bioRxiv preprint doi: https://doi.org/10.1101/136101; this version posted May 9, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Pushing the limits of detection of weak binding using fragment based drug discovery: identification of new cyclophilin binders

Charis Georgiou,[1,2] Iain McNae,[1] Martin Wear,[1] Harris Ioannidis,[2] Julien Michel,[2] and Malcolm Walkinshaw[1]

1. University of Edinburgh, Michael Swann Building, Max Born Crescent, Edinburgh, Scotland, EH9 3BF, UK

2. University of Edinburgh, Joseph Black Building, David Brewster Road, Edinburgh, Scotland, EH9 3FJ, UK

Corresponding Authors: Malcolm Walkinshaw, m.walkinshaw@ed.ac.uk, +44(0)131 650 7056 (for biophysical experiments), Julien Michel, julien.michel@ed.ac.uk, +44(0)131 650 4797 (for computational chemistry)

Highlights

- FBDD is a popular method but weak binding is difficult to detect
- There is a need for pushing the limits of weak binding detection
- Combination of X-ray, SPR and MD methodologies increases successful characterization of weak binding events
- Several novel Cyclophilin fragment binders were identified

bioRxiv preprint doi: https://doi.org/10.1101/136101; this version posted May 9, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Abstract

Fragment Based Drug Discovery (FBDD) is an increasingly popular and successful method to identify novel small-molecule drug candidates. One of the limitations of the approach is the difficulty of accurately characterizing weak binding events. This work reports a combination of X-ray diffraction, surface plasmon resonance (SPR) experiments and molecular dynamics (MD) simulations, for the characterisation of binders to different isoforms of the cyclophilin (Cyp) protein family. Although several Cyp inhibitors have been reported in the literature, it has proven challenging to achieve high binding selectivity for different isoforms of this protein family. The present studies have led to the identification of several structurally novel fragments that bind to diverse Cyp isoforms in distinct pockets with low millimolar dissociation constants. A detailed comparison of the merits and drawbacks of the experimental and computational techniques is presented, and emerging strategies for designing ligands with enhanced isoform specificity are described.

Keywords: PPIases, cyclophilin inhibitors, fragment based drug discovery, free energy calculations, protein – ligand X-ray crystallography

Abbreviations: Cyclophilins (Cyps), cyclophilin A (CypA), peptidyl-prolyl isomerases (PPIases), cyclosporin A (CsA), human immunodeficiency virus 1 (HIV-1), hepatitis C virus (HCV), structure-based drug design (SBDD), fragment based drug discovery (FBDD), molecular dynamics (MD), surface plasmon resonance (SPR), hydrogen bond donor (HBD), hydrogen bond acceptor (HBA), dimethyl sulfoxide (DMSO), phosphate buffered saline (PBS)

bioRxiv preprint doi: https://doi.org/10.1101/136101; this version posted May 9, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Introduction

Fragment based drug discovery (FBDD) is frequently used to identify small organic molecules (fragments) as starting points for further structure-based drug design (SBDD) programs that aim to deliver drug-like molecules suitable for clinical studies. Fragments can be described using the "rule of 3" [1] [2,3]. According to this rule, a fragment is typically an organic molecule with molecular weight (Mw) ≤ 300 Da, number of H-bond acceptors ≤ 3, number of H-bond donors ≤ 3 and clogP (computed partition coefficient) ≤ 3 [1]. Fragments typically exhibit a dissociation constant in the micromolar to low millimolar range and the success of FBDD can be linked to steady improvements in robust biophysical characterisation of weak binding [2,4].

A number of problems are associated with the effective screening of fragment libraries. It is often difficult to solubilise fragments at concentrations required to saturate a protein target. Additionally the presence of aggregates, impurities and/or reactive intermediates can also lead to false positives or negatives. These issues are exacerbated for particularly small fragments (≤ 150 Da) that are likely to exhibit at best mM dissociation constants, and there is a need for pushing the limits of detection of weak binding to broaden the scope of FBDD.

The past decade has seen rapid developments in the application of molecular simulations to structure-based drug design [5]. Molecular simulations and free energy calculations are now being used to complement experimental approaches for a wide range of protein – ligand complexes, including estimation of binding energies of drug-like molecules [6,7], and fragments [8]. This report focuses on the combination of biophysical measurements and molecular simulation methods to characterize weak Cyclophilins (Cyps) binders present within a library of small fragments.

Cyclophilins (Cyps) are a family of peptidyl-prolyl isomerases (PPIases) that catalyze the isomerization of proline residues, promoting and facilitating protein folding [9]. The human cyclophilin family counts seventeen members with the prototype and most abundant being cyclophilin A (CypA) [10]. Cyclophilin orthologues can also be found in most plants,

bioRxiv preprint doi: https://doi.org/10.1101/136101; this version posted May 9, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

parasites and animals [11,12]. The amino acid sequences of all human Cyps are well established, as well as their secondary and tertiary structures [10]. All Cyps share a highly conserved PPIase domain. Cyps are also members of the immunophilins class of proteins [13–15] and many Cyps are inhibited by the natural immunosuppressant cyclosporin A (CsA) [10,15]. CsA is used in organ transplantation to prevent immune response and organ rejection. Cyps are also involved in (mis)regulation of several biological signaling pathways including damage-induced cell death [16], RNA splicing [17] and different types of cancer. Cyps are also known to be involved in the life cycle of different viruses such as Human Immunodeficiency Virus 1 (HIV-1) and Hepatitis C Virus (HCV) [18,19].

Because of their diverse biological roles Cyps are recognized as potential biological targets for the treatment of HCV [19–21], HIV [22–24], cancer [25–27] and neurodegenerative diseases such as Parkinson's and Alzheimer's [28–30]. Originally, efforts for the identification of cyclophilin inhibitors were focused on cyclic peptides, analogues of CsA [20,21,31,32]. In recent years a number of small-molecule Cyp inhibitors have been reported in the literature as [33–41]. A major unsolved challenge is to achieve robust binding affinity and specificity to distinct Cyp isoforms. The difficulty arises from the shallow, solvent-exposed nature of the active site of Cyps. Yet high binding specificity is widely thought to be necessary to produce chemical probes able to elucidate the biological roles of different Cyp isoforms, and to pave the way for next-generation Cyp drugs with reduced side-effects in comparison with CsA analogues.

The present work used Surface Plasmon Resonance, X-ray diffraction and molecular dynamics simulations to screen a focused library of small fragments against the most common Cyp isoforms Cyp A, B and D. This combination of multiple methodologies lead to the characterization of novel Cyp fragments that are suitable starting points for further optimization into more potent and isoform specific lead molecules.

bioRxiv preprint doi: https://doi.org/10.1101/136101; this version posted May 9, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Results

A focussed fragment library to interrogate ligandability of the Abu pocket in Cyclophilins

The starting point for this focused library design was 2,3-diaminopyridine, for which SPR and X-ray data had been generated previously, demonstrating the stoichiometric binding of this fragment to the Abu pocket of CypA (Figure 1) [42]. Comparative analysis of primary protein sequences has suggested that the Abu pocket in Cyps may be utilised to engineer isoform-selective ligands owing to small variations in amino-acids that line up the edge of the pocket between the common isoforms Cyp A-D. This contrasts with the nearby Pro pocket that is more conserved across the Cyp family. Thus small fragment analogues of 2,3-diaminopyridine were selected to study the chemical diversity that would be tolerated by the Abu pocket. A total of one hundred small fragments, structurally distinct from the ones previously tested, were chosen based on chemical similarity and commercial availability. Analogues include substituted aromatic rings such as pyridines, pyrazines, pyrimidines, as well as non-aromatic rings. A full list is provided in the supplementary Table 1. The relatively low average molecular weight (ca. 150 g mol^{-1}) of the library members is dictated by the small size of the Abu pocket, hence any binders are expected to exhibit at best high micromolar to low millimolar binding constants.

SPR suggests several of the fragments may bind weakly to diverse Cyp isoforms

All the fragments were initially screened by surface plasmon resonance at 1 mM, using conditions described previously [42], on a high-density (to account for the mass ratio of the fragments to His-CypA, -B and -D) surface of 3200, 3000 and 3000 RU covalently stabilized His-CypA, -B and –D, respectively. Apparent Cyp specific hits were further analysed with a 2-fold concentration series from 0.015 mM to 1 mM.

This exercise provided evidence for specific binding in the range of 0.5 – 20 mM against at least one isoform for approximately 15 fragments. However it was not possible to derive reliable K_d estimates range due to limitations in compounds solubility, and difficulties

bioRxiv preprint doi: https://doi.org/10.1101/136101; this version posted May 9, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

in rigorously removing potential refractive index artefacts at the analyte concentrations tested. Thus the outcome of the SPR screen was deemed overall encouraging, but further characterisation of potential binders was pursued via other means.

Free energy calculations estimate fragment binding energies in line with experimental data, and suggest binding pocket preferences

The standard (absolute) binding free energies of 86 of the 100 experimentally tested fragments were estimated by alchemical free energy calculations.[43] The 14 excluded compounds were either carrying out a net charge, or were too structurally dissimilar to other compounds in the library to allow straightforward computation of relative free energies of binding (see methods). The overall objective of the project was to discover new Abu pocket binders, but fragments could in principle bind to either the Abu or Pro pockets. Therefore two sets of calculations were executed with fragments docked in the Abu or Pro pockets respectively. Figure 2 depicts a histogram that summarizes the results of these calculations.

The data shows that there is a clear preference for the fragments to bind to the Abu pocket of Cyp A, with a mean and standard deviation of binding energies of -3.0±0.7 kcal.mol^{-1} *vs* -0.7±0.5 kcal.mol^{-1} for the Pro pocket. This was deemed encouraging and reflective of the strategy used to assemble diverse fragments into a focussed library targeting the Abu pocket. Conversion of the calculated absolute binding free energies into dissociation constants indicate that the range of predicted K_ds to the Abu pocket of CypA is between 0.5 – 30 mM, whereas for the Pro pocket the range is 30 mM – 2 mol L^{-1}. The computed binding energies preferences for the Abu pocket were in line with the 0.5 – 20 mM binding constant range that was estimated from the SPR screen. Altogether the molecular dynamics data was supportive of preferential weak mM binding to the Abu pocket for several fragments in the focussed library.

X-ray crystallography provides evidence of binding for several fragments

bioRxiv preprint doi: https://doi.org/10.1101/136101; this version posted May 9, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

A total of 40 compounds were selected for further X-ray crystallographic studies on CypA, CypB and CypD. The selection of this subset was guided by a combination of factors including: evidence of binding to a Cyp surface in the preceding SPR screen; computed interactions with the Abu/Pro pockets; desire for producing co-crystal structures of fragments bounds to multiple isoforms; evidence of sufficient solubility; available beam-time.

Unfortunately no crystal structures of fragments in complex with CypB or CypD were obtained. Although well diffracting crystals were obtained for CypB, the crystal packing arrangement blocked access to the active site and soaking experiments were thus not successful. CypD yielded crystals that were too fragile for soaking and data collection.

Over thirty compounds that were soaked into CypA crystals diffracted well. Refinement of the structural data led to the detection of 10 fragments in complex with CypA. The chemical structures of all such fragments are given in Table 1 and data collection, refinement and Ramachandran plot statistics for each of the CypA complexes are reported in the supplementary Table 2. Fragments that were not found in the active site of CypA are listed in the supplementary Table 3. The majority of the observed compounds bind in the Abu pocket, in line with the expectations from the library design strategy and MD calculations, however some fragments were found to bind to the neighbouring Pro pocket. The pocket preference appears to be dictated by the nature of H-bond donors and acceptors, and also the hydrophobicity of each ligand.

Ligands with hydrogen bond donors and acceptors such as **3**, **5**, **56**, **89**, **97**, **98** and **99** bind to the more hydrophilic Abu pocket of CypA. This is in agreement with the MD simulations that predicted a preference for the Abu over the Pro pocket for all 7 fragments (mean preference -2.6±1.0 kcal.mol^{-1}). These compounds adopt broadly similar binding poses and interactions with the Abu pocket. Compounds **89**, **98** and **99** have a very similar binding pose, and the electron density is well defined (Supplementary Fig. 2). Figure 3 depicts fragment **98** and shows how the 4-amino group forms hydrogen bonds with two water molecules, W1 (violet-purple) and W2 (sky-blue). Both W1 and W2 have B-factors values (ca. 15 Å^2) lower than for most other water molecules present in the crystal structure.

bioRxiv preprint doi: https://doi.org/10.1101/136101; this version posted May 9, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

This was taken as indication that these two hydration sites are well ordered. By contrast the 3-amino group is H-bonded to two additional mobile water molecules W3 (pink) and W4 (green). These water molecules have higher B-factors values (ca. 30 $Å^2$) and are thus expected to be less ordered. Overall this somewhat unusual cluster of water mediated protein-fragment interactions play a key role for fragment binding to the Abu pocket. W1 forms H-bond interactions with three protein residues; Ala101, Gln111 and Gly109; W2 with Gly74, Ser110 and Gln111; W3 with Glu81, Gly75 and W4 with Thr73.

Additionally a diverse range of small functional groups (halides, amino, methyl) is tolerated at a third position on the aromatic ring. Fragments **3, 5, 56** and **97** also bind in the Abu pocket of CypA and occupy the same position as **89**, **98** and **99**, though their electron density is less well defined. An unambiguous assignment of their binding poses is not possible and, in light of the molecular dynamics simulation results (see below), the electron density map was interpreted as the result of these ligands adopting multiple poses.

Fragment **56** seems to be able to adopt a slightly different orientation in the Abu pocket, with the chloro group of the ring buried in the pocket, whereas the amino group is more water exposed. This alternative pose is stabilised by two other water molecules that interact with the amine and the pyrimidine nitrogen atoms of the fragments and residues Gly63 and Thr73. The structure of all compounds bound to the Abu pocket, as well as their electron density and distances between fragments and protein/water molecules can be seen in Supplementary Fig. 2.

Some fragments were found to also bind out of a pocket in the vicinity of the 80's loop. Figure 4 shows fragment **3** as an example. This second binding site lies between two symmetry related Cyp molecules in the crystal CypA with halogen – carbonyl contacts to Asn106 and Glu84 and an aromatic interaction with Trp121 of another CypA molecule. This site was deemed to be a crystallographic artefact and ignored for further studies.

Three of the ten fragments, **16**, **60** and **61** were found to bind the Pro pocket of CypA instead of the Abu pocket. This contrasts with the MD simulations that predicted a preference for binding to the Abu pocket (mean preference -1.7±0.7 $kcal.mol^{-1}$). Figure 5

bioRxiv preprint doi: https://doi.org/10.1101/136101; this version posted May 9, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

shows that crystallography indicates the three fragments establish H-bond interactions in the Pro pocket with the conserved CypA residues, Arg55, Gln63 and Asn102. In contrast to the Abu pocket binders that all contain one aromatic ring, fragments **16**, **60**, **61** are made of non-planar six membered rings. Thus a combination of lack of planarity and increased hydrophobicity seems to bias fragments towards binding to the conserved Pro pocket.

Molecular dynamics simulations suggest that some fragments adopt multiple binding modes and help guide the interpretation of X-ray diffraction measurements

The structures of fragments derived by X-ray crystallographic analyses revealed a conserved binding mode that appears to be adopted by all fragments binding to the Abu pocket (Fig. 3). However for some fragments (e.g. **3**, **5, 56** and **97**) refinement of the structures was ambiguous and suggested that these fragments may adopt two different binding modes in the Abu pocket. This was investigated further by carrying out a detailed analysis of MD simulations of **3** bound to Cyp A. Since **3** shows a high degree of structural similarity to the other fragments found to the Abu pocket, it is anticipated that the observed interactions are also relevant for the other fragments. Figure 6 shows RMSD distributions of **3** from one X-ray crystallographic pose calculated from the MD simulations. A total of three different clusters are apparent in the distribution. Monitoring of hydrogen bonding distances between ligand/protein donor/acceptor groups also indicates three clusters whose structure and population match the results produced by RMSD analyses (Supplementary Fig. 3). The first cluster, denoted ‘’MD pose A’’, corresponds to a binding mode very similar to one X-ray pose (‘’X-ray pose A’’) with a RMSD of 1 ± 1 Å. The second cluster, ‘’MD pose B’’, corresponds to a binding mode very similar to another X-ray pose (‘’X-ray pose B’’) with a RMSD of 0.8± 0.3 Å. The last cluster (‘’MD pose C’’) deviates more significantly from the X-ray data, with RMSD values of 3.8 ± 0.5 Å and 2.5 ± 0.5 Å to X-ray pose A and X-ray pose B respectively. MD pose A accounts for approximately 30% of the conformations, MD pose pose B accounts for 55% of the conformations, and MD pose C accounts for the remaining 15%.

bioRxiv preprint doi: https://doi.org/10.1101/136101; this version posted May 9, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Detailed visualisation of the MD trajectories indicates the interactions that account for the stability of each pose. Figure 7 shows four representative trajectory snapshots from MD pose A, indicating that compound **3** remains flexible, with its amino group interconverting rapidly between alternative hydrogen-bonding arrangements involving Thr107 (pale-green), Ala101 (light yellow), W1 (violet-purple) and W2 (sky blue). Figure 8a shows a representative snapshot from MD pose B. Compound **3** is rotated by 90° in the Abu pocket with respect to MD pose A, and is H-bonded to Thr73 (cyan) and W2 (orange). This is the main interaction pattern observed in MD pose B. Figure 8b depicts a representative trajectory snapshot of compound **3** in MD pose C. This pose is further rotated by 90° degrees with respect to pose B to position the amino group of **3** in the direction of the Pro pocket. This pose features mainly hydrogen bonding interactions with the backbone carbonyl oxygen of Gly72 (red). Detailed statistics about hydrogen-bonding interactions in each poses are given in supplementary figure 3 and 4.

In all MD trajectories water molecules W1 and W2 are very tightly bound in the Abu pocket and stay in the same orientation for the duration of the simulations. This is in agreement with the X-ray data which showed W1 and W2 present in the Abu pocket in all structures with low B-factors. In contrast to W1 and W2, water molecules W3 and W4 are more mobile, and easily displaced by other water molecules during the MD simulations. This is again in agreement with the X-ray results which showed W3 and W4 to have higher B-factor values compared to the W1 and W2 crystal water sites.

Overall this detailed analysis increased confidence in the interpretation of electron density maps derived from X-ray crystallography, and ligands **3** and **97** were ultimately refined in two different poses with 50% occupancy each (Supplementary Figure 2).

Discussion and Conclusion

How confidently can weak mM fragment binders be determined?

The combined data from SPR, X-ray crystallography measurements and MD simulations provides solid evidence that several of the fragments from the present in-house

bioRxiv preprint doi: https://doi.org/10.1101/136101; this version posted May 9, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

bespoke library interact with Cyp surfaces with dissociation constants in the low millimolar range. Figure 9 depicts a Venn diagram demonstrating overlap between 'hits' detected between X-ray analyses and MD simulations in the case of Cyp A. The SPR data was deemed too ambiguous to enable meaningful evaluation of hit rates and is thus not included in Figure 9. Given the criteria used here for declaring a compound a 'hit', X-ray and MD detected 10 and 33 binders respectively. This corresponds to a hit rate of 25-33% for X-ray (10 out of 40, or 10 out 30 if poorly diffracting crystals are excluded from the dataset), and 38% for MD (33 out of 86). For X-ray the definition of a hit is relatively clear (structure of a bound fragment refined against observed electron density), but many samples could not be analysed to detect ligands due to low quality diffraction data. Additionally, the methodology as used here doesn't yield K_d values and other techniques must be used to estimate binding constants for the fragments. By contrast for MD the criteria for declaring a hit is less clear cut since binding constant estimates are obtained for all compounds tested. Given the observed distribution of computed binding energies in Figure 2, a cutoff in standard binding free energies of -3.5 kcal.mol^{-1} was deemed reasonable. This corresponds to dissociation constants in line with the SPR estimates, and ligand efficiencies in the range of 0.3-0.5. Of course the number of hits can be decreased merely by setting more stringent requirements for the MD binding energy estimates.

Overall 3 compounds were found active in both techniques. Out of these 3 binders, compound **98** was deemed of particular interest (Table 2). MD simulations suggest a CypA binding energy of -3.7 ± 0.4 kcal mol^{-1}. The corresponding ligand efficiency is reasonable given the challenges posed by the Abu pocket in Cyclophilins. Compound **98** was also observed by X-ray crystallography in the Abu pocket of CypA, establishing H-bond interactions with water molecules and protein residues according to the canonical pattern observed for other Abu pocket fragments.

How can molecular dynamics complement biophysical methods in fragment-based drug discovery?

bioRxiv preprint doi: https://doi.org/10.1101/136101; this version posted May 9, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

The binding free energies computed via molecular dynamics simulations were able to reproduce the preference for the binding of the fragments to the Abu pocket instead of the Pro pocket of Cyp A. Additionally the MD binding affinity estimates were in line with the range inferred from SPR analysis. These observations suggest that MD simulations can be used to guide the design of focussed fragment libraries enriched in structures more likely to bind a target pocket with a reasonable binding affinity. This contrasts with the more common strategy that involves assembling a generic library of structurally diverse fragments as a primary screening library.

Nevertheless a significant number of MD 'hits' were not observed via X-ray diffraction experiments. Arguably, reliably rank-ordering compounds by potency is particularly challenging for this dataset since by design the focussed library contains compounds that are structurally similar, and thus can all be reasonably modelled in the Abu pocket. Most of the library compounds show a computed standard binding energy that falls within the interval of -3 ± 1 kcal mol^{-1} (Figure 2). The relatively small dynamic range of calculated binding energies poses challenges given the precision of the free energy estimates (ca. 0.4 kcal mol^{-1}), and the typical accuracy of classical biomolecular force-fields (typical systematic errors around 1 kcal mol^{-1} in favourable cases). It is likely that greater performance may be observed for the scoring of structurally diverse fragment libraries. This hypothesis could be tested by blinded-predictions, in a fashion similar to competitions for blinded estimations of standard binding energies of host/guest and protein/drug-like complexes have been organised recently. [44,45] It also cannot be ruled out that the conditions used to solve protein structures in X-ray crystallography experiments disfavour the binding of some of the fragments predicted by MD simulations to bind in solution.

Analysis and visualisation of MD trajectories for selected fragments proved to be a rich source of structural insights. The simulations revealed that several fragments remained highly dynamic when bound to the Abu pocket. In the case of compound **3**, three different poses were characterised in details (Fig .7, Fig .8, Supplementary Fig. 3 and Supplementary Fig. 4). This may explain why the electron density maps of a number of fragments (e.g. **3**, **5**

bioRxiv preprint doi: https://doi.org/10.1101/136101; this version posted May 9, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

and **56**) were ambiguous and could not be fitted to a single binding pose. This suggests that it may be valuable to carry out MD simulations to assist model refinement process of X-ray diffracted crystals. The additional binding modes inferred from MD may also generate ideas for growing fragments out of the Abu pocket that may have been overlooked if design considerations had relied on a X-ray diffracted crystal structure showing a single pose. For instance the binding poses B and C of compound **3** suggest that Gly72 and Thr73 residues can form additional H-bonds interactions with Cyp binders.

What did the fragments reveal about Cyclophilin inhibition strategies?

The binding modes of the various fragments to the Abu pocket of Cyp A were found to be remarkably similar to the interactions of compounds recently reported by the Pawlotsky and Guichou groups [38,46]. In these studies, several compounds were described as Cyp inhibitors and were co-crystallized in the active site of CypD. Gelin *et al.* and Ahmed-Belkacem *et al.* reported in particular that a para-aniline moiety was one of the very few substituents tolerated in the Abu pocket of CypD. The amino group forms hydrogen bonding interactions with Thr107 [38,46]. The X-ray structure of one inhibitor (PDB ID 4ZSC) was aligned with **3** from the present study (Supplementary Fig. 5). The fragment aligns well with the ligand and forms similar patterns of interactions (even though the isoforms differ). Considerable efforts have been invested by Ahmed-Belkacem *et al.* to replace the para-aniline moiety by an analogue that presents fewer toxicity liabilities for *in vivo* studies. The present results suggest that renewed chemistry efforts to merge the presently disclosed Abu pocket fragments with the scaffold from the lead series of Ahmed-Belkacem et al. may produce superior ligands for further drug development or *in vivo* investigations.

Finally, the three Cyp A Pro pocket fragments **16**, **60** and **61** (Fig .5) were found to form interactions similar to those observed in crystal structures of Cyclosporine A in complex with Cyp A [10]. Specifically these three fragments are able to establish H-bond interactions with key structural and catalytic active site residues including Arg55, Gln63 and Asn102.

bioRxiv preprint doi: https://doi.org/10.1101/136101; this version posted May 9, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

This suggests that the binding modes of these fragments provide a good template for the elaboration of more potent lead-like molecules.

In conclusion, this work explored new ways of combining molecular dynamics simulations with biophysical measurement methods for fragment-based drug design. This work led to the discovery of several novel fragments that bind to Cyclophilin surfaces, and the resulting structural information will impact on future efforts to discover novel classes of potent and selective Cyclophilin ligands.

bioRxiv preprint doi: https://doi.org/10.1101/136101; this version posted May 9, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Materials and Methods

Protein expression purification and characterization

N-terminal hexa-histidine tagged human Cyclophilin (CypA, CypB and wt CypD) were provided by the Edinburgh Protein Production Facility (EPPF), University of Edinburgh. K133I-CypD plasmid was provided by Dr Jacqueline Dornan (Walkinshaw group, Edinburgh). Expression and purification for all isoforms was performed as described [42]. Dynamic light scattering (DLS) was used to verify the proteins' quaternary structure and to examine any possible aggregation between Cyp monomers. DLS was performed in triplicates for each protein sample at 20 °C, on a Zetasizer APS instrument from Malvern using a 384 well plate with the total final volume of each well used to be 60µL. Prior DLS protein samples were buffer exchanged to PBS, concentrated to 1mg ml^{-1} and spun gently at 13,500 g for 20 minutes at 4 °C. Polydispersity index for all proteins was calculated as: 0.15 CypA, 0.16 CypB and 0.21 for CypD. Moreover, the hydrodynamic radius (R_H) for each protein was also determined (4.3 ± 1.7 d.nm for CypA, 4.6 ± 1.8 d.nm for CypB and 4.3 ± 1.1 d.nm for CypD) and was in accordance with expected values for all Cyps, given existing X-ray crystallographic co-ordinates

SPR

Compounds were initially screened at 1mM on a surface of 3200, 3000 and 3000 RU covalently stabilized His-CypA, -B and –D, respectively, in 10 mM PBS, 137 mM NaCl, 2.7 mM KCl, pH 7.4, 0.005 % v/v P20; 0.5 % v/v DMSO (2% v/v ethanol present to increase solubility of some compounds), at 30 $\mu l \cdot min^{-1}$ with a 15 sec contact and dissociation time at 25˚C. The surface was regenerated between each measurement and the compounds were washed off the surface of the sensor by running an excess of buffer solution at 40 $\mu l.min^{-1}$ for 30 sec followed by a further 5 sec stabilization period. Selected compounds were further analysed with a 2-fold concentration series from 0.015 mM to 1 mM in 10 mM PBS, 137 mM NaCl, 2.7 mM KCl, pH 7.4, 0.005 % v/v P20; 0.5 % v/v DMSO (2% v/v ethanol present to

bioRxiv preprint doi: https://doi.org/10.1101/136101; this version posted May 9, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

increase solubility of some compounds) at 30 µl·min^{-1} with a 15 sec contact and dissociation time.

Crystallization and data collection

Purified and his-tag cleaved CypA was concentrated to 30.4 mg ml^{-1} in PBS buffer. The vapor diffusion by hanging drop method at 6 ºC was used for the crystallization of CypA. The well volume was 1 mL, the precipitation solution consisted of 100 mM Tris-HCl, pH 8.0 and 21 – 24 % v/v PEG 8,000 and the drop consisted of 1.5 µL of CypA and 1.5 µL of well solution. Each ligand was soaked into an *apo* CypA crystal using 2 µl of soaking solution (30 % v/v PEG 8,000, 100 mM Tris-HCl,pH 8.0) and saturated concentrations of ligands (50 – 100 mM), followed by flash freezing with liquid nitrogen. Unliganded crystals were also prepared for obtaining *apo* CypA structures. X-ray intensity data were collected at the Diamond synchrotron-radiation facility in Oxfordshire, England. Intensity data for *apo* CypA and CypA-ligand complexes were collected from single crystals flash cooled in liquid nitrogen at 100 K. Data were processed with MOSFLM [47] and scaled with SCALA [48].

Structure determination and refinement

All structures were originally processed with DIMPLE [49]. DIMPLE was followed by multiple cycles of restrained refinement using REFMAC 9 [50,51]. The side chains of the models were manually adjusted, while ligands and water molecules were added where appropriate using Coot 10 [52]. Further iterative cycles of restrained refinements and manual adjustments of ligands, side chains and water molecules, were carried out until the overall quality of the electron density maps could no longer be further improved.

Preparation of proteins and ligands for free energy calculations

CypA, CypB and CypD structures were taken from the X-ray crystal structures with PDB IDs: 1CWA, 3ICH and 2BIT respectively. All water molecules were removed from the structures except the three tightly bound water molecules in the Abu pocket of cyclophilins

bioRxiv preprint doi: https://doi.org/10.1101/136101; this version posted May 9, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

[53]. All proteins were capped at the C-terminal and N-terminal with an N-methyl and acetyl groups, respectively. Protonation of the histidine residues was predicted by PROPKA, as implemented in Maestro Protein Preparation Wizard [54,55]. Specifically, the protonation states of CypA's histidines were: HIE54, HIE70, HID92, HID126; in CypB: HIE54, HID92, HID126 and in CypD: HIE54, HIE70, HID92, HID126 and HID131. Ligands were prepared using Maestro as distributed by Schrödinger package [54]. Most ligands were neutral, but a number were expected to be charged in assay conditions. Since correcting for net-charge changes is non-trivial in free energy calculations, these compounds were not included in the dataset for free energy calculations. [56,57]

Docking calculations

All ligands were docked into the Abu and Pro pockets of CypA, CypB and CypD using the Autodock Vina plugin of Pymol [58,59]. For the docking calculations in the Abu pocket, the grid was centered in the protein's Abu pocket. The box size of the grid was set at $x = y = z = 13.12$ Å and the spacing between the grid points ($x = y = z = 35$) was 0.375 Å. The whole protein was used as a receptor and only one pose was generated for each ligand. Moreover, three water molecules that are tightly bound in Abu pocket [53] were kept and treated as part of the protein for docking purposes.

All ligands were docked in the Pro pocket of CypA, CypB and CypD, in a similar way. The grid was centered in the Pro pocket of the protein while the box size of the grid was set at $x = y = z = 7.50$ Å and the spacing between the grid points ($x = y = z = 20$) was 0.375 Å. As was previously described, the whole protein was used as a receptor and only one pose was generated for each ligand.

Free energy calculations set up

The strategy used here consisted in using series of relative free energy calculations to connect all compounds in the dataset to a reference compound, and then to compute a

bioRxiv preprint doi: https://doi.org/10.1101/136101; this version posted May 9, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

standard (absolute) binding free energy calculation for the reference compound. This two-step procedure yields standard binding free energies for all compounds in the dataset.

For the calculation of relative binding free energies, ligands should be perturbed from one ligand to another in complex with the protein and alone in a water box. For this reason, a perturbation map was generated (Supplementary Fig. 6), by manually connecting all the ligands via multiple transformations. Ligand **3** was set as the center of the map, due to its high chemical similarity to the rest of the compounds. All ligands were connected by keeping the number of perturbation steps to **3** to less than four. Mappings featured compounds with high structural similarity, in order to reduce the number of perturbed atoms in each relative free energy calculation. Input files for all subsequent free energy simulations were set up using FESetup1.1 software [60].

Protein parametrization was done using the Amber ff14SB force field [61], while ligands were parametrized using GAFF force field as implemented in amber14 [62,63] that uses AM1-BCC charges [64,65]. All systems were solubilized in a rectangular box of TIP3P water molecules, with a box length of 10 Å away from the edge of the solute, and Na^+ or Cl^- ions were added to neutralize the net charge of the system. The systems were energy minimized for 200 steps to alleviate any steric clashes, followed by a heating step to 300 K for 200 ps, with harmonic potential restraints on all non-solvent atoms using a 10 kcal mol^{-1} $Å^2$ force constant. Systems were then equilibrated by running short (200 ps) MD runs using NVT ensemble with the same restraints as before. Finally, to stabilize the density a final run of 5 ns using an NPT ensemble at 1 atm was performed. The last trajectory snapshot was used as the starting point for the free energy calculations.

Alchemical free energy calculations

Relative free energy changes for transforming a ligand L1 into a ligand L2 ($\Delta\Delta G_{(L1\rightarrow L2)}$) were calculated as the difference in the free energy change of transforming L1 to L2 in a water box ($\Delta G_{w(L1\rightarrow L2)}$) and in complex with the protein ($\Delta G_{p(L1\rightarrow L2)}$) as shown by equation1:

bioRxiv preprint doi: https://doi.org/10.1101/136101; this version posted May 9, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

$$\Delta\Delta G_{(\mathrm{L1}\rightarrow\mathrm{L2})} = \Delta G_{w(\mathrm{L1}\rightarrow\mathrm{L2})} - \Delta G_{p(\mathrm{L1}\rightarrow\mathrm{L2})} \quad \textbf{(1)}$$

The MBAR implementation in the software pymbar was used for the calculation of both $\Delta G_{w(\mathrm{L1}\rightarrow\mathrm{L2})}$ and $\Delta G_{p(\mathrm{L1}\rightarrow\mathrm{L2})}$ free energy changes [66]. Relative binding free energy calculations were only performed once, and statistical errors were propagated from each step, with the final reported error calculated using equation 2

$$\mathrm{err}\left(\Delta\Delta G_{(\mathrm{L1}\rightarrow\mathrm{L2})}\right) = sqrt\left(err\left(\Delta G_{W,(\mathrm{L1}\rightarrow\mathrm{L2})}\right)^2 + err\left(\Delta G_{P,(\mathrm{L1}\rightarrow\mathrm{L2})}\right)^2\right) \quad \textbf{(2)}$$

where the terms on the right hand side are the statistical errors associated with $\Delta G_{w(\mathrm{L1}\rightarrow\mathrm{L2})}$ and $\Delta G_{p(\mathrm{L1}\rightarrow\mathrm{L2})}$ respectively.

Absolute binding free energies were calculated instead with equation 3:

$$\Delta G^{o}_{bind}(\mathrm{L3}) = \Delta G^{C}_{W} + \Delta G^{LJ}_{W} - \Delta G^{C}_{P} - \Delta G^{LJ}_{P} - \Delta G^{o\ restr}_{P} \quad \textbf{(3)}$$

where $\Delta G^{o}_{bind}(\mathrm{L3})$ is the absolute binding free energy of ligand **3** (L3). ΔG^{C}_{W} and ΔG^{C}_{P} are the free energies of turning off the partial charges of L3 in water and in complex respectively, while ΔG^{LJ}_{W} and ΔG^{LJ}_{P} are the free energies of turning off the Lenard Jones parameters of the uncharged ligand. $\Delta G^{o\ restr}_{P}$ is the free energy cost of releasing the restraints to reach standard state conditions, as explained in ref [67]. This quantity was computed with a custom script. Each absolute binding free energy calculation was repeated in quadruplicate and the reported value is the average of the quadruplicate. The final error reported for the absolute binding free energy was calculated as the standard error of the mean, from the four repeats.

Free energy simulations were performed using the Sire – OpenMM framework [68,69] on GPUs (GeForce GTX465 and Tesla/M2090/K20 graphic cards). Twelve equidistant λ windows (0.0000, 0.0909, 0.1818, 0.2727, 0.3636, 0.4545, 0.5455, 0.6364, 0.7372, 0.8182, 0.9091, 1.00) were used and the systems were energy minimized for 1000 steps and then re-equilibrated at the appropriate *λ* value for 20 ps. The total length of each simulation (at each *λ*) was 5 ns, and the perturbed energies were saved every 200 fs. The perturbed energies where post-processed using the MBAR estimator [70]. A softcore potential was used in all simulations [71]. The hydrogen mass repartitioning (HMR)

bioRxiv preprint doi: https://doi.org/10.1101/136101; this version posted May 9, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

technique [72] was used to increase the timestep of the simulation and decrease the time needed to run each simulation. A 4 fs timestep and a repartitioning factor of 4 were used and all bonds were constrained. Simulations were performed in the NPT ensemble, where temperature control was achieved by using the Andersen thermostat, [73] and pressure control was achieved by a Monte Carlo Barostat. Periodic boundary conditions were applied with a 10 Å cut-off for the non-bonded interactions.

In relative binding free energy simulations interatomic flat-bottom distance restraints of 7.5 ± 3.0 Å in the Abu and Pro pockets were applied with a force constant of 5 kcal mol^{-1}. One atom of each ligand was selected and restrained to one protein atom in the Abu pocket (Gln111/backbone O atom) in the case of ligands present in the Abu pocket. For the simulations of ligands in the Pro pocket, the protein atom was Met61/CA and the ligand atom was the same as in the Abu pocket. Although restraints are not usually needed in relative binding free energy simulations, these were considered necessary in the present simulations. This is because fragments weakly bound to a shallow pocket can drift away from their starting binding position into bulk after a few ns of simulation, leading to the evaluation of free energy changes that no longer correspond to interactions with the desired pocket. The selected atomic distance restraints and the force constant were sufficient to restrain our fragments in the binding pocket and prevent them from drifting away, while at the same time, avoiding unduly penalizing motions in the Abu or Pro pockets.

For reference compound **3,** absolute binding free energy calculations were performed to both the Abu and Pro pockets of CypA. Ligand **3** was restrained in the pocket of the protein using a force constant of 5 kcal mol^{-1}. One atom of ligand **3** was distance restrained to two atoms of the protein in the Abu pocket (GLN111 and ASN102 CA backbone atoms) and two atoms of the protein in the Pro pocket (HID126 and ILE57 CB atoms). In the Abu pocket the distance restraint tolerance was set to 5.5 ± 1.5 Å and in the Pro pocket to 7.5 ± 1.0 Å. The free energy cost of setting up this restraint was corrected by post-processing the simulations to yield a standard free energy of binding.

bioRxiv preprint doi: https://doi.org/10.1101/136101; this version posted May 9, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

All input files for the FESetup and somd codes, and a summary of the calculated free energies obtained by post-processing of the generated trajectories can be downloaded from the github repository located at https://github.com/michellab/cypafragmentsfreeenergy.

Accession numbers

Coordinates and structure factors for all ligands binding CypA were deposited in the protein data bank with accession numbers: CypA_lig3 (**5NOQ**), CypA_lig5 (**5NOR**), CypA_lig16 (**5NOS**), CypA_lig56 (**5NOT**), CypA_lig60 (**5NOU**), CypA_lig61 (**5NOV**), CypA_lig89 (**5NOW**), CypA_lig97 (**5NOX**), CypA_lig98 (**5NOY**), CypA_lig99 (**5NOZ**).

Acknowledgements

Gratitude is expressed to BBSRC and the EaStBio DTP for funding this project. Julien Michel is supported by a University Research Fellowship from the Royal Society.

References

[1] M. Congreve, R. Carr, C. Murray, H. Jhoti, A “Rule of Three” for fragment-based lead discovery?, Drug Discov. Today. 8 (2003) 876–877. doi:10.1016/S1359-6446(03)02831-9.

[2] N. Fuller, L. Spadola, S. Cowen, J. Patel, H. Schönherr, Q. Cao, A. McKenzie, F. Edfeldt, A. Rabow, R. Goodnow, An improved model for fragment-based lead generation at AstraZeneca, Drug Discov. Today. 21 (2016) 1272–1283. doi:10.1016/j.drudis.2016.04.023.

[3] D.A. Erlanson, S.W. Fesik, R.E. Hubbard, W. Jahnke, H. Jhoti, Twenty years on: the impact of fragments on drug discovery, Nat. Rev. Drug Discov. 15 (2016) 605–619. doi:10.1038/nrd.2016.109.

[4] J.-P. Renaud, C. Chung, U.H. Danielson, U. Egner, M. Hennig, R.E. Hubbard, H. Nar, Biophysics in drug discovery: impact, challenges and opportunities., Nat. Rev. Drug

bioRxiv preprint doi: https://doi.org/10.1101/136101; this version posted May 9, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Discov. (2016) 1–20. doi:10.1038/nrd.2016.123.

[5] J. Michel, Current and emerging opportunities for molecular simulations in structure-based drug design., Phys. Chem. Chem. Phys. 16 (2014) 4465–4477. doi:10.1039/c3cp54164a.

[6] G. Calabrò, C.J. Woods, F. Powlesland, A.S.J.S. Mey, A.J. Mulholland, J. Michel, Elucidation of Nonadditive Effects in Protein-Ligand Binding Energies: Thrombin as a Case Study, J. Phys. Chem. B. 120 (2016) 5340–5350. doi:10.1021/acs.jpcb.6b03296.

[7] S.K. Mishra, G. Calabró, H.H. Loeffler, J. Michel, J. Koča, Evaluation of Selected Classical Force Fields for Alchemical Binding Free Energy Calculations of Protein-Carbohydrate Complexes, J. Chem. Theory Comput. 11 (2015) 3333–3345. doi:10.1021/acs.jctc.5b00159.

[8] T.B. Steinbrecher, M. Dahlgren, D. Cappel, T. Lin, L. Wang, G. Krilov, R. Abel, R. Friesner, W. Sherman, Accurate Binding Free Energy Predictions in Fragment Optimization, J. Chem. Inf. Model. 55 (2015) 2411–2420. doi:10.1021/acs.jcim.5b00538.

[9] G. Fischer, B. Wittmann-Liebold, K. Lang, T. Kiefhaber, F.X. Schmid, Cyclophilin and peptidyl-prolyl cis-trans isomerase are probably identical proteins., Nature. 337 (1989) 476–478. doi:10.1038/337476a0.

[10] T.L. Davis, J.R. Walker, V. Campagna-Slater, P.J. Finerty, P.J. Finerty, R. Paramanathan, G. Bernstein, F. Mackenzie, W. Tempel, H. Ouyang, W.H. Lee, E.Z. Eisenmesser, S. Dhe-Paganon, Structural and biochemical characterization of the human cyclophilin family of peptidyl-prolyl isomerases, PLoS Biol. 8 (2010). doi:10.1371/journal.pbio.1000439.

[11] A. Marín-Menéndez, P. Monaghan, A. Bell, A family of cyclophilin-like molecular chaperones in Plasmodium falciparum, Mol. Biochem. Parasitol. 184 (2012) 44–47. doi:10.1016/j.molbiopara.2012.04.006.

[12] A. Bell, P. Monaghan, A.P. Page, Peptidyl-prolyl cis-trans isomerases (immunophilins)

bioRxiv preprint doi: https://doi.org/10.1101/136101; this version posted May 9, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

and their roles in parasite biochemistry, host-parasite interaction and antiparasitic drug action, Int. J. Parasitol. 36 (2006) 261–276. doi:10.1016/j.ijpara.2005.11.003.

[13] A. Galat, Peptidylprolyl cis/trans isomerases (Immunophilins): biological diversity-targets-functions., 2003. doi:10.2174/1568026033451862.

[14] J. Dornan, P. Taylor, M.D. Walkinshaw, Structures of immunophilins and their ligand complexes., Curr. Top. Med. Chem. 3 (2003) 1392–409. doi:10.2174/1568026033451899.

[15] N. Takahashi, T. Hayano, M. Suzuki, Peptidyl-prolyl cis-trans isomerase is the cyclosporin A-binding protein cyclophilin., Nature. 337 (1989) 473–5. doi:10.1038/337473a0.

[16] P.C. Waldmeier, K. Zimmermann, T. Qian, M. Tintelnot-Blomley, J.J. Lemasters, Cyclophilin D as a drug target., Curr. Med. Chem. 10 (2003) 1485–1506. doi:10.2174/0929867033457160.

[17] D.S. Horowitz, E.J. Lee, S.A. Mabon, T. Misteli, A cyclophilin functions in pre-mRNA splicing, EMBO J. 21 (2002) 470–480. doi:10.1093/emboj/21.3.470.

[18] A. Iwasaki, Innate Immune Recognition of HIV-1, Immunity. 37 (2012) 389–398. doi:10.1016/j.immuni.2012.08.011.

[19] L.A. Gaither, J. Borawski, L.J. Anderson, K.A. Balabanis, P. Devay, G. Joberty, C. Rau, M. Schirle, T. Bouwmeester, C. Mickanin, S. Zhao, C. Vickers, L. Lee, G. Deng, J. Baryza, R.A. Fujimoto, K. Lin, T. Compton, B. Wiedmann, Multiple cyclophilins involved in different cellular pathways mediate HCV replication, Virology. 397 (2010) 43–55. doi:10.1016/j.virol.2009.10.043.

[20] P.A. Gallay, Cyclophilin inhibitors: A novel class of promising host-targeting anti-HCV agents, Immunol. Res. 52 (2012) 200–210. doi:10.1007/s12026-011-8263-5.

[21] S. Hopkins, P. Gallay, Cyclophilin inhibitors: An emerging class of therapeutics for the treatment of chronic hepatitis C infection, Viruses. 4 (2012) 2558–2577. doi:10.3390/v4112558.

[22] C. Liu, J.R. Perilla, J. Ning, M. Lu, G. Hou, R. Ramalho, B.A. Himes, G. Zhao, G.J.

bioRxiv preprint doi: https://doi.org/10.1101/136101; this version posted May 9, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Bedwell, I.-J. Byeon, J. Ahn, A.M. Gronenborn, P.E. Prevelige, I. Rousso, C. Aiken, T. Polenova, K. Schulten, P. Zhang, Cyclophilin A stabilizes the HIV-1 capsid through a novel non-canonical binding site., Nat. Commun. 7 (2016) 10714. doi:10.1038/ncomms10714.

[23] A. De Iaco, J. Luban, Cyclophilin A promotes HIV-1 reverse transcription but its effect on transduction correlates best with its effect on nuclear entry of viral cDNA, Retrovirology. 11 (2014) 11. doi:10.1186/1742-4690-11-11.

[24] J. DeBoer, C.J. Madson, M. Belshan, Cyclophilin B enhances HIV-1 infection, Virology. 489 (2016) 282–291. doi:10.1016/j.virol.2015.12.015.

[25] Q. Yao, M. Li, H. Yang, H. Chai, W. Fisher, C. Chen, Roles of cyclophilins in cancers and other organ systems, in: World J. Surg., 2005: pp. 276–280. doi:10.1007/s00268-004-7812-7.

[26] J. Lee, S. Kim, Current implications of cyclophilins in human cancers, J. Exp. Clin. Cancer Res. 29 (2010) 97. doi:10.1186/1756-9966-29-97.

[27] P.T.M. Lavin, M.M. Mc Gee, Cyclophilin function in Cancer; lessons from virus replication., Curr. Mol. Pharmacol. 9 (2015) 148–64. doi:10.2174/1874467208666150519115443.

[28] H. Fliri, Neurodegenerative diseases: The potential of cyclophilin inhibition, Drug Target Rev. (2016).

[29] H. Du, L. Guo, F. Fang, D. Chen, A.A. Sosunov, G.M. McKhann, Y. Yan, C. Wang, H. Zhang, J.D. Molkentin, F.J. Gunn-Moore, J.P. Vonsattel, O. Arancio, J.X. Chen, S. Du Yan, Cyclophilin D deficiency attenuates mitochondrial and neuronal perturbation and ameliorates learning and memory in Alzheimer's disease., Nat. Med. 14 (2008) 1097–105. doi:10.1038/nm.1868.

[30] R. Banerjee, A.A. Starkov, M.F. Beal, B. Thomas, Mitochondrial dysfunction in the limelight of Parkinson's disease pathogenesis., Biochim. Biophys. Acta. 1792 (2009) 651–63. doi:10.1016/j.bbadis.2008.11.007.

[31] J. Kallen, V. Mikol, P. Taylor, M.D. Walkinshaw, X-ray structures and analysis of 11

bioRxiv preprint doi: https://doi.org/10.1101/136101; this version posted May 9, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

cyclosporin derivatives complexed with cyclophilin A., J. Mol. Biol. 283 (1998) 435–449. doi:10.1006/jmbi.1998.2108.

[32] A. Scribner, D. Houck, Z. Huang, S. Mosier, M. Peel, B. Scorneaux, Synthesis and biological evaluation of [D-lysine]8cyclosporin A analogs as potential anti-HCV agents, Bioorganic Med. Chem. Lett. 20 (2010) 6542–6546. doi:10.1016/j.bmcl.2010.09.036.

[33] C.J. Dunsmore, K.J. Malone, K.R. Bailey, M.A. Wear, H. Florance, S. Shirran, P.E. Barran, A.P. Page, M.D. Walkinshaw, N.J. Turner, Design and Synthesis of Conformationally Constrained Cyclophilin Inhibitors Showing a Cyclosporin-A Phenotype in C. elegans, ChemBioChem. 12 (2011) 802–810. doi:10.1002/cbic.201000413.

[34] Y. Yang, E. Moir, G. Kontopidis, P. Taylor, M.A. Wear, K. Malone, C.J. Dunsmore, A.P. Page, N.J. Turner, M.D. Walkinshaw, Structure-based discovery of a family of synthetic cyclophilin inhibitors showing a cyclosporin-A phenotype in Caenorhabditis elegans, Biochem. Biophys. Res. Commun. 363 (2007) 1013–1019. doi:10.1016/j.bbrc.2007.09.079.

[35] S. Daum, M. Schumann, S. Mathea, T. Aumuller, M.A. Balsley, S.L. Constant, B.F. De Lacroix, F. Kruska, M. Braun, C. Schiene-Fischer, Isoform-specific inhibition of cyclophilins, Biochemistry. 48 (2009) 6268–6277. doi:10.1021/bi9007287.

[36] J. Li, J. Zhang, J. Chen, X. Luo, W. Zhu, J. Shen, H. Liu, X. Shen, H. Jiang, Strategy for discovering chemical inhibitors of human cyclophilin A: Focused library design, virtual screening, chemical synthesis and bioassay, J. Comb. Chem. 8 (2006) 326–337. doi:10.1021/cc0501561.

[37] K. Chen, Z. Tan, M. He, J. Li, S. Tang, I. Hewlett, F. Yu, Y. Jin, M. Yang, Structure-activity relationships (SAR) research of thiourea derivatives as dual inhibitors targeting both HIV-1 capsid and human cyclophilin A, Chem. Biol. Drug Des. 76 (2010) 25–33. doi:10.1111/j.1747-0285.2010.00981.x.

[38] A. Ahmed-Belkacem, L. Colliandre, N. Ahnou, Q. Nevers, M. Gelin, Y. Bessin, R.

bioRxiv preprint doi: https://doi.org/10.1101/136101; this version posted May 9, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Brillet, O. Cala, D. Douguet, W. Bourguet, I. Krimm, J.-M. Pawlotsky, J.-F. Guichou, Fragment-based discovery of a new family of non-peptidic small-molecule cyclophilin inhibitors with potent antiviral activities, Nat. Commun. 7 (2016) 12777. doi:10.1038/ncomms12777.

[39] J.-F. Guichou, J. Viaud, C. Mettling, G. Subra, Y.-L. Lin, A. Chavanieu, Structure-based design, synthesis, and biological evaluation of novel inhibitors of human cyclophilin A., J. Med. Chem. 49 (2006) 900–10. doi:10.1021/jm050716a.

[40] S. Ni, Y. Yuan, J. Huang, X. Mao, M. Lv, J. Zhu, X. Shen, J. Pei, L. Lai, H. Jiang, J. Li, Discovering potent small molecule inhibitors of cyclophilin A using de novo drug design approach, J. Med. Chem. 52 (2009) 5295–5298. doi:10.1021/jm9008295.

[41] B.M. Dunyak, J.E. Gestwicki, Peptidyl-Proline Isomerases (PPIases): Targets for Natural Products and Natural Product-Inspired Compounds, J. Med. Chem. 59 (2016) 9622–9644. doi:10.1021/acs.jmedchem.6b00411.

[42] M.A. Wear, M.W. Nowicki, E.A. Blackburn, I.W. Mcnae, M.D. Walkinshaw, Thermo-kinetic analysis space expansion for cyclophilin-ligand interactions – identification of a new non- peptide inhibitor using Biacore™ T200, FEBS Open Bio. (2017) 0–3. doi:10.1002/1878-0261.12201.

[43] J. Michel, N. Foloppe, J.W. Essex, Rigorous free energy calculations in structure-based drug design, Mol. Inform. 29 (2010) 570–578. doi:10.1002/minf.201000051.

[44] S. Bosisio, A.S.J.S. Mey, J. Michel, Blinded predictions of host-guest standard free energies of binding in the SAMPL5 challenge, J. Comput. Aided. Mol. Des. (2016) 1–10. doi:10.1007/s10822-016-9933-0.

[45] A.S.J.S. Mey, J. Juárez-Jiménez, A. Hennessy, J. Michel, Blinded predictions of binding modes and energies of HSP90-α ligands for the 2015 D3R grand challenge., Bioorg. Med. Chem. (2016). doi:10.1016/j.bmc.2016.07.044.

[46] M. Gelin, V. Delfosse, F. Allemand, F. Hoh, Y. Sallaz-Damaz, M. Pirocchi, W. Bourguet, J.-L. Ferrer, G. Labesse, J.-F. Guichou, Combining `dry' co-crystallization and it in situ diffraction to facilitate ligand screening by X-ray crystallography, Acta

bioRxiv preprint doi: https://doi.org/10.1101/136101; this version posted May 9, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Crystallogr. Sect. D. 71 (2015) 1777–1787. doi:10.1107/S1399004715010342.

[47] T.G.G. Battye, L. Kontogiannis, O. Johnson, H.R. Powell, A.G.W. Leslie, iMOSFLM: A new graphical interface for diffraction-image processing with MOSFLM, Acta Crystallogr. Sect. D Biol. Crystallogr. 67 (2011) 271–281. doi:10.1107/S0907444910048675.

[48] P. Evans, Scaling and assessment of data quality, in: Acta Crystallogr. Sect. D Biol. Crystallogr., 2006: pp. 72–82. doi:10.1107/S0907444905036693.

[49] M.D. Winn, C.C. Ballard, K.D. Cowtan, E.J. Dodson, P. Emsley, P.R. Evans, R.M. Keegan, E.B. Krissinel, A.G.W. Leslie, A. McCoy, S.J. McNicholas, G.N. Murshudov, N.S. Pannu, E.A. Potterton, H.R. Powell, R.J. Read, A. Vagin, K.S. Wilson, Overview of the CCP4 suite and current developments, Acta Crystallogr. Sect. D Biol. Crystallogr. 67 (2011) 235–242. doi:10.1107/S0907444910045749.

[50] G.N. Murshudov, A.A. Vagin, E.J. Dodson, Refinement of macromolecular structures by the maximum-likelihood method, Acta Crystallogr. Sect. D Biol. Crystallogr. 53 (1997) 240–255. doi:10.1107/S0907444996012255.

[51] G.N. Murshudov, P. Skubák, A.A. Lebedev, N.S. Pannu, R.A. Steiner, R.A. Nicholls, M.D. Winn, F. Long, A.A. Vagin, REFMAC5 for the refinement of macromolecular crystal structures, Acta Crystallogr. Sect. D Biol. Crystallogr. 67 (2011) 355–367. doi:10.1107/S0907444911001314.

[52] P. Emsley, B. Lohkamp, W.G. Scott, K. Cowtan, Features and development of Coot, Acta Crystallogr. Sect. D Biol. Crystallogr. 66 (2010) 486–501. doi:10.1107/S0907444910007493.

[53] C.M. Stegmann, D. Seeliger, G.M. Sheldrick, B.L. De Groot, M.C. Wahl, The thermodynamic influence of trapped water molecules on a protein-ligand interaction, Angew. Chemie - Int. Ed. 48 (2009) 5207–5210. doi:10.1002/anie.200900481.

[54] Schrödinger LLC, Schrödinger Release 2013-3: Maestro, version 9.6, (2013).

[55] G. Madhavi Sastry, M. Adzhigirey, T. Day, R. Annabhimoju, W. Sherman, Protein and ligand preparation: Parameters, protocols, and influence on virtual screening

bioRxiv preprint doi: https://doi.org/10.1101/136101; this version posted May 9, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

enrichments, J. Comput. Aided. Mol. Des. 27 (2013) 221–234. doi:10.1007/s10822-013-9644-8.

[56] M.M. Reif, C. Oostenbrink, Net charge changes in the calculation of relative ligand-binding free energies via classical atomistic molecular dynamics simulation, J. Comput. Chem. 35 (2014) 227–243. doi:10.1002/jcc.23490.

[57] G.J. Rocklin, D.L. Mobley, K.A. Dill, P.H. Hunenberger, Calculating the binding free energies of charged species based on explicit-solvent simulations employing lattice-sum methods: An accurate correction scheme for electrostatic finite-size effects, J. Chem. Phys. 139 (2013). doi:10.1063/1.4826261.

[58] Schrödinger LLC, The PyMOL Molecular Graphics System, Version 1.4.1, (n.d.).

[59] O. Trott, A.J. Olson, AutoDock Vina, J. Comput. Chem. 31 (2010) 445–461. doi:10.1002/jcc.21334.

[60] H.H. Loeffler, J. Michel, C. Woods, FESetup: Automating Setup for Alchemical Free Energy Simulations, J. Chem. Inf. Model. 55 (2015) 2485–2490. doi:10.1021/acs.jcim.5b00368.

[61] J.A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K.E. Hauser, C. Simmerling, ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB, J. Chem. Theory Comput. 11 (2015) 3696–3713. doi:10.1021/acs.jctc.5b00255.

[62] J. Wang, R.M. Wolf, J.W. Caldwell, P.A. Kollman, D.A. Case, Development and testing of a general Amber force field, J. Comput. Chem. 25 (2004) 1157–1174. doi:10.1002/jcc.20035.

[63] J. Wang, W. Wang, P.A. Kollman, D.A. Case, Automatic atom type and bond type perception in molecular mechanical calculations, J. Mol. Graph. Model. 25 (2006) 247–260. doi:10.1016/j.jmgm.2005.12.005.

[64] A. Jakalian, D.B. Jack, C.I. Bayly, Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation, J. Comput. Chem. 23 (2002) 1623–1641. doi:10.1002/jcc.10128.

bioRxiv preprint doi: https://doi.org/10.1101/136101; this version posted May 9, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

[65] A. Jakalian, B.L. Bush, D.B. Jack, C.I. Bayly, Fast, efficient generation of high-quality atomic charges. AM1-BCC model: I. Method, J. Comput. Chem. 21 (2000) 132–146. doi:10.1002/(sici)1096-987x(20000130)21:2<132::aid-jcc5>3.0.co;2-p.

[66] M.R. Shirts, J.D. Chodera, K.R. Beauchamp, pymbar v2.1.0-beta, (2014). https://github.com/choderalab/pymbar.

[67] D.L. Mobley, J.D. Chodera, K.A. Dill, On the use of orientational restraints and symmetry corrections in alchemical free energy calculations., J. Chem. Phys. 125 (2006) 084902. doi:10.1063/1.2221683.

[68] J. Woods, C.; Calabro, C.; Michel, www.siremol.org, (n.d.). www.siremol.org (accessed April 29, 2016).

[69] P. Eastman, V. Pande, OpenMM: A Hardware-Independent Framework for Molecular Simulations, Comput. Sci. Eng. 12 (2010) 34–39. doi:10.1109/MCSE.2010.27.

[70] M.R. Shirts, J.D. Chodera, Statistically optimal analysis of samples from multiple equilibrium states, J. Chem. Phys. 129 (2008). doi:10.1063/1.2978177.

[71] J. Michel, M.L. Verdonk, J.W. Essex, Protein-ligand complexes: Computation of the relative free energy of different scaffolds and binding modes, J. Chem. Theory Comput. 3 (2007) 1645–1655. doi:10.1021/ct700081t.

[72] C.W. Hopkins, S. Le Grand, R.C. Walker, A.E. Roitberg, Long-time-step molecular dynamics through hydrogen mass repartitioning, J. Chem. Theory Comput. 11 (2015) 1864–1874. doi:10.1021/ct5010406.

[73] H.C. Andersen, Molecular dynamics simulations at constant pressure and/or temperature, J. Chem. Phys. 72 (1980) 2384. doi:10.1063/1.439486.

bioRxiv preprint doi: https://doi.org/10.1101/136101; this version posted May 9, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Table captions

Table 1. Fragments that bind to CypA based on X-ray crystallography data.

Table 2 Summary of structural data for ligand **98**.

bioRxiv preprint doi: https://doi.org/10.1101/136101; this version posted May 9, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Figure legends

Fig.1. 3D surface structure of CypA colour coded according to degree of residue conservation across isoforms A, B, C, D, F (blue – strictly conserved, red – different in every isoform). Abu and Pro pockets and key active side residues are also highlighted.

Fig.2. Histogram of calculated binding free energy of fragments to Abu and Pro pockets of CypA. Compounds are binned (bin width 0.5 kcal mol^{-1}) based on their calculated K_d (bottom *x*-axis) and calculated $\Delta G°_{bind}$ values (top *x*-axis). The coloured numbers indicate the total number of compounds in each bin. The calculated binding energies for the Abu and Pro pockets are depicted in blue and red respectively. The SPR range of K_d estimates for fragments bound to CypA is depicted by the green arrow.

Fig.3. Binding of **98** in the Abu pocket of CypA. **A)** 3D surface structure of **98** in the Abu pocket of CypA. **B)** The inset shows a *Fo – Fc* omit map of **98** in the Abu pocket CypA. A 3σ contour is shown in green. H-bonds between ligand and water molecules are highlighted and colour coded as; W1 in violet-purple, W2 in sky blue W3 in green, W4 in light pink and W5 in orange (distances in Angstrom). H-bond interactions between water molecules and CypA residues are also highlighted in black.

Fig.4. Binding of fragments out of the 80's loop of CypA. **A)** Overlay of **3**, **5**, **56** and **97** bound to a 3D surface representation of CypA. A symmetric CypA structure can be seen in pink. **B)** The inset shows interactions between two alternative conformations of **3** (light orange and light pink) with Asn106 and Glu84 residues of CypA and a close van der Waals interactions with Trp121 in the symmetric CypA unit. A *Fo – Fc* omit map of **3** contoured at 3σ is shown in green.

bioRxiv preprint doi: https://doi.org/10.1101/136101; this version posted May 9, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Fig.5. Bound poses of **16**, **60** and **61** in the Pro pocket of CypA. **A)** Overlay of **16**, **60** and **16** bound to Cyp A (3D surface representation). The three other panels depict interactions of **16**, **60** and **61** with Arg55, Gln63 and Asn102 residues of CypA in the Pro pocket. *Fo – Fc* omit maps contoured at 3σ are shown in green. Distances between ligands and CypA residues are highlighted in black and given in Angstrom.

Fig.6. Histogram of RMSD of **3** from the X-ray crystallographic pose A computed from MD simulation trajectories.

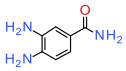
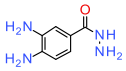
Fig.7. Representative conformations of ligand **3** in MD pose A sampled by the MD simulations. Protein residues or water molecules that are involved in H-bonding with **3** are colour coded separately. The distances between hydrogen bond donors and acceptors are given in Angstrom.

Fig.8. Representative conformation of ligand **3** in MD pose B (**A**) and MD pose C (**B**). Protein residues or water molecules that are involved in H-bonding with **3** are colour coded separately. The distances between hydrogen bond donors and acceptors are given in Angstrom.

Fig.9. Venn diagram showing the number of hit compounds produced by X-ray crystallography and molecular dynamics techniques.

bioRxiv preprint doi: https://doi.org/10.1101/136101; this version posted May 9, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

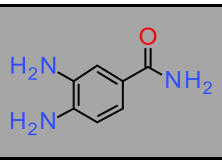
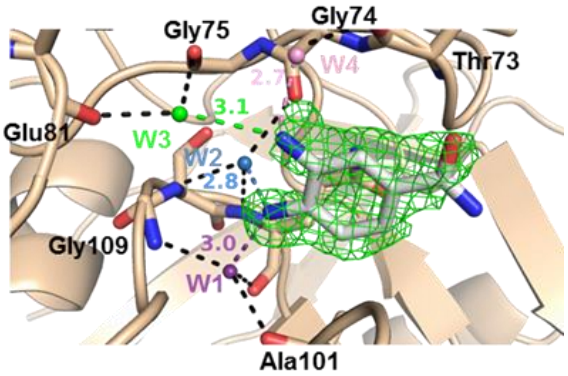
Table 1

Abu pocket			
Ligand Number	Structure	Ligand Number	Structure
3		89	
5		97	
56		98	
		99	

Pro pocket			
Ligand Number	Structure	Ligand Number	Structure
16		61	
60			

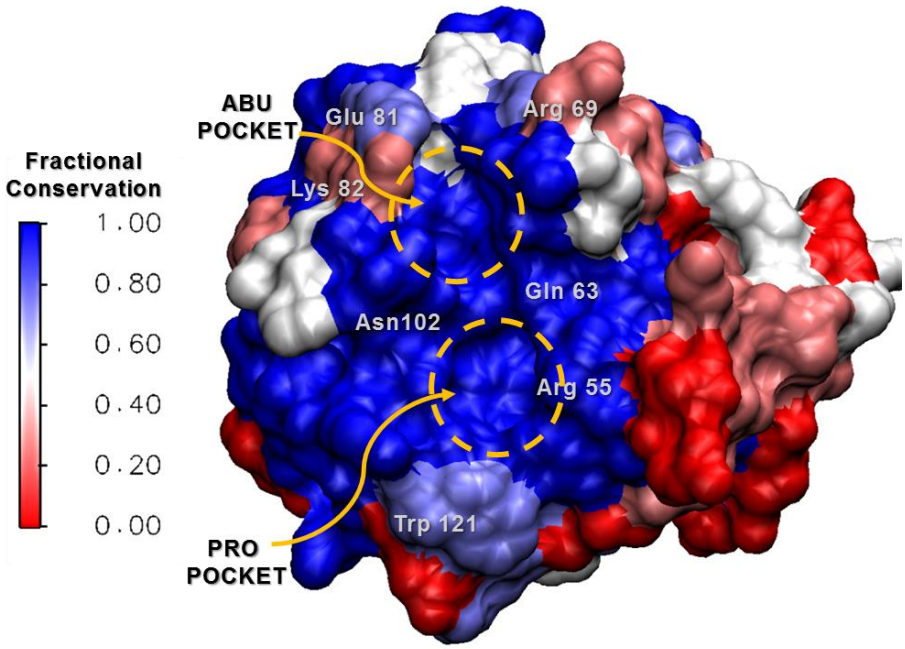
bioRxiv preprint doi: https://doi.org/10.1101/136101; this version posted May 9, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Table 2

Ligand 98 / Technique	
X-ray	
MD (ΔG^{o}_{bind})	-3.7 ± 0.4 kcal mol^{-1}
Ligand efficiency	0.34 ± 0.04

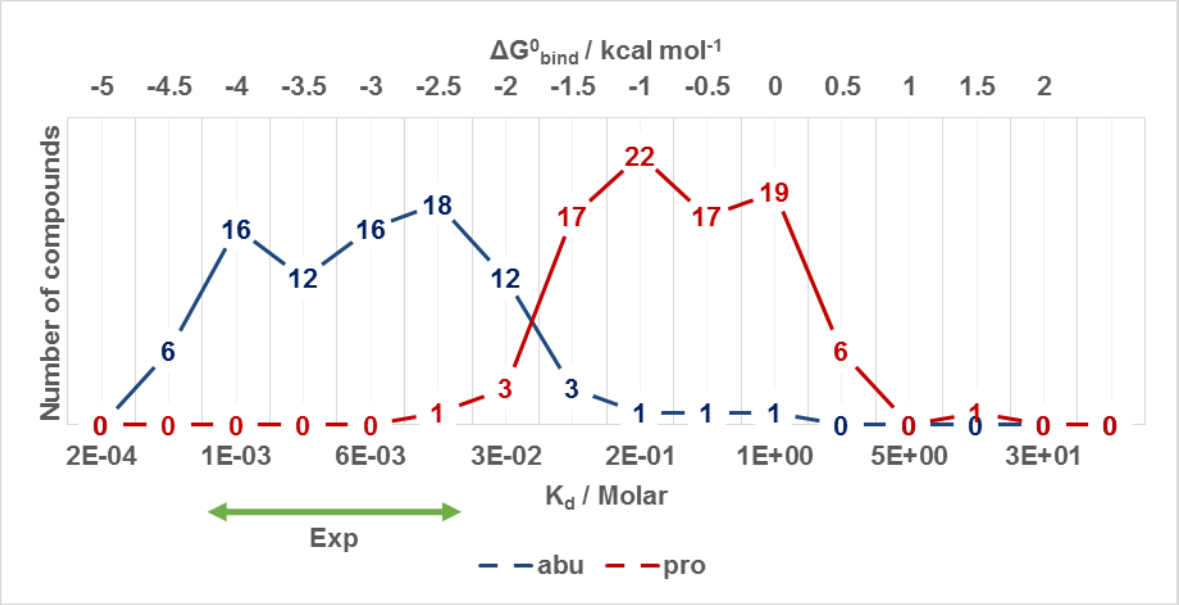
bioRxiv preprint doi: https://doi.org/10.1101/136101; this version posted May 9, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Fig.1



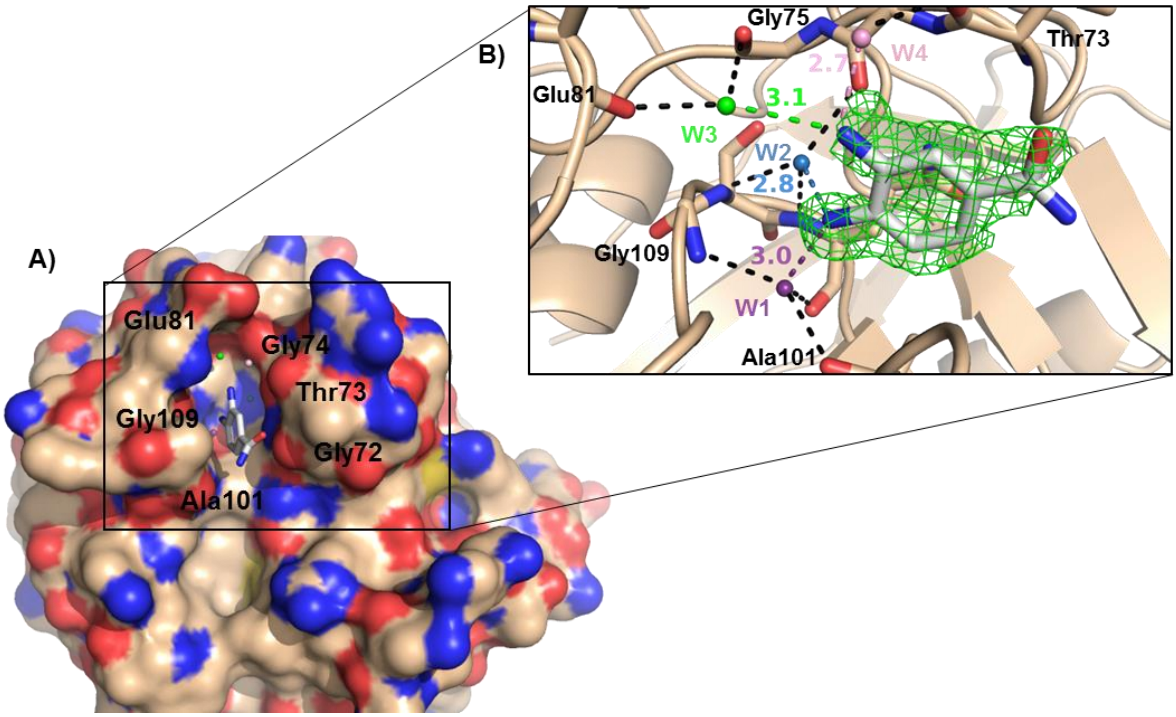
bioRxiv preprint doi: https://doi.org/10.1101/136101; this version posted May 9, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Fig.2



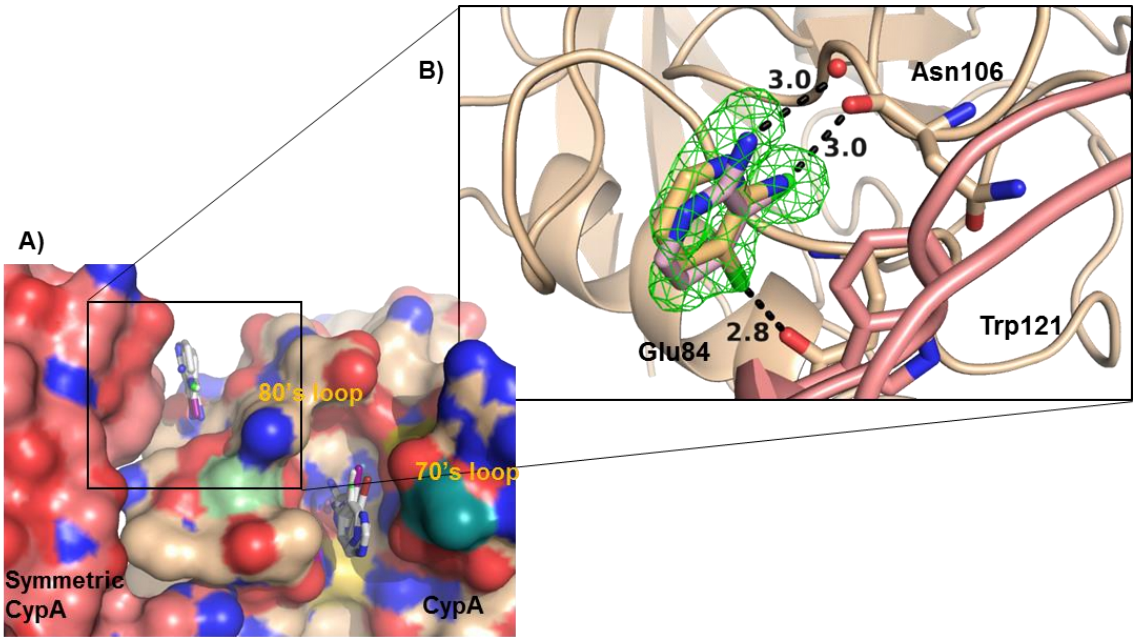
bioRxiv preprint doi: https://doi.org/10.1101/136101; this version posted May 9, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Fig.3



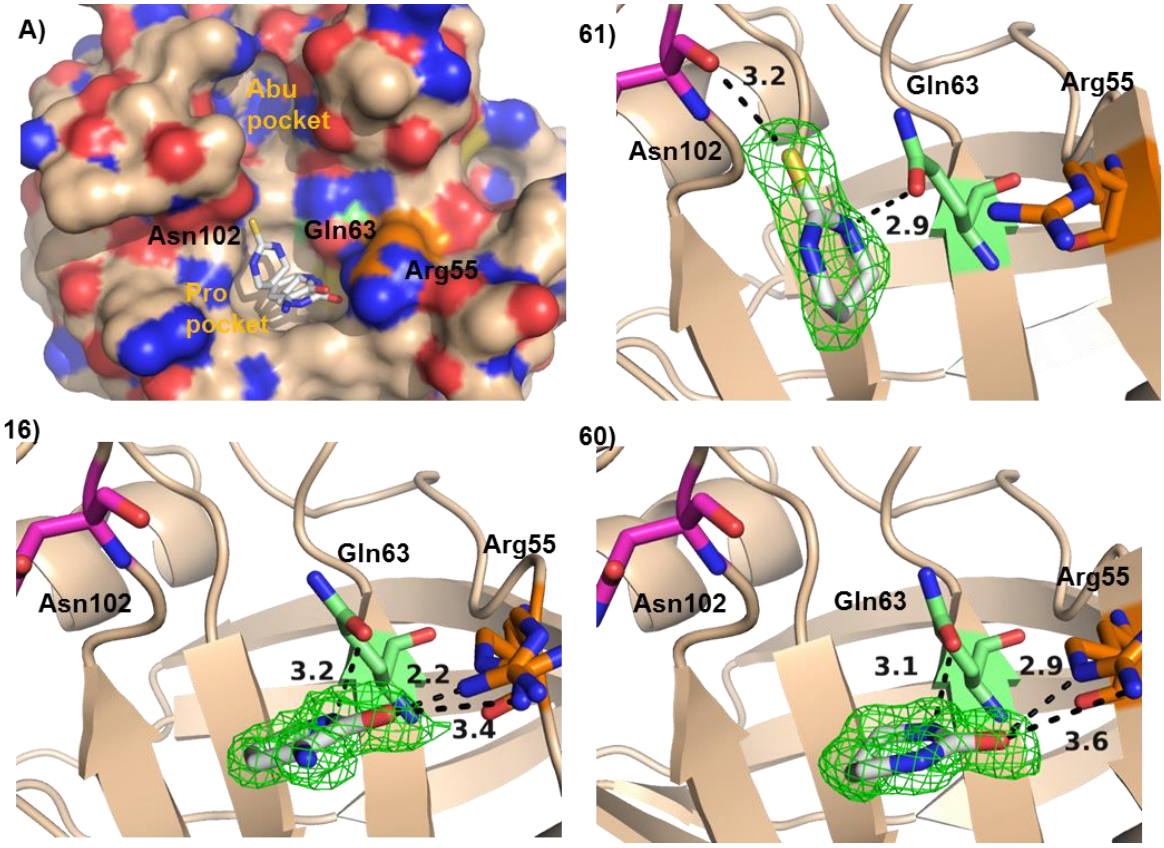
bioRxiv preprint doi: https://doi.org/10.1101/136101; this version posted May 9, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Fig.4



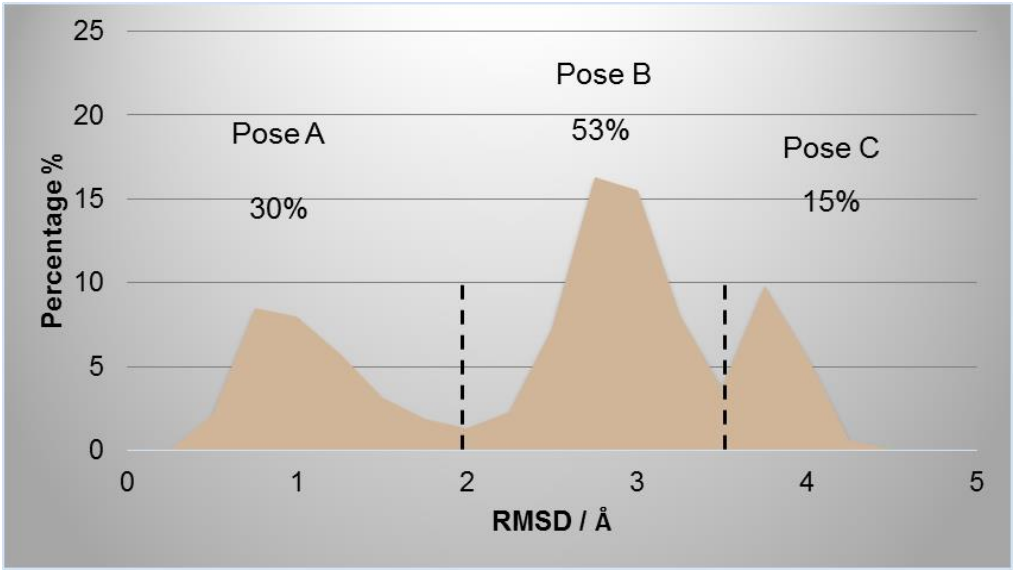
bioRxiv preprint doi: https://doi.org/10.1101/136101; this version posted May 9, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Fig.5



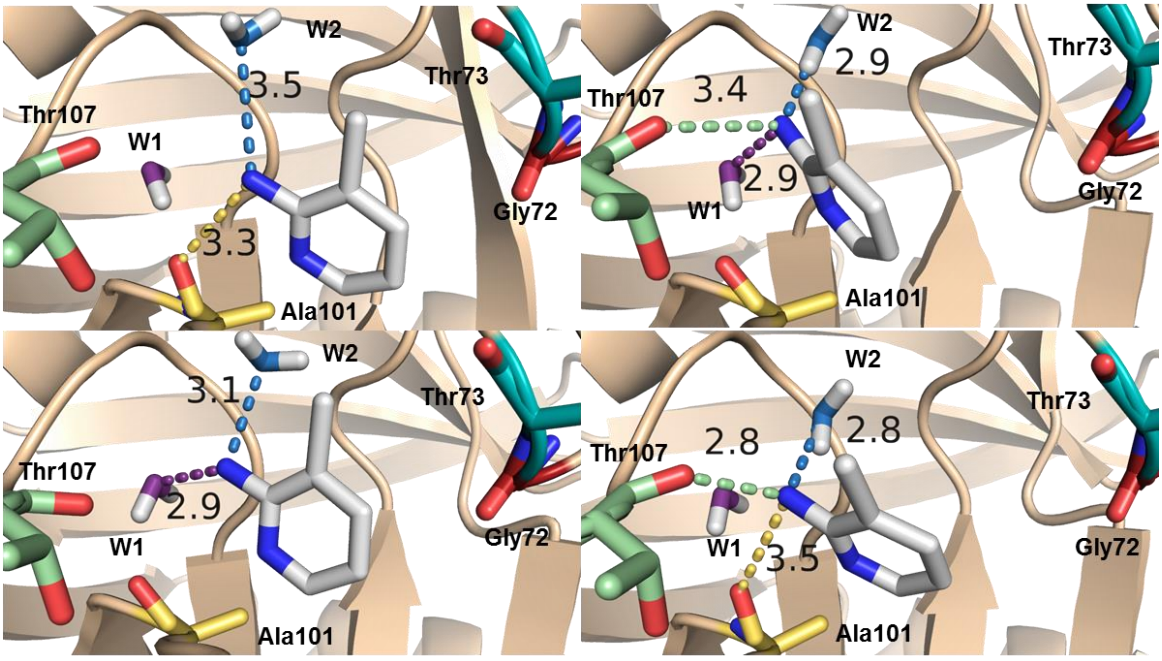
bioRxiv preprint doi: https://doi.org/10.1101/136101; this version posted May 9, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Fig.6



bioRxiv preprint doi: https://doi.org/10.1101/136101; this version posted May 9, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

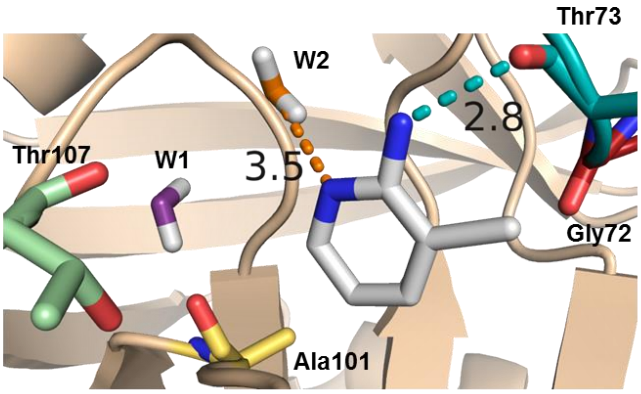
Fig.7



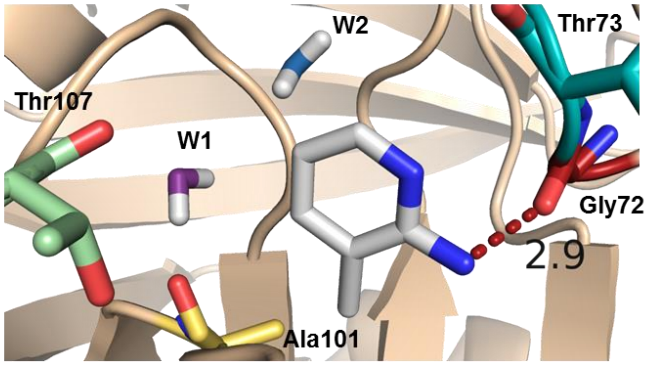
bioRxiv preprint doi: https://doi.org/10.1101/136101; this version posted May 9, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Fig.8

A)

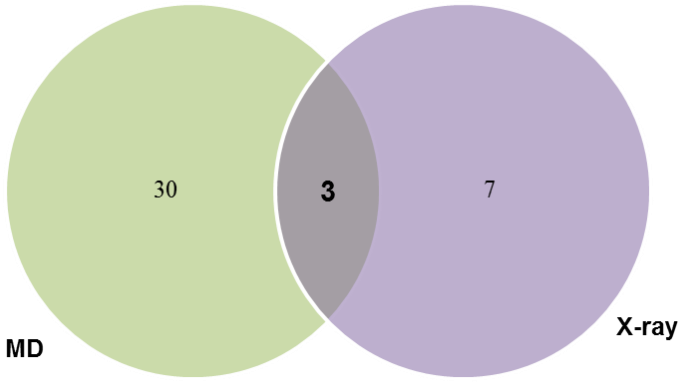


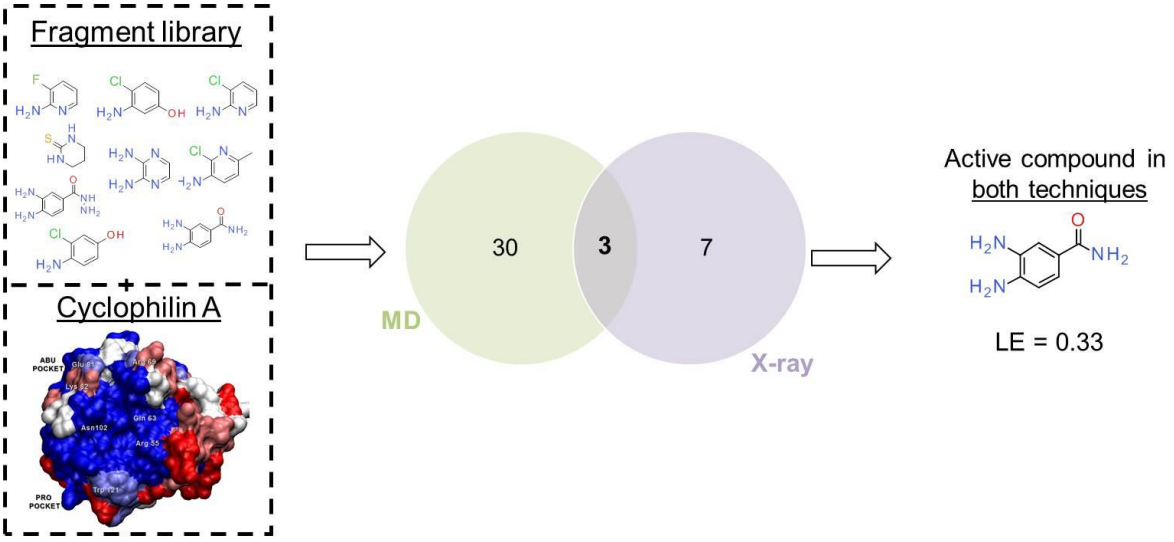
B)



bioRxiv preprint doi: https://doi.org/10.1101/136101; this version posted May 9, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Fig.9



bioRxiv preprint doi: https://doi.org/10.1101/136101; this version posted May 9, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Graphical Abstract