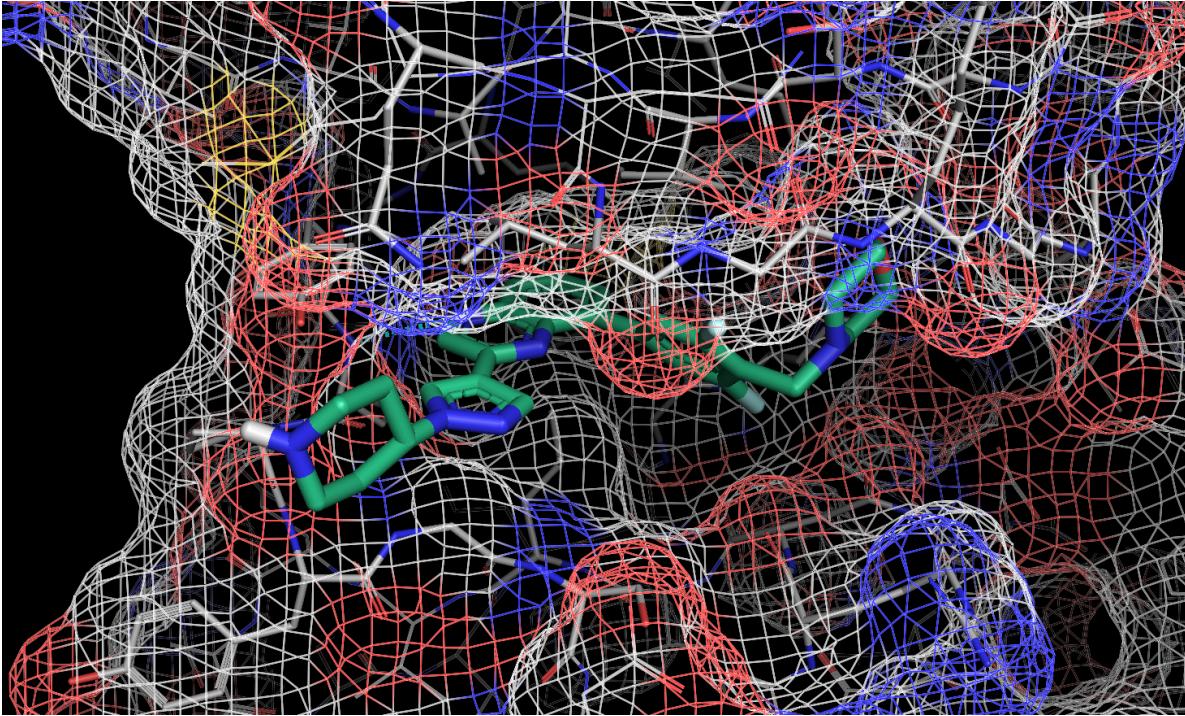


Figure 3: Zoomed view of the ATP-binding pocket showing the bound quinoxaline inhibitor.

The structure was imported into GRID 2025 in PDB format. Standard PDB fixing procedures were applied, while crystallographic water molecules were retained.

Visual inspection of the PDB structure reveals the kinase domain with the ligand deeply inserted inside the ATP-binding pocket. The quinoxaline is oriented toward the hinge region of JAK.

GRID protein MIFs and key ligand interactions

Protein GRID Molecular Interaction Fields (MIFs) were computed for the 3KRR complex using a pocket definition centered on the bound ligand. In this analysis, three probes were selected in order to capture the main interaction types expected in an ATP-binding site: (i) C1= (aromatic carbon probe) to highlight aromatic hotspots, (ii) O (carbonyl oxygen probe) to identify regions favourable for hydrogen-bond acceptors on the ligand, and (iii) N1 (amide nitrogen probe) to identify regions favourable for hydrogen-bond donors on the ligand. After computation, the MIFs are displayed as isocontours around the binding site. In GRID, the energy threshold of each isocontour can be adjusted (more negative values correspond to more favourable interactions). In this work the MIFs were visualised in wireframe mode.

Figure 4 shows the three protein MIFs (C1=, O and N1) overlaid with the bound inhibitor. Overall, the ligand exhibits good complementarity with the pocket: most of its scaffold is embedded within favourable regions of the fields. The aromatic ring nitrogen N16 forms a plausible hydrogen bond with the backbone NH of Leu932 ($\text{N}\cdots\text{N}$ distance $\approx 3.1 \text{ \AA}$). This interaction is consistent with a typical kinase binding mode, where hinge backbone groups provide directional hydrogen bonding that helps lock the inhibitor in place.

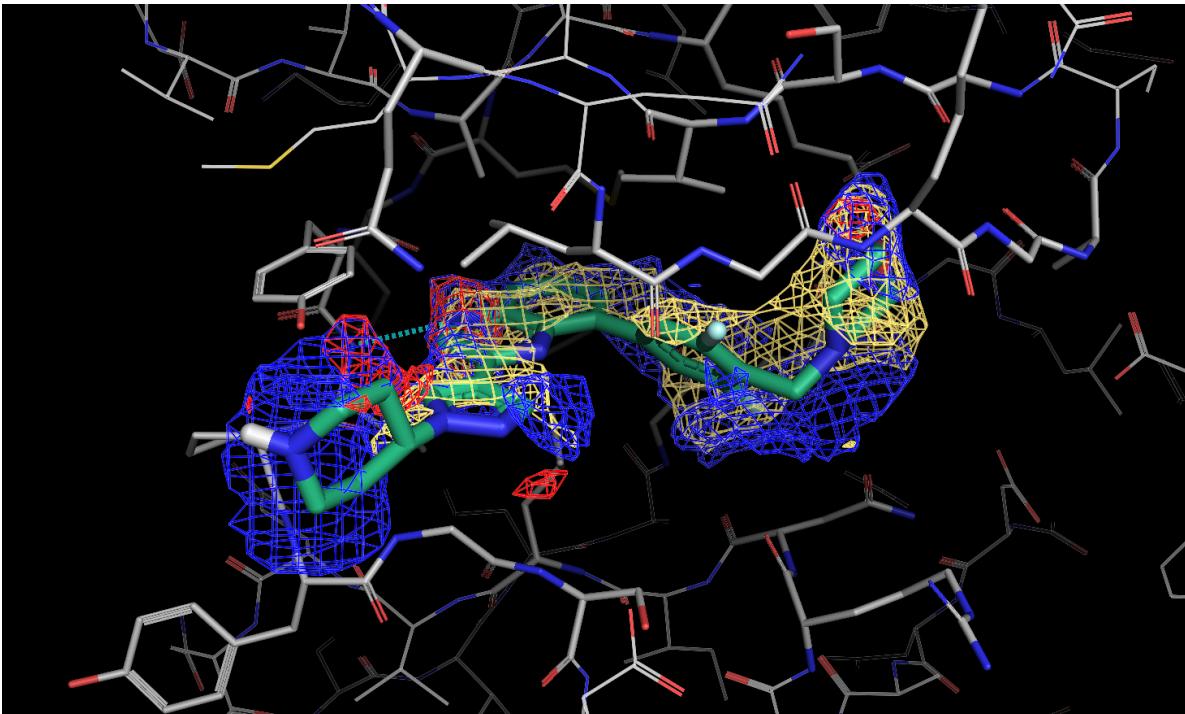


Figure 4: Protein GRID MIFs computed for the 3KRR complex using the C1= (aromatic carbon), O (carbonyl oxygen) and N1 (amide nitrogen) probes and displayed as wireframe isocontours around the bound ligand. The dashed line highlights the hinge anchoring interaction involving the ligand ring nitrogen (N16) and the backbone NH of Leu932 ($N \cdots N \approx 3.1 \text{ \AA}$).

In addition to the hinge anchor, the distribution of C1= hotspots around aromatic regions of the ligand suggests that lipophilic packing plays an important role in stabilising the complex. Likewise, the presence of O and N1 hotspots near polar functionalities indicates opportunities for directional hydrogen bonding. Regions where the ligand overlaps with these GRID hotspots highlight the key interactions already exploited. Unlike ligand parts that appear partially exposed or less optimally aligned with hotspots, ligand parts that can be considered as design places for subsequent optimization to improve binding or tune physicochemical properties. However, inspection of the protein MIFs indicated that the inhibitor already shows strong complementarity with the ATP-binding pocket, with most of the ligand occupying favourable regions of the C1=, O and N1 fields. In particular, the ligand is anchored at the hinge region through an interaction involving the aromatic ring nitrogen (N16) and the hinge backbone of **Leu932**, consistent with a typical kinase binding mode.

Ligand modification: hydroxyl addition

Despite this good initial fit, GRID Generate Ideas was used to explore whether a small local modification could further improve binding complementarity.

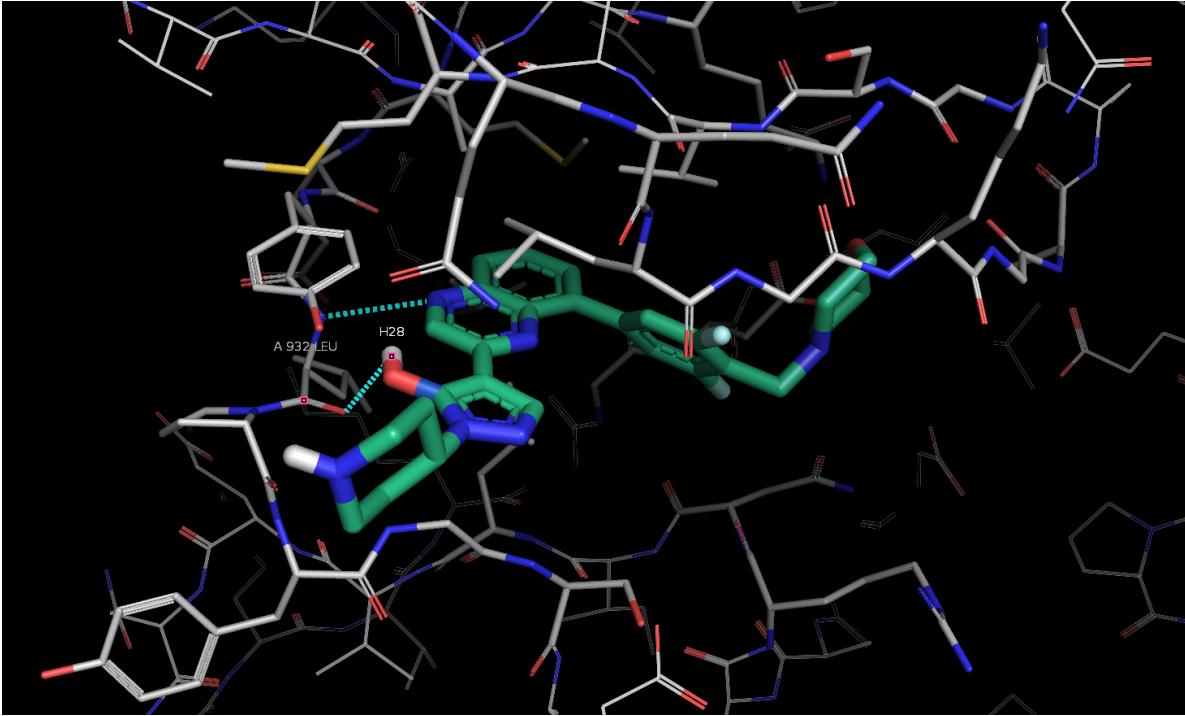


Figure 5: Ligand modification proposed by GRID *Generate Ideas*: addition of a hydroxyl group at C25. The new OH group forms an additional hydrogen bond to the hinge backbone carbonyl oxygen of Leu932 ($\text{H}\cdots\text{O} \approx 2.8 \text{ \AA}$), reinforcing hinge anchoring and potentially increasing complex stability.

The suggested modification consisted of introducing a hydroxyl (OH) group at C25. This change creates an additional hydrogen-bond interaction, the hydroxyl hydrogen acts as a donor toward the backbone carbonyl oxygen of Leu932. The measured $\text{H}\cdots\text{O}$ distance is approximately 2.8 \AA , which is consistent with a strong hydrogen bond.

This newly introduced interaction strengthens the hinge anchoring and may increase the overall stability of the protein–ligand complex by adding an extra polar contact without disrupting the existing binding pose.

1.1 FLAP 2 Water Analysis for 3KRR

To further analyse the binding site of JAK2 in the 3KRR structure, FLAP 2 was used to predict the pseudo-apo water network around the ligand. The analysis was performed considering a region of approximately 8 \AA around the bound inhibitor, following the recommended protocol described in the TCCM tutorials. Two complementary analyses were carried out: water prediction and water perturbation analysis.

In the water prediction step, FLAP places water molecules in the binding pocket according to the energetic favourability of water–protein interactions in the absence of the ligand. The predicted waters are classified into three categories: favourable (blue), neutral (yellow), and unfavourable (red).

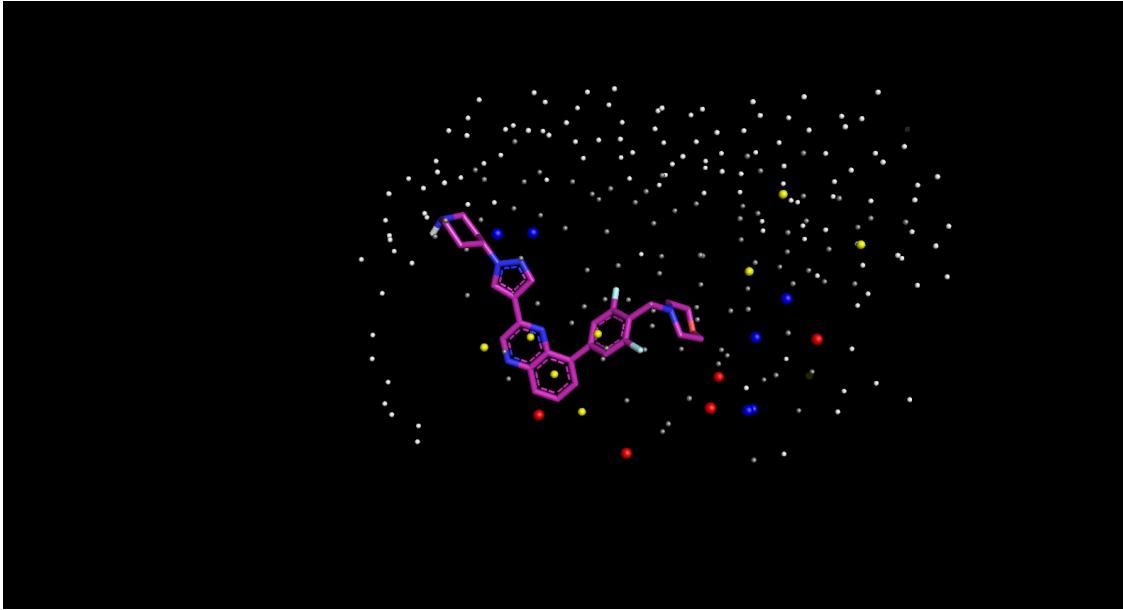


Figure 6: Pseudo-apo water network predicted by FLAP 2 for the 3KRR complex within an 8 Å radius around the ligand. Waters are classified as happy (blue), unhappy (yellow), very unhappy (red) and solvent-like (grey).

For the 3KRR system, the predicted water network contains 6 unfavourable (red) waters, 11 neutral (yellow) waters, and 6 favourable (blue) waters. These waters are mainly located around the ligand core and the surrounding pocket, with a noticeable clustering close to the aromatic and heterocyclic regions of the inhibitor.

The presence of several unfavourable waters close to the ligand suggests that this region of the pocket is unstable for water molecules and is therefore well suited to be occupied by hydrophobic or aromatic ligand moieties.

The water perturbation analysis evaluates how the presence of the ligand affects the predicted water network, identifying which waters are displaced, stabilised, or remain unaffected upon ligand binding. This step is crucial to distinguish between productive and non-productive water displacement.

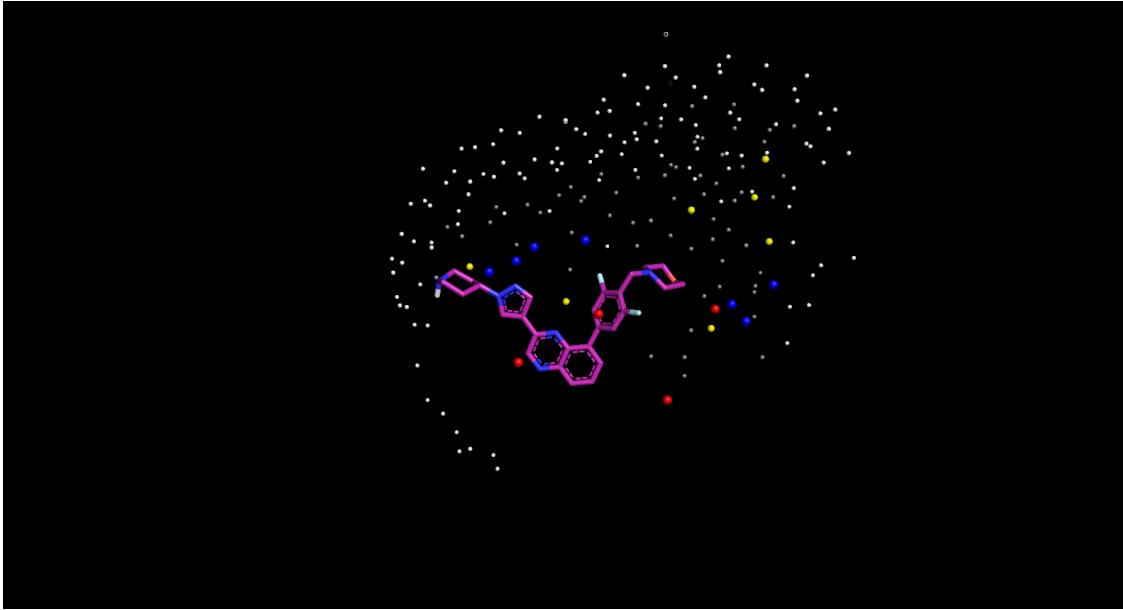


Figure 7: Pseudo-apo water network predicted by FLAP 2 for the 3KRR complex within an 8 Å radius around the ligand. Waters are classified as happy (blue), unhappy (yellow), very unhappy (red) and solvent-like (grey).

In the perturbation analysis for 3KRR there are 4 unfavourable (red), 10 neutral (yellow), and 7 favourable (blue) waters. Compared to the initial prediction, a reduction in the number of unfavourable waters is observed, indicating that the ligand displaces several energetically poor waters upon binding.

This behaviour suggests that the ligand binding is thermodynamically favourable, as it replaces unstable waters. The slight increase in the number of favourable waters after perturbation indicates that some water molecules remain stabilised in the pocket.

Overall, the FLAP water analysis for 3KRR shows that the inhibitor efficiently displaces unfavourable and neutral waters from the binding site, contributing positively to binding affinity.

2 3HS4

PDB entry **3HS4** is the X-ray crystal structure of the human enzyme *carbonic anhydrase II* (hCA II) in complex with the inhibitor *acetazolamide* (AZM), solved at very high resolution (\sim 1.1 Å). [2, 3] Carbonic anhydrase II is a zinc-dependent enzyme that catalyses the reversible conversion of carbon dioxide and water into bicarbonate and a proton. Because this reaction is important for pH control and fluid balance, carbonic anhydrase inhibitors are clinically relevant and have been used as diuretics and in other indications. [2]

In the 3HS4 structure, AZM binds in the catalytic pocket and anchors directly to the active-site Zn^{2+} ion, which is a typical binding mode for sulfonamide carbonic anhydrase inhibitors. [2] Figure 8 shows the overall protein surface with the ligand located inside the active-site cavity, while Figure 9 highlights the pocket environment around AZM.

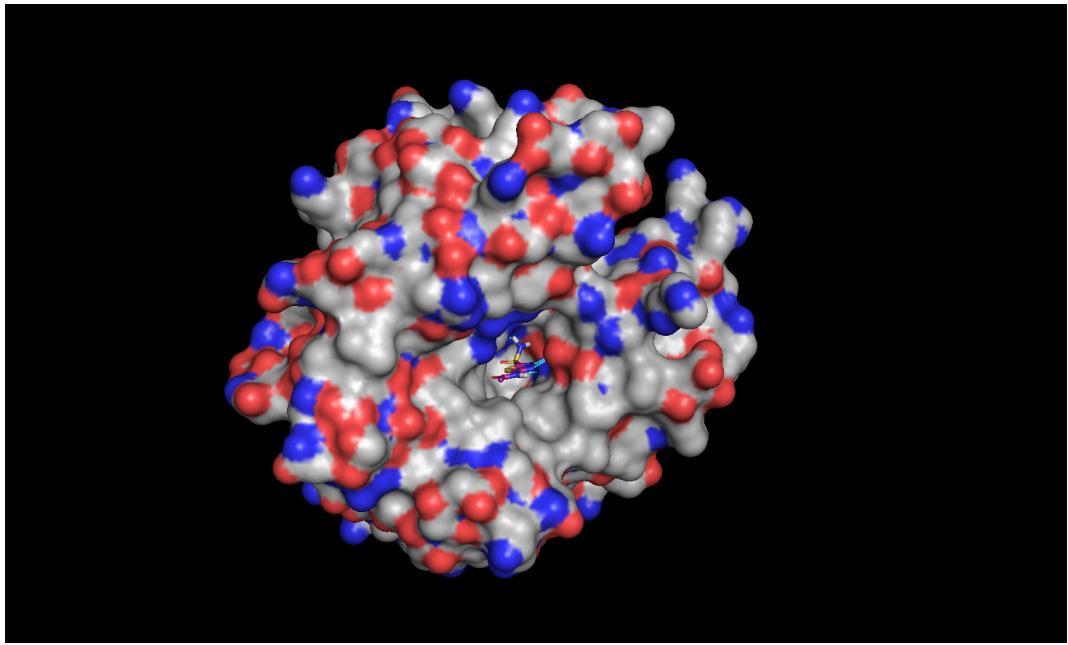


Figure 8: Overall view of human carbonic anhydrase II (3HS4) shown as a surface representation. The inhibitor acetazolamide (AZM) is located in the active-site cavity.

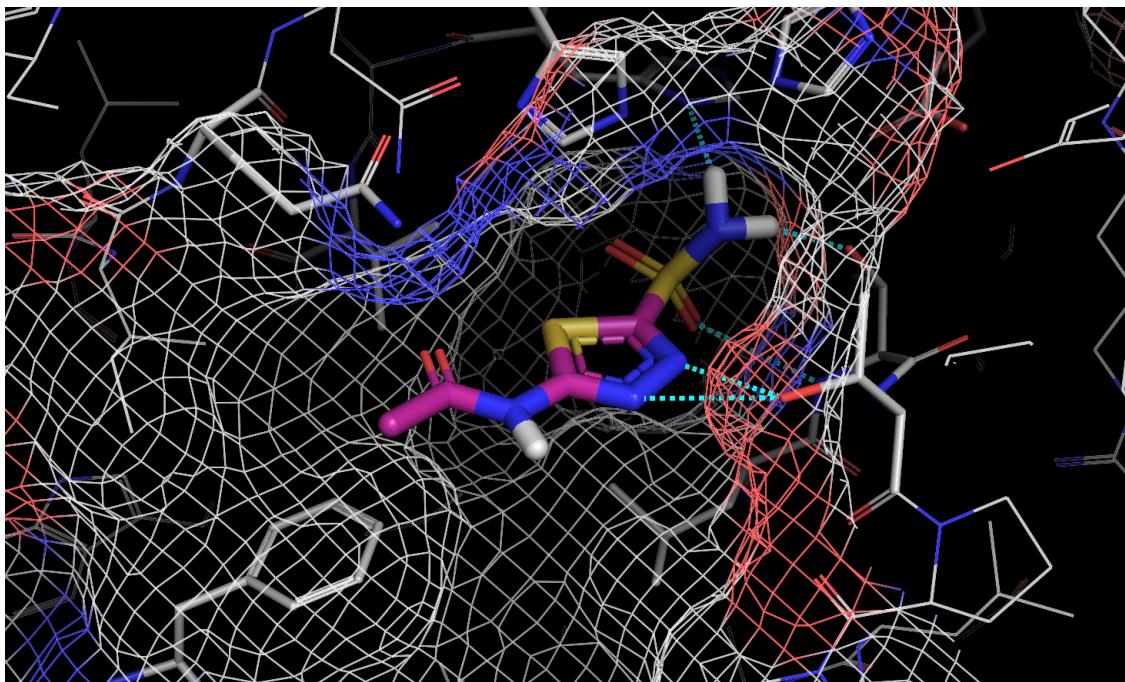


Figure 9: Close-up of the 3HS4 active site showing acetazolamide (AZM) bound in the catalytic pocket near the Zn^{2+} centre. This is the ligand copy selected for subsequent GRID and FLAP analyses.

Acetazolamide (AZM) is a small inhibitor with a sulfonamide group. Its 2D structure is

shown in Figure 10. In the crystal unit of 3HS4, several copies of AZM are present (labelled as different residue numbers in the PDB file). However, only one copy is positioned inside the catalytic pocket of the protein. The additional AZM molecules are located on the protein surface and are most consistent with crystal packing or non-specific contacts rather than true active-site binding.

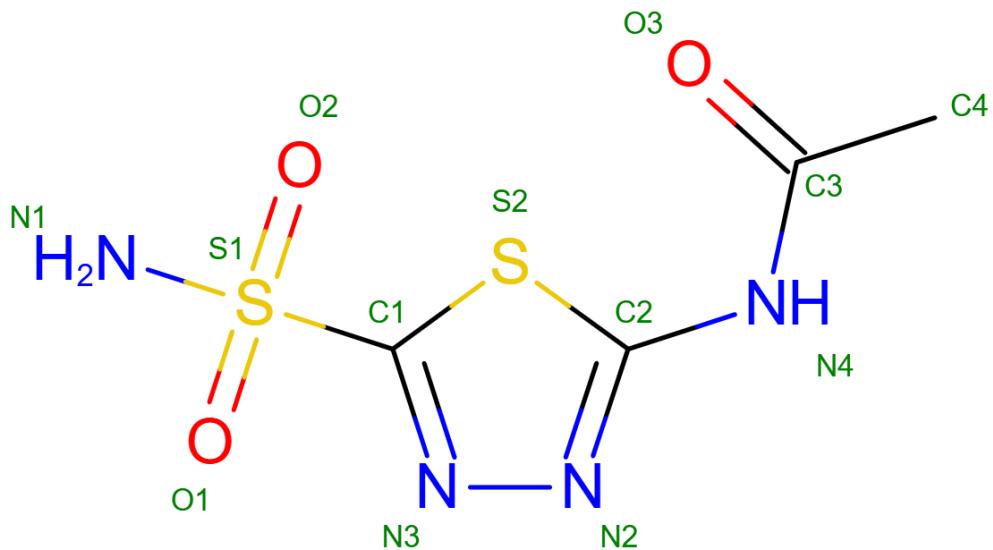


Figure 10: 2D structure of acetazolamide (AZM), the inhibitor co-crystallised with human carbonic anhydrase II in PDB 3HS4.

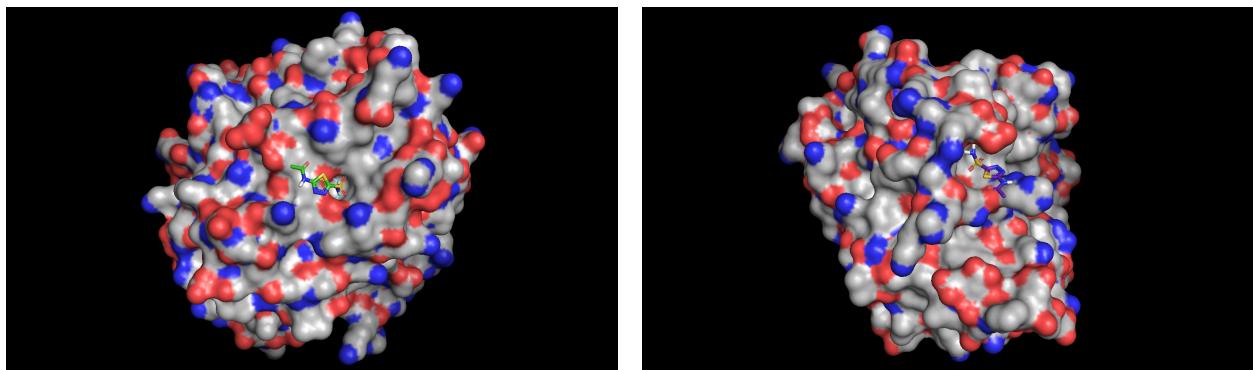


Figure 11: *
 (a) Additional AZM copy on the protein surface

Figure 12: *
 (b) Additional AZM copy on the protein surface

Figure 13: Other acetazolamide (AZM) molecules present in the 3HS4 crystal unit. These copies do not occupy the catalytic pocket and were not used for the binding-site analyses.

GRID protein MIFs and key ligand interactions

Protein molecular interaction fields (MIFs) were calculated using GRID with the C1=, O and N1 probes, focusing on an 8 Å radius around the bound acetazolamide ligand. The resulting MIFs are shown in Figure 14. A clear localisation of favourable interaction regions is observed around the sulfonamide moiety of the ligand, while the opposite side of the molecule shows very limited MIF density.

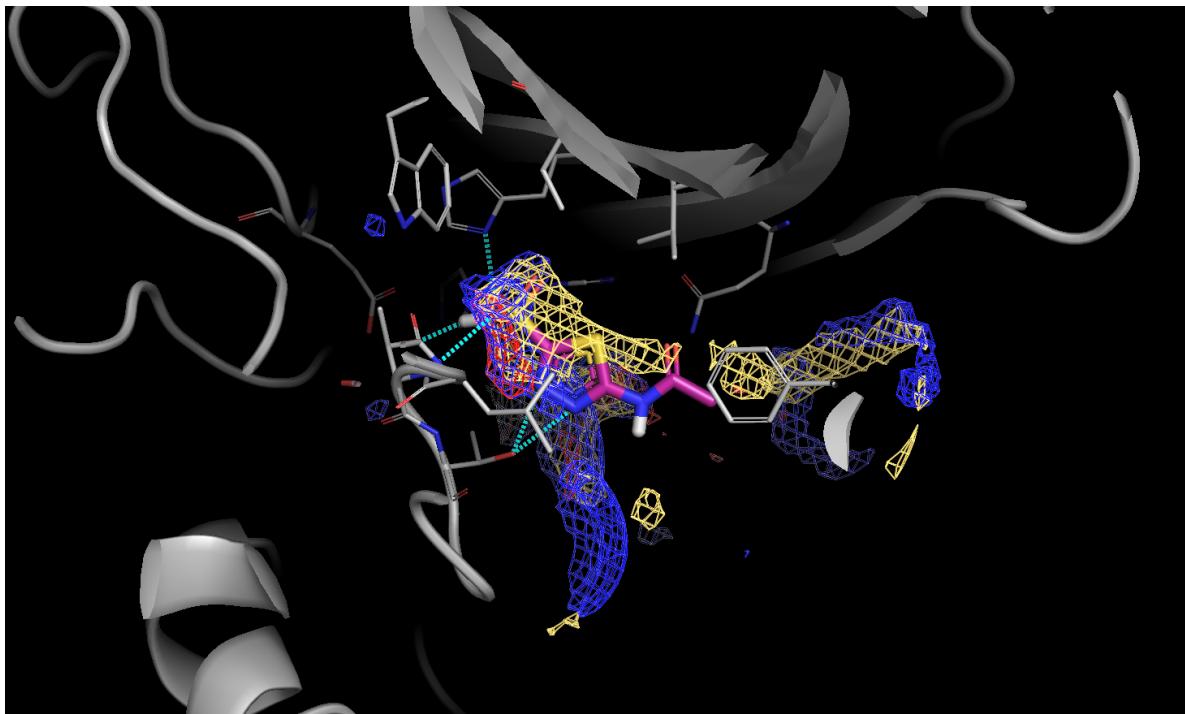


Figure 14: Protein MIFs calculated with the C1= (lipophilic), O (hydrogen-bond acceptor) and N1 (hydrogen-bond donor) probes around acetazolamide in the active site of carbonic anhydrase II (3HS4).

This distribution indicates that the sulfonamide group is the main contributor to the binding affinity in this complex, whereas the remainder of the ligand plays a more secondary role in stabilising the interaction. In particular, the dense clustering of O and N1 MIFs around the sulfonamide highlights the importance of hydrogen-bonding and polar interactions in this region of the pocket.

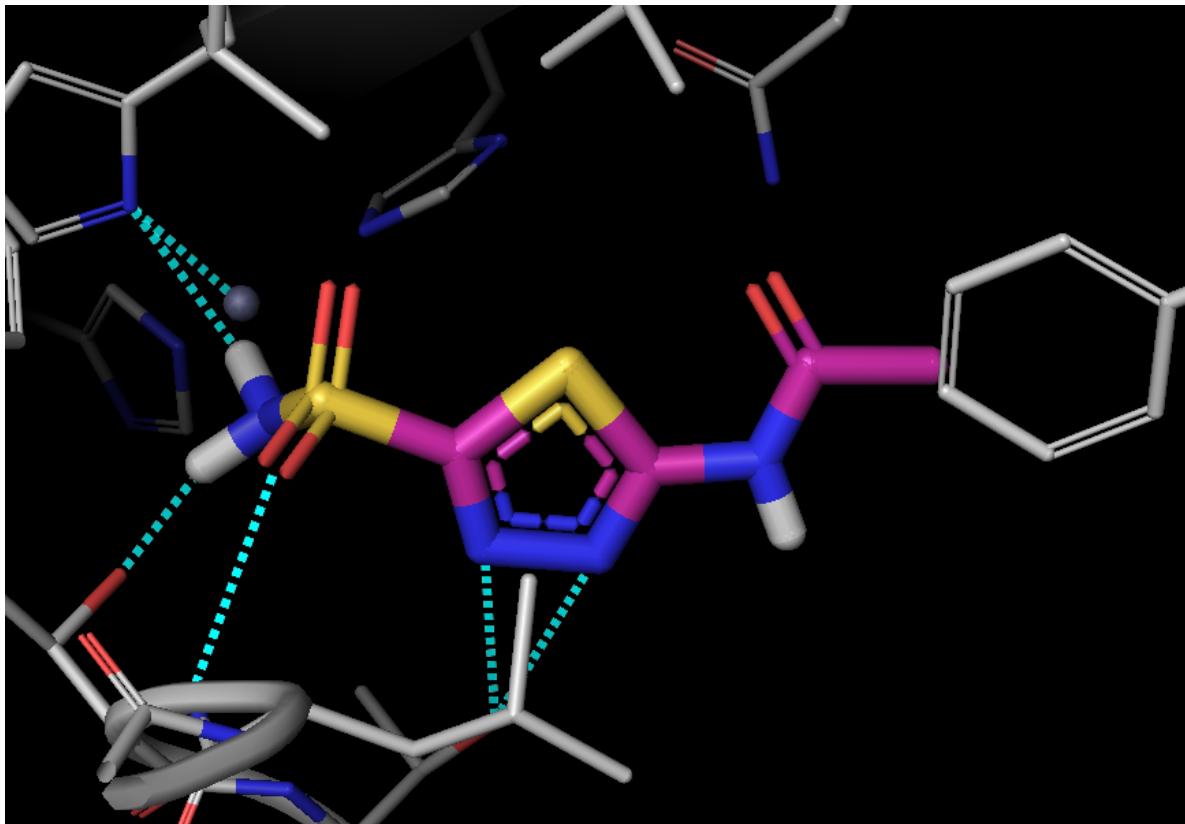


Figure 15: Protein-acetazolamide interactions in the active site of carbonic anhydrase II (3HS4).

Several specific interactions between acetazolamide and the protein active site can be identified. One of the nitrogen atoms of residue A 119 HIS coordinates the catalytic Zn^{2+} ion at a distance of approximately 2.0 Å, consistent with its role in metal stabilisation. The same histidine residue also forms a hydrogen bond with the hydrogen attached to the N1 atom of the sulfonamide group, with a distance of 2.4 Å. The second hydrogen of the N1 sulfonamide nitrogen forms a strong hydrogen bond with the oxygen atom of residue A 199 THR, at a distance of 1.8 Å.

In addition, the O1 oxygen of the sulfonamide interacts with the backbone nitrogen of A 199 THR at a distance of approximately 3.0 Å. Further stabilising interactions are observed between the heterocyclic nitrogens N3 and N2 of the ligand and the backbone oxygen of residue A 200 THR, with distances of 2.9 Å and 3.1 Å, respectively. Together, these interactions explain the strong concentration of MIFs around the sulfonamide group and its key role in anchoring the ligand within the active site.

By contrast, the opposite end of the acetazolamide molecule does not significantly overlap with any strong GRID hotspots, suggesting that this region is less optimised for interaction with the protein. This observation indicates a potential opportunity for ligand optimisation by introducing additional functional groups in this area.

Ligand optimisation

The GRID Generate Ideas module was used to explore simple modifications of acetazolamide aimed at improving its overlap with the strongest interaction fields. The MIF analysis suggested that most favourable regions were concentrated around the sulfonamide group, while the opposite side of the ligand showed fewer strong hotspots. For this reason, the proposed modification focused on the less engaged end of the molecule.

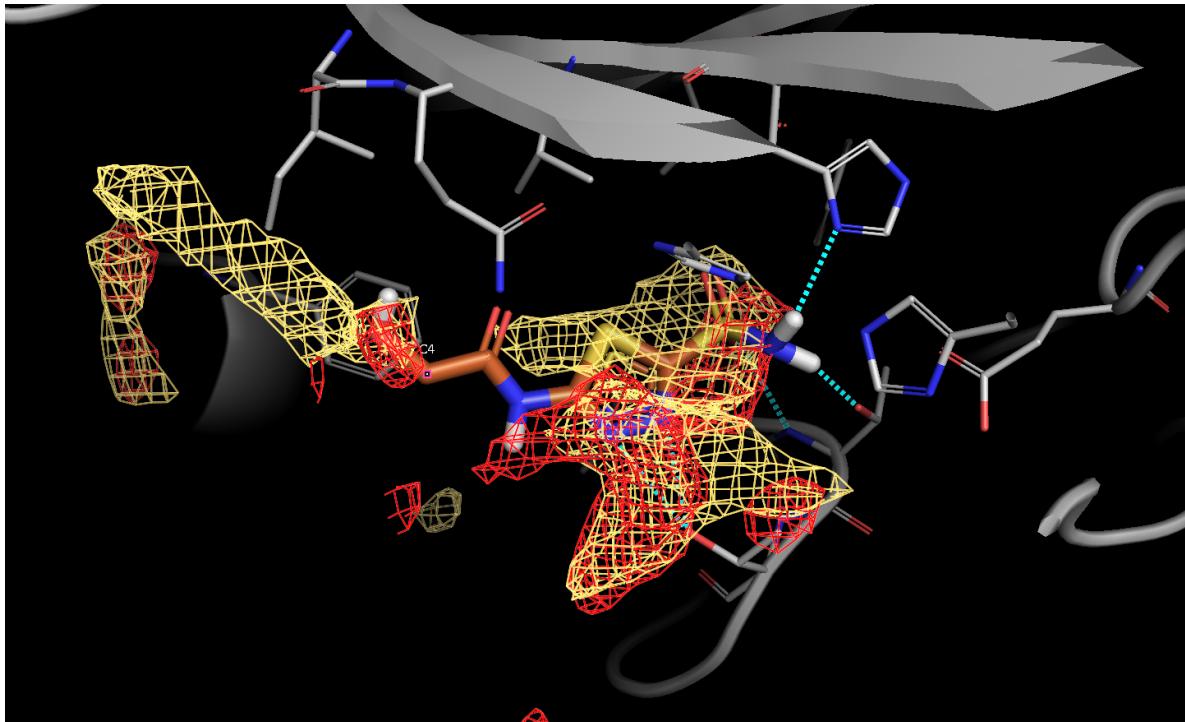


Figure 16: Proposed ligand modification generated by GRID: addition of an OH group at C4. The new substituent lies within regions highlighted by the O and C1= MIFs, consistent with improved local complementarity.

The top-ranked suggestion was the addition of a hydroxyl group at atom C4 (Figure 16). This change produced a small but favourable predicted energy improvement of approximately $\Delta E = -3.37$. No clear new direct hydrogen bonds to protein residues were observed after introducing the OH group. However, the added hydroxyl is positioned within regions highlighted by the O and C1= probes. This indicates that the local pocket environment can accommodate an oxygen atom in that position (O hotspot) while still maintaining favourable van der Waals and lipophilic contacts (C1= hotspot). In contrast, no strong N1 donor hotspot was present in this area.

Overall, this modification slightly improves complementarity with the pocket fields without disrupting the original binding mode.

FLAP 2 pseudo-apo water network and perturbation analysis

FLAP2 was used to predict the pseudo-apo water network in the binding pocket, considering a sphere of 8 Å around the bound ligand.

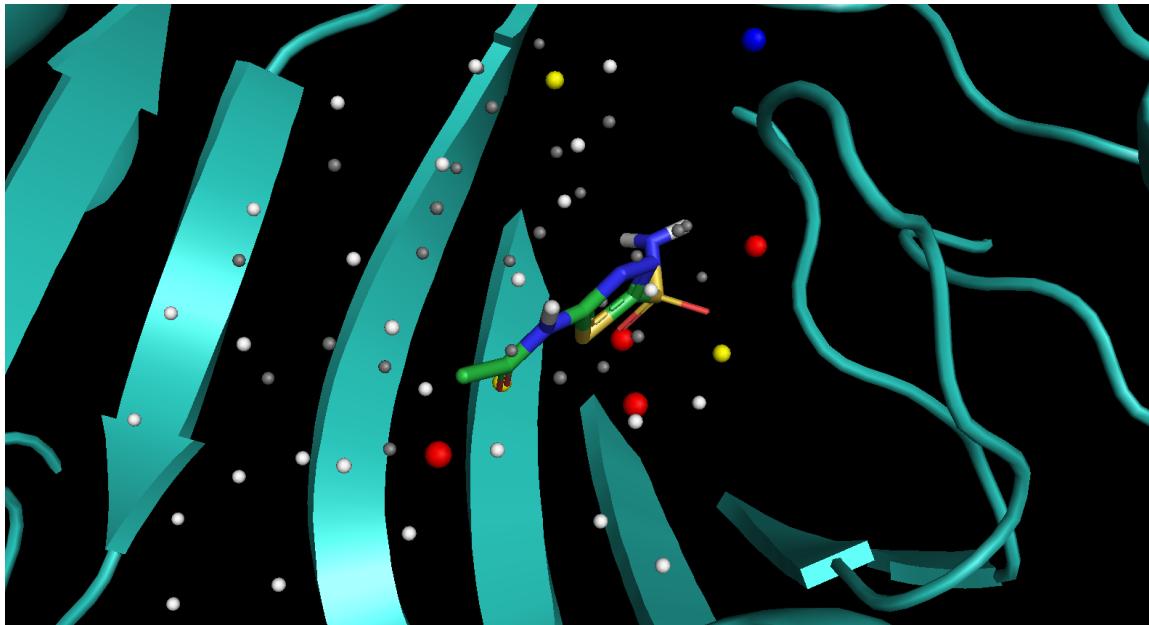


Figure 17: FLAP2 pseudo-apo water prediction within 8 Å of acetazolamide in the 3HS4 pocket. The predicted network contains 4 red waters, 3 yellow waters and 1 blue water.

The predicted water sites are shown in Figure 17. In this model, the pocket contains a small set of structured waters with different stability classes: **4 red** waters (very unhappy / energetically unfavourable), **3 yellow** waters (intermediate), and **1 blue** water (happy / favourable). In the predicted water network, several waters classified as unhappy and very unhappy are located near the sulfonamide region, while only a limited number of happy waters are present close to the ligand. This suggests that the ligand is positioned in a region where water molecules are energetically disfavoured and are likely displaced upon binding, which is generally beneficial for ligand affinity.

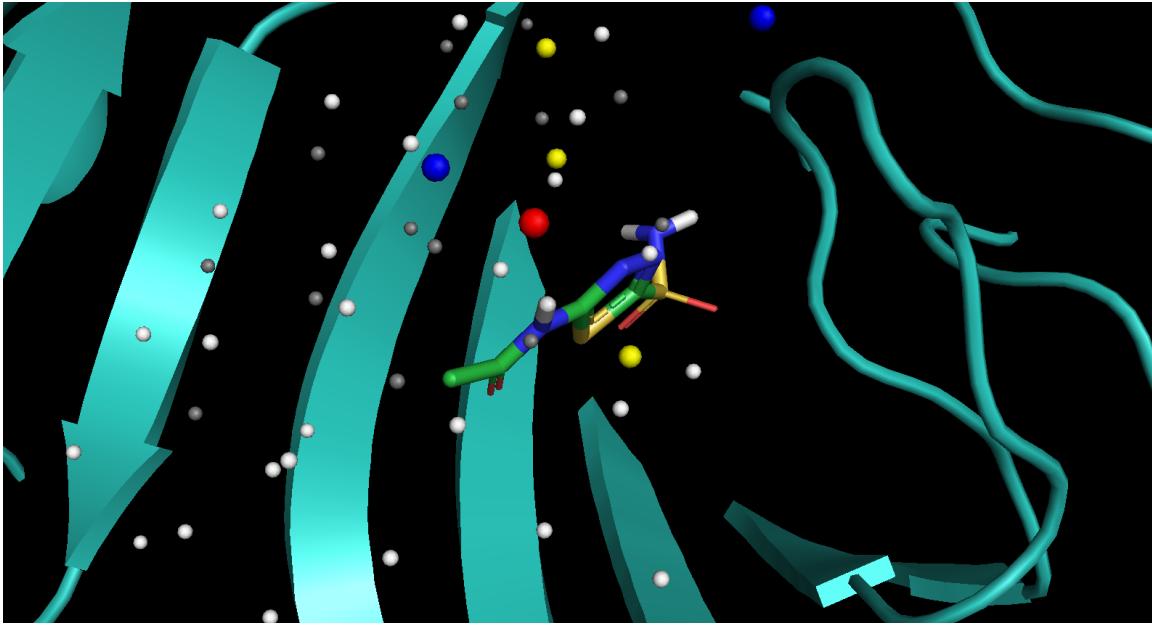


Figure 18: FLAP2 water perturbation analysis for the 3HS4 binding pocket (8 \AA around the ligand). After perturbation, the network shifts to 2 blue waters, 3 yellow waters and 1 red water, indicating that several initially unhappy sites are not strongly conserved.

Water perturbation analysis was then performed starting from the predicted apo water network. This analysis highlights how the presence of the ligand affects the local water environment. Compared to the initial prediction, the refined water network shows a redistribution of water stability classes, with fewer very unhappy waters remaining close to the ligand and some waters becoming stabilised by ligand–protein interactions. This indicates that the ligand not only displaces energetically unfavourable waters but also contributes to reorganising the remaining water network within the pocket.

References

- [1] F. Baffert, C. H. Regnier, A. De Pover, C. Pissot-Soldermann, G. A. Tavares, F. Blasco, J. Brueggen, P. Chene, P. Drueckes, D. Erdmann, P. Furet, M. Gerspacher, M. Lang, D. Ledieu, L. Nolan, S. Ruetz, J. Trappe, E. Vangrevelinghe, M. Wartmann, L. Wyder, F. Hofmann, and T. Radimerski, *Molecular Cancer Therapeutics*, vol. 9, no. 7, pp. 1945–1955, 2010.
- [2] K. H. Sippel, A. H. Robbins, J. Domsic, C. Genis, M. Agbandje-McKenna, and R. McKenna, “High-resolution structure of human carbonic anhydrase II complexed with acetazolamide reveals insights into inhibitor drug design,” *Acta Crystallographica Section F*, vol. 65, pp. 992–995, 2009. DOI: [10.1107/S1744309109036665](https://doi.org/10.1107/S1744309109036665).
- [3] RCSB Protein Data Bank, “PDB entry 3HS4: Human carbonic anhydrase II complexed with acetazolamide.” (Accessed 2026). DOI: [10.2210/pdb3HS4/pdb](https://doi.org/10.2210/pdb3HS4/pdb).