# Are protein-ligand complexes robust structures?

*Maciej Majewski,*[†] *Sergio Ruiz-Carmona,*[†] *and Xavier Barril*.*[†‡]

[†]Institut de Biomedicina de la Universitat de Barcelona (IBUB) and Facultat de Farmacia, Universitat de Barcelona, Av. Joan XXIII 27-31, 08028 Barcelona, Spain

[‡]Catalan Institution for Research and Advanced Studies (ICREA), Passeig Lluis Companys 23, 08010 Barcelona, Spain

Drug Design, Dynamic Undocking, Hydrogen Bond, Molecular Dynamics, Structural Stability.

The predominant view in structure-based drug design is that small-molecule ligands, once bound to their target structures, display a well-defined binding mode. While this is convenient from a design perspective, it ignores the fact that structural stability (robustness) is not necessary for thermodynamic stability (binding affinity). In fact, any potential benefit of a rigid binding mode will have to be balanced against the entropic penalty that it entails. Surprisingly, little is known about the causes, consequences and real degree of robustness of protein-ligand complexes. Here we investigate a diverse set of 77 drug-like structures, focusing on hydrogen bonds as they have been described as essential for structural stability. We find that most ligands combine a single anchoring point with looser regions, thus balancing order and disorder. But some ligands appear untethered and may form fuzzy complexes, while others are completely constricted. Structural stability analysis reveals a hidden layer of complexity in protein-ligand complexes that should be considered in ligand design.

## INTRODUCTION

Biomolecular systems present a large number of degrees of freedom and must find a suitable balance between order and disorder. In the particular case of non-covalent complexes, they can exist in a continuum spectrum of possibilities, ranging from the lock-and-key model to extreme disorder.[1,2] While the importance of target flexibility is well-appreciated in drug discovery,[3] the flexibility of small-molecule ligands in their bound state has attracted much less attention. Detailed analyses reveal that ligands often retain residual mobility.[4–6] However, changes in binding mode are more the exception than the norm[7,8] and ligand design based on rigid crystallographic geometries has been remarkably successful.[9] Perhaps for this reason, little is known about the molecular mechanisms that control structural stability, to what extend do ligands preserve flexibility or what are the energetic and functional consequences of rigidity.

It is important to note that structural stability (robustness) is fundamentally different from thermodynamic stability (i.e. binding free energy; $\Delta G_{bind}$). This is eloquently exemplified in the recent work by Borgia et al., where a protein-protein complex with picomolar affinity is shown to lack structure.[2] While $\Delta G_{bind}$ has been the center of attention of scientific research for decades, little attention has been paid to the factors that determine if a complex will be tight or loose. The source of structural robustness must be sought on sharp (and possibly transitory) energetic barriers that keep the atoms in their positions of equilibrium. Such hypothetical barriers, like the ones that determine binding kinetics, could have their origin in intramolecular (i.e. conformational rearrangement), bimolecular (e.g. repulsive transitional configurations) or many-body effects (e.g. desolvation).[10] But they will only provide structural stability if the barriers are steep and located very close to the position of minimum energy. In that respect, HBs are ideal candidates because they have strict distance and angular dependencies[11] and are one of the most

frequent interaction types in protein-ligand complexes.[12] The contribution of HBs to $\Delta G_{bind}$ has been largely debated in the literature.[13–17] The current consensus is that it is highly variable and context dependent, but their contribution to thermodynamic stability is 1.8 kcal mol-1 at the most.[14] However, due to desolvation, the transitional penalty of breaking a HB can be much larger.[18] Indeed, we have shown that this is the case for water-shielded HBs, which can even act as kinetic traps.[19] More recently, we have also shown that formation of structurally robust intermolecular HBs at specific positions is a necessary condition for binding, and have developed a method to assess the robustness of individual HBs that is very effective in virtual screening applications.[20]

With this background, we decided to perform a systematic investigation of the possible role of HBs as structural anchors of protein-ligand complexes. Our findings not only confirm a general role of HBs as source of structural stability, but also offer a new perspective to understand and design ligand-receptor complexes.

## RESULTS AND DISCUSSION

Using Dynamic Undocking (DUck), an MD-based computational procedure,[20] we have assessed the robustness of every HB in a set of 77 drug-like protein-ligand complexes from the Iridium Data Set.[21] Detailed information about the data set and the selection criteria is presented in Materials and Methods and Table S1. Each HB was pulled to a distance of 5 Å, according to the DUck protocol reported previously.[20,22] In this way, we obtain a work value ($W_{QB}$) that reflects the cost of breaking each HB. Based on our previous research, we define HBs as robust (i.e. capable of providing structural stability) if $W_{QB} > 6$ kcal mol$^{-1}$, labile if $W_{QB} < 4$ kcal mol$^{-1}$ and medium otherwise.

The distribution of work values for the entire set of 342 HBs ranges from 0 to 26 kcal $mol^{-1}$, with a region of maximum probability in the 0-6 kcal $mol^{-1}$ region and a gradual decrease thereafter (Fig.1a). Noteworthy, more than half HBs (58%) are robust, and three quarters of all complexes (75%) contain at least one such HB. Considering that structural stability is not a requisite for tight binding and that HBs may not the only mechanism capable of providing structural stability, it is striking that such a large proportion of the complexes in this set are anchored through HBs. A further 13% of complexes present medium values and only in 9 cases (12%) all their HBs are labile (Fig.S1). Two of those cases are very low affinity complexes. In the remaining cases, structural stability might be provided by other mechanisms, such as water-mediated hydrogen bonds or cation-pi interactions, or may be lacking (see examples in Fig. 2).

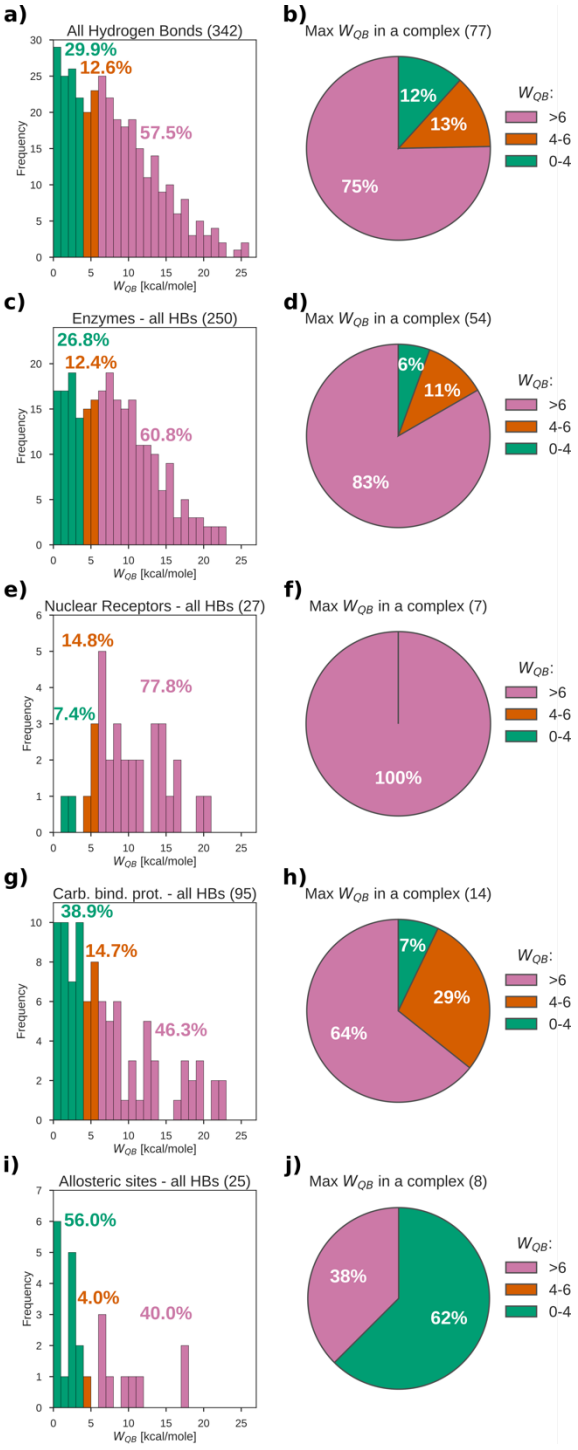


Figure 1 Histograms of frequency of HBs by $W_{QB}$ value for: a) all simulated HBs (342), c) HBs in enzymes (250), e) HBs in nuclear receptors (27), g) HBs in carbohydrate binding proteins (95), i) HBs in allosteric sites (25). Pie charts showing share of complexes with at least one robust HB

($W_{QB}$ > 6 kcal mol$^{-1}$, pink), all labile HBs ($W_{QB}$ < 4 kcal mol$^{-1}$, green) or intermediate situations (red) for: b) all simulated complexes (77), d) enzymes (54), f) nuclear receptors (7), h) carbohydrate binding proteins (14), j) allosteric sites (8).

Splitting this analysis by protein class (Fig.1c-j) provides strong indication that the behavior is dictated by the nature of the target. The proportion of robust complexes increases to 83% in the case of enzymes, which speaks about the need of keeping the substrate in place for efficient catalysis. Nuclear receptors form fewer HBs with their ligands, but most of them (78%) are robust and all ligands (100%) are well anchored. In this case, forming a rigid structure may be necessary to stabilize the AF2 co-regulatory protein binding surface in an optimal conformation for co-activator binding.[23] Carbohydrate binding proteins, on the other hand, form many more HBs, but a lower proportion of robust ones (46%). Finally, in the case of allosteric ligands, only 40% of complexes are robust, suggesting that non-functional sites tend to yield looser complexes. As demonstrated in the case of HIV reverse transcriptase inhibitors (Fig.2C), lack of robust HBs does not preclude tight binding. In fact, a multiplicity of binding modes might be beneficial to preserve binding affinity when the target is mutated, thus averting resistance.[24,25]

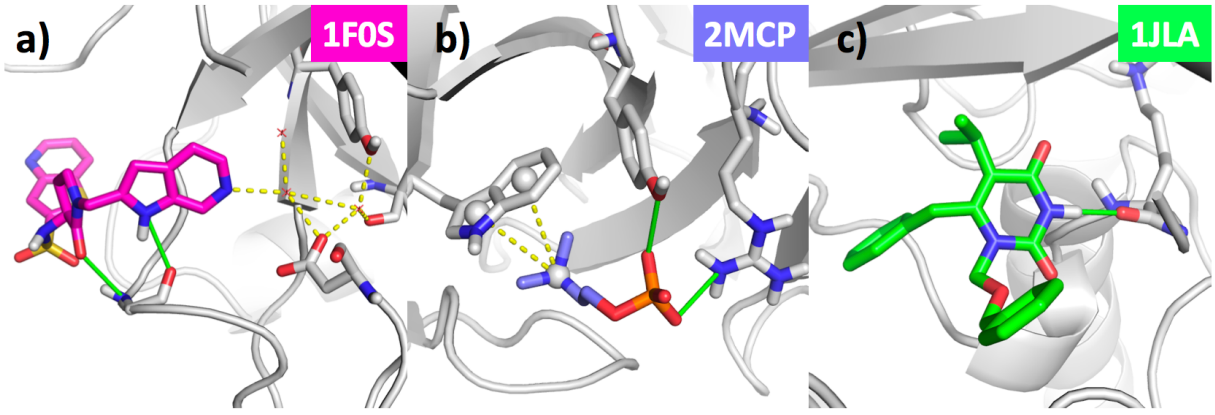

Figure 2 Structures of protein-ligand complexes that form potentially labile structures (all HBs weaker than 4 kcal/mole). a) Complex of FXa with inhibitor RPR208707 (PDB id 1F0S; $K_i$ = 18 nM) forms two direct, but labile, HBs with the protein. An additional water-mediated HB with the catalytic residues (yellow dotted lines) might provide structural stability. b) An antibody that recognises phosphocholine (PDB id 2MCP) forms two charge-reinforced but labile HBs. A cation-pi interaction (yellow dotted lines) might provide structural stability. c) Reverse transcriptase inhibitor (PDB id 1JLA; $IC_{50}$ = 6 nM) forms a single but labile HB with the protein. No other source of structural stability is apparent.

We analyzed the distribution of robust HBs and found that they tend to concentrate on one part of the ligand (Fig.S2). To better understand this observation, all HBs in each complex were clustered, based on their distance in space, into fragment-sized group of atoms (Fig.S3). In the majority of complexes (61%) robust HBs were located in a single group, forming a strong structural anchor (Fig.3, Tab.S3). The concentration of robust interactions on a single site, allowing a some degree of movement to the other parts, minimises the entropic costs and can be desirable from a binding affinity perspective.[6] It also ties in with the observation that fragment screening hits - in spite of their low binding affinity - already form the key interactions and have a well-defined binding mode that serves as a foundation from which to spread and catch additional interactions.[26] Only 23% of ligands form two structural anchors on separate regions, though this is more common in the case of carbohydrate-binding proteins (Tab.S4). Three exceptional ligands manage to form 3 distinct stable anchors. Interestingly, they have completely unrelated functions, chemical structures and physical properties. At least in two of those cases there is a possible functional explanation for the extreme robustness (Fig.4).

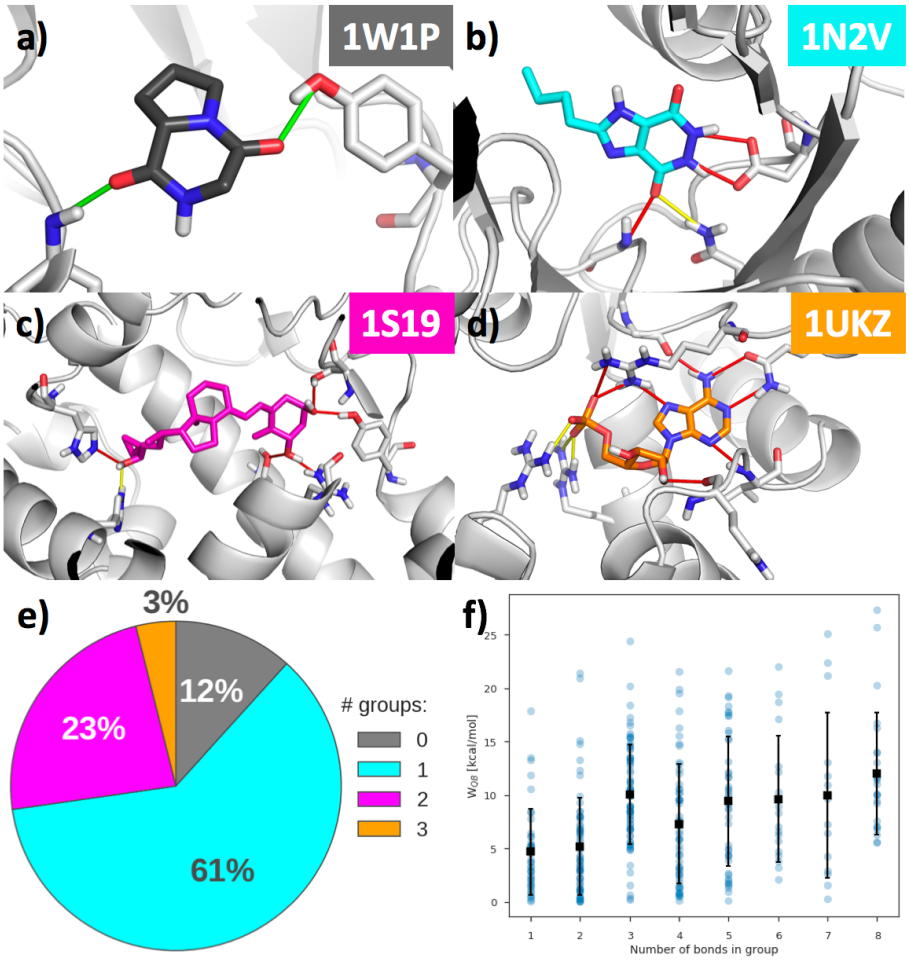


Figure 3 Division of complexes based on the number of structural anchors. Representative of each group is presented in the following image: a) 0 anchors: Chitinase B with inhibitor (PDB id 1W1P; $IC_{50}$ = 5 mM); b) 1 anchor: Queuine tRNA-ribosyltransferase with inhibitor (PDB id 1N2V; $K_i$ = 83 $\mu$M), c) 2 anchors: Vitamin D3 receptor with calcipotriol (PDB id 1S19; $K_d$ = 0.31 nM) and d) 3 anchors: Uridylate kinase - AMP (PDB id 1UKZ). e) Pie chart presenting distribution of number of anchors across the data set. f) Distribution of strength of HBs ($W_{QB}$) versus the number of HBs per group of atoms.

The distribution of $W_{QB}$ per number of HBs in a local group (Fig.3f) is suggestive of cooperative behavior. HBs in isolation usually do not form robust interactions (mean and median values: 4.7 and 3.5 kcal mol$^{-1}$, respectively), although in exceptional cases they can reach values above 10 kcal mol$^{-1}$. By contrast, when three or more HBs cluster together, formation of robust complexes is the most common outcome (mean and median values: 9.4 and 9.0 kcal mol$^{-1}$, respectively). The HBs within these clusters tend to present similar $W_{QB}$ values, suggesting that breaking the weakest bond leads to rupture of the whole network. This all-or-nothing behavior not only ensures higher barriers to dissociation, but is also well-suited to provide selectivity, as small changes in the composition or geometry of the ligand may result in large changes in magnitude of $W_{QB}$ (see example in Fig.S4).

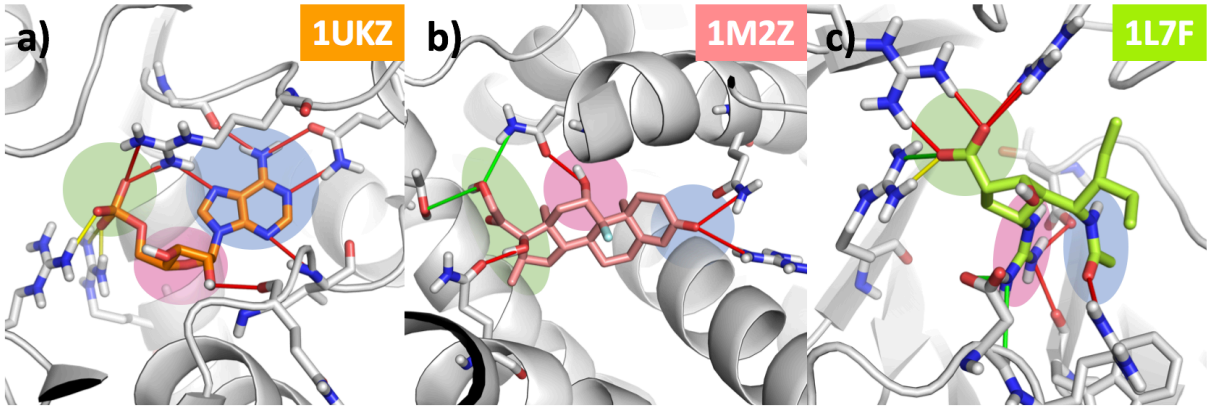


Figure 4 Structures of complexes with three binding anchors (shaded areas). a) Uridilate kinase with AMP (PDB id 1UKZ) where the base, ribose and phosphate of the nucleotide are forming three distinctive centres of interactions. b) Glucocorticoid receptor ligand-binding domain bound to dexamethasone (PDB id 1M2Z; $K_d$ = 19 nM). The ligand has three regions that form robust interaction, well separated in space but located on the steroid core, thus behaving as a single rigid block. c) Influenza virus neuraminidase with inhibitor BCX-1812 (PDB id 1L7F; $K_i$ single digit

nM for various virus strains). Three different functional groups branching out of the pentane scaffold form robust interactions in this extremely polar and solvent exposed binding site.

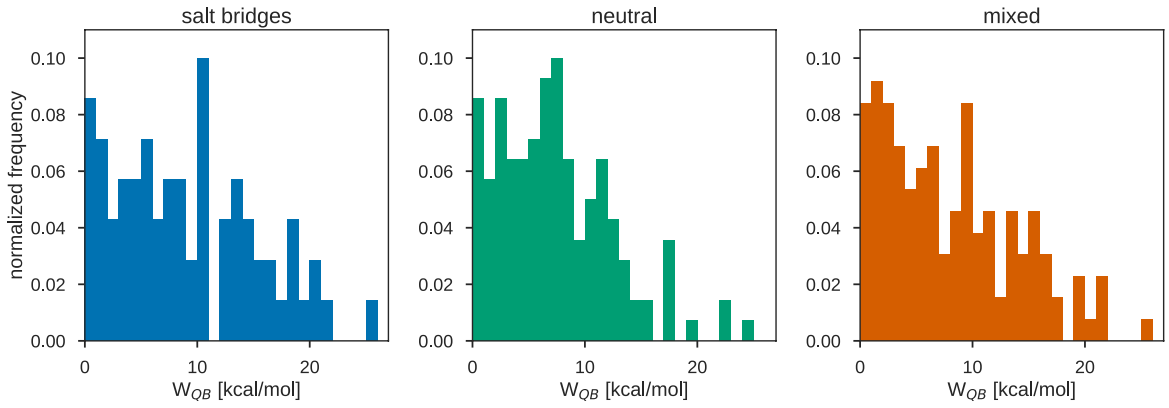


Figure 5 Histograms presenting distribution of $W_{QB}$ values in the set of HBs from Iridium DS, divided into salt bridges, neutral and mixed (ion-neutral) interactions

In order to assess the effect of charge reinforcement on HBs, we have classified them into neutral, mixed (ionic-neutral) and salt bridges (Fig.5). We find that salt bridges are slightly skewed towards more robust interactions than neutral HBs (p-value = 0.08). However, mixed types are indistinguishable from neutral ones (p-value = 0.39) and, unexpectedly, the maximal values are equal across all three categories. Theoretically, ionic species could provide even larger energetic barriers, but there may be no biological use for them, as the maximal $W_{QB}$ values observed here already ensure very robust and long-lived structures.

Knowing that a HB has a large $W_{QB}$ value can be likened to knowing the $k_{off}$ of a compound without knowing the $k_{on}$ nor $\Delta G_{bind}$: larger values may indicate that it has a higher transition state (if $\Delta G_{bind}$ remains the same; Fig.6a), that the complex is thermodynamically more stable (if $k_{on}$ remains the same; Fig.6b), or a combination thereof. In this data set, we find that anchoring sites

often correspond to binding hot spots. This is indeed the case for all kinases and proteases, which have a well-known binding hot spot (Tab.S5, Fig.S3). In such cases, $\Delta G_{bind}$ must be a component of $W_{QB}$, but there is no correlation between both magnitudes (Fig.S5), as already noted.[20] Thus, we conclude that $W_{QB}$ must be largely dominated by a transitory dissociation penalty. The origin of this penalty can be explained by a physical decoupling between HB rupture and resolvation.[19] Likewise, several studies of the reverse event have identified desolvation of the binding pocket as the rate-limiting step in ligand association.[18,27,28] Indeed, solvent exposed HBs invariably lead to low $W_{QB}$ values (but note that they can be thermodynamically stable),[29] whereas water-shielding is a necessary but not sufficient condition of robust HBs (Fig.S6).

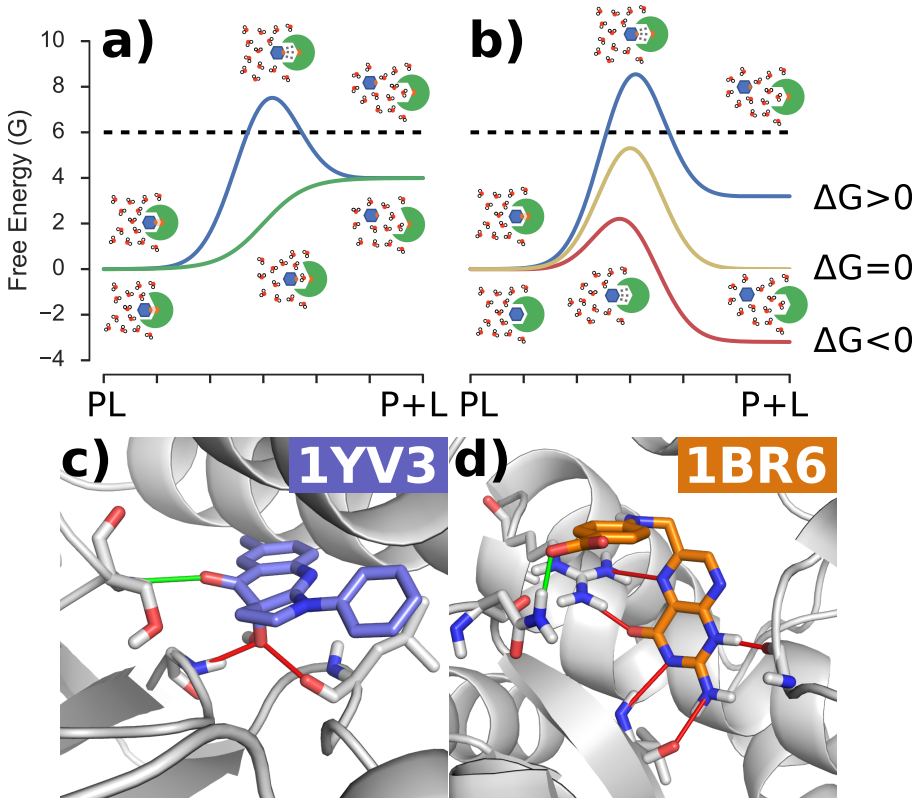

Figure 6 Ways of achieving structural robustness. a) Idealized representation of two dissociation pathways for complexes with the same $\Delta G_{bind}$ and different desolvation costs. b) Likewise for two complexes with the same desolvation cost but different $\Delta G_{bind}$ c) Example of a complex with high dissociation cost due to extreme water-shielding. d) Example of a complex with high dissociation cost due to a tight network of multiple HBs.

## CONCLUSION

Taken together, our results show that structural stability cannot be taken for granted. Examples in our data set range from the very tight (e.g. nuclear receptor ligands) to the very loose (e.g. HIV-RT allosteric inhibitors), while most complexes combine a firm anchor site with more relaxed side interactions. Each one of these solutions entails important consequences that have, so far, been neglected in drug design. First of all, a firm anchor provides a framework from which to grow and capture additional interactions, and the preservation of a common binding mode helps interpreting structure-activity relationships. Secondly, structural robustness can have functional implications, particularly in the case of receptors, where flexibility has been linked to the agonist/antagonist response.[23,30] Thirdly, structural stability implies an entropic penalty and must be balanced to avoid loss of potency.[6,31] Finally, the deep and narrow energetic minima that cause rigidity also imply large penalties for small recognition defects, thus increasing the fidelity of the recognition event. This has been shown for protease-substrate pairs[32] and HIV-protease inhibitors.[33] In conclusion, this work opens up the possibility of understanding and designing structural robustness in ligand-receptor complexes. Qualitatively, water-shielded HBs (Fig.6c) and HB clusters (Fig.6d) are tell-tale signs of robustness. Quantitatively, DUck simulations offer an inexpensive and automated protocol to calculate $W_{QB}$. While HBs appear to be the most

common means of achieving structural robustness, other interaction types (e.g. cation-pi, water-mediated HBs, halogen bonds) will be investigated in the future.

## AUTHOR INFORMATION

**Corresponding Author**

*E-mail: xbarril@ub.edu

## ACKNOWLEDGMENT

We thank C. Galdeano and F. J. Luque for helpful discussions and manuscript revision. The research was funded under the EU Horizon 2020 programme, Marie Skłodowska-Curie grant agreement No. 675899 (FRAGNET). Spanish Ministerio de Economia (SAF2015-68749-R). Catalan government (2014 SGR 1189). Barcelona Supercomputing Center (computational resources).

## AUTHOR CONTRIBUTIONS

X.B. designed the project; M.M. performed calculations; S.R.C. contributed new analytic tools; M.M. and X.B. analysed data and wrote the manuscript.

## REFERENCES

1. Tompa, P. & Fuxreiter, M. Fuzzy complexes: polymorphism and structural disorder in protein-protein interactions. *Trends Biochem. Sci.* **33,** 2–8 (2008).

2. Borgia, A. *et al.* Extreme disorder in an ultrahigh-affinity protein complex. *Nature* **555,** 61–66 (2018).

3. Cozzini, P. *et al*. Target Flexibility: An Emerging Considertaion in Drug Discovery. *J. Med. Chem*. **51,** 6237–6255 (2008).

4. van Zundert, G. C. P. *et al*. qFit-ligand reveals widespread conformational heterogeneity of drug-like molecules in X-ray electron density maps. *bioRxiv* 253419 (2018).

5. Klebe, G. Applying thermodynamic profiling in lead finding and optimization. *Nature Reviews Drug Discovery* **14,** 95–110 (2015).

6. Glas, A., Wamhoff, E.-C., Kruger, D. M., Rademacher, C. & Grossmann, T. N. Increased Conformational Flexibility of a Macrocycle – Receptor Complex Contributes to Reduced Dissociation Rates. *Chem. Eur. J*. **23,** 16157–16161 (2017).

7. Malhotra, S. & Karanicolas, J. When does chemical elaboration induce a ligand to change its binding mode? *J. Med. Chem*. **60,** 128–145 (2017).

8. Kuhnert, M. *et al*. Tracing binding modes in hit-to-lead optimization: Chameleon-like poses of aspartic protease inhibitors. *Angew. Chemie - Int. Ed*. **54,** 2849–2853 (2015).

9. Sliwoski, G., Kothiwale, S., Meiler, J. & Lowe, E. W. Computational methods in drug discovery. *Pharmacol. Rev*. **66,** 334–95 (2014).

10. Pan, A. C., Borhani, D. W., Dror, R. O. & Shaw, D. E. Molecular determinants of drug – receptor binding kinetics. *Drug Discov. Today* **18,** 667 (2013).

11. Bissantz, C., Kuhn, B. & Stahl, M. A Medicinal Chemist's Guide to Molecular Interactions. *J. Med. Chem*. **53,** 5061–5084 (2010).

12. Ferreira de Freitas, R. & Schapira, M. A systematic analysis of atomic protein–ligand interactions in the PDB. *Med. Chem. Commun.* **8,** 1970–1981 (2017).

13. Fersht, A. R. The hydrogen bond in molecular recognition. *Trends Biochem. Sci.* **12,** 301–304 (1987).

14. Pace, C. N. Energetics of protein hydrogen bonds. *Nat. Struct. Mol. Biol.* **16,** 681–682 (2009).

15. Pace, C. N. *et al*. Contribution of hydrogen bonds to protein stability. *Protein Sci.* **23,** 652–661 (2014).

16. Gao, J., Bosco, D. A., Powers, E. T. & Kelly, J. W. Localized thermodynamic coupling between hydrogen bonding and microenvironment polarity substantially stabilizes proteins. *Nat. Struct. Mol. Biol.* **16,** 684–690 (2009).

17. Nick Pace, C., Martin Scholtz, J. & Grimsley, G. R. Forces stabilizing proteins. *FEBS Lett.* **588,** 2177–2184 (2014).

18. Mondal, J., Friesner, R. A. & Berne, B. J. Role of Desolvation in Thermodynamics and Kinetics of Ligand Binding to a Kinase. *J. Chem. Theory Comput.* **10,** 5696–5705 (2014).

19. Schmidtke, P., Javier Luque, F., Murray, J. B. & Barril, X. Shielded hydrogen bonds as structural determinants of binding kinetics: Application in drug design. *J. Am. Chem. Soc.* **133,** 18903–18910 (2011).

20. Ruiz-carmona, S. *et al*. Dynamic undocking and the quasi-bound state as tools for drug discovery. *Nat. Chem.* **9,** 201 (2017).

21. Warren, G. L., Do, T. D., Kelley, B. P., Nicholls, A. & Warren, S. D. Essential considerations for using protein-ligand structures in drug discovery. *Drug Discov. Today* **17,** 1270–1281 (2012).

22. Majewski, M., Ruiz-Carmona, S. & Barril, X. in *Rational Drug Design: Methods and Protocols* (eds. Mavromoustakos, T. & Kellici, T. F.) 195–215 (Springer New York, 2018). doi:10.1007/978-1-4939-8630-9_11

23. Mayer-Wrangowski, S. C. & Rauh, D. Monitoring ligand-induced conformational changes for the identification of estrogen receptor agonists and antagonists. *Angew. Chemie - Int. Ed.* **54,** 4379–4382 (2015).

24. Das, K. *et al.* Roles of Conformational and Positional Adaptability in Structure-Based Design of TMC125-R165335 ( Etravirine ) and Related Non-nucleoside Reverse Transcriptase Inhibitors That Are Highly Potent and Effective against Wild-Type and Drug-Resistant HIV-1 Vari. *J. Med. Chem.* **47,** 2550–2560 (2004).

25. Lee, W. G., Chan, A. H., Spasov, K. A., Anderson, K. S. & Jorgensen, W. L. Design, Conformation, and Crystallography of 2-Naphthyl Phenyl Ethers as Potent Anti-HIV Agents. *ACS Med. Chem. Lett.* **7,** 1156–1160 (2016).

26. Ferenczy, G. G. & Keseru, G. M. Thermodynamics of fragment binding. *J. Chem. Inf. Model.* **52,** 1039–1045 (2012).

27. Dror, R. O. *et al.* Pathway and mechanism of drug binding to G-protein-coupled receptors. *Proc. Natl. Acad. Sci. U. S. A.* **108,** 13118–23 (2011).

28. Schuetz, D. A. *et al*. Ligand Desolvation steers on-rate and impacts Drug Residence Time of Heat shock protein 90 (Hsp90) Inhibitors. *J. Med. Chem*. **90,** (2018).

29. Ciulli, A., Williams, G., Smith, A. G., Blundell, T. L. & Abell, C. Probing hot spots at protein-ligand binding sites: A fragment-based approach using biophysical methods. *J. Med. Chem*. **49,** 4992–5000 (2006).

30. Ghanouni, P. *et al*. Functionally Different Agonists Induce Distinct Conformations in the G Protein Coupling Domain of the β2Adrenergic Receptor. *J. Biol. Chem*. **276,** 24433–24436 (2001).

31. Brandt, T. *et al*. Congeneric but still distinct: How closely related trypsin ligands exhibit different thermodynamic and structural properties. *J. Mol. Biol*. **405,** 1170–1187 (2011).

32. Fuchs, J. E. *et al*. Cleavage Entropy as Quantitative Measure of Protease Specificity. *PLoS Comput. Biol*. **9,** e1003007 (2013).

33. Shen, Y., Radhakrishnan, M. L. & Tidor, B. Molecular mechanisms and design principles for promiscuous inhibitors to avoid drug resistance: Lessons learned from HIV-1 protease inhibition. *Proteins Struct. Funct. Bioinforma*. **83,** 351–372 (2015).