

UCSF

UC San Francisco Electronic Theses and Dissertations

Title

Model systems for molecular docking: understanding molecular recognition in polar and charged binding sites

Permalink

<https://escholarship.org/uc/item/81j4m12t>

Author

Boyce, Sarah Emily

Publication Date

2009

Peer reviewed|Thesis/dissertation

Model systems for molecular docking:
Understanding molecular recognition in polar and charged binding sites

by

Sarah Emily Boyce

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Chemistry and Chemical Biology

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

Copyright 2009

by

Sarah E. Boyce

I would like to dedicate this to my friends and family, and especially to Patrick J. Roland. Without his support this work would not have been possible.

Acknowledgements

I'd like to thank my family— my husband Patrick Roland, my parents Stephen and JoAnn Boyce, and my siblings Chris and Cassie Boyce, and Kayla McMillan-Schmidt, for their unrelenting support and interest in my work. I'd also like to acknowledge my grandmother Maria A.G. Boyce, who inspired me with her natural curiosity and the continued desire to accumulate knowledge at every stage in her life.

Thanks especially to the following people: my thesis advisor Brian Shoichet for taking in a chemist with zero computational experience and giving me the opportunity to work on this project; my thesis committee members Matt Jacobson and Ken Dill; my many collaborators— David Goodin, Stephan Vetter, Rich Wilson, Dave Case, Devleena Shivakumar and David Mobley; Patsy Babbitt, Linda Brinen and Sue Miller for support and advice; CCB program administrator Christine Olson for her help with every aspect of graduate school; Ruth Brenk, my mentor on this project, and Gabe Rocklin, who inherits it; my friends (and moral support)— Holly Atkinson, Andrea McReynolds, Kerim Babaoglu, Alan Graves, Veena Thomas and Brian Feng. These people, and many other members of the Shoichet lab, helped guide my progress as a PhD student with good advice, technical expertise and friendship. I am extremely grateful to them all.

The text of Chapter 1 is a reprint of the material as it appears in:

Brenk R, Vetter SW, Boyce SE, Goodin DB, Shoichet BK. Probing molecular docking in a charged model binding site. *J Mol Biol* **357** (5), 1449-70 (2006).

It appears here with permission from the authors. The supplementary material from this paper has been included as Appendix A.

The text of Chapter 2 is a reprint of the material as it appears in:

Graves AP*, Shivakumar DM*, Boyce SE, Jacobson MP, Case DA, Shoichet BK.

Rescoring docking hit lists for model cavity sites: predictions and experimental testing. *J Mol Biol* **377** (3), 914-34 (2008).

It appears here with permission from the authors. The supplementary material from this paper has been included as Appendix B.

The text of Chapter 3 is a reprint of the material as it appears in:

Boyce SE*, Mobley DL*, Rocklin GJ, Graves AP, Dill KA, Shoichet BK. Predicting Ligand Binding Affinity with Alchemical Free Energy Methods in a Polar Model Binding Site. *J Mol Biol*. Available online Sept. (2009.)

It appears here with permission from the authors. The supplementary material from this paper has been included as Appendix C.

Abstract

New model systems for molecular docking: Understanding molecular recognition in polar and charged binding sites

Virtual screening is a powerful tool in drug discovery, with the potential to find novel ligands for therapeutically relevant target structures. However, the field is plagued by both false positive and false negative predictions. This is due to approximations within the scoring functions, leading to the failure to distinguish between true ligands and high-ranking nonbinders (decoys). To compound the problem, in a typical target the complexity of the ligand-receptor interactions prevents us from unraveling the many components of the binding energy that lead to the incorrect predictions. Model binding sites provide simpler systems in which individual terms can be isolated and studied.

In Chapter 1, cytochrome *c* peroxidase (CCP) W191G, an anionic, wet, and buried cavity is introduced. This cavity primarily binds aromatic monocations; dications and most neutral molecules do not bind detectably. In Chapter 2, CCP W191G is included in a series of model systems (the T4 lysozyme L99A hydrophobic and L99A/M102Q polar cavities) to evaluate MM-GBSA rescoring of docking hit lists; both chapters consider the case for CCP W191G where the scoring function must balance the cost of ligand desolvation with the favorable electrostatic interaction energy between ligand and protein.

Chapter 3 returns to the T4 lysozyme L99A/M102Q polar models system for absolute and relative binding free energy predictions. This system proved to be difficult for the free energy methods, but not due to the additional polarity, as we had initially

predicted. Instead, protein conformational change, sampling of reasonable ligand orientations and methodological failures proved a challenge to accurate predictions.

In the final chapter, a new open cavity in cytochrome *c* peroxidase, created by the W191G\ P190G Δ G192-A193 deletion mutant is introduced. The cavity contains multiple ordered waters and an interface to bulk solvent. This more complicated cavity presents an opportunity to investigate displacing individual ordered waters, and the potential for neutral ligands in a charged cavity. Implications for this new charged, open cavity and preliminary results are discussed in Chapter 4: Future Directions.

Table of Contents

ACKNOWLEDGEMENTS	IV
ABSTRACT.....	VI
TABLE OF CONTENTS	VIII
LIST OF TABLES	XI
LIST OF FIGURES.....	XIII
INTRODUCTION	1
I. Understanding molecular recognition	1
II. Computational methods for drug discovery	2
III. Polar and charged model systems for molecular docking	4
IV. Guide to the chapters	5
V. References.....	8
GLOSS TO CHAPTER 1.....	16
CHAPTER 1: Probing molecular docking in a charged model binding site	22
1.1 Abstract.....	23
1.2 Introduction.....	26
1.3 Results.....	30
1.4 Discussion.....	54
1.5 Conclusions.....	60
1.6 Methods.....	61
1.7 Acknowledgements.....	68
1.8 References.....	69
GLOSS TO CHAPTER 2.....	79
CHAPTER 2: Rescoring docking hit lists for model cavity sites: predictions and experimental testing	83

2.1 Abstract.....	84
2.2 Introduction.....	88
2.3 Results.....	93
2.4 Discussion.....	116
2.5 Materials and Methods.....	122
2.6 Acknowledgements.....	128
2.7 Supporting Information Available	128
2.8 References.....	129
GLOSS TO CHAPTER 3.....	138
CHAPTER 3: Predicting ligand binding affinity with alchemical free energy methods in a polar model binding site	142
3.1 Abstract.....	143
3.2 Introduction.....	145
3.3 Results.....	147
3.4 Discussion.....	164
3.5 Conclusion	173
3.6 Materials and Methods.....	175
3.7 Acknowledgments.....	183
3.8 References.....	184
GLOSS TO CHAPTER 4: FUTURE DIRECTIONS	190
CHAPTER 4: CCP W191G gateless model binding site: charged, polar, and open to bulk solvent.....	191
4.1 Abstract.....	192
4.2 Introduction.....	194
4.3 Preliminary Results and Discussion.....	195
4.4 Future Directions	203

4.5 Methods.....	206
4.6 Acknowledgments.....	207
4.7 References.....	208
APPENDIX A: SUPPLEMENTARY MATERIAL FOR CHAPTER 1	211
A.1 Supplementary figures for Chapter 1	211
A.2 References.....	214
APPENDIX B: SUPPLEMENTARY MATERIAL FOR CHAPTER 2	215
B.1 PLOP side chain rotamer search and minimization algorithm.....	216
B.2 AMBERDOCK parameters and optimization.....	216
B.3 Supplementary Tables for Chapter 2.....	220
B.4 Supplementary Figures for Chapter 2	221
APPENDIX C: SUPPLEMENTARY MATERIAL FOR CHAPTER 3	224
C.1 Supplementary Methods.....	224
C.2 Supplementary Figures for Chapter 3	232
C.3 Supplementary Tables for Chapter 3.....	234
C.4 References	240
Publishing Agreement.....	242

List of Tables

Chapter 1

TABLE 1. Rmsd for the top scoring docking pose compared to the previously determined structure	31
TABLE 2. New docking-derived hits, tested for binding to CCP W191G	35
TABLE 3. Crystallographic data	41

Chapter 2

TABLE 1. Compounds predicted by AMBERDOCK and PLOP to bind to T4 Lysozyme L99A	100
TABLE 2. Crystallographic measurement and the RMSD values for predicted and crystallographic ligand geometries in the L99A and L99A/M102Q sites	104
TABLE 3. Compounds predicted by AMBERDOCK and PLOP to bind to T4 Lysozyme L99A/M102Q	105
TABLE 4. Compounds predicted to bind by AMBERDOCK and PLOP to CCP W191G	111
TABLE 5. Crystallographic measurement and the RMSD values for predicted and crystallographic ligand geometries in the CCP site.....	114
TABLE 6. Likely ligands and decoys among the top 100 ranked ligands by docking and MM-GBSA	115

Chapter 3

TABLE 1. Prospective test: absolute binding free energy predictions.....	149
TABLE 2. Experimentally determined binding free energy for phenol derivatives.....	161

Chapter 4

TABLE 1. Neutral ligands preferentially bind to CCP-GA at pH 6.0.....	200
--	------------

Appendix B

TABLE S1. Top 100 hits predicted by DOCK3.5.54 for L99A	220
TABLE S2. Top 100 hits predicted by PLOP for L99A.....	220
TABLE S3. Top 100 hits predicted by AMBERDOCK for L99A.....	220
TABLE S4. Top 100 hits predicted by DOCK3.5.45 for L99A/M102Q	220
TABLE S5. Top 100 hits predicted by PLOP for L99A/M102Q.....	220
TABLE S6. Top 100 hits predicted by AMBERDOCK for L99A/M102Q.	220
TABLE S7. Top 100 hits predicted by DOCK3.5.54 for CCP.....	220
TABLE S8. Top 100 hits predicted by PLOP for CCP	220
TABLE S9. Top 100 hits predicted by AMBERDOCK for CCP	220
TABLE S6. Top 100 hits predicted by AMBERDOCK for L99A/M102Q.	220

Appendix C

TABLE S1. Retrospective Absolute Free Energy Calculations.....	234
TABLE S2. X-ray data collection and refinement.....	235
TABLE S3. Comparison of preferred rotameric state predicted by confine-and-release protocol, and the crystallographic results	236
TABLE S4. Comparison of absolute free energy predictions beginning from the Apo vs. Holo structure (with corrected long range dispersion term and corrected partial charges)	237
TABLE S5. X-ray data collection and refinement.....	238
TABLE S6. Relative binding free energy results recalculated with corrected long range dispersion term	239

List of Figures

Chapter 1

FIGURE 1. The cavity in CCP W191G	28
FIGURE 2. Retrospective enrichment of previously known, “test set” ligands for the W191G cavity in CCP	32
FIGURE 3. Binding of cationic ligands induces a red shift in the Soret band of the unbound protein.....	39
FIGURE 4. Crystal structures of selected ligands from Table 2 bound to CCP W191G	40
FIGURE 5. Superposition of the apo- structure (carbon atoms colored in cyan) with the phenol complex	49
FIGURE 6. The variation of ligand enrichment (continuous lines) and decoy downgrading (broken lines) with protein dielectric	52
FIGURE 7. Ranks of the CCP W191G cavity ligands scored using Gaussian charges and desolvation energies plotted against the ranks obtained using AMSOL charges and desolvation energies	53

Chapter 2

FIGURE 1. The model cavity sites.....	91
FIGURE 2. Retrospective enrichment of ligands and decoys	95
FIGURE 3. Predicted and experimental ligand orientations for the hydrophobic L99A cavity site.....	101
FIGURE 4. Predicted and experimental ligand orientations for the polar L99A/M102Q cavity site.....	107
FIGURE 5. Predicted and experimental ligand orientations for the anionic CCP cavity	112
FIGURE 6. The topologically similar ligands and decoys.....	119

Chapter 3

FIGURE 1. T4 Lysozyme L99A/M102Q binding site shown in complex with phenol and one ordered water molecule.....	148
FIGURE 2. Representative ITC measurements and fit.....	151
FIGURE 3. Comparison of predicted to experimental binding modes for the absolute binding free energy predictions.....	154
FIGURE 4. Comparison of predicted to experimental binding modes for the relative binding free energy predictions starting from reference compounds catechol & phenol	162
FIGURE 5. Relative binding free energy predictions from reference compounds Catechol and Phenol.....	164
FIGURE 6. Crystallographic orientations of the reference ligands phenol and catechol overlaid on the apo reference structure	172

Chapter 4

FIGURE 1. The open cavity mutant, CCP W191G-GA	191
FIGURE 2. Comparison of enrichment of actives and decoys in CCP wt* and CCP-GA	196
FIGURE 3. Comparison of the enrichment for CCP-GA with four solvation methods	199
FIGURE 4. Crystal structures of neutral ligands.....	201
FIGURE 5. UV-Vis titration of the heme Soret band in CCP-GA by 3-fluoro-catechol	202
FIGURE 6. Docking vs crystal pose of 4-hydroxybenzaldehyde in CCP-GA.....	205

Appendix A

FIGURE S1. Scatter plot of experimentally determined binding constants for the previously known, “test set” ligands1-3 compared to $ E_{tot} $, the absolute value of the binding energy calculated by DOCK	211
FIGURE S2. Docking-based Enrichment of ligands and downgrading of decoys in the cavities of CCP W191G, T4 lysozyme L99A, and T4 lysozyme L99A/M102Q	212
FIGURE S3. Enrichment of ligands and downgrading of decoys docked in the L99A and the L99A/M102Q cavities of T4 lysozyme with the desolvation energy calculated according to equation 2 and equation 3	213

Appendix B

FIGURE S1. Distribution of pair-wise Tanimoto similarities among the top-ranked docked and rescored molecules, calculated using ECFP4 fingerprints	221
FIGURE S2. Flowchart and parameters for AMBERDOCK rescoring.....	222
FIGURE S3. Preliminary scoring protocols for AMBERDOCK	223

Appendix C

FIGURE S1. The L99A/Gln102 binding site. Sidechains with dihedrals that are explicitly sampled by confine-and-release protocol are shown in yellow	232
FIGURE S2. Thermodynamic cycle for computing relative binding free energies.....	232
FIGURE S3. Representative plots of convergence with simulation time of free energy estimates for alchemical transformation component calculations	233

Introduction

I. Understanding molecular recognition

When I applied to the chemistry and chemical biology program at UCSF I knew that my interests were in the area of developing tools to understand biology, in particular for drug discovery and rational drug design. Having graduated with a B.S. in chemistry, my inclination was to think of the problem from the chemistry side of the equation, designing “chemical tools”—small molecules specifically designed to interact with a particular protein target—to investigate the recognition of a substrate by a protein, perhaps to elucidate the mechanism of catalysis, or as an inhibitor of protein function for a therapeutically relevant target. In the Shoichet group, I was introduced to an alternative approach; to understand molecular recognition in a biological context we would instead design the receptor, in this case an artificial binding site in a protein, which could selectively bind small molecules based on properties we engineered into the cavity. The knowledge we then gained from these systems would be leveraged into advances in computational methods for drug discovery.

Molecular recognition is a broadly used term, but in general it describes the interaction of two molecules through non-covalent bonding, such as van der Waals forces, electrostatics, hydrogen bonding, and hydrophobic burial of non-polar groups. The definition of molecular recognition also implies a best-fit between the receptor and ligand, where one molecule finds an optimal conformation in the binding site of another molecule. Fischer first proposed this view of the complementarity between enzyme and substrate with the lock-and-key model in 1894.¹ Although he assumed a rigid receptor,

which has long since been replaced with knowledge of protein flexibility and induced-fit upon ligand binding², the original connection between selectivity for a receptor and complementarity of the ligand fit still stands as a milestone to the way we think about the interaction of biological molecules.

Some of the earliest examples of experimental models for molecular recognition were the cyclic polyether host-guest systems; work resulting in the 1987 Nobel Prize in Chemistry, awarded to Charles Pedersen, Donald J. Cram and Jean-Marie Lehn for the discovery, synthesis and uses of crown ethers and cryptands.³⁻⁶ As early as 1967, Charles Pedersen determined that crown ethers (the “host” molecule) had the ability to bind alkali metal cations (the “guest”). Cryptands, analogs of crown ethers, bury the “guest” within the center of the host. The cryptand host-guest systems are more selective, and the enclosure of the ligand within the structure also allows discrimination between size of the guest molecule that can bind.⁷ Since then, host-guest systems have become popular as a means to investigate particular properties of molecular recognition. One intriguing example are the class of host-guest systems known as molecular tweezers, which are open “host” molecules that bind “guests” using non-covalent bonding, such as hydrogen bonds or van der Waals interactions.⁸⁻¹⁵ These systems have proved useful for computational studies of molecular recognition¹⁶⁻¹⁸; however, they lack the biological context that is critical to the development of computational methods for drug discovery.

II. Computational methods for drug discovery

The goal of molecular docking is to computationally recapitulate molecular recognition between two molecules of interest, the protein target (receptor) and small molecule (ligand). There are many types of programs that attempt this, from DOCK,

which was among the very first docking program developed by Kuntz *et al*, which was simply a shape-based descriptor method¹⁹⁻²¹, to more sophisticated molecular mechanics or molecular dynamics methods, using generalized born or Poisson Boltzmann methods for implicit solvent treatment²²⁻²⁸, and finally, full simulations of the ligand and receptor in implicit or explicit water.^{16-18,29-44} As the methods become more physically realistic the computational cost increases; MM-GBSA methods are orders of magnitude slower than simple docking methods, but still allow a relatively large number of molecules to be screened compared to full free energy calculations in explicit water, which can only realistically be used for small sets of compounds.^{45,46}

In drug discovery, computationally screening large libraries of compounds has become relatively inexpensive, thanks to advances in computing (the advent of modern processors and parallel computing); however, virtually screening a large database of small molecules against a protein target of interest requires speed which generally translates into sacrificing accuracy. The reason that there is still interest in this approach is that many more molecules can be considered, compared to the similar experimental assay approach, a high through-put screen (HTS), and the hit rate from the virtual screen is typically at least 10 fold better, in cases where the two approaches have been compared.⁴⁷⁻⁵⁰ Also, the focus on screening commercially available compounds⁵¹ enables feasible follow-up of interesting hits. But the lack of accuracy results in a high number of failures. These failures, false negatives—true ligands missed by the docking methods—and the generally more exasperating case, false positives—compounds that don't in fact bind, but score well and flood the top of the docking hit lists—are expensive failures, because they waste both time and resources. Rescoring with more sophisticated

algorithms is one way to try and “rescue” false negatives and improve enrichment of likely ligands; these methods are considered in this work (Chapter 2), but have their own limitations and inaccuracies. Therefore, understanding why the docking method fails at molecular recognition is the very heart of the problem, and part of the solution lies in specially designed model systems.

III. Polar and charged model systems for molecular docking

Two model systems in T4 lysozyme, the hydrophobic cavity created by the L99→A mutation, and the slightly polar cavity created by the addition M102→Q mutation, were already well established when I joined the Shoichet lab. They had been used successfully to evaluate ligand charge models and desolvation⁵², and docking against a “flexible” receptor.⁵³ However, both systems were primarily hydrophobic, although the L99A/M102Q cavity was slightly polar and had hydrogen bonding potential; what was missing was a cavity with an isolated and buried charge. So to expand our suite of model systems we needed a charged cavity, preferably one that could potentially bind charged ligands. Work on a new charged mutant in T4 lysozyme, M102→E by our collaborators in Brian Matthews group at the University of Oregon, was proving to be extremely difficult; placing a charge in a hydrophobic pocket destabilized the protein, even with additional stability mutations engineered into the protein, and recent evidence suggests the glutamate is protonated.⁵⁴ So we turned to a new protein, cytochrome *c* peroxidase (CCP), an anionic cavity created by the W191→G mutation by our collaborators in the Goodin lab at Scripps, which would selectively bind aromatic monocations.⁵⁵⁻⁵⁷

What sold me on this project was that failure was not only acceptable in these systems, it was often preferred; the elegance of these models was that we could learn as much, or more, from failed predictions as from successful ones. Like the T4 lysozyme cavities, the importance of our approach in the CCP W191G was two-fold: 1) while slightly more complicated, the charged cavity was still simple enough so that individual aspects of the molecular recognition process could be isolated and studied—in particular the balance between ligand desolvation and the favorable electrostatic interaction between a charged ligand and anionic protein—and 2) we could use what we learned about molecular recognition and apply it to computational methods, from our simple DOCK scoring function to MM-GBSA rescoring, and eventually all the way to full free energy calculations. The other advantages of this binding site was that the assay was extremely simple⁵⁷ so that we could extensively test our computational predictions in a prospective manner, by picking molecules from the docking hit lists and experimentally assaying them for binding. Finally, protein crystallography was even easier in this protein than T4 lysozyme, making it incredibly easy to compare the predicted pose to actual experimentally determined binding modes.

There were drawbacks to this system however; when we began to consider doing alchemical free energy calculations with this binding site reality hit. A previous study in the L99A binding site had been extremely successful.⁵⁸ However, charged ligands, the presence of ordered displaceable waters, and finally, a heme group that formed the back of the cavity made the CCP model system too complicated to go to straight from the very simple L99A cavity. The third chapter of my thesis is performed instead in the polar T4 lysozyme L99A/M102Q model system, where we attempted to predict absolute and

relative binding free energies of small molecules to a polar cavity. The following section is a guide to the chapters of my dissertation.

IV. Guide to the Chapters

In *Chapter 1: Probing Molecular Docking in a Charged Model Binding Site*, we described the development of the CCP W191G binding site as a model system to evaluate our own DOCK scoring function and found, surprisingly, that DOCK outperformed our expectations. Of the 16 compounds picked from the top of the hit list, 15 were determined to bind. However, interesting and notable failures were the two neutral molecules that were found to weakly bind but were missed by docking, and 3 ligands with highly localized charges that scored poorly, due to a high desolvation penalty, but did in fact bind; for all three a water molecule contacting the charged amino group on the ligand is observed, helping to mitigate the cost of desolvation.

Chapter 2: Rescoring Docking Hit Lists for Model Cavity Sites: Predictions and Experimental Testing, details the successes and failures inherent in rescoring docking hit lists using MM-GBSA methods; while some molecules were “rescued” from the bottom of the docking hit lists and new chemotypes emerged, partial atomic charge distribution on the compounds, failure to sample large protein conformational changes, and the constraint to sampling near to the starting dock pose proved to be detrimental in several cases to accurate predictions.

In *Chapter 3: Predicting absolute and relative binding free energies in a polar model binding site*, we returned to the T4 lysozyme L99A/M102Q model system for absolute and relative binding free energy calculations, in which we attempted to prospectively predict binding energy, pose, and rank-ordering for a set of unknown

molecules using alchemical free energy methods. Although we set out to investigate free energy calculations for a polar site, the defining characteristic of all the ligands, with only one exception, was that they enlarged the cavity to varying extent. Accurate prediction of absolute binding free energies and even rank-ordering of the compounds proved to be difficult in this system for many reasons, discussed in this chapter in detail. The relative free energy calculations fared slightly better.

In *Chapter 4: Future Directions*, a new model system is introduced in CCP, and some preliminary results are discussed, for an anionic cavity that is open to bulk solvent and contains multiple ordered waters that are displaced upon ligand binding.⁵⁹

V. References

1. Fischer, E. (1894). Einfluss der Configuration auf die Wirkung der Enzyme. *Ber. Dt. Chem. Ges.* **27**, 2985-2993.
2. Koshland, D. E. (1958). Application of a Theory of Enzyme Specificity to Protein Synthesis. *Proc. Natl. Acad. Sci.* **44**, 98-104.
3. Pedersen, C. J. (1967). *J. Am. Chem. Soc.* **89**, 7017-7036.
4. Pedersen, C. J. (1967). *J. Am. Chem. Soc.* **89**, 2495-2496.
5. Pedersen, C. J. (1988). Macrocyclic Polyethers: Dibenzo-18-Crown-6 Polyether and Dicyclohexyl-18-Crown-6 Polyether. *Org Synth* **6**, 395.
6. Lehn, J. M. (1995). *Supramolecular Chemistry: Concepts and Perspectives*, Wiley-VCH, Weinheim.
7. Dietrich, B., Ed. (1996). "Cryptands" in *Comprehensive Supramolecular Chemistry*. Vol. 1. Edited by Gokel, G. W. Oxford: Elsevier.
8. Petitjean, A., Khoury, R. G., Kyritsakas, N. & Lehn, J. M. (2004). Dynamic Devices. Shape Switching and Substrate Binding in Ion-Controlled Nanomechanical Molecular Tweezers. *J. Am. Chem. Soc.* **126**, 6637–6647.
9. Sygula, A., Fronczek, F. R., Sygula, R., Rabideau, P. W. & Olmstead, M. M. (2007). A Double Concave Hydrocarbon Buckycatcher. *J. Am. Chem. Soc.* **129**, 3842–3843.
10. Frank-Gerrit, K. & Björn, K. (2003). Molecular Tweezers and Clips as Synthetic Receptors. Molecular Recognition and Dynamics in Receptor-Substrate Complexes. *Acc. Chem. Res.* **36**, 919–932.

11. Chen, C. W. & Whitlock, H. W. (1978). Molecular Tweezers - A Simple-Model of Bifunctional Intercalation. *J. Am. Chem. Soc.* **100**, 4921.
12. Zimmerman, S. C. (1993). Rigid molecular tweezers as hosts for the complexation of neutral guests. *Top. Curr. Chem.* **165**, 71.
13. Zimmerman, S. C. & VanZyl, C. M. (1987). Rigid molecular tweezers: synthesis, characterization, and complexation chemistry of a diacridine. *J. Am. Chem. Soc.* **109**, 7894.
14. Zimmerman, S. C. & Wu, W. (1989). A rigid molecular tweezers with an active site carboxylic acid: exceptionally efficient receptor for adenine in an organic solvent. *J. Am. Chem. Soc.* **111**, 8054.
15. Huang, X., Fujioka, N., Pescitelli, G., Koehn, F., Williamson, R. T., Nakanishi, K. & Berova, N. (2002). Absolute Configurational Assignments of Secondary Amines by CD-sensitive Dimeric Zinc Porphyrin Host. *J. Am. Chem. Soc.* **124**, 10320–10335.
16. Chang, C. E. & Gilson, M. K. (2004). Free energy, entropy, and induced fit in host-guest recognition: calculations with the second-generation mining minima algorithm. *J. Am. Chem. Soc.* **126**, 13156-64.
17. Moghaddam, S., Inoue, Y. & Gilson, M. K. (2009). Host-guest complexes with protein-ligand-like affinities: computational analysis and design. *J. Am. Chem. Soc.* **131**, 4012-21.
18. Rekharsky, M. V., Mori, T., Yang, C., Ko, Y. H., Selvapalam, N., Kim, H., Sobransingh, D., Kaifer, A. E., Liu, S., Isaacs, L., Chen, W., Moghaddam, S., Gilson, M. K., Kim, K. & Inoue, Y. (2007). A synthetic host-guest system

- achieves avidin-biotin affinity by overcoming enthalpy-entropy compensation. *Proc. Natl. Acad. Sci. U S A* **104**, 20737-42.
19. Shoichet, B. K. & Kuntz, I. D. (1991). Protein docking and complementarity. *J Mol. Biol.* **221**, 327-346
 20. Shoichet, B. K., Bodian, D. L. & Kuntz, I. D. (1992). Molecular docking using shape descriptors. *J Comp. Chem.* **13**, 380-397
 21. Meng, E. C., Shoichet, B. K. & Kuntz, I. D. (1992). Automated docking with grid-based energy evaluation. *J Comp. Chem.* **13**, 505-524
 22. Wang, J., Kang, X., Kuntz, I. D. & Kollman, P. A. (2005). Hierarchical database screenings for HIV-1 reverse transcriptase using a pharmacophore model, rigid docking, solvation docking, and MM-PB/SA. *J Med Chem* **48**, 2432-44.
 23. Kalyanaraman, C., Bernacki, K. & Jacobson, M. P. (2005). Virtual screening against highly charged active sites: identifying substrates of alpha-beta barrel enzymes. *Biochemistry* **44**, 2059-71.
 24. Huang, N., Kalyanaraman, C., Irwin, J. J. & Jacobson, M. P. (2006). Physics-Based Scoring of Protein-Ligand Complexes: Enrichment of Known Inhibitors in Large-Scale Virtual Screening. *J. Chem. Inf. Model.* **46**, 243-253.
 25. Perola, E. (2006). Minimizing false positives in kinase virtual screens. *Proteins: Structure, Function, and Bioinformatics* **64**, 422-435.
 26. Lee, M. R. & Sun, Y. (2007). Improving Docking Accuracy through Molecular Mechanics Generalized Born Optimization and Scoring. *J. Chem. Theory Comput.* **3**, 1106-1119.

27. Lyne, P. D., Lamb, M. L. & Saeh, J. C. (2006). Accurate Prediction of the Relative Potencies of Members of a Series of Kinase Inhibitors Using Molecular Docking and MM-GBSA Scoring. *J. Med. Chem.* **49**, 4805-4808.
28. Jacobson, M. P., Kaminski, G. A., Friesner, R. A. & Rapp, C. S. (2002). Force Field Validation Using Protein Side Chain Prediction. *J. Phys. Chem. B* **106**, 11673-11680.
29. Jayachandran, G., Shirts, M. R., Park, S. & Pande, V. S. (2006). Parallelized-over-parts computation of absolute binding free energy with docking and molecular dynamics. *J. Chem. Phys.* **125**, 084901.
30. Mobley, D. L., Chodera, J. D. & Dill, K. A. (2006). On the use of orientational restraints and symmetry corrections in alchemical free energy calculations. *J. Chem. Phys.* **125**, 084902.
31. Warshel, A., Sussman, F. & King, G. (1986). Free energy of charges in solvated proteins: microscopic calculations using a reversible charging process. *Biochemistry* **25**, 8368-8372, ISSN 0006-2960, 1986.
32. Deng, Y. & Roux, B. (2006). Calculation of standard binding free energies: Aromatic molecules in the T4 lysozyme L99A mutant. *J Chem Theory Comput* **5**, 1255-1273.
33. Wang, J., Deng, Y. & Roux, B. (2006). Absolute Binding Free Energy Calculations Using Molecular Dynamics Simulations with Restraining Potentials. *Biophysical Journal* **91**, 2798 -814.

34. Bash, P. A., Singh, U. C., Brown, F. K., Langridge, R. & Kollman, P. A. (1986). Calculation of the Relative Change in Binding Free Energy of a Protein-Inhibitor Complex *Science* **235**, 574-576.
35. Wong, C. F. & McCammon, A. J. (1986). Dynamics and design of enzymes and inhibitors. *J. Am. Chem. Soc.* **108**, 3830-3832.
36. Hermans, J. & Subramaniam, S. (1986). The Free Energy of Xenon Binding to Myoglobin from Molecular Dynamics Simulation. *Isr. J. Chem.* **27**, 225-227.
37. Nicholls, A., Mobley, D. L., Guthrie, P. J., Chodera, J. D. & Pande, V. S. (2008). Predicting Small-Molecule Solvation Free Energies: An Informal Blind Test for Computational Chemistry. *J. Med. Chem.*
38. Kamerlin, S. C., Haranczyk, M. & Warshel, A. (2009). Are mixed explicit/implicit solvation models reliable for studying phosphate hydrolysis? A comparative study of continuum, explicit and mixed solvation models. *Chem Phys Chem* **10**, 1125-34.
39. Lee, F. S., Chu, Z. T., Bolger, M. B. & Warshel, A. (1992). Calculations of antibody-antigen interactions: microscopic and semi-microscopic evaluation of the free energies of binding of phosphorylcholine analogs to McPC603. *Protein Eng.* **5**, 215-28.
40. Rosta, E., Haranczyk, M., Chu, Z. T. & Warshel, A. (2008). Accelerating QM/MM free energy calculations: representing the surroundings by an updated mean charge distribution. *J. Phys. Chem. B.* **112**, 5680-92.
41. Hermans, J. & Wang, L. (1997). Inclusion of the Loss of Translational and Rotational Freedom in Theoretical Estimates of Free Energies of Binding.

- Application to a Complex of Benzene and Mutant {T4 Lysozyme}. *J Am Chem Soc* **119**, 2707-2714.
42. Mann, G. & Hermans, J. (2000). Modeling Protein-small Molecule Interactions: Structure and Thermodynamics of Noble Gases Binding in a Cavity in Mutant Phage T4 Lysozyme L99A. *J. Mol. Biol.* **302**, 979-989.
43. Warshel, A., Sharma, P., Chu, Z. & Aqvist, J. (2007). Electrostatic Contributions to Binding of Transition State Analogues Can Be Very Different from the Corresponding Contributions to Catalysis: Phenolates Binding to the Oxyanion Hole of Ketosteroid Isomerase. *Biochemistry* **46**, 1466-1476.
44. Kim, J. T., Hamilton, A. D., Bailey, C. M., Domoal, R. A., Wang, L., Anderson, K. S. & Jorgensen, W. L. (2006). FEP-guided selection of bicyclic heterocycles in lead optimization for non-nucleoside inhibitors of HIV-1 reverse transcriptase. *J. Am. Chem. Soc.* **128**, 15372-15373.
45. Huang, N. & Jacobson, M. P. (2007). Physics-based methods for studying protein-ligand interactions. *Current Opinion in Drug Discovery & Development* **10**, 325-331.
46. Shirts, M. R., Mobley, D. L. & Chodera, J. D. (2007). Alchemical Free Energy Calculations: Ready for Prime Time? In *Annual Reports in Computational Chemistry*, Vol. 3, pp. 41-53. Elsevier.
47. Babaoglu, K., Anton Simeonov, A., Irwin, J. J., Nelson, M. E., Feng, B., Thomas, C. J., Cancian, L., Costi, M. P., Maltby, D. A., Jadhav, A., James Inglese, J., Austin, C. P. & Shoichet, B. K. (2008). Comprehensive Mechanistic Analysis of

- Hits from High-Throughput and Docking Screens against Beta-Lactamase. *J. Med. Chem.* **51**, 2502-2511.
48. Doman, T. N., McGovern, S. L., Witherbee, B. J., Kasten, T. P., Kurumbail, R., Stallings, W. C., Connolly, D. T. & Shoichet, B. K. (2002). Molecular docking and high-throughput screening for novel inhibitors of protein tyrosine phosphatase-1B. *J Med Chem* **45**, 2213-2221.
 49. Evensen, E., Eksterowicz, J. E., Stanton, R. V., Oshiro, C., Grootenhuis, P. D. & Bradley, E. K. (2003). Comparing performance of computational tools for combinatorial library design. *J Med Chem* **46**, 5125-8.
 50. Paiva, A. M., Vanderwall, D. E., Blanchard, J. S., Kozarich, J. W., Williamson, J. M. & Kelly, T. M. (2001). Inhibitors of dihydrodipicolinate reductase, a key enzyme of the diaminopimelate pathway of *Mycobacterium tuberculosis*. *Biochim Biophys Acta* **1545**, 67-77.
 51. Irwin, J. J. & Shoichet, B. K. (2005). ZINC - A Free Database of Commercially Available Compounds for Virtual Screening. *J Chem Inf Model* **45**, 177-182.
 52. Wei, B. Q., Baase, W. A., Weaver, L. H., Matthews, B. W. & Shoichet, B. K. (2002). A model binding site for testing scoring functions in molecular docking. *J. Mol. Biol.* **322**, 339-355.
 53. Wei, B. Q., Weaver, L., Ferrari, A. M., Matthews, B. M. & Shoichet, B. K. (2004). Testing a flexible-receptor docking algorithm in a model binding site. *J. Mol. Biol.* **337**, 1161-82.

54. Liu, L., Baase, W.A., Michael, M. M. & Matthews, B.W. (2009). Use of stability mutations to engineer a charged group within a ligand-binding hydrophobic cavity in T4 lysozyme. *Biochemistry* **48**, 8842-8851.
55. Fitzgerald, M. M., Churchill, M. J., McRee, D. E. & Goodin, D. B. (1994). Small molecule binding to an artificially created cavity at the active site of cytochrome c peroxidase. *Biochemistry* **33**, 3807-18.
56. Fitzgerald, M. M., Trester, M. L., Jensen, G. M., McCree, D. E. & Goodin, D. B. (1995). The role of aspartate-235 in the binding of cations to an artificial cavity at the radical site of cytochrome c peroxidase. *Protein Science* **4**, 1844-1850.
57. Musah, R. A., Jensen, G. M., Bunte, S. W., Rosenfeld, R. J. & Goodin, D. B. (2002). Artificial protein cavities as specific ligand-binding templates: characterization of an engineered heterocyclic cation-binding site that preserves the evolved specificity of the parent protein. *J Mol Biol* **315**, 845-57.
58. Mobley, D. L., Graves, A. P., Chodera, J. D., McReynolds, A. C., Shoichet, B. K. & Dill, K. A. (2007). Predicting absolute ligand binding free energies to a simple model site. *J. Mol. Biol.* **371**, 1118-1134.
59. Rosenfeld, R. J., Hays, A. A., Musah, R. A. & Goodin, D. B. (2002). Excision of a proposed electron transfer pathway in cytochrome c peroxidase and its replacement by a ligand-binding channel. *Protein Science* **11**, 1251-1259.

Gloss to chapter 1

The following chapter chronicles the introduction of the first singly charged model binding site used successfully to evaluate the DOCK scoring function in our lab, cytochrome c peroxidase W191G, (CCP W191G). This project and the paper that came out of it were almost entirely due to two sources of expertise and my co-authors: Ruth Brenk, then a postdoctoral researcher in the Shoichet lab, was my mentor on this project. While we were both new to this model system and learned its idiosyncrasies together, she taught me everything that I applied to this project (and to my future work): docking, molecular biology (protein expression and purification), protein crystallography, and how to do a binding assay and determine ligand affinity. This project was also a collaboration with the David Goodin's group, in particular Stefan Vetter, who sent us the initial plasmid and also lent his technical expertise to every experiment, from the initial binding assay, to protein preparation, and finally to protein crystallography, resulting in 6 of the 15 ligand co-complex structures published in the paper.

Starting this project we had several pieces of information about this site, due to the extensive work done by the Goodin lab where this artificial cavity was first engineered.¹ For one, we knew that the cavity primarily bound small, aromatic, cationic heterocycles.² This made sense from the standpoint of the native function of cytochrome c peroxidase; the characteristics of the cavity itself were designed to stabilize a cationic tryptophan radical.² Removal of the cationic tryptophan left a polar cavity lined with carbonyl groups and an unfulfilled charged at the catalytic aspartate, Asp235. At the time we began this work, nearly two dozen crystal structures for this particular binding site in

complex with ligands had been solved;^{1,2} in all of these the charge on the ligand interacted with the charge on the Asp235, forming a salt bridge. In the apo structure, five ordered waters and one potassium ion were observed; binding of ligands displaced the waters and potassium, with the exception of one water, wat308, which was conserved in most ligand bound structures.² We also knew, thanks to work by Goodin et. al., that the site was titratable by potassium ion; the cavity had weak, but measurable affinity for K⁺, which could compete for binding with charged ligands.³ Finally, and perhaps initially most interesting to us, the charge of the known ligands was limited to monocations, both dications and neutral (but polar) molecules were decoys for this site.^{2,4}

It was this subtle difference that we were intrigued by; we knew from work done in the T4 lysozyme polar cavity, L99A/M102Q, DOCK had trouble distinguishing polar ligands from only slightly more polar decoys (witness catechol as a ligand compared to the isosteric decoy 2-aminophenol that DOCK clearly predicts will bind).⁵ This polar and charged cavity seemed ideal to address the balance in the docking algorithm between the favorable electrostatic interactions of charged ligand and protein and the penalty such charged ligands must pay in the desolvation energy upon binding to the cavity. The questions we sought to answer were the following:

- 1) Could DOCK successfully distinguish monocations from neutral and dicationic decoys?
- 2) Could we accurately predict new ligands and their correct binding modes?
- 3) Are neutral compounds really decoys for this site? Could we find neutral ligands that would bind detectably to this cavity?

- 4) Could we find new chemotypes for this site? Many of the ligands had similar scaffolds, based around the aminopyridine or thiazole structure.

The first questions were addressed quite clearly; yes, DOCK could distinguish monocation ligands from dication and neutral decoys quite well. We showed this both by retrospective enrichment of ligands compared to known decoys and by prospectively testing compounds from the top of the docking hit list and by solving the crystal structures of the new ligands in complex with the cavity. We also found two neutral molecules that bound to the cavity, and although they ranked poorly by docking, this was something that had previously been unobserved and the poor rank seemed at least to correspond with their weak affinity.

However, when we submitted the manuscript to the Journal of Molecular Biology, the last point came up in the reviewers' comments; the new ligands we discovered appeared, at a cursory level, to be very much like the known ligands. This is a point that comes up frequently when working in the model cavity sites, particularly in CCP; the size of the site precludes large ligands from binding. Therefore, with fewer atoms, and the requirement for a charged group to interact with the aspartate, it would seem logical that the diversity of the ligands would be restricted. However, we were able to show that the new ligands were sufficiently different by Tanimoto analysis and that they would not have been found using simple similarity searching methods. In particular, the two weakly binding neutral molecules, phenol and 3-fluorocatechol, would not have been found using these methods, although admittedly they were false negatives by docking as well.

The most interesting result though, was most likely the realization that DOCK itself did exceptionally well by any standard at distinguishing ligands and decoys. Of the 16 molecules tested from the top of the hit list 15 bound, and not one of the top 100 molecules was a dication or neutral, a red flag for a decoy; in fact, the prevalence of monocations, all seemingly likely ligands, in the top of hit list was completely unprecedented, based on our expectations from both the model binding sites in T4 lysozyme and in the typical low hit rate from virtual screening campaigns against real drug targets.⁶⁻⁹

Overall the enormous success we had in this cavity was both elating and unnerving; after all, these model systems are carefully crafted so that while successes are all well and good, to truly learn from this system we need failures that we can trace back to the source and understand. The notable failures in this set were the two neutrals and the ligands with localized charges that were false negatives; these last ligands brought in additional waters, mitigating the high desolvation penalty they would otherwise pay upon binding (and did pay in the docking score). From our standpoint they were not unexpected; we didn't model any new waters during docking, so predicting a new water interaction would be impossible for DOCK. However, these failures gave us interesting areas to pursue down the road, themes that reoccur in Chapter 2 & 4. The Supplementary Material for Chapter 1 appears in Appendix A.

References

1. Fitzgerald, M. M., Churchill, M. J., McRee, D. E. & Goodin, D. B. (1994). Small molecule binding to an artificially created cavity at the active site of cytochrome c peroxidase. *Biochemistry* **33**, 3807-18.
2. Musah, R. A., Jensen, G. M., Bunte, S. W., Rosenfeld, R. J. & Goodin, D. B. (2002). Artificial protein cavities as specific ligand-binding templates: characterization of an engineered heterocyclic cation-binding site that preserves the evolved specificity of the parent protein. *J Mol Biol* **315**, 845-57.
3. Fitzgerald, M. M., Trester, M. L., Jensen, G. M., McCree, D. E. & Goodin, D. B. (1995). The role of aspartate-235 in the binding of cations to an artificial cavity at the radical site of cytochrome c peroxidase. *Protein Science* **4**, 1844-1850.
4. Rosenfeld, R., Goodsell, D., Musah, R., Garret, M., Goodin, D. & Olson, A. (2003). Automated docking of ligands to an artificial active site: augmenting crystallographic analysis with computer modeling. *J Comput Aided Mol Des.*
5. Graves, A. P., Brenk, R. & Shoichet, B. K. (2005). Decoys for docking. *J Med Chem* **48**, 3714-28.
6. Wang, R., Lu, Y., Fang, X. & Wang, S. (2004). An extensive test of 14 scoring functions using the PDBbind refined set of 800 protein-ligand complexes. *J Chem Inf Comput Sci* **44**, 2114-25.
7. Shoichet, B. K. (2004). Virtual screening of chemical libraries. *Nature* **432**, 862-5.

8. Perola, E., Walters, W. P. & Charifson, P. S. (2004). A detailed comparison of current docking and scoring methods on systems of pharmaceutical relevance. *Proteins* **56**, 235-49.
9. Doman, T. N., McGovern, S. L., Witherbee, B. J., Kasten, T. P., Kurumbail, R., Stallings, W. C., Connolly, D. T. & Shoichet, B. K. (2002). Molecular docking and high-throughput screening for novel inhibitors of protein tyrosine phosphatase-1B. *J Med Chem* **45**, 2213-2221.

Chapter 1:

Probing molecular docking in a charged model binding site

Ruth Brenk^{1†}, Stefan W. Vetter^{2‡}, Sarah E. Boyce¹⁺, David B. Goodin^{*2}, Brian K. Shoichet^{*1}

¹ University of California San Francisco, QB3 Building, Dept. of Pharmaceutical Chemistry, 1700 4th Street, San Francisco, CA, 94143, USA

² The Scripps Research Institute, Department of Molecular Biology, 10550 North Torrey Pines Road, La Jolla, CA 92037, USA

⁺ The Graduate Group in Chemistry and Chemical Biology

[†] Current address: Division of Biological Chemistry & Molecular Microbiology, University of Dundee, Dow Street, Dundee DD1 5EH, United Kingdom

[‡] Current address: Florida Atlantic University, Department of Chemistry, 777 Glades Road, Boca Raton, FL 33431

* Corresponding authors:

David B. Goodin; phone (858) 784-9892; fax: (858) 784-2857; e-mail: dbg@scripps.edu

Brian K. Shoichet; phone: (415) 514-4126; fax: (415) 514-4260; e-mail: shoichet@cgl.ucsf.edu

1.1 Abstract

A model binding site was used to investigate charge-charge interactions in molecular docking. This simple site, a small (180 Å³) engineered cavity in cytochrome *c* peroxidase (CCP), is negatively charged and completely buried from solvent, allowing us to explore the balance between electrostatic energy and ligand desolvation energy in a system where many of the common approximations in docking do not apply. A database with about 5300 molecules was docked into this cavity. Retrospective testing with known ligands and decoys showed that overall the balance between electrostatic interaction and desolvation energy was captured. More interesting were prospective docking screens that looked for novel ligands, especially those that might reveal problems with the docking and energy methods. Based on screens of the 5300 compound database, both high-scoring and low-scoring molecules were acquired and tested for binding. Out of 16 new, high-scoring compounds tested, 15 were observed to bind. All of these were small heterocyclic cations. Binding constants were measured for a few of these— they ranged between 20 to 60 μM. Crystal structures were determined for ten of these ligands in complex with the protein. The observed ligand geometry corresponded closely to that predicted by docking. Several low-scoring alkyl amino cations were also tested and found to bind. The low docking score of these molecules owed to the relatively high charge density of the charged amino group and the corresponding high desolvation penalty. When the complex structures of those ligands were determined, a bound water molecule was observed interacting with the amino group and a backbone carbonyl group of the cavity. This water molecule mitigates the desolvation penalty and improves the interaction energy relative to that of the “naked” site used in the docking screen. Finally, six low-scoring neutral

molecules were also tested, with a view to looking for false negative predictions.

Whereas most of these did not bind, two did (phenol and 3-fluorocatechol). Crystal structures for these two ligands in complex with the cavity site suggest reasons for their binding. That these neutral molecules do, in fact bind, contradicts previous results in this site and, along with the alkyl amines, provides instructive false negatives that help identify weaknesses in our scoring functions. Several improvements of these are considered.

Keywords

molecular docking; electrostatic; solvation; cytochrome *c* peroxidase: X-ray
crystallography

1.2 Introduction

Molecular docking is widely used to discover new ligands for biological targets with a known 3D structure.^{1;2} Notwithstanding important successes,^{1;3-10} docking screens remain hampered by the prediction of false positives and negatives.¹¹⁻¹⁷ This is tolerated for two reasons: docking focuses on easily available compounds and hit rates are often higher than those obtained by random high throughput screening.¹⁸⁻²⁰ Nevertheless, it is clear that improved scoring functions would have considerable impact. Isolating the effects of particular changes in scoring functions is difficult because of the entanglement of various energetic contributions in ligand-receptor binding. These include receptor and ligand desolvation, other entropic contributions, polar and non-polar interactions, the hydrophobic effect, and receptor flexibility, among others.^{21;22} Therefore it would be useful to have model systems that are simple enough to allow one to separate the different energetic contributions via experimental measurements and to isolate modifications in a new scoring function from other aspects of binding.

Examples of such simple systems are cavities engineered in the core of T4 lysozyme. The cavity created by the substitution Leu99→Ala is completely buried from solvent, uniformly hydrophobic and contains no ordered water molecules.²³ The ligands that bind to this pocket are small hydrophobic compounds like benzene or indene.²⁴ The cavity does not tolerate ligand polarity well: toluene binds to the cavity, but there is no evidence that phenol does. By the additional substitution Met102→Gln, a single polar atom was introduced in the wall of this cavity.²⁵ This cavity accommodates the hydrophobic ligands of the L99A mutant cavity, but also polar compounds like phenol or 3,5-difluoroaniline. The simplicity of these sites, combined with well established binding

assays and crystallization conditions, makes these pockets good model systems to test scoring functions both retro- and prospectively, and to guide their improvement.^{12; 25-27} In recent work, Gilson and colleagues have taken this approach one step further using organic host-guest complexes as model systems to explore enthalpy-entropy compensation.²⁸ The motivation behind each of these systems is to simplify molecular recognition to the point where individual driving forces can be isolated and studied.

The aspects of scoring functions and docking algorithms that can be probed in a model system are determined by its properties. The T4 lysozyme cavities provide systems to examine ligand binding in a hydrophobic and a slightly polar environment^{12; 25} and to investigate limited receptor flexibility.^{26; 27} To simulate other aspects of ligand binding, such as charge-charge interactions, new systems are needed.

A model site well-suited for this purpose is an engineered pocket in cytochrome *c* peroxidase (CCP) that was created by the substitution Trp191→Gly (Figure 1).²⁹ This substitution creates a small pocket that in some ways resembles those of the T4 lysozyme cavities. It has roughly the same volume as the lysozyme cavities (180 Å³ vs. 150 Å³) and it, too, is completely buried from solvent. Unlike the lysozyme cavities, the CCP W191G cavity is negatively charged and “wet”, containing five ordered water molecules and a potassium ion. The charge owes to the presence of Asp235, and the water molecules and the potassium ion ligate both the carboxylate group of this residue and several exposed backbone carbonyl groups. Twenty-three ligands and seventeen compounds that don't bind to this pocket are known.²⁹⁻³¹ Most ligands are small heterocycles bearing a single positive charge (Table 1). For 18 of these, the x-ray crystal structures of the cavity-complexes were determined. Typically, non-ligands, which we will refer to as

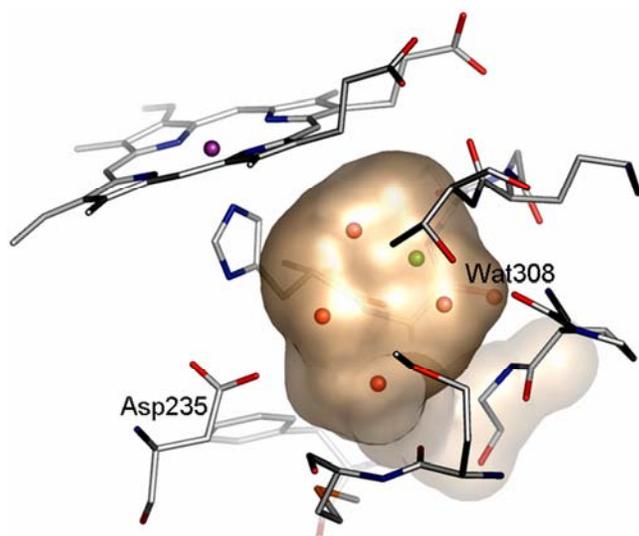


Figure 1. The cavity in CCP W191G. A transparent surface is displayed showing four ordered water molecules (red) and one potassium ion (green) in the cavity of the apo structure. (This Figure was made using PyMOL (www.pymol.org, as were Figures 4 and 5).)

“decoys”,¹² are small enough to fit in the pocket, but have the wrong net charge (0 or +2). This model site was used previously for two retrospective studies related to inhibitor design. Brooks and colleagues tested their λ -dynamics approach to predict binding affinities.^{32; 33} Olson and colleagues tested the ability of AutoDock³⁴ to reproduce crystallographically observed binding modes and to predict binding affinities of the known ligands.³¹

Here, we use the CCP W191G pocket for studying charge-charge and charge-polar interactions in docking screens of large compound databases. These electrostatic interactions are common in protein-ligand binding, but can be difficult to model using physics-based scoring functions, such as the one we use in this work.⁴¹ This scoring function, implemented in DOCK3.5.54,^{25; 35} includes van der Waals (E_{vdw}) and electrostatic terms (E_{elec}) and is corrected for ligand desolvation (ΔG_{solv}):

$$E = E_{elec} + E_{vdw} - \Delta G_{solv} \quad (1)$$

For charge-charge interactions, the large gain in electrostatic energy must be balanced against the corresponding large desolvation energy penalty. An additional complication is that the absolute error in calculating desolvation energies for charged compounds is usually higher than for neutral compounds.^{36, 37}

Running a virtual screening campaign against this relatively simple model system allowed us to address several questions that only emerge in database screens, when not only potential ligands, but also a vast number of “decoy” molecules are fit into the site and ranked. First, how well balanced are electrostatic and desolvation energy in the docking screen? Are molecules with the “right” overall charge picked out as likely ligands from among the decoys that dominate the database, or do either electrostatic or desolvation energy terms dominate? Second, can we discover any new chemotypes for this cavity? The known ligands were picked based on chemical intuition, and most resemble one-another. Screening a large database of compounds might allow us to find new classes of ligands. We docked a database with about 5300 neutral, single and double positively charged molecules, small enough to fit in the cavity, against this pocket, and tested high ranking compounds. Third, we were curious as to why no neutral molecules were found as ligands for this cavity. Such neutral molecules can form a charged-dipole hydrogen bond with Asp235 and would be easier to desolvate relative to charged ligands. Fourth, we investigated how the docking predictions change when we used a higher level of theory for calculating partial charges and desolvation energies of the docked molecules, or when the value of the dielectric constant in the binding pocket is changed.

To see how docking results for binding sites with different properties are affected by these changes, we included the T4 lysozyme cavities in the comparison study. Finally, we consider the false positive and negative predictions of the database screen against the CCP W191G cavity as a guide to future improvements of docking scoring functions.

1.3 Results

Retrospective tests

We began by evaluating the ability of the docking program to predict binding modes of known ligands and to recognize them as high-scoring “hits” in retrospective database screens. Eighteen known ligands and fifteen known decoys (test set) were seeded into a database of about 5300 neutral or positively charged molecules small enough to fit into the cavity in CCP W191G. Each database molecule was docked into the cavity in multiple orientations and conformations, scored for van der Waals and electrostatic complementarity and penalized for ligand desolvation energy. Because the CCP W191G cavity is small and completely buried, we did not consider differential receptor desolvation. The conformation of the cavity was held rigid, the potassium ion and all ordered water molecules except Wat308, which is conserved in all previous structures, were removed (Figure 1). Performance was evaluated based on the prediction of binding modes, enrichment of known ligands and downgrading of known decoys.

First, we checked the ability of the docking program to predict the binding modes for the ligands in the test set for which an unambiguous binding mode had been determined (Table 1).^{29; 30; 38} With AMSOL partial charges and desolvation energies for

Table 1. (legend following page)

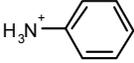
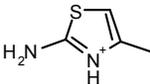
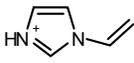
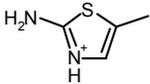
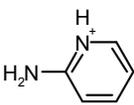
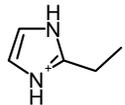
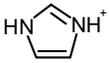
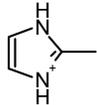
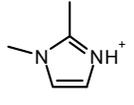
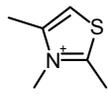
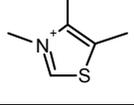
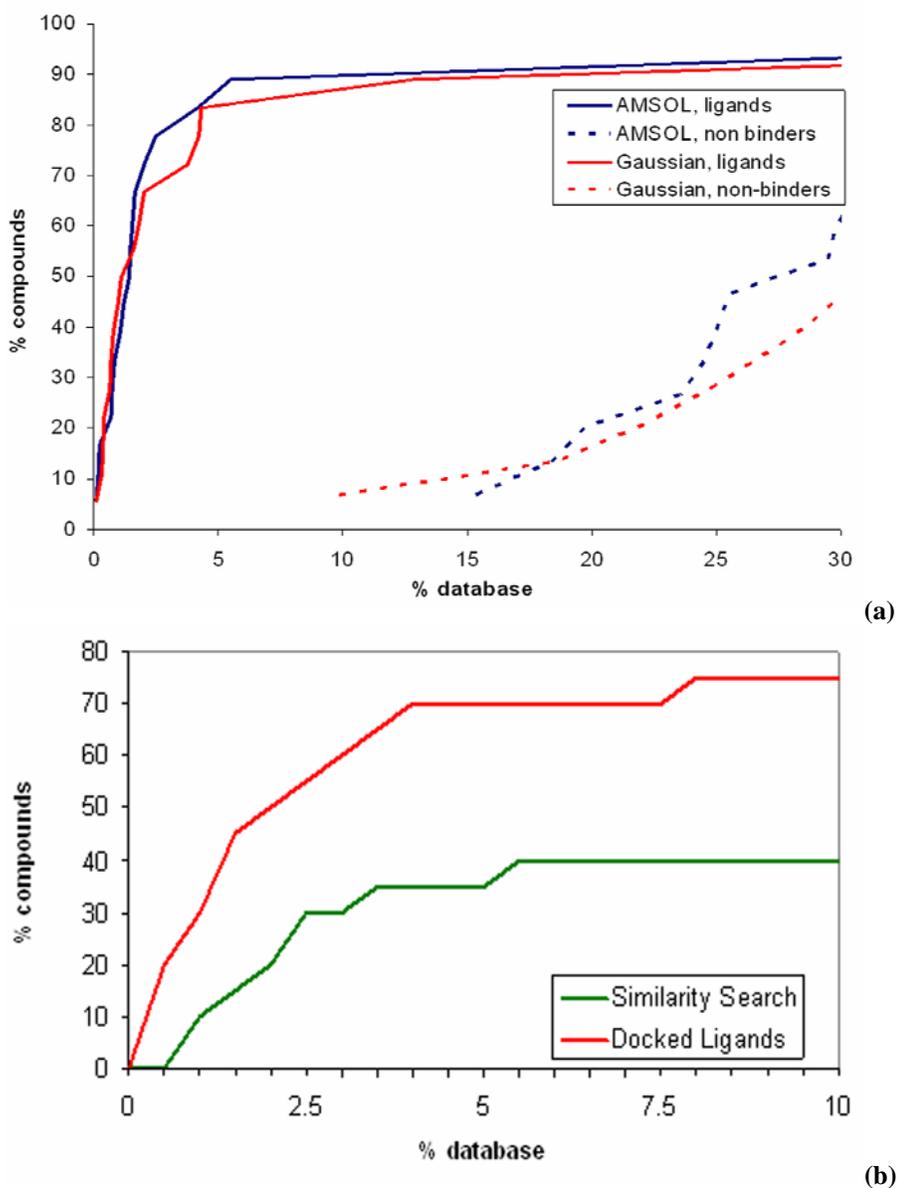
#	Structure	PDB Code	RMSD [Å]	
			AMSOL ^a	Gaussian ^b
1		1AEB	1.68 (12)	0.44
2		1AED	3.01 (31)	2.95 (5)
3		1AEE	0.35	0.45
4		1AEH	3.02 (29)	3.52 (16)
5		1AEJ	0.54	0.63
6		1AEN	1.73 (5)	1.74 (4)
7		1AEO	0.37	0.34
8		1AEQ	0.61	0.35
9		1AES	0.61	0.7
10		1AEU	0.89	0.96
11		1CMP	0.45	0.29
12		1AC4	2.46 ^c	2.46 ^c
13		1AC8	2.4 ^c	2.32 ^c

Table 1. rmsd for the top scoring docking pose compared to the previously determined structure. Where the best scoring pose has an rmsd $>1\text{\AA}$, the best rank for a pose with an rmsd $<1\text{\AA}$ is given in parentheses. ^a Using AMSOL to calculate ligand partial charges and desolvation energies. ^b Using Gaussian to calculate ligand partial charges and desolvation energies. ^c These ligands make a steric clash in the cavity.

the small molecules (our standard procedure²⁵), seven of thirteen ligands had a binding mode close to that found in the crystal structure (RMSD $< 1\text{\AA}$, Table 1). If the correct binding mode, i.e., having an RMSD $< 1\text{\AA}$, among the top 10 poses is considered success, eight correct predictions were made. For the ligands **12** and **13** no correct binding modes can be predicted. These ligands have van der Waals violations even when docked back into their own receptors, probably owing to lack of full refinement of the complex structures³⁰ and should therefore be discounted.

We next turned to enrichment of ligands and downgrading of decoys. The test set was seeded into the 5300 compound database, docked into CCP W191G, and ranked by score. As expected, little correlation was observed when we compared the dock energies to the experimental binding constants (Figure S1, supplementary material). For docking, a less ambitious and more reasonable concern is the enrichment of known ligands among the top ranking docked molecules. Using AMSOL partial charges and desolvation energies, 72% of the ligands ranked in the top 2% of the database, an enrichment of 36, and no known decoys were found in the top 15% of the database (Figure 2). The best scoring neutral molecule, 3,5-difluoroaniline, ranked 147th; the best scoring dicationic compound, pyrimidine-2,4,5,6-tetraamine, ranked 3295th. The structure-based enrichment was much better than what would have been achieved based on simple ligands (Figure 2(b)).

Figure 2. Retrospective enrichment of previously known, “test set” ligands for the W191G cavity in CCP.³⁵⁻³⁷ (a) Using molecular docking, looking at enrichment of known ligands (solid lines) and downgrading of known decoys (dashed-lines), using either AMSOL-based (blue curve) or Gaussian-based ligand partial charges and desolvation energies. (b) Comparing ligand enrichment using chemical similarity the known ligands vs. docking the same database against the cavity structure.



Prospective predictions

A more compelling series of experiments involved prospective testing for new ligands and chemotypes. Twenty-four compounds from the database screen were picked for experimental testing of what we thought might be strengths and weaknesses of our scoring function (Table 2). Compounds **14 - 26** and **28 - 30** were chosen based on their high ranks and chemical diversity, i.e. we chose them based on standard docking criteria. The alkyl amines **33 - 35** were included to assess the limits between desolvation energy penalty and gain in electrostatic energy. These latter compounds ranked poorly in the screen because their desolvation energies are relatively high in magnitude. Consequently, the sums of their electrostatic and desolvation energies, i.e. the net electrostatic contribution to binding, average only -3.6 kcal/mol, whereas the average of those two terms is -14 kcal/mol for the known ligands in the test set. Therefore it seemed likely to us that these were true negative predictions. Similarly, we also wanted to test neutral compounds such as **27, 31, 32, 36,** and **37** which had a good steric fit with the pocket and would give us a chance to probe the previous finding that neutral compounds do not bind to this cavity.²⁹

All of the high-ranking charged compounds tested, except for compound 30, bind to CCP W191G when assayed at 0.5 mM or lower concentration (Table 2). To ensure that the compounds were protonated as modeled in the docking screen, the assay was performed at pH 4.5. Compound 30 gave no evidence of binding at 10 mM in the UV assay, and soaking CCP crystals at 50 mM did not reveal electron density for this compound. Therefore we consider it to be a decoy. For selected ligands (**14, 16, 18,** and **21**), we measured binding constants with full titration curves (Figure 3). These ranged

Table 2. New docking-derived hits, tested for binding to CCP-W191G

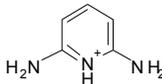
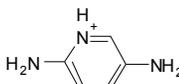
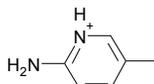
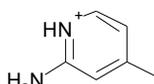
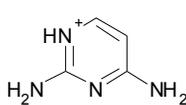
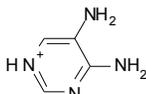
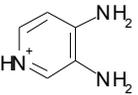
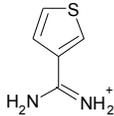
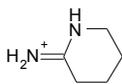
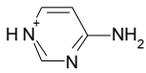
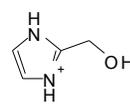
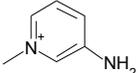
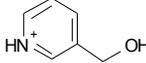
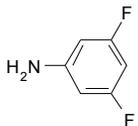
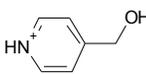
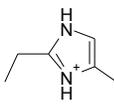
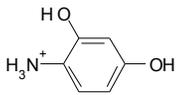
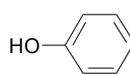
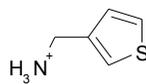
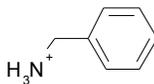
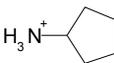
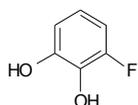
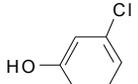
#	Structure	Conc. used in UV assay [mM] ^a	Crystal structure	Binding (K _a [mM]) ^b	Rank		RMSD [Å]	
					AMSOL ^c	AMSOL ^c	AMSOL ^c	Gaussian ^d
14			yes	yes (0.06)	4	0.34	0.36	
15		0.25	yes	yes	10	0.31	0.22	
16			yes	yes (0.04)	11	0.43	0.44	
17		0.25	yes	yes	23	0.39	0.50	
18			yes	yes (0.05)	26	0.30	0.33	
19		0.50	no	yes	31	NA ^e	NA ^e	
20		0.5	no	yes	42	NA ^e	NA ^e	
21			yes	yes (0.02)	55	0.39	0.50	
22		0.13	yes	yes	65	0.49	0.52	
23		0.25	no	yes	73	NA ^e	NA ^e	
24		0.50	yes	yes	95	0.85	0.88	
25		0.25	yes	yes	111	0.46	0.42	
26		0.50	no	yes	140	NA ^e	NA ^e	

Table 2. (continued)

27		50.00 ^h	yes ^g	no	147	NA ^e	NA ^e
28		0.50	yes	yes	187	0.87	0.89
29		1.00	no	yes	198	NA ^e	NA ^e
30		10.00	yes ^g	no	351	NA ^e	NA ^e
31		20.00	no	no	410	NA ^e	NA ^e
32			yes	yes ^f 4.10	420	2.60	2.58
33			yes	yes (0.05)	614	0.56	1.66
34		0.25	yes	yes	998	0.62	0.66
35		0.50	yes	yes	1122	0.64	0.66
36			yes	yes ^f 7.70	1152	2.60	0.62
37		20.00	no	no	1518	NA ^e	NA ^e

^a Concentration is only given, if no binding constant was determined.

^b The error of the binding constants is 30%.

^c Using AMSOL to calculate ligand partial charges and desolvation energies.

^d Using Gaussian to calculate ligand partial charges and desolvation energies.

^e Non applicable, because no complex crystal structure was determined.

^f Binding of these compounds results in a blue shift.

^g No difference electron density for the ligand was obtained.

^h To assure the compound is neutral, the assay was done at pH 6.

from 20 μM to 60 μM . For ten of the new ligands (**14** - **18**, **21**, **22**, **24**, **25**, and **28**), we determined crystal structures in complex with CCP W191G by x-ray crystallography. The resolution of these structures ranged from 1.12 to 1.70 \AA (Table 3). All were extensively refined leading to R_{cryst} and R_{free} values that ranged from 14.4 to 19.3 and from 15.2 to 22.6, respectively. The $|F_o| - |F_c|$ omit electron density allowed us to position the ligands unambiguously (Figure 4). Typically, the docking predicted binding mode agreed well with the crystallographically determined one (< 1 \AA rmsd, Table 2).

Complex structures for isosteric ligands

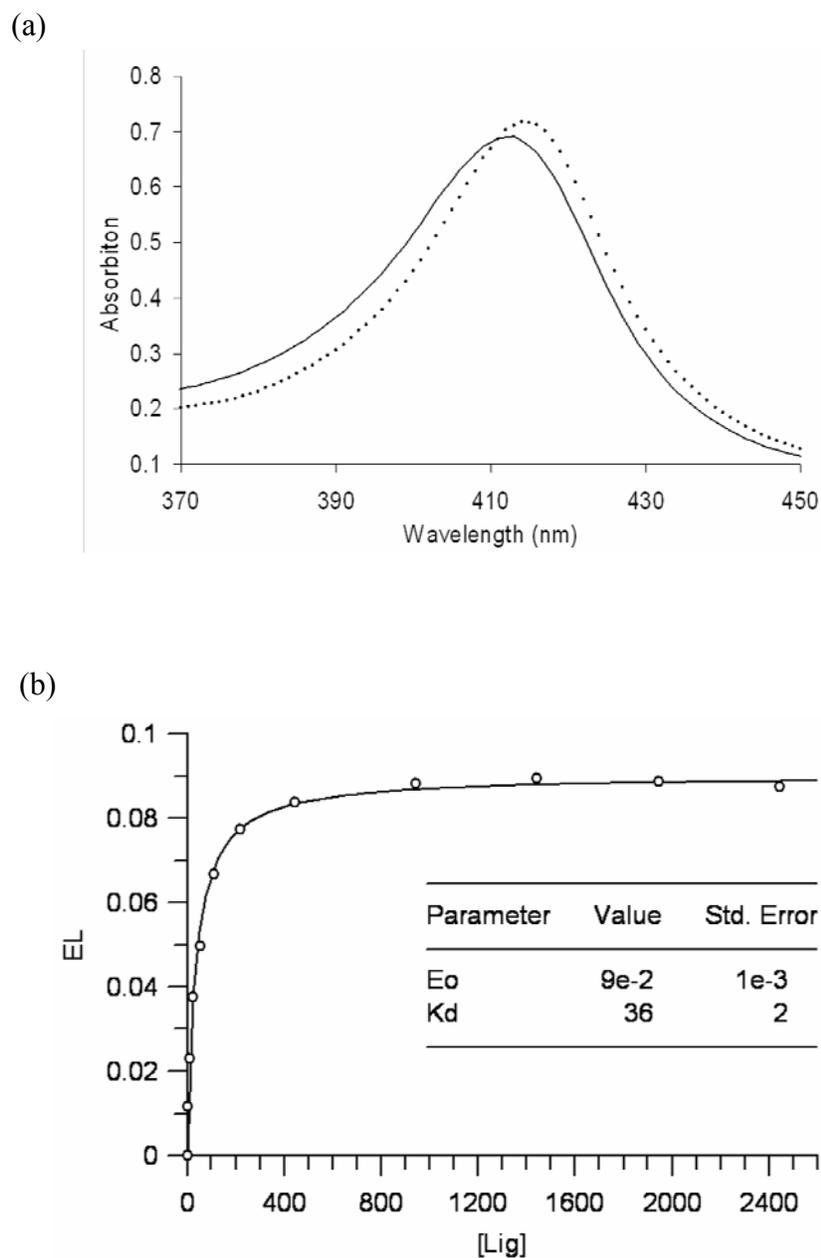
Ligands **14**, **17**, **18**, and **25**, as well as **15** and **16**, have the same shape, but differ in charge distribution and spatial arrangement of hydrogen bond donors and hydrophobic groups. 2,4-diaminopyrimidine (**18**) forms a double hydrogen bond to Asp235 (Figure 4i, j). In 2,6-diaminopyridine (**14**), a carbon atom replaces the ring nitrogen of **18**, which in the complex structure interacts with Asp235. Interestingly, 2,6-diaminopyridine does not adopt a binding mode which would allow for a double hydrogen bond via its remaining ring nitrogen and an exocyclic amino group. Instead its binding mode resembles that of 2,4-diaminopyrimidine, allowing for only one hydrogen bond with Asp235 (Figure 4a, b). Despite the loss of this hydrogen bond, the binding constant of 2,6-diaminopyridine is similar to 2,4-diaminopyrimidine (0.05 vs. 0.06 mM). In 2-amino-4-picoline (**17**, Figure 4g, h), the amino group of 2,6-diaminopyrimidine (**18**), which interacts with Leu177 and Wat308 (Figure 4j), is replaced by a methyl group. Superposition of both complexes reveals that this methyl group is further away from Leu177, resulting in the displacement of Lys179 and Thr180. 2,5-diaminopyridine (**15**) and 2-amino-5-picoline (**16**) also differ only by the replacement of an amino group with a methyl group. Whereas in CCP

W191G·**15** the ligand forms a hydrogen bond with Leu177 (Figure 4c, d), in CCP W191G·**16** (Figure 4e, f) the ligand is shifted away from Leu177. In this complex, in contrast to CCP W191G·**17**, Lys179 and Thr180 are not displaced. The binding constant of **16** is 0.02 mM. Due to high optical density, the binding constant of **15** could not be determined with the UV assay. Another isostere is **25**, an N-methylated pyridinium in which the ring nitrogen is no longer available for direct hydrogen bonding. The position of **25** is defined unambiguously in the $|F_o|-|F_c|$ electron density map with electron density for the pyridinium nitrogen still visible when contoured as high as 9σ (Figure 4q, r). The ligand does not interact with Asp235 via a hydrogen bond to Asp235 through its amino group, but via an ion-dipole interaction that some might classify as a CH-hydrogen bond (distance CH...O 3.2 Å, angle C-H-O 152 °).³⁸⁻⁴⁰

Complex structures for amidiniums

All previously discovered cyclic ligands are aromatic with their positive charge delocalized over the aromatic ring system (Table 1). The amidiniums **21** and **22** seemed interesting because they explore a new cationic functionality, and in the case of **22** the ligand is not even aromatic. Piperidinylideneamine (**22**) adopts a similar binding mode as 2-aminopyridine (**7**), forming two hydrogen bonds with Asp235 (Figure 4m, n). Thiopheneamidine (**21**) does not orient both nitrogens of its charged group to Asp235 to form a double hydrogen bond, as do most ligands, but instead forms a double hydrogen bond with Met230 and only a single hydrogen bond with Asp235 (Figure 4k, l). Thiopheneamidine has the lowest (best) K_d value in the series of ligands measured for this paper (0.02 mM) and is among the better ligands discovered for this site to date.³⁰

Figure 3. (a) Binding of cationic ligands induces a red shift in the Soret band (solid line: spectra of the unbound protein, dashed line: spectra if a ligand is bound (here 18)). **(b)** Titration curve for 18. The solid line represents the least square fit of the data according to the equation for single site binding described in Methods.



Complex structures for ligands with rotatable bonds

Most known ligands for CCP W191G are rigid (Table 1). Binding mode predictions for ligands with rotatable bonds are more challenging because of the increased search space. Therefore we selected two flexible ligands, imidazolmethanol (**24**) and pyridinylmethanol (**28**), to test how well their binding mode is predicted (we note that these ligands are only slightly flexible, with one rotatable bond each—the cavity constraints tilt against much more flexible ligands). In the crystal structure the hydroxyl group of imidazolmethanol orients towards Asp235 in agreement with the docking prediction (Figure 4o and p). In contrast, pyridinylmethanol hydrogen bonds with Asp235 with its ring nitrogen and its hydroxyl group interacts with the backbone carbonyl group of Leu177 (Figure 4s and t). Whereas the former interaction was predicted, the latter was not (Figure 4t). This result reflects the procedure used for preparing the database; the conformer found in the crystal structure was not generated. If the required conformer is added manually, the binding mode is predicted correctly. This is thus a failure of database preparation. Whereas database preparation is a critical challenge in virtual screening,⁴¹ this problem is not one of docking and scoring per se, the foci of this work.

The crystal structure of CCP W191G·**24** revealed that an unmodeled water molecule mediates the contact between the ligand and the protein. This water molecule was also found in the previously determined CCP complex with 2-ethylimidazole⁴² and coincides with the position of the potassium ion in the apo-structure (Figure 4p). Despite the fact that this water molecule was not considered in the docking screen, imidazolmethanol ranks in the top 3% of the database.

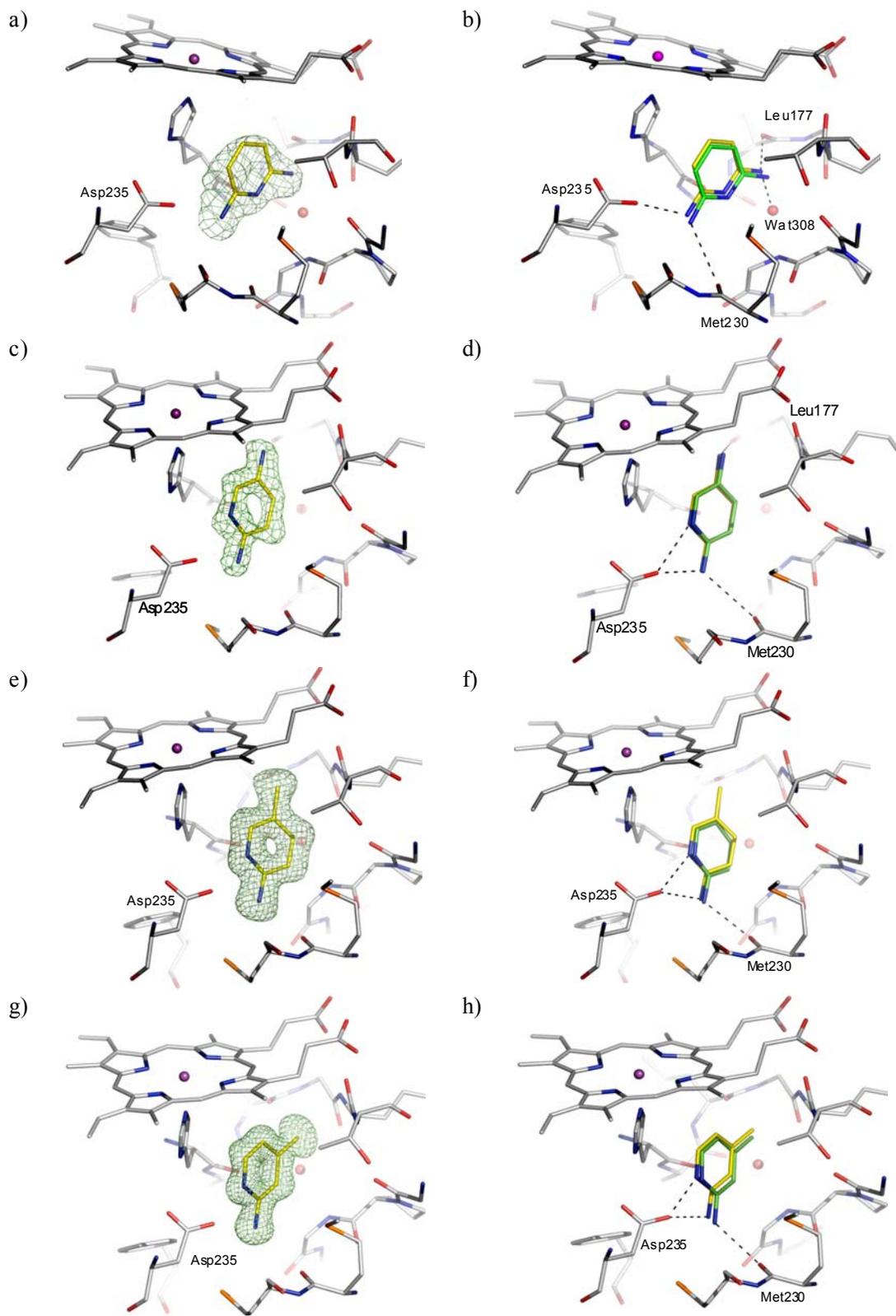
Table 3. Crystallographic data

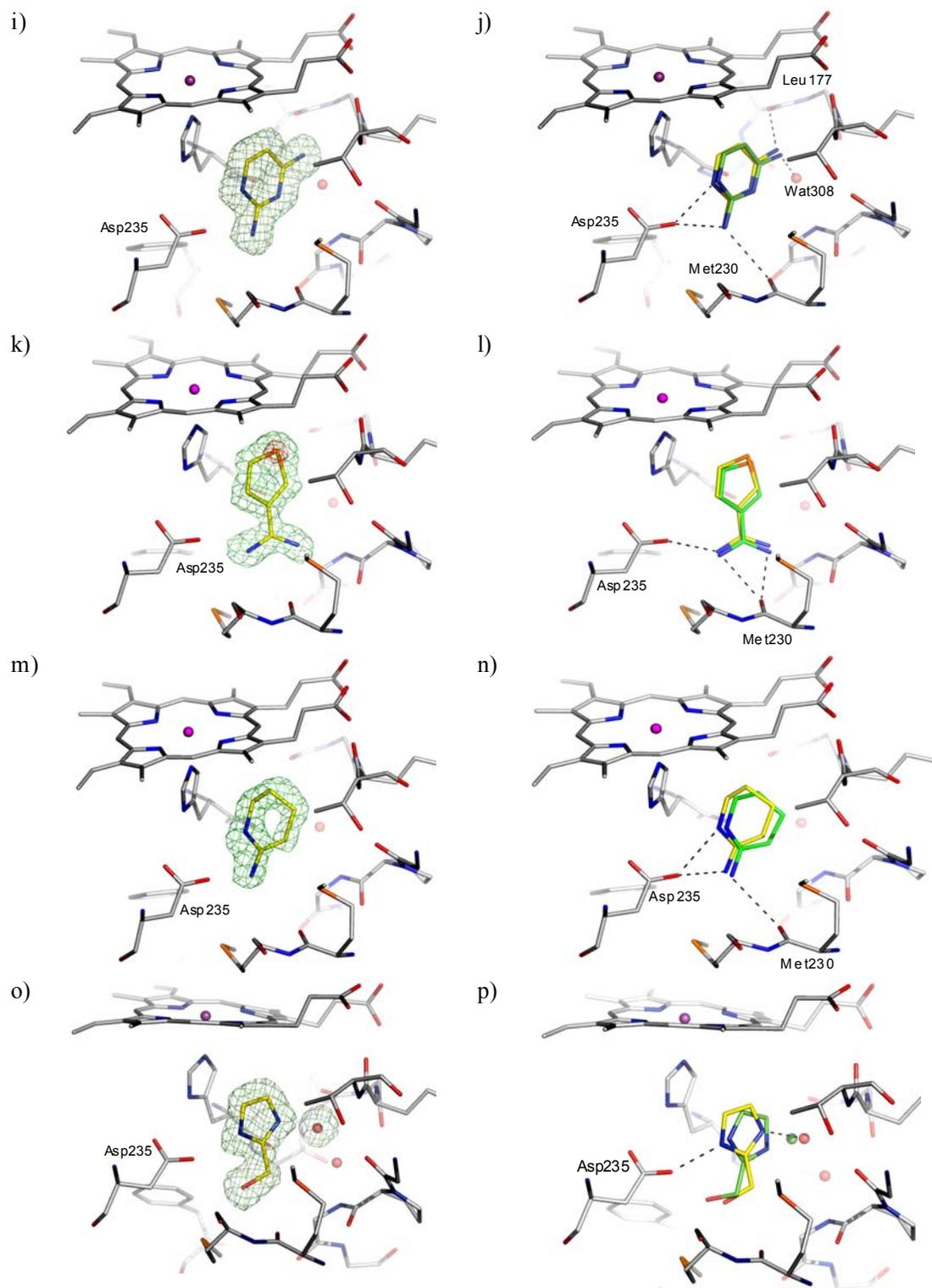
Complex with	14	15	16	17	18	21	22	24	25	28	32	33	34	35	36
pH of soaking buffer	4.5	7.0	6.0	6.0	6.0	7.0	7.0	6.0	7.0	6.0	4.5	6.0	6.0	7.0	4.5
Resolution (Å)	40.0 - 1.75 (1.81 - 1.75)	10.0 - 1.35 (1.40 - 1.35)	50.0 - 1.40 (1.45 - 1.40)	100 - 1.12 (1.16 - 1.12)	36.8 - 1.49 (1.54 - 1.49)	10.0 - 1.55 (1.61 - 1.55)	10.0 - 1.45 (1.50 - 1.45)	50.0 - 1.45 (1.50 - 1.45)	50.0 - 1.45 (1.50 - 1.45)	50.0 - 1.39 (1.49 - 1.39)	10.0 - 1.40 (1.45 - 1.40)	50.0 - 1.30 (1.35 - 1.30)	50.0 - 1.55 (1.61 - 1.55)	10.0 - 1.45 (1.50 - 1.45)	10.0 - 1.30 (1.35 - 1.30)
No. of unique reflections	41,392 (4,210)	90,551 (8,687)	62,656 (4,696)	124,366 (10,410)	53,147 (3,215)	61,090 (5,938)	72,346 (6,447)	59,787 (5,272)	57,355 (4,697)	60,114 (9,364)	81,465 (7,826)	80,407 (6,207)	47,799 (3,984)	72,355 (6,117)	99,539 (7,998)
R_{merge} (%)	6.4 (39.5) ^a	3.3 (23.5)	3.8 (33.5)	7.2 (36.7)	6.7 (31.7)	4.1 (32.7)	3.8 (28.1)	4.2 (37.1)	3.7 (28.3)	6.4 (32.0)	3.3 (36.1)	3.5 (25.8)	6.5 (37.1)	2.8 (23.8)	3.9 (31.5)
Completeness (%)	96.9 (99.9)	99.5 (96.7)	94.2 (71.1)	95.9 (81.6)	94.2 (57.8)	99.3 (98.0)	98.6 (89.1)	98.5 (87.9)	96.0 (79.7)	86.4 (68.4)	99.4 (96.7)	95.9 (74.9)	94.6 (80.1)	97.6 (83.8)	97 (78.9)
I / σ	18.3 (3.5)	28.4 (4.3)	24.3 (2.4)	36.4 (2.2)	8.1 (1.5)	23.2 (3.2)	22.4 (3.3)	21.4 (2.2)	17.3 (2.2)	15.8 (1.9)	27.2 (2.6)	26.2 (2.8)	16.9 (2.5)	30.1 (3.3)	31.3 (3.1)
R_{free} (%) ^b	19.9	16.5	17.1	15.2	22.6	20.0	17.7	16.1	18.1	17.1	18.4	16.6	17.8	18.3	13.8
R-factor (%)	18.1	13.4	14.8	14.3	19.3	15.2	13.7	14.5	15.2	14.7	14.6	14.6	14.8	14.2	17.5
average B-factor of protein atoms (Å ²)	16.4	16.3	17.0	12.4	20.3	15.3	17.0	15.7	14.0	14.8	18.8	11.8	11.5	16.4	19.3
average B-factor of ligand atoms (Å ²)	14.5	15.2	13.6	10.6	16.3	14.3	14.6	14.3	11.3	14.6	24.7	9.2	9.2	17.2	28.9

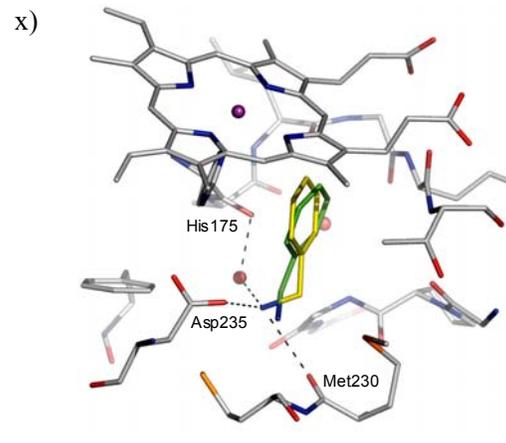
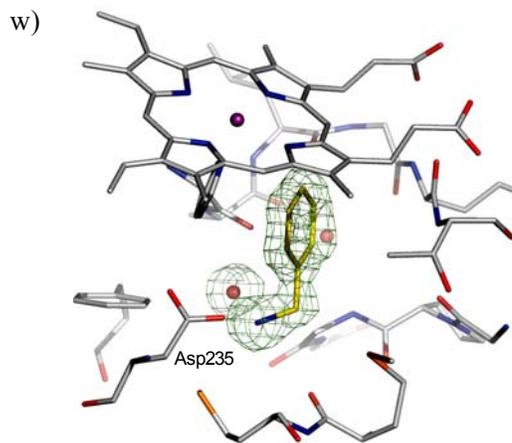
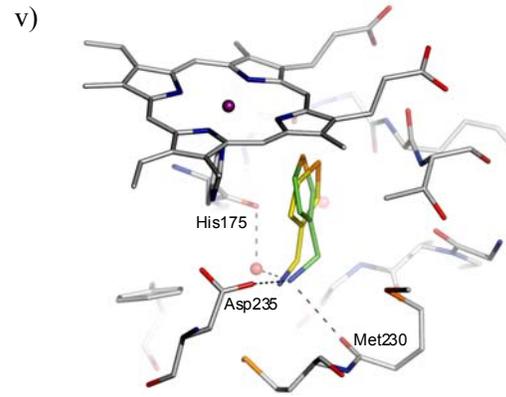
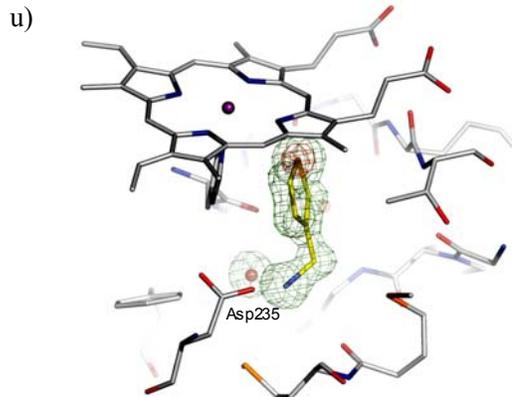
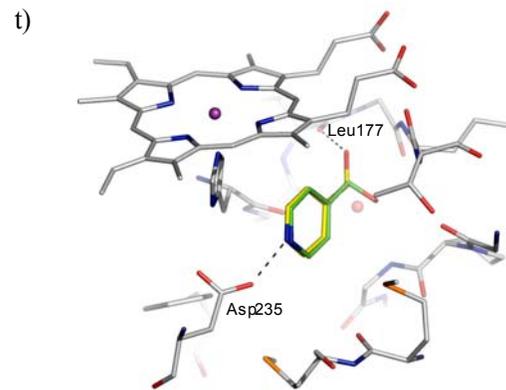
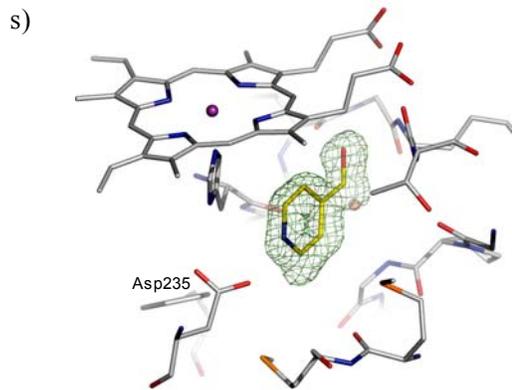
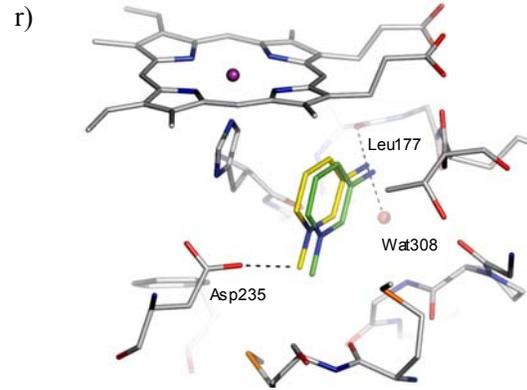
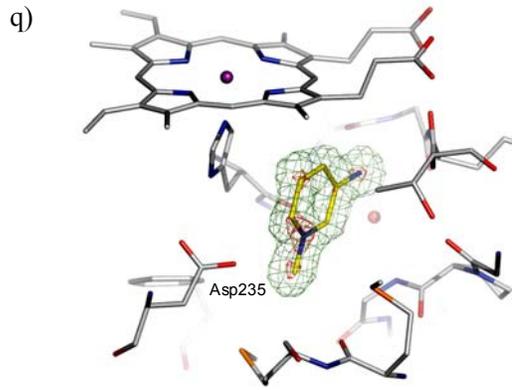
^a Values in parentheses are for the highest resolution shell

^b R_{free} was calculated from a random selection of reflections constituting 5% of the data. The R factor was calculated with the remaining intensities.

Figure 4. (legend page 45)







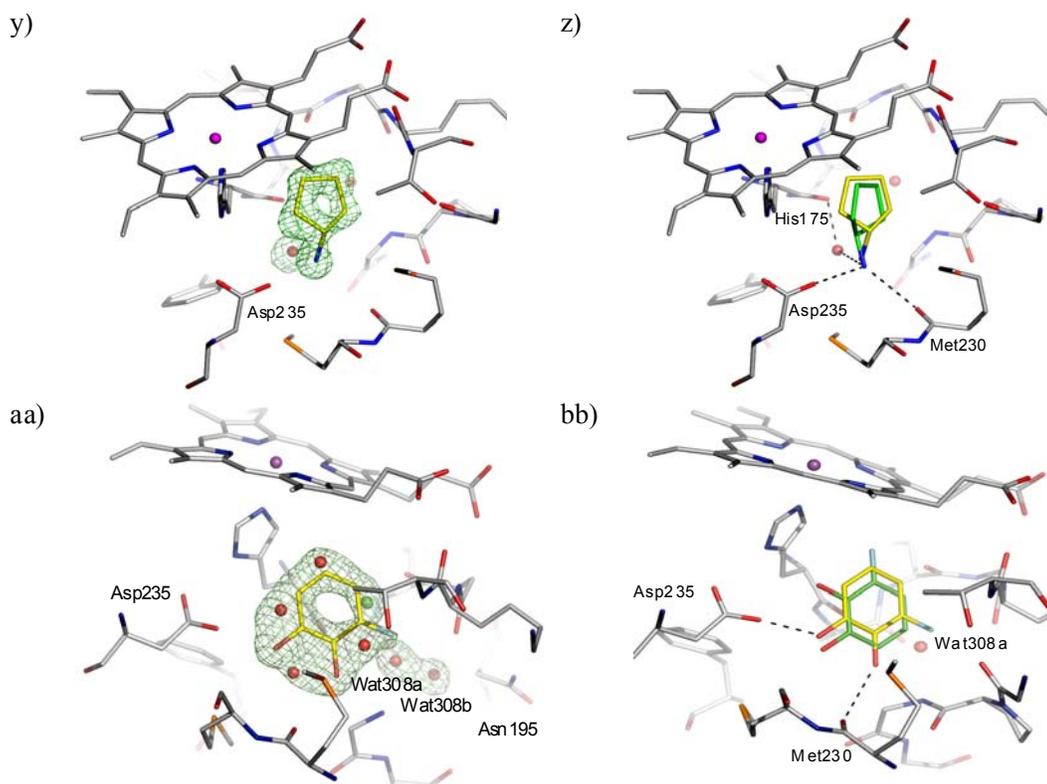


Figure 4. Crystal structures of selected ligands from Table 2 bound to CCP W191G. Left column: $|F_o|-|F_c|$ omit map for the refined complexes, except for a, c, k, m, and y where the map of the unrefined complex is shown, contoured at 2.5σ (green) with the ligand left out of the calculation, but shown in the figure of clarity. Right column: Superposition of the highest ranking dock pose (green carbon atoms) with the crystallographically determined binding mode (yellow carbon atoms). Hydrogen bonds are drawn as dashed lines. a,b) **14**; c,d) **15**; e,f) **16**; g,h) **17**; i,j) **18**; k,l) **21**, the $|F_o|-|F_c|$ map, contoured at 10σ (red) is also shown; m,n) **22**; o,p) **24**; q,r) **25**, the $|F_o|-|F_c|$ map, contoured at 9σ (red) is also shown; s,t) **28**; u,v) **33**, the $|F_o|-|F_c|$ map, contoured at 14σ (red) is also shown; w,x) **34**; y,z) **35**; aa,bb) **36**.

Complex structures with false negative alkyl amines

Surprisingly, the alkyl amines **33** – **35** also bind to this cavity. These alkyl amines rank not even in the top 10 % of the database. Their low score is due to their localized charge which leads to a less favorable desolvation energy for these compounds compared to the desolvation energies of ligands with a delocalized charge (**1** - **26** and **28** - **30**). A

good example of this is thiophenylmethylamine (**33**), whose localized charge makes it harder to desolvate than thiopheneamidinium (**21**), a close analog with a delocalized charge. Nevertheless, the K_d of thiophenylmethylamine (**33**) is 0.05 mM, only slightly worse than that of thiopheneamidinium (**21**), which is 0.02 mM. Accordingly, the alkyl amines are clear false negatives. An explanation is provided by the complex structures; an unexpected water molecule mediates an additional contact between the alkyl amino group of the ligands and His175 (Figure 4u, w, y). This water molecule was previously only observed in CCP W191·**3**. Since predicting the binding modes of most of the ligands in the test set (Table 1) was not possible, when this water molecule was present in the receptor, we did not consider it for the database screen. In the docking screen, the correct orientation of the alkyl amino group with respect to Asp235 is not predicted correctly (Figure 4v, x, z). When the alkyl amines are docked with the water molecule added to the receptor, the right orientation of the amino group is found for **33** and **34** (not shown). Also, the scores of these ligands improve by about 7 kcal/mol, which would result in rank 140 for thiophenylmethylamine (**33**), 234 for benzylamine (**34**) and 306 for cyclopentylamine (**35**).

Complex structures with neutral ligands

Most of the neutral molecules did not bind, consistent with previous expectations.²⁹ Surprisingly, two did, though not in the predicted geometry. As expected, the apolar and neutral molecule toluene (**31**) did not bind to CCP W191G when tested in the UV assay, nor did 3,5-difluoroaniline (**27**) and 3-chlorophenol (**37**). As a further test, we soaked CCP crystals in 50 mM 3,5-difluoroaniline in 25% MPD; no difference in electron density for the compounds was observed. Soaking of phenol (**32**) at neutral pH

was unsuccessful, but at pH 4.5 difference electron density suggested ligand binding and the presence of a partially occupied new water molecule (Wat308b, Figure 5a.) Also observable in this structure at partial occupancy are the water molecules and the potassium ion that fill the apo cavity. The occupancy of phenol and Wat308b was refined to 65%, and the occupancies of the apo-water molecules and the potassium ion correspondingly to 35%. As a consequence of the displacement of Wat308 by phenol, part of a loop (Gly191 to Asn195) is also displaced (Figure 5b). Surprisingly, phenol does not hydrogen bond with Asp235 but rather with the carbonyl group of with Leu177. The unsuccessful soaking at neutral pH, and the absence of a hydrogen bond between phenol and Asp235 suggests that Asp235 is protonated in the pH 4.5 complex. The binding constant of phenol is 4.1 mM at pH 4.5 and 3.3 mM at pH 6.0 Based on the crystal structures it is unclear why the binding constants of phenol at pH 6.0 and 4.5 are so similar.

A second new neutral ligand that binds to the cavity in CCP W191G is 3-fluorocatechol (**36**). Like phenol, the binding constant is in the low millimolar range (7.7 mM). Soaking of this ligand was successful at neutral pH (electron density not shown) and pH 4.5 (Figure 4aa). The ligand is present at a partial occupancy of 77%. Also observed in this structure are the water molecules and the potassium ion associated with the apo cavity. As in the phenol complex, Wat308b is present, but at lower occupancy than the ligand. In contrast to the phenol complex, the conformation of part of a loop from Gly191 to Asn195 is unchanged relative to the apo-structure. The distance between Wat308b and C β of Asn195 is only 2.0 Å, and between Wat308a and Wat308b 1.9 Å. This suggests that Wat308b is present alternatively to Wat308a and the side chain

conformation of Asn195, as defined by the electron density. Refining the occupancies of this residue and the water molecules resulted in 77% for Asn195 and Wat308a and 23% for Wat308b.

Because 3-fluorocatechol is symmetric, if the atom types are not considered, there is some difficulty assigning the interactions in the complex unambiguously. At the resolution of the complex (1.3 Å), it is impossible to distinguish oxygen from fluorine atoms based on the electron density. In one binding mode that is consistent with the difference electron density, the ligand hydrogen bonds with Asp235 and Met235 (Figure 4bb). Due to the geometry of the hydrogen bond, Asp235 must be deprotonated (distance $O_{3\text{-fluorocatechol}} \cdots O_{\text{Asp}}$ 2.4 Å, angle O-H- O_{Asp} 142 °). This configuration seems the more likely to us, but we cannot rule out the possibility that there is an alternative binding mode in which the positions of the oxygen atom interacting with Asp235 and the fluorine atom are switched. In this binding mode, one oxygen of the ligand would be in close distance with Wat308 (2.6 Å) without being able to hydrogen bond with it for geometric reasons. It might therefore be the case that if the ligand adopts this binding mode, Wat308a is displaced and Wat308b is present. Based on the occupancies, the latter binding mode would be adopted in 23% of the unit cells, the former in 54% of the unit cells, and in the remaining unit cells the apo-water molecules and the potassium ion would be present. Neither of the possible binding modes was predicted by DOCK using AMSOL partial charges and desolvation energies (Figure 4bb).

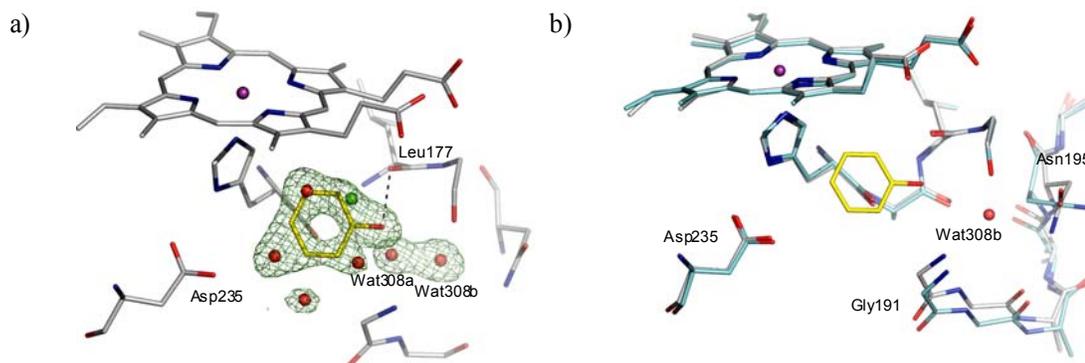


Figure 5. (a) $|F_o|-|F_c|$ omit map of the refined phenol-CCP W191G complex contoured at 3.0σ , calculated with the ligand and the potassium ion and the cavity water molecules left out. The occupancy of the ligand was refined to of 62%, that of Wat308b to 64%, that of Wat308a to 36%, and the occupancies of the remaining water molecules to 38%. (b) Superposition of the apo-structure (carbon atoms colored in cyan) with the phenol complex (carbon atoms colored in gray); water molecules which are not present when the ligand is bound are removed for clarity. In the complex the region from Gly191 to Asn195 is displaced relative to the apo-structure.

Quantum mechanically calculated partial charges and desolvation energies

Both docked geometries and molecule rankings depend upon ligand partial atomic charges and desolvation energies. These were calculated by the semi-empirical quantum mechanical method AMSOL.^{43; 44} This method had served us well in previous studies,²⁵ but it seemed possible that in this charged cavity a higher level of theory would be appropriate. We recalculated the partial charges and desolvation energies for the entire database at the HF level using the 6-31G(d) basis set for neutral molecules and the 6-31+G(d) basis set for charged compounds, with the conductor-like polarizable continuum

model (CPCM) as implemented in Gaussian 03.⁴⁵ These combinations were chosen based on a recent benchmark study.³⁶

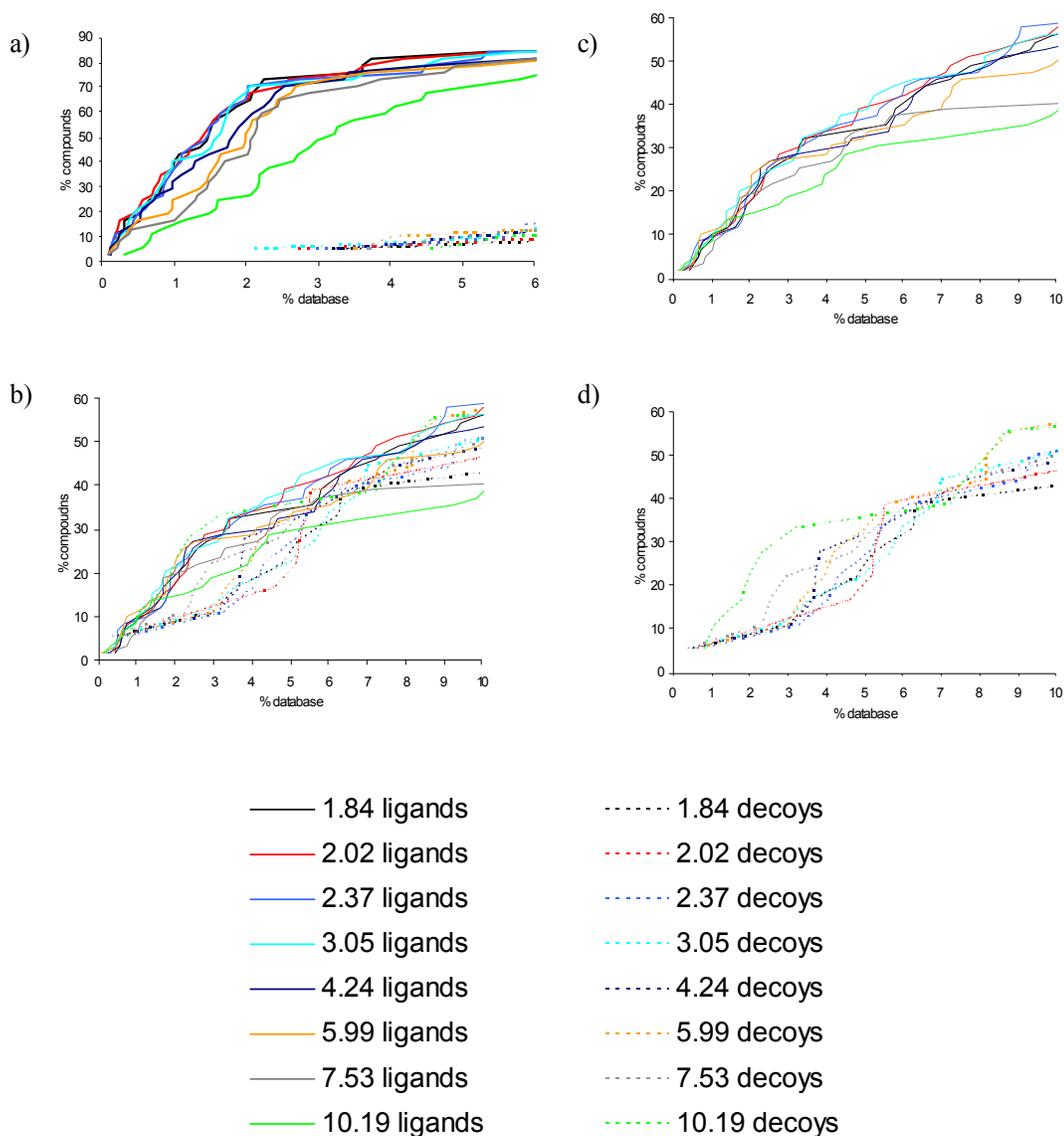
There were no significant differences in binding mode predictions between ligands charged using AMSOL or those charged using Gaussian (Table 1, Table 2). With the Gaussian partial charges, the binding mode of 3-fluorocatechol (**36**) and the position of the sulfur atom of compound **1** (Table 1) is correctly predicted, unlike the predictions using AMSOL partial charges. The binding mode of **33** (Table 2) is only predicted correctly with the AMSOL partial charges. The overall enrichment of the compounds is also about the same (Figure 2) with differences only in the ranking of individual compounds. Interestingly, the neutral compounds of the prospective test (**27**, **31**, **32**, **36**, **37** in Table 2) all rank better with the Gaussian partial charges and desolvation energies, irrespective of whether they bind or not.

Probing the dielectric constant

There is no consensus on which value of the dielectric constant should be used for rigid protein binding sites; estimates vary from 1 to 2046-48 and this range leads to large differences in predicted binding energies. In all calculations described above, we assumed a dielectric constant of 78 for the aqueous buffer and a dielectric constant of 2 for the protein. To test if a different dielectric constant would give us better results, we recalculated desolvation energies and partial charges of the small molecules in the database using dielectric constants ranging from 1.84 to 10.19 (these values were chosen based on defined solvent parameters for AMSOL). We then redocked the database against the cavities in CCP W191G, T4 lysozyme L99A and L99A/M102Q using the same dielectric constant for calculating the electrostatic potential of the receptor as used

for calculating the properties of the small molecules. In all three systems, no significant change in the enrichment is obtained if the dielectric constant in the binding pocket is varied from 1.84 to 3.04, when judged by the number of ligands found in the top 2% of the database for the CCP W191G pocket and top 10% for the T4 lysozyme systems (Figure 6). If the dielectric constant is increased further, enrichment drops in all three systems. A worse enrichment can reflect two effects: either more decoys get enriched or unknown ligands show up in the top ranks. Based on previous results, only hydrophobic compounds can bind to the T4 lysozyme L99A cavity. If a dielectric of 10.19 is assumed for the binding pocket, 55 of the top 100 molecules contain nitrogen or oxygen atoms compared to 25 of the top 100 molecules for a dielectric constant of 2.02. This indicates that if the dielectric constant is increased, polar decoys are enriched. The same is true for the slightly polar cavity in T4 lysozyme L99A/M102Q. Only 59 out the 100 top scoring molecules contain one or less nitrogen or oxygen atoms when a dielectric constant of 10.19 is assumed, compared to 85 for a dielectric constant of 2.02. For CCP W191G, all 100 top scoring molecules have a total charge of +1 when a dielectric constant of 2.02 is used. After increasing the dielectric constant to 10.19, one molecule in the top 100 has a charge of +2, and 12 have a total charge of 0. Most of these have no polar atoms, which makes it unlikely that they bind in this cavity. Taking these results together, increasing the dielectric constant to 10.19 led to enrichment of more decoys and consequently worse results in all three simple cavities.

Figure 6. The variation of ligand enrichment (continuous lines) and decoy downgrading (broken lines) with protein dielectric constant when docking into: **(a)** the charged cavity of CCP W191G; **(b)** the hydrophobic cavity of T4 lysozyme L99A; **(c)** and **(d)** the slightly polar cavity of T4 lysozyme L99A/M102Q (for clarity, ligands and decoys are separated). For calculating the score, the dielectric constant in the pocket was varied from 1.84 to 10.19. The ligands and decoys are the corresponding “test set” compounds (see Methods) except for CCP W191G, where the test set was augmented with the newly-discovered docking hits (Table 2).



Partial charges

In all calculations described above, the partial charges for the molecules in the database were calculated in the medium of low dielectric. Intuitively, this might be the obvious way to proceed, because this is the same dielectric assumed for the cavity. However, the partial charges of the ligands might be polarized upon ligand binding. To simulate this process, we calculated the partial charges of the compounds in the database in water and redocked them in the cavities of the model systems. No change in enrichment was obtained in any system (Figure S2, supplementary material).

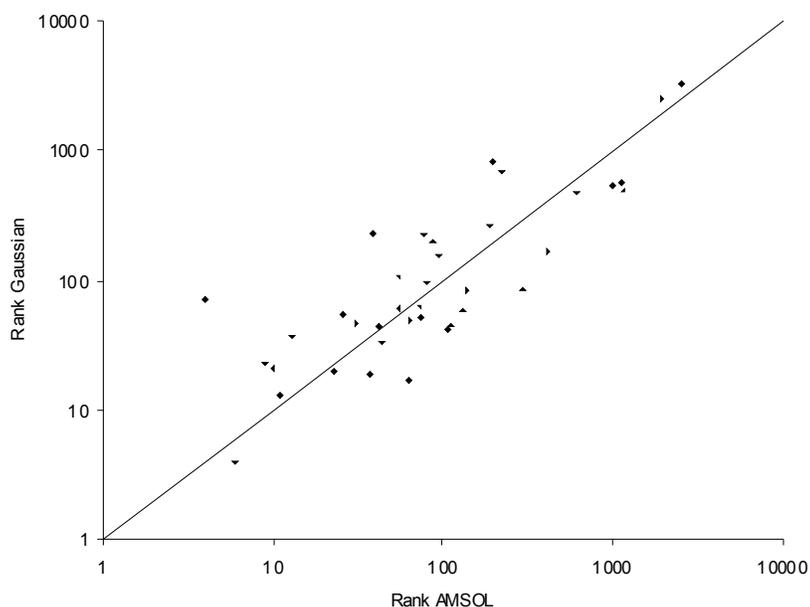


Figure 7. Ranks of the CCP W191G cavity ligands (test set ligands and the new ligands in Table 2) scored using Gaussian charges and desolvation energies plotted against the ranks obtained using AMSOL charges and desolvation energies.

1.4 Discussion

Modeling charge-charge interactions in docking is challenging because the gain in electrostatic energy upon ligand binding has to be balanced against desolvation energies. Both values are high in magnitude, as are their errors in computer simulations. This study allowed us to probe charge-charge interactions in a controlled environment, a small pocket completely buried from solvent. If we are able to get the balance right anywhere, it should be in such a relatively simple site. Correspondingly, mispredictions are particularly informative because they come much less entangled by the approximations necessary in more complicated sites. Five points stand out from this study. First, overall electrostatic and desolvation energy appear to be balanced well in the physics-based scoring function. No neutral compound ranked among the top 100 molecules, and the first dicationic molecule scores poorly at rank 3295. Second, from a practical standpoint, virtual screening with this cavity was successful. Fifteen of sixteen chemically diverse compounds, which ranked in the top 5% of the database, did actually bind to the site when tested experimentally. For all 10 high-ranking ligands for which the crystal structures in complex with CCP W191G were determined, the binding modes were predicted within $< 1 \text{ \AA}$ rmsd. Third, neither using a higher level of theory for calculating partial charges and desolvation energies, nor changing the dielectric constant in the cavity improves these results. Fourth, although the overall performance is good, problems exist for neutral compounds. The only neutral ligand found that interacts with the deprotonated Asp135 (**36**) ranks poorly (1152nd), whereas the best scoring neutral decoy **27** ranks 147th. Fifth, analyzing false negative predictions points to weaknesses in current docking protocols and can guide the improvement of scoring functions and docking algorithms. Examples of such instructive false negatives are the alkyl amines **33** - **35**. Their poor

ranking owes to an inadequate handling of explicit water molecules during docking. Similarly, the binding mode of phenol was not predicted correctly, because pK_a shifts were not considered. We consider these points further below.

The physics-based scoring function used here (eq. 1) was surprisingly effective at enriching new ligands and predicting their binding geometries. We had expected the scoring function to have trouble balancing the interaction energy and desolvation terms, finding either more high-scoring neutral or dicationic hits than was warranted. Instead, the top scoring hits were dominated by singly charged cationic heterocycles, with the first neutral ligand ranked 147th (top 2.8% of the database) and the first dicationic molecule ranked 3295th (top 62.2% of the database). Of the 17 high-scoring molecules tested experimentally for binding, only two, 3,5-difluoroaniline (**27**) and aminoresorcin (**30**) were not observed to bind (Table 2). It is debatable if aminoresorcin is really a false positive or rather a true negative prediction, since it does not even rank in the top 6% of the database. For four of the new high-ranking ligands binding constants were determined. They range from 20 to 60 μ M putting them among the better ligands known for this cavity³⁰ with a “ligand efficiency” for the best ligand close to the projected maximum.⁴⁹⁻⁵¹

The geometric fidelity of the docking predictions was also high (Table 1 and 2). At a first glance, predicting the correct pose might seem trivial, since most of the ligands are rigid and the pocket is small, but even in this simple system it can be a challenge. For instance, 2,6-diaminopyridine (**14**) does not form a double hydrogen bond with Asp235 via its ring and exocyclic nitrogens as one might expect, and as it is actually observed for 2-aminopyridine³⁰, 2,5-diaminopyridine (**15**, Figure 4c, d), 2-amino-5-picoline (**16**,

Figure 4e, f), 2-amino-4-picoline (**17**, Figure 4g, h) and 2,4-diaminopyrimidine (**18**, Figure 4i, j). Instead 2,6-diaminopyridine only forms one hydrogen bond via the exocyclic amine group, and the protonated ring nitrogen does not have a hydrogen bonding partner at all (Figure 4a and b). Although this is not the binding mode we might intuitively predict for this ligand, it *is* correctly predicted in the docked geometry (Figure 4b). Also the binding mode of thiopheneamidinium (**21**) is predicted correctly, despite the absence of steric constraints to guide the position of the sulfur atom of the thiophene ring (Figure 4k, l). In summary, the quality of the docked geometries was typically high for the novel ligands, even in cases where distinguishing between the correct and incorrect pose involved a subtle balance of forces; even in a simple site, such balanced forces are often in play.

Along with the high hit rates came new and interesting chemotypes as ligands. Considering their small size, the enriched ligands are diverse and include disubstituted pyridines (**14 - 17, 20, 25**), pyrimidines (**18, 19, 23**), amidines (**21, 22**), alcohols (**24, 26, 28**), and non-aromatic ligands (**22**) (Table 2), none of which had previously been discovered. That said, all of these molecules are small and cationic; could they have been found by simpler methods, such as simply chemical similarity? Using Daylight fingerprints, only three of the sixteen new ligands have a Tanimoto coefficient of 0.85 or better to the previously known ligands (described in refs. 29 to 31). Another way to pose this question is to ask how many of the supposedly novel docking hits would have been found by screening the database by similarity to the previously known ligands? Again using topological similarity as a metric, the enrichment of the docking-derived ligands from the similarity search was considerably lower than the structure-based docking

enrichment (Figure 2b); most of the new chemotypes would not have been discovered solely by using a similarity search. Thus the docking-derived ligands seem genuinely novel, which is important for a model binding site as diverse ligands will avoid bias towards a particular type of chemotype when testing and improving scoring functions.

Taken together, the high enrichment of monocationic ligands and the high fidelity of the binding mode predictions suggest that the relatively simple, physics-based scoring function represented by eq. 1 can at least separate likely from unlikely ligands, getting the overall balance between electrostatic interactions and desolvation correct. On closer inspection, however, problems with the predictions do emerge. Not all ligand interactions were correctly predicted (Figure 4t), one high-ranking docking hit did not bind (**27**), and two neutral ligands (**32** and **36**) were ranked poorly. What do these problems tell us about weaknesses in our scoring functions and how might they be overcome?

We had previously found, in the neutral lysozyme cavities, that docking could be improved by moving to a higher level of theory in modeling ligand desolvation and partial atomic charges.²⁵ Here, we investigated moving one step further, from a semi-empirical quantum mechanical method to a fully quantum mechanical method to calculate ligand partial charges and desolvation energies. Overall, moving to higher theory had little effect, with changes only in the relative ranking of the ligands and decoys (Figure 2 and 6, Table 1 and 2). The desolvation energies calculated by both methods can differ by several kcal/mol. The consequence for docking is that different ranks are predicted for specific ligands, without changing overall performance. Indeed, we may be reaching a limit on how well we can hope to do with even fairly sophisticated methods for calculating ligand desolvation. The error in the calculated energy for the

transfer for a cation from water to vacuum with these methods is 3 - 4 kcal/mol.^{36; 37} With our scoring function, a change of 3 kcal/mol can make a difference of about 200 rank units. To have a significant impact on molecular docking for virtual screening, a new method to calculate charges and desolvation energies must have a smaller error than this 3 – 4 kcal/mol uncertainty level that most of the current methods have for simple solvent transfer free energies.

There is no consensus as to what is the best dielectric constant to model electrostatics in a protein binding pocket.⁴⁶⁻⁴⁸ Based on strictly electronic effects, we used a dielectric constant of 2.²⁵ This may be an extreme choice, given that we are docking to a rigid receptor. Also, changing the dielectric constant is a way to influence the weighting between van der Waals term, ligand desolvation energy and electrostatic energy, and so, from a pragmatic standpoint, it seemed interesting to explore. We therefore repeated the docking screens using different dielectric constants for the protein binding site leaving the external dielectric fixed at 78 (Figure 6). In addition to the negatively charged CCP W191G, we also docked the database against the hydrophobic cavity in T4 lysozyme L99A and the slightly more polar cavity in T4 lysozyme L99A/M102Q. In all three systems, the best enrichment is obtained for values between 1.84 and 3.04. We also compared the performance obtained when the partial charges are either calculated in water or cyclohexane (Figure S2, supplementary material). In all three systems, the enrichment is not influenced by these small changes. The similar behavior of these systems, in which the properties of the binding pocket range from completely hydrophobic to polar to charged, indicates that the physics-based scoring function used here is not grossly biased towards a particular type of interaction.

An attractive feature of model binding sites, such as CCP W191G, is that false positive and false negative predictions are often more informative than true predictions. We were thus almost disappointed by the high initial hit rate of the prospective docking screen. As we dug further, however, interesting problems did emerge. The neutral compound 3,5-difluoroaniline (**27**, Table 2) ranks well, but does not bind to the cavity, whereas another neutral compound, 3-fluorocatechol (**36**) scores badly, but does bind. Also the alkyl amines **33** - **35** rank poorly, but bind to the cavity. These ligands form one hydrogen bond to a water molecule which was not considered during docking (Figure 4v, x and z). If this water molecule is considered during docking, the scores of these alkyl amines improve leading to a difference in more than 500 ranks. Unfortunately, simply adding a water molecule to the target is not a panacea. There are some molecules, like imidazolmethanol (**24**), that can dock with or without a water molecule despite the fact that it hydrogen bonds with it in the crystal structure (Figure 4p). Worse, most of the ligands won't bind with either of those water molecules present. To improve docking, algorithms are needed that treat the water structure flexibly, and that can balance the energetic costs and benefits of either binding or displacing ordered water molecules.⁵²⁻⁵⁵

Perhaps the most interesting mispredicted molecules are phenol (**32**) and 3-fluorocatechol (**36**), which are the first neutral ligands for this cavity (Table 2, Figure 4aa, bb, Figure 5). Neither of these molecules ranks well, in the docking hit list. Admittedly, these neutral compounds are weaker ligands than many of the cationic ligands, though it is also true that **12** (Table 1) binds in the millimolar range (1.5 mM).³⁰ Nevertheless, the poor ranking of the neutral compared to the charged ligands points to weaknesses in the docking protocol. Most likely, binding of phenol is associated with a

pK_a shift of Asp235 (Figure 5). Such pK_a shifts of either the ligand or the protein are not uncommon, but are not considered routinely in current docking protocols. Reliably modeling these changes would lead to better predictions. 3-fluorocatechol interacts with the deprotonated Asp235 as modeled during docking, but still does not rank well. Thus even though the docking screen performed well overall, there is room to improve the balance between electrostatic and desolvation energy; such an imbalance, obvious in this simple system, will become more deleterious, if harder to see, in more complicated “drug-like” sites.

1.5 Conclusions

The cavity site in CCP W191G is the third model system that we have studied for docking, adding a charge-dominated cavity to the hydrophobic and slightly polar sites represented by T4 lysozyme L99A and L99A/M101Q. CCP W191G allows us to explore the critical balance between electrostatic interaction energy and ligand desolvation in a site where many of the common approximations in docking do not apply. Docking was able to predict novel ligands at a surprisingly high hit rate, suggesting at least gross features of the desolvation-electrostatic balance were correct. That said, there were important and interesting failures – some neutral compounds rank low, but bind, others rank high, but don't bind, and the charged alkyl amines rank poorly, but also bind. The reasons for these failures are the same as observed in complex binding pockets: an inadequate handling of water molecules, neglect of pK_a shifts and insufficient treatment of ligand desolvation energies. In this model system we can hope to study these problems in detail without the entanglement of other effects that occur in complex binding sites.

We suspect that this charged model binding site, in conjunction with the cavity sites in T4 lysozyme, will provide illuminating model binding sites not only for docking methods, but also for much more sophisticated theoretical techniques. The simplicity of these cavities, the dominance of particular terms in each of them, the atomic resolution structures available for multiple ligands and the ability to test new predictions prospectively, makes these sites interesting test cases for many molecular simulation methods.

1.6 Methods

Receptor preparation

Polar hydrogens were added to CCP W191G (PDB code 1AC4) using MOLOC and their positions minimized using the MAB force field.⁵⁶ Since water molecule 308 was observed in all complexes determined to date,³⁰ it was kept as a rigid part of the receptor. All other water molecules in the pocket and the potassium ion were removed. AMBER charges⁵⁷ were assigned to the protein atoms and to Wat308. Partial charges for the heme cofactor were calculated in cyclohexane using Jaguar (Schrödinger Inc.) with the 3-21G basis set for the Fe (III) atom and the 6-31+G(d) basis set for all other atoms. Grid-based excluded volume and van der Waals energy maps, the latter based on the AMBER potential function, were calculated for the cavity using the DOCK utilities DISTMAP and CHEMGRID. DelPhi⁵⁸ was used to calculate an electrostatic potential for the receptor, using an internal dielectric of 2 and an external dielectric of 78, unless explicitly described otherwise in the text. To approximate the effect of ligand binding, the effective dielectric of the binding site was reduced by identifying the volume expected to

be occupied by ligand atoms as a low dielectric region.²⁵ Ligand atoms from the crystal structures, augmented with SPHGEN spheres,⁵⁹ were used as receptor matching positions to dock molecules in the site. The cavities in T4 lysozyme L99A and L99A/M102Q were prepared as described.²⁵

Test sets

The test set for the CCP pocket was composed of the ligands and decoys described previously.³⁰ Several ligands do not interact directly with Asp235 but instead form a water mediated contact; since we did not attempt to model explicit water molecules, we excluded these ligands from our test set. Two ligands, indoline and imidazo(1,2-a)pyridine, alter the protein conformation. They were therefore not considered, nor was quinoline, for which no complex structure exists, but which is even larger than these compounds. Including tautomers, there were 18 ligands and 15 decoys in the test set. The test set for the T4 lysozyme cavities was composed of previously published ligands and decoys.^{12; 24; 25} Since no attempt was made to model receptor flexibility, ligands which could not pass the DISTMAP filter for simple steric fit were not included. Altogether, there were 44 ligands and 31 decoys for the L99A cavity and 59 ligands and 18 decoys for L99A/M102Q cavity in the test sets. All of these are available free of charge from our laboratory site (<http://shoichetlab.compbio.ucsf.edu/take-away.php>)

Database preparation

With a python script based on OpenEye's OEChem library, duplicates in the Available Chemicals Directory (ACD) 2003 were removed and the remaining compounds

filtered for molecules with a maximum of 15 heavy atoms and at least one ring. Subsequently, LigPrep (Schrödinger Inc.) was used to convert the molecules from 2D to 3D, enumerate stereoisomers, tautomers and protonation states. In the latter step, a pH of 5 +/- 2 was assumed resulting in all titrable groups with an assigned pK_a lower than 3.0 as deprotonated, above 7.0 as protonated, and both states were represented for the remaining groups. Conformations were sampled using Omega (OpenEye) and stored in a hierarchical flexibase.³⁵ Partial atomic charges, desolvation energies and van der Waals parameters were calculated as described with one exception related to the treatment of the cavity terms in AMSOL.^{25; 60} The desolvation energy in AMSOL is composed of two terms: the change in solute-electronic and solvent-polarization free energy (ΔG_{EP}) and the cavity-dispersion-solvent-structure free energy (G_{CPS}).³⁷ The first term accounts for the electrostatic interactions of the solute molecule and the solvent, the second term accounts for forming a cavity in the solvent into which the solute is transferred. In our previous study on the T4 lysozyme systems, the desolvation penalty of the small molecules was calculated as:²⁵

$$\Delta G_{solv} = \Delta G_{EP}^{water} - \Delta G_{EP}^{cyclohexane} + \Delta G_{CPS}^{water} \quad (2)$$

This was based on the assumption that the cavities in the apo-structure are preformed and free of solvent. Whereas this assumption is sensible, it might be problematic from a practical point of view. AMSOL is a parameterized semi-empirical method. During parameterization no attempt was made to get both terms correct, but only the overall desolvation energy. Thus, the G_{CPS} term was also designed to make up for

systematic deficiencies and intrinsic uncertainties in ΔG_{EP} .²⁵ Based on these considerations, the desolvation energy must be calculated as:

$$\Delta G_{solv} = \Delta G_{EP}^{water} - \Delta G_{EP}^{cyclohexane} + (\Delta G_{CPS}^{water} - \Delta G_{CPS}^{cyclohexane}) \quad (3)$$

We therefore docked the small database (see below) with ligand desolvation energies calculated with both equations in the T4 lysozyme pockets (L99A and L99A/M102Q). Ligands were better enriched and decoys further downgraded in the top 10% of the database with a scoring function based on eq. 3 (Figure S3, supplementary material). With the scoring function based on eq. 2, all of the top scoring ligands contain several fluorine atoms together with polar groups (data not shown). In our experience, these molecules most likely do not bind to these rather hydrophobic pockets.^{24; 25} In contrast, with the scoring function based on eq. 3, fluorinated compounds are no longer enriched and the top scoring molecules closely resemble known ligands. Thus, in this study we calculated desolvation energies as the difference between the total desolvation calculated in water minus the total desolvation calculated in a solvent with lower dielectric constant.

To reduce the size of the database and to ensure that the compounds in the database have similar properties as the ligands and decoys in the test set,⁶¹ all molecules were docked into the CCP W191G pocket, and only those with a negative van der Waals score and a net charge of zero or higher were kept. Ligands of the test sets not present in the ACD were added manually. The final database contained about 5300 compounds, 131 of them were +2 charged, 996 were +1 charged and the remaining molecules are neutral. For these molecules, partial charges were also assigned according to the Merz-Singh-Kollman scheme,⁶² with desolvation energies for the transfer from water to cyclohexane

calculated based on the CPCM method^{63; 64} using GAUSSIAN 03⁴⁵ with the HF 6-31G(d) basis set for neutral molecules and HF 6-31+G(d) for charged molecules. If the dielectric constant was varied in the pocket, a solvent with the same dielectric constant was used for recalculating desolvation energies and partial charges with AMSOL.^{43; 44}

Docking protocol

DOCK3.5.54^{25; 35} was used to dock a multi-conformer database of small molecules into the cavities. To sample ligand orientations, ligand, receptor and overlap bins were set to 0.2 Å; the distance tolerance for matching ligand atoms to receptor matching was set to 0.75 Å. Each docking pose was evaluated for steric fit. Compounds passing this filter were scored for electrostatic and van der Waals complementarity and corrected for desolvation.

Similarity Search

A similarity search was performed with the test set of ligands as the reference structures, using Daylight fingerprints. Each ligand was compared to the full database used in the docking study. A Tanimoto-index of 0.85 was used as the cutoff for when two molecules were considered similar⁶⁵. The enrichment plot for the similarity search was made by using the test set ligands to search the full database with the Tanimoto-index threshold at zero and using the top Tanimoto-coefficient for each compound in the database to rank the database as a whole by similarity to the known ligands. Ranks of the new binders from the similarity search were then compared to the ranks of the new binders from the docking run. Smiles strings for the ligands in the test set and the full database were generated using a python script based on OpenEye's OEchem software

version 1.3.4. Daylight fingerprints were built from the smiles strings using the Fingerprint Toolkit in Daylight version 4.83 distributed by Chemical Information Systems, Inc (CIS Inc). The similarity search was performed utilizing a Tanimoto coefficient calculation derived from code in CACTVS subset 1.0 (CIS Inc).

Protein expression and purification

CCP W191G was expressed and purified as described.^{29; 66}

Ligand-binding measurements

Compound **25** was from Specs, **21**, **24**, and **33** were from Maybridge and all other compounds from Aldrich. Ligand binding was measured in 500 mM acetate buffer at pH 4.5, except **27**, which was assayed at pH 6.0 to ensure that the compound was neutral. To avoid competition in ligand binding with small cations like potassium,²⁹ the pH of the buffer was adjusted with Bis-Tris. The compounds were dissolved in either buffer or DMSO. Binding was monitored by the red shift and increase of absorbance of the heme Soret band, except for the neutral ligands where a blue shift was observed.²⁹ Binding constants were obtained by plotting the difference in absorbance at 418 nM and fitting the data with GraFit (Erithacus Software Limited) to equation $EL = \frac{-(L_o + E_o + K_d) \pm \sqrt{(L_o + E_o + K_d)^2 - 4E_oL_o}}{2}$, where E_o is the total enzyme concentration, L_o is the total ligand concentration, EL is the concentration of the bound complex, which is proportional to the observed change in the Soret band, and K_d is the binding constant.

Structure determination

Crystals were grown as described.³⁰ Compounds **16 - 18, 24, 25, 28, 33, and 34** were soaked overnight by adding 1 μ L of 100 mM stock solution dissolved in water to the mother liquor. Compounds **15, 21, 22, 25, and 27** were soaked for one hour at a concentration of 50 mM in 25% MPD, and compounds **14** and **30** for one hour at a concentration of 50 mM in 125 mM acetate buffer (pH 4.5) containing 25% MPD. Compounds **32** and **36** were soaked in both the MPD buffer and the acetate buffer. Diffraction data for the complex with **14** was collected at University of California San Francisco and for the complex with **18** at the Scripps Research Institute, San Diego, using a Rigaku X-ray generator equipped with a rotating copper anode and a Raxis IV image plate. Data for the complexes with **24** and **27** was collected on Beamline 5.0.1 of the Advanced Light Source (ALS) at Lawrence Berkley National Laboratory using an ADSC-CCD detector and for all remaining complexes on Beamline 8.3.1 of the ALS using an ADSC-CCD detector. All data sets were collected at 100 K. Data for the complex with **18** were reduced and scaled with CrystalClear and d*trek⁶⁷ and for all other complexes with HKL2000.⁶⁸ The complex with **14** was refined using CNS⁶⁹ and the complexes with **15, 21, 22, 32, 35** and **36** were refined using SHELX.⁷⁰ Parameters for these ligands were generated using PRODRG.⁷¹ The remaining complexes were refined using CNS and the CCP4 software package.⁷² Interactive model building was performed using O⁷³ and Xtalview.⁷⁴

Protein Data Bank accession code

The crystallographic coordinates for the complex structures presented in this work have been deposited with the RCSB Protein Data Bank (<http://www.rcsb.org>) wit

accession codes 2ANZ, 2AQD, 2AS1, 2AS2, 2AS3, 2AS4, 2AS6, 2EUN, 2EUP, 2EUQ, 2EUO, 2EUR, 2EUS, 2EUT, 2EUU.

1.7 Acknowledgments

This work was supported by NIH grants GM59957 (to BKS) and GM41049 (to DBG). RB was supported by a fellowship from the Ernst Schering Research Foundation. We thank OpenEye Scientific Software (Santa Fe, NM) for the use of Omega, OEChem and other tools, and MDL (San Leandro, CA) for use of the ACD database and ISISBase. We thank Kerim Babaoglu for help in data collection, Michael Keiser for his help with the similarity search, BinQing Wei for many conversations and Alan Graves for reading this manuscript.

1.8 References

1. Kitchen, D. B., Decornez, H., Furr, J. R. & Bajorath, J. (2004). Docking and scoring in virtual screening for drug discovery: methods and applications. *Nat Rev Drug Discov* 3, 935-49.
2. Shoichet, B. K. (2004). Virtual screening of chemical libraries. *Nature* 432, 862-5.
3. Alvarez, J. C. (2004). High-throughput docking as a source of novel drug leads. *Curr Opin Chem Biol* 8, 365-70.
4. Abagyan, R. & Totrov, M. (2001). High-throughput docking for lead generation. *Curr Opin Chem Biol* 5, 375-82.
5. Huang, D., Luthi, U., Kolb, P., Edler, K., Cecchini, M., Audetat, S., et al. (2005). Discovery of cell-permeable non-peptide inhibitors of beta-secretase by high-throughput docking and continuum electrostatics calculations. *J Med Chem* 48, 5108-11.
6. Meiering, S., Inhoff, O., Mies, J., Vincek, A., Garcia, G., Kramer, B., et al. (2005). Inhibitors of *Trypanosoma cruzi* trypanothione reductase revealed by virtual screening and parallel synthesis. *J Med Chem* 48, 4793-802.
7. Carbone, V., Ishikura, S., Hara, A. & El-Kabbani, O. (2005). Structure-based discovery of human L-xylulose reductase inhibitors from database screening and molecular docking. *Bioorg Med Chem* 13, 301-12.
8. Verras, A., Kuntz, I. D. & Ortiz de Montellano, P. R. (2004). Computer-assisted design of selective imidazole inhibitors for cytochrome p450 enzymes. *J Med Chem* 47, 3572-9.

9. Li, C., Xu, L., Wolan, D. W., Wilson, I. A. & Olson, A. J. (2004). Virtual screening of human 5-aminoimidazole-4-carboxamide ribonucleotide transformylase against the NCI diversity set by use of AutoDock to identify novel nonfolate inhibitors. *J Med Chem* 47, 6681-90.
10. Kraemer, O., Hazemann, I., Podjarny, A. D. & Klebe, G. (2004). Virtual screening for inhibitors of human aldose reductase. *Proteins* 55, 814-23.
11. Cummings, M. D., DesJarlais, R. L., Gibbs, A. C., Mohan, V. & Jaeger, E. P. (2005). Comparison of automated docking programs as virtual screening tools. *J Med Chem* 48, 962-76.
12. Graves, A. P., Brenk, R. & Shoichet, B. K. (2005). Decoys for docking. *J Med Chem* 48, 3714-28.
13. Ferrara, P., Gohlke, H., Price, D. J., Klebe, G. & Brooks, C. L., 3rd. (2004). Assessing scoring functions for protein-ligand interactions. *J Med Chem* 47, 3032-47.
14. Perola, E., Walters, W. P. & Charifson, P. S. (2004). A detailed comparison of current docking and scoring methods on systems of pharmaceutical relevance. *Proteins* 56, 235-49.
15. Wang, R., Lu, Y., Fang, X. & Wang, S. (2004). An extensive test of 14 scoring functions using the PDBbind refined set of 800 protein-ligand complexes. *J Chem Inf Comput Sci* 44, 2114-25.
16. Stahl, M. & Rarey, M. (2001). Detailed analysis of scoring functions for virtual screening. *J Med Chem* 44, 1035-1042.

17. Bissantz, C., Folkers, G. & Rognan, D. (2000). Protein-based virtual screening of chemical databases. 1. Evaluation of different docking/scoring combinations. *J Med Chem* 43, 4759-4767.
18. Doman, T. N., McGovern, S. L., Witherbee, B. J., Kasten, T. P., Kurumbail, R., Stallings, W. C., et al. (2002). Molecular docking and high-throughput screening for novel inhibitors of protein tyrosine phosphatase-1B. *J Med Chem* 45, 2213-2221.
19. Evensen, E., Eksterowicz, J. E., Stanton, R. V., Oshiro, C., Grootenhuis, P. D. & Bradley, E. K. (2003). Comparing performance of computational tools for combinatorial library design. *J Med Chem* 46, 5125-8.
20. Paiva, A. M., Vanderwall, D. E., Blanchard, J. S., Kozarich, J. W., Williamson, J. M. & Kelly, T. M. (2001). Inhibitors of dihydrodipicolinate reductase, a key enzyme of the diaminopimelate pathway of *Mycobacterium tuberculosis*. *Biochim Biophys Acta* 1545, 67-77.
21. Zavodszky, M. I. & Kuhn, L. A. (2005). Side-chain flexibility in protein-ligand binding: the minimal rotation hypothesis. *Protein Sci* 14, 1104-14.
22. Sims, P. A., Wong, C. F., Vuga, D., McCammon, J. A. & Sefton, B. M. (2005). Relative contributions of desolvation, inter- and intramolecular interactions to binding affinity in protein kinase systems. *J Comput Chem* 26, 668-81.
23. Eriksson, A. E., Baase, W. A., Wozniak, J. A. & Matthews, B. W. (1992). A cavity-containing mutant of T4 lysozyme is stabilized by buried benzene. *Nature* 355, 371-373.

24. Morton, A., Baase, W. A. & Matthews, B. W. (1995). Energetic origins of specificity of ligand binding in an interior nonpolar cavity of {T4 } lysozyme. *Biochemistry* 34, 8564-8575.
25. Wei, B. Q., Baase, W. A., Weaver, L. H., Matthews, B. W. & Shoichet, B. K. (2002). A model binding site for testing scoring functions in molecular docking. *J Mol Biol* 322, 339-355.
26. Wei, B. Q., Weaver, L. H., Ferrari, A. M., Matthews, B. W. & Shoichet, B. K. (2004). Testing a Flexible-receptor Docking Algorithm in a Model Binding Site. *J Mol Biol* 337, 1161-82.
27. Ferrari, A. M., Wei, B. Q., Costantino, L. & Shoichet, B. K. (2004). Soft docking and multiple receptor conformations in virtual screening. *J Med Chem* 47, 5076-84.
28. Chang, C. E. & Gilson, M. K. (2004). Free energy, entropy, and induced fit in host-guest recognition: calculations with the second-generation mining minima algorithm. *J Am Chem Soc* 126, 13156-64.
29. Fitzgerald, M. M., Churchill, M. J., McRee, D. E. & Goodin, D. B. (1994). Small molecule binding to an artificially created cavity at the active site of cytochrome c peroxidase. *Biochemistry* 33, 3807-18.
30. Musah, R. A., Jensen, G. M., Bunte, S. W., Rosenfeld, R. J. & Goodin, D. B. (2002). Artificial protein cavities as specific ligand-binding templates: characterization of an engineered heterocyclic cation-binding site that preserves the evolved specificity of the parent protein. *J Mol Biol* 315, 845-57.

31. Rosenfeld, R., Goodsell, D., Musah, R., Garret, M., Goodin, D. & Olson, A. (2003). Automated docking of ligands to an artificial active site: augmenting crystallographic analysis with computer modeling. *J Comput Aided Mol Des.*
32. Banba, S., Guo, Z. Y. & Brooks, C. L. (2000). Efficient sampling of ligand orientations and conformations in free energy calculations using the lambda-dynamics method. *J Phys Chem B* 104, 6903-6910.
33. Banba, S. & Brooks, C. L. (2000). Free energy screening of small ligands binding to an artificial protein cavity. *J Chem Phys* 113, 3423-3433.
34. Morris, G. M., Goodsell, D. S., Halliday, R. S., Huey, R., Hart, W. E., Belew, R. K., et al. (1998). Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J Comput Chem* 19, 1639-1662.
35. Lorber, D. M. & Shoichet, B. K. (2005). Hierarchical docking of databases of multiple ligand conformations. *Curr Top Med Chem* 5, 739-49.
36. Takano, Y. & Houk, K. N. (2004). Benchmarking the Conductor-like Polarizable Continuum Model (CPCM) for Aqueous Solvation Free Energies of Neutral and Ionic Organic Molecules. *J Chem Theory & Comp* 1, 70-77.
37. Li, J. B., Zhu, T. H., Hawkins, G. D., Winget, P., Liotard, D. A., Cramer, C. J., et al. (1999). Extension of the platform of applicability of the SM5.42R universal solvation model. *Theoretical Chemistry Accounts* 103, 9-63.
38. Musah, R. A., Jensen, G. M., Rosenfeld, R. J., McRee, D. E., Goodin, D. B. & Bunte, S. W. (1997). Variation in strength of an unconventional C-H to O hydrogen bond in an engineered protein cavity. *Journal of the American Chemical Society* 119, 9083-9084.

39. Taylor, R. & Kennard, O. (1982). Crystallographic Evidence for the Existence of C-H...O, C-H...N, and C-H...Cl Hydrogen-Bonds. *Journal of the American Chemical Society* 104, 5063-5070.
40. Pierce, A. C., Sandretto, K. L. & Bemis, G. W. (2002). Kinase inhibitors and the case for CH...O hydrogen bonds in protein-ligand binding. *Proteins* 49, 567-76.
41. Brenk, R., Irwin, J. & Shoichet, B. (2005). Here be dragons: Docking and Screening in an Uncharted Region of Chemical Space. *J Biomol Screen* in press.
42. Fitzgerald, M. M., Musah, R. A., McRee, D. E. & Goodin, D. B. (1996). A ligand-gated, hinged loop rearrangement opens a channel to a buried artificial protein cavity. *Nat Struct Biol* 3, 626-31.
43. Li, J. B., Zhu, T. H., Cramer, C. J. & Truhlar, D. G. (1998). New class IV charge model for extracting accurate partial charges from wave functions. *J Phys Chem A* 102, 1820-1831.
44. Chambers, C. C., Hawkins, G. D., Cramer, C. J. & Truhlar, D. G. (1996). Model for aqueous solvation based on class IV atomic charges and first solvation shell effects. *J Phys Chem* 100, 16385-16398.
45. Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., et al. (2004). Gaussian 03. Gaussian, Inc., Wallingford CT.
46. Schutz, C. N. & Warshel, A. (2001). What are the dielectric "constants" of proteins and how to validate electrostatic models? *Proteins* 44, 400-17.
47. Honig, B. & Nicholls, A. (1995). Classical electrostatics in biology and chemistry. *Science* 268, 1144-1149.

48. Simonson, T., Archontis, G. & Karplus, M. (1999). A Poisson-Boltzmann study of charge insertion in an enzyme active site: the effect of dielectric relaxation. *J. Phys. Chem. B* 193, 6142-6156.
49. Kuntz, I. D., Chen, K., Sharp, K. A. & Kollman, P. A. (1999). The maximal affinity of ligands. *Proc Natl Acad Sci U S A* 96, 9997-10002.
50. Carr, R. A., Congreve, M., Murray, C. W. & Rees, D. C. (2005). Fragment-based lead discovery: leads by design. *Drug Discov Today* 10, 987-92.
51. Hopkins, A. L., Groom, C. R. & Alex, A. (2004). Ligand efficiency: a useful metric for lead selection. *Drug Discov Today* 9, 430-1.
52. Minke, W. E., Diller, D. J., Hol, W. G. & Verlinde, C. L. (1999). The role of waters in docking strategies with incremental flexibility for carbohydrate derivatives: heat-labile enterotoxin, a multivalent test case. *J Med Chem* 42, 1778-88.
53. de Graaf, C., Pospisil, P., Pos, W., Folkers, G. & Vermeulen, N. P. (2005). Binding mode prediction of cytochrome p450 and thymidine kinase protein-ligand complexes by consideration of water and rescoring in automated docking. *J Med Chem* 48, 2308-18.
54. Rarey, M., Kramer, B. & Lengauer, T. (1999). The particle concept: placing discrete water molecules during protein-ligand docking predictions. *Proteins* 34, 17-28.
55. Osterberg, F., Morris, G. M., Sanner, M. F., Olson, A. J. & Goodsell, D. S. (2002). Automated docking to multiple target structures: incorporation of protein mobility and structural water heterogeneity in AutoDock. *Proteins* 46, 34-40.

56. Gerber, P. R. (1998). Charge distribution from a simple molecular orbital type calculation and non-bonding interaction terms in the force field MAB. *J Comput Aided Mol Des* 12, 37-51.
57. Cornell, W. D., Cieplak, P., Bayly, C. I. & Kollman, P. A. (1993). Application of Resp Charges to Calculate Conformational Energies, Hydrogen-Bond Energies, and Free-Energies of Solvation. *J Am Chem Soc* 115, 9620-9631.
58. Nicholls, A. & Honig, B. (1991). A Rapid Finite-Difference Algorithm, Utilizing Successive over-Relaxation to Solve the Poisson-Boltzmann Equation. *J Comput Chem* 12, 435-445.
59. Kuntz, I. D., Blaney, J. M., Oatley, S. J., Langridge, R. & Ferrin, T. E. (1982). A Geometric Approach to Macromolecule-Ligand Interactions. *Journal of Molecular Biology* 161, 269-288.
60. Meng, E. C., Shoichet, B. K. & Kuntz, I. D. (1992). Automated docking with grid-based energy evaluating. *J Comb Chem* 13, 505-524.
61. Verdonk, M. L., Berdini, V., Hartshorn, M. J., Mooij, W. T., Murray, C. W., Taylor, R. D., et al. (2004). Virtual screening using protein-ligand docking: avoiding artificial enrichment. *J Chem Inf Comput Sci* 44, 793-806.
62. Besler, B. H., Merz, K. M. & Kollman, P. A. (1990). Atomic Charges Derived from Semiempirical Methods. *J Comput Chem* 11, 431-439.
63. Cossi, M., Rega, N., Scalmani, G. & Barone, V. (2003). Energies, structures, and electronic properties of molecules in solution with the C-PCM solvation model. *J Comput Chem* 24, 669-681.

64. Barone, V. & Cossi, M. (1998). Quantum calculation of molecular energies and energy gradients in solution by a conductor solvent model. *J Phys Chem A* 102, 1995-2001.
65. Martin, Y. C., Kofron, J. L. & Traphagen, L. M. (2002). Do structurally similar molecules have similar biological activity? *J Med Chem* 45, 4350-8.
66. Goodin, D. B., Davidson, M. G., Roe, J. A., Mauk, A. G. & Smith, M. (1991). Amino acid substitutions at tryptophan-51 of cytochrome c peroxidase: effects on coordination, species preference for cytochrome c, and electron transfer. *Biochemistry* 30, 4953-62.
67. Pflugrath, J. W. (1999). The finer things in X-ray diffraction data collection. *Acta Crystallogr D Biol Crystallogr* 55 (Pt 10), 1718-25.
68. Otwinowski, Z. & Minor, W. (1997). Processing of X-ray Diffraction Data Collected in Oscillation Mode. In *Methods in Enzymology* (Carter, C. W. J. & Sweet, R. M., eds.), Vol. 276, pp. 307-326. Academic Press, New York.
69. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., et al. (1998). Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr* 54, 905-921.
70. Sheldrick, G. & Schneider, T. (1997). SHELXL: High Resolution Refinement. *Methods in Enzymologie* 277, 319-343.
71. Schuttelkopf, A. W. & van Aalten, D. M. F. (2004). PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallographica Section D-Biological Crystallography* 60, 1355-1363.

72. Bailey, S. (1994). The Ccp4 Suite - Programs for Protein Crystallography. *Acta Crystallogr D* 50, 760-763.
73. Jones, T. A. & Kjeldgaard, M. (1997). Electron-density map interpretation. *Methods in Enzymology* 277, 173—208.
74. McRee, D. E. (1999). XtalView/Xfit--A versatile program for manipulating atomic coordinates and electron density. *J Struct Biol* 125, 156-65.

Gloss to chapter 2.

In Chapter 2 we set out to test the ability of molecular mechanics-generalized Born surface area (MM-GBSA) methods to rescore top ranking molecules from a docking hit list. The DOCK scoring function, while physically derived, is at best an approximation of the forces that determine protein-ligand interactions. As a brute force method it is useful to screen many molecules against a protein target, but it fails in many instances, due in part to the aspects of the algorithm that make it successful for screening large databases, it's very simplicity. The scoring function lacks treatment of water in the cavity (both ordered waters and receptor desolvation), entropic terms, ligand internal energies, and changes in the protein (receptor flexibility) upon ligand binding are not considered.

The MM-GBSA methods considered in this study, PLOP^{1,2} from Matt Jacobson and Amberdock from Dave Case's group, are more physically realistic methods, with both minimization and the potential for sampling of protein-ligand complexes, the ability to capture ligand and receptor strain, and to account for ligand and receptor solvation using Poisson Boltzmann (PB) or generalized Born (GB) methods. However, they are orders of magnitude slower than docking and are therefore not suitable for screening large databases in a reasonable timeframe. Instead, we sought to apply the higher level of theory as a post-processing or "rescoring" step after the initial docking run to a percentage of the top hits. In this way we would gain the advantages of using the higher level of theory without having to sacrifice too deeply in computational cost.

The questions we were interested in ranged from the ability of MM-GBSA methods to find new chemotypes overlooked by DOCK, to “recover” both known and new ligands ranked poorly in the initial screen and, of course, to accurately predict ligand binding modes. But precisely because we were doing this study in our model systems, where we can extensively test the methods prospectively, we were also interested in finding the limitations of the MM-GBSA methods. Although they represent a higher level of theory than docking, MM-GBSA methods still include many approximations, areas ripe for failure and therefore places to learn and to make improvements.

CCP W191G, now an established model system in the group, was included in a panel of three model systems; these systems ranged from the least complicated hydrophobic T4 lysozyme L99A site, to the similar L99A/M102Q hydrophobic but slightly polar site to the polar and anionic CCP W191G cavity. The value of testing against a panel of model systems was enormous; for example, weaknesses in the solvation estimation or ligand charge model, not observed in a simple site such as L99A, were visible in CCP W191G, reflected in the unusually large number of neutral and dication molecules predicted in the top of the rescored hit lists. Additionally, the degree to which the MM-GBSA methods relaxed the individual cavities allowed larger ligands to dock. Depending on the cavity and the initial dock pose this could result in a recovered ligand rank and pose relative to the original docking result, for example 2-ethoxyphenol rescued as a ligand for L99A. Or it could lead to a recovered ligand with a decoy pose, such as 2,4,6-triaminopyrimidine, which was predicted with an incorrect binding mode due to the limited protein sampling in the MM-GBSA methods which could not account for the large P190-Asn195 loop movement in CCP W191G.

The differences in the MM-GBSA methods' performance in these three systems highlighted the importance of including all three in this study. To that end I would like to thank Alan Graves and Devleena Shivakumar, co-first authors on this paper, for their foresight in recognizing the significant contribution that CCP W191G could add to the project. Besides doing the initial docking for L99A and M102Q, and rescoring with PLOP the top hit lists for the model systems L99A, M102Q and CCP W191G, Alan also selected and experimentally tested 19 compounds for binding and determined protein crystal structures of 10 ligands in complex with the model systems L99A and M102Q. Devleena Shivakumar, from Dave Case's group, was the driving force in the development of Amberdock; she rescored the top molecules from the L99A, M102Q and the entire CCP W191G hit list. From both the PLOP and Amberdock rescored hit lists for CCP I selected 14 compounds for testing and crystallography, resulting in 10 new ligands, 4 decoys, and 10 ligand co-complex structures for CCP W191G. The Supplementary Materials for this chapter are included in Appendix B.

References

1. Jacobson, M. P., Kaminski, G. A., Friesner, R. A. & Rapp, C. S. (2002). Force Field Validation Using Protein Side Chain Prediction. *J. Phys. Chem. B* **106**, 11673-11680.
2. Lorber, D. M. & Shoichet, B. K. (1998). Flexible ligand docking using conformational ensembles. *Protein Sci* **7**, 938-50.

Chapter 2:

Rescoring docking hit lists for model cavity sites: predictions and experimental testing

Alan P. Graves,^{a,b,†} Devleena M. Shivakumar,^{c,†,‡} Sarah E. Boyce,^{a,d} Matthew P.

Jacobson,^{a,*} David A. Case,^{c,*} and Brian K. Shoichet^{a,*}

^a Department of Pharmaceutical Chemistry,

^b Graduate Group in Biophysics,

^c Department of Molecular Biology, The Scripps Research Institute, 10550 N. Torrey Pines Road, La Jolla, CA 92037, and

^d Graduate Group in Chemistry and Chemical Biology, University of California San Francisco, 1700 4th Street, San Francisco, CA 94158-2330

[†] These authors contributed equally to this work

[‡] Current address: Institute for Molecular Pediatric Sciences, University of Chicago, 929 E. 57th Street, Chicago, IL 60637

* Email addresses of corresponding authors:

BKS: shoichet@cgl.ucsf.edu, MPJ: matt.jacobson@ucsf.edu, DAC: case@scripps.edu

2.1 Abstract

Molecular docking computationally screens thousands to millions of organic molecules against protein structures, looking for those with complementary fits. Many approximations are made, often resulting in low “hit rates.” A strategy to overcome these approximations is to rescore top-ranked docked molecules using a better but slower method. One such is afforded by Molecular Mechanics-Generalized Born Surface Area (MM-GBSA) techniques. These more physically realistic methods have improved models for solvation and electrostatic interactions and conformational change compared to most docking programs. To investigate MM-GBSA rescoring, we re-ranked docking hit lists in three small, buried sites: a hydrophobic cavity that binds apolar ligands, a slightly polar cavity that binds aryl and hydrogen-bonding ligands, and an anionic cavity that binds cationic ligands. These sites are simple; consequently incorrect predictions can be attributed to particular errors in the method, and many likely ligands may actually be tested. In retrospective calculations, MM-GBSA techniques with binding site minimization better distinguished the known ligands for each cavity from the known decoys, compared to the docking calculation alone. This encouraged us to test rescoring prospectively on molecules that ranked poorly by docking but that ranked well when rescored by MM-GBSA. A total of 33 molecules highly ranked by MM-GBSA for the three cavities were tested experimentally. Of these, 23 were observed to bind—these are docking false negatives rescued by rescoring. The ten remaining molecules are true negatives by docking and false positives by MM-GBSA. X-ray crystal structures were determined for 21 of these 23 molecules. In many cases, the geometry prediction by MM-

GBSA improved the initial docking pose and more closely resembled the crystallographic result; yet in several cases, the rescored geometry failed to capture large conformational changes in the protein. Intriguingly, rescoring not only rescued docking false positives, but also introduced several new false positives into the top-ranking molecules. We consider the origins of the successes and failures in MM-GBSA rescoring in these model cavity sites and the prospects for rescoring in biologically relevant targets.

Abbreviations

L99A, Leu99 → Ala mutant of T4 lysozyme

L99A/M102Q, Leu99 → Ala and Met102 → Gln double mutant of T4 lysozyme

CCP, Trp191 → Gly mutant of Cytochrome C Peroxidase

MM-GBSA, molecular mechanics with generalized Born surface area approximation

PLOP, Protein Local Optimization Program

RMSD, root mean square deviation

ACD, the Available Chemicals Directory

CD, circular dichroism

UV-VIS, ultraviolet visible

PDB, the Protein Data Bank

HTS, high throughput screening

Keywords

Decoys, molecular docking, virtual screening, MM-GBSA, cavity, drug design

2.2 Introduction

Molecular docking computationally screens large databases of small molecules against a macromolecular binding site of defined structure. The technique is often used to find novel ligands for drug discovery. Notwithstanding important successes,^{1; 2; 3; 4; 5; 6; 7; 8} docking continues to struggle with many methodological deficits. Many approximations are made to screen many molecules in a timely fashion. These include using only one conformation of the protein, neglecting the internal energies of the docking molecules, using simplified models of ligand solvation energies, typically ignoring protein desolvation, and ignoring most entropic terms entirely. These and other short-cuts lead to the high false positive and false negative rates for which docking screens are notorious. Docking methods are unreliable for affinity prediction and, except in domains of highly related compounds, even for rank ordering the likely hits that emerge from the virtual screens.

To overcome these deficits, several groups have combined disparate scoring functions in a consensus fashion to capitalize on the strengths and overcome the deficiencies of individual methods.^{9; 10; 11; 12} This “consensus scoring” approach is attractive when it has worked, but its theoretical underpinnings are slim.¹³ An alternative approach involves using a higher level of theory to re-score the docking hit lists after the docking calculation has completed. The goal is to re-evaluate the top docking hits for energetic complementarity to the target after including more terms and degrees of freedom than modeled by the docking program. Because more terms are considered, rescoring is typically much slower than docking, so much so that only the top-scoring docking pose of the best scoring docked molecules are often considered. This approach

has been adopted by versions of the program GLIDE.¹⁴ Here ligands are first docked using simplified and relaxed criteria and are then refined by more sophisticated and stringent evaluation of the energies of binding. Similarly, Kollman used a hierarchical technique that begins with initial database screening and progresses to Molecular Mechanics-Poisson-Boltzmann Surface Area (MM-PBSA) rescoring to find HIV-1 Reverse Transcriptase inhibitors.¹⁵ The combination of an initial docking screen with subsequent re-scoring by a Molecular Mechanics-Generalized Born Surface Area (MM-GBSA) method has been used to improve enrichment of known ligands for several enzymes in retrospective studies and even to identify substrates.^{16; 17; 18; 19; 20}

Such MM-PBSA and MM-GBSA methods involve minimization and often dynamic sampling of the protein-ligand complexes, and include ligand and receptor conformational energies and strain. They evaluate the electrostatics and solvation components of the binding energy by PB or GB methods, including both ligand *and* receptor desolvation. The MM-GBSA binding energy is determined by

$E_{complex} - E_{receptor} - E_{ligand}$ where E is an MM-GBSA estimate and solute configurational entropy effects are ignored. In this paper, we focus on relative binding energies of different ligands to the same receptor, so the free receptor energy ($E_{receptor}$) does not affect the results. Because the MM-GBSA function includes both internal energies and solvation free energies, and because we explicitly subtract complex ($E_{complex}$) and ligand (E_{ligand}) contributions, desolvation effects upon complex formation for both the ligand and the receptor are included, at least in principle. There are three main limitations: (1) the force fields and solvation energies are not uniformly accurate; (2) for reasons of computational efficiency, only a small part of configuration space near the DOCK

starting pose is really explored; and (3) configurational entropy effects are ignored. Notwithstanding these limitations, the MM-GBSA methods represent a substantially higher level of theory than that encoded by most docking programs and are attractive alternatives to a more complete treatment of the energies of interaction by free energy perturbation (FEP) and thermodynamic integration (TI),²¹ which remain the gold standard but are very slow.

In this study, we set out to test MM-GBSA rescoring of docking hit lists in simple model cavity sites. These sites have been engineered into the buried cores of proteins and bind multiple small organic molecules. In contrast to most drug targets, these cavities are small (150-180 Å³), buried from bulk solvent, and are dominated by a single interaction term. The L99A (Leu99→Ala) cavity in T4 lysozyme²² is almost entirely apolar, the L99A/M102Q (Leu99→Ala/Met102→Gln)²³ cavity in the same protein has a single hydrogen-bond acceptor (the introduced Gln102), whereas the W191G (Trp191→Gly) cavity in Cytochrome C Peroxidase (CCP)^{24; 25} has a single anionic residue, Asp235 (Figure 1). The ligands recognized by these sites correspond to these features: the hydrophobic L99A binds small, typically aromatic non-polar molecules; the slightly polar L99A/M102Q binds both apolar molecules but also those bearing one or two hydrogen-bond donors; whereas, the anionic W191G cavity almost exclusively binds small monocations. The simplicity of these sites is conducive to disentangling the energetic terms of ligand binding, which are so often convoluted in drug targets with their larger, more complex binding sites. It should be noted that previous work with solvent exposed sites has suggested that a major advantage of MM-GBSA scoring functions is calculating partial receptor desolvation upon ligand binding.¹⁷ This benefit with complex solvent

exposed binding sites may be less relevant in the buried cavity sites, especially the hydrophobic L99A and polar L99A/M102Q sites, which are mostly desolvated. (It is our experience that the cavity sites, in fact, impose a greater strain on the GBSA solvent models to fully desolvate the pockets.)

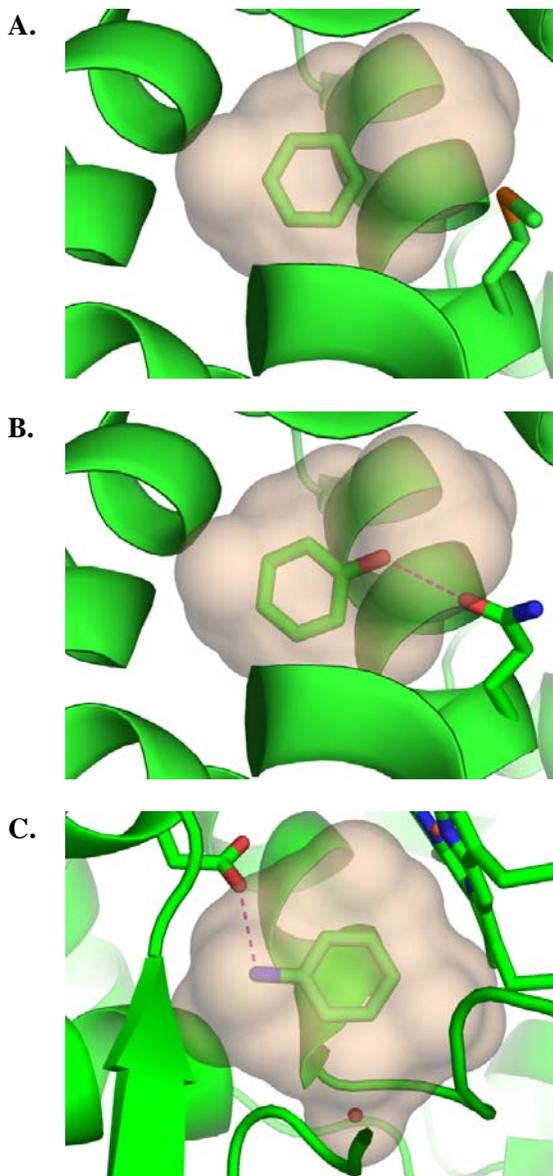


Figure 1. The model cavity sites. A. Cavity binding site in T4 lysozyme L99A with benzene bound. B. Cavity binding site in T4 lysozyme L99A/M102Q with phenol bound; the hydrogen bond with the O ϵ 2 oxygen of Gln102 is represented by a dashed line. C. Cavity binding site of cytochrome C peroxidase W191G with aniline bound; the hydrogen bond with Asp235 is represented by a dashed line. The heme and an ordered water molecule are also depicted. In A., B., and C. the cavities are represented by a tan molecular surface and the protein ribbons are colored green. Rendered with the program PyMOL.⁶³

In the cavity sites, as in other simplified sites,²⁶ an incorrect prediction is often informative, identifying a single problematic term in a scoring function; we have used these cavities as model binding sites to identify problems in molecular docking^{23; 27; 28; 29} and, more recently, thermodynamic integration.²¹ Others have found them attractive test systems for methods development studies.^{30; 31; 32; 33} An important advantage of these cavity sites is that they are experimentally tractable for detailed, prospective testing of ligand predictions. Because the ligands they bind are small—in the 70 to 150 amu range—many possible ligands are readily available commercially, which is rarely true of drug targets.³⁴ The binding of these predicted ligands may be tested by direct binding assays, and the structures of the ligand-protein complexes may be routinely determined by x-ray crystallography to resolutions better than 2 Å. Extensive study in the Matthews, Goodin, and our own laboratories has resulted in many tens of diverse ligands for each cavity, as well as tens of “decoys,” which are molecules that were predicted to bind to the sites but for which no binding was observed at concentrations as high as 10 mM on experimental testing.^{21; 23; 27; 28; 29}

We thus used these three simple model cavity sites, L99A, L99A/M102Q, and W191G, as templates to measure the strengths and weaknesses of MM-GBSA rescoring of docking hit lists. We used two rescoring programs: PLOP^{35; 36}, with binding site side chain rotamer search and minimization, and AMBERDOCK, using short MD steps and minimization of binding site residues (Materials and Methods). Molecular docking was used to screen compound libraries that contained between 5000 and 60,231 fragment-like molecules from the Available Chemicals Directory (ACD); the library size was chosen to partly mitigate issues of size and charge bias from the library alone, and to be consistent

with earlier studies in these sites (Results).^{27,28} The single best pose for each compound that ranked among the top 5000 or 10000 compounds by docking was then rescored by both MM-GBSA programs. Multiple known ligands and decoys were among the molecules rescored for all three sites' rescored sets. In retrospective calculations, MM-GBSA rescoring improved the separation of ligands from decoys in each of the cavities. We then tested 33 new ligands that were predicted to bind by the MM-GBSA methods that docking alone ranked poorly—generally much worse than the top 500. To investigate the detailed basis of the MM-GBSA predictions, we determined crystal structures for 21 of these new ligands and compared them to the geometries predicted by theory. These studies suggest areas where MM-GBSA methods can contribute to the success of virtual screening, and areas where this method faces important challenges.

2.3 Results

Retrospective Docking and Rescoring in the Hydrophobic Cavity.

Approximately 60,000 small molecules were docked into the hydrophobic cavity L99A using DOCK3.5.54^{23,37} (Figure 1a). The compounds in this set were selected from a much larger library so as not to exceed 25 non-hydrogen atoms, as previously described.²⁸ This reduced the enrichment-factor bias that would have otherwise occurred by the trivial ability of the docking program to remove compounds that were simply too large to fit in the cavities. We note that reducing the number of molecules to 60,000 from the several million that are in the ACD or ZINC³⁸ databases has the effect of *reducing* our enrichment factors. Among the top-scoring 10,000 molecules were 39 known ligands and 40 experimentally tested decoys. DOCK found 44% (17 molecules) of these ligands

and 43% (17 molecules) of these decoys among the top 500 molecules (Figure 2a).

Ligands such as toluene (DOCK rank 32), benzene (DOCK rank 151), and ethylbenzene (DOCK rank 301) are small, aromatic and hydrophobic compared to known decoys such as nitrosobenzene (DOCK rank 125), phenol (DOCK rank 234), and 3-methylpyrrole (DOCK rank 435). Like the ligands, these decoys are also small and aromatic, but are presumably too polar for the hydrophobic cavity to overcome their desolvation penalty (Figure 1a).

The top-ranking 10,000 docking hits for the hydrophobic cavity were re-ranked by PLOP and the top-ranking 5,000 docking hits were re-ranked by AMBERDOCK; fewer molecules were treated by AMBERDOCK simply because it was much more computationally intensive than PLOP. For both methods, the enrichment of the ligands actually decreased slightly relative to that achieved by docking alone; that is to say, fewer ligands were found among the very best scoring molecules (Figure 2a). Rescored by PLOP, 41% (16 molecules) of the known ligands were found among the top 500 molecules, whereas 28% (11 molecules) were found by AMBERDOCK. Both enrichment factors were lower than those found by docking alone. On the other hand, the enrichment of the known decoys was lower still (Figure 2a). Only 5% of the decoys (2 molecules) were ranked among the top 500 molecules by PLOP and only 13% (5 molecules) were so ranked by AMBERDOCK. This represents a substantial improvement on docking alone, one that reflects a significant change in the relative energies of the ligands and decoys. For instance, in the L99A cavity the average differential energy between the first ten ligands and the first ten decoys was only 0.7 kcal/mol by docking. Meanwhile, the average total energy for the top ten docked ligands was -15.8 kcal/mol and the difference

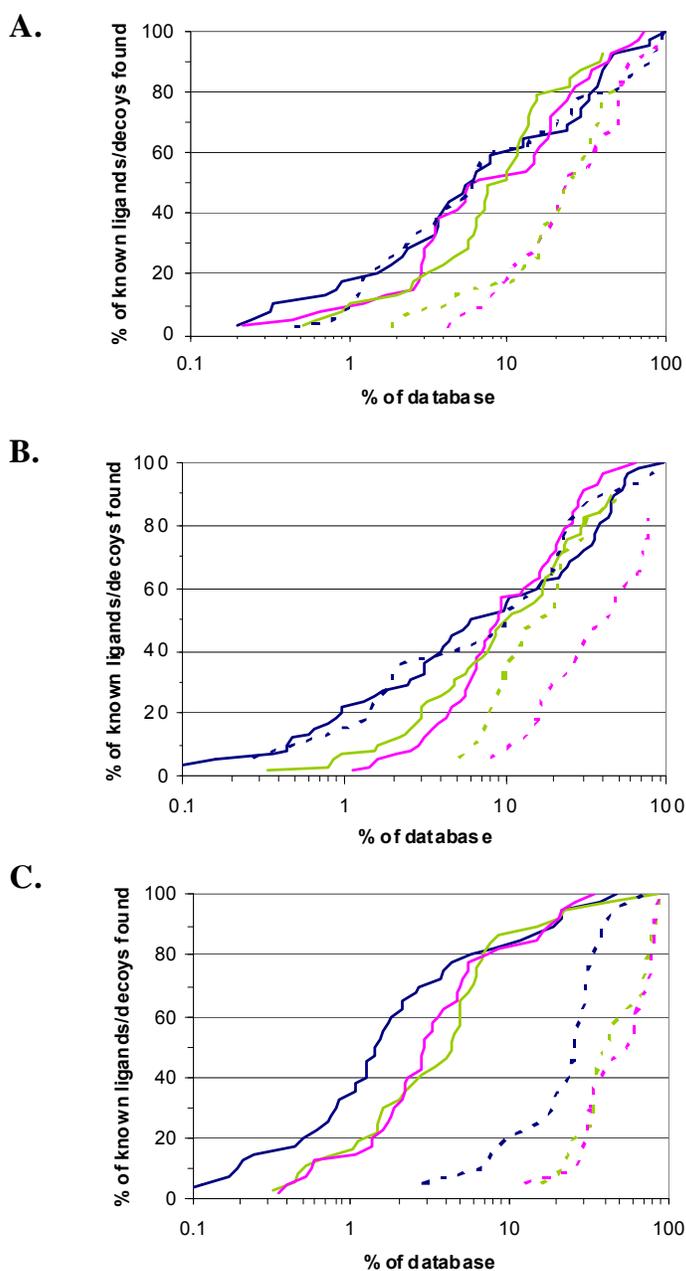


Figure 2. Retrospective enrichment of ligands and decoys for **(a)** the hydrophobic L99A cavity, **(b)** the polar L99A/M102Q cavity, and **(c)** the anionic W191G cavity. The plots depict the percentage of known ligands (continuous lines) or decoys (dashed lines) found (y-axis) at each percentage level of the ranked database using the top 10,000 best scoring docking hits (x-axis) for L99A **(a)** and L99A/M102Q **(b)** and the 5400 best scoring docking hits (x-axis) for CCP **(c)**. Docking enrichment of known ligands (continuous lines) and decoys (dashed lines) are represented by the dark blue curves. PLOP enrichment of known ligands (continuous lines) and decoys (dashed lines) are represented by the pink curves. AMBERDOCK enrichment of known ligands (continuous lines) and decoys (dashed lines) are represented by green curves.

between 1st and the 10th ranked ligand is 2.9 kcal/mol; the ligands and decoys were essentially indistinguishable by docking energy. For the PLOP rescored molecules, conversely, the average difference in energies for the top ten ligands and decoys was 4.0 kcal/mol. Meanwhile, average energy for the top ten ligands was -21.7 kcal/mol and the difference between top ranked ligand the 10th was 5.5 kcal/mol; the best ligands and decoys are separated significantly by rescored energy. We should note that both the ligand enrichment and the decoy enrichment are strongly biased for docking—many of the ligands and almost all of the decoys were originally tested based on docking predictions^{23; 28; 29}—so it is reasonable to expect that the enrichment of ligands will be higher by docking, as will the decoys. Perhaps more informative than is the separation of the ligands from the decoys, as measured by the ratios of their enrichment factors. These were improved eight-fold by PLOP and two-fold for AMBERDOCK, relative to that of DOCK in this hydrophobic cavity.

Retrospective Docking and Rescoring in the Polar Cavity.

The same 60,000 molecules were docked into the polar cavity L99A/M102Q (Figure 1b). Among the top-scoring 10,000 molecules were 58 ligands and 17 experimentally tested decoys. DOCK found 45% (26 molecules) of these ligands and 35% (6 molecules) of these decoys among the top 500 molecules (Figure 2b). The increased polarity from O ϵ of the Gln102 side chain in the cavity accommodates the binding of phenol (DOCK rank 354) and 3-methylpyrrole (DOCK rank 307), which are decoys for the L99A cavity, as well as hydrophobic ligands such as toluene (DOCK rank 16) and benzene (DOCK rank 78). The increased polarity of the site only goes so far, however, and it cannot accommodate decoys such as 1-vinylimidazole (DOCK rank 136)

or 2-aminophenol (DOCK rank 208), whose polarity is presumably still too great for the single carbonyl oxygen of the site to overcome the attendant desolvation terms.

The top 10,000 docking hits for the polar cavity were re-ranked by PLOP and the top 5,000 re-ranked by AMBERDOCK. For both methods, the enrichment of the ligands again decreased slightly relative to the docking enrichment factor (Figure 2b). Rescored by PLOP, 22% (13 molecules) of the known ligands were found among the top 500 molecules, whereas 34% (20 molecules) were found by AMBERDOCK. However, the enrichment of the known decoys was lower still. None of the decoys were ranked among the top 500 molecules by PLOP or AMBERDOCK, in contrast to DOCK where 35% (6 molecules) of the known decoys were scored among the top 500 molecules. As in the hydrophobic site, despite the decrease in overall ligand enrichment, the separation of the ligands from the decoys was improved substantially for the polar cavity: by 20 fold for PLOP and 35 fold for AMBERDOCK.

Retrospective Docking and Rescoring in the Anionic Cavity.

Approximately 5400 molecules were docked in the charged cavity of CCP (Figure 1c). This library was also selected from a much larger set to reduce enrichment-factor bias from trivial physical non-complementarity between library molecules and the CCP cavity.²⁷ Thus, any molecules from the larger ACD that had unfavorable van der Waals scores (i.e., simply did not fit), or that bore an anionic charge, were removed from the larger library. As with the lysozyme cavities, this smaller library of more physically plausible ligands reduces the enrichment factors we would otherwise achieve with docking. Within this database were 40 known ligands and 20 experimentally tested decoys. DOCK found 78% (31 molecules) of these ligands and 20% (4 molecules) of

these decoys among the top 500 molecules (Figure 2b). The anionic cavity typically binds cationic ligands such as 2-aminopyridine (DOCK rank 6) and imidazole (DOCK rank 227). Most neutral polar compounds, such as 3,5-difluoroaniline (DOCK rank 148), and apolar compounds, such as toluene (DOCK rank 411), are decoys for this cavity, as are anionic compounds or those bearing a formal charge greater than +1.

All of the 5400 docking hits for the anionic cavity were re-ranked by PLOP and AMBERDOCK. Rescored by PLOP, 83% (33 molecules) of the known ligands were found among the top 500 molecules, and 80% (32 molecules) were found by AMBERDOCK (Figure 2c). Both enrichment factors are comparable to those found by docking alone, which found 83% (33 molecules) of the known ligands among the top 500 molecules. On the other hand, fewer of the known decoys were enriched by the MM-GBSA methods. None of the known decoys were ranked among the top 500 molecules by PLOP or AMBERDOCK, and the best scoring decoy ranked 655 for PLOP and 785 for AMBERDOCK compared to 145 for docking. Thus, whereas the overall enrichment of the ligands relative to the rest of the database molecules remained unchanged, the separation of the ligands from the decoys was improved by four-fold for PLOP and AMBERDOCK.

Prediction and Experimental Testing of New Ligands.

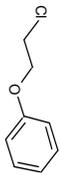
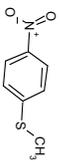
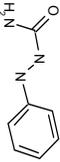
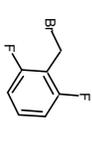
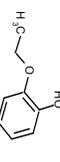
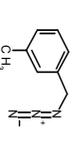
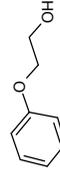
A more robust test, one less biased by previous knowledge, involves prospective prediction of new ligands. For each of the three cavities, we looked for molecules that had been poorly ranked by docking but that ranked well by either PLOP or AMBERDOCK or both. We note that our use of “well” and “poorly” ranked is inexact because there is no fully reliable way to separate molecules based on docking energies

alone. We therefore looked for molecules where the ranking's changed substantially—typically rising from ranks lower than 1500 to ranks in the top 200. Of the 33 molecules selected, 24 were ranked worse than 1500th by docking, seven were ranked between 500 and 1500, and two were ranked between 300 and 500. The rankings of all 33 rose to be among the top 200 on rescoring. Our choice of 200 was purely pragmatic, as it is a reasonable number of top ranking hits to visualize and consider for testing, which is often done when picking docking hits; another reasonable cutoff would have been top 500. Nine compounds were picked and tested for the hydrophobic L99A cavity, ten were tested for the polar L99A/M102Q cavity, and fourteen were tested for the anionic W191G cavity. Structures for 21 of these 33 molecules in complex with the cavities were determined by protein crystallography, allowing us to compare the predicted and experimental geometries in detail. In the following discussion, we report whether binding was detected at a single concentration tested. The actual affinities were not measured but will often be substantially better than the concentration reported.

New L99A Ligands Predicted by Rescoring.

All of the nine ligands predicted by PLOP and AMBERDOCK were relatively large compounds that do not easily fit into the unminimized cavity into which they were docked, explaining their poor docking ranks, but they fit well upon receptor relaxation by MM-GBSA. Binding was detected at millimolar concentrations by temperature of melting (T_m) upshift experiments for seven of these nine compounds; however, for two no binding was detected (Table 1). AMBERDOCK correctly predicted binding for five ligands and incorrectly predicted binding for 1-phenylsemicarbazide (3) and 2-phenoxyethanol (9)

Table 1. Compounds predicted by AMBERDOCK and PLOP to bind to T4 Lysozyme L99A.

Structure	Compound (ID)	Score and Rank ^a			C ^b	pH	ΔH	ΔT_m	Binding detected	Structure determined
		DOCK	AMBER	PLOP						
	β -chlorophenetole (1)	-4.89 (3786)	-22.38 (5)	-26.31 (15)	10	3	31	6.5	Yes	Yes
	4-(methylthio)nitrobenzene (2)	-5.69 (3358)	-22.36 (6)	-16.22 (12.43)	<10	3	6.2	1.3 ^c	Yes	Yes
	1-phenylsemicarbazide (3)	-4.49 (3965)	-22.03 (8)	-7.69 (5290)	10	6.8	1.6	-0.9 ^c	No	No
	2,6-difluorobenzyl bromide (4)	-10.59 (1046)	-22.01 (9)	-21.1 (186)	<10	3	10	1.6 ^c	Yes	Yes
	2-ethoxyphenol (5)	-6.74 (2806)	-21.54 (12)	-15.19 (1642)	5	3	12	1.2	Yes	Yes
	3-methylbenzylazide (6)	-10.54 (1061)	-19.58 (57)	-25.19 (27)	10	3	8.1	1.5	Yes	Yes
	Cis-3-hexenyl formate (7)	3.61 (7746)	NR ^d	-25.17 (28)	10	3	5.1	1.5	Yes	No
	6-methyl-1,5-heptadiene (8)	1.78 (7035)	NR ^d	-24.92 (30)	10	3	14.9	2.8	Yes	No
	2-phenoxyethanol (9)	-5.76 (3323)	-19.64 (55)	-23.13 (68)	10	3	2.7	0	No	No

^a Compound scores and ranks (in parenthesis) for DOCK, AMBERDOCK, and PLOP.

Scores and ranks in bold font indicate ligands which rank in the top 200 for the respective scoring function. ^b Concentration at which ligand was tested. ^c ΔT_m monitored using fluorescence, exciting at $\lambda=283\text{nm}$ and measuring the integrated emission above 300 nm. ^d NR is not ranked.

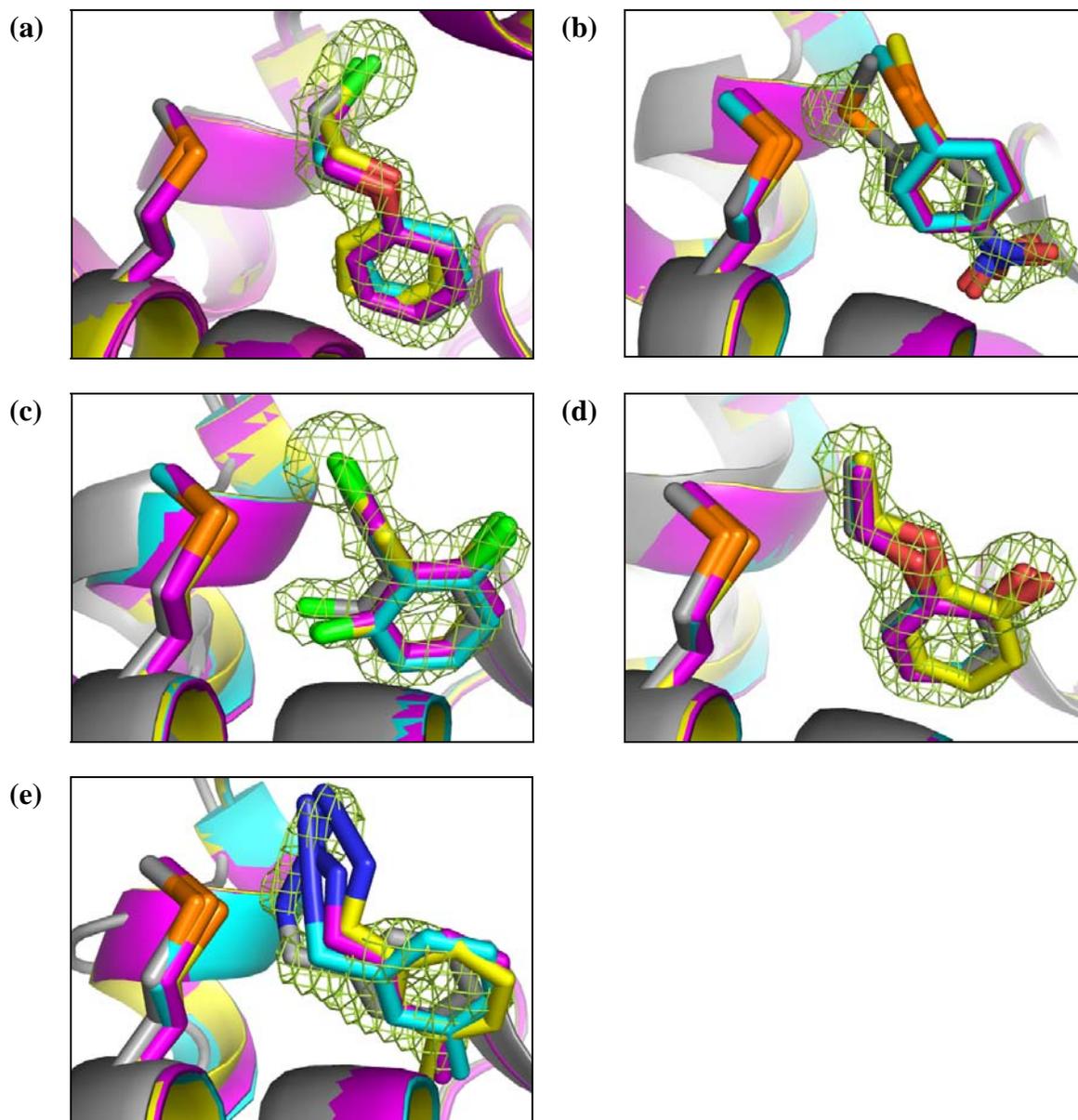


Figure 3. Predicted and experimental ligand orientations for the hydrophobic L99A cavity. The carbons of the crystallographic pose, the DOCK predicted pose, the AMBERDOCK predicted pose, and the PLOP predicted pose are colored grey, yellow, cyan, and magenta, respectively. The $f_o - f_c$ omit electron density maps (green mesh) are contoured at $2.5-3.0\sigma$ **(a)** β -chlorophenetole (1), **(b)** 4-(methylthio)nitrobenzene (2), **(c)** 2,6-difluorobenzylbromide (4), **(d)** 2-ethoxyphenol (5), and **(e)** 3-methylbenzylazide (6) bound to L99A. Rendered with the program PyMOL.²⁶

(two of the prospectively tested molecules were not rescored by AMBERDOCK because docking ranked them worse than 5000). PLOP correctly predicted binding for five ligands, while incorrectly predicted binding for 2-phenoxyethanol (9). PLOP agreed with docking on the remaining three molecules that had been prioritized by AMBERDOCK, ranking them worse than 1000. Two of these, 4-(methylthio)nitrobenzene (2) and 2-ethoxyphenol (5), were true ligands and so are false negatives for PLOP.

Five high resolution (better than 2 Å) protein-ligand crystal structures were obtained for these new L99A ligands to compare experimental to predicted poses (Figure 3). In each case, electron density for the ligands was unambiguous, allowing us to model their positions in the site. Docking and MM-GBSA methods predicted the binding geometry for three of the five ligands to within 0.3 to 0.8 Å RMSD (Table 2). Conversely, the docked pose of 3-methylbenzylazide (6) was 1.4 Å RMSD from the crystallographic pose. The PLOP minimized prediction had a slightly improved RMSD of 1.1 Å, but the refined ligand also had a non-linear azide group, highlighting a failure in ligand parameterization. In addition, docking and MM-GBSA methods predicted poses which were approximately 1.5 Å RMSD from the crystallographic pose of 4-(methylthio)nitrobenzene (2). The crystallographic poses of these two ligands would have been within 2 Å of the Val111 side chain in the conformation of the cavity used for the docking calculation, a steric conflict that is relieved by conformational expansion of the cavity in the experimental structures. Indeed, for all complexes, with the exception of β -chlorophenetole (1), the F-helix of lysozyme (residues 108-113) that forms one wall of the cavity reorients by about 2 Å and swings Val111 further out of the cavity to accommodate the ligands.³⁹ The protein conformations seen in these structures more

closely resemble the larger isobutylbenzene bound cavity site (PDB id 184L) than the smaller benzene bound cavity site (PDB id 181L) used for docking and rescoring. Whereas the MM-GBSA methods do not capture this helix motion, receptor and ligand minimization reduces the steric clash sufficiently to improve the ranks of what were docking false negatives. Higher level calculations using free energy methods and molecular dynamics have captured the F-helix motion and explained discrepancies in free energies upon ligand binding due to its displacement.^{21; 30}

New L99A/M102Q ligands Predicted by Rescoring.

Ten representative compounds that scored well by the MM-GBSA methods were experimentally tested for binding to the polar cavity (Table 3). These compounds were ranked poorly by docking, again typically because they were too large for the conformation of the cavity targeted by docking. Binding was detected at millimolar concentrations by T_m upshift for six of these ten compounds; for the remaining four binding was not observed (Table 3). We note, however, that for one of these four, 2-(n-propylthio) ethanol (12), we were able to determine a crystal structure in complex with the ligand by soaking a crystal of L99A/M102Q with 100 mM of compound, suggesting that it is a weak ligand for this cavity. AMBERDOCK correctly predicted binding for four of the six ligands that it suggested should bind, while incorrectly predicted binding for o-benzylhydroxylamine (14) and 1-phenylsemicarbazide (3). Of the remaining two hits tested, prioritized by a high PLOP ranking, AMBERDOCK missed one real ligand but correctly distinguished one real decoy, ranking both compounds worse than 500. Two of the prospectively tested molecules were not rescored by AMBERDOCK because docking ranked them worse than 5000. PLOP correctly predicted binding for five of the

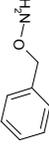
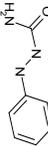
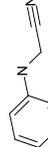
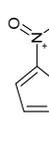
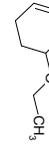
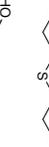
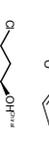
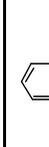
Table 2. Crystallographic measurement and the RMSD values for predicted and crystallographic ligand geometries in the L99A and L99A/M102Q sites.

	L99A Ligands (ID)						L99A/M102Q Ligands (ID)					
Resolution (Å)	1.80 (1.84)	1.64 (1.68)	1.84 (1.89)	1.70 (1.74)	1.46 (1.50)	1.29 (1.32)	1.29 (1.32)	1.47 (1.51)	1.63 (1.68)	1.43 (1.47)	1.56 (1.60)	
Reflections	18414 (1314)	24246 (1780)	13875 (772)	21923 (1615)	33797 (2337)	48474 (3576)	48915 (3563)	33813 (2446)	24034 (1662)	35537 (2423)	28172 (1930)	
R _{merge} (%)	7.0 (50.0)	6.1 (45.4)	7.7 (38.9)	7.4 (63.2)	9.3 (34.8)	8.0 (56.9)	6.7 (62.2)	7.1 (46.6)	6.4 (45.1)	7.6 (37.6)	10.6 (36.1)	
Completeness (%)	99.7 (98.9)	99.8 (99.9)	80.0 (60.6)	99.2 (99.9)	99.5 (94.9)	98.8 (100.0)	99.9 (100.0)	99.8 (99.6)	97.0 (93.1)	99.4 (93.1)	99.3 (93.6)	
$\langle I \rangle / \langle \sigma(I) \rangle$	23.2 (3.4)	22.8 (3.3)	11.8 (2.6)	15.2 (2.4)	13.8 (3.6)	21.0 (3.6)	26.9 (3.3)	24.3 (4.5)	14.5 (2.9)	17.3 (4.2)	14.6 (5.0)	
R-factor (%)	18.7 (27.8)	19.1 (34.9)	19.6 (32.3)	19.1 (32.4)	18.0 (25.6)	17.8 (21.4)	17.2 (23.5)	18.1 (24.6)	20.8 (41.4)	18.1 (23.4)	18.3 (19.6)	
R-free (%)	21.2 (30.2)	22.1 (44.3)	23.3 (34.3)	23.0 (38.9)	21.3 (32.3)	19.1 (21.6)	19.1 (28.6)	20.8 (28.6)	24.0 (55.0)	20.1 (27.1)	20.5 (25.6)	
Absord lengths (Å)	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.02	0.01	0.01	
Absord anisotropy (°)	1.25	1.15	0.99	1.23	1.08	1.07	1.22	1.1	1.58	1.13	1.13	
PDB code	2RAY	2RAZ	2RB0	2RB1	2RB2	2RBN	2RBO	2RBP	2RBQ	2RBR	2RBS	
DOCK					1.42	1.29	2.04/1.12 ^b	0.97	1.44	1.16	1.93/1.84 ^b	
RMSD (Å)	0.82	1.44	0.6	0.62	0.83	0.91	2.00/0.81 ^b	NA	0.87	0.93	NA	
AMBER												
RMSD (Å)	0.64	1.46	0.52	0.28	1.08	1.1	2.00/0.65 ^b	0.63	1.61	1.02	1.80/1.70 ^b	
PLOP												
RMSD (Å)	0.54	1.49	0.58	0.37								

All crystals belong to space group $P3_22_1$

^a Values in parentheses are for the highest resolution shell. ^b Two conformations of the crystallographic ligand were modeled.

Table 3. Compounds predicted by AMBERDOCK and PLOP to bind to T4 Lysozyme L99A/M102Q.

Structure	Compound (ID)	Score and Rank ^a			C ^b	pH	ΔT_m (°C)	Binding detected	Structure determined
		DOCK	AMBER	PLOP					
	O-benzylhydroxylamine (14)	-11.35 (647)	-28.05 (1)	-14.14 (2271)	10	6.8	-0.6	No	No
	1-phenylsemicarbazide (3)	-3.76 (3783)	-26.79 (4)	-16.42 (1354)	10	6.8	0.0°	No	No
	N-phenylglycynitrile (10)	-8.6 (1556)	-25.47 (11)	-40.17 (11)*	<10	3	5.1	Yes	Yes
	2-nitrothiophene (11)	-12.82 (318)	-24.52 (13)	-16.94 (1165)	<10	3	4.4	Yes	Yes
	2-ethoxy-3,4-dihydro-2H-pyran (15)	-7.14 (2215)	-24.21 (14)	-15.18 (1824)	10	3	1.3	Weak	No
	2-(n-propylthio)ethanol (12)	6.02 (6847)	NR ^d	-27.2 (20)	10	3	0.1	No	Yes
	Cis-2-hexen-1-ol (16)	-1.58 (4291)	-10.25 (2260)	-27.19 (21)	10	3	0	No	No
	3-methylbenzylazide (6)	-5.35 (2740)	-20.51 (116)	-25.87 (35)	10	3	1.9	Yes	Yes
	2-phenoxyethanol (9)	-4.08 (3270)	-16.53 (551)	-25.82 (36)	10	3	1.2	Yes	Yes
	(R)-(+)-3-chloro-1-phenyl-1-propanol (13)	3.6 (6074)	NR ^d	-25.65 (37)	10	3	7.8	Yes	Yes

^a Compound scores and ranks (in parenthesis) for DOCK, AMBERDOCK, and PLOP. Scores and ranks in bold font indicate ligands which rank in the top 200 for the respective scoring function. ^b Concentration at which ligand was tested. ^c ΔT_m monitored using fluorescence at $\lambda=291.5\text{nm}$ and measuring the integrated emission above 300 nm. ^d NR is not ranked.

six ligands that it suggested should bind but incorrectly predicted binding for cis-2-hexenol (16). Of the remaining hits tested, prioritized for testing by AMBERDOCK, PLOP missed two true ligands but correctly distinguished two decoys by ranking them worse than 1000.

Crystal structures of six L99A/M102Q ligand complexes were determined to compare predicted and experimental poses of these new ligands (Figure 4). Electron density for each ligand was unambiguous and was detailed enough to suggest two binding modes for 2-nitrothiophene (11) and 3-chloro-1-phenyl-1-propanol (13). Docking predicted the pose of one ligand, 2-(n-propylthio)ethanol (12), to within 1 Å RMSD, while AMBERDOCK further minimized five of its six ligands and PLOP minimized three of its six ligands to within 1 Å RMSD (Table 2). Although the MM-GBSA methods collectively improved the binding mode predictions of all but one ligand, the key hydrogen bond interaction was missed in three of these structures (Figure 4a, e, and f). In addition, the azide group of 3-methylbenzylazide (6) was incorrectly parameterized by both AMBERDOCK and PLOP, as was also observed in the L99A cavity. Neither DOCK nor the MM-GBSA rescoring correctly predicted the binding mode for 3-chloro-1-phenyl-1-propanol (13), with RMSD values of 1.9 and 1.7 Å, respectively. In three structures—2-nitrothiophene (11), 3-methylbenzylazide (6), and 3-chloro-1-phenyl-1-propanol (13)—the F-helix of the cavity moves to accommodate the ligands while keeping the cavity still buried from solvent. In the complexes with 2-(n-propylthio)ethanol (12) and 2-phenoxyethanol (9), there is evidence of a second conformation of residue Phe114 within the cavity that rotates and opens a water channel

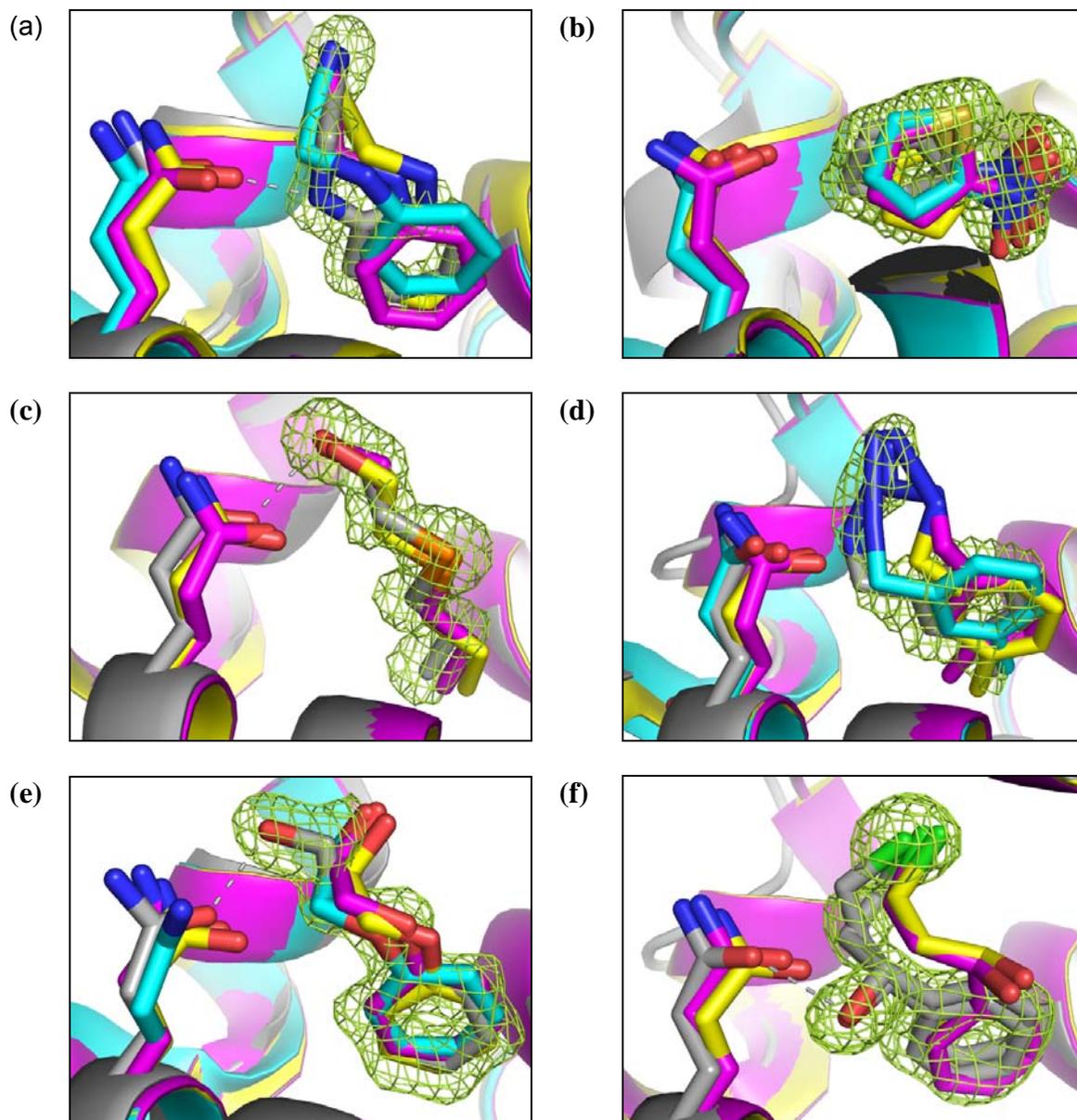


Figure 4. Predicted and experimental ligand orientations for the polar L99A/M102Q cavity site. The carbons of the crystallographic, DOCK, AMBERDOCK, and PLOP predicted poses are colored grey, yellow, cyan, and magenta, respectively. Hydrogen bonds are depicted with dashed lines. The $f_o - f_c$ electron density omit maps (green mesh) are contoured at 2.5-3.0 σ . **(a)** n-phenylglycinonitrile (10), **(b)** 2-nitrothiophene (11), **(c)** 2-(n-propylthio)ethanol (12), **(d)** 3-methylbenzylazide (6), **(e)** 2-phenoxyethanol (9), and **(f)** (R)-(+)-3-chloro-1-phenyl-1-propanol (13) bound to L99A/M102Q. Rendered with the program PyMOL.²⁶

to the surface of the protein. Neither the helix movement nor the Phe114 rotation was sampled by the MM-GBSA methods.

New W191G ligands Predicted by Rescoring.

Fourteen representative compounds reprioritized to score well by the MM-GBSA methods but scored poorly by docking were experimentally tested for binding by measuring perturbation of the heme Soret band in CCP (Table 4).²⁴ Binding was detected for ten of these compounds at concentrations ranging from 50 μ M to 10 mM. Of the eleven compounds that AMBERDOCK predicted to bind with ranks better than 500, binding was detected for eight. Of the remaining prospective hits tested, AMBERDOCK correctly distinguished one compound as a decoy but missed two ligands by ranking them worse than 500. Of the nine compounds that PLOP predicted to bind with ranks better than 500, binding was detected for eight. Of the remaining prospective hits tested, PLOP missed two ligands but correctly distinguished three decoys, ranking them worse than 500.

Crystal structures of CCP in complex with the ten new ligands were obtained (Figure 5). The electron density for the ligands was unambiguous. Docking predicted three structures to within 1 Å of the crystallographic result whereas the MM-GBSA methods did so for seven structures, typically with improved hydrogen bonding interactions (Table 5). For three ligands, the docking poses were over 1.9 Å away from the crystallographic results, and MM-GBSA refinement did little to improve these structures. In four of the complex structures—cyclopentane-carboximidamide (19), 1,2-dimethyl-1H-pyridine-5-amine (22), pyrimidine-2,4,6-triamine (24), and 1-methyl-2-

vinyl-pyridinium (30)—the loop composed of residues 190-195 flips out by nearly 12 Å opening the cavity to bulk solvent. This large loop motion was not sampled by MM-GBSA.

Overall Performance in Predicting Top 100 Hits.

The simplicity of these model cavity sites, the number of known ligands and decoys, and our experience with their ligands^{21; 23; 27; 28; 29} often allow us to predict what turn out to be true ligands and true decoys from among top-scoring molecules, based on their physical properties. We examined the top 100 hits predicted to bind by docking and MM-GBSA, compared property distributions, and made educated guesses as to whether or not they will bind. The 100 top ranking MM-GBSA rescored compounds for the L99A and L99A/M102Q cavities were larger, more flexible, and more polar, with more hydrogen bond acceptors and lower ClogP values per heavy atom compared to the top 100 hits from docking. For the anionic W191G cavity there was a similar trend towards larger molecules and also a drift away from the singly charged cations favored by DOCK, with more dications and neutral molecules prioritized among the top ranking 100 molecules by the MM-GBSA methods. The increased size and greater differences in polarity of the molecules in the MM-GBSA hit lists resulted in lower mean pair-wise similarities among the molecules, and consequently, an increase in the diversity of the rescored hit lists relative to the docking hit lists. Thus, using ECFP_4 fingerprints (SciTegic, Inc.), the average pair-wise Tanimoto coefficient among the 100 top docking molecules for the L99A cavity with DOCK, AMBERDOCK, and PLOP was 0.17, 0.12, and 0.10, respectively (full distributions of pairwise similarities are given in Supplementary Figure S1). Similar trends were observed in the other two cavities. The same tendencies that led

to greater diversity in ligands and their properties, however, reduced the raw hit rates we anticipate from among the top 100 ranking MM-GBSA ligands compared to those predicted by docking (Table 6). For example, among the top 100 docking hits for the CCP cavity there were 29 true ligands and no experimentally determined decoys. Of the remaining molecules—all untested—were what we predict to be 79 likely ligands and 7 likely decoys, based on their similarity to known ligands and decoys and their physical properties such as size and charge complementarity. Conversely, among the top 100 PLOP hits for the anionic cavity were only 15 experimentally tested ligands and 1 experimental decoy. Among the untested molecules were what we suspect are 53 further ligands and 22 further decoys. Among the top AMBERDOCK hits for this cavity were 19 true ligands and 3 experimental decoys. Among the untested molecules prioritized by this program, we suspect that there are 67 further ligands and 14 more decoys. Similar trends were observed in the other two cavities (Table 6). Admittedly, these numbers reflect guesses only, but we suspect that the overall trends would be born out by experiment (the interested reader may draw their own conclusions from the full lists in Supplementary Materials Tables 1-9) Thus, whereas the MM-GBSA methods rescued many docking false negatives and sampled a more diverse chemical space among the top hits, they also suggested more false positives among the very top-scoring molecules and, we suspect, have a lower overall hit-rate in this segment of the molecules prioritized for testing.

Table 4. Compounds predicted to bind by AMBERDOCK and PLOP to CCP W191G.

Structure	Compound (ID)	Score and Rank ^a			C ^b (mM)	Binding detected	Structure determined
		DOCK	AMBER	PLOP			
	N-methyl-1,2-phenylene diamine (17)	-120.6 (618)	347.03 (30)	-38.08 (530)	10	No	No
	N-methylbenzylamine (18)	-18.59 (942)	347.85 (38)	-16.21 (952)	1	Yes	Yes
	Cyclopentane-carboximidamide (19)	-13.38 (2134)	347.86 (39)	-44.39 (389)	1	Yes	Yes
	(1-methyl-1h-pyrrol-2-yl)-methylamine (20)	-14.74 (1830)	348.17 (49)	-31.88 (796)	0.05	Yes	Yes
	5-nitro-6-aminouracil (21)	-12.14 (2435)	348.49 (62)	-31.47 (7827)	1	No	No
	1,2-dimethyl-1h-pyridin-5-amine (22)	-22.95 (362)	349.34 (87)	-54.67 (59)	0.05	Yes	Yes
	2-aminobenzylamine (23)	-12.62 (2316)	349.34 (96)	-34.19 (671)	10	No	No
	Pyrimidine-2,4,6-triamine (24)	-36.54 (7)	344.29 (12)	-59.87 (53)	1	Yes	Yes
	1,3-dimethyl-2-oxo-2,3-dihydro-pyrimidin-1-ium (25)	-8.52 (43093)	363.47 (1901)	-56.65 (32)	10	No	No
	1-methyl-5-imidazolecarboxaldehyde (26)	-21.14 (4291)	358.53 (746)	-57.12 (28)	10	Yes	Yes
	3-methoxypyridine (27)	-23.05 (2665)	355.17 (393)	-55.31 (44)	10	Yes	Yes
	2-imino-4-methylpiperidine (28)	-17.3 (1695)	349.17 (82)	-52.43 (119)	10	Yes	Yes
	2,4,5-trimethyl-3-oxazoline (29)	-13.96 (1962)	355.98 (455)	-52.32 (124)	0.25	Yes	Yes
	1-methyl-2-vinylpyridinium (30)	-15.17 (1716)	363.6 (1938)	-52.32 (125)	0.5	Yes	Yes

^a Compound scores and ranks (in parenthesis) for DOCK, AMBERDOCK, and PLOP. Scores and ranks in bold font indicate ligands which rank in the top 200 for the respective scoring function. ^b Concentration at which ligand was tested.

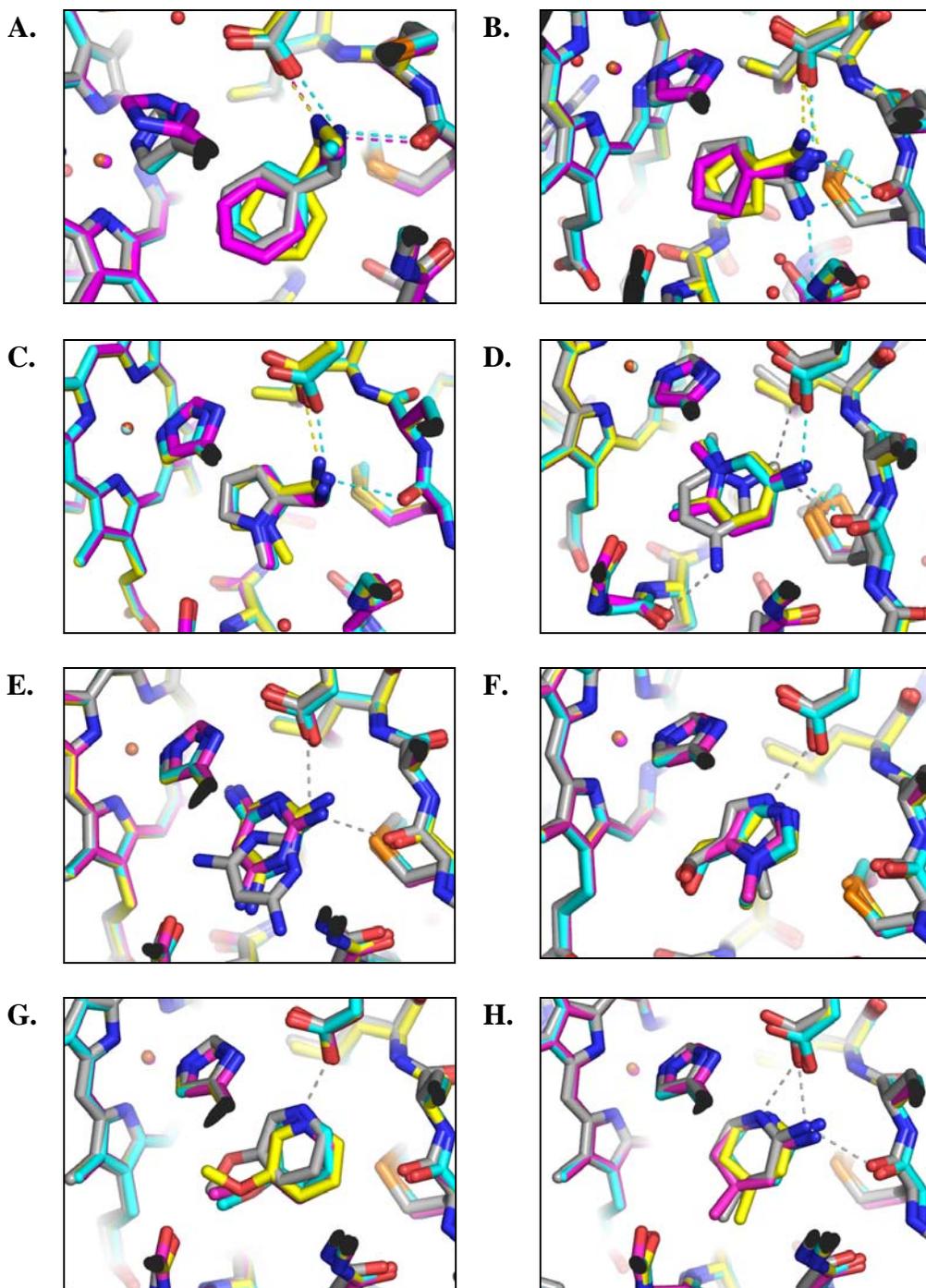


Figure 5. Predicted and experimental ligand orientations for the anionic CCP cavity. The carbons of the crystallographic, the DOCK, AMBERDOCK, and PLOP predicted poses are colored grey, green, cyan, and orange, respectively. (a) *n*-methylbenzylamine (18), (b) cyclopentane carboximidamide (19), (c) (1-methyl-1H-pyrrol-2-yl)-methylamine (20), (d) 1,2-dimethyl-1H-pyridin-5-amine (22), (e) pyrimidine-2,4,6-triamine (24), (f) 1-methyl-

5-imidazolecarboxaldehyde (26), (g) 3-methoxypyridine (27), (h) 2-imino-4-methylpiperidine (28), (i) 2,4,5-trimethyl-3-oxazoline (29), and (k) 1-methyl-2-vinylpyridinium (30). Rendered with the program PyMOL.²⁶

Origins of False Positive Hits Suggested by MM-GBSA Rescoring.

In these simple cavities, false-positive hits often identify specific pathologies in a scoring function. For example, the MM-GBSA methods seemed distracted by compounds bearing what is almost certainly the wrong net charge for the W191G cavity, which extensive testing has shown preferentially binds mono-cations over neutral molecules (few of which have been observed to bind, and then only weakly) and dications (none of which have been observed to bind). For instance, among the top 100 ranking molecules predicted by PLOP, there were 13 dications. Whereas AMBERDOCK predicted only one dication, it prioritized five neutral molecules among the top 100 hits. The dications will pay too high a desolvation penalty to be compensated by the interaction with the single anion in the site (Asp235), and the neutral compounds desolvate the same aspartate without recouping enough in interaction energy. Balancing polar and ionic interactions with concomitant solvation penalties is a challenge for the field, one clearly faced by these methods as well. On the other hand, many of the top ranked PLOP ligands for L99A (47 out of the top 50) and L99A/M102Q contained one or more nitriles. While some of these compounds may well be ligands, as in the case of n-phenylglycinonitrile (10) for L99A/M102Q (Table 3), we suspect that this represents a ligand parameterization problem as opposed to a genuinely meaningful enrichment. Indeed, the PLOP solvation energies for the top 47 nitriles actually rewarded desolvation, rather than

Table 5. Crystallographic measurement and the RMSD values for predicted and crystallographic ligand geometries in the CCP site (compound IDs in parens).

	n-methyl-benzylamine (18)	cyclopentane carboximidamide (19)	(1-methyl-1H-pyrrol-2-yl)-methylamine (20)	1,2-dimethyl-1H-pyridin-5-amine (22)	pyrimidine-2,4,6-triamine (24)	1-methyl-5-imidazole carboxaldehyde (26)	3-methoxy pyridine (27)	2-imino-4-methyl piperidine (28)	2,4,5-trimethyl-3-oxazoline (29)	1-methyl-2-vinylpyridinium (30)
Resolution (Å)	1.24 (1.27)	1.80 (1.85)	1.39 (1.43)	1.50 (1.54)	1.50 (1.54)	1.50 (1.54)	1.80 (1.85)	1.50 (1.54)	2.49 (2.56)	1.50 (1.54)
Reflections	104081 (4744)	36504 (2670)	78561 (5188)	57782 (4371)	63009 (4512)	63445 (4489)	36874 (2587)	63783 (4467)	11042 (680)	63108 (3364)
R _{range} (%)	6.1 (40.7)	5.3 (10.9)	4.5 (20.2)	5.1 (22.4)	5.2 (38.5)	3.8 (22.1)	6.4 (23.3)	5.0 (33.4)	2.7 (5.9)	4.2 (19.6)
Completeness (%)	93.1 (58.0)	99.4 (99.9)	99.0 (90.5)	92.0 (95.9)	99.8 (97.9)	99.7 (96.3)	99.5 (96.2)	99.7 (96.5)	82.3 (69.7)	99.3 (95.4)
<I>/<σ(I)>	29.9 (1.8)	51.7 (30.1)	43.2 (5.6)	41.7 (6.0)	30.6 (2.8)	37.8 (5.2)	34.8 (9.2)	31.0 (3.5)	26.1 (12.9)	39.2 (5.9)
R-factor (%)	12.2 (24.8)	15.9 (19.3)	13.8 (20.7)	14.4 (17.8)	15.2 (21.0)	14.5 (16.9)	15.0 (18.4)	14.3 (17.9)	17.7 (21.9)	14.6 (16.3)
R-free (%)	14.6 (24.4)	19.5 (24.0)	15.4 (24.1)	16.7 (25.5)	17.3 (25.7)	16.4 (21.0)	19.1 (24.0)	16.9 (23.9)	22.9 (29.4)	16.7 (21.7)
Absol lengths (Å)	0.01	0.02	0.01	0.01	0.01	0.01	0.02	0.01	0.02	0.01
Absol angles (°)	1.59	1.39	1.19	1.22	1.14	1.24	1.5	1.22	1.95	1.14
PDB code	2RBT	2RBU	2RBV	2RBW	2RBX	2RBY	2RBY	2RBY	2RC1	2RC2
DOCK										
RMSD (Å)	1.05	1.34	0.58	2.97/3.16 ^b	1.91	1.07	0.89	0.55	1.32	2.58
AMBER										
RMSD (Å)	0.37	0.31	0.34	2.84/3.01 ^b	1.7	1.06	0.81	0.5	0.83	2.55
PLOP										
RMSD (Å)	0.45	0.99	0.3	2.98/3.15 ^b	1.91	0.94	0.77	0.33	0.81	2.59

All crystals belong to space group $P2_1 2_1 2_1$

^a Values in parentheses are for the highest resolution shell. ^b Two conformations of the crystallographic ligand were modeled.

penalizing it as is almost always the case otherwise, suggesting that there is an issue with determining the correct self-energy for this functional group. Not wishing this term to dominate our analysis, we excluded these compounds from the PLOP rescored hit lists for L99A and L99A/M102Q; they do not contribute to the accounting described in this work). This highlights the importance of good ligand parameterization for database screening—which is a considerable challenge for hundreds-of-thousands of molecules typically screened by docking—the lack of which can undermine any improvement in theory.

Table 6.

Likely ligands and decoys among the top 100 ranked ligands by docking and MM-GBSA.

Method	True ligands in top hits	True decoys in top hits	Likely ligands in top 100 hits ^a	Likely decoys in top 100 hits ^b	Ambiguous ^c
L99A cavity					
DOCK	7	3	63	23	14
PLOP	6	1	35	22	43
AMBERDOCK	8	2	54	25	21
L99A/M102Q cavity					
DOCK	13	2	73	12	15
PLOP	5	1	31	8	61
AMBERDOCK	7	2	43	22	35
W191G cavity					
DOCK	29	0	79	7	14
PLOP	15	1	53	22	25
AMBERDOCK	19	3	67	14	19

^a Molecules that, based on their physical properties and similarity to known ligands, are likely to be cavity ligands (a full list is given in Tables S1-S9 in Supplementary Materials). ^b Molecules that, based on their physical properties and similarity to known decoys, are likely not to bind. ^c Molecules that are sufficiently different from known ligands and decoys, and whose physical properties are not sufficiently distinctive, such that no prediction was made (for L99A and L99A/M102Q molecules). For W191G, molecules that were mis-protonated during database preparation relative to the expected

protonation at pH 4.5 are not counted to measure the performance of the scoring function.

2.4 Discussion

In principle, the most important improvements of MM-GBSA over docking, certainly over the program used in this study, DOCK3.5.54, are the better representation of electrostatic interactions, ligand and protein desolvation energies, and relaxation of the ligand-protein complex. The simplicity of the model cavity sites allows us to explore how these terms influence docking results in detail and to make prospective predictions for ligands that we can, in fact, acquire and test. Many investigators will be unsurprised to see that the MM-GBSA methods can rescue molecules that rank poorly in the docking calculation owing to the rigid-receptor approximation used in docking. Ligands that were too big to be accommodated well in the original docking are well-fit by a binding site that has been allowed to relax by energy minimization and, in the case of AMBERDOCK, short MD simulations. This was true both in retrospective calculations as well as in prospective predictions. The ability to relax the site also resulted in rescored hit lists that were more diverse with a wider range of likely ligands. Perhaps less anticipated was the cost of allowing such conformational change—some of the rescued, high-scoring molecules by MM-GBSA do not, in fact, bind to the cavity sites. These molecules are new false-positives introduced by the higher level of theory. Indeed, the overall hit rates at the very top of the ranked lists are arguably better by simple docking than by MM-GBSA rescoring, at least when evaluated simplistically by the raw number of hits and likely hits (this is arguably offset by the greater diversity of the MM-GBSA hit lists). Partly this reflects problems in ligand parameterization, and partly difficulties in the treatment of the electrostatics in the binding sites. The most important challenge for MM-

GBSA and for flexible receptor models in general is balancing the opportunities to find new ligands as receptor geometries are relaxed with the introduction of new false positives as the need to consider large receptor internal energies is introduced. Specific examples of these opportunities and problems are apparent in the three cavity sites studied here.

The principal improvement conferred by MM-GBSA rescoring in the model cavity sites over docking was the inclusion of receptor binding site relaxation, which improved the ranks of larger ligands that rigid receptor docking missed. AMBERDOCK, for example, correctly predicted 2-ethoxyphenol (5) to bind to L99A (Table 1, Figure 3d). This compound is too large for the unrelaxed conformation of this cavity targeted by docking, but minimization and MD simulations allow the ligand to be well accommodated by effectively expanding the site. Often, this relaxation led not only to improved rankings but also improved geometries. For many ligands, RMSD values between the MM-GBSA predictions and the crystallographic results declined relative to those of the docking predictions and, especially in the W191G anionic cavity, many ligands refined by MM-GBSA had improved hydrogen bonding to the site. Examples of this include the new W191G cavity ligands n-methylbenzylamine (18) and cyclopentane-carboximidamide (19) (Table 4, Figures 5a and 5b, respectively).

The structural relaxation with MM-GBSA performed well when the initial docking geometry resembled the crystallographic pose, but did little when large protein conformational changes were provoked by ligand binding. For instance, F-helix unwinding and rotamer change by Val111 in L99A and L99A/M102Q were never captured by the method, nor was the extensive loop flipping observed in several of the W191G-ligand complexes. When such movements occurred, MM-GBSA rescoring could not rescue substantially incorrect docking poses, such as that adopted by 3-chloro-1-

phenyl-propanol (13) for L99A/M102Q (Table 3, Figure 4f) and pyrimidine-2,4,6-triamine (24) predicted for CCP (Table 4, Figure 5e), notwithstanding the large improvement in their rankings conferred by the rescoring. These large movements are outside the radius of convergence of the local relaxation undertaken by the MM-GBSA methods. Indeed, even more time-consuming thermodynamic integration methods are hard put to sample such changes without explicit “confine-and-release” strategies, which depend on a foreknowledge that such movements are likely.⁴⁰ And whereas loop sampling methods have had encouraging successes in predicting such large movements,⁴¹ this remains a frontier challenge for ligand and structure prediction methods.

Pragmatically, the inability to predict the structural accommodations provoked by some large ligands is offset by the correct re-prioritization of what were docking false-negatives as ligands. The same comfort is not afforded by the ten false negatives introduced by the MM-GBSA methods, nor by the lower overall hit rates compared to docking among the very top scoring ligands (Table 6). By allowing the receptor to respond to ligand binding, one allows for new and potentially unfavorable receptor conformations. These must be distinguished by the MM-GBSA energy functions from the true low-energy conformations that may be sampled in solution. This is challenging as the receptor conformational energies are large, and the errors in these calculations are typically on the same order of the net interaction energy of the protein-ligand complex. Although some of the errors are cancelled by subtraction of the internal energies before and after ligand binding, one is still subtracting two large numbers with relatively large errors to find a small one, the net binding free energy. Consistent with this view, ligands achieved their maximal advantage over decoys on rescoring when we allowed only a 5 Å region around the binding site to relax. Allowing the full protein to relax, or even an 8 Å

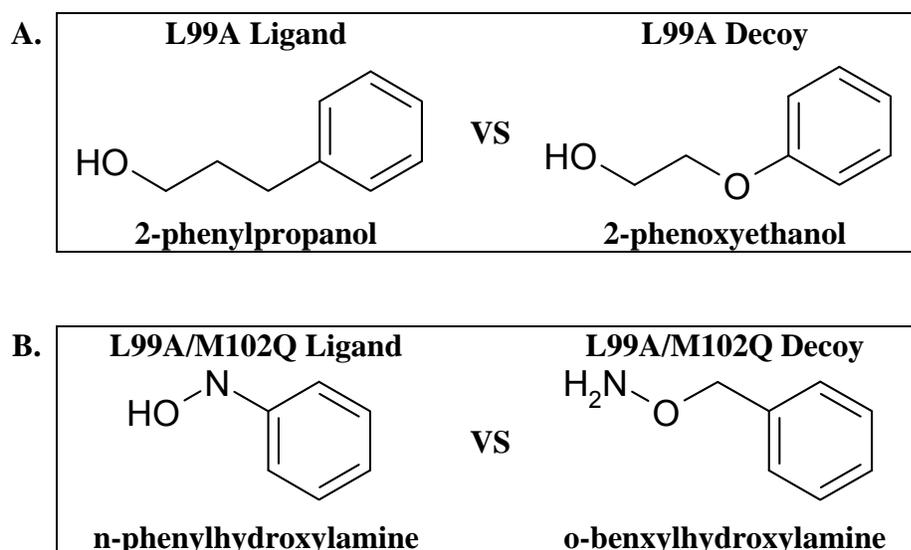


Figure 6. The topologically similar ligands and decoys of A. 2-phenylpropanol and 2-phenoxyethanol (9) for L99A and B. n-phenylhydroxylamine and o-benzyloxyamine (14) to L99A/M102Q.

region around the binding site, diminished the discrimination of known ligands from decoys. Of course, relaxing the entire system is the more physically correct way to calculate these energies. Falling back on limited relaxation speaks to a larger methodological issue.

The three cavity sites targeted here are contrivances of human design and ligands discovered for them have no intrinsic value other than for testing methods. Indeed, in these simple model systems the failures are often more interesting than the successes, as they can illuminate a specific methodological problem.^{21; 23; 27; 28; 29} Examples are the ten false positives predicted for the cavity sites by MM-GBSA rescoring. Some of these reflect ligand parameterization problems. For instance, we suspect that the many nitrile containing decoys predicted by PLOP for L99A and L99A/M102Q reflect failures in ligand parameterization. Such mechanical failures may be addressed by close attention to particular ligand groups and improved partial atomic charge models; admittedly, this can

be a daunting task for screening databases containing hundreds of thousands of disparate molecules. More interesting are the eight false-positives that are true energy function decoys. Several of these highlight difficulties in the treatment of electrostatics and solvation in the binding sites. 2-phenoxyethanol (9), for example, was predicted to bind by both PLOP and AMBERDOCK to L99A (Table 1). This decoy has a similar topology to 2-phenylpropanol, a known ligand²² (Figure 6a); however, the ether of 2-phenoxyethanol (9) increases its polarity and presumably its solvation energy, which is not fully captured by the MM-GBSA implicit solvent model (another possibility would be that the 2-phenoxyethanol is docked in a high-energy conformation, one that is not recognized by the rescoring methods, but this turns out not be the case, with both the decoy and the ligand 2-phenylpropanol adopting similar and low energy conformations). Similarly, o-benzylhydroxylamine (14) was the top-ranking AMBERDOCK hit for L99A/M102Q, but is a decoy (Table 3). The terminal -ONH₂ of this compound is too polar for the site, stranding one unpaired polar hydrogen from the NH₂ group in this largely hydrophobic site. Interestingly, the polar cavity does bind n-phenylhydroxylamine (unpublished data), which has the same hydrogen bond accounting as o-benzylhydroxylamine (14) and topologically resembles it closely (Figure 6b). The difference between these nearly identical molecules is that in the former the two hydrogen bond donors from the ligand can both be accommodated by the carbonyl of the receptor glutamine, whereas in the decoy both hydrogen bond donors originate from the same atom—the nitrogen of the o-benzylhydroxylamine (14)—and only one can be accommodated by the carbonyl oxygen.

The challenges of balancing ligand electrostatic interaction energies and desolvation penalties were also apparent in the anionic, W191G cavity. Most obvious were those molecules that did not bear the correct mono-cationic charge state. The 13

molecules that were doubly charged among the top scoring PLOP hits are almost certainly decoys, and this is also the case for the AMBERDOCK false positive 5-nitro-6-aminouracil (21), which is neutral and cannot make the ion-pair interaction with Asp235 (Table 4). More subtly, whereas 1,3-dimethyl-2-oxo-2,3-dihydropyrimidin-1-ium (25) is charged, this charge is shared between the two cyclic nitrogens and results in a compound with reduced electrophilicity compared to a compound with a localized charge. The AMBERDOCK false-positives n-methyl-1,2-phenylene-diamine (17) and 2-aminobenzylamine (23) (Table 4) most likely do not bind because of steric clashes that inhibit optimal positioning of the charge-charge interaction. These failures point to specific directions for improved treatment of the balance between electrostatic interaction and desolvation energies in the MM-GBSA methods.

Overall, the results of MM-GBSA rescoring of docking hit lists on the model binding sites seem conflicted. On the one hand, rescoring rescued many docking false negatives, improved the geometric fidelity of most of the predicted structures, and increased the diversity of the hit lists. On the other hand, rescoring introduced more false-positives, especially among the very top ranking ligands, compared to the simpler docking protocol. These observations may be reconciled by recognizing that what is probably the greatest advantage of the MM-GBSA methods over docking for the model sites, the relaxation of the protein-ligand complex, also presents the greatest challenge to discrimination. To allow a flexible receptor, one must consider the relative energies of the different protein conformations explored. This implicates the pair-wise interactions of thousands of protein atoms, as opposed to the tens of atoms involved in the immediate protein-ligand complex. To properly rank the energies of the complexes, one must also properly account for the larger uncertainties that accompany the much higher magnitude energies of the overall system. Whereas this is the thermodynamically correct approach,

it introduces many interactions that have little bearing on the intimacies of the protein-ligand complex itself. Rigid receptor docking, for all the calumny poured upon it, can ignore these large magnitude yet low relevance interactions. Of course, this leads to many false-negatives, but it avoids many of the false positives to which the MM-GBSA methods are prone. Pragmatically, this suggests that hits derived from docking to a rigid experimental receptor conformation—and ideally more than one^{29; 42}—and hits prioritized by rescoring after MM-GBSA refinement with binding site minimization will provide good candidates for experimental testing. Despite its greater sophistication, MM-GBSA rescoring has a harder task, and its predictions will not, by every criterion, be better than those of a modern docking program; rather, our results suggest they will complement and add to them. Still, MM-GBSA is a higher level of theory, and because it is grounded in physics, they can be built upon and improved in a regular way. They are thus on a path to fundamental improvement in molecular docking and structure-based screening, which is so actively sought.⁴³

2.5 Materials and Methods

Docking against Cavity Sites

DOCK3.5.54^{23; 37} was used to dock a multi-conformer database of small molecules into the model cavity sites. The receptors, grids, spheres, and ligand databases were prepared as described for the T4 Lysozyme²³ and CCP²⁷ cavities, respectively. Briefly, to sample ligand orientations, ligand, receptor, and overlap bins were set to 0.2 Å; the distance tolerance for matching ligand atoms to receptor was set to 0.75 Å. Each docking pose was evaluated for steric fit. Compounds passing this filter were scored for electrostatic and van der Waals complementarity and assigned the full penalty for transfer

from a dielectric of 80 to one of 2, as calculated by AMSOL.^{44; 45} Sampling and scoring required less than a second per ligand on a single 3.2 GHz Xeon processor. The best scoring conformation of each of the 10,000 top scoring molecules against L99A and L99A/M102Q and the 5400 top scoring molecules against CCP were saved and rescored by the MM-GBSA protocols.

Rescoring with PLOP

The rescoring procedure with Protein Local Optimization Program (PLOP)^{35; 36} was essentially as described.¹⁷ Ligand parameters were calculated with IMPACT.⁴⁶ The partial atomic charges of the ligands were replaced by the AM1-CM2 charges calculated by AMSOL (v6.5.3) as these were the same charges used during the initial docking.²³ The same protein structure file used in docking was used for rescoring. Protein parameters were defined by IMPACT with the exception of the partial charges for the heme cofactor in CCP W191G, which were the same as used in the docking method.²⁷ All energy minimizations were performed using PLOP with the all-atom OPLS force field (OPLS-AA)⁴⁷ and the Surface Generalized Born (SGB) implicit solvent model.⁴⁸ PLOP implements a multiscale truncated-Newton (MSTN) minimization algorithm as described.⁴⁹ For receptor minimization and calculation of $E_{complex}$ and $E_{receptor}$, a pre-specified list of residues within 5 Å of the binding site were minimized after an initial sidechain rotamer search. (Residues 78, 84, 85, 87, 88, 91, 98-100, 102, 103, 106, 111, 118, 121, 133, and 153 for L99A and L99A/M102Q and residues 174-180, 189-192, 202, 230-232, 235, and water 308 for CCP). The rotamer search algorithm is as described in supplementary materials.

Preliminary PLOP calculations of the hydrophobic and polar cavities were performed with a rigid receptor and resulted in very little separation of ligands and

known decoys. On the other hand, PLOP calculations in which a larger set of residues (those within 8 Å of the binding site) were minimized and resulted in worse overall enrichments of known ligands and a decreased separation of known ligands and decoys relative to minimizing a smaller 5 Å pocket. To approximate a fully desolvated ligand and cavity for the hydrophobic L99A and polar L99A/M102Q sites, only the SGB solvation term of the free ligand was included in the calculation of the total PLOP binding energies. Initial PLOP calculations including the SGB solvation terms for the calculation of the complex and free protein energies resulted in poor enrichments of known ligands, decreased separation of ligands and known decoys, as well as an enrichment of hits with increased polarity and electrostatic interactions. For the more solvated CCP cavity, the SGB terms were included in the calculation of the complex, free protein, and free ligand energies for the total binding energy.

Rescoring with AMBERDOCK

AMBERDOCK is based on the `amber_score()` scoring module in DOCK6. The ligand structures were modified using the *antechamber* suite of programs to create input files that could be read by Leap to generate the parameter and topology files for AMBERDOCK. Antechamber⁵⁰ has been developed to be used with the general AMBER force field (GAFF) for small molecules.⁵¹ Charges for the ligands were generated using three charge methods in Antechamber—PEOE,⁵² AM1-BCC,⁵³ and HF/6-31G* RESP.⁵⁴ The protonation states of the ligands were kept the same as the previous docking run for consistency in rescoring. AMBER *ff94* parameters were assigned to all the protein atoms. The standard parameters for the heme cofactor as implemented in the Amber 9 program was used for the CCP cavity.⁵⁵ The protonation states of Histidine residues were predicted based on their close neighbors. The GB model corresponding to `igb=5` in the

AMBER 9 program was used.⁵⁶ The surface area term was calculated using the LCPO model.⁵⁷ A non-bonded cutoff of 18 Å was used for the calculations.

The starting structures were taken from the docked pose. The structures were subjected to 100 steps of conjugate gradient minimization, 3000 steps of MD simulation with a 1 fs time step at a temperature of 300K, followed by 100 steps of minimization. During the minimization and MD, only the ligand and the protein residues within 5 Å of the ligand were allowed to move. To expedite the scoring process, we calculated the energy of the receptor ($E_{receptor}$) once, and used this energy as a constant term during the subsequent energy evaluations for the rest of the ligands in the database. Binding free energy calculations with AMBERDOCK follows a scheme as described in supplementary materials (Supplementary Figure S2). Several AMBERDOCK rescoring protocols with slight variations were retrospectively tested and results are described in supplementary materials (Supplementary Figure S3).

Protein Preparation and Expression

T4 Lysozyme mutants L99A and L99A/M102Q and CCP mutant W191G were expressed and purified as described.^{23; 24}

Binding Detection of Ligands to T4 Lysozyme Cavities by Upshift of Thermal Denaturation Temperature

To detect binding, L99A and L99A/M102Q were denatured reversibly by temperature in the presence and absence of the putative ligand. Molecules that bind preferentially to the folded cavity-containing protein should stabilize it relative to the *apo* protein, raising its temperature of melting.²² All thermal melts were conducted in a Jasco J-715 spectropolarimeter as described.²² Each compound was screened in its neutral

form. All compounds tested against L99A and L99A/M102Q were assayed in a pH 3 buffer containing 25 mM KCl, 2.9 mM phosphoric acid, and 17 mM KH_2PO_4 with the exception of 1-phenylsemicarbazide (3) and o-benzylhydroxylamine (14). To maintain compound neutrality, these two were assayed at pH 6.8 in a 50 mM potassium chloride and 38% (v/v) ethylene glycol buffer.²² Thermal melts were monitored by far UV circular dichroism, except for melts in the presence of 4-(methylthio)nitrobenzene (2), 1-phenylsemicarbazide (3), and 2,6-difluorobenzylbromide (4), which absorb strongly in the far UV region. For these three, thermal denaturation was measured by the intensity of the integrated fluorescence emission for all wavelengths above 300 nm, exciting at 283 to 292 nm, using a fluorescence PMT on the Jasco instrument. Thermal melts were performed at a temperature ramp rate of 2 K/min. A least-squares fit of the two-state transition model was performed with the program EXAM⁵⁸ to calculate T_m and van't Hoff ΔH values for the thermal denaturations. The ΔC_p was set to 8 $\text{KJ mol}^{-1} \text{K}^{-1}$ (1.94 $\text{kcal mol}^{-1} \text{K}^{-1}$).

Binding Detection of Ligands to CCP W191G

Ligand binding was measured in 50 mM acetate buffer pH 4.5. To avoid competition in ligand binding with small cations like potassium,²⁴ the pH of the buffer was adjusted with Bis-Tris propane. The compounds were dissolved in dimethyl sulfoxide (DMSO). Binding of compounds to CCP was monitored by the red shift and increase of absorbance of the heme Soret band²⁴ at 10 °C.

Structure Determination

Crystals for L99A and L99A/M102Q were grown as described²³ and the resulting crystals belonged to space group $P3_22_1$. Crystals were soaked overnight to one week in

crystallization buffer containing as much as 100 mM compound. In addition to soaking, drops of neat compound were added to the cover slip surrounding the drop containing the crystal. After soaking, the crystals were cryoprotected with a 50:50 Paraton-N (Hampton Research, Aliso Viejo, CA), mineral oil mix. Crystals for CCP W191G were grown as described²⁵ and the resulting crystals belonged to space group $P2_12_12_1$. Crystals were soaked in 25% MPD with 1 to 50 mM compound for 4 hours or overnight with the exception of pyrimidine-2,4,6-triamine (24), which was soaked for 15 minutes.

Diffraction data for the complexes of L99A with β -chlorophenetole (1), 4-(methylthio)nitrobenzene (2), 2,6-difluorobenzylbromide (4) and the complex of L99A/M102Q with 3-methylbenzylazide (6) were collected using a Rigaku X-ray generator equipped with a rotating copper anode and a Raxis IV image plate. Data for the complexes of L99A/M102Q with n-phenylglycinonitrile (10) and 2-nitrothiophene (11) and the complex of CCP with n-methylbenzylamine (18) were collected on Beamline 9-1 at the Stanford Synchrotron Radiation Laboratory (SSRL) using an ADSC-CDD detector. Data for all other complexes were collected on Beamline 8.3.1 of the Advanced Light Source (ALS) at Lawrence Berkeley National Laboratory using an ADSC-CCD detector. All data sets were collected at 100 K. Reflections were indexed, integrated, and scaled using HKL2000.⁵⁹ Parameters for ligands were generated with PRODRG.⁶⁰ Complexes were refined using the CCP4 software package.⁶¹ Interactive model building was performed using Coot.⁶²

Protein Data Bank Accession Codes

The crystallographic coordinates for the complex structures presented in this work have been deposited with the RCSB Protein Data Bank with accession codes 2RAY,

2RAZ, 2RB0, 2RB1, 2RB2, 2RBN, 2RBO, 2RBP, 2RBQ, 2RBR, 2RBS, 2RBT, 2RBU, 2RBV, 2RBW, 2RBX, 2RBY, 2RBZ, 2RC0, 2RC1, and 2RC2.

2.6 Acknowledgments

This work supported by GM59957 (to BKS), AI035707 (to MPJ) and GM56531 (to DAC). MPJ is a consultant to Schrodinger Inc. We thank Niu Huang for advice using PLOP, Michael Keiser and Jerome Hert for help with chemical similarity calculations, Michael Mysinger, Veena Thomas and Michael Keiser for reading this manuscript, and MDL for providing the ACD database.

2.7 Supporting Information Available

A description of the PLOP side chain rotamer search and minimization algorithm, AMBERDOCK parameters and optimization, and structures for the top 100 hits predicted by DOCK, PLOP, and AMBERDOCK for the three cavity sites.

2.8 References

1. Powers, R. A., Morandi, F. & Shoichet, B. K. (2002). Structure-Based Discovery of a Novel, Noncovalent Inhibitor of AmpC beta-Lactamase. *Structure* **10**, 1013-1023.
2. Evers, A. & Klebe, G. (2004). Successful Virtual Screening for a Submicromolar Antagonist of the Neurokinin-1 Receptor Based on a Ligand-Supported Homology Model. *J. Med. Chem.* **47**, 5381-5392.
3. Grzybowski, B. A., Ishchenko, A. V., Kim, C.-Y., Topalov, G., Chapman, R., Christianson, D. W., Whitesides, G. M. & Shakhnovich, E. I. (2002). Combinatorial computational method gives new picomolar ligands for a known enzyme. *PNAS* **99**, 1270-1273.
4. Li, S., Gao, J., Satoh, T., Friedman, Thea M., Edling, Andrea E., Koch, U., Choksi, S., Han, X., Korngold, R. & Huang, Z. (1997). A computer screening approach to immunoglobulin superfamily structures and interactions: Discovery of small non-peptidic CD4 inhibitors as novel immunotherapeutics. *PNAS* **94**, 73-78.
5. Huang, N., Nagarsekar, A., Xia, G., Hayashi, J. & MacKerell, A. D., Jr. (2004). Identification of non-phosphate-containing small molecular weight inhibitors of the tyrosine kinase p56 Lck SH2 domain via in silico screening against the pY + 3 binding site. *J Med Chem* **47**, 3502-11.
6. Song, H., Wang, R., Wang, S. & Lin, J. (2005). A low-molecular-weight compound discovered through virtual database screening inhibits Stat3 function in breast cancer cells. *PNAS* **102**, 4700-4705.

7. Lyne, P. D., Kenny, P. W., Cosgrove, D. A., Deng, C., Zabudoff, S., Wendoloski, J. J. & Ashwell, S. (2004). Identification of compounds with nanomolar binding affinity for checkpoint kinase-1 using knowledge-based virtual screening. *J Med Chem* **47**, 1962-8.
8. Cavasotto, C. N., Oritz, M. A., Abagyan, R. A. & Piedrafita, F. J. (2006). In silico identification of novel EGFR inhibitors with antiproliferative activity against cancer cells. *Biorganic and Medicinal Chemistry Letters* **16**, 1969-1974.
9. Charifson, P. S., Corkery, J. J., Murcko, M. A. & Walters, W. P. (1999). Consensus scoring: A method for obtaining improved hit rates from docking databases of three-dimensional structures into proteins. *Journal of Medicinal Chemistry* **42**, 5100-5109.
10. Bissantz, C., Folkers, G. & Rognan, D. (2000). Protein-based virtual screening of chemical databases. 1. Evaluation of different docking/scoring combinations. *Journal of Medicinal Chemistry* **43**, 4759-4767.
11. Gohlke, H. & Klebe, G. (2001). Statistical potentials and scoring functions applied to protein-ligand binding. *Curr Opin Struct Biol* **11**, 231-5.
12. Stahl, M. & Rarey, M. (2001). Detailed analysis of scoring functions for virtual screening. *J Med Chem* **44**, 1035-42.
13. Wang, R. X. & Wang, S. M. (2001). How does consensus scoring work for virtual library screening? An idealized computer experiment. *Journal of Chemical Information and Computer Sciences* **41**, 1422-1426.
14. Friesner, R. A., Murphy, R. B., Repasky, M. P., Frye, L. L., Greenwood, J. R., Halgren, T. A., Sanschagrin, P. C. & Mainz, D. T. (2006). Extra Precision Glide: Docking and Scoring Incorporating a Model of Hydrophobic Enclosure for Protein-Ligand Complexes. *J. Med. Chem.* **49**, 6177-6196.

15. Wang, J., Kang, X., Kuntz, I. D. & Kollman, P. A. (2005). Hierarchical database screenings for HIV-1 reverse transcriptase using a pharmacophore model, rigid docking, solvation docking, and MM-PB/SA. *J Med Chem* **48**, 2432-44.
16. Kalyanaraman, C., Bernacki, K. & Jacobson, M. P. (2005). Virtual screening against highly charged active sites: identifying substrates of alpha-beta barrel enzymes. *Biochemistry* **44**, 2059-71.
17. Huang, N., Kalyanaraman, C., Irwin, J. J. & Jacobson, M. P. (2006). Physics-Based Scoring of Protein-Ligand Complexes: Enrichment of Known Inhibitors in Large-Scale Virtual Screening. *J. Chem. Inf. Model.* **46**, 243-253.
18. Emanuele Perola. (2006). Minimizing false positives in kinase virtual screens. *Proteins: Structure, Function, and Bioinformatics* **64**, 422-435.
19. Lee, M. R. & Sun, Y. (2007). Improving Docking Accuracy through Molecular Mechanics Generalized Born Optimization and Scoring. *J. Chem. Theory Comput.* **3**, 1106-1119.
20. Lyne, P. D., Lamb, M. L. & Saeh, J. C. (2006). Accurate Prediction of the Relative Potencies of Members of a Series of Kinase Inhibitors Using Molecular Docking and MM-GBSA Scoring. *J. Med. Chem.* **49**, 4805-4808.
21. Mobley, D. L., Graves, A. P., Chodera, J. D., McReynolds, A. C., Shoichet, B. K. & Dill, K. A. (2007). Predicting Absolute Ligand Binding Free Energies to a Simple Model Site. *J Mol Biol* **371**, 1118-1134.
22. Morton, A., Baase, W. A. & Matthews, B. W. (1995). Energetic origins of specificity of ligand binding in an interior nonpolar cavity of T4 lysozyme. *Biochemistry* **34**, 8564-75.

23. Wei, B. Q., Baase, W. A., Weaver, L. H., Matthews, B. W. & Shoichet, B. K. (2002). A model binding site for testing scoring functions in molecular docking. *J Mol Biol* **322**, 339-355.
24. Fitzgerald, M. M., Churchill, M. J., McRee, D. E. & Goodin, D. B. (1994). Small molecule binding to an artificially created cavity at the active site of cytochrome c peroxidase. *Biochemistry* **33**, 3807-18.
25. Musah, R. A., Jensen, G. M., Bunte, S. W., Rosenfeld, R. J. & Goodin, D. B. (2002). Artificial protein cavities as specific ligand-binding templates: characterization of an engineered heterocyclic cation-binding site that preserves the evolved specificity of the parent protein. *J Mol Biol* **315**, 845-57.
26. Chang, C. E. & Gilson, M. K. (2004). Free energy, entropy, and induced fit in host-guest recognition: calculations with the second-generation mining minima algorithm. *J Am Chem Soc* **126**, 13156-64.
27. Brenk, R., Vetter, S. W., Boyce, S. E., Goodin, D. B. & Shoichet, B. K. (2006). Probing molecular docking in a charged model binding site. *J Mol Biol* **357**, 1449-70.
28. Graves, A. P., Brenk, R. & Shoichet, B. K. (2005). Decoys for docking. *J Med Chem* **48**, 3714-28.
29. Wei, B. Q., Weaver, L. H., Ferrari, A. M., Matthews, B. W. & Shoichet, B. K. (2004). Testing a Flexible-receptor Docking Algorithm in a Model Binding Site. *J Mol Biol* **337**, 1161-82.
30. Deng, Y. & Roux, B. (2006). Calculation of Standard Binding Free Energies: Aromatic Molecules in the T4 Lysozyme L99A Mutant. *J. Chem. Theory Comput.* **2**, 1255-1273.

31. Boresch, S., Tettinger, F., Leitgeb, M. & Karplus, M. (2003). Absolute Binding Free Energies: A Quantitative Approach for Their Calculation. *J. Phys. Chem. B* **107**, 9535-9551.
32. Machicado, C., Lopez-Llano, J., Cuesta-Lopez, S., Bueno, M. & Sancho, J. (2005). Design of ligand binding to an engineered protein cavity using virtual screening and thermal up-shift evaluation. *Journal of Computer-Aided Molecular Design* **19**, 421-443.
33. Rosenfeld, R., Goodsell, D., Musah, R., Garret, M., Goodin, D. & Olson, A. (2003). Automated docking of ligands to an artificial active site: augmenting crystallographic analysis with computer modeling. *J Comput Aided Mol Des.*
34. Elowe, N. H., Blanchard, J. E., Cechetto, J. D. & Brown, E. D. (2005). Experimental Screening of Dihydrofolate Reductase Yields a "Test Set" of 50,000 Small Molecules for a Computational Data-Mining and Docking Competition. *J Biomol Screen* **10**, 653-657.
35. Jacobson, M. P., Pincus, D. L., Rapp, C. S., Day, T. J., Honig, B., Shaw, D. E. & Friesner, R. A. (2004). A hierarchical approach to all-atom protein loop prediction. *Proteins* **55**, 351-67.
36. Jacobson, M. P., Kaminski, G. A., Friesner, R. A. & Rapp, C. S. (2002). Force Field Validation Using Protein Side Chain Prediction. *J. Phys. Chem. B* **106**, 11673-11680.
37. Lorber, D. M. & Shoichet, B. K. (1998). Flexible ligand docking using conformational ensembles. *Protein Sci* **7**, 938-50.
38. Irwin, J. J. & Shoichet, B. K. (2005). ZINC--a free database of commercially available compounds for virtual screening. *J Chem Inf Model* **45**, 177-82.

39. Morton, A. & Matthews, B. W. (1995). Specificity of ligand binding in a buried nonpolar cavity of T4 lysozyme: linkage of dynamics and structural plasticity. *Biochemistry* **34**, 8576-88.
40. Mobley, D. L., Chodera, J. D. & Dill, K. A. (2007). Confine-and-Release Method: Obtaining Correct Binding Free Energies in the Presence of Protein Conformational Change. *J. Chem. Theory Comput.* **3**, 1231-1235.
41. Song, L., Kalyanaraman, C., Fedorov, A. A., Fedorov, E. V., Glasner, M. E., Brown, S., Imker, H. J., Babbitt, P. C., Almo, S. C., Jacobson, M. P. & Gerlt, J. A. (2007). Prediction and assignment of function for a divergent N-succinyl amino acid racemase. *Nature Chemical Biology* **3**, 486-491.
42. Claussen, H., Buning, C., Rarey, M. & Lengauer, T. (2001). FlexE: efficient molecular docking considering protein structure variations. *J Mol Biol* **308**, 377-95.
43. Leach, A. R., Shoichet, B. K. & Peishoff, C. E. (2006). Prediction of Protein-Ligand Interactions. Docking and Scoring: Successes and Gaps. *J. Med. Chem.* **49**, 5851-5855.
44. Li, J. B., Zhu, T. H., Cramer, C. J. & Truhlar, D. G. (1998). New class IV charge model for extracting accurate partial charges from wave functions. *Journal of Physical Chemistry A* **102**, 1820-1831.
45. Chambers, C. C., Hawkins, G. D., Cramer, C. J. & Truhlar, D. G. (1996). Model for aqueous solvation based on class IV atomic charges and first solvation shell effects. *Journal of Physical Chemistry* **100**, 16385-16398.
46. Banks, J. L., Beard, H. S., Cao, Y., Cho, A. E., Damm, W., Farid, R., Felts, A. K., Halgren, T. A., Mainz, D. T., Maple, J. R., Murphy, R., Philipp, D. M., Repasky, M. P., Zhang, L. Y., Berne, B. J., Friesner, R. A., Gallicchio, E. & Levy, R. M.

- (2005). Integrated Modeling Program, Applied Chemical Theory (IMPACT). *J Comput Chem* **26**, 1752-80.
47. Kaminski, G. A., Friesner, R. A., Tirado-Rives, J. & Jorgensen, W. L. (2001). Evaluation and Reparametrization of the OPLS-AA Force Field for Proteins via Comparison with Accurate Quantum Chemical Calculations on Peptides. *J. Phys. Chem. B* **105**, 6474-6487.
48. Gallicchio, E., Zhang, L. Y. & Levy, R. M. (2002). The SGB/NP hydration free energy model based on the surface generalized born solvent reaction field and novel nonpolar hydration free energy estimators. *Journal of Computational Chemistry* **23**, 517-529.
49. Zhu, K., Shirts, M. R., Friesner, R. A. & Jacobson, M. P. (2007). Multiscale Optimization of a Truncated Newton Minimization Algorithm and Application to Proteins and Protein-Ligand Complexes. *J. Chem. Theory Comput.* **3**, 640-648.
50. Wang, J., Wang, W., Kollman, P. A. & Case, D. A. (2006). Automatic atom type and bond type perception in molecular mechanical calculations. *J Mol Graph Model* **25**, 247-260.
51. Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A. & Case, D. A. (2004). Development and testing of a general amber force field. *J Comput Chem* **25**, 1157-74.
52. Gasteiger, J. & Marsili, M. (1980). Iterative partial equalization of orbital electronegativity: a rapid access to atomic charges. *Tetrahedron* **36**, 3219-3288.
53. Jakalian, A., Jack, D. B. & Bayly, C. I. (2002). Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. *Journal of Computational Chemistry* **23**, 1623-1641.

54. Bayly, C. I., Cieplak, P., Cornell, W. D. & Kollman, P. A. (1993). A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges - the Resp Model. *Journal of Physical Chemistry* **97**, 10269-10280.
55. Giammona, D. A. (1984). An examination of conformational flexibility in porphyrins and bulky-ligand binding in myoglobin. Ph.D. Dissertation, University of California, Davis.
56. Onufriev, A., Bashford, D. & Case, D. A. (2004). Exploring protein native states and large-scale conformational changes with a modified generalized born model. *Proteins: Structure, Function, and Bioinformatics* **55**, 383-394.
57. Weiser, J., Shenkin, P. S. & Still, W. C. (1999). Approximate atomic surfaces from linear combinations of pairwise overlaps (LCPO). *Journal of Computational Chemistry* **20**, 217-230.
58. Kirchhoff, W. (1993). EXAM: A two-state thermodynamic analysis program, NIST, Gaithersburg, Maryland.
59. Otwinowski, Z. & Minor, W. (1997). Processing of X-ray Diffraction Data Collected in Oscillation Mode. In *Methods in Enzymology* (Carter, C. W. J. & Sweet, R. M., eds.), Vol. 276, pp. 307-326. Academic Press, New York.
60. Schuttelkopf, A. W. & van Aalten, D. M. F. (2004). PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallographica Section D-Biological Crystallography* **60**, 1355-1363.
61. Bailey, S. (1994). The Ccp4 Suite - Programs for Protein Crystallography. *Acta Crystallogr D* **50**, 760-763.
62. Emsley, P. & Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126-32.

63. DeLano, W. L. (2002). The PyMOL Molecular Graphics System. DeLano Scientific, San Carlos, CA, USA.

Gloss to chapter 3

In the previous chapters, my work dealt primarily with the CCP W191G model system to assess molecular docking methods. In chapter 1, CCP W191G was introduced as a new polar and charged cavity, which we subsequently used to evaluate the DOCK scoring function, a physically derived, but approximate method best used for brute force screening of large virtual libraries of small molecules. In chapter 2, CCP W191G was included in a series of model systems to evaluate more sophisticated scoring functions, MM-GBSA methods. Although these methods are more physically realistic, including some protein relaxation and molecular mechanics treatment of bonded terms (torsions, etc), still there are limitations. For one, the solvent is treated implicitly (GBSA) and does not include solute entropic effects. Although protein dynamics are considered, we had found that relaxing the protein worked best within a limit of 5-8Å from the binding site, allowing more of the protein to relax introduced decoy receptor conformations (Chapter 2). Finally, due to the restraints on computational efficiency, sampling of the ligand is limited to conformational space near the DOCK starting pose.

So, the question we set out to address was, if we aren't limited by computational cost and can use the highest level of theory available to us, can we successfully predict ligands (and decoys) in a polar model system? We already had good evidence that we could be successful in another simpler system, the hydrophobic L99A cavity in T4 lysozyme, using alchemical free energy methods to predict the absolute binding free energy of small molecules to this cavity.¹ Unfortunately, using CCP W191G as the polar model system was too great a leap from the L99A cavity. The presence of ordered displaceable waters, a flexible loop², finding accurate parameters for the heme, and

accounting for charged ligands proved to be, for the time being, a hurdle that was too great.

Instead, we took a step back and looked to the less complicated, but still slightly polar, model cavity T4 lys L99A/M102Q. The advantages to using this system were, as in L99A, the availability of several ligands with known affinities and crystal structures for retrospective analysis. We also could anticipate certain protein motions, sidechain rotamers and slight unwinding of helix F based on information from L99A and from M102Q crystal structures.³⁻⁵ Finally, both protein crystallography and ligand affinity measurements by low *c*-value ITC were well established for the T4 lysozyme model systems.⁶

The question we initially sought to address was whether alchemical free energy methods could also predict the absolute binding free energy in a polar cavity that binds more diverse ligands. The distinction between what constitutes too polar is subtle; both benzene, phenol and catechol bind, but 2-aminophenol is a decoy. What we didn't anticipate was that every compound we picked, with the exception of *n*-phenylglycinonitrile would enlarge the cavity to some extent. Although polarity (and the potential for hydrogen bonding) was what we initially set out to test, this was not the defining characteristic of the compounds we picked; instead unknowingly it was ligand size, which translated to protein conformational change upon ligand binding.

Another challenging aspect for of this study was sampling ligand orientations. A good example was our decision, based purely on computational expediency, to restrict the relative free energy calculations only to starting orientations with the hydroxyl pointing towards the glutamine (Gln102). This seemed reasonable at the time, but could have been a disastrous failure if we had only picked phenol as our reference ligand. The predictions from phenol were random at best, with no rank ordering and generally poor prediction of

binding mode. Luckily, by including catechol, with two hydroxyl positions (therefore 4 starting orientations), we were able to correctly rank order our predictions by relative affinity and accurately predict the correct binding mode as well. Restricting sampling of the ligand starting orientation too much resulted in nearly complete failure; simply doubling the possible orientations from 2 to 4 was enough to recover rank-ordering and pose prediction, at least in the relative calculations.

Overall, by comparison to the results for the L99A cavity¹, those from this study in M102Q are less impressive as far as predictive ability (no rank-ordering in the absolute free energy predictions, generally overestimated free energies, and a fair number of mispredicted poses). However, we gained a good deal of knowledge to apply to the next round of free energy predictions, soon to be underway in the CCP model systems. Furthermore, the new ligands and decoys from this study, 14 total compounds for the absolute free energy calculations and 6 new ligands from the relative free energy calculations, including full characterization of new ligands from affinity prediction by ITC to crystal structures of the 13 ligands in complex with M102Q, provides an invaluable test set for future studies.

This work was a collaboration with David Mobley, who handled the computational side of the project, both the absolute and relative free energy calculations, and with whom I share co-first authorship. Gabe Rocklin provided an absolutely invaluable critical eye to all aspects of the project and helped with a great deal of the analysis. Alan Graves introduced me to this project, picked the initial compounds for the absolute free energy calculations, did several of the initial binding assays, and was responsible for the crystal structures of two ligands, *n*-phenylglycinonitrile and 2-nitrothiophene (Chapter 2). The Supplementary Material for this chapter is in Appendix C.

References

1. Mobley, D. L., Graves, A. P., Chodera, J. D., McReynolds, A. C., Shoichet, B. K. & Dill, K. A. (2007). Predicting absolute ligand binding free energies to a simple model site. *J. Mol. Biol.* **371**, 1118-1134.
2. Fitzgerald, M. M., Musah, R. A., McRee, D. E. & Goodin, D. B. (1996). A ligand-gated, hinged loop rearrangement opens a channel to a buried artificial protein cavity. *Nat Struct Biol* **3**, 626-31.
3. Wei, B. Q., Baase, W. A., Weaver, L. H., Matthews, B. W. & Shoichet, B. K. (2002). A model binding site for testing scoring functions in molecular docking. *J. Mol. Biol.* **322**, 339-355.
4. Wei, B. Q., Weaver, L., Ferrari, A. M., Matthews, B. M. & Shoichet, B. K. (2004). Testing a flexible-receptor docking algorithm in a model binding site. *J. Mol. Biol.* **337**, 1161-82.
5. Graves, A. P., Brenk, R. & Shoichet, B. K. (2005). Decoys for Docking. *J. Med. Chem.* **48**, 3714 - 372.
6. Turnbull, W. B. & Daranas, A. H. (2003). On the Value of c : Can Low Affinity Systems Be Studied by Isothermal Titration Calorimetry? *J. Am. Chem Soc.* **125**, 14859-14866.

Chapter 3:

Predicting ligand binding affinity with alchemical free energy methods in a polar model binding site

Sarah E. Boyce^{1†}, David L. Mobley^{2†}, Gabriel Rocklin³, Alan P. Graves^{3‡},

Ken A. Dill^{4*}, and Brian K. Shoichet^{4*}

¹ Graduate Group in Chemistry and Chemical Biology, University of California- San Francisco

San Francisco, CA 94158-2518

² Department of Chemistry, University of New Orleans

New Orleans, LA 70148

³ Graduate Group in Biophysics, University of California- San Francisco

San Francisco, CA 94158-2518

⁴ Department of Pharmaceutical Chemistry, University of California- San Francisco

San Francisco, CA 94158-2518

* Corresponding authors

† These authors contributed equally to this work.

‡ Current address: GlaxoSmithKline Pharmaceuticals, 709 Swedeland Road,

King of Prussia, Pennsylvania 19406

3.1 Abstract

We present a combined experimental and modeling study of organic ligand molecules binding to a slightly polar engineered cavity site in T4 lysozyme (L99A/M102Q). For modeling, we computed alchemical absolute binding free energies. These were blind tests performed prospectively on 13 diverse, previously untested candidate ligand molecules. We predicted that eight compounds would bind to the cavity and five would not; 11 of 13 predictions were correct at this level. The RMS error to the measurable absolute binding energies was 1.8 kcal/mol. In addition, we computed *relative* binding free energies for six phenol derivatives starting from two known ligands: phenol and catechol. The average RMS error in the relative free energy prediction was 2.5 (phenol) and 1.1 (catechol) kcal/mol. To understand these results at atomic resolution, we obtained x-ray co-complex structures for nine of the diverse ligands and for all six phenol analogs. The average RMSD of the predicted pose to the experiment was 2.0Å (diverse set), 1.8Å (phenol derived predictions) and 1.2Å (catechol derived predictions). We found that to predict accurate affinities and rank-orderings required near-native starting orientations of the ligand in the binding site. Unanticipated binding modes, multiple ligand binding, and protein conformational change all proved challenging for the free energy methods. We believe these results can help guide future improvements in physics-based absolute binding free energy methods.

Keywords

Alchemical free energy, free energy calculation, T4 lysozyme, model cavity site,
hydrophobic and polar

3.2 Introduction

A longstanding goal in structure based drug discovery is to predict ligand binding free energies accurately. The most sophisticated simulation-based methods, such as free energy perturbation (FEP) and thermodynamic integration (TI), calculate the binding free energy of small molecules to proteins using a thermodynamic cycle. These approaches step through a series of non-physical (alchemical) overlapping states from free ligand + protein to the protein-ligand bound complex. Nearly all such approaches, with few exceptions¹⁻³, require a bound structure as input. The calculations are time consuming, due in part to the explicit treatment of solvent, as well as the need for many intermediate states. Also, each step of the transformation requires equilibration of the system. Finally, the many degrees of freedom in these systems make it difficult to achieve convergence of the free energy estimate.

Despite these concerns, investigators have been able to correlate calculated absolute and relative binding free energies of small molecules with experimental results, some to within 1-2 kcal/mol in retrospective studies^{1,3-19}. There are far fewer examples of prospective predictions, a more rigorous test for any computational method. Most successful prospective tests to date have involved relative free energies, where a ligand's affinity is predicted based on the measured affinity of a similar ligand. An example of such prospective prediction is the work of Jorgensen & colleagues, who used relative free energy calculations for lead optimization of potent non-nucleoside HIV-1 reverse transcriptase inhibitors²⁰. In this case, knowledge of the affinities and structural information for a set of ligands was leveraged to systematically improve potency relative to the starting compounds. A more ambitious goal is to predict absolute binding free energies: these calculations would not require knowledge of other ligands already tested.

In a previous study, we used a simple model cavity to explore the strengths and weaknesses of alchemical free energy methods. Owing to their simplicity, tractability, and the many untested plausible ligands, it is possible to test prospective predictions and analyze both successes and failures in detail for such sites. We therefore predicted the absolute binding free energy of five unknown compounds to the hydrophobic model binding site T4 lysozyme L99A ². The simplicity of the site made it possible to isolate errors due to convergence and force field issues that might otherwise have been masked by pKa shifts, complex protein motions, and other factors. In the hydrophobic cavity, free energy methods correctly separated ligands from non-binders, rank-ordered and accurately calculated the new ligand affinities, and predicted the correct binding geometries.

Still, the simplicity of this site and the few compounds tested left several questions unanswered, specifically how alchemical free energy methods would perform with the introduction of polarity into the binding site and with the corresponding increase in the diversity of ligands that would then bind. To explore these questions we turned to a second model binding site in T4 lysozyme L99A/M102Q, essentially the same hydrophobic cavity but with polarity and the potential for hydrogen bonding introduced by the Met102 → Gln substitution (Fig. 1.A). As in the apolar L99A cavity, this site binds small aromatic molecules (benzene, toluene etc), but it also binds more polar molecules such as phenol. Not all polar molecules bind, for example whereas catechol is a ligand, the isosteric 2-aminophenol does not bind detectably. The balance between the solvation free energy of the ligand and the requirement of correctly orienting molecules for hydrogen bonding to Gln102 introduces complexities absent from the apolar cavity, without abandoning the essential simplicity of the cavity. And, like the original site, the L99A/M102Q model system lends itself to prospective studies, with both experimental

assays and protein crystallization conditions well established²¹⁻²³. Also – and this is a key point – because the cavity is so small ($\sim 150\text{\AA}^3$) there are many hundreds of likely ligands to be found within commercially available libraries that are easy to acquire and test.

3.3 Results

Retrospective analysis of ligands and decoys

To begin this study, seven known ligands and two related non-binders were selected for retrospective free energy calculations. Results from the previous study in the hydrophobic cavity, L99A, suggested that accurate prediction of ΔG_b required forced umbrella sampling of Val111 sidechain rotamers (Fig.1 Supplementary Material)^{2,24}. This movement had been observed to occur with ligand binding, sometimes coupled with a motion of helix F, residues 108-114 (Fig 1.B) , enlarging the cavity to accommodate bigger ligands. For the polar cavity, L99A/M102Q, alternate rotamers of Leu118 and Val103 were also observed in the complex crystal structures with catechol and 3-chlorophenol (PDB IDs 1XEP, 1LI3); we therefore sampled all these rotamers explicitly in this study using umbrella sampling (Fig 2 Supplementary Material)²⁴. Also, unlike common approaches that use bound structures as inputs, we attempt to predict the actual binding mode by sampling many possibilities. Because different possible orientations and conformations can be separated by large kinetic barriers, we sample each “pose”, corresponding to a kinetically stable orientation, separately.

The retrospective absolute free energy calculations for the seven ligands and two decoys had an RMS error of 1.2 kcal/mol relative to experiment (Table 1 Supplementary Material)²³. Six of the seven ligands were correctly identified as measurable binders (K_d values $< 10\text{mM}$) and one of the two non-binders was also correctly identified as such.

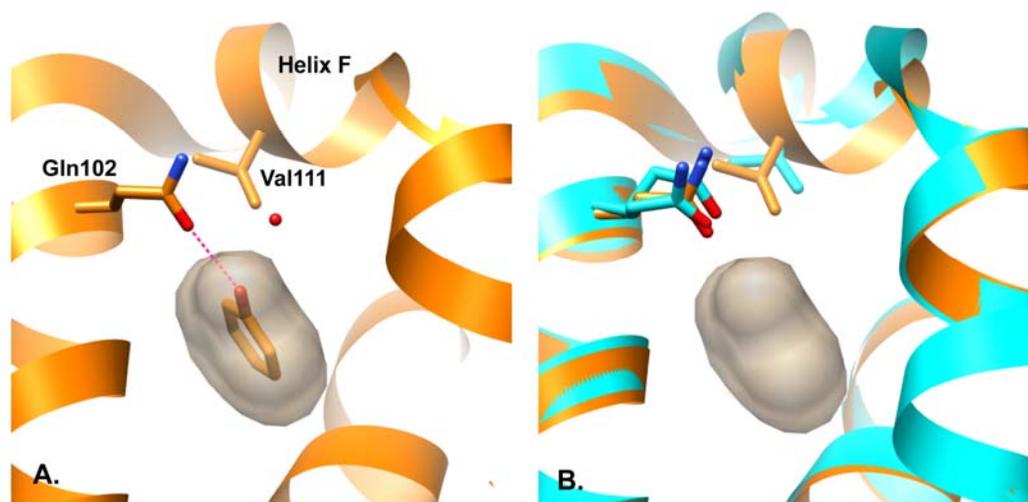


Figure 1. (a) T4 Lysozyme L99A/M102Q binding site shown in complex with phenol and one ordered water molecule 23. (b) Unwinding of Helix-F upon binding of certain ligands (cyan) yields an enlarged binding site relative to apo (orange).

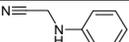
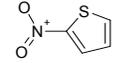
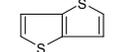
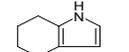
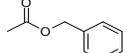
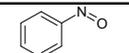
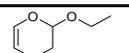
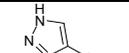
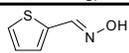
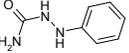
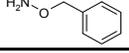
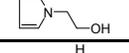
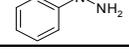
There were two failures in this set: 2-aminophenol and 4-vinylpyridine. 2-aminophenol is a non-binder that was calculated to have a ΔG_b of -3.93 kcal/mol, indicating a weak but measurable affinity ($K_d \approx 1\text{mM}$), while 4-vinylpyridine is also a non-binder with a calculated ΔG_b of -3.39 kcal/mol (very weakly binding, $K_d \approx 3\text{mM}$). The prediction of the binding geometries for this retrospective set was satisfactory. Most ligand orientations were predicted to within 2\AA of the crystallographic binding mode. Despite the small size of the test set, these results suggested that prospective tests for this site would be feasible.

Prospective predictions: Absolute binding free energy

To choose candidate, previously untested, ligands for L99A/M102Q we docked a large library of organic small molecules and selected compounds, keeping in mind that we wanted a range of affinities, from non-binders to those with good ligand efficiencies. Thirteen compounds were chosen, encompassing both compounds we believed might be too polar or too large for the cavity to those that we (SEB & BKS) expected to bind. The

experiment was set up in a blind fashion; the calculation of the absolute binding free energies was performed independently (DLM, KAD) without knowledge of the experimental results, whereas the measurement of K_d values (by ITC) and the

Table 1. Prospective test: absolute binding free energy predictions

Structure	$\Delta G_{b,exp}$ (Kcal/mol) a.	$\Delta G_{b,calc}$ (Kcal/mol) b.	RMSD (Å) c.	PDB ID
	-5.52 ± 0.18	-5.63 ± 0.38	0.87	2RBO
	-4.85 ± 0.25	-5.73 ± 0.13	1.09(A); 2.86(B)	2RBN
	-4.67 ± 0.17	-6.8 ± 0.23	0.73(A) 0.44(B)	3HUQ
	-4.61 ± 0.09	-5.4 ± 0.45	0.66 & 1.78**	3HUA
<i>benzyl acetate</i> 	-4.48 ± 0.16	-1.31 ± 0.44	>10	3HUK
	weak	-5.55 ± 0.23	3.24(A), 3.22(B)	3HU9
	weak	-3.8 ± 0.18	4.35	3HTG
	weak	-7.86 ± 0.12	2.07	3HTF
	NB	-2.30 ± 0.09	---	---
	NB	0.45 ± 0.24	---	---
	NB	-2.58 ± 0.13	---	---
<i>1-2-hydroxyethylpyrrole</i> 	NB	-5.72 ± 0.12	---	---
	NB	-2.51 ± 0.44	---	---

Compound names of mispredictions are in bold italics. ** indicates RMSD for one crystal pose calculated to two predictions. ^aFree energy of binding determined by ITC at 10°C. For ligands designated weak, binding was established by T_m upshift but ΔG_b could not be determined. NB, nonbinder. $\Delta T_m \approx 0^\circ\text{C}$ at concentrations between 1 and 10mM. ^bCalculated free energy of binding. ^cRMSD of predicted ligand geometry to experimentally observed crystal pose. If multiple ligand orientations were present in the crystal (designated A and B), only the best RMSD to the prediction is reported.

determination of x-ray structures (both in the Shoichet lab) were done before the predictions were made known.

Of the 13 compounds, five were not detected to bind and eight were ligands (Table 1). Of the eight new ligands, three had weak to moderate affinity: nitrosobenzene, 4-chloro-1H-pyrazole, and 2-ethoxy-3,4-dihydropyran; binding was only detected at concentrations higher than 1mM by thermal denaturation upshift (ΔT_m). Due to this and to solubility limitations, the actual binding affinity could not be measured for these ligands and was instead estimated to be between 1-10mM (-3.88 to -2.59 kcal/mol). Of the remaining five new ligands, the binding free energies, determined by ITC, ranged from -4.48 kcal/mol (benzylacetate) to -5.52 kcal/mol (n-phenylglycinonitrile), placing this last ligand among the best known for this cavity (Fig 2).

The complex crystal structures of the ligands bound to L99A/M102Q were determined for all eight new ligands, with resolutions between 1.26Å and 1.85Å (Fig. 3.A-R, Table 2 Supplementary Material). Three structures had multiple binding modes: benzylacetate, thieno[3,2-b]thiophene and 2-nitrothiophene. Two others had adopted geometries in the site dissimilar to all previously known ligands. Thiophene-2-carboxaldoxime, purchased as the 98% pure *E* isomer, was not observed to bind in the CD assay at 5mM, but electron density for the *Z* isomer was observed at a crystal soak concentration of 50mM (PDB entry 3HTD). Apparently the *Z* isomer was present as an impurity, making (*Z*)-thiophene-2-carboxaldoxime the ninth new ligand discovered in this study. The other compound, 4-chloro-1H-pyrazole, had two molecules present in the cavity, both at 100% occupancy. This is unprecedented for this site. To accommodate this, helix F unwinds from residue 108 to 114. This enlarges the cavity more than observed with any other ligand and opens it to bulk solvent. The new water channel that

results interacts with both the helix F backbone and with one molecule of 4-chloro-1H-pyrazole.

Movement of helix F was observed in *all* of the complex structures with only one exception, n-phenylglycinonitrile (Fig. 3.K-L). This movement displaces the Val111 sidechain, allowing larger ligands to fit in the cavity. Conversely, few alternate sidechain rotamers were observed on ligand binding. The only example for Val111 was the 2-nitrothiophene bound structure, in which an alternate rotamer conformation was observed in addition to the apo rotamer. Val103 had two alternate conformations in the benzylacetate and 2-ethoxy-3,4-dihydro-2h-pyran structures. Leu118 had alternate rotamer conformations in two structures: 4,5,6,7-tetrahydroindole and (Z)-thiophene-2-carboxaldoxime. All other rotamers observed corresponded to the apo conformation (Table 3 Supplementary Material).

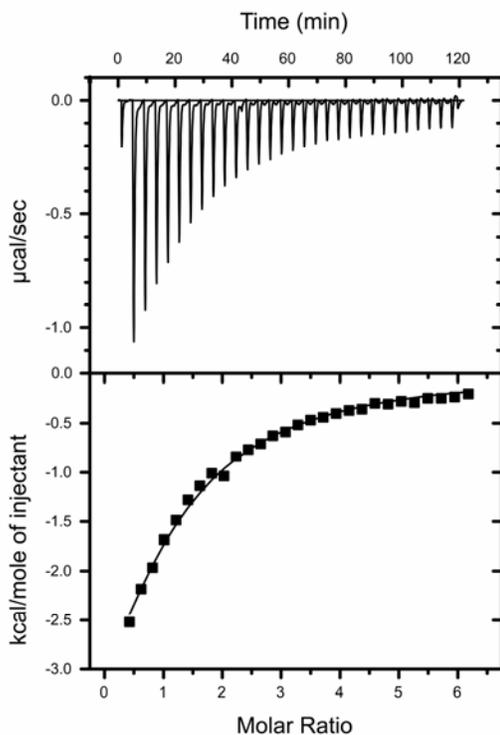


Figure 2. Representative ITC measurements and fit: n-phenylglycinonitrile titration (stock concentration 1.4mM) into L99A/M102Q (initial concentration 42.5µM). An initial injection of 2µL was followed by 29 x10µL injections of ligand to a final ligand concentration in the reaction cell of 237.9µM.

Comparing Prediction to Experiment

The criteria we use to assess the accuracy of the predictions are as follows: 1) whether the free energy methods correctly distinguish ligands from non-binders, 2) whether the predictions are accurate (RMS error to experimental ΔG_b) and whether the compounds can be rank-ordered by affinity, 3) RMSD of the predicted pose(s) to the experimentally determined binding mode. In this system, the free energy approach was unable to rank-order the compounds by affinity. For seven of the 13 compounds, the affinity is overestimated, and even when the ligand was correctly predicted as a binder the affinity was overestimated six out of nine times (Table 1). Three compounds were completely mispredicted (either as a binder or non-binder when the reverse was true): 1,2-hydroxyethylpyrrole, benzylacetate and n-phenylglycinonitrile. In the last of these, a trivial error led to incorrect partial charges. With the corrected result for n-phenylglycinonitrile, the overall RMS error in the predicted absolute binding free energies was 1.8 kcal/mol for the five compounds with measurable affinities (by ITC). If we estimate the affinity of the 3 weak ligands based on the concentration where significant T_m upshift was observed (nitrosobenzene $\approx 1\text{mM}$, 4-chloro-1H-pyrazole $\approx 5\text{mM}$, 2-ethoxy-3,4-dihydro-2H-pyran $\approx 10\text{mM}$), the overall RMS error is 2.3 kcal/mol. Of the two remaining mispredictions, 1,2-hydroxyethylpyrrole is a non-binder, but it is predicted to be a low micromolar binder, with $\Delta G_{b,\text{calc}} = -5.72$ kcal/mol. Benzylacetate is a ligand for this site, yet the predicted ΔG_b is unfavorable.

Comparing the predicted ligand poses to the experimentally determined binding modes, the overall average RMSD to the crystallographic results was 2.0\AA . In six of the nine structures the RMSD was less than 2\AA , thieno[3,2-b]thiophene, 4,5,6,7-tetrahydropyran, 2-nitrothiophene, thiophene-2-carboxaldoxime, 4-chloro-1H-pyrazole,

and *n*-phenylglycinonitrile. The predictions that corresponded most closely to the experimental results were thieno[3,2-*b*]thiophene, 2-nitrothiophene and 4,5,6,7-tetrahydropyran. For these three ligands, not only did the free energy methods accurately predict the affinity, the binding pose prediction was also accurate (Table 1). Two equally contributing orientations were predicted for all 3 compounds, although for 4,5,6,7-tetrahydropyran only one binding mode was observed crystallographically (Fig. 3.M-L). All three ligands induced a slight conformational change in helix F, as observed in the ligand-protein complex structure, but this was not observed in the predicted structure.

Instead, alternate rotamers of Val111 and Leu118 were predicted in order to avoid a steric clash with the sidechain atoms. These rotamer predictions did not agree with the experimental result, with the exception of Val111 for 2-nitrothiophene (Table 3 Supplementary Material). Despite this, we count these three cases as successes for pose prediction. Although thiophene-2-carboxaldoxime was purchased as the *E* isomer, the *Z* isomer is the actual ligand observed to bind in the x-ray complex structure. The prediction agreed with the experimental result for both isomers for binding and predicted orientation; the *E* isomer was predicted to be a non-binder, $\Delta G_{b,calc} = -2.30$ kcal/mol and the *Z* isomer a binder, $\Delta G_{b,calc} = -7.13$ kcal/mol. Although the RMSD of the prediction was satisfactory to both observed binding modes, 1.9Å for each, the predicted structure was the mirror image (Fig. 3.C-D); however, the hydrogen bond to Gln102 is maintained in spite of this. Additionally, an alternate rotamer was predicted for Leu118 for (*Z*)-thiophene-2-carboxaldoxime, which corresponded with the experimental result (Table 3 Supplementary Material).

Figure 3. Comparison of predicted to experimental binding modes for the absolute binding free energy predictions.

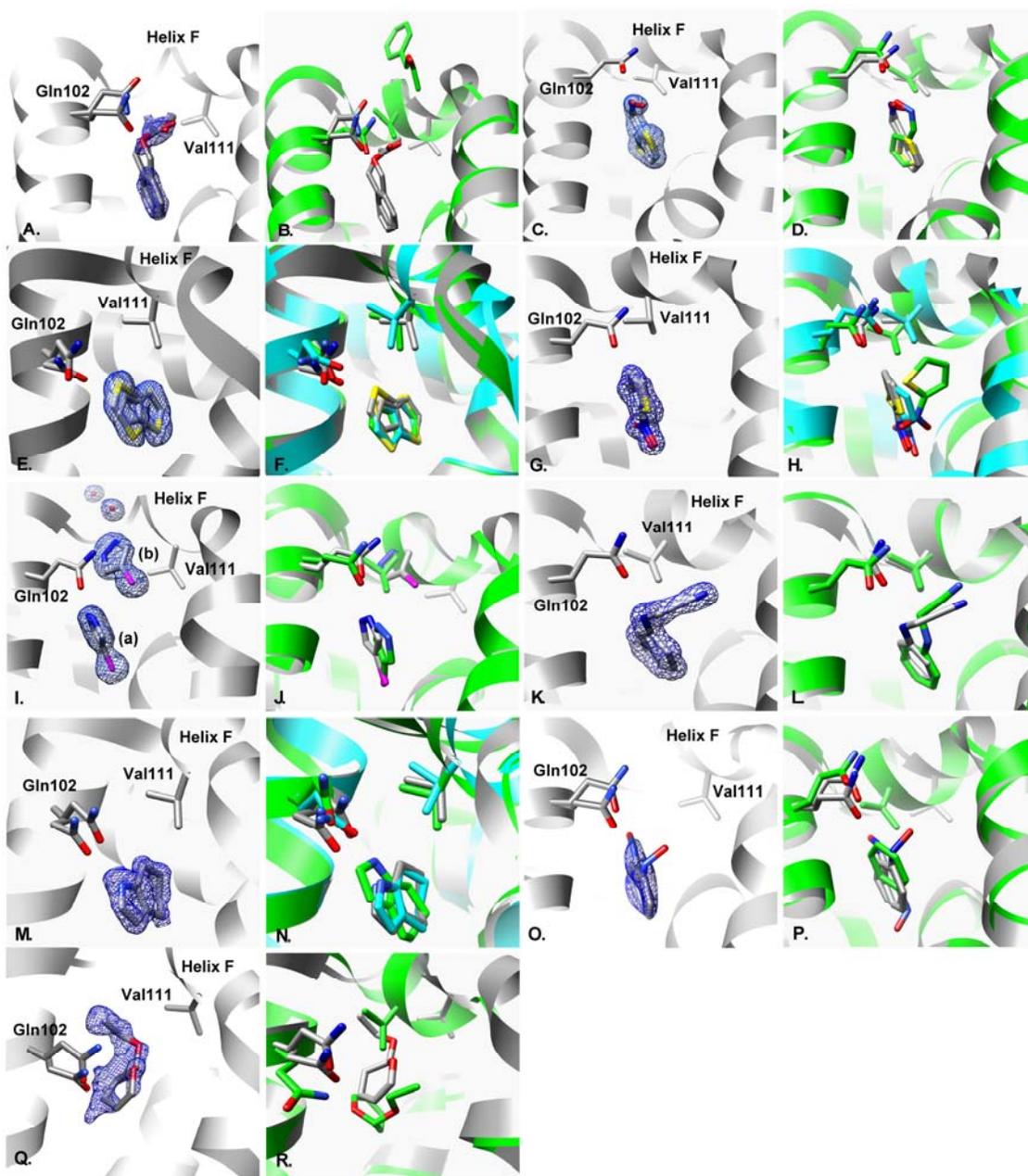


Figure 3. (a and b) Two conformations of benzylacetate: (a) X-ray result, 25:25 occupancy. $2F_o - F_c$ electron density map displayed at 1.0σ . (b) Overlay of x-ray result (gray) with predicted geometry (green). (c and d) Two conformations of thiophene-2-carboxaldoxime: (c) X-ray result, 50:50 occupancy. $2F_o - F_c$ electron density map displayed at 1.5σ . (d) Overlay of x-ray result (gray) with predicted geometry (green). (e and f) Two conformations of thieno[3,2-b]thiophene: (e) X-ray result, 50:50 occupancy. $2F_o - F_c$ electron density map displayed at 1.5σ . (f) Overlay of x-ray result (gray) with

predicted geometries (green & cyan). (g and h) Two conformations of 2-nitrothiophene: (g) X-ray result, 50:50 occupancy. $2F_o-F_c$ electron density map displayed at 1.5σ . (h) Overlay of x-ray result (gray) with predicted geometries (green & cyan). (i and j) Two molecules of 4-chloro-1H-pyrazole: (i) X-ray result, two molecules (A and B) bound at 100% occupancy with two ordered water molecules. $2F_o-F_c$ electron density map displayed at 1.5σ . (j) Overlay of x-ray result (gray) with predicted geometry (green). (k and l) n-phenylglycinonitrile: (k) X-ray result, 100% occupancy. $2F_o-F_c$ electron density map displayed at 1.5σ . (l) Overlay of x-ray result (gray) with predicted geometry (green). (m and n) 4,5,6,7-tetrahydroindole: (m) X-ray result, 50% occupancy. $2F_o-F_c$ electron density map displayed at 1.0σ . (n) Overlay of x-ray result (gray) with predicted geometries (green & cyan). (o and p) Two conformations of nitrosobenzene: (o) X-ray result, 25:25 occupancy. $2F_o-F_c$ electron density map displayed at 1.0σ . (p) Overlay of x-ray result (gray) with predicted geometry (green). (q and r) 2-ethoxy-3,4-dihydro-2h-pyran: (q) X-ray result, 25% occupancy. $2F_o-F_c$ electron density map displayed at 0.5σ . (r) Overlay of x-ray result (gray) with predicted geometry (green).

4-chloro-1H-pyrazole was also a partial success, due to the presence of two ligands bound simultaneously at 100% occupancy in the binding site; predicting both molecules is an impossible case for the free energy methods. Admittedly, the calculated ΔG_b was greatly overestimated for even one molecule, yet it was still correctly predicted to bind. Also, when we compare the prediction to the molecule bound in the canonical L99A/M102Q binding site (pose A) the RMSD is 2.1\AA , which seems high, but the key interactions are captured (Table 1, Fig. 3.I-J). However, the helix F conformation change is not predicted, although it is the largest movement seen for this set of compounds. Instead, alternate rotamers of Val111 and Leu118 are predicted, inconsistent with the experimental result (Table 3 Supplementary Material).

There are three cases, 2-ethoxy-3,4-dihydro-2h-pyran, nitrosobenzene, and benzylacetate, in which the RMSD is greater than 3\AA . In all three, the ligand clearly cannot fit in the binding site unless helix F moves out of the way, enlarging the cavity.

For example, the prediction for nitrosobenzene occupies the same overall location in the cavity as the crystallographically observed ligand, but is flipped over in the binding site relative to the actual binding mode to avoid a steric clash with the Val111 sidechain atoms (Fig. 3.O-P). This allows the ligand to fit in the cavity with what is essentially the apo conformation of the helix F, but the interaction between the nitroso substituent and Gln102 is missed. Benzylacetate and 2-ethoxy-3,4-dihydro-2h-pyran are more extreme cases in which the ligand itself occupies space normally taken by the Val111 sidechain and helix F backbone atoms in the apo conformation (Fig. 3.A-B, Q-R, respectively). Neither predicted binding pose is accurate, because in both helix F remains in its apo conformation, leaving no space for the ligand. Consequently, benzylacetate starts in an unfavorable orientation and over the course of the simulations actually leaves the binding site, which explains why it was predicted to be a non-binder. Although the predicted orientation(s) of 2-ethoxy-3,4-dihydro-2h-pyran and nitrosobenzene are incorrect, both were correctly predicted to be ligands for the cavity and are therefore not complete failures. Benzylacetate, on the other hand, is a hard failure. Both predicted pose and affinity are incorrect, owing to its departure from the cavity during the simulation. The last ligand, n-phenylglycinonitrile, is also a failure, albeit due to a scripting error. Although the RMSD to the crystal pose was 0.9Å, the key hydrogen bond to Gln102 is missed, and it was predicted to be a non-binder, yet experimentally it has the highest affinity of all the ligands. This was determined to be the result of incorrect partial atomic charges in the ligand parameter file for one portion of the initial calculation. Correction of the parameter file led to the recovery of the ligand, with a predicted affinity $\Delta G_b = -5.63 \pm 0.38$ kcal/mol.

Analysis of the predictions after the experiments

As noted, some failures can be explained by errors in the calculations. After we learned the results of the affinity and crystallographic experiments, we re-analyzed the initial prospective calculations and uncovered problems with several. These were:

- (1) A bug in GROMACS caused long range dispersion corrections to be calculated incorrectly, introducing errors in the range of 0-3 kcal/mol for all calculations.
- (2) A bug in Antechamber caused some partial charges to be computed incorrectly, introducing errors of up to 1 kcal/mol for some more polar molecules
- (3) A scripting error led to jumbled partial charges for atoms in some parameter files (for example for n-phenylglycinonitrile), introducing errors up to 3 kcal/mol.

To fix these errors, we (1) recomputed the long-range dispersion correction, which is always applied in a post-processing step. For problems (2)-(3), we performed a set of calculations where we computed the free energy of switching the ligand partial charges to the correct set for each ligand and orientation. In the case of (2), because Antechamber gave incorrect partial charges (see Supplementary Material) we obtained AM1-BCC partial charges directly from Christopher Bayly²⁵.

For the corrected results on the prospective compounds, the RMS error was 1.6 kcal/mol; a full table of corrected results for the prospective compounds is available in the Supplementary Material (Table S4). To preserve the prospective nature of the test, we have not changed the results presented in the tables and figures discussing prospective

data (Table 1-2, Fig 3-4) to reflect these corrections. However, both the results for the retrospective set (Table S1 Supplementary Material) and the calculations from the holo structure, discussed below, include these corrections.

Absolute binding free energy calculations from the holo structure

For ligands that induce some degree of unwinding in helix F we saw two distinct trends, the predicted free energy was far too favorable in most cases and the protein conformational change in helix F was not predicted. We were concerned that the original calculations had not converged and to investigate this we restarted the free energy calculations using the holo structures of the 9 ligands from the prospective study. It is important to mention here that we would expect the energies to be even more favorable if the movement of helix F had been accounted for in the original calculation; therefore, although not relevant to true predictions – a major goal of this study – the holo calculations are useful to understand where the predictions go awry, a point to which we will return later.

The RMS error of the holo result from the experimental results was 3.0 kcal/mol for the five ligands with measurable affinities (compared to 1.6 kcal/mol for apo); with estimated affinities for the weak ligands (see Results: Comparing prediction to experiment) the RMS error was 2.6 and 1.8 kcal/mol for the holo and apo results, respectively. The high RMS error in the holo result was due to one compound whose affinity was grossly overestimated, 2-nitrothiophene. Discounting this outlier, the RMS error calculations decreased to 1.4 kcal/mol. Additionally, compared to the experimental result the affinity from the holo structures was only overestimated in two out of the nine ligands. Indeed, the holo structure calculations overestimated the affinity *less* frequently than the apo structure calculations. This observation was initially perplexing to us, but

may be reconciled by several sources of convergence problems discovered during the calculations. This a point to which we will return (Discussion).

The RMSD of the predicted pose(s) to the crystallographic binding modes was greatly improved over the apo calculation (Table 4 Supplementary Material). Of particular note, the prediction for benzylacetate, which left the binding site when starting from the apo structure, was found with an RMSD $< 1 \text{ \AA}$ to the experimental binding mode (orientation A). The only exception was the prediction for 2-ethoxy-3,4-dihydro-2h-pyran, which did not improve over the apo result.

Predicting the relative binding free energy

One might hope that comparing relative binding strengths of related inhibitors, starting from one or more ligands with known affinities and structures, would be a less challenging computational problem. We selected six previously untested derivatives of the ligand phenol to predict affinities for: 2-methylphenol, 2-ethylphenol, 2-propylphenol, 2-methoxyphenol, 2-ethoxyphenol and 5-chloro-2-methylphenol. For all the compounds there is one substituent ortho to the hydroxyl group and, in many cases, the difference between one compound and the next is only one heavy atom. The binding free energy of each ligand was then calculated relative to two previously known reference ligands, phenol and catechol.

The affinity and ΔG_b for the six unknowns and for catechol were determined by ITC and ranged from -5.51 kcal/mol to -4.02 kcal/mol (Table 2). Due to limited solubility and low affinity the full binding isotherm for 2-methoxyphenol could not be measured, but was estimated from the incomplete ITC curve to be -2 kcal/mol, making this an extremely weak ligand for this cavity. X-ray crystal structures of each ligand in complex with the protein were also determined, with resolutions between 1.59 \AA and 2.02 \AA (Table

2, Fig. 4.A-L, Table 5 Supplementary Material). The structures reveal two important and unanticipated results; all ligands induce movement in helix F, and two of the more polar ligands, 2-methoxyphenol and 2-ethoxyphenol, have intramolecular hydrogen bonds with the ether oxygen (Fig. 4.A, C, respectively). For all other compounds and one orientation of 2-methoxyphenol, the hydroxyl group hydrogen bonds with Gln102, as do the reference compounds phenol and catechol.

Two of the six ligands have multiple binding modes in the crystal structures: 2-methylphenol and 2-methoxyphenol (Fig. 4.E, A). In the complex structure with 2-methoxyphenol, the two binding modes are related by a rotation around the C2 axis of the ortho substituent, pointing the hydroxyl away from helix F and Gln102. In this second mode there appears to be an intramolecular hydrogen bond for the ligand. Only one orientation of 2-ethoxyphenol is observed, but the hydroxyl is oriented away from Gln102 similar to the alternate orientation observed for 2-methoxyphenol. This is the only ligand for which there appears to be no hydrogen bond with Gln102. Indeed this represents one of the few high quality x-ray structures for which an H-bond group is clearly left unfulfilled ²⁶.

The predictions correlated well with the experimental results for the relative calculations starting from catechol, but less well for those beginning from phenol (Table 2, Fig. 5). For the transformations from phenol, the direction of the change in the binding free energy was correctly predicted for five of six compounds, but the magnitude was overestimated (Table 2, Fig. 5.B), with an RMS error in $\Delta G_{b,calc} = 2.51$ kcal/mol. The transformation of phenol to the other reference ligand catechol resulted in a relative $\Delta G_{b,calc} = 0.36$ kcal/mol. For the catechol transformations, all $\Delta G_{b,calc}$ were in the correct direction. Also, the predictions from catechol are correctly rank-ordered by $\Delta G_{b,calc}$, with the exception of 5-chloro-2-methylphenol, and the overall RMS error between

Table 2. Experimentally determined binding free energy for phenol derivatives

Phenol analogs	$\Delta G_{b,exp}$ (Kcal/mol)	catechol	phenol	PDB IDs
		$\Delta G_{b,calc}$ (Kcal/mol)	$\Delta G_{b,calc}$ (Kcal/mol)	
2-propylphenol	-5.33 ± 0.05	-6.13	-3.05	3HTB
<i>phenol</i>	-5.24	-	-	1LI2
5-chloro-2-methylphenol	-5.04 ± 0.07	-7.54	-1.32	3HT8
2-ethylphenol	-4.57 ± 0.11	-5.15	-1.14	3HT7
2-methylphenol	-4.44 ± 0.11	-4.31	-2.55	3HT6
<i>catechol</i>	-4.16 ± 0.03	-	-4.88 [†]	1XEP
2-ethoxyphenol	-4.02 ± 0.03	-4.16	-4.25	3HU8
2-methoxyphenol	$> -2.0^*$	-2.76	-3.84	3HT9

Reference ligands are in italics. Free energy of binding determined by ITC at 10°C; * indicates the binding free energy was estimated as $\Delta G_b \geq -2.0$. The binding free energies for the six phenol derivatives are shown; for each compound, $\Delta G_{b,calc}$ was calculated from the predicted relative binding free energy and the experimentally determined ΔG_b of the reference compounds catechol or phenol. [†]The corrected retrospective result for the phenol to catechol transformation is given (described in “Analysis of predictions of the predictions after the experiments”).

experiment and prediction was 1.1 kcal/mol (Table 2, Fig. 5.A). It must be mentioned that the same dispersion correction errors that affected the absolute free energy calculations (discussed previously in Retrospective analysis of predictions) also affected the relative free energy calculations. However, recalculating the relative energies did not change the results significantly except for the catechol to phenol transformation (Table 6 Supplementary Material).

The average RMSD for the predicted poses to the crystallographic result was 1.8Å and 1.2Å beginning from phenol and catechol, respectively (Table 2, Fig 4). Consistent with the quality of the binding energy predictions, the predicted poses for

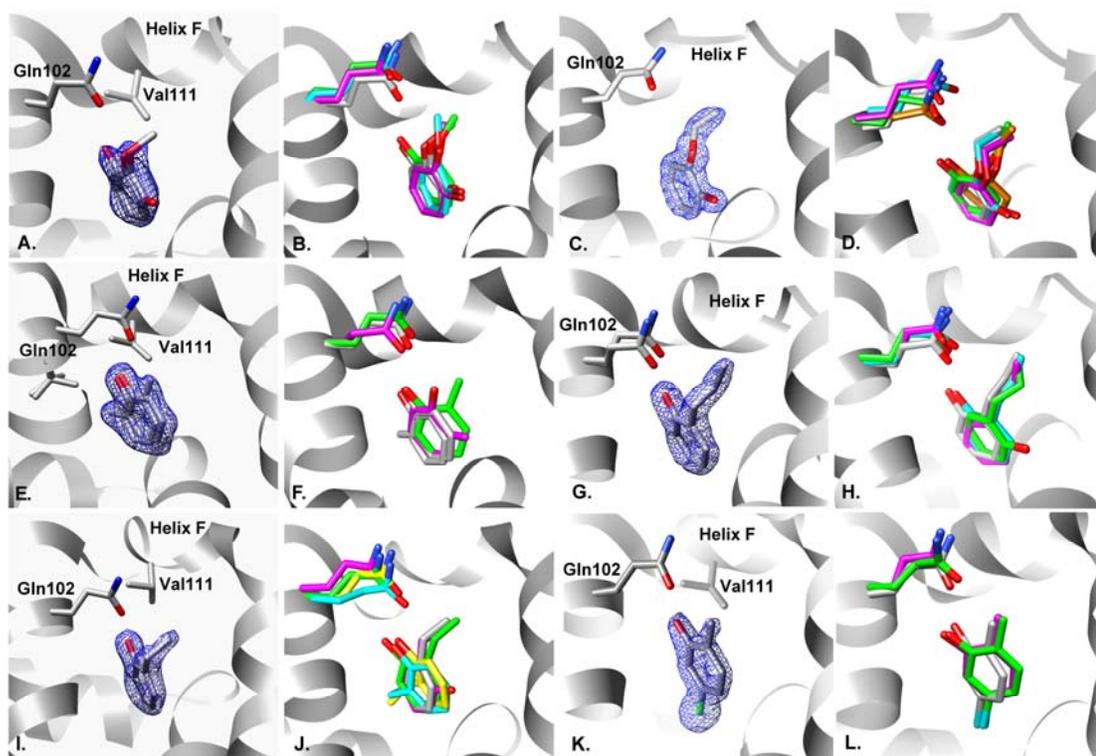


Figure 4. Comparison of predicted to experimental binding modes for the relative binding free energy predictions starting from reference compounds catechol & phenol. (a and b) Two conformations of 2-methoxyphenol: (a) X-ray result, 50:50 occupancy. $2F_o-F_c$ electron density map displayed at 1.0σ . (b) Overlay of x-ray result (gray), prediction from catechol (green & cyan), RMSD 0.52 and 0.85Å, and prediction from phenol (magenta), RMSD 0.65 Å. (c and d) 2-ethoxyphenol: (c) X-ray result, 100% occupancy. $2F_o-F_c$ electron density map displayed at 1.5σ . (d) Overlay of x-ray result (gray) and prediction from catechol (green & cyan), RMSD 0.58 and 0.76Å, and prediction from phenol (magenta), RMSD 0.86Å. (e and f) Two conformations of 2-methylphenol: (e) X-ray result, 50:50 occupancy. $2F_o-F_c$ electron density map displayed at 1.0σ . (f) Overlay of x-ray result (gray) and prediction from catechol (green), RMSD 1.02Å, and prediction from phenol (magenta), RMSD 2.03Å. (g and h) 2-propylphenol: (g) X-ray result, 100% occupancy. $2F_o-F_c$ electron density map displayed at 1.5σ . (h) Overlay of x-ray result (gray) and prediction from catechol (green & cyan), RMSD 0.40 and 1.08Å, and prediction from phenol (magenta), RMSD 2.23Å. (i and j) 2-ethylphenol: (i) X-ray result, 100% occupancy. $2F_o-F_c$ electron density map displayed at 1.5σ . (j) Overlay of x-ray result (gray) and prediction from catechol (green, cyan, yellow), RMSD 0.64Å, 3.10Å, 3.08Å and prediction from phenol (magenta), RMSD 2.25Å. (k and l) 5-chloro-2-methylphenol: (k) X-ray result, 100% occupancy. $2F_o-F_c$ electron density map displayed

at 1.5σ . (l) Overlay of x-ray result (gray) and prediction from catechol (green), RMSD 0.87Å, and prediction from phenol (magenta), RMSD 0.66Å.

catechol are typically accurate. Four of the six new ligands were predicted to have two or more contributing orientations to the overall relative $\Delta G_{b,calc}$: 2-ethoxyphenol, 2-ethylphenol, 2-propylphenol, and 2-methoxyphenol. However, only 2-methoxyphenol had multiple orientations observed crystallographically (Fig 4). For 2-methylphenol, only one of the experimentally observed poses is predicted. The alternate crystallographic orientation, a rotation around the hydroxyl of the ligand, which correlates with the experimentally observed Val103 alternate rotamer configuration, is missed.

2-methoxy- and 2-ethoxyphenol are interesting cases because they make intramolecular hydrogen bonds instead of contacting Gln102. This is something that we were initially surprised by, and we anticipated that the free energy calculations would miss this.

However, for 2-methoxyphenol both orientations of the ligand are correctly predicted, and for the structure where the hydroxyl points away from Gln102 it makes an intramolecular hydrogen bond, as observed experimentally. For 2-ethoxyphenol the intramolecular hydrogen bond is indicated in both predicted orientations, as Gln102 has moved away from the ligand in the simulation.

Overall, the results from the catechol relative free energy calculations are more accurate in both prediction of relative free energy and rank-ordering of the ligands, by comparison to the phenol transformations and also the absolute free energy calculations for the same model system, which did not perform nearly so well. Possible reasons for the success of the relative calculations with respect to the absolute free energy calculations, and what can be learned from this study will be discussed in the following section.

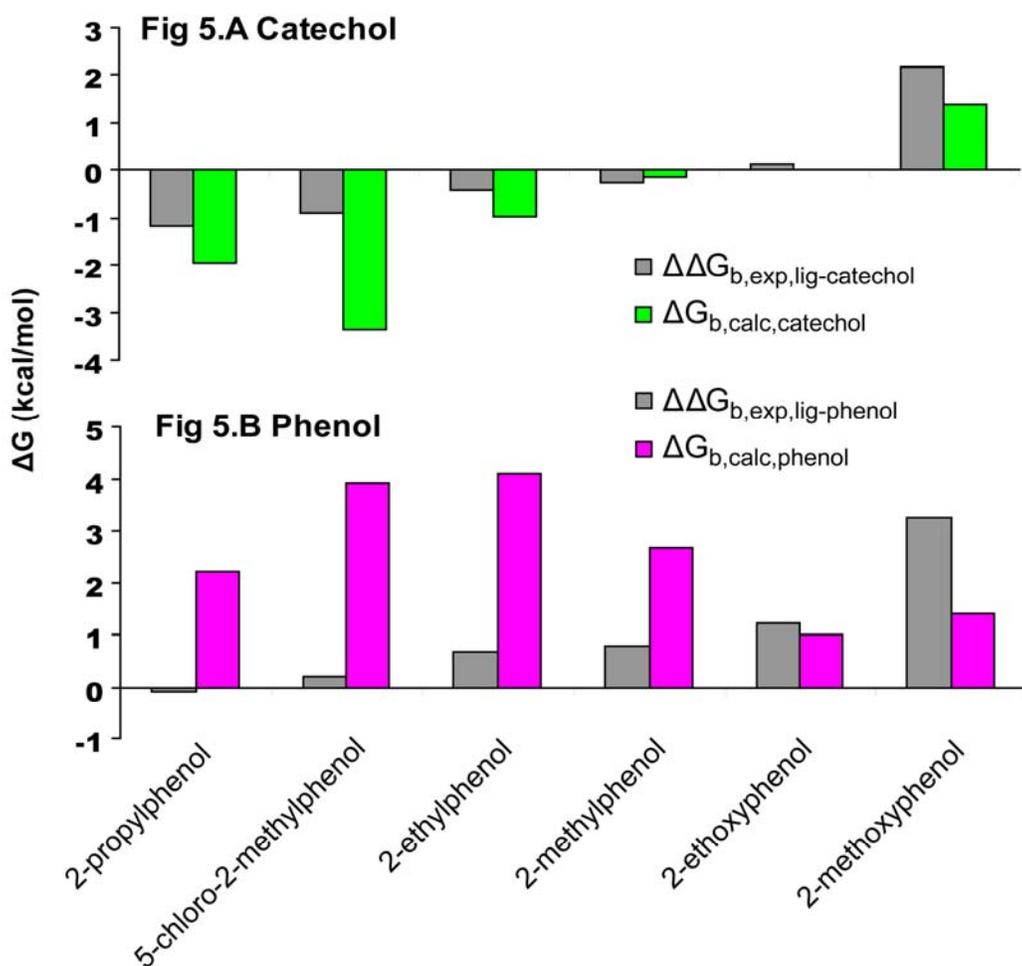


Figure 5. Relative binding free energy predictions from reference compounds Catechol and Phenol. The experimental free energy of binding was determined by ITC at 10°C for each compound; the relative binding free energy for each ligand to the experiment is shown in gray. The calculated relative free energy of binding relative to the reference compound catechol is shown in green and to phenol in magenta.

3.4 Discussion

There are two novel aspects to our study. First, prospective tests were undertaken with 20 new molecules not previously measured for this site. We calculated absolute binding free energies for 14 of them (including a fortuitous impurity) and relative binding

free energies for six. As far as we know, this exceeds the sum total of all prospective tests of these methods in the literature over the last 20 years. Second, these experiments are conducted in a simple cavity site, where we can hope to learn systematically from successes and especially failures of the method by detailed comparisons of prediction to experiment.

To understand what we may learn from this model system, it is useful to summarize how well the calculations corresponded to experiment. We used three criteria, each successively more stringent, to assess the predictions. 1) Did the free energy methods distinguish ligands from non-binders? 2) Were the free energy predictions accurate and can the compounds be rank-ordered by affinity? 3) Are the crystallographic binding modes, and possible alternate ligand orientations, predicted? From the initial prospective calculations, predicting ligand versus non-binder was successful in 10 of 13 cases. Including thiophene-2-carboxaldoxime, for which there were, as we discovered, two relevant isomers to consider, raises these numbers to 11 of 14 correct. Including *n*-phenylglycinonitrile, for which a trivial error led to the assignment of the wrong partial charges, raises the number to 12 correct predictions out of 14 total at the binder/non-binder level. At the more stringent level of energetic accuracy, the RMS error in $\Delta G_{b,calc}$ was 1.8 kcal/mol (1.6 kcal/mol after bugfixes described above). This number does not include the four weakly binding ligands whose affinity could not be measured; if we include estimates of affinity of these weak ligands compared to the calculated energies, the RMS error rises 2.3 kcal/mol (see Results). Although not entirely wrong, these predicted energies are too inaccurate for affinity progression in congeneric series, for instance. Rank-ordering of the entire test set by affinity was certainly unsuccessful. Finally, the lack of energetic accuracy was due in large part to the substantially overestimated affinities, which in turn can be traced to problems with convergence and to

an error in the long range van der Waals correction energy (below). The predictions of ligand orientations and alternate conformations were mixed; six of the nine ligands had a predicted pose with an RMSD of $\sim 2\text{\AA}$ or better, but three predictions were entirely wrong. Interestingly, eight of the nine ligands induced a protein conformational change in helix F that was not predicted, because of a failure to sample relevant protein motions.

These discrepancies between the theoretical predictions and the experimental results may have three sources, putting aside the chance that the experiments themselves may be wrong (for this study, based as it is on full binding affinity measurements and crystallography, the experimental observables seem reliable). The first sort of error is purely mechanical, relating to improper working of the computational methods or incorrect user choices. These are rarely observed in retrospective studies, as they can be caught and fixed before they are communicated, but in prospective studies, of the sort described here, they can appear; indeed, their occurrence can reflect how easily a method is, in practice, to perform. Second, there can be problems with the sampling of relevant states in the simulation, such as a particular ligand orientation or protein conformation, and convergence to the proper equilibrium distribution among these states, as defined by the force field (we distinguish between “sampling” and “convergence” for this reason). Third, the force field itself is an approximation, and even once the simulation has converged to free energy estimates that are “correct” according to the force field, these may still disagree with experiment. Force field errors can only be detected *after* results are converged, so before considering force field errors we assess the quality of convergence.

As it happens, technical and scripting problems, such as the bugs in scripts and simulation packages and tools, or even the choice of the wrong stereo-isomer from the initial poses calculated by DOCK 6²⁷, had a substantial impact; each led to specific

failures (n-phenylglycinonitrile, thiophene-2-carboxaldoxime). An error in the calculation of the long range van der Waals forces also resulted in too favorable binding free energies, in some cases by as much as 2-3 kcal/mol. Another avoidable error occurred during the calculation of ligand restraining energies; this is an involved, if technically important point, and so we take a short detour to explain it here.

Distance and orientational restraints are used to maintain the ligand inside the binding site as its interactions with the protein are decoupled in the free energy simulation. The first step in the alchemical transformation is to compute the free energy cost of applying these restraints, which owes to the loss of translational and rotational freedom (Methods & Supplemental Material). The cost of applying these restraints is calculated in steps of increasing restoring forces, the integration of which amounts to the total restraining free energy. For some calculations, the free energy of restraining the ligand to a single orientation (typically 1-2 kcal/mol) was unusually large (4-12 kcal/mol). Upon examining these simulations retrospectively, we found that particular ligands (for example 2-ethoxy-3,4-dihydro-2h-pyran) had erroneously begun their restraining simulations in an unfavorable, kinetically trapped orientation that was different from the target orientation of the restraints. Given the length of the simulations, this can present a serious convergence problem. The ligand may remain in that unfavorable orientation at steps in the restraining process where the restraints are not strong enough to pull the ligand across the kinetic barrier into the target orientation. As long as the ligand remains in the trapped orientation, the restoring forces are being integrated, at ever increasing levels, into the overall restraining energy. When the barrier is significant, a large restraining energy is thus calculated. Thus by misidentifying an unfavorable trap as a favorable thermodynamic state, these unconverged restraining energies ultimately make binding appear artificially too favorable. Finally, These give

some indication of how intricate these calculations remain, with many opportunities for the introduction of essentially trivial human error. Aggravating as this is, these errors can be traced to their origin and results typically improve when the errors are fixed.

In principle, an algorithmic origin of error comes from the calculation of ligand solvation energy, which might typically be laid at the door of the force field since sampling is less of a concern here. However, recent tests on computed hydration free energies of tens to hundreds of small molecules gave RMS errors to experiment that were typically in the 1.0 to 1.8 kcal/mol range, depending on compound polarity^{13,28-30}. For example, phenol, one of the ligands here, had a computed hydration free energy that differed by only 0.9 kcal/mol to that observed experimentally. This leads us to believe modeling of ligand interactions with bulk solvent is probably not the dominant source of error in this study.²⁹ Of course, force field errors may also play a role in ligand-protein interaction energies, but as we will argue in the following sections, these errors, to the extent that they occurred, were largely obscured by what we believe to be problems with sampling and convergence.

Protein Conformational Change, Sampling Protein Motions & Convergence of the Calculations

Rerunning the calculations starting with the holo structures was an attempt to determine whether or not the simulations reached convergence and if under-sampling of helix F led to errors in the predicted ΔG_b values. Several scenarios may be considered. When apo and holo results agree, it suggests that either there is no convergence problem or there is a cancellation of errors. When they disagree, and the holo result is more favorable, it suggests a sampling problem in the apo state, possibly related to protein conformational changes on binding, as these will typically make binding appear more

favorable^{2,24}. The case where the apo result is more favorable than the holo is counter-intuitive, but owing to poor choices in initial ligand orientations, and consequent convergence problems in the restraining energies (see above), this did in fact occur.

Because we knew that the helix motion had not been sampled over the course of the apo calculation, we assumed that the second case would predominate, that is, the holo result would be more favorable for the eight ligands that induced protein conformational change upon binding. Instead, the opposite was true, with the holo result *less* favorable than the apo result in nearly every case (Table 4 Supplementary Material). This suggests that under-sampling of helix F in the apo calculation was not the dominant cause of errors observed in the ΔG_b values; instead several other sources of convergence problems, including ligand restraints discussed above, were responsible for these errors (Fig. 3 Supplementary Material). Still, since F-helix closure was rarely observed in the course of the holo simulations following ligand decoupling as well, the energetic costs associated with breaking the helix during binding were poorly accounted for in these computed free energies; had it been, the holo results might have even been *less* favorable.

Despite being on the whole less favorable than the apo predictions, the holo results are less frequently overestimated, and are therefore in better overall agreement with the experimental binding free energies. Also, two of the ligands with the highest RMSD values to the crystallographic binding mode beginning from the apo prediction, nitrosobenzene and benzylacetate, were correctly predicted starting from the holo structure. This improvement reflects the prediction of a near-native orientation in the holo vs. an incorrect one in the apo pose prediction, and is consistent with large kinetic barriers separating different ligand binding modes^{3,24}. For 2-ethoxy-3,4-dihydro-2h-pyran, however, even starting from the holo structure, the correct ligand orientation was not generated during the setup stage of the calculation (discussed in detail in Methods),

and therefore was never sampled in the simulation. The end result is that the prediction for this ligand does not improve relative to the absolute binding free energy calculation from the apo structure.

Sampling the Ligand: Starting with the Wrong Orientation

Previously we had shown that decomposing configuration space by orientation, i.e. generating multiple starting orientations by docking and using simulations beginning from these to identify candidate binding modes, which we then consider separately, would allow adequate sampling (and also convergence) of ligand states separated by large kinetic barriers³. But this approach requires that docking generate candidate starting orientations which are close enough to the true binding mode that it can be found on nanosecond timescales. However, in L99A/M102Q a protein conformational change alters the size of the small binding site to allow larger ligands to bind and, in some cases, adversely affects the quality of the docking poses. This presents a problem for the alchemical free energy methods, which rely on having starting orientations that are reasonable. This proved to be the key to several of the mispredicted ligands in the apo calculations, most notably benzylacetate, which cannot be docked into the apo binding site with a good initial orientation (Fig. 3.A-B) and is a hard failure, both predicted pose and affinity are incorrect.

Relative Free Energy Calculations

We examined a different problem, looking for prediction of relative affinities within a congeneric series of phenol derivatives. Since we already knew affinity and structures for two such ligands, phenol and catechol, we began relative binding free energy calculations for six others beginning with both of these two, separately. We

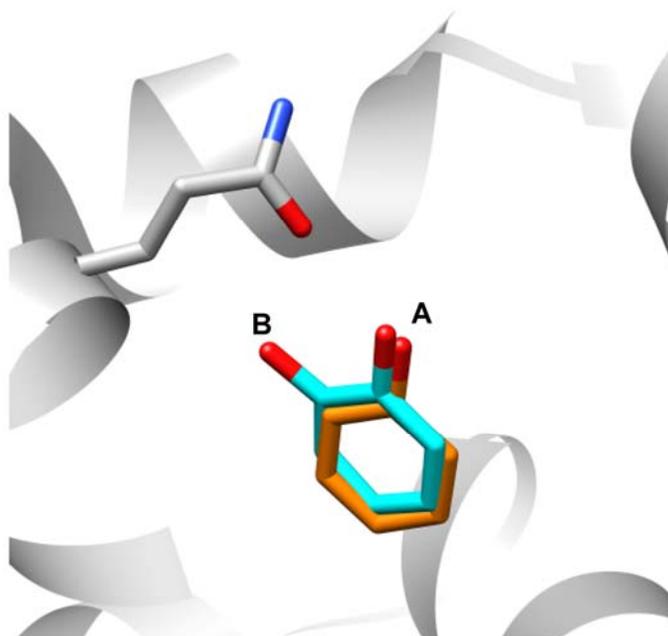
further simplified the experiment by restricting the starting orientations of the ligands to symmetric binding modes about the ligand hydroxyl group that contacts Gln102. Starting from catechol, these calculations did well (Table 2, Fig 5.A) for all six compounds. For the same six compounds, performance diminished when we began with phenol. How can these results be reconciled?

For the relative free energy predictions, it was necessary to generate one or more starting orientations of each new ligand. In the absolute calculations, we consider many candidate binding modes for each ligand, but this is not a typical for relative free energy calculations. We therefore assumed that the position of the hydroxyl would typically be preserved, but we were unsure of the location for additional substituents. For example, for 2-methylphenol overlaid onto phenol, the methyl group could be on the left or right side of the binding site while preserving the hydroxyl position. Because of kinetic barriers to rotation of ligands within the binding site we considered both (see Material & Methods).

The results of the relative calculations reflect this initial sampling of the ligand orientation. There are only two potential ligand starting orientations for phenol. Also, because of the restriction of the hydroxyl position, phenol starts with only one position for the hydroxyl group (Fig. 6). If the simulations could be run long enough to sample all ligand orientations, the starting orientations would not substantially affect the results. But timescales for interconversion of orientations in this binding site are slow, so the choice of starting orientations affects the results. Indeed, it happens that the position of the hydroxyl position in the derivatives is not conserved, in several cases adopting the ortho hydroxyl position in catechol, (Fig 6, B). Hence, for phenol the prospective predictions of the relative affinity are little better than random (Table 2).

By the same logic, the performance of the catechol transformation is substantially better. Catechol has two hydroxyl groups, which translates to four starting orientations for the ligand, two reflections around each hydroxyl axis. Relative to phenol, the sampling of the hydroxyl position and, consequently, the ortho substituent, is better. The result is that one starts with

Figure 6. Crystallographic orientations of the reference ligands phenol (orange; PDB ID 1LI2) and catechol (cyan; PDB ID 1XEP) overlaid on the apo reference structure (gray; PDB ID 1LGU). The two alternate hydroxyl positions are labeled A and B.



reasonable sampling of ligand orientations, improving the predictions. Besides accurately predicting the relative free energy change (RMS error 1.1 kcal/mol), the compounds are also correctly rank-ordered, with the exception of 5-chloro-2-methylphenol. Pose fidelity to the crystallographic binding mode is high, and in three cases alternate orientations were correctly predicted. 5-chloro-2-methylphenol is the caveat to this set of

transformations; its pose is accurately predicted (RMSD of 0.9Å), but the relative free energy is overestimated, possibly due to force field issues. For this ligand, and indeed all the free energy results, a good starting orientation is necessary but not sufficient for accurate prediction of both affinity and pose.

3.5 Conclusion

Predicting the absolute free energy of binding remains challenging, even in a simple model system. There were several technical issues that arose during the study (incorrectly assigned ligand parameters, unconverged restraining energies, errors with the calculation of long range dispersion forces) due to either the complexities of the calculations, bugs, or human error, highlighting the difficulty of this work. Still these sorts of errors are entirely avoidable. Conversely, problems with convergence and sampling of ligand and receptor configurations present more fundamental problems for the field, and in this study prevented even rank ordering of ligands by affinity.

Both the absolute and the relative results highlight the importance of predicting a near-native orientation for the ligand in advance, rather than expecting the ligand to find its optimal orientation through dynamics. The additional polarity in this binding site appears to have introduced additional roughness to the energy landscape, relative to the apolar cavity, exacerbating this problem. The need to start with a near-native orientation may have contributed to the improvement in relative results using catechol as a reference instead of phenol. In absolute calculations, better algorithms are needed to define unique ligand orientations to avoid convergence problems in restraining calculations and overestimation of binding affinities. This may be addressable in the future through kinetics-based partitioning of ligand orientations.

Undersampling of protein conformational changes also caused several failures. The energy barriers to these changes proved too large to be overcome on timescales reasonable for alchemical calculations, and at times the protein reacted to oversized ligands by forcing them out of the binding site rather than expanding to accommodate them. In principle, these conformational changes can be sampled using our confine-and-release approach that has been successful in sampling side-chain rotations. However, this requires that the range of possible backbone motions be identified in advance, and sampling these is much more difficult than side chain torsions. Addressing this will be a goal of future work as well.

What emerges are results that, even for simple ligands in a simple binding site, are less convincing than what might have been hoped for based on our initial study with the apolar cavity² and also based on over 20 years of development of these techniques. Still, the method also had substantial successes: 11 of 14 compounds were correctly predicted from the absolute binding free energy calculations, the geometries of all but three of the ligands from this set were predicted better than $\sim 2\text{\AA}$, subtle contributions to the binding free energy were correctly predicted – multiple binding modes and internal hydrogen bonding, and when the right configurations are captured, as was the case for relative affinities based on catechol, true rank ordering can be achieved. We therefore believe that substantial progress can be made with alchemical free energy methods. To show that these methods are pragmatic, and to illuminate opportunities for their optimization, it remains important to couple theoretical developments with genuine prospective prediction, in sites that allow detailed experiment and analysis on multiple ligands.

3.6 Material & Methods

Software

In simulations and setup, we drew on several different software packages; we specify the versions here and refer to them simply by name later. We used version 3.3.1 of GROMACS^{31,32}. We also used version 1.2.7 of the Antechamber package³³. We used version 6.0 of the DOCK software package³⁴ for docking. We also used OpenEye's OEChem toolkit (Python implementation), version 1.3.33.

Protein/system preparation and simulation parameters/protocols

System preparation and parameters always followed the same procedure, although several starting structures were used. Unless otherwise specified, the starting protein structure was the *apo* conformation of L99A/M102Q (PDB: 1LGU)²³. Setup began with the protein PDB structure and the protein setup procedure was the same regardless of the starting structure. Our setup procedure for the protein and the simulation parameters/protocols were essentially identical to our previous study (see Supplementary Methods)².

Docking

We used DOCK 6.2 to generate candidate bound orientations for each potential ligand for absolute free energy calculations²⁷. Docking was done on protein structures that had been pre-equilibrated in water for 1 ns as discussed above, after then stripping hydrogens and waters. Docking poses were clustered by RMSD by DOCK. We manually examined the resulting clusters and retained non-redundant orientations in the binding site for further simulations. Symmetric orientations were not retained, nor were clusters that placed the ligand outside of the cavity. For simulations beginning with the holo

protein structure, we used docking to generate ligand orientations, but manually selected the one that best corresponded to the crystallographic ligand orientation

Absolute free energy calculations

Basic strategy: Absolute binding free energy calculations were done beginning from both the unbound structure (1LGU) and a variety of different holo structures as described previously²; these calculations followed unrestrained simulations to identify reference orientations for restraints (details given in Supplementary Material).

Unrestrained simulations and selecting reference orientations

Prior to performing binding free energy studies, candidate ligand orientations were identified in the binding site. For the absolute binding free energy studies, potential bound orientations resulting from docking were used as input for unrestrained simulations^{2,3}. This number ranged from two to seven different potential orientations. One unrestrained simulation was run beginning from each of these (see Supplementary Material).

After the unrestrained simulations, the value for each of the six degrees of freedom at every stored snapshot over the course of the simulation was computed (snapshots were stored every 1 ps), and histograms of each degree of freedom were made, as well as plots of the timeseries of each degree of freedom. To select reference orientations for restraint, we manually examined the histograms and timeseries. For orientations that did not interconvert, we defined the reference orientation by restraining each degree of freedom to its peak histogram value; redundant orientations were discarded. When a ligand interconverted between several orientations over the course of the simulation, the most populated orientation was retained. Other orientations were

retained only if they rarely interconverted with the most populated orientation (i.e. less than three times over the course of the simulation). In a few cases two (slowly interconverting) orientations were identified from a single unrestrained simulation. This procedure resulted in one or more reference orientations to restrain each ligand to (typically 2-4).

Long Range Dispersion Corrections

Long-range dispersion interactions between ligands and proteins can contribute to the binding affinity, even at distances larger than 0.7 to 0.9 nm³⁵. For lysozyme ligands this contribution is typically in the 0.4 to 1.0 kcal/mol range, though it can be larger with larger ligands. Though a contribution of 0.4 to 1.0 kcal/mol is not that large, including these corrections is also important to ensure consistency across research groups using different protocols³⁵. Therefore, we used a reweighting strategy as in our previous study^{2,35}. Initially, we used the WHAM-LR approach³⁵, but occasionally this gave large uncertainties or otherwise questionable results; in these instances we also applied the EXP-LR approach and checked for consistency, using the value that appeared more reasonable. Following the predictions, we discovered that a bug in GROMACS resulted in incorrect LR corrections for the WHAM-LR approach, and a scripting error resulted in a sign error for the EXP-LR approach. All of the LR corrections were recalculated and EXP-LR was used throughout, except for 3-chlorophenol, where WHAM-LR was used instead.

In some systems, very long range electrostatic interactions can also play an important role³⁶, but these are unlikely to be important for these largely apolar ligands in this predominantly apolar cavity.

Oriental decomposition

In the absolute binding free energy calculations, we computed contributions of different metastable orientations for ligands separately, and then combined the effective binding free energies for these different orientations into a total binding free energy as described previously (see Supplementary Material)^{2,3}.

Confine-and-Release Approach

The contribution to the binding free energy of slow sidechain conformational changes in the binding site was calculated using the confine-and-release approach, as described previously^{2,24}. Here we included three rotamers for umbrella sampling and the confine-and-release approach, Val111, Val103 and Leu118 (see Supplementary Material). The sidechains of these residues are reasonably distant from one another, so we assumed that couplings between the sidechain conformations were negligible and performed umbrella sampling simulations for each sidechain independently, in the presence and absence of each ligand. These results were used in simulations beginning from the *apo* structures.

Data analysis

Data analysis and plots were done with custom Python scripts based on the matplotlib and Numarray libraries. Free energies were computed using the Bennett Acceptance Ratio (BAR)^{37,38} approach as done previously^{2,3,24}. Uncertainties were computed using the block bootstrap method – breaking the simulations up into blocks the length of the autocorrelation time of the potential energy, selecting random blocks and composing new timeseries of the same length as the original, and computing a new estimated free energy. We repeated this process 40 times and took the uncertainty as the

standard error over these trials, as in our previous work². Long range dispersion corrections were analyzed as described previously^{2,35}.

Water removal

To address the possibility that waters might occupy the cavity in the absence of ligand, we used absolute binding free energy techniques to compute the free energy of removing waters from the binding site (Supplementary Material). The results indicated that removal of water was always favorable; thus, it was unnecessary to include a free energy correction for desolvation of the binding site.

Relative free energy calculations

We performed two sets of relative free energy calculations, one using phenol and one using catechol, to calculate differences in binding free energies between a reference compound and a series of phenol derivatives. Mutations were also done between catechol and phenol. This provided redundant information to use in assessing error.

The calculations began using the unbound (1LGU) protein structure, with the reference ligand (catechol or phenol) in an orientation roughly corresponding to that in the co-crystal structure (Fig 6). This was a simulation snapshot from an unrestrained simulation of the ligand in the binding site and was manually compared with the crystal structure.

To set up the relative free energy calculations, we used a maximal common substructure search to identify common atoms between the reference and target ligands. Deletion of atoms is the most difficult part of the transition, so to minimize the number of deletions we mutated both the reference and target ligand into a scaffold resembling the

maximal common substructure rather than directly into one another. The scaffold had the same atoms as the maximal common substructure; for example, the maximal common substructure for the transformation of phenol to catechol is phenol, with one hydrogen atom removed. In the scaffold, partial charges on shared atoms are taken to be the average of those for the overlaid shared atoms. After this averaging, an equal adjustment is applied to all shared atoms to ensure net neutrality of the ligand. A similar averaging approach is applied for bonded and Lennard-Jones interactions involving shared atoms.

Free energy calculations for the transition of each ligand to the scaffold are carried out in two steps. First, there is a charging calculation, where charges on all deleted atoms are turned to zero, and charges on shared atoms are mutated to match those in the scaffold. This is followed by a van der Waals calculation, where any atoms being deleted are turned into dummy atoms (with Lennard-Jones interactions turned to zero) and adjustments are made to Lennard-Jones parameters for shared atoms. In this step the bonded interactions are modified to match the scaffold. Soft core potentials are used for the van der Waals calculation but not the charging calculation, and the two calculations are done separately for the reasons discussed elsewhere (<http://www.alchemistry.org>). Each relative binding free energy calculation involves mutating each ligand to the scaffold once in the protein and once in the binding site (Fig. 2 Supplementary Material).

To generate potential bound structures, we began with the structure of the reference ligand bound, and then overlaid the target ligand onto the reference ligand manually in Pymol; consequently, the aromatic ring remains in-plane. For simplicity, we chose to preserve the position of the hydroxyl (which hydrogen bonds with the protein). For the transition from phenol this still left two possible starting orientations, and for catechol four possible orientations (two for each hydroxyl group).

Since we began with one bound structure for each reference ligand, and 2-4 for each target ligand, each relative free energy calculation involved mutating the reference ligand to the scaffold just once, and each target ligand to the scaffold once per orientation. In the results section, we report a single relative free energy for each reference-to-target transformation. For this value, we combine the contributions of different orientations as we did previously in absolute free energy calculations³, but without any discarding of snapshots (see Supplementary Material). The van der Waals and charging components of the relative free energy calculations were done using the same protocol, but unlike the absolute calculations, these used no restraints.

Experimental measurements of ligand affinity: Binding detection by Tm-upshift monitoring folded to unfolded CD spectra

The thermal denaturation experiments, using a Jasco J-715 spectropolarimeter with a Jasco PTC-348WI Peltier-effect in-cell temperature control device and in-cell stirring, were performed as described^{2,39}. Transition from folded to unfolded protein in the presence of ligand was monitored by circular dichroism at a wavelength of 223 to 233.5nm, depending on the absorbance of the ligand in the same region of interest. Whereas 233.5 is far from the helical maximum, helicity may still be measured reliably here. The Tm and van't hoff ΔH were analyzed using the program EXAM⁴⁰. Compounds were assayed in their neutral form at concentrations ranging from 0.5 to 10mM, solubility permitting. Nitrosobenzene, n-phenylglycinonitrile, 2-nitrothiophene, 4-chloro-1H-pyrazole, 4, 5, 6, 7-tetrahydroindole, 2-ethoxy-3,4-dihydro-2h-pyran, 1-2-hydroxyethylpyrrole, 2-ethoxyphenol, 2-ethylphenol and 5-chloro-2-methylphenol were assayed at pH 3.07 in a 17mM KH₂PO₄, 25mM KCl, 2.9mM phosphoric acid buffer³⁹. Benzylacetate, 1-phenylsemicarbazide, o-benzylhydroxylamine hydrochloride,

thiophene-2-carboxaldoxime and thieno[3,2-b]thiophene were assayed at pH 6.8 in a 50mM KPi, 200mM KCl, 38% ethylene glycol buffer³⁹. Phenylhydrazine and benzylacetate were also assayed at pH 5.4 in a 8.6mM sodium acetate, 100mM NaCl, 1.6mM acetic acid buffer³⁹.

ITC measurements

Isothermal titration calorimetry data was obtained using a Microcal VP-ITC model calorimeter⁴¹. Experiments were performed at 10°C². Ligands were assayed at pH 6.8 in a 50mM KPi, 200mM KCl buffer. Protein concentrations ranged from 0.03mM to 0.06mM.

Protein production and Structure determination

T4 lysozyme L99A/M102Q was overexpressed and purified as previously described⁴². Crystals belonging to space group P3₂2₁ were grown at pH 6.5 in a 2.4M potassium phosphate buffer. Compounds were soaked into crystals for 1 hour to 2 days. Thiophene-2-carboxaldoxime and 4-chloro-1h-pyrazole were soaked at 50mM. 2-methylphenol was soaked at 20mM; 2-propylphenol and 2-ethylphenol at ~1mM. 2-methoxyphenol, 5-chloro-2-methylpheno and 2-ethoxyphenol were soaked at saturating conditions (<50mM). Benzylacetate was soaked at 100mM for 2 days. Due to insolubility in aqueous buffer, the following ligand soaks were performed at a concentration of 50mM in the cryoprotectant mixture of 50:50 mineral oil to Paratone-N (Hampton Research, Aliso Viejo, CA): thieno[3,2-b]thiophene, nitrosobenzene, 4,5,6,7-tetrahydroindole, 2-ethoxy-3,4-dihydro-2h-pyran and 1,2-hydroxyethylpyrrole. X-ray diffraction data was measured at 100K with an ADSC-CCD detector on Beamline 8.3.1 of the Advanced Light Source (ALS) at Lawrence Berkley National Laboratory. The data

were processed using the HKL⁴³ software package with the exception of 3HTF, 3HTD, 3HTG, 3HUQ, 3HTB, 3HU8, 3HT9, which were processed with XDS⁴⁴. Complex structure refinement was performed with Refmac5⁴⁵. Model building and placement of waters was done using Coot⁴⁶. Ligand parameters were generated using PRODRG⁴⁷.

ACCESSION NUMBERS

Coordinates and structure factors have been deposited in the Protein Data Bank with accession numbers: 3HUQ, 3HUA, 3HUK, 3HU9, 3HTG, 3HTF, 3HTB, 3HT8, 3HT7, 3HT6, 3HU8, 3HT9, 3HTD.

3.7 Acknowledgements

We thank Michael Mysinger for reading the manuscript. This work was supported by NIH grants GM59957 (to BKS) and GM034993 & GM063592 (to KAD), and an award from the Sandler Family Trust (to KAD), and by the Louisiana Optical Network Initiative Institute (DLM), supported by the Louisiana Board of Regents Post-Katrina Support Fund Initiative grant LEQSF(2007-12)-ENH-PKSFI-PRS-01. This work was performed in part with the UCSF QB3 Shared Computing Facility.

3.8 References

1. Jayachandran, G., Shirts, M. R., Park, S. & Pande, V. S. (2006). Parallelized-over-parts computation of absolute binding free energy with docking and molecular dynamics. *J. Chem. Phys.* **125**, 084901.
2. Mobley, D. L., Graves, A. P., Chodera, J. D., McReynolds, A. C., Shoichet, B. K. & Dill, K. A. (2007). Predicting absolute ligand binding free energies to a simple model site. *J. Mol. Biol.* **371**, 1118-1134.
3. Mobley, D. L., Chodera, J. D. & Dill, K. A. (2006). On the use of orientational restraints and symmetry corrections in alchemical free energy calculations. *J. Chem. Phys.* **125**, 084902.
4. Warshel, A., Sussman, F. & King, G. (1986). Free energy of charges in solvated proteins: microscopic calculations using a reversible charging process. *Biochemistry* **25**, 8368-8372, ISSN 0006-2960, 1986.
5. Deng, Y. & Roux, B. (2006). Calculation of standard binding free energies: Aromatic molecules in the T4 lysozyme L99A mutant. *J Chem Theory Comput* **5**, 1255-1273.
6. Wang, J., Deng, Y. & Roux, B. (2006). Absolute Binding Free Energy Calculations Using Molecular Dynamics Simulations with Restraining Potentials. *Biophysical Journal* **91**, 2798 -814.
7. Bash, P. A., Singh, U. C., Brown, F. K., Langridge, R. & Kollman, P. A. (1986). Calculation of the Relative Change in Binding Free Energy of a Protein-Inhibitor Complex *Science* **235**, 574-576.

8. Wong, C. F. & McCammon, A. J. (1986). Dynamics and design of enzymes and inhibitors. *J. Am. Chem. Soc.* **108**, 3830-3832.
9. Hermans, J. & Subramaniam, S. (1986). The Free Energy of Xenon Binding to Myoglobin from Molecular Dynamics Simulation. *Isr. J. Chem.* **27**, 225-227.
10. Chang, C. E. & Gilson, M. K. (2004). Free energy, entropy, and induced fit in host-guest recognition: calculations with the second-generation mining minima algorithm. *J. Am. Chem. Soc.* **126**, 13156-64.
11. Rekharsky, M. V., Mori, T., Yang, C., Ko, Y. H., Selvapalam, N., Kim, H., Sobransingh, D., Kaifer, A. E., Liu, S., Isaacs, L., Chen, W., Moghaddam, S., Gilson, M. K., Kim, K. & Inoue, Y. (2007). A synthetic host-guest system achieves avidin-biotin affinity by overcoming enthalpy-entropy compensation. *Proc. Natl. Acad. Sci. U S A* **104**, 20737-42.
12. Moghaddam, S., Inoue, Y. & Gilson, M. K. (2009). Host-guest complexes with protein-ligand-like affinities: computational analysis and design. *J. Am. Chem. Soc.* **131**, 4012-21.
13. Nicholls, A., Mobley, D. L., Guthrie, P. J., Chodera, J. D. & Pande, V. S. (2008). Predicting Small-Molecule Solvation Free Energies: An Informal Blind Test for Computational Chemistry. *J. Med. Chem.*
14. Kamerlin, S. C., Haranczyk, M. & Warshel, A. (2009). Are mixed explicit/implicit solvation models reliable for studying phosphate hydrolysis? A comparative study of continuum, explicit and mixed solvation models. *Chem Phys Chem* **10**, 1125-34.
15. Lee, F. S., Chu, Z. T., Bolger, M. B. & Warshel, A. (1992). Calculations of antibody-antigen interactions: microscopic and semi-microscopic evaluation of

- the free energies of binding of phosphorylcholine analogs to McPC603. *Protein Eng.* **5**, 215-28.
16. Rosta, E., Haranczyk, M., Chu, Z. T. & Warshel, A. (2008). Accelerating QM/MM free energy calculations: representing the surroundings by an updated mean charge distribution. *2008* **112**, 5680-92.
 17. Hermans, J. & Wang, L. (1997). Inclusion of the Loss of Translational and Rotational Freedom in Theoretical Estimates of Free Energies of Binding. Application to a Complex of Benzene and Mutant {T4 Lysozyme}. *J. Am. Chem. Soc.* **119**, 2707-2714.
 18. Mann, G. & Hermans, J. (2000). Modeling Protein-small Molecule Interactions: Structure and Thermodynamics of Noble Gases Binding in a Cavity in Mutant Phage T4 Lysozyme L99A. *J. Mol. Biol.* **302**, 979-989.
 19. Warshel, A., Sharma, P., Chu, Z. & Aqvist, J. (2007). Electrostatic Contributions to Binding of Transition State Analogues Can Be Very Different from the Corresponding Contributions to Catalysis: Phenolates Binding to the Oxyanion Hole of Ketosteroid Isomerase. *Biochemistry* **46**, 1466-1476.
 20. Kim, J. T., Hamilton, A. D., Bailey, C. M., Domoal, R. A., Wang, L., Anderson, K. S. & Jorgensen, W. L. (2006). FEP-guided selection of bicyclic heterocycles in lead optimization for non-nucleoside inhibitors of HIV-1 reverse transcriptase. *J. Am. Chem. Soc.* **128**, 15372-15373.
 21. Graves, A. P., Brenk, R. & Shoichet, B. K. (2005). Decoys for Docking. *J. Med. Chem.* **48**, 3714 - 372.
 22. Wei, B. Q., Weaver, L., Ferrari, A. M., Matthews, B. M. & Shoichet, B. K. (2004). Testing a flexible-receptor docking algorithm in a model binding site. *J. Mol. Biol.* **337**, 1161-82.

23. Wei, B. Q., Baase, W. A., Weaver, L. H., Matthews, B. W. & Shoichet, B. K. (2002). A model binding site for testing scoring functions in molecular docking. *J. Mol. Biol.* **322**, 339-355.
24. Mobley, D. L., Chodera, J. D. & Dill, K. A. (2007). Confine-and-release method: Obtaining correct binding free energies in the presence of protein conformational change. *J. Chem. Theory Comput.* **3**, 1231-1235.
25. Bayly, C. I. (2009). (Mobley, D., ed.).
26. Hann, M. M., Leach, A. R. & Harper, G. (2001). Molecular Complexity and Its Impact on the Probability of Finding Leads for Drug Discovery. *J. Chem. Inf. Comput. Sci.* **41**, 856-864.
27. Lang, P. T., Brozell, S. R., Mukherjee, S., Pettersen, E. F., Meng, E. C., Thomas, V., Rizzo, R. C., Case, D. A., James, T. L. & Kuntz, I. D. (2009). DOCK 6: Combining techniques to model RNA-small molecule complexes. *RNA* **15**.
28. Mobley, D. L., Bayly, C. I., Cooper, M. D. & Dill, K. A. (2009). Predictions of hydration free energies from all-atom molecular dynamics simulations. *J. Phys. Chem. B* **113**, 4533-4537.
29. Mobley, D. L., Bayly, C. I., Cooper, M. D., Shirts, M. R. & Dill, K. A. (2009). Small molecule hydration free energies in explicit solvent: An extensive test of fixed-charge atomistic simulations. *J. Chem. Theory Comput.*
30. Mobley, D. L., Dumont, I., Chodera, J. D. & Dill, K. A. (2007). Comparison of charge models for fixed-charge force fields: Small-molecule hydration free energies in explicit solvent. *J. Phys. Chem. B* **111**, 2242-2254.
31. van der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A. E. & Berendsen, H. J. C. (2005). GROMACS: Fast, flexible, and free. *J. Comp. Chem.* **26**, 1701-1718.

32. Lindahl, E., Hess, B. & van der Spoel, D. (2001). GROMACS 3.0: A Package for Molecular Simulation and Trajectory Analysis. *J. Mol. Modeling* **7**, 306-317.
33. Wang, J., Wang, W., Kollman, P. A. & Case, D. A. (2006). Automatic atom type and bond type perception in molecular mechanical calculations. *J Mol Graphics Model* **25**, 247-260.
34. Moustakas, D. T., Lang, P. T., Pegg, S., Pettersen, E., Kuntz, I. D., Brooijmans, N. & Rizzo, R. C. (2006). Development and validation of a modular, extensible docking program: DOCK 5. *J. Comput. Aided Mol. Des.* **20**, 601-619.
35. Shirts, M. R., Mobley, D. L., Chodera, J. D. & Pande, V. S. (2007). Accurate and Efficient Corrections for Missing Dispersion Interactions in Molecular Simulations. *J. Phys. Chem. B* **111**, 13052-13063.
36. Warshel, A., Sharma, P., Kato, M. & Parson, W. (2006). Modeling electrostatic effects in proteins. *Biochim Biophys Acta.* **1764**, 1647-76.
37. Bennett, C. H. (1976). Efficient Estimation of Free Energy Differences from Monte Carlo Data. *J. Comp. Phys.* **22**, 245-268.
38. Shirts, M. R., Bair, E., Hooker, G. & Pande, V. S. (2003). Equilibrium free energies from nonequilibrium measurements using maximum-likelihood methods. *Phys. Rev. Letters* **91**, 140601.
39. Morton, A. & Matthews, B. W. (1995). Specificity of ligand binding in a buried non-polar cavity of T4 lysozyme: linkage of dynamics and structural plasticity. *Biochemistry* **34**, 8576–8588.
40. Kirchhoff, W. (1993). EXAM: A Two-State Thermodynamic Analysis Program. In *NIST*, Gaithersburg, Maryland
41. Plotnikov, V. V., Brandts, J. F. & Lin, L. N. (1997). A new ultrasensitive scanning calorimeter. *Anal. Biochem.* **250**, 237–244.

42. Poteete, A. R., Dao-Pin, S., Nicholson, H. & Matthews, B. W. (1991). Second-site revertants of an inactive t4 lysozyme mutant restore activity by restructuring the active site cleft. *Biochemistry* **30**, 1425–1432.
43. Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326.
44. Kabsch, W. (1993). Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. Appl. Cryst.* **26**, 795-800.
45. Murshudov, G. N., Vagin, A. A. & Dobson, E. J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr., D Biol. Crystallogr* **53**, 240–255.
46. Emsley, P. & Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr., D Biol. Crystallogr* **60**, 2126–2132.
47. Schuettelkopf, A. W. & van Aalten, D. M. F. (2004). PRODRG - a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallographica* **D60**, 1355--1363.

Gloss to Chapter 4: Future Directions

Chapter 4 pertains to a project that has spanned four years of my work in the Shoichet lab. Much of the work described is preliminary, but the model system that is introduced in the following chapter is unique in that it is an open cavity, contacting bulk solvent at one end and at the other is a buried charge. In between is a narrow channel lined with carbonyl groups and filled with eight ordered waters that span the length of the cavity, contacting one another and the channel walls. This cavity is derived from the anionic CCP W191G cavity, but the mutation that opened the cavity, on one hand adding another complication to an already complicated model system, also eliminated the flexible loop. The trade off then is an open and wet but fairly rigid cavity; one that is extremely relevant for molecular docking by beginning to bridge the gap between the closed model systems and true drug targets.

Chapter 4: Future Directions

CCP W191G gateless model binding site: charged, polar, and open to bulk solvent

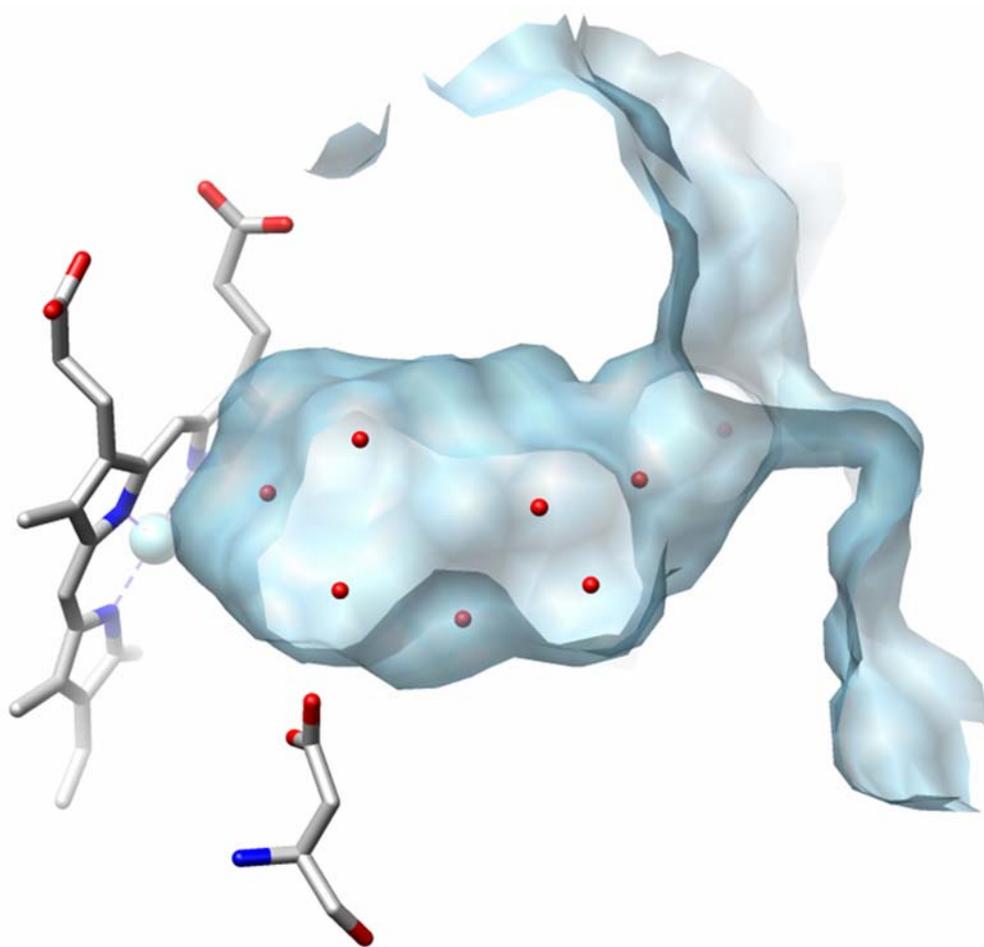


Figure 1. The open cavity mutant, CCP W191G-GA. Shown in blue, the cut-away surface of the apo cavity with a hydrogen bonding network of eight ordered water molecules (red) out to bulk solvent. Aspartate233 is shown in the lower left.

4.1 Abstract

There are several critical questions in molecular docking which have not been possible to address with the current closed cavity model systems: interaction/displacement with ordered waters, contact with bulk solvent, a higher internal cavity dielectric and the resulting effect on ligand binding. Here I introduce a new open model cavity, CCP W191G-gateless (CCP-GA), an anionic cavity that contains both ordered waters and an interface with bulk solvent. The results shown are very preliminary, but they confirm our initial suspicions that this cavity will be a vital intermediary between the closed cavities and true drug targets. Preliminary docking results indicate that larger molecules dominate the hit list, resulting in a decrease in the enrichment of known actives (small by comparison). Neutral molecules show up with greater frequency in the dock hit lists as well, and 3 neutral ligands have been shown to bind to this cavity. More importantly, they appear to do so only when the critical aspartate233 is charged. Crystal structures of the neutral ligands bound to the cavity have been solved and the observed binding modes also confirm their interaction with Asp233. In addition, this cavity has been used to evaluate both our current ligand desolvation and a new desolvation method (SEV) developed by Michael Mysinger. Several molecules have been selected for prospective testing, and although the assays are not finished, it appears that some of these ligands do in fact bind; these larger ligands have the potential to displace individual *and* different ordered waters within the cavity, paving the way to evaluate the benefit of including a penalty for ordered water displacement during flexible docking with waters.

Abbreviations

Proximal or open cavity mutant CCP W191G: CCP W191G-gateless or CCP-GA

CCP W191G: CCP wt* or wt*

CCP W191G/D235N: D235N

DUD: A Directory of Useful Decoys

4.2 Introduction

The closed cavity model systems developed in T4 lysozyme and CCP have proven to be extremely useful at isolating particular weaknesses within our DOCK scoring function and others¹⁻⁴; in Chapter 1, CCP W191G was used to look at the subtle balance between ligand desolvation and the electrostatic interaction energy between the ligand and protein in a system where monocations bind, but dications and neutrals do not. Both these model systems have also been invaluable at evaluating higher levels of theory discussed in Chapters 2 & 3.^{5,6} However, as closed cavities their greatest strength, isolating individual components of molecular recognition, is also a weakness; these cavities are far from resembling actual protein binding sites, which are wet, contain many properties that contribute to ligand binding (charge interactions, hydrogen bonding, hydrophobic burial, entropic effects), and also are open to bulk solvent. To begin to tackle some of these complicating factors requires an intermediary binding site, still simple enough so that only a few degrees of freedom are addressed at one time and also still easily testable, both retrospectively and prospectively.

The proximal cavity mutant in CCP-W191G, created by the mutations P190→G, ΔG192-A193, was initially designed by the Goodin lab to test the hypothesis that this loop region was the electron transfer pathway from cytochrome c.⁷ Fortuitously, the shortened loop created a cavity with similar properties to the original anionic cavity, but also a constitutively open pocket with a channel of ordered waters extending from the original pocket (Fig 1) all the way to the cavity entrance, the interface with bulk solvent. This new “open” cavity bridges the gap between the closed apolar, polar and anionic cavities with real protein binding sites, and allows us to ask new specific questions.

Will the higher internal dielectric of the cavity and larger binding site alter the type of ligands that bind to this cavity and their affinity compared to the wt*? Several consequences can be imagined when enlarging and opening the cavity to bulk solvent. Larger ligands may bind to this cavity, but will they dominate the docking hit lists? If we assume that the internal dielectric of the cavity is higher, will neutral ligands bind to this cavity? And a related question, will Asp233, remain charged or will it be neutral? If these are the most likely changes we predict, what challenges do they present to the DOCK scoring function? A bias towards larger ligands might imply a van der Waals bias in DOCK. If neutral ligands do bind, will DOCK track this change? And if they bind better than to the wt* cavity, will we successfully capture this behavior in the scoring function? What follows are preliminary results that attempt to begin to address these questions.

4.3 Preliminary Results and Discussion

Larger molecules dominate the docking hit lists

The subset37 ZINC database⁸ of ~650,000 molecules, itself a superset of the ZINC fragment database with the only requirement that molecular weight be between 250 and 40, was docked to CCP-GA and for comparison to the wt* cavity. Enrichment was calculated for the known actives for the wt* and decoys (Fig 2); the actives and decoys for the wt* were considered to be the same for the CCP-GA cavity as well, based on data for a subset of wt* ligands from Goodin *et al* that were shown to bind to CCP-GA.⁷ Unlike the high enrichment of actives observed for CCP wt*, CCP-GA experienced a steep drop in overall enrichment of actives; whereas 40% of known compounds were found at 0.1% of the database for CCP wt*, enrichment of actives did not start for CCP-

GA until the 1% mark, and many ligands were not found until the very bottom of the database.

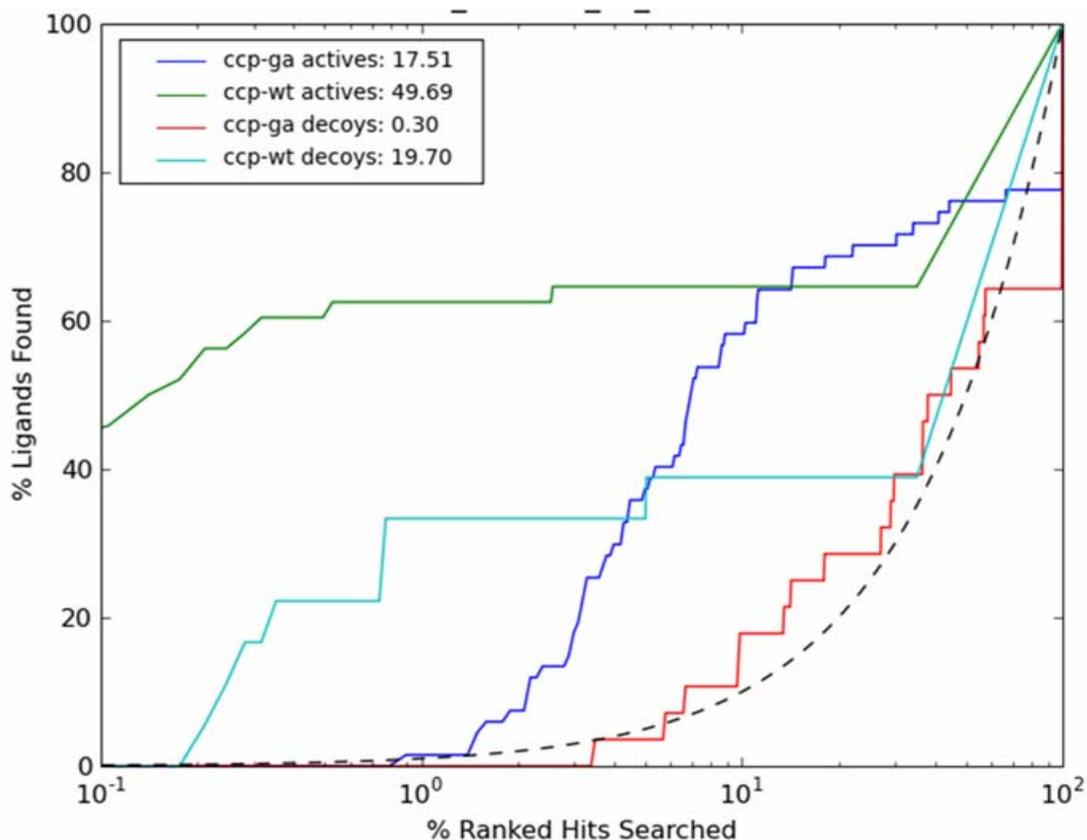


Figure 2. Comparison of enrichment of actives and decoys in CCP wt* (actives: green, decoys: cyan) and CCP-GA (actives: blue, decoys: red).

Examining the top 100 hits for CCP-GA it became apparent that larger compounds dominated. Whereas the average size of currently known actives for CCP-GA is between 6-9 heavy atoms, most of the top 100 hits had > 11 heavy atoms. From a docking and scoring standpoint this makes sense, as long as the molecule fits it will receive a favorable van der Waals score. Larger ligands will, not surprisingly, have the advantage in a cavity that can accommodate them. Perhaps especially in an anionic

cavity that binds monocations, larger +1 charged molecules pay essentially the same desolvation penalty as a smaller charged ligand, but make this energy back in the van der Waals score.

What was interesting, however, was that a larger percentage of these ligands than anticipated, although very polar, were uncharged. We turned at this point a new more accurate method to estimate ligand desolvation developed within the lab by Michael Mysinger to see if we could explain this unexpected result.

Comparing ligand desolvation methods in CCP-GA

In the closed cavity model systems, typically the cost of desolvating a ligand upon binding is taken as the full desolvation penalty, because the ligand is totally enclosed by the protein once it binds. In CCP-GA, because the cavity is open at one end, we decided to switch to using volume-based desolvation², typically used in docking campaigns in our lab. The volume based desolvation is then a percentage of the full desolvation, estimated by using a probe molecule (typically with a radius of 1.4Å) to determine the burial of the ligand atoms by the protein. We knew there were errors in this method; using a probe radius of 1.4Å led to gaps between ligand and protein too small to fit a molecule of water, yet these remain high dielectric. Taking it one step further, using a probe radius of 0.0Å eliminated the gaps by making all space between the protein and ligand molecular surface high dielectric; this helped improve enrichment in some charged systems in the DUD set^{9,10}, but in CCP wt* this “quick fix”—meant to alleviate a high desolvation penalty for charged molecules—exacerbated the problem. Dications began to show up in CCP wt* hit lists. The SEV or Solvent Excluded Volume method (Mysinger, unpublished) calculates the volume based desolvation and includes the re-entrant surface, eliminating these high dielectric gaps.

The new SEV method had already tested on the DUD test set where it appeared to make large improvement in enrichments⁹; however, the CCP-GA cavity offered the first chance to both retrospectively *and* prospectively test this method. The results at this stage are preliminary, but still surprising. While enrichment of actives greatly improves using SEV solvation by comparison to volume based and full desolvation methods (Fig 3), there were two outcomes that are worth noting. First, the number of neutral molecules in the top of the hit list increased with SEV solvation. This would be consistent with charged ligands paying a greater penalty for desolvation, and our hypothesis that in an open cavity neutral molecules had a greater chance of being true ligands for this binding site, but as yet we were unsure whether experimental results would confirm this prediction. The second, and more surprising, result was that the poses for known ligands did not improve; in many cases the charge still did not interact with Asp233. This was true for most of the charged compounds in the hit list. The charged moiety preferred instead to interact with backbone carbonyls lining the walls of the cavity. The cause for this last observation is still unresolved, although our current thought is that it may have to do with the distribution of partial charges on the ligands; this still bears testing out. A selection of neutral and charged, large and small compounds from the two hit lists, volume based desolvation and SEV, have been purchased and are in the beginning stages of testing. However, the question of whether neutral molecules in general would bind to CCP-GA has progressed further and the answer is both encouraging and perplexing.

Neutral ligands bind to CCP-GA

Before we even began docking to CCP-GA, while the ZINC subset³⁷ database was curated, we selected several neutral compounds to test against the open cavity. Phenol and 3-fluorocatechol we picked because they were known ligands for CCP wt*,

albeit weak binders with millimolar affinities at best (Chapter 1). 4-hydroxybenzaldehyde was selected because it had recently been discovered as the first known ligand for the neutral version of the CCP wt* cavity (SEB, unpublished results), created by the additional mutation D235→N, replacing the charged aspartate with neutral asparagine.¹¹

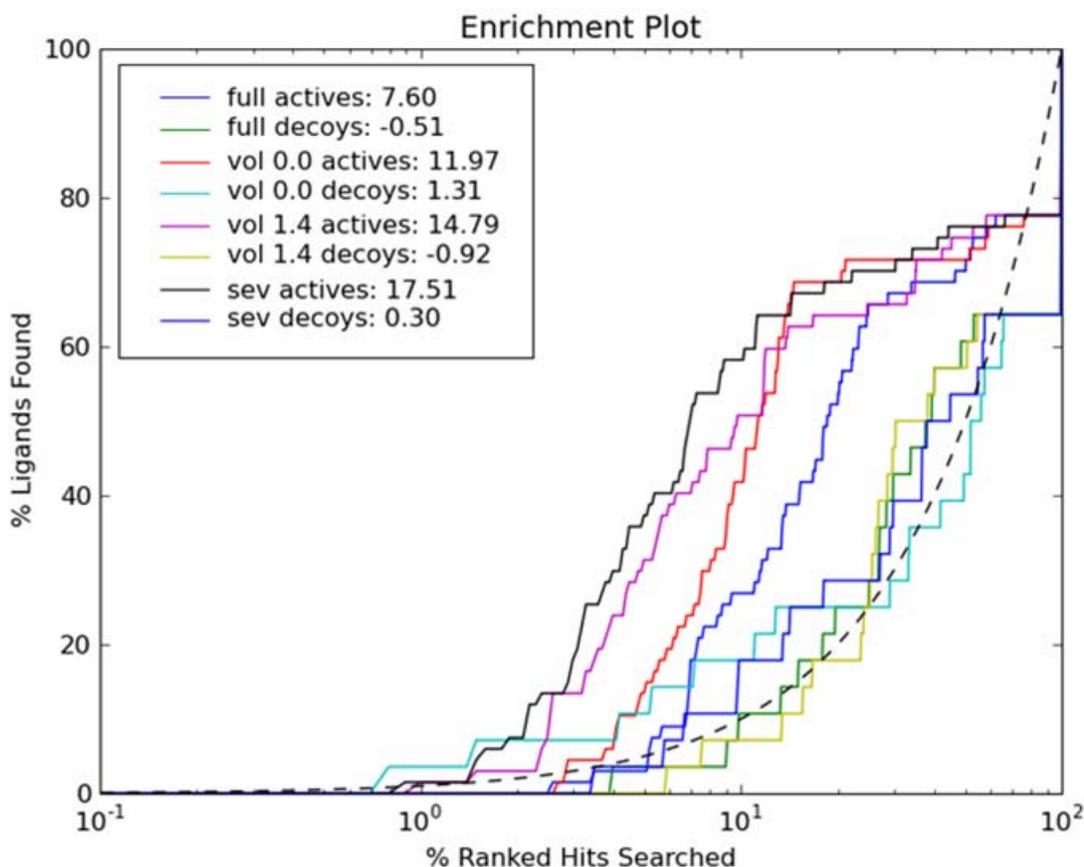


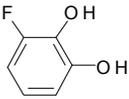
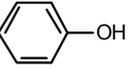
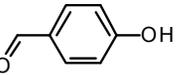
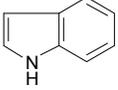
Figure 3. Comparison of the enrichment for CCP-GA with four solvation methods: full, volume based desolvation with probe radius 0Å, volume based desolvation with probe radius 1.4Å, and the new SEV method (actives: black, decoys: blue). SEV performs as well, or better than volume_1.4.

Finally we selected indole, as it was known to be a neutral decoy for CCP wt*, but was isosteric with the wt* ligands benzimidazole and imidazo-[1,2- α]pyridine¹², also ligands

for CCP-GA with co-complex crystal structures available for comparison (in-house results).⁷

The four ligands were assayed against CCP-GA by UV-Vis titration of the heme Soret band¹³ and three of the four showed binding (Table 1). In addition, all four compounds were soaked into CCP-GA crystals and 2F_o-F_c electron density for three compounds, phenol, 3-fluorocatechol and 4-hydroxybenzaldehyde, clearly showing that they bound to CCP-GA. More importantly the crystal data showed that the compounds interacted with Asp233 (analogous to Asp235 in the wt* and Asn235 in D235N). 3-fluorocatechol bound in essentially the same orientation as in the wt*. However, phenol was refined to two orientations (Fig 4.a), one consistent with the wt* binding mode, in which the ligand does not interact with the aspartate and instead flips around to interact with a backbone carbonyl of Leu 177 in the wt* this had led us to believe a pKa shift had occurred in the cavity, and that the aspartate was no longer charged.

Table 1. Neutral ligands preferentially bind to CCP-GA at pH 6.0

		pH 6.0	pH 4.5
3-fluorocatechol		1mM	10mM
phenol		1mM	no binding detected 10mM
4-hydroxybenzaldehyde		1mM	no binding detected 10mM
indole		no binding detected 10mM	no binding detected 10mM

However, in CCP-GA the alternate orientation does contact Asp233. Finally, the 4-hydroxybenzaldehyde binding mode is the exact reverse of that observed in D235N (Fig 4.b); in the D235N structure the ligand's aldehyde group contacts Asn235, consistent with the observation that the donor amine is pointing into the cavity (unpublished result, observed at 1.2Å in the apo crystal structure the difference between oxygen and nitrogen atoms is apparent). In the CCP-GA complex structure the hydroxyl group is oriented towards the aspartate. Although this

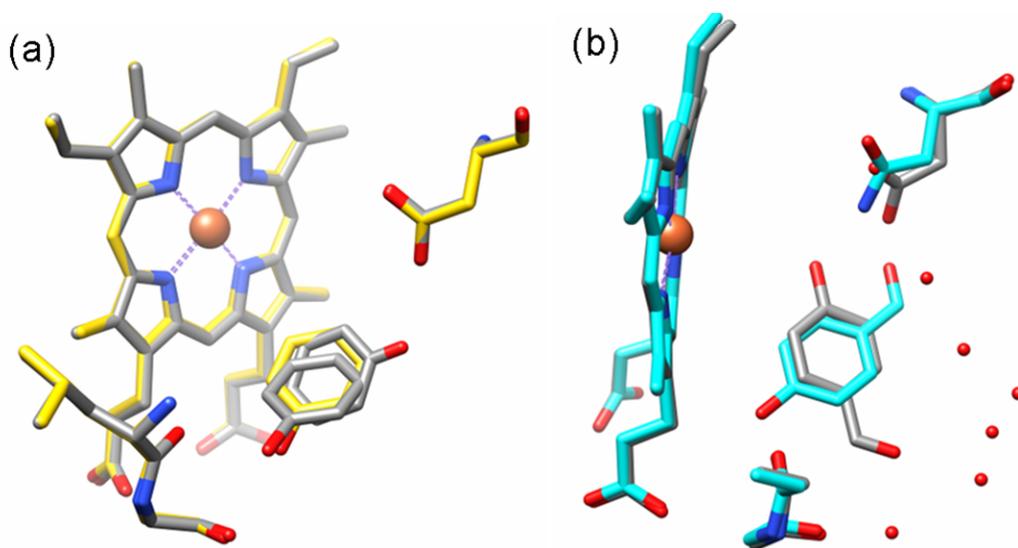


Figure 4. Crystal structures of neutral ligands. **(a)** phenol binding to CCP-GA shown in gray and wt* in yellow. **(b)** 4-hydroxybenzaldehyde binding to CCP-GA shown in gray and CCP W191G/D235N in cyan. The waters (red) are observed in the CCP-GA complex structure.

does not definitively indicate that the aspartate is charged, as the hydroxyl group of the ligand can function as an acceptor as well, the fact that the ligand did not adopt the alternate pose seen in the D235N structure leaves us with an interesting possibility: Are neutral ligands binding to a charged aspartate?

Ligands bind to a charged Asp233

The question of whether the neutral ligands were binding to the charged aspartate was answered by performing the Soret band titration assays at pH 4.5 (essentially the pKa of aspartate) and at pH 6.0 (where the aspartate should be fully charged). The initial results are compelling; all three ligands showed a stronger shift of the heme Soret band at lower concentration at pH 6.0 than pH 4.5 (Table 1). In fact, phenol and 4-hydroxybenzaldehyde did not appear to bind detectably at pH 4.5 even at 10mM. 3-fluorocatechol showed a slight shift at higher concentrations at pH 4.5, but clearly binds at 1mM at pH 6.0 (Fig 5).

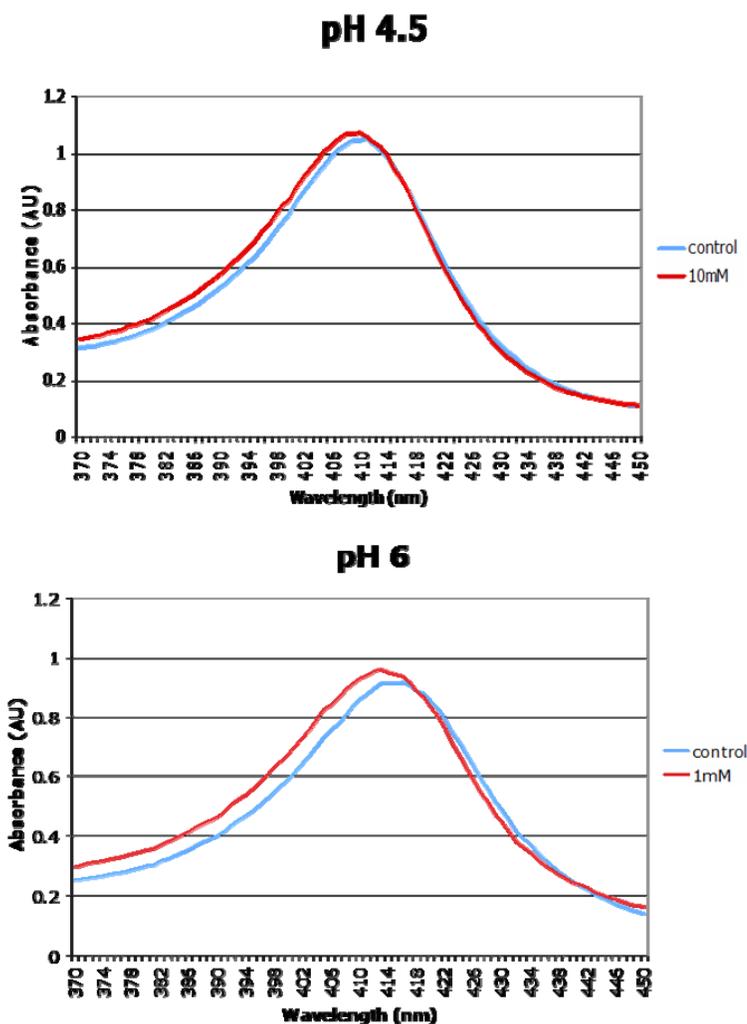


Figure 5. UV-Vis titration of the heme Soret band in CCP-GA by 3-fluoro-catechol. At 10mM concentration only a slight blue-shift is observed at pH 4.5 (red line), compared to the intense shift seen at 1mM at pH 6.0.

4.4 Future Directions

Determining the impact of a solvent exposed cavity on ligand binding

This preliminary data appears to support the binding of neutral ligands to the charged cavity, consistent with the dock predictions of more neutral compounds in the top of the hit list. This of course requires more testing; as stated above several neutral ligands from docking have been purchased. However, another interesting question that also has not yet been addressed, although it has been hinted at by the preliminary binding data, is that neutral ligands may bind with a higher affinity to CCP-GA than the wt* cavity; conversely the smaller charge ligands may bind with a decreased affinity based on work done by Rosenfeld *et al.*⁷ To test this, we plan to select both constitutively charged small known ligands and neutral molecules for the wt* cavity and determine the affinity for these compounds against both CCP-GA and CCP wt* using low c-value isothermal titration calorimetry.¹⁴ The assays will be done at both pH 4.5 and pH 6.0 as a double difference experiments to answer both questions; in which state is the Asp233 (or 235) upon ligand binding, and whether the higher dielectric in the open cavity enhances affinity of the neutral ligands while dampening the interaction between the charge ligands and the aspartate.

Ordered waters and bulk solvent interface

The importance of considering ordered water displacement for ligand optimization and rational drug design is not a new idea; Jorgensen *et al* have developed several methods to predict water position and interaction with a protein or ligand-protein complex (refs).^{15,16} Our goal would be to apply a similar methodology during docking. Code developed by Niu Huang to flexibly treat waters has been implemented in DOCK

and shown to be successful in several cases against targets in the DUD set.¹⁷ The method only considers crystallographically observed waters and their interaction with the docked molecules. However, there are two key components which have been neglected in this implementation; the cost (or benefit) of stripping the waters from the binding site and the potential for new waters to bind with different ligands.. The ability of “new” waters to come in with a ligand has already been observed in the wt* cavity for ligands with localized charges, and is extremely important in mitigating the desolvation penalty the ligand would otherwise pay (Chapter 1).

Determining the cost of removing known waters and also predicting potential new waters could be possible using alchemical free energy methods, applied in Chapter 3 to predict ligand binding; there are plans already underway to repeat the free energy predictions in Chapter 3 with this system—including free energy calculations for cavity solvation would already be considered in these calculations and therefore is not a farfetched idea. In fact, a recent result indicates that during free energy simulations of the apo CCP-GA with water molecules allowed to enter and leave the cavity, the correct water positions for all 8 waters were found.¹⁸

A more accessible goal, for the time being, is to use a new method developed in the Dill group using semi-explicit solvation free energy method to predict “hotspots” for waters in the pocket and to calculate a desolvation penalty for removing waters from those locations.¹⁹ While the first option, alchemical free energy methods, is by far more rigorous, the second option may be more feasible as a fast method to predict waters. While this work has not yet been attempted in CCP-GA, there is good evidence that determining alternative water positions will be critical to accurate binding mode prediction and ranking of known ligands, for example the structure of 4-hydroxybenzaldehyde bound to CCP-GA, in which there is an extensive network of

hydrogen bonds between waters and the ligand (Fig 4.b). The docking prediction (using both volume based and SEV solvation methods) just barely misses the interaction with Asp233, but with waters present may have found the correct orientation (Fig 6).

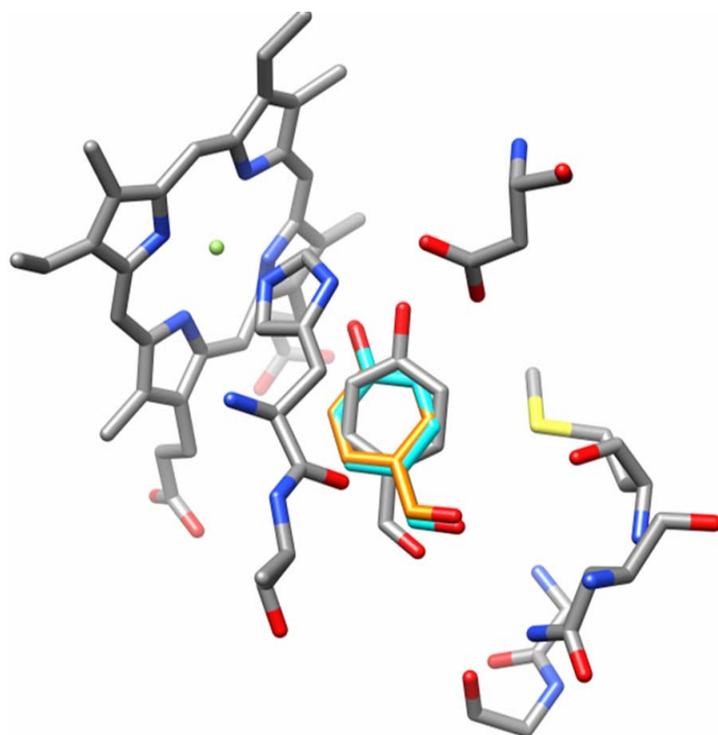


Figure 6. Docking vs. crystal pose of 4-hydroxybenzaldehyde in CCP-GA. Gray: the xtal pose contacts Asp233; cyan: volume based desolvation with 1.4Å probe radius; orange: SEV.

And finally, another direction to pursue with this cavity that has direct implications for virtual screening in true drug targets is the interaction with both the bulk solvent interface. It has been a longstanding goal in the lab to investigate including receptor desolvation in our scoring function; this cavity may provide a system in which to do this. It has already been shown by Goodin *et al* that a probe molecule, designed as a peptide mimic of the mutated flexible loop reaches out to and interacts with bulk solvent.²⁰ Although this ligand binds with micromolar affinity, it indicates the feasibility of finding molecules that span the binding site and interact with the solvent interface.

While this particular project may be the furthest on the horizon, this cavity presents a golden opportunity to prospectively test receptor desolvation.

4.5 Methods

Protein Prep

The plasmid for the CCP-GA mutant protein was received from Goodin lab and protein expression and purification followed the same protocol used for the wt* protein (Chapter 1).

UV-Vis titration assay

Ligand stocks were made up to 1M in DMSO, solubility permitting. Titration of the heme Soret band followed the same protocol described in Chapter 1 and also Musah *et al.*¹³ The following buffers were used for the assay: pH 4.5 100mM citric acid with bis-tris propane (BTP) counter ion and 500mM MES buffer, adjusted with BTP to pH 6.0.

Crystallography

Ligands were soaked into crystals grown under previously published conditions (Chapter 1) at concentrations up to 50mM in MES pH 6.0 buffer or, if solubility was an issue, in the cryoprotectant, 25% MPD.

Docking

Docking was performed as previously (Chapter 1 & 2) to the wt* cavity with the following exceptions. The receptor was prepared using Dockblaster.²¹ Spheres were generated from a file that combined the original docking spheres for the wt* cavity with new ordered waters observed in the apo CCP-GA cavity, waters observed in the

benzimidazole bound structure (1KXM) and the benzimidazole ligand atoms. Ligand desolvation was calculated using the volume based desolvation method or the new SEV method, as described in Preliminary Results.

4.6 Acknowledgements

I would like to acknowledge the following people for their help with this project: Dave Goodin and Rich Wilson for providing the plasmid and other expertise with the open cavity; Michael Mysinger and John Irwin for help with docking, ZINC and interesting conversations about the model system cavities; Gabe Rocklin, Chris Fennell, Charlie Kehoe, and Ken Dill for their willingness to try free energy and semi-explicit solvation calculations in the (tricky) CCP model system; and Jason Fernandez, Nick Hertz and Rocco Varela, who were all rotation students on this project and contributed to docking, discovery and crystallography of many of the new ligands for the CCP-GA binding site.

4.7 References

1. Wei, B. Q., Baase, W. A., Weaver, L. H., Matthews, B. W. & Shoichet, B. K. (2002). A model binding site for testing scoring functions in molecular docking. *J. Mol. Biol.* **322**, 339-355.
2. Wei, B. Q., Weaver, L., Ferrari, A. M., Matthews, B. M. & Shoichet, B. K. (2004). Testing a flexible-receptor docking algorithm in a model binding site. *J. Mol. Biol.* **337**, 1161-82.
3. Graves, A. P., Brenk, R. & Shoichet, B. K. (2005). Decoys for Docking. *J. Med. Chem.* **48**, 3714 - 372.
4. Rosenfeld, R., Goodsell, D., Musah, R., Garret, M., Goodin, D. & Olson, A. (2003). Automated docking of ligands to an artificial active site: augmenting crystallographic analysis with computer modeling. *J Comput Aided Mol Des.*
5. Mobley, D. L., Graves, A. P., Chodera, J. D., McReynolds, A. C., Shoichet, B. K. & Dill, K. A. (2007). Predicting absolute ligand binding free energies to a simple model site. *J. Mol. Biol.* **371**, 1118-1134.
6. Mobley, D. L., Chodera, J. D. & Dill, K. A. (2006). On the use of orientational restraints and symmetry corrections in alchemical free energy calculations. *J. Chem. Phys.* **125**, 084902.
7. Rosenfeld, R. J., Hays, A. A., Musah, R. A. & Goodin, D. B. (2002). Excision of a proposed electron transfer pathway in cytochrome c peroxidase and its replacement by a ligand-binding channel. *Protein Science* **11**, 1251-1259.
8. Irwin, J. J. & Shoichet, B. K. (2005). ZINC - A Free Database of Commercially Available Compounds for Virtual Screening. *J Chem Inf Model* **45**, 177-182.
9. Mysinger, M. M. (2009). New ligand based solvation methods (Boyce, S. E., ed.).

10. Huang, N., Shoichet, B. K. & Irwin, J. J. (2006). Benchmarking sets for molecular docking. *J Med Chem* **49**, 6789-6801
11. Fitzgerald, M. M., Trester, M. L., Jensen, G. M., McCree, D. E. & Goodin, D. B. (1995). The role of aspartate-235 in the binding of cations to an artificial cavity at the radical site of cytochrome c peroxidase. *Protein Science* **4**, 1844-1850.
12. Fitzgerald, M. M., Musah, R. A., McRee, D. E. & Goodin, D. B. (1996). A ligand-gated, hinged loop rearrangement opens a channel to a buried artificial protein cavity. *Nat Struct Biol* **3**, 626-31.
13. Musah, R. A., Jensen, G. M., Bunte, S. W., Rosenfeld, R. J. & Goodin, D. B. (2002). Artificial protein cavities as specific ligand-binding templates: characterization of an engineered heterocyclic cation-binding site that preserves the evolved specificity of the parent protein. *J Mol Biol* **315**, 845-57.
14. Turnbull, W. B. & Daranas, A. H. (2003). On the Value of c: Can Low Affinity Systems Be Studied by Isothermal Titration Calorimetry? *J. Am. Chem Soc.* **125**, 14859-14866.
15. Michel, J., Tirado-Rives, J. & Jorgensen, W. L. (2009). Energetics of displacing water molecules from protein binding sites: consequences of ligand optimization. *J. Am. Chem Soc.* **131**, 15403-15411.
16. Michel, J., Tirado-Rives, J. & Jorgensen, W. L. (2009). Prediction of water content in protein binding sites. *J Phys Chem B* **113**, 13337-13346.
17. Huang, N. & Shoichet, B. K. (2008). Exploiting ordered waters in molecular docking. *J Med Chem* **51**, 4862-4865.
18. Rocklin, G. J. (2009). Free energy calculations in CCP-GA with water (Boyce, S. E., ed.).
19. Fennell, C. & Kehoe, C. (2009). Semi-explicit solvation (Boyce, S. E., ed.).

20. Hays, A. A., Gray, H. B. & Goodin, D. B. (2003). Trapping of peptide-based surrogates in an artificially created channel of cytochrome c peroxidase. *Protein Science* **12**, 278-287.
21. Irwin, J. J., Shoichet, B. K., Mysinger, M. M., Huang, N., Colizzi, F., Wassam, P. & Cao, Y. (2009). Automated docking screens: a feasibility study. *J. Med. Chem.* **52**, 5712-5720.

Appendix A:

Supplementary Material for Chapter 1

Probing molecular docking in a charged model binding site

Ruth Brenk^{1†}, Stefan W. Vetter^{2‡}, Sarah E. Boyce¹⁺, David B. Goodin^{*2}, Brian K. Shoichet^{*1}

A.1 Supplementary Figures for Chapter 1

Figure S1. Scatter plot of experimentally determined binding constants for the previously known, “test set” ligands¹⁻³ compared to $|E_{tot}|$, the absolute value of the binding energy calculated by DOCK (the larger the score the better in this plot).

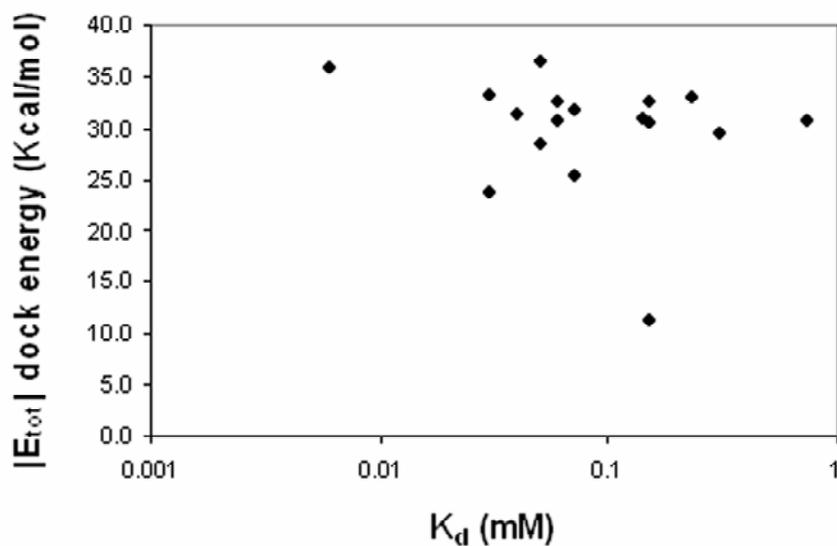


Figure S2. Docking-based Enrichment of ligands (solid lines) and downgrading of decoys in the cavities of CCP W191G, T4 lysozyme L99A, and T4 lysozyme L99A/M102Q. The partial charges of the ligands were either calculated in a medium of high dielectric (water, HD) or low dielectric (cyclohexane, LD). The known compounds for CCP W191G are the those in the “test set”¹⁻³ augmented with the new compounds (Table2).

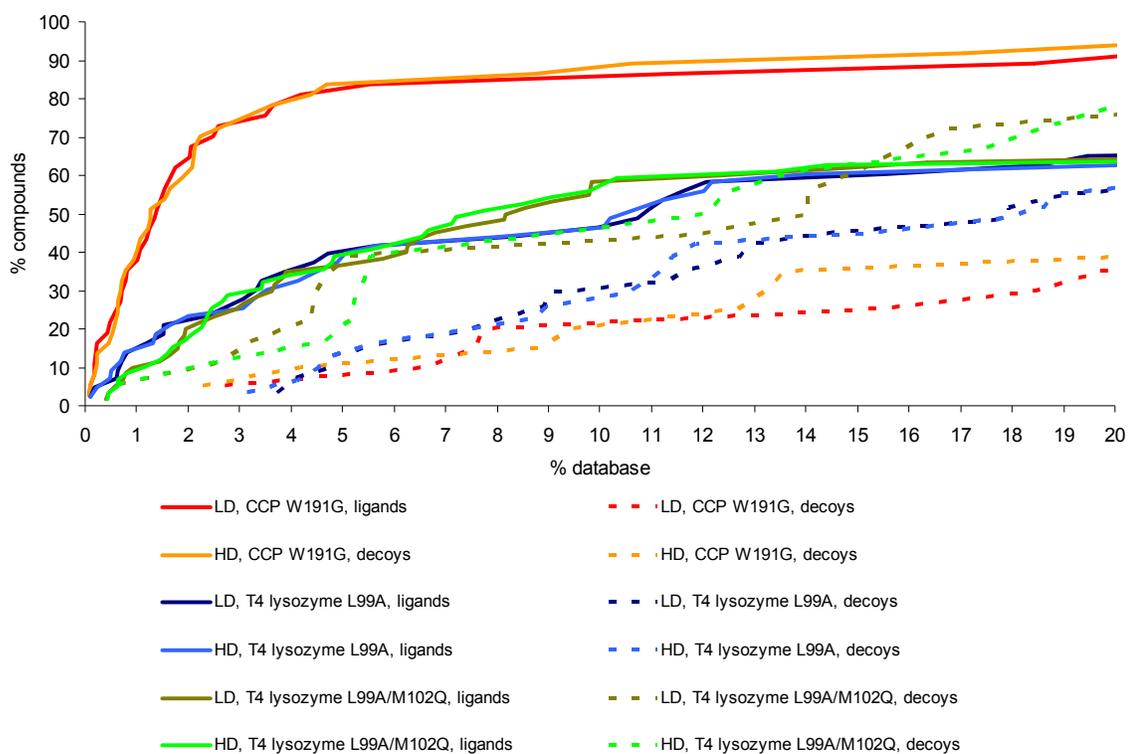
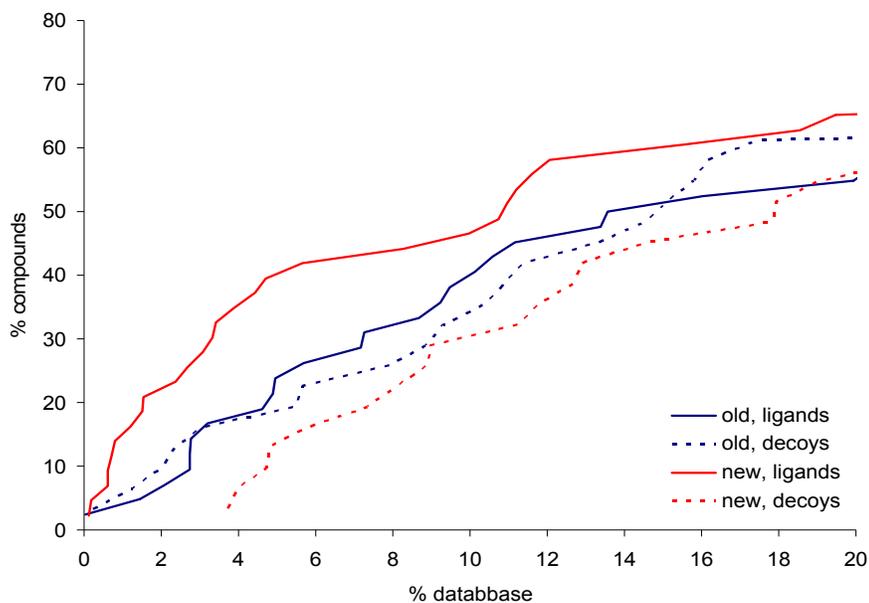
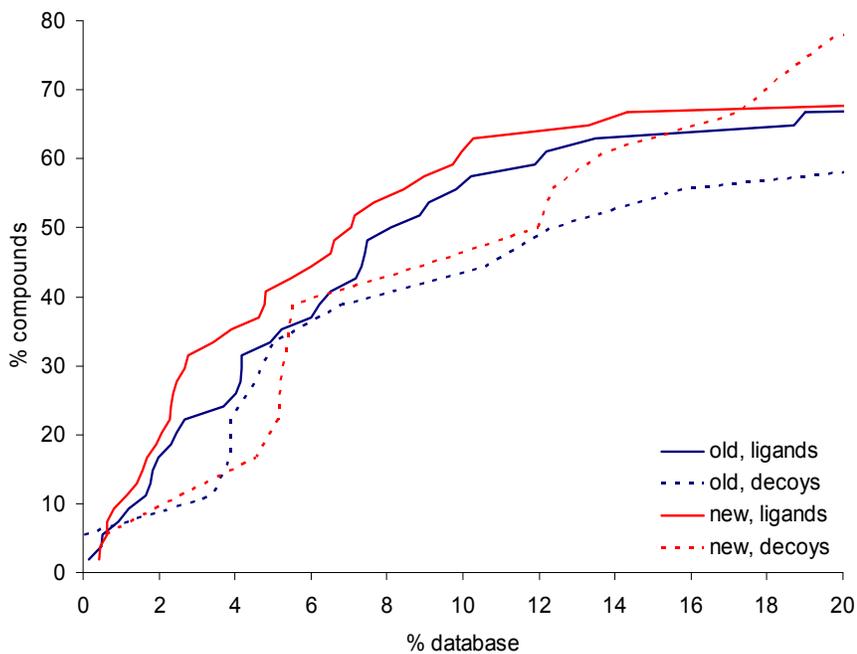


Figure S3. Enrichment of ligands (solid lines) and downgrading of decoys (dashed-lines) docked in the L99A (a) and the L99A/M102Q (b) cavities of T4 lysozyme with the desolvation energy calculated according to equation 2 (blue, old method) and equation 3 (red, new method).

a)



b)



A.2 References

1. Rosenfeld, R., Goodsell, D., Musah, R., Garret, M., Goodin, D. & Olson, A. (2003). Automated docking of ligands to an artificial active site: augmenting crystallographic analysis with computer modeling. *J Comput Aided Mol Des.*
2. Musah, R. A., Jensen, G. M., Bunte, S. W., Rosenfeld, R. J. & Goodin, D. B. (2002). Artificial protein cavities as specific ligand-binding templates: characterization of an engineered heterocyclic cation-binding site that preserves the evolved specificity of the parent protein. *J Mol Biol* **315**, 845-57.
3. Fitzgerald, M. M., Churchill, M. J., McRee, D. E. & Goodin, D. B. (1994). Small molecule binding to an artificially created cavity at the active site of cytochrome c peroxidase. *Biochemistry* **33**, 3807-18.

Appendix B:

Supplementary Material for Chapter 2

Rescoring docking hit lists for model cavity sites: predictions and experimental testing

Alan P. Graves,^{a,b,†} Devleena M. Shivakumar,^{c,†,‡} Sarah E. Boyce,^{a,d} Matthew P.
Jacobson,^{a,*} David A. Case,^{c,*} and Brian K. Shoichet^{a,*}

Contents:

B.1 PLOP side chain rotamer search and minimization algorithm.

B.2 AMBERDOCK parameters and optimization.

B.3 Tables S1-S9.

B.4 Figures S1-S3.

B.1 PLOP side chain rotamer search and minimization algorithm.

The side chain rotamer search algorithm proceeds by optimizing one side chain at a time (keeping the others fixed), and keeps looping through the residues until they stop changing conformation. During the side chain rotamer search, the default resolution of 10° for the rotamer library was used, and the initial side chain conformations were utilized by specifying the parameter “randomize no”. The initial (ofac_init) and minimal (ofac_min) overlap factors, which define steric clashes, were set to 0.75 and 0.5, respectively. The calculation of the complex energy ($E_{complex}$) involved an initial short minimization of the ligand with a maximum number of steps (mxitn) set to 15 followed by the sidechain rotamer search, minimization of residues with a maximum number of steps (mxitn) set to 30, and finally minimization of ligand to convergence. Minimization and calculation of the free receptor energy ($E_{receptor}$) began with the initial side chain rotamer search followed by the minimization of residues with a maximum number of steps (mxitn) set to 30.

B.2 AMBERDOCK parameters and optimization.

The traditional AMBER force field has been parameterized to work with biological macromolecules, and has limited parameters for organic molecules. The general AMBER force field (GAFF), which has been specifically designed to cover most pharmaceutical molecules, is compatible with the existing AMBER force fields in such a way that the two can be mixed during a simulation. It uses a simple harmonic energy function as the additive AMBER force fields, but the atom types used in GAFF are much

more general such that they cover most of the organic chemical space. The current implementation of the GAFF force field consists of 33 basic atom types and 22 special atom types covering almost all of the organic chemical space that is made up of C, N, O, S, P, H, F, Cl, Br and I atoms. The input ligand files for AMBERDOCK can be generated automatically using the perl script (prepare_amber.pl) provided with the DOCK6.1 program. This perl script calls for programs such as *antechamber* to calculate the charges for the ligands and tleap to assign the parameter set for protein and ligand atoms. The newer version of the DOCK6 program bypasses this external perl step, and the input files are generated internally when the AMBER score routine is called.

Binding free energy calculations with AMBERDOCK follows the flowchart in Figure S1. In Step 1, the input files (pdb, inpcrd, prmtop) in AMBER format are read. In Step 2, the frozen atoms are defined based on the choice of the user. The user can specify the atoms that are allowed to relax in the DOCK6 input file (dock.in). A brief minimization is carried out using the GB implicit solvent model in Step 2. In Step 3, the molecular dynamics (MD) simulation is carried out on the minimized system. In step 4, a brief minimization is carried out to minimize the structure generated from the previous MD step. Steps 2-4 are performed with the frozen atom approximation in GB implicit solvent model. When frozen atoms are specified, the NAB program calculates the energy for only those atoms that are allowed to move. This significantly speeds up the calculations and uses less memory. In Step 5, a single point energy is calculated on the structure obtained from Step 4. In this step, the energy of the full system is calculated without any frozen atoms and non-bonded cutoffs. Also, the surface area term is added to this final step to get a more accurate energy term. The Steps 1-5 were repeated for complex ($E_{complex}$), ligand (E_{ligand}) and receptor ($E_{receptor}$). To expedite the scoring process, we calculated the energy of the receptor for the first ligand and used this energy

as a constant term during the energy evaluations for the rest of the ligands in the database. It is suggested to repeat these steps with various combinations of molecular mechanics (MM) options to optimize the variables.

Several scoring protocols were tested on the 58 known ligands and 17 known decoys for L99A/M102Q to find the optimal set of conditions for AMBERDOCK rescoring. The initial attempt was to use AMBERDOCK with only minimization and no molecular dynamics simulation. For example, *score1* (Figure S2b) involved scoring the test set with Gasteiger-Marselli PEOE charges for the ligands and AMBER parm 94 parameter set for protein atoms; the Hawkins-Cramer-Truhlar pairwise GB model (equivalent to *igb=1* in AMBER); the LCPO method to calculate the surface area term; and minimization of nine binding site residues (78, 84, 88, 91, 102, 111, 118, 121, and 153). The results clearly show that the minimization alone is not sufficient to separate out the ligands from decoys with several known decoys scoring better than known ligands. In fact, the results from DOCK (Figure S2a) is adversely affected by running *score1*. We also tried rescoring the same test set with a series of AMBERDOCK scores with different combinations of the molecular mechanics input parameters, using large or small flexible receptor regions, GB models (*igb=1, 2, 5*) with or without the surface area term, number of minimization steps, etc. We found that all of them predicted several decoys among the top 20 scored ligands with only minimization of the system, and no molecular dynamics (MD) simulations. As a logical next step, we introduced MD simulations along with minimization in the system. The initial receptor-ligand configuration obtained from docking was subjected to a few steps of minimization followed by a few picoseconds of MD simulations, and then a final minimization run to obtain the total energy. Selection of charge models is also very important. The AMBERDOCK scores obtained when using the AM1-BCC charge model provided better

results compared to the Gasteiger-Marselli PEOE charge model. Using *score24G* with Gasteiger- Marselli PEOE charges with MD and minimization, 6 decoys scored as well as the top 20 ligands (Figure S2c). Using *score24* with AM1-BCC charges with MD and minimization, only 1 decoy scored as well as the top 20 scoring ligands (Figure S2d). This scoring protocol produced good results for most of our test sets of known ligands and decoys compared to the other scoring protocols that we tested. The specifications of *score24* are: GB model of 5 (corresponding to $igb=5$ in the AMBER program); surface area term calculated using the LCPO model; a non-bonded cutoff of 18 Å; 100 steps of conjugate gradient minimization; and 3000 steps of MD simulation with a 1 fs time step at a temperature of 300K, followed by 100 steps of minimization. During the minimization and MD, only the ligand and the protein residues within 5 Å of the ligand were allowed to move.

B.3 Supplementary Tables for Chapter 2

The following tables are available online at:

[doi:10.1016/j.jmb.2008.01.049](https://doi.org/10.1016/j.jmb.2008.01.049)

Table S1. Top 100 hits predicted by DOCK3.5.54 for L99A.

Table S2. Top 100 hits predicted by PLOP for L99A.

Table S3. Top 100 hits predicted by AMBERDOCK for L99A.

Table S4. Top 100 hits predicted by DOCK3.5.45 for L99A/M102Q.

Table S5. Top 100 hits predicted by PLOP for L99A/M102Q.

Table S6. Top 100 hits predicted by AMBERDOCK for L99A/M102Q.

Table S7. Top 100 hits predicted by DOCK3.5.54 for CCP.

Table S8. Top 100 hits predicted by PLOP for CCP.

Table S9. Top 100 hits predicted by AMBERDOCK for CCP.

B.4 Supplementary Figures for Chapter 2

Figure S1. Distribution of pair-wise Tanimoto similarities among the top-ranked docked and rescored molecules, calculated using ECFP4 fingerprints.

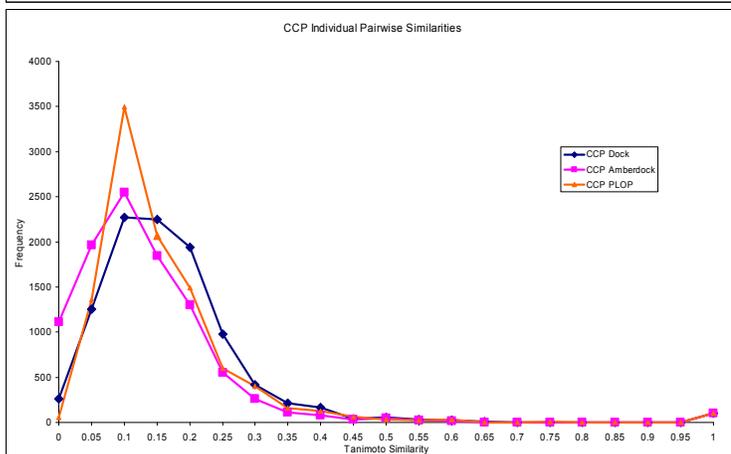
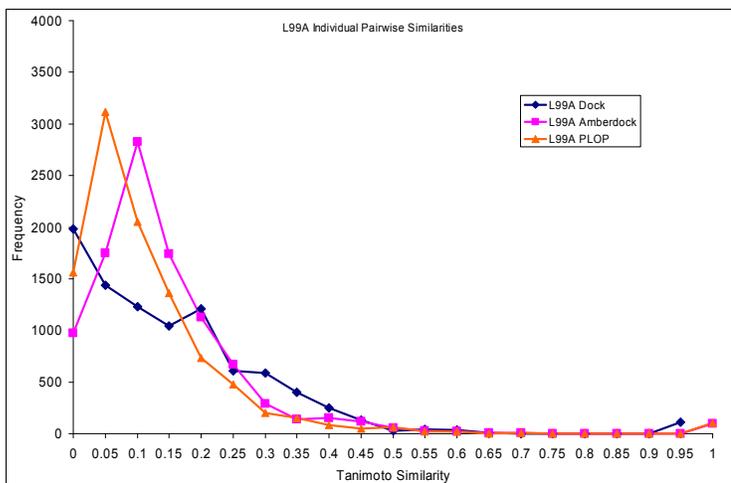
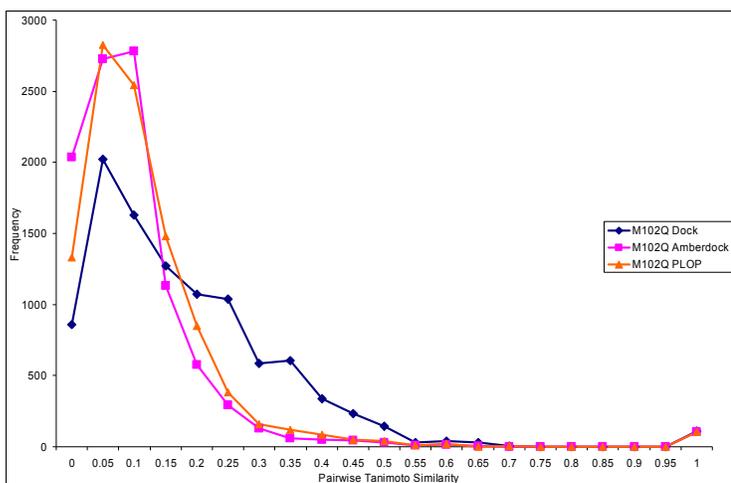
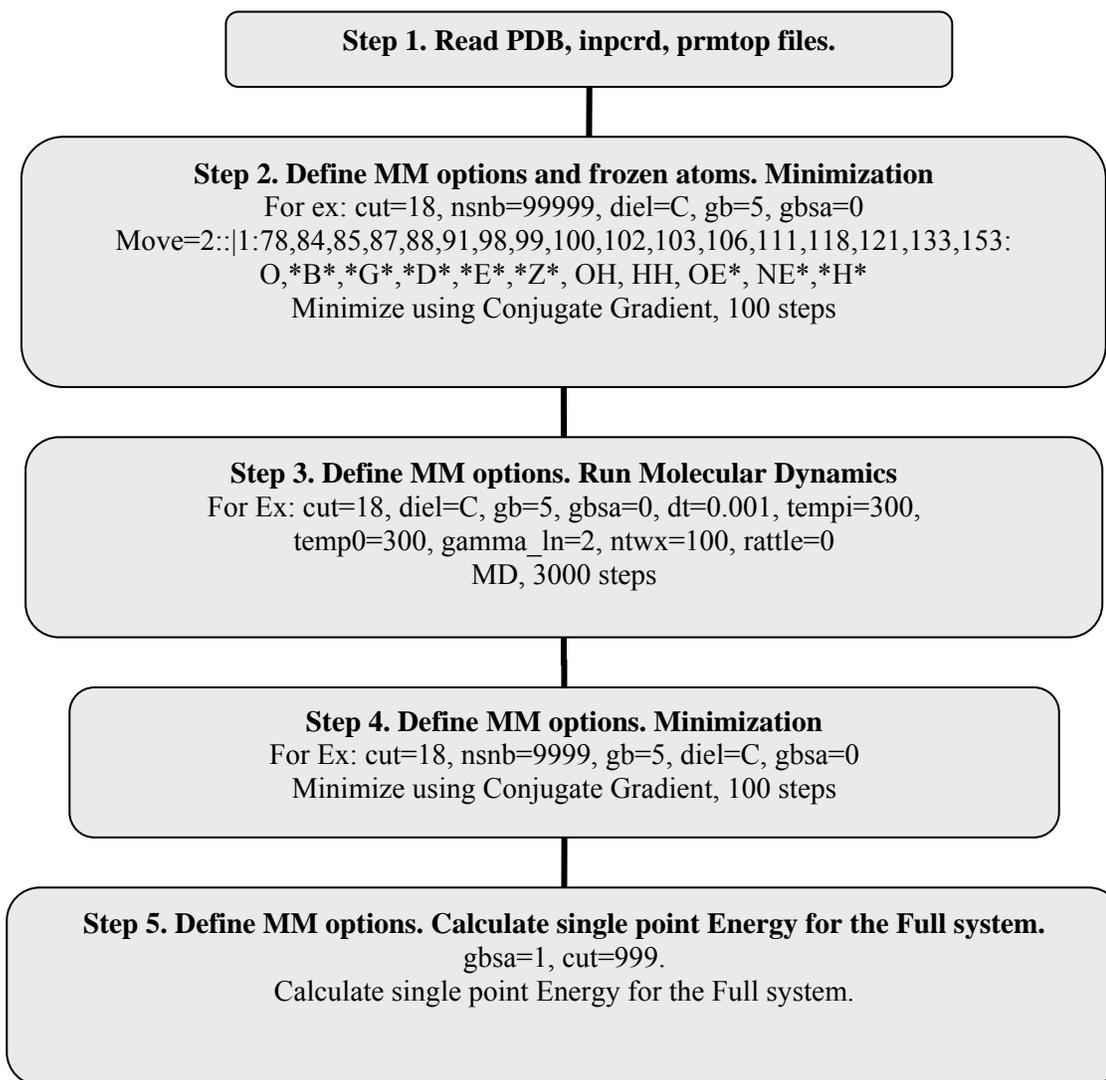
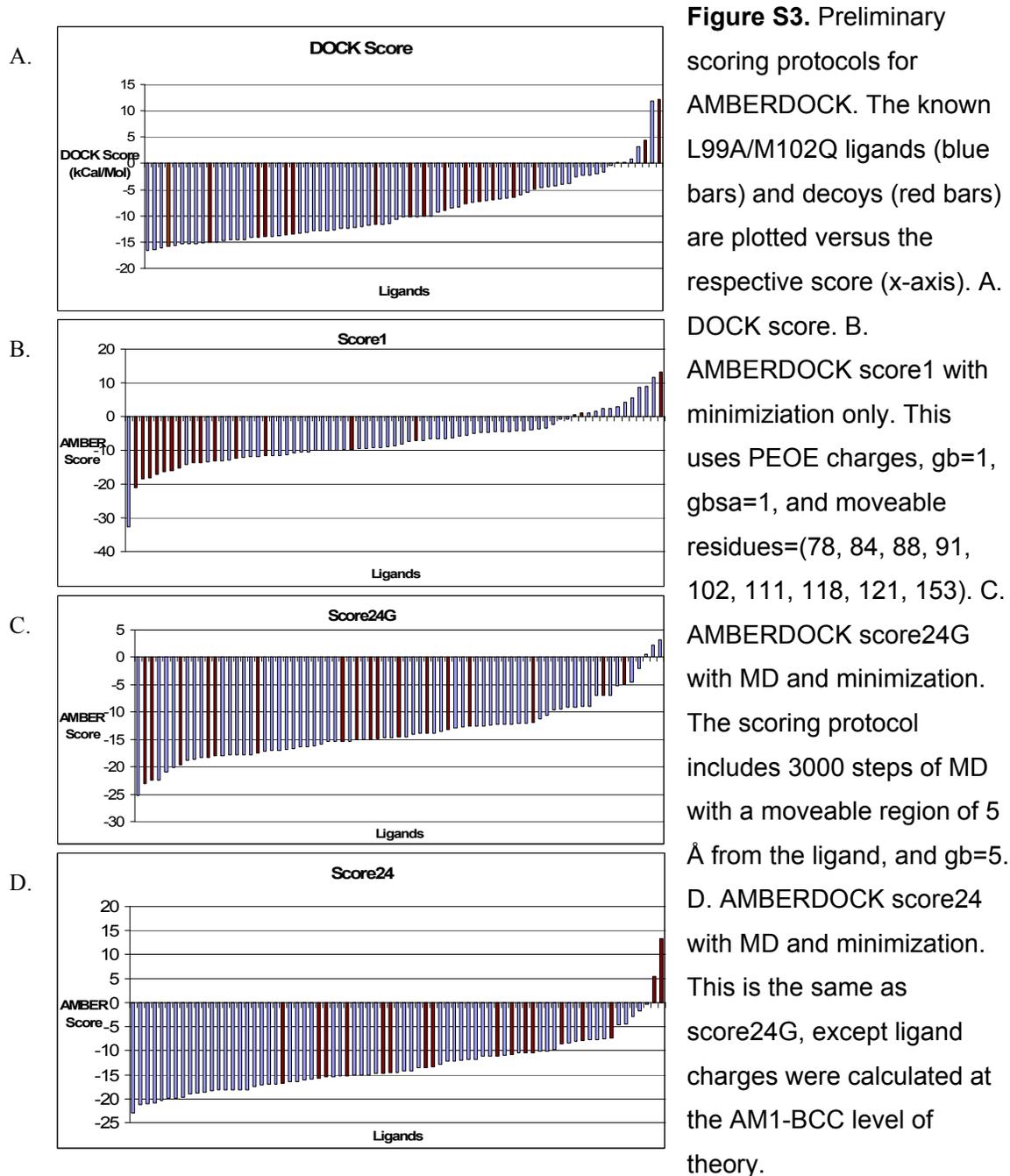


Figure S2. Flowchart and parameters for AMBERDOCK rescoring.



Description of the keywords: gb=2, Onufriev, Bashford, Case (OBC) variant of GB, with $\alpha=0.8$, $\beta=0.0$, $\gamma=2.909$. Similar to igb=2 in AMBER.; gb=5, with $\alpha=1.0$, $\beta=0.8$, $\gamma=4.85$; diel=C, uses constant dielectric; dt, time step, ps; tautp, temperature coupling parameter, in ps; gamma_ln, collision frequency for Langevin dynamics, in ps^{-1} ; temp0, target temperature; gbsa, add a surface-area dependent energy.



Appendix C:

Supplementary Material for Chapter 3

Predicting ligand binding affinity with alchemical free energy methods in a polar model binding site

Sarah E. Boyce^{1†}, David L. Mobley^{2†}, Gabriel Rocklin³, Alan P. Graves^{3‡}, Ken A. Dill^{4*},
Brian K. Shoichet^{4*}

C.1 Supplementary Methods

Protein/system preparation and parameters

Following setup, the protein was embedded in a periodic dodecahedral simulation box with a minimum distance of 1.0 nm from the protein, at the center of the box, to the nearest box edge. TIP3P¹ water molecules were added to fill the box, ensuring that no waters were inserted nearer than 0.3 nm from any protein atom². Following this, we equilibrated the water for 1 ns while holding heavy atoms in the protein fixed, using our standard simulation protocol/parameters (below). We used these equilibrated systems as input for further simulations.

Ligands were parameterized as previously²⁻⁶: We used the GAFF small molecule parameter set⁷ and AM1-BCC partial charges computed with Antechamber; however, Antechamber gave inconsistent partial charges when using `-j 1` from using the default `-j` option, even when atom and bond types did not change. These were substantially different in some cases from those provided by Bayly and apparently in error. To correct

for this we obtained AM1-BCC partial charges directly from Christopher Bayly⁸ and used these in some additional tests.

Initial ligand conformations were generated from the chemical names using OpenEye's Lexichem and Omega, and these were used as input for the partial charge calculation. The resulting mol2 files with partial charges were used as both input for docking and for input to tleap to generate AMBER format parameter and coordinate files, which we then converted to GROMACS format using our amb2gmx conversion tool⁹. When simulated separately from the protein, ligands were placed in a dodecahedral simulation box with at least 1.2 nm from the molecule to the nearest box edge, and solvated in TIP3P water.

Standard simulation parameters/protocols

Our standard simulation parameters were to use particle mesh Ewald¹⁰ for treatment of long range electrostatics, with a real space cutoff of 1.0 nm, a Fourier grid spacing of as near as possible to 0.1 nm, a spline order of 6, and a relative tolerance of 10^{-6} . A timestep of 2 fs was used for all simulations unless otherwise noted. The Langevin integrator was used for temperature control with a friction coefficient of 1 ps^{-1} , at a temperature of 300 K. van der Waals interactions were calculated with a switched cutoff between 0.8 and 0.9 nm, and a neighbor list of 1.0 nm was used, updated every 10 steps. We used the analytical long range van der Waals dispersion correction implemented in GROMACS to approximately account for the effects of truncating long range dispersion interactions on the energy and pressure.

For each simulation the system was first minimized with up to 5,000 steps of steepest descents minimization, then run 10 ps of isothermal molecular dynamics,

followed by 100 ps of isothermal-isobaric dynamics pressure regulated using the Berendsen weak-coupling scheme¹¹. The pressure regulation used a time constant of 0.5 ps, a reference pressure of 1.0 atm, and an isothermal compressibility of 4.5×10^{-5} bar. Subsequently, production simulations were run using isothermal molecular dynamics, typically for 1 ns for simulations of protein-ligand complexes and 5 ns for ligands in water/vacuum.

Vacuum simulations differed from the standard protocol in that lattice sum electrostatics were turned off, and cutoffs were increased to ensure that all interactions were computed.

Unrestrained simulations and selecting reference orientations

These unrestrained simulations play an important role in the subsequent absolute binding free energy calculations. Our absolute binding free energy protocol involves restraining the ligand in the binding site to a particular orientation (and computing the restraining free energy) then gradually turning off interactions between the ligand and the protein in a series of separate simulations, computing this free energy as well. The ligand is then transferred to water, the free energy of removing the restraints is accounted for analytically, and the interactions are restored in water, completing a thermodynamic cycle equivalent to transferring the ligand from the binding site to the standard state in water. Details of this cycle have been discussed previously² and are recapped below. For the restraints, we restrain the six relative degrees of freedom between the protein and the ligand, in this case using two angles, three dihedrals, and one distance, to measure the ligand position relative to reference atoms in the protein, as discussed previously^{2,9,12}. However, this approach requires selecting a ligand orientation to restrain the ligand to.

This choice is in principle arbitrary, but in practice, a good choice can ease convergence of the calculations. Additionally, large kinetic barriers between different ligand orientations mean that considering multiple candidate orientations and combining the effective binding free energies of different orientations into a total binding free energy for each ligand can also greatly speed convergence^{2,9}.

Ligand orientations were defined by the six relative degrees of freedom between the ligand and the protein – $\theta_A, \theta_B, \phi_A, \phi_B, \phi_C, r_{aA}$, as discussed previously^{2,9}. Reference atoms A, B, and C in the protein were taken to be the C, Calpha, and N of Tyrosine 88, as discussed previously⁹ and the reference atoms in the ligand were typically taken to be the first three heavy atoms in the ligand for algorithmic simplicity.

Absolute free energy calculations: Basic strategy

Each absolute binding free energy calculation involves three different sets of calculations with the ligand in the binding site, and then two with the ligand in water. Each set of calculations involves a final and an end state, and then several artificial “alchemical” intermediate states which serve to ensure phase space overlap; each of these states involves a separate simulation. For example, one component calculation is the discharging of the ligand in the binding site (turning off the partial charges on the ligand). In addition to running simulations with original partial charges on the ligand, and simulations with the charges turned entirely off, we run several intermediate simulations where ligand electrostatics interactions are reduced. These intermediate states are associated with λ values which describe the interaction strength; λ runs from 0 (full interactions) to 1 (no interactions).

The binding component calculations here are as follows. We first restrain the ligand in the binding site to the reference orientation, using a potential of the form $U = \frac{k\lambda}{2}(\xi - \xi_0)^2$ where ξ denotes the specific degree of freedom, ξ_0 the reference value, and k the base spring constant. k is chosen as previously to be 10 kcal/(mol angstrom²) for distance restraints and 10 kcal/(mol rad²) for angle/torsional restraints. To prevent large forces, distance restraints grow only linearly if the distance exceeds the reference distance by 0.2 nm, with the slope chosen to ensure continuous first derivatives. Lambda values were 0.01, 0.025, 0.05, 0.075, 0.1, 0.2, 0.35, 0.50, 0.75, 1.0. Reference atoms are as described above.

The next component calculation is discharging the ligand in the protein. For this we scaled back the ligand electrostatic potential by multiplying it by $(1 - \lambda)$ over a series of simulations. Lambda values were 0, 0.25, 0.5, 0.75, 1.0.

The final component calculation involves turning off the Lennard-Jones interactions between the ligand and the protein. We turned off just the interactions with the protein (leaving the intramolecular interactions) and used the modified soft core functional form for this transformation, as described previously^{2,9,13}. Lambda values were 0, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.65, 0.7, 0.75, 0.8, 0.85, 0.9, 0.95, and 1. For the ligand alone in water, only the charging and Lennard-Jones component calculations were done, using the same λ values as above. Simulation lengths were longer as noted above. Instead of performing simulations of removing the restraints, we computed the free energy of removing the restraints analytically using the approach of Boresch *et al.*¹².

Orientational Decomposition

In the absolute binding free energy calculations, we computed contributions of different metastable orientations for ligands separately, and then combined the effective binding free energies for these different orientations into a total binding free energy using the formalism described and applied previously^{2,9}. A key requirement for this formalism is that the orientations not cover the same regions of phase space, hence we had to define the orientations. We used the six relative degrees of freedom between the ligand and the protein as a starting point, we took the histograms of the degrees of freedom from the unrestrained simulations of each ligand in its different orientations and defined orientations using the single degree of freedom that separated the orientations well (based on the full-width at half maximum for the largest peak in the probability distribution); boundaries between orientations were chosen to split the distance between peaks. We then manually inspected all decompositions. In some cases this approach failed, and we used the histograms from the restrained simulations instead. In some cases (especially those with relatively more orientations) no single degree of freedom could separate orientations well, in which case we manually defined decompositions based on an inspection of the histograms from the unrestrained and fully restrained simulations.

Once our orientations were defined, we filtered the simulations for each orientation and discarded snapshots that did not fall within the specified orientation. Filtering typically had significant effects only at the early stages of the restraining calculations, when restraints were weak, and effects were minimal for the charging and van der Waals components of the calculations. In all cases, unless otherwise noted, at least half of the (2000) coordinate snapshots remained after filtering. Energies were stored more frequently (5000 snapshots). Energy snapshots taken between coordinate

snapshots that remained in the orientation of interest were assumed to fall in the orientation of interest (that is, transitions were assumed to be slow compared to the snapshot frequency), and energy snapshots taken between two coordinate snapshots that showed a transition were assumed to be outside the region of interest.

Confine-and-Release Approach

The umbrella sampling simulations here used 36 windows, spaced every 10 degrees, with spring constants of 400 kJ/(mol rad²) and production simulations were 100 ps per window, with equilibration and other settings as described above. Data was analyzed as in our previous study³. Umbrella sampling was done separately for each ligand orientation.

To apply the confine-and-release procedure, we need to be sure that we do not doublecount the free energy associated with conformational changes, which would happen if we sampled the motion during the binding calculations and also applied the free energies from the PMFs. Therefore, we analyzed all of the simulations to ensure the absolute free energy analysis excluded all simulation snapshots where the sidechain rotamers for these sidechains changed rotameric state relative to the apo starting structure.

Water removal

We used absolute binding free energy techniques to compute the free energy of removing waters from the binding site. We placed one, two, three, and four waters into the binding site, and computed the free energy of removing one water molecule from the binding site and reinserting it into bulk using our standard protocols². These free

energies were always favorable (typically by roughly 4 kcal/mol) indicating that it is unfavorable to have waters in the binding site, and thus that it will typically be vacant even in the absence of ligand. This is consistent with crystallographic data for small apolar binding sites¹⁴.

Relative Free Energy Calculations

As mentioned in the main text, we retained all snapshots when combining contributions of different orientations in the relative free energy calculations. We believe retaining all snapshots when combining contributions of different orientations is justified in this case for three reasons: (1) Results from different orientations were typically different, suggesting that ligands typically did not sample the same orientations when beginning from different orientations, and a visual inspection of some trajectories supported this; (2) Relatively few cases had multiple orientations within kT of one another, so in most cases a single orientation predominated; (3) Past experience in this binding site suggests that ring flips out of the plane are extremely rare for steric reasons, suggesting transitions between orientations differing by rotation of the ring around the axis of the hydroxyl would be very rare.

C.2 Supplementary Figures for Chapter 3

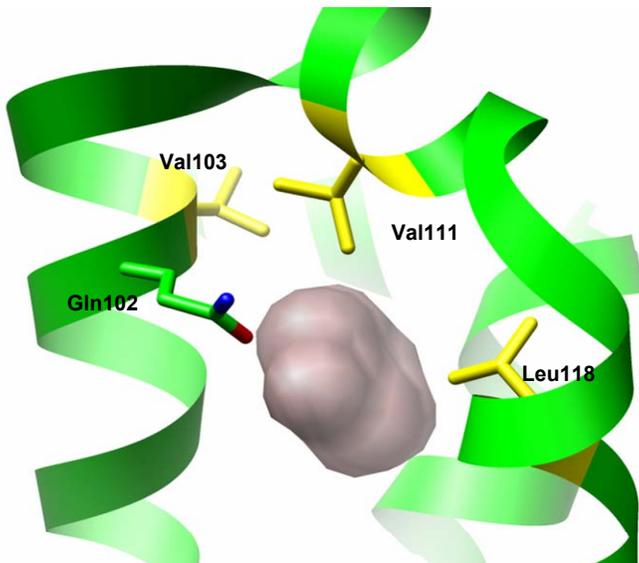


Figure S1. The L99A/Gln102 binding site. Sidechains with dihedrals that are explicitly sampled by confine-and-release protocol are shown in yellow.

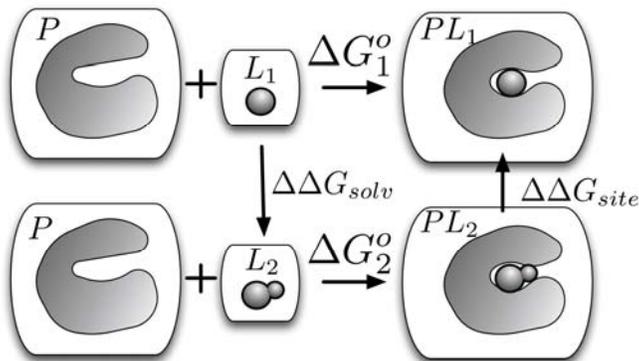


Figure S2. Thermodynamic cycle for computing relative binding free energies. Ligand 1 (L_1) is mutated into ligand 2 (L_2) in solvent (left) and L_2 into L_1 in the binding site (right), to calculate free energy changes $\Delta\Delta G_{solv}$ and $\Delta\Delta G_{site}$, respectively. The relative binding free energy we are interested in is $\Delta\Delta G_{1,2} = \Delta G_2 - \Delta G_1 = -(\Delta\Delta G_{solv} + \Delta\Delta G_{site})$. Further details are given in the methods section.

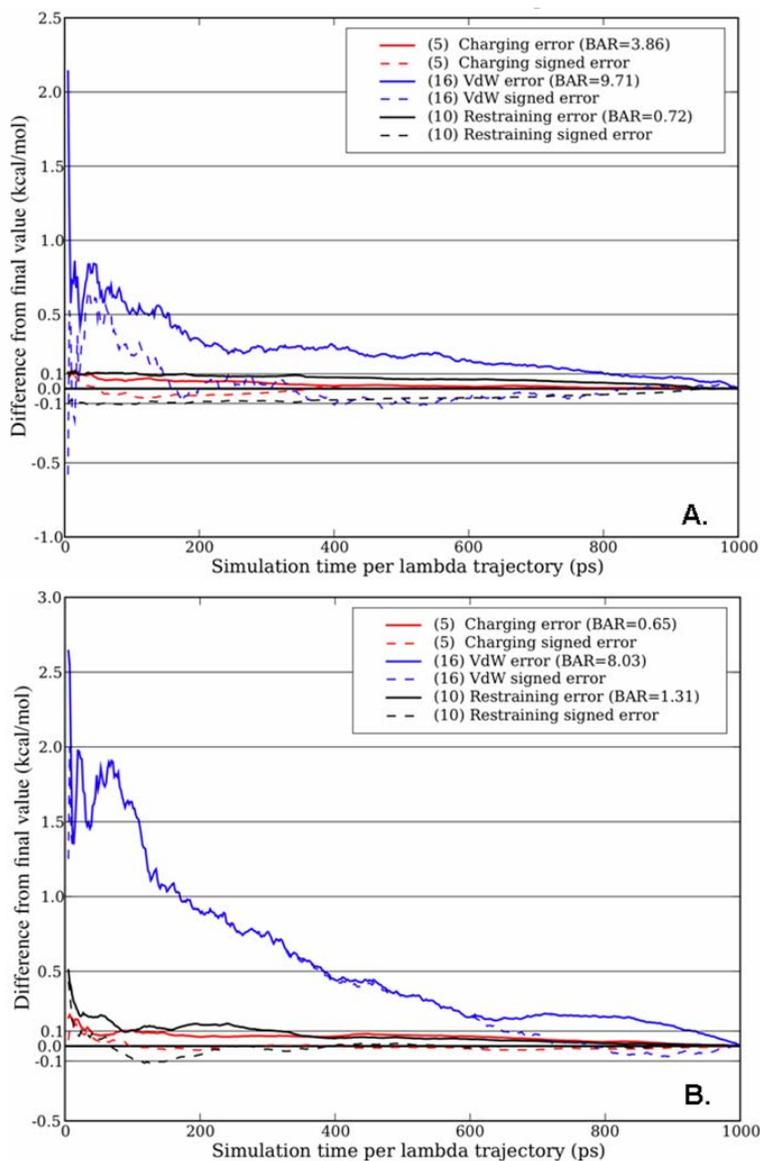


Fig. S3. Representative plots of convergence with simulation time of free energy estimates for alchemical transformation component calculations. Dashed lines represent the actual convergence error relative to complete 1-ns trajectories for each component of the alchemical transformation. Solid lines represent an artificial, worse-case estimate of the possible convergence error, if the error from each individual lambda step of a given component had been in the same direction. **A.** Thieno-32b-thiophene's dominant orientation, showing rapid

convergence. **B.** Nitrosobenzene's single orientation, showing slower convergence of the van der Wals decoupling due to cavity side chain motions. Many simulations had unique convergence issues in all three of the transformation component steps.

C.3 Supplementary Tables for Chapter 3

Table S1. Retrospective Absolute Free Energy Calculations

		Binder a.	Prediction b.	$\Delta G_{b,exp}$ (Kcal/mol) c.	$\Delta G_{b,calc}$ (Kcal/mol) d.	RMSD (Å) e.
3-chlorophenol		Yes	Yes	-5.51	-4.07 ± 0.23	2.843, 0.469*
phenol		Yes	Yes	-5.23	-4.39 ± 0.09	1.563, 2.788*
2-fluoroaniline		Yes	Yes	-5.18	-4.72 ± 0.43	0.565, 1.460
toluene		Yes	Yes	-4.93	-3.47 ± 0.06	-
3-methylpyrrole		Yes	Yes	-4.92	-3.28 ± 0.10	0.822(A), 1.870(B)
catechol		Yes	Yes	-4.16	-3.14 ± 0.13	0.997(A), 2.196(B)
3,5-difluoroaniline		Yes	Yes	1.7 [†]	-4.63 ± 0.15	0.730(A), 0.647(B)
4-vinylpyridine		No	weak	-	-3.39 ± 0.14	-
2-aminophenol		No	Yes	-	-3.93 ± 0.10	-

A. Experimentally determined binder (yes) or non-binder (no).

B. Prediction; binder (yes) or non-binder (no).

C. Free energy of binding determined by ITC at 10°C. [†] ΔT_m (°C) measured by CD at concentrations between 1-10mM at pH 3.0.

D. Calculated free energy of binding.

E. RMSD of prediction ligand geometry to experimentally observed crystal pose. If multiple ligand orientations were present in the crystal (designated A and B) and only one pose was predicted by the free energy methods, the RMSD for both poses is reported.

Table S2. X-ray data collection and refinement

	4-chloro-1h-pyrazole	thiophene-2-carboxaldehyde	2-ethoxy-3,4-dihydro-2h-pyran	benzylacetate	nitrosobenzene	4,5,6,7-tetrahydroindole	thieno[3,2-b]thiophene
pH of soaking buffer	6.50	6.50	7.00	6.50	7.00	7.00	7.00
Cell dimensions:							
a=b (Å)	59.960	60.260	60.110	60.024	60.579	60.327	60.060
c (Å)	96.600	96.940	96.440	94.468	97.032	97.438	96.990
Resolution (Å)	1.85 (1.95) ^a	1.40 (1.48) ^a	1.26 (1.32) ^a	1.29 (1.34) ^a	1.46 (1.51) ^a	1.35 (1.385) ^a	1.45 (1.54) ^a
Number of Unique Reflections	17,665 (2532) ^a	40,713 (6131) ^a	53,794 (7247) ^a	51,404 (5080) ^a	36,042 (3559) ^a	40,838 (4019) ^a	36,574 (5891) ^a
R _{merge} %	9.8 (30.1) ^a	8.7 (57.6) ^a	6.5 (25.7) ^a	5.5 (48.7) ^a	8.0 (34.5) ^a	7.1 (50.6) ^a	4.8 (14.3) ^a
Completeness %	99.6 (99.2) ^a	99.8 (99.7) ^a	95.3 (76.7) ^a	100 (100) ^a	99.7 (99.9) ^a	100 (100) ^a	99.7 (98.7) ^a
<I>/<σ(I)>	16.01 (8.86) ^a	11.36 (2.67) ^a	15.46 (4.80) ^a	54.9 (43.8) ^a	59.8 (6.98) ^a	53.7 (4.88) ^a	53.13 (24.53) ^a
R _{free} %	20.8	19.0	18.1	18.5	20.2	19.2	18.9
R-factor %	16.3	16.7	16.4	16.6	16.9	17.3	17.0
Resolution Range (Å)	30.0 - 1.85	30.0 - 1.40	30.0 - 1.26	51.99 - 1.29	30.0 - 1.46	52.27 - 1.35	30 - 1.45
Abondlengths (Å)	0.011	0.008	0.006	0.006	0.011	0.007	0.008
Abondangles (°)	1.273	1.264	1.084	1.094	1.257	1.140	1.308
Average B-factor protein atoms (Å ²)	14.106	15.251	12.662	14.40	18.88	15.49	11.73
Average B-factor ligand atoms (Å ²)	15.69	19.62	18.82	18.73	17.03	15.31	13.33

^a Values in parentheses are for the highest resolution shell. Proteins are crystallized in space group P32₁2₁.

Table S3. Comparison of preferred rotameric state predicted by confine-and-release protocol, and the crystallographic results

	<i>prediction / experiment</i>		
	Val111	Val103	Leu118
thieno-3,2b-thiophene	alternate (0.7) / apo	apo	equal / apo
nitrosobenzene	partial alt / apo	apo	equal / apo
2-nitrothiophene	alternate	apo	alternate / apo
4-chloro-1h-pyrazole	partial alt / apo	apo	partial alt / apo
4,5,6,7-tetrahydroindole	alternate (0.7) /apo	apo	equal
(Z)-thiophene-2-carboxaldoxime	equal / apo	apo	partial alt / equal
n-phenylglycinonitrile	apo	apo	apo
benzylacetate	partial alt / apo	apo / equal	partial alt / apo
2-ethoxy-3,4-dihydro-2h-pyran	apo	apo / equal	apo

Agreement between experiment and calculation is indicated by ***bold italics***. Observed sidechain rotamers of Val111, Val103, Leu118 are designated the following: **Apo**- the rotamer occupied is the same as the one observed in apo crystal structure; **Alternate**- an alternate rotamer is observed contributing greater than 0.7 kcal/mol to the binding free energy; **Partial alternate**- low energetic contribution from an alternate rotamer 0.04-0.3 kcal/mol; **Equal**- an alternate rotamer is equiprobable contributing typically 0.3-0.5 kcal/mol.

Table S4. Comparison of absolute free energy predictions beginning from the Apo vs. Holo structure with corrected long range dispersion term and corrected partial charges.

	$\Delta G_{b, calc}$ (kcal/mol)		RMSD to xtal pose (Å)	
	apo a.	holo b.	apo c.	holo d.
1-phenylsemicarbazide	0.21 ± 0.10	---	---	---
o-benzylhydroxylamine	-1.91 ± 0.11	---	---	---
1-2-hydroxyethylpyrrole	-6.16 ± 0.37	---	---	---
phenylhydrazine	-1.86 ± 0.25	---	---	---
(E)-thiophene-2-carboxaldoxime	-2.52 ± 0.13	---	---	---
(Z)-thiophene-2-carboxaldoxime	-6.81 ± 0.12	-5.06 ± 0.09	1.88(A); 1.88(B)*	0.73(A), 0.81(B)
4-chloro-1h-pyrazole	-5.61 ± 0.10	-4.61 ± 0.10 [†]	2.07	1.80(D), 2.06(E)
2-ethoxy-3,4-dihydro-2h-pyran	-4.53 ± 0.18	-4.34 ± 0.20	4.35	3.87, 5.41
nitrosobenzene	-5.36 ± 0.12	-3.93 ± 0.18	3.24(A), 3.22(B)	0.41(A), 2.14(B)
benzyl acetate	-1.45 ± 0.50	-3.89 ± 0.18	>10	1.28(A), 0.69(B)
4,5,6,7-tetrahydroindole	-3.02 ± 0.14	-2.45 ± 0.19	0.66 & 1.78**	0.64
thieno-3,2b-thiophene	-5.43 ± 0.09	-4.33 ± 0.08 (1) -6.86 ± 0.59 (2)	0.73(A), 0.45(B)	0.57 & 0.73(1); 0.75 & 0.81(2)
2-nitrothiophene	-5.27 ± 0.12	-11.40 ± 0.05	1.09(A), 2.86(B)	0.66(A), 1.16(B)
n-phenylglycinonitrile	-5.82 ± 0.37	-6.09 ± 0.10	0.87	1.02

A. Calculated free energy of binding from the apo structure with corrected long range dispersion term and charges for all compounds.

B. Calculated free energy of binding from the holo structure for the nine ligands. Thieno-3,2b-thiophene has two starting holo structures and therefore two predictions each, designated (1) & (2).

C-D. RMSD of prediction ligand geometry to experimentally observed crystal pose from apo (D) and holo (E) result. If multiple ligand orientations were present in the crystal (designated A and B) only the best RMSD to the prediction is reported. ** indicates RMSD for one crystal pose calculated to two predictions.

Table S5. X-ray data collection and Refinement

	2-methylphenol	2-ethylphenol	2-propylphenol	2-ethoxyphenol	2-methoxyphenol	5-chloro-2-methylphenol
pH of soaking buffer	6.50	6.50	6.50	6.50	6.50	6.50
Cell dimensions:						
a=b (Å)	60.281	60.266	59.950	60.140	60.090	60.480
c (Å)	98.110	97.140	96.610	96.160	96.450	97.680
Resolution (Å)	1.59 (1.67) ^a	1.70 (1.77) ^a	1.81 (1.86) ^a	1.80 (1.93) ^a	2.02 (2.15) ^a	1.60 (1.64) ^a
Number of Unique Reflections	26,865 (4027) ^a	21,860 (2897) ^a	17,966 (1463) ^a	19,222 (3565) ^a	13,727 (2309) ^a	26,490 (3574) ^a
R _{merge} %	8.1 (49.6) ^a	7.5 (47.9) ^a	8.1 (66.1) ^a	8.9 (42.5) ^a	9.9 (37.3) ^a	7.1 (56.1) ^a
Completeness %	99.75 (100) ^a	99.78 (100) ^a	95.1 (100.0) ^a	99.9 (100) ^a	99.9 (99.8) ^a	99.89 (97.1) ^a
$\langle I \rangle / \langle \sigma(I) \rangle$	33.2 (4.24) ^a	39.6 (4.84) ^a	22.48 (5.61) ^a	17.37 (4.68) ^a	18.18 (5.88) ^a	37.6 (2.46) ^a
R _{free} %	22.0	22.2	24.5	21.2	21.9	21.4
R-factor %	18.2	18.6	19.7	18.0	17.1	18.8
Resolution Range (Å)	52.20 - 1.59	52.20 - 1.70	30.0 - 1.81	30.0 - 1.80	30.0 - 2.02	52.34 - 1.60
Abondlengths (Å)	0.009	0.011	0.024	0.012	0.015	0.011
Abondangles (°)	1.133	1.218	1.966	1.273	1.423	1.193
Average B-factor protein atoms (Å ²)	18.41	20.98	15.44	16.16	17.67	21.81
Average B-factor ligand atoms (Å ²)	15.02	20.75	15.55	26.83	11.81	27.84

^a Values in parentheses are for the highest resolution shell. Proteins are crystallized in space group P32₁2₁.

Table S6. Relative binding free energy results recalculated with corrected long range dispersion term.

Compound	catechol (kcal/mol):		phenol (kcal/mol):			
	uncorrected	corrected	uncorrected	corrected	uncorrected	corrected
	$\Delta\Delta G_{\text{exp,lig-ref}}$	$\Delta G_{\text{calc,lig-ref}}$	$\Delta G_{\text{calc,lig-ref}}$	$\Delta\Delta G_{\text{exp,lig-ref}}$	$\Delta G_{\text{calc,lig-ref}}$	$\Delta G_{\text{calc,lig-ref}}$
2-propylphenol	-1.17	-1.97 ± 0.20	-2.15 ± 0.23	-0.09	2.19 ± 0.07	2.06 ± 0.10
phenol	-1.08	-	-	-	-	-
5-chloro-2-methylphenol	-0.88	-3.38 ± 0.13	-2.95 ± 0.18	0.2	3.92 ± 0.18	3.86 ± 0.24
2-ethylphenol	-0.41	-0.99 ± 0.09	-1.26 ± 0.11	0.67	4.10 ± 0.33	4.46 ± 0.42
2-methylphenol	-0.28	-0.15 ± 0.14	-0.46 ± 0.16	0.8	2.69 ± 0.08	2.80 ± 0.09
catechol	-	-	-	1.08	6.99 ± 0.34	0.36 ± 0.20
2-ethoxyphenol	0.14	0.00 ± 0.09	0.02 ± 0.10	1.22	0.99 ± 0.12	1.09 ± 0.08
2-methoxyphenol	2.16	1.40 ± 0.19	1.77 ± 0.19	3.24	1.40 ± 0.19	2.17 ± 0.29

Column 1. Experimentally determined binding free energy of phenol derivatives relative to catechol. Column 2 & 3. Uncorrected and corrected calculated free energy of binding relative to the reference compound catechol. Column 4. Experimentally determined binding free energy of phenol derivatives relative to phenol. Column 5 & 6. Uncorrected and corrected calculated free energy of binding relative to the reference compound phenol.

C.4 References

1. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. (1983). Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926-935.
2. Mobley, D. L., Graves, A. P., Chodera, J. D., McReynolds, A. C., Shoichet, B. K. & Dill, K. A. (2007). Predicting absolute ligand binding free energies to a simple model site. *J. Mol. Biol.* **371**, 1118-1134.
3. Mobley, D. L., Chodera, J. D. & Dill, K. A. (2007). Confine-and-release method: Obtaining correct binding free energies in the presence of protein conformational change. *J. Chem. Theory Comput.* **3**, 1231-1235.
4. Mobley, D. L., Bayly, C. I., Cooper, M. D., Shirts, M. R. & Dill, K. A. (2009). Small molecule hydration free energies in explicit solvent: An extensive test of fixed-charge atomistic simulations. *J. Chem. Theory Comput.* **5**, 350-358.
5. Mobley, D. L., Dumont, I., Chodera, J. D. & Dill, K. A. (2007). Comparison of charge models for fixed-charge force fields: Small-molecule hydration free energies in explicit solvent. *J. Phys. Chem. B* **111**, 2242-2254.
6. Nicholls, A., Mobley, D. L., Guthrie, P. J., Chodera, J. D. & Pande, V. S. (2008). Predicting Small-Molecule Solvation Free Energies: An Informal Blind Test for Computational Chemistry. *J. Med. Chem.*
7. Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A. & Case, D. A. (2004). Development and testing of a general amber force field. *J. Comp. Chem.* **25**, 1157-1174.
8. Bayly, C. I. (2009). (Mobley, D., ed.).

9. Mobley, D. L., Chodera, J. D. & Dill, K. A. (2006). On the use of orientational restraints and symmetry corrections in alchemical free energy calculations. *J. Chem. Phys.* **125**, 084902.
10. Essman, U., Perera, L., Berkowitz, M., Darden, T., Lee, H. & Pedersen, L. (1995). A smooth particle mesh Ewald method. *J. Chem. Phys.* **103**, 8577-8593.
11. Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, A. & Haak, J. R. (1984). Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **81**, 3584-3690.
12. Boresch, S., Tettinger, F., Leitgeb, M. & Karplus, M. (2003). Absolute binding free energies: A quantitative approach for their calculation. *J. Phys. Chem. A* **107**, 9535-9551.
13. Shirts, M. R. & Pande, V. S. (2005). Solvation free energies of amino acid side chains for common molecular mechanics water models. *J. Chem. Phys.* **122**, 134508.
14. Matthews, B. W. & Liu, L. (2008). A review about nothing: Are apolar cavities in proteins really empty? *Protein Science* **18**, 494-502.

Publishing Agreement

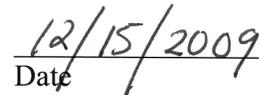
It is the policy of the University to encourage the distribution of all theses, dissertations, and manuscripts. Copies of all UCSF theses, dissertations, and manuscripts will be routed to the library via the Graduate Division. The library will make all theses, dissertations, and manuscripts accessible to the public and will preserve these to the best of their abilities, in perpetuity.

Please sign the following statement:

I hereby grant permission to the Graduate Division of the University of California, San Francisco to release copies of my thesis, dissertation, or manuscript to the Campus Library to provide access and preservation, in whole or in part, in perpetuity.



Author Signature



Date