bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516467; this version posted February 22, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity.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-NC 4.0 International license.

Discovery and validation of the binding poses of allosteric fragment hits to PTP1b: From molecular dynamics simulations to X-ray crystallography

Jack B. Greisman,[1,*] Lindsay Willmore,[1,*] Christine Y. Yeh,[1,*] Fabrizio Giordanetto,[1] Sahar Shahamadtar,[1] Hunter Nisonoff,[1] Paul Maragakis,[1,†] and David E. Shaw[1,2,†]

[1] D. E. Shaw Research, New York, NY 10036, USA.

[2] Department of Biochemistry and Molecular Biophysics, Columbia University, New York, NY 10032, USA.

[*] These authors contributed equally to the manuscript.

† To whom correspondence should be addressed.

Paul Maragakis

E-mail: Paul.Maragakis@DEShawResearch.com

Phone: (212) 478-0414

Fax: (212) 845-1414

David E. Shaw

E-mail: David.Shaw@DEShawResearch.com

Phone: (212) 478-0260

Fax: (212) 845-1286

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516467; this version posted February 22, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity.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-NC 4.0 International license.

Abstract

Fragment-based drug discovery has led to six approved drugs, but the small size of the chemical fragments used in such methods typically results in only weak interactions between the fragment and its target molecule, which makes it challenging to experimentally determine the three-dimensional poses fragments assume in the bound state. One computational approach that could help address this difficulty is long-timescale molecular dynamics (MD) simulation, which has been used in retrospective studies to recover experimentally known binding poses of fragments. Here, we present the results of long-timescale MD simulations that we used to prospectively discover binding poses for two series of fragments in allosteric pockets on a difficult and important pharmaceutical target, protein-tyrosine phosphatase 1b (PTP1b). Our simulations reversibly sampled the fragment association and dissociation process. One of the binding pockets found in the simulations has not to our knowledge been previously observed with a bound fragment, and the other pocket adopted a very rare conformation. We subsequently obtained high-resolution crystal structures of members of each fragment series bound to PTP1b, and the experimentally observed poses confirmed the simulation results. To the best of our knowledge, our findings provide the first demonstration that MD simulations can be used prospectively to determine fragment binding poses to previously unidentified pockets.

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516467; this version posted February 22, 2023. 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-NC 4.0 International license.

Fragment-based drug discovery (FBDD) aims to identify small molecules of low molecular weight (fragments) that bind a biomolecule of interest. Such fragments typically bind weakly themselves, but are used as starting points for the design of more potent compounds. FBDD often involves the experimental evaluation of binding for many fragments (a process known as "fragment screening"). Many pharmaceutical companies use fragment screening in their early-stage drug discovery pipelines to identify new chemical matter, complement high-throughput screening, or refine the understanding of a particular binding site.[1,2] Such laboratory fragment screens commonly make use of biophysical assays, including nuclear magnetic resonance spectroscopy (NMR), X-ray crystallography, and surface-plasmon resonance (SPR)[1,2] to identify fragments that bind their intended target (hits). These efforts have already resulted in six approved drugs that were designed starting from a fragment hit.[3]

To validate a fragment hit and improve its drug-like properties, it is often important to understand the binding site and pose of the fragment.[1,2] It remains challenging, however, to do so using experimental means: By definition, fragments have low molecular weight (typically <300 Da) and associate with their target with millimolar to micromolar affinity,[1,2] and the various biophysical assays used to identify fragment hits typically have low signal-to-noise ratios for low-molecular-weight or weak binders, meaning that these techniques often give inconclusive, and in some cases, contradictory results about which fragments interact with the target. Indeed, a requirement that two biophysical techniques agree with each other often results in the elimination of the majority of all fragment hits.[4]

Molecular dynamics (MD) simulation can model the motions of all the atoms of a biomolecular system of interest, and is thus a potentially powerful tool to identify the binding modes of fragment hits. The most direct MD approach is to simulate the pharmaceutical target of interest and the surrounding solvent containing one or more fragments, in order to observe the diffusion

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516467; this version posted February 22, 2023. 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-NC 4.0 International license.

of the fragments in solution and their eventual association with the target. Such an approach is a potentially powerful one for identifying binding modes, because it does not require any a priori knowledge of the bound conformation, in particular the location of the binding site, information which is often unavailable in practice. MD simulations have already been used for the related problem of identifying new interactions that could be used to enhance ligand-binding affinity,[5,6] and several MD simulation studies have been used to reproduce the previously known crystallographic binding poses of ligands to their target proteins.[7–15] To date, however, we are not aware of any published MD simulation study that has predicted the pose of a ligand in a previously unknown binding site, and that has subsequently been verified by X-ray crystallography experiments.[16]

Here, we present the results of long-timescale, all-atom MD simulations that we used to discover the binding sites and poses of two fragments to protein-tyrosine phosphatase 1b (PTP1b), an important, but difficult pharmaceutical target. Prior to our simulations, no structures were available for either of these fragments bound to PTP1b. In our simulations, we observed repeated binding to and unbinding from their respective bound poses in two distinct allosteric sites (pockets 1 and 2). For the fragment that bound in pocket 2, we observed the rearrangement within the pocket of two phenylalanine side chains adjacent to the fragment that led to a PTP1b conformation rarely observed in crystal structures. We then obtained high-resolution crystal structures of PTP1b bound to each fragment and found that these crystal structures confirmed the poses observed in simulation. To provide additional evidence that these binding poses are not artifacts, we synthesized fragments that were chemically closely related to the hits and obtained high-resolution crystal structures of PTP1b bound to these related fragments in poses that overlapped with the poses of the original hits. These results demonstrate, for the first time, MD simulations being used in the successful prediction of the previously unknown binding sites and poses of fragments to their target proteins.

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516467; this version posted February 22, 2023. 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-NC 4.0 International license.

PTP1b—a negative regulator of the insulin and leptin pathways and of unfolded protein response in the endoplasmic reticulum—has been a prototypical single-domain protein for experimentation with technologies for allosteric-binding-site discovery. It was previously considered a potentially attractive target for treating type II diabetes and obesity, though concerns regarding skin lesions in animal models, and difficulties encountered in finding potent and selective compounds that are active in the cell, have limited recent efforts in PTP1b drug discovery.[17–19] PTP1b is also involved in the development of Her2-positive breast cancers, and a product derived from dogfish liver that is currently an early-stage clinical candidate for this indication is believed to act on PTP1b by an allosteric mechanism.[20] Because of the conserved and very polar nature of the active site of PTP1b, it is widely believed that only allosteric inhibitors of PTP1b have a realistic chance of becoming useful drugs, and for this reason PTP1b has become a workhorse for methods that attempt to elucidate allosteric pockets.[21–27]

As part of our in-house drug discovery efforts, we ran an SPR screen of two 384-well plates of the Maybridge library (Thermo Fisher Scientific) and identified two fragments (referred to here as DES-4799 and DES-4884) that showed stronger binding to PTP1b than to five off-target proteins that are structurally unrelated to PTP1b (see Supplementary Information). We then used MD simulations of these two fragments in the presence of PTP1b to predict their binding sites and poses. The fragments were not subject to any artificial biasing forces, and were free to diffuse about the simulation box and sample potential binding modes with PTP1b. (See SI for further details regarding the setup of these MD simulations.)

DES-4799, an anilinic fragment, bound reversibly in our simulations to pocket 1, a novel allosteric binding site located diametrically opposite from the active site. Fig. 1 shows that this aniline reversibly associated with a small number of sites on the surface of the protein over the course of a 100-μs MD simulation (see Movie S1). At the fragment concentration in this

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516467; this version posted February 22, 2023. 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-NC 4.0 International license.

simulation, the fragment was unbound for the majority of the simulation time, suggesting an affinity >3 mM. The single dominant binding site was defined by residues Ser80, Ser205, His208, and Gly209; the conformation of these residues was stabilized upon binding (Fig. S3). The location of this binding site, along with those of three additional minor binding sites, are highlighted in Fig. 1b. In the main site, the aniline bound in three different orientations (Fig 1c), with a strong preference for one of the three poses (see histogram in Fig 1a).

DES-4884, a piperazinylquinoline fragment, bound in our simulations to pocket 2, which overlaps with the allosteric pocket discovered by Sunesis.[22] In our simulations, however, binding of the fragment involved a rearrangement of aromatic side chains not previously observed in PTP1b X-ray structures. Fig. 2 shows that DES-4884 reversibly bound to its main binding site several times in two different orientations. It also visited a small number of other binding sites, but spent at least an order of magnitude more time at the primary binding site than at any of these secondary sites (see Movie S2). The simulated binding affinity to the main site was ~1 mM. Fig. 2b shows the locations of the secondary PTP1b binding sites visited by the piperazinylquinoline fragment; the site colored in orange overlaps with a pocket recently identified by Keedy et al.[23]

Fig. 2c shows the two distinct binding orientations of DES-4884. In the dominant orientation, the piperazine formed a hydrogen bond with Asn193. In both orientations, the pocket rearranged to a conformation distinct from most previously reported crystal structures: Both Phe280 (top right) and Phe196 swung out to open a pocket that allowed for hydrophobic contacts with the quinoline core (Fig. 2d). Phe196 does not adopt the swung-out configuration in the vast majority of crystal structures published to date (Fig. 2e). The only published structures in which Phe196 adopts the swung-out configuration can be explained by crystal contacts,[24] nearby mutations,[28] or fragment binding.[23] In simulation, the side chain remained swung out while the fragment was

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516467; this version posted February 22, 2023. 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-NC 4.0 International license.

bound, and then rearranged to the dominant apo configuration while the fragment was not bound (Fig 2e) (see Movie S3).

Following binding characterization by MD simulation, we validated the predicted binding poses by obtaining high-resolution crystal structures of PTP1b in complex with each of the two fragments. To further validate the binding poses we identified for these fragments, we obtained additional high-resolution crystal structures of two variants of DES-4799 and one variant of DES-4884 that recapitulated their respective binding modes. Fig. 3 shows that the dominant binding poses from simulation very closely match the high-resolution crystal structures, including structurally important details such as the location of water molecules bound to each of the fragments.

The crystal structure that includes DES-4799 overlaps with the dominant bound pose in simulation (the heavy-atom RMSD of the pocket is 0.99 Å, and that of the fragment is 0.57 Å). The fragment pose observed in the crystal structure suggested two plausible directions in which the fragment could be extended, so we synthesized ethyl variants (DES-5742 and DES-5743) of the original DES-4799 fragment and obtained high-resolution X-ray structures. These structures reproduce the binding mode of the original aniline fragment.

The crystal structure that includes DES-4884 overlaps well with the structure of the dominant bound pose in MD simulation (the heavy-atom RMSD of the pocket residues is 0.49 Å, and the heavy-atom RMSD of the fragment is 0.70 Å). The crystal structure recovers the observed hydrogen bond to Asn193 and the reorientation of the Phe196 and Phe280 side chains, as seen in the MD simulations (Movie S3). We synthesized DES-6016, a hydroxymethyl derivative of the DES-4884 fragment that extends beyond the hydrogen bond to Asn193 in the direction covered by the Sunesis compound.[22] The high-resolution crystal structure of DES-6016 bound to PTP1b

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516467; this version posted February 22, 2023. 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-NC 4.0 International license.

has the core of the fragment in essentially the same pose as in the crystal structure of DES-4884, adding further experimental support to the DES-4884 binding mode predicted by simulation.

In this communication, we have demonstrated that MD simulation can discover binding poses of fragments at allosteric sites, which were previously unknown, of a challenging drug discovery target. Our simulations incorporated no *a priori* information about possible binding sites. The simulations were also able to sample possible conformational rearrangements of the fragments and the protein during the binding process. The two fragments bound at very different locations, and high-resolution crystal structures of each of these fragments and their variants strongly support the binding modes observed in simulation.

Significantly, one of the binding pockets was occluded by two aromatic side chains in the crystal structure used to initiate the MD simulations, and the open pocket that resulted from the rearrangement of these side chains is rarely observed in published crystal structures. Our results also show that MD simulation can be used to predict binding sites and poses even when fragments have very weak binding strength: The aniline fragment was computationally predicted to have a very weak affinity (>3 mM), and yet the dominant binding pose in our simulations matched well the pose in the subsequent crystal structure. Perhaps counterintuitively, long-timescale MD can be particularly useful when simulating weakly binding fragments, as the long simulations enable a sampling of many different binding modes and multiple reversible binding events at the dominant site.

MD simulations can serve as a tool for obtaining atomically detailed insights about binding modes of weak allosteric binders—a task that is difficult to achieve using traditional biophysical methods. One remaining uncertainty lies in the accuracy with which the underlying physical models used in such simulations (force fields) describe the protein and the small-molecule

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516467; this version posted February 22, 2023. 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-NC 4.0 International license.

fragments. With the progress that has been made in the development of force fields, this source of uncertainty has now become manageable for proteins with well-resolved folded-state structures and for fragment-sized small molecules. We anticipate that with the availability of increased computational power and further improvements in force field accuracy, the use of long-timescale, unbiased MD simulation will become increasingly common in fragment-hit validation.

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516467; this version posted February 22, 2023. 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-NC 4.0 International license.

References

1. Klon, A. E. Fragment-Based Methods in Drug Discovery; Methods in Molecular Biology 1289; Humana Press: New York, 2015.

2. Erlanson, D. A.; Jahnke, W.; Mannhold, R.; Kubinyi, H.; Folkers, G. Fragment-based Drug Discovery: Lessons and Outlook; Methods and Principles in Medicinal Chemistry; Wiley: Weinheim, Germany, 2016.

3. Bon, M.; Bilsland, A.; Bower, J.; & McAulay, K. Fragment-based drug discovery-the importance of high-quality molecule libraries. *Molecular Oncology* **2022**, *16(21)*, 3761–3777.

4. Schiebel, J.; Radeva, N.; Krimmer, S. G.; Wang, X.; Stieler, M.; Ehrmann, F. R.; Fu, K.; Metz, A.; Huschmann, F. U.; Weiss, M. S.; Mueller, U.; Heine, A.; Klebe, G. Six biophysical screening methods miss a large proportion of crystallographically discovered fragment hits: a case study. *ACS Chem. Biol.* **2016**, *11(6)*, 1693–1701.

5. Tan, Y. S.; Śledź, P.; Lang, S.; Stubbs, C. J.; Spring, D. R.; Abell, C.; Best, R. B. Using ligand-mapping simulations to design a ligand selectively targeting a cryptic surface pocket of polo-like kinase 1. *Angew. Chemie Int. Ed. Engl.* **2012**, *51(40)*, 10078–10081.

6. Matricon, P.; Ranganathan, A.; Warnick, E.; Gao, Z. G.; Rudling, A.; Lambertucci, C.; Marucci, G.; Ezzati, A.; Jaiteh, M.; Dal Ben, D.; Jacobson, K. A.; Carlsson, J. Fragment optimization for GPCRs by molecular dynamics free energy calculations: Probing druggable subpockets of the A_{2A} adenosine receptor binding site. *Sci. Rep.* **2017**, *7(1)*, 6398.

7. Ahmad, M.; Gu, W.; Helms, V. Mechanism of fast peptide recognition by SH3 Domains. *Angew. Chemie Int. Ed.* **2008**, *47(40)*, 7626–7630.

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516467; this version posted February 22, 2023. 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-NC 4.0 International license.

8. Shan, Y.; Kim, E. T.; Eastwood, M. P.; Dror, R. O.; Seeliger, M. A.; Shaw, D. E. How does a drug molecule find its target binding site? *J. Am. Chem. Soc.* **2011**, *133(24)*, 9181–9183.

9. Dror, R. O.; Pan, A. C.; Arlow, D. H.; Borhani, D. W.; Maragakis, P.; Shan, Y.; Xu, H.; Shaw, D. E. Pathway and mechanism of drug binding to G-protein-coupled receptors. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108(32)*, 13118–13123.

10. Bisignano; P.; Doerr, S.; Harvey, M. J.; Favia, A. D.; Cavalli, A.; De Fabritiis, G. Kinetic characterization of fragment binding in AmpC β-lactamase by high-throughput molecular simulations. *J. Chem. Inf. Model.* **2014**, *54(2)*, 362–366.

11. Pan, A. C.; Xu, H.; Palpant, T.; Shaw, D. E. Quantitative characterization of the binding and unbinding of millimolar drug fragments with molecular dynamics simulations. *J. Chem. Theory Comput.* **2017**, *13(7)*, 3372–3377.

12. Linker, S. M.; Magarkar, A.; Köfinger, J.; Hummer, G.; Seeliger, D. Fragment binding pose predictions using unbiased simulations and Markov-state models. *J. Chem. Theory Comput*. **2019**, *15(9)*, 4974–4981.

13. Ferruz, N.; Harvey, M. J.; Mestres, J.; De Fabritiis, G. Insights from fragment hit binding assays by molecular simulations. *J. Chem. Inf. Model.* **2015**, *55*, 2200−2205.

14. Martinez-Rosell, G.; Harvey, M. J.; De Fabritiis, G. Molecular-simulation-driven fragment screening for the discovery of new CXCL12 inhibitors. *J. Chem. Inf. Model*. **2018**, 58(3), 683–691.

15. Bernetti, M.; Masetti, M.; Recanatini, M.; Amaro, R. E.; Cavalli, A. Integrated Markov-state model and path metadynamics approach to characterize drug binding processes. *J. Chem. Theory Comput.* **2019**, *15*, 5689−5702.

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516467; this version posted February 22, 2023. 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-NC 4.0 International license.

16. Hollingsworth; S. A.; Dror, R. O. Molecular dynamics simulation for all. *Neuron* **2018**, *99(6)*, 1129–1143.

17. Zabolotny; J. M.; Bence-Hanulec, K. K.; Stricker-Krongrad, A.; Haj, F.; Wang, Y.; Minokoshi, Y.; Kim, Y. B.; Elmquist, J. K.; Tartaglia, L. A.; Kahn, B. B.; Neel, B. G. PTP1B regulates leptin signal transduction in vivo. *Dev. Cell* **2002**, *2(4)*, 489–495.

18. Cheng, A.; Uetani, N.; Simoncic, P. D.; Chaubey, V. P.; Lee-Loy, A.; McGlade, C. J.; Kennedy, B. P.; Tremblay, M. L. Attenuation of leptin action and regulation of obesity by protein tyrosine phosphatase 1B. *Dev. Cell* **2002**, *2(4)*, 497–503.

19. Feldhammer, M.; Uetani, N.; Miranda-Saavedra, D.; Tremblay, M. L. PTP1B: A simple enzyme for a complex world. *Crit. Rev. Biochem. Mol. Biol.* **2013**, *48(5)*, 430–445.

20. Krishnan, N.; Koveal, D.; Miller, D. H.; Xue, B.; Akshinthala, S. D.; Kragelj, J.; Jensen, M. R.; Gauss, C. M.; Page, R.; Blackledge, M.; Muthuswamy, S. K.; Peti, W.; Tonks, N. K. Targeting the disordered C terminus of PTP1B with an allosteric inhibitor. *Nat. Chem. Biol.* **2014**, *10(7)*, 558–566.

21. Szczepankiewicz, B. G.; Liu, G, Hajduk, P. J.; Abad-Zapatero, C.; Pei, Z.; Xin, Z.; Lubben, T. H.; Trevillyan, J. M.; Stashko, M. A.; Ballaron, S. J. Discovery of a potent, selective protein tyrosine phosphatase 1B inhibitor using a linked-fragment strategy. *J. Am. Chem. Soc.* **2003,** *125(14)*, 4087–4096.

22. Wiesmann, C.; Barr, K. J.; Kung, J.; Zhu, J.; Erlanson, D. A.; Shen, W.; Fahr, B. J,; Zhong, M.; Taylor, L.; Randal, M.; McDowell, R. S.; Hansen, S. K. Allosteric inhibition of protein tyrosine phosphatase 1B. *Nat. Struct. Mol. Biol.* **2004,** *11(8)*, 730–737.

23. Keedy, D. A.; Hill, Z. B.; Biel, J. T.; Kang, E.; Rettenmaier, T. J.; Brandão-Neto, J.; Pearce, N. M.; von Delft, F.; Wells, J. A.; Fraser, J. S. An expanded allosteric network in PTPB1 by multitemperature crystallography, fragment screening, and covalent tethering. *eLife* **2018,** *7*, e36307.

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516467; this version posted February 22, 2023. 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-NC 4.0 International license.

24. Asante-Appiah, E.; Patel, S.; Dufresne, C.; Roy, P.; Wang, Q.; Patel, V.; Friesen, R. W.; Ramachandran, C.; Becker, J. W.; Leblanc, Y.; Kennedy, B. P.; Scapin, G. The structure of PTP-1B in complex with a peptide inhibitor reveals an alternative binding mode for bisphosphonates. *Biochemistry* **2002**, *41(29)*, 9043–9051.

25. Choy, M. S.; Li Y; Machado, L. E. S. F.; Kunze, M. B. A.; Connors, C. R.; Wei, X.; Lindorff-Larsen, K.; Page, R.; Peti, W. Conformational rigidity and protein dynamics at distinct timescales regulate PTP1B activity and allostery. *Mol. Cell* **2017,** *65(4)*, 644–658.

26. Hjortness, M. K.; Riccardi. L.; Hongdusit, A.; Zwart, P. H.; Sankaran, B; De Vivo, M.; Fox, J. M. Evolutionarily conserved allosteric communication in protein tyrosine phosphatases, *Biochemistry* **2018**, *57(45)*, 6443–6451.

27. Cui, D. S., Beaumont, V., Ginther, P. S., Lipchock, J. M., Loria, J. P. Leveraging reciprocity to identify and characterize unknown allosteric sites in protein tyrosine phosphatases. *J. Mol. Biol.* **2017,** *429(15)*, 2360–2372.

28. Torgeson, K. R., Clarkson, M. W., Granata, D., Lindorff-Larsen, K., Page, R., & Peti, W. Conserved conformational dynamics determine enzyme activity. *Science Advances* **2022,** *8(31)*, eabo5546.

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516467; this version posted February 22, 2023. 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-NC 4.0 International license.

Figures

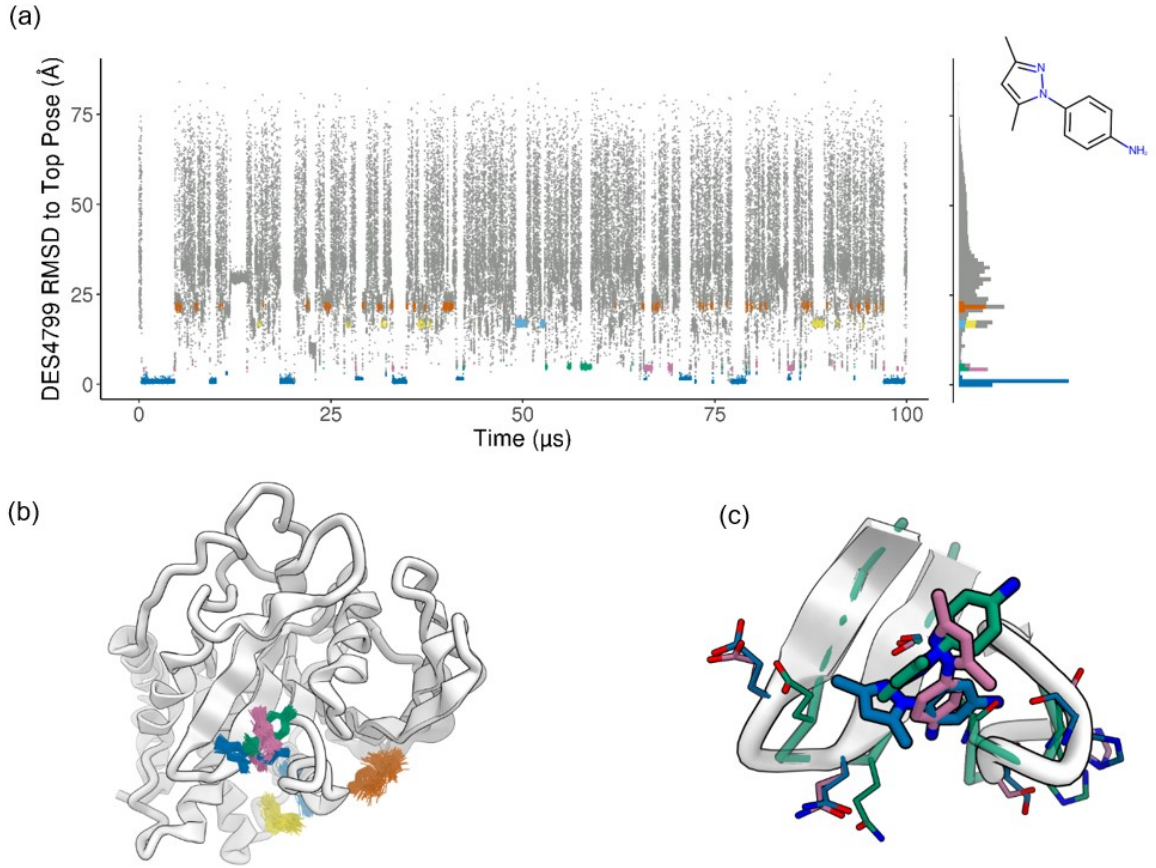

Figure 1. Summary of a representative 100-µs MD simulation of DES-4799. (a) The time series and probability densities of the RMSD of the small molecule to the dominant binding pose with the protein backbone aligned to the initial structure of the simulation. The data from the top six binding clusters of the fragment density are shown in color. (b) The locations of the top binding clusters on the surface of the protein, represented by 100 snapshots closest to their respective centroids. (c) The orientation of the fragment in the three clusters that bind in the main pocket.

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516467; this version posted February 22, 2023. 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-NC 4.0 International license.

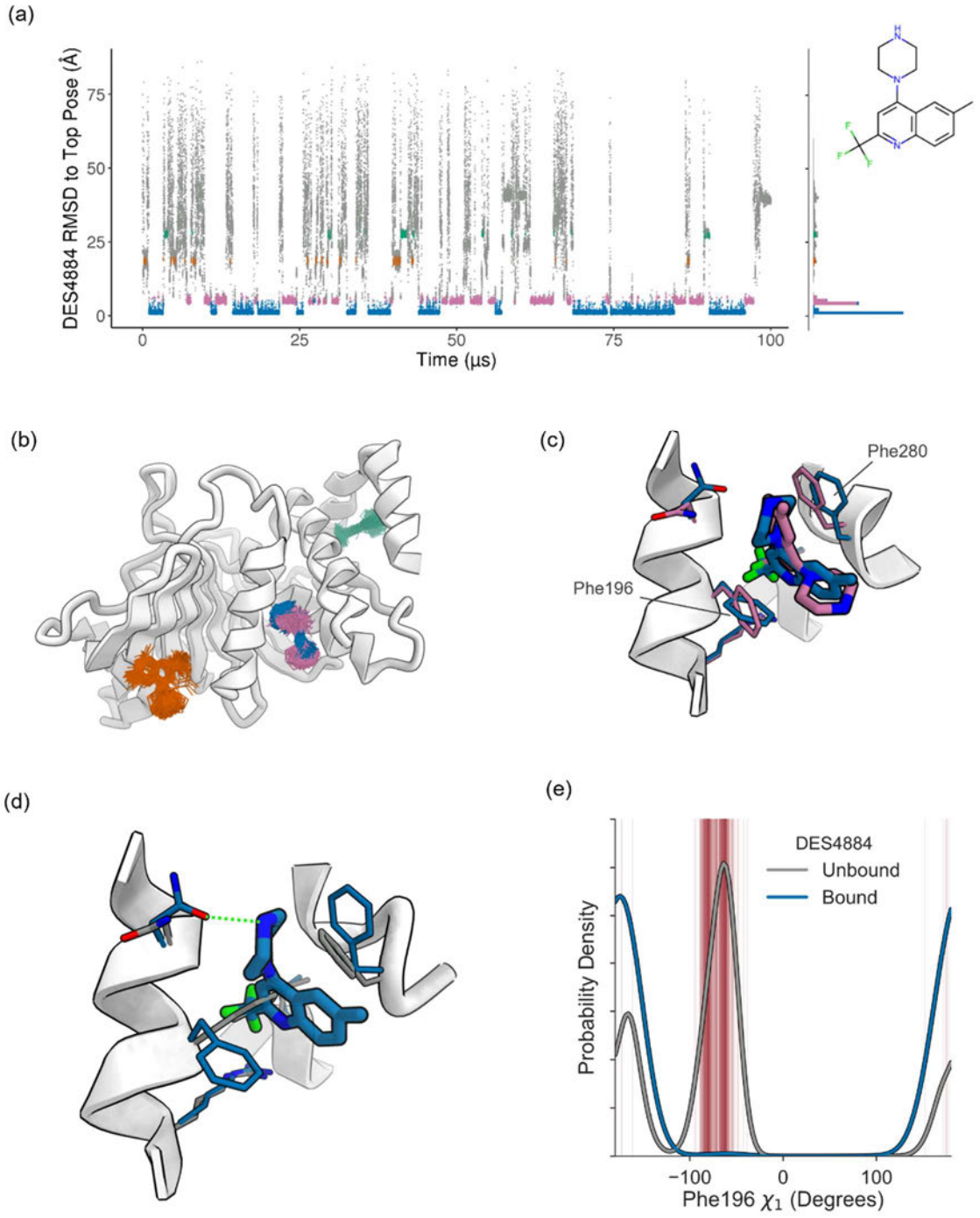

Figure 2. Summary of a representative 100-μs MD simulation of DES-4884. (a) The time

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516467; this version posted February 22, 2023. 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-NC 4.0 International license.

series and probability densities of the RMSD of the small molecule to the dominant binding pose with the protein backbone aligned to the initial structure of the simulation. The data from the top four binding clusters of the fragment density are shown in color. (b) The locations of the top four binding clusters, represented by 100 snapshots closest to their respective centroids. (c) The orientation of the fragment in the two clusters that bind in the main pocket. (d) The main binding pose of DES-4884; the fragment forms a hydrogen bond with Asn193 and the side chains of Phe280 and Phe196 swing out (shown in blue) compared to their apo orientations (shown in gray). (e) The density of the χ_1 angle of Phe196 during the snapshots in which the fragment is bound to the main pocket (blue) and the snapshots in which the fragment is not bound (gray). The red lines highlight the χ_1 angle of Phe196 in all crystal structures of PTP1b published in the PDB, including a line for each independent chain or alternate conformation.

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516467; this version posted February 22, 2023. 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-NC 4.0 International license.

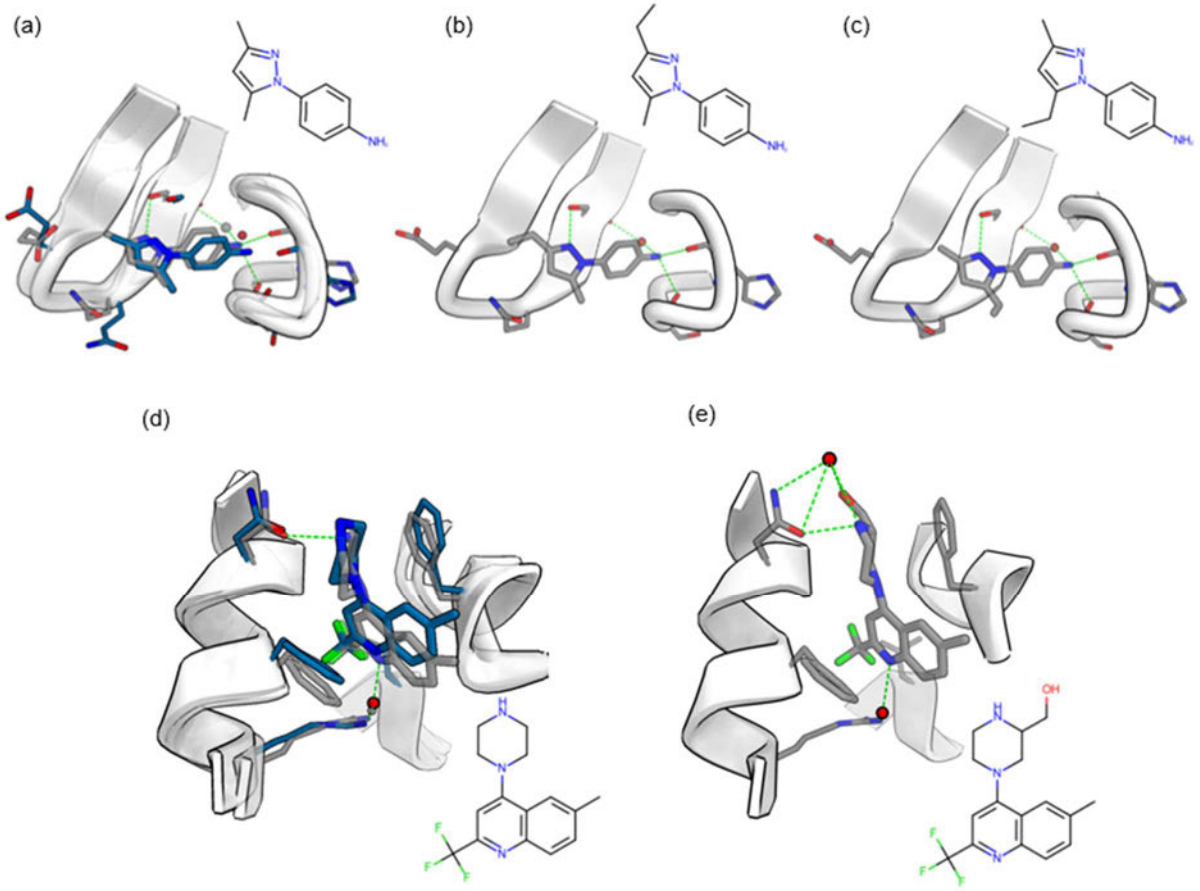

Figure 3. X-ray structures of DES-4799, DES-4884, and respective variants bound to PTP1b. (a) The top binding pose of DES-4799 predicted by MD simulation (blue) overlaid on its high-resolution X-ray crystal structure (gray). The oxygen atom of water is represented by a red sphere (simulation) or by a grey sphere (crystal structure). (b) The binding mode of the crystal structure of DES-5742, a single-methyl variant of DES-4799. (c) The binding mode of the crystal structure of DES-5743, a single-methyl variant of DES-4799. (d) The top binding pose of DES-4884 and pocket conformational changes predicted by MD simulation (blue) overlaid on the high-resolution X-ray crystal structure of DES-4884 (gray). The oxygen atom of water is represented by a red sphere (simulation) or by a grey sphere (crystal structure; partially

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516467; this version posted February 22, 2023. 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-NC 4.0 International license.

obscured by the red sphere). (e) The binding mode of DES-6016, a variant of DES-4884, resolved by high-resolution X-ray crystallography. In all structures, green dashed lines are used to highlight hydrogen bonds.