

UCSF

UC San Francisco Electronic Theses and Dissertations

Title

Computational methods as applied to molecular recognition in RNA

Permalink

<https://escholarship.org/uc/item/8s6231jj>

Author

Lang, P. Therese

Publication Date

2006

Peer reviewed|Thesis/dissertation

COMPUTATIONAL METHODS AS APPLIED TO
MOLECULAR RECOGNITION IN RNA

by

P. Therese Lang

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

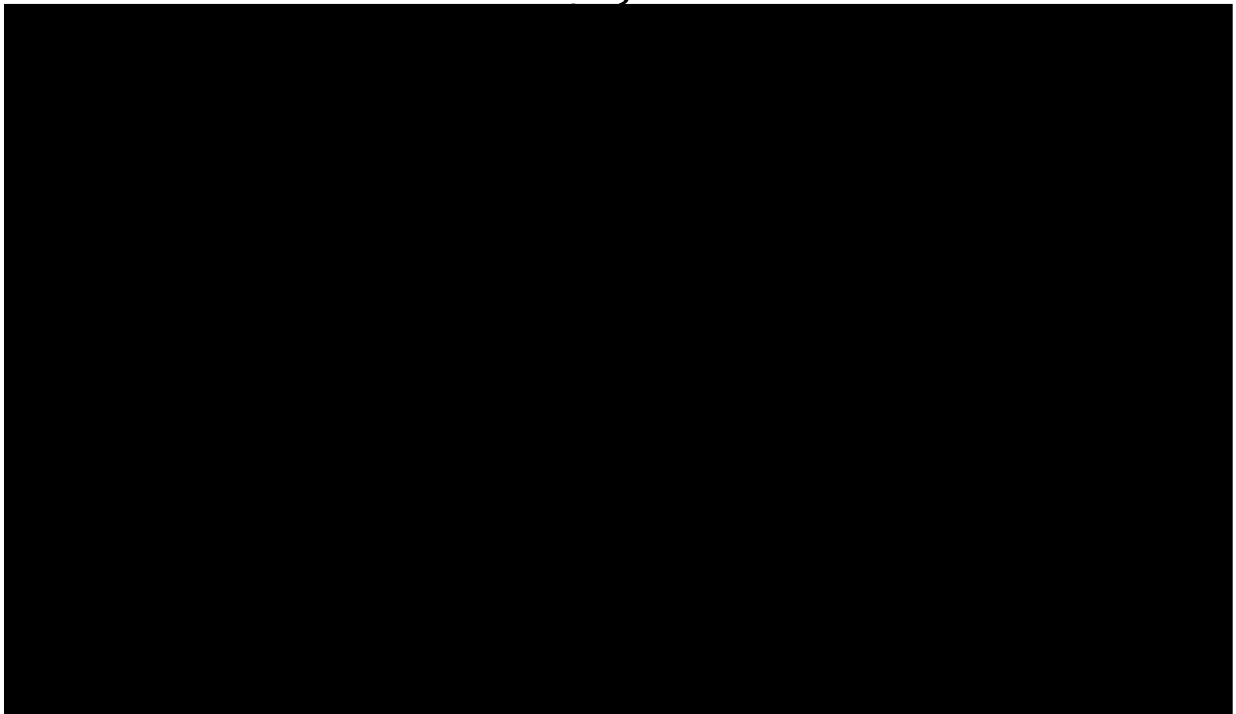
Chemistry and Chemical Biology

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, SAN FRANCISCO



“There is science, logic, reason; there is thought verified by experience. And then there is California.”

--Edward Abbey

PREFACE

Getting to this point in my life has been so much more than 5 ½ years. It is the culmination of over 2 ½ decades of study and the support and guidance of hundreds of friends, teachers, and loved ones. Thank you to all the people who have touched my life.

To Tack, thank you for taking me on as your last graduate student. For your insights into drug design at every level. For continually reminding me that part of collaborating is knowing when to adjust to their working style and when to push a little. For teaching me the fine art of being critical in a tactful way. For giving me enough rope to explore and for providing with a way out when I got tied into knots.

To Tom, thank you for welcoming me into your lab. For your insights into the trials and tribulations of NMR data and working with RNA. For waiting patiently while I work on Tack’s projects (I will get to that RNA stuff soon, really). For guidance on internal politics and collaborations. For giving me the freedom to design my own projects and trusting me to run with them.

To Jim Frazine and Song Ling, thank you for making the computers, and thus my life, run as smoothly as possible. You have both been sorely missed.

To all former and current members of the James and Kuntz labs, thank you for your insight, suggestions, and companionship.

To **Matt**, thank you providing guidance from a younger perspective. For continually challenging me and my ideas. For allowing me to tag along as an honorary member of your group.

To **Kip**, thank you for our continuing collaborations. For teaching me to manage a wide range of people and projects. For always making me feel that I deserved to be involved and that computation, while not always perfect, is a critical part of the evolution of drug design.

To **Ken Dill**, **Martin Shetlar**, and **Tom Cheatham**, thank you for helping me through orals. To **Tom** in particular, thank you for providing an outside perspective so early in my career.

To **Charles Craik** and **Christine Olson**, thank you for organizing and running the **CCB** program. I have learned a great deal about both chemical biology and drug design as a result of the environment you have created.

To **Ian**, **John**, **Vince**, **Ben**, **Arjun**, **Scott**, **Eric**, **Brian**, **Rob**, **Amy**, **Kristi**, **Kat**, **Petri**, and all the other friends I have made while at **UCSF**. Thank for enriching my life inside the lab and out with both meaningful conversations and meaningless diversions.

To **Dr. Shalhoub**, thank you for introducing me to the field of computation. For continually challenging me and guiding me to continually challenge myself. For showing me that teaching is so much more than just presenting information.

To **Dr. Jones**, thank you for suggesting I try an independent study. For convincing me from the first semester of undergrad that chemistry is far superior

to biology. For providing me with a role model of a scientist, a teacher, a mother, and a woman.

To my Nana and Poppop, thank you for teaching me that throwing risk to the wind is sometimes worth it, particularly if you get to take a warm bath and a nap after. For encouraging my imagination through the exploration of ruins, both old and new, and through tea parties, and card houses, and easter egg hunts, and hikes, and many other life-threatening adventures.

To my Dadoo, thank you for warm memories of crab bakes and Musikfest and a feeling of community.

To Dorothy, thank you being the older sister I never had. For talks that are never quite long enough and mountains of useful advice.

To the remainder of my extended family, thank you for the many ways you have cared for and supported me through the years.

To Jack, thank you for keeping me and my ego grounded. For always being ready to make a joke or inspire an eye roll.

To Colleen, thank you for being the rebel I never could be. For commiserating over life's little dips. For being ready to jump on a plane when I really need you.

To my father, thank you for laughing with me when life gets too serious and laughing at me when I take my life too seriously. For going toe-to-toe with all those nuns over the years and trying to convince them that little girls should be encouraged in math and science no matter what. For taking me to all those

museum exhibits. For Blizzards and hitting golf balls, for late nights and Hershey kisses, for really big hugs.

To my mother, thank you for continuing to try to understand me both in my career and my life. For help with all those science posters, which it turns out I DO still have to put together. For driving back to school to get the books I forgot...again. For letting me dissect weird animals in your backyard and encouraging my interests in both traditional and less-than-traditional pursuits for my gender. For boy walks, for head scratches, for finding my ten-year old high school biology notebook so I could cite it in my intro.

To my husband, Ed, thank you for your love and support. For coming on this adventure to the left coast with me. For never letting me give up no matter how tough things got. For always encouraging me to stand up for myself. For reminding me that things are never as bad as I think they are. For making sure I did not starve to death when I was taking classes or studying for orals. For exploring with me, for continuing to surprise me, for letting me hold the remote control. I never could have made it through without you.

More specific to the work in my thesis, I would like to additionally acknowledge the contributions of the various authors I have listed for each chapter. The first chapter was truly a group effort, with each author contributing individual sections. In addition to my section, I smoothed over some of the transitions and added references. For chapter two, the other authors provided a great deal of guidance and advice both when designing the judging criterion and when drafting the paper.

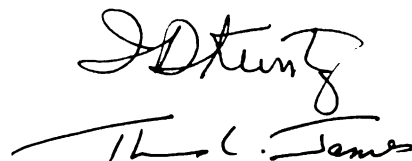
Chapter three was the culmination of many student years of work. Demer wrote the majority of the DOCK 5 code with a bit of help from Scott and me. Natasja put together the original protein test set, which was used for the paper with only minor modifications. Eric coaxed Chimera into successfully working as an all-in-one structure preparation tool. Rob kindly provided his computer cluster for testing and debugging the parallel code.

In chapter four, I served more of an advisory role. Peter and Sabine synthesized all the compounds. Moriz screened the majority of the compounds, with Irene providing a few follow-up experiments to address reviewer concerns. I was originally involved in the project to provide modeling, which was not used in the final paper. However, I also helped analyze the experimental results and developed the final version of the structure-activity relationship that appeared in the paper.

The methods in chapter five represent a large coding effort performed primarily by Scott, and Kaushik, and myself with help from Sudipto. Veena assisted with the preparation of the original binding mode test set, which was

later updated for the final version of the chapter. Alan provided invaluable assistance with the preparation procedures for the DOCK 3.5 scoring function, and Scott ran and analyzed the calculations for the protein tests.

Chapters six and seven were, once again, a group effort. For the first, John provided assistance both with the replica exchange procedure and the clustering algorithm. Nick and Chris ran the miniCarlo and PEDC simulations, respectively, and provided the structures for analysis. In the final chapter, David performed all the synthesis, Irene screened the compounds, and Anang ran all Pipeline Pilot calculations.

A handwritten signature in black ink, appearing to read "T. C. James". The signature is written in a cursive style with a long horizontal stroke extending to the left.

ABSTRACT

The aim of this thesis was to improve computational methods for structure based drug design, particularly for RNA targets. For the first portion, we present a critical evaluation of various computational drug design algorithms for their ability to predict experimental binding poses and rank libraries of small molecules against protein targets. In particular, we characterize the strengths and weaknesses of the ligand sampling method for the DOCK suite of programs. In the second portion of the thesis, we apply the lessons learned from protein targets to the disruption of protein-RNA interactions critical to the life cycle of the HIV virus. As a class, RNA historically has presented a difficult computational challenge both due to its highly localized charge and flexibility. Therefore, we have extended protocols and added new protocols in existing software packages such as DOCK and AMBER to predict experimental binding poses, once again validating our results using experimental data. Finally, we apply these protocols both to develop libraries of small molecules against druggable RNA targets and to establish a fragment-based library designed to find new scaffolds for RNA.

TABLE OF CONTENTS

Introduction	1
Chapter 1: Molecular Docking and Structure-Based Design.....	12
Abstract	13
Introduction.....	14
Molecular Docking.....	18
<i>Overview.....</i>	<i>18</i>
<i>Receptor Site Identification.....</i>	<i>19</i>
<i>Receptor Site Characterization.....</i>	<i>22</i>
<i>Orientation of the Ligand in the Target Site</i>	<i>24</i>
<i>Evaluation of Ligand Orientations.....</i>	<i>25</i>
Ligand Structure Generation	28
Description of Docking Programs	30
Tests of Docking and Structure-Based Design.....	30
Conclusions and Future Directions.....	40
Acknowledgments	41
References	41
Chapter 2: Evaluating the High-Throughput Screening Computations	47
Abstract	48
Introduction.....	49
Evaluating the Entries	52
Implications for Experimental Design	55

Implications for Computational Methods	56
Recommendations for Future Competitions	57
Conclusions.....	58
Acknowledgements	58
References	59
Chapter 3: Development and Validation of a Modular, Extensible Docking	
Program: DOCK 5.....	60
Abstract	61
Introduction.....	62
<i>DOCK Background</i>	63
<i>Overview of Test Set</i>	69
Methods.....	72
<i>DOCK 4 to DOCK 5 Conversion</i>	72
<i>Conversion of the DOCK Codebase from C to C++</i>	74
<i>Test Set Preparation</i>	75
<i>Optimized Hydrogen Locations for Test Set Receptors</i>	77
<i>Selection of Active Site Waters</i>	78
<i>DOCK Parameter Optimization</i>	79
<i>Greedy Clustering of Conformational Ensembles</i>	79
<i>Evaluation of MPI Functionality</i>	80
Results	81
<i>Rigid Ligand DOCKing</i>	82
<u>Overall Performance</u>	82

<u>Dependence on Ligand Conformation</u>	82
<u>Analysis of Total Orientational Ensemble</u>	83
<u>Geometric Clustering of Poses</u>	87
<i>Flexible Ligand DOCKing</i>	87
<u>Overall Performance</u>	87
<u>Dependence on Anchor Position</u>	87
<u>Analysis of Total Conformational Ensemble</u>	88
<u>Geometric Clustering of Poses</u>	89
<i>Comparison to DOCK 4</i>	89
<i>Comparison to Other Docking Methods</i>	90
<i>Analysis of Successes and Failures of DOCKing Protocols</i>	90
<u>Failures Resulting from Receptor Modeling/Structural</u>	
<u>Problems</u>	92
<u>Failures Resulting from Ligand Flexibility</u>	94
<u>Sampling Versus Scoring Failures</u>	96
<i>Analysis of DOCK Score for DOCKing Protocols</i>	98
<i>Database DOCKing using MPI</i>	101
Discussion	101
Conclusions.....	104
Acknowledgements	104
Appendix	105
<i>Rigid DOCKing Parameter Optimization</i>	105
<i>Flexible DOCKing Parameter Optimization</i>	107

References	113
Chapter 4: Synthesis and Testing of a Focused Phenothiazine Library for Binding to HIV-1 TAR RNA	116
Summary	117
Introduction.....	118
Results and Discussion	120
<i>Chemistry</i>	120
<i>Library Screening</i>	124
Significance	135
Experimental Procedures	136
<i>NMR Experiments</i>	136
<i>Synthesis of Compounds</i>	139
Acknowledgements	142
References	142
Chapter 5: Optimization of DOCK for RNA Targets	145
Abstract	146
Introduction.....	147
<i>New Features of DOCK</i>	148
<i>Development of Test Set</i>	151
Methods Applied to Date	154
<i>Test Set Preparation</i>	154
<i>Modification of Pruning Algorithm</i>	156
<i>Scaling of Repulsive VDW Radii</i>	158

<i>Modification of Bump Filter</i>	159
<i>Optimization of Parameters for DOCKing</i>	161
Results and Discussion of Current Status	162
<i>Rigid Ligand Sampling Optimization</i>	162
<i>Flexible Ligand Sampling Optimization</i>	164
<i>Success as a Function of the Number of Rotatable Bonds</i>	171
<i>Comparison Between Versions of DOCK</i>	172
<i>Examination of Ensemble of Generated Orientations</i>	175
Conclusions.....	177
Future Directions	178
References	180

Chapter 6: Steps Toward Fully Flexible Docking to RNA Targets:

Strengths and Weaknesses of Various Sampling Algorithms in Generating Ensembles of RNA Structures	184
Abstract	185
Introduction.....	186
Methods Employed to Date	192
<i>Selection of TAR Structure</i>	192
<i>Molecular Dynamics Simulations</i>	193
<i>Path Exploration with Distance Constraints</i>	196
<i>MiniCarlo</i>	197
<i>Culling the Conformations</i>	198
<i>TAR RNA Small Molecule Test Set</i>	200

Results and Discussion of Current Status	200
<i>Filtering</i>	200
<i>Replica Exchange Simulation</i>	209
Preliminary Conclusions	212
Future Directions	214
<i>Clustering</i>	214
<i>Analysis of Conformational Ensembles</i>	214
<i>Cross Docking</i>	215
References	215
 Chapter 7: A Fragment-Based Screening Method Designed for RNA	
Targets	219
Abstract	220
Introduction.....	221
Methods Employed to Date	226
<i>Design of Fragment Library</i>	226
<i>Detection of Binding by STD NMR</i>	228
<i>Initial Optimization of Synthetic Conditions</i>	228
<i>Synthesis of Scaffolds from Fragment Library</i>	230
Discussion of Current Status	232
<i>Synthesis</i>	232
<i>Binding of Fragments</i>	232
Preliminary Conclusions	233
Future Directions	233

<i>Fragment Library</i>	233
<i>Synthesis</i>	234
<i>Modeling</i>	234
<i>Binding</i>	235
References	236
Conclusions	239

LIST OF FIGURES

Chapter 1

1. Example flow of drug design process from both experimental and computational perspectives..... 15
2. Predicted complexes versus X-ray crystallographic structures that were subsequently determined 34
3. Docking predicted ligands from virtual screening against simple cavity sites..... 35
4. Predicted vs. experimental structures from virtual screening 37

Chapter 2

1. Timeline and major events in the McMaster University High-Throughput Data-Mining and Docking Competition..... 51

Chapter 3

1. The “anchor and grow” conformational search algorithm 66
2. The major DOCK 5 classes and their interconnections..... 68
3. Rigid and flexible docking success rate as function of conformation and anchor perturbation, respectively 84
4. Rigid and flexible docking success rate as function of energy gap and ranked conformers 85
5. Rigid and flexible docking success rates as function of number of cluster heads examined..... 86

6. Correlation of flexible ligand success and failure rates with crystallographic resolution and experimental B-factor.....	93
7. Rigid and flexible docking success and failure rates as a function of the number of rotatable bonds in each ligand	93
8. Score from the top-ranking pose for both rigid and flexible ligand docking that were successful against the complex crystal structure score	99
9. Speedup for docking libraries of small molecules	100
10. Optimization of parameters for rigid ligand docking	106
11. Optimization of parameters for flexible ligand docking	111

Chapter 4

1. Binding site of TAR RNA with acetylpromazine bound.....	121
2. Synthesis of a 10 <i>H</i> -phenothiazine library.....	123
3. Binding between phenothiazine ring derivatives and TAR RNA.....	126
4. Phenothiazine scaffolds with side chain substitutions tested for binding to TAR RNA	127
5. Representative NMR spectra of phenothiazine compounds	131
6. Binding affinities as determined by chemical shift monitoring of TAR RNA imino proton resonances for selected compounds.....	133
7. Representative imino proton NMR spectra of RNA in the presence of compound	137

Chapter 5

1. Diagram of identification of rigid anchor (layer 1) and flexible layers for growth	157
2. Influence of different methods of softening the repulsive force for the VDW term.....	160
3. Optimization of parameters for rigid ligand docking	163
4. Optimization of parameters for flexible ligand docking	167
5. Success rate and average length of calculation as a function of the number of rotatable bonds in the ligand for flexible ligands	176

Chapter 6

1. HIV-1 TAR RNA	187
2. General scheme of receptor conformation generation and selection	191
3. Hydrogen bonding distances of base pairs	194
4. Average hydrogen bond distances for base pairs over course of unrestrained REMD simulation	210
5. Example of optimization of number of clusters for the replica exchange simulation.....	213

Chapter 7

1. Scheme for RNA-biased fragment based screening method	222
2. Synthesis of scaffolds from fragment library	229

LIST OF TABLES

Chapter 1

1. Examples of Commonly Used Structure-based Drug Design Packages	31
2. Recent Examples of Novel Inhibitor Discovery Using Molecular Docking	33
3. Hit Rates and Drug-Like Properties for Inhibitors Discovered with High-Throughput and Virtual Screening Against the Enzyme PTP-1B	38

Chapter 2

1. Number of Active Compounds Identified in Each Group's Ranked List.....	54
---	----

Chapter 3

1. Summary of Scoring Functions and Sampling Algorithms for Commonly Used Docking Programs.....	64
2. Complexes Used in the Test Set.....	71
3. Zinc VDW Parameters Used to Generate Grids.....	71
4. Success Based on DOCK Version	91
5. Average Length of Time in Seconds for Docking Calculation Using the Optimized Parameter Set.....	91
6. Comparison of DOCK Success Rates to Other Docking Programs for Flexible Ligand Docking	91
7. Flexible Ligand Success as Function of Active Site Cofactor.....	95
8. Flexible Ligand Success as Function of CF Test Set Preparation.....	95

9. Comparison of Success and Failure Cases of Both Rigid and Flexible Docking for Complexes in Test Set with Cofactors in Active Site.....	95
10. Comparison of Success and Failure Cases of Both Rigid and Flexible Docking for Complexes in CF Test Set	95
11. Description of and Optimized Default Values for Parameters that Affect Rigid Ligand Docking	106
12. Description of and Optimized Default Values for Parameters that Affect Flexible Ligand Docking	112

Chapter 5

1. List of PDB Codes for All RNA-Ligand Complexes in Test Set	153
2. Description of and Optimized Default Values for Parameters that Affect Rigid Ligand Docking for RNA Test Set	163
3. Description of and Optimized Values for Parameters that Affect Flexible Ligand Docking for RNA Test Set	166
4. Average Score, Success Rate, and Length of Calculation for Various Subsets of Test Set Using Different Versions of DOCK.....	173

Chapter 6

1. Total Number of Conformations at Each Stage of the Culling Process	201
2a. Heavy Atom RMSD and Standard Deviation (Å) Between Structures from REMD Simulations and All Experimental Structures	203
2b. Active Site Heavy Atom RMSD and Standard Deviation (Å) Between Structures from REMD Simulations and All Experimental Structures.....	203

3a. Heavy Atom RMSD and Standard Deviation (Å) Between Structures from MD Simulations and All Experimental Structures	204
3b. Active Site Heavy Atom RMSD and Standard Deviation (Å) Between Structures from MD Simulations and All Experimental Structures.....	204
4a. Heavy Atom RMSD and Standard Deviation (Å) Between Structures from miniCarlo Simulations and All Experimental Structures	205
4b. Active Site Heavy Atom RMSD and Standard Deviation (Å) Between Structures from miniCarlo Simulations and All Experimental Structures	205
5a. Heavy Atom RMSD and Standard Deviation (Å) Between Structures from PEDC Simulations and All Experimental Structures.....	206
5b. Active Site Heavy Atom RMSD and Standard Deviation (Å) Between Structures from PEDC Simulations and All Experimental Structures	206
6. Average RMSD and Standard Deviation (Å) Between from All Experimental Structures.....	207

Chapter 7

1. Binding of Original Fragment Library to TAR RNA.....	227
2. Percent Yield from Synthesis Using Optimized Conditions	231

"Where the world ceases to be the scene of our personal hopes and wishes, where we face it as free beings admiring, asking and observing, there we enter the realm of Art and Science."

--Albert Einstein

INTRODUCTION

Humans have been looking for ways to understand and treat disease for centuries. In the past, doctors would use methods that we consider to be barbaric to alleviate the suffering of their patients. Now, we understand that most diseases have a specific molecular basis or are the result of an outside pathogen disrupting normal biological behavior. We have drug treatments that target the molecular basis of these diseases with exquisite specificity. However, the reality of medical science and basic biological research is that the amount of information we actually know about how the human body works in both the healthy and diseased state is still a tiny, tiny drop in the bucket of the vastness of biology. As an example, science is in the midst of revolution in the way it thinks about the most basic principle of molecular biology.

When I was in high school, I learned the Central Dogma of Molecular Biology—DNA begets RNA begets proteins—using a model that likened DNA to the original architectural blueprint, the RNA to a photocopy, and the proteins to the completed building (Downing PT, 1996). This original model, proposed by Crick in 1958 (without the photocopier idea obviously), implied two main conclusions (Crick FHC, 1958). First, the flow of information in cells goes monodirectionally from DNA to proteins and, second, that DNA and proteins are the main players, whereas RNA simply transports information from the nucleus to

the cytoplasm. Later, in college and then more in-depth in graduate school, I learned that both of these conclusions were not only over-simplified, but basically incorrect.

In 1975, the Nobel Prize in Medicine was awarded to Drs. David Baltimore, Renato Dulbecco and Howard Temin for determining that a protein, later identified and named reverse transcriptase, was capable of transforming viral RNA into DNA that could then be incorporated into the host cell genomes (Baltimore D, 1970; Temin HM and Mizutani S, 1970). In 1986, Nobel Prize winner Walter Gilbert proposed the "RNA World" hypothesis, which proposes a stage in evolution in which all biological processes of cells are governed by RNA without the need for DNA or proteins (Gilbert W, 1986). Between the years 2000 and 2001, when I was entering graduate school, the structure of the ribosome was solved to atomic resolution, in which it was found that the protein component of the machinery serves to stabilize the machinery structure, while the RNA component performs that actual peptide synthesis reaction through a mechanism that is still under debate (Ban N et al., 2000; Harms J et al., 2001; Schluenzen F et al., 2000; Wimberly BT et al., 2000; Yusupov MM et al., 2001).

These data, in conjunction with many others, contradict the models above and point toward a Central Dogma in which information flows predominantly from DNA through RNA to proteins, but where functionally important countercurrents also exist. In addition, these data strongly suggest that RNA is not simply a copy of DNA. Rather, RNA is a functionally critical part of many portions of the cellular life cycle. As more information about RNA and its involvement in cellular

function, emerge, more interest arises in the possibility of utilizing RNA as a drug target in diseases that have developed resistance to protein therapies, in particular the human immunodeficiency virus (HIV).

According to World Health Organization estimates, 38.6 million people are living with HIV/AIDS worldwide as of the end of 2005, with an estimated 2.8 million people newly infected each year. While the proportion of people living with HIV/AIDS appears to have stabilized in the 1990s, due both to effective drug treatment and social programs, the raw number of people living with HIV/AIDS continues to increase with the population and the ability of therapies to prolong life (UNAIDS/WHO, 2006). Historically, HIV has been notoriously difficult to treat. Drug resistant strains are rapidly generated due to the highly mutable character of its genome. This results in failure of drug therapy and eventual death of the patient. If we as a society hope to continue these trends, we need to continue to search for new treatments for the disease. One potential solution is to exploit new RNA targets that are essential to the life cycle of the virus.

The Tat-TAR complex has been identified as an attractive target for the inhibition of HIV (Hsu MC et al., 1991). In the first stages of HIV replication, the Tat protein facilitates viral transcription from the promoter region of the provirus incorporated into the DNA of the host cell. In order to form this interaction, Tat binds specifically to an RNA hairpin known as trans-activating response element (TAR) at the 5' end of the newly formed viral transcripts (Calnan BJ et al., 1991). Once transcription is complete, the viral proteins are produced in the cell, virus assembly and budding occur, and progeny virions are released to infect other

cells. The Tat-TAR complex has been found to enhance the overall rate of viral mRNA production by as much as 100-fold (Calnan BJ et al., 1991; Frankel AD and Young JA, 1998). It has been shown that disruption of this complex prevents elongation of the RNA genome by RNA polymerase, reducing viral replication (Karn J, 1999).

As the experimental interest in RNA as a drug target for HIV and other diseases has increased, so has the need for computational tools that can be used to model RNA in similar ways to proteins. Traditionally, as a class, RNA presents a difficult computational challenge as compared to proteins due to its electrostatic density and flexibility. In my naiveté as a young graduate student, I interpreted this to mean that using the long history of physics-based methods on a new system would be relatively simple and had not been done in the past simply from lack of practical interest. In line with this philosophy, the first several projects of my graduate career, and thus the first several chapters of my thesis, deal with the application and optimization of physics-based computational methods to more traditional systems.

Chapter 1, entitled "Molecular Docking and Structure-based Design," was written as a review of the current status of the field of computational structure-based drug design. The various components of molecular docking, including receptor site identification and characterization, orientation of ligands in the active site, and scoring of the ligand-receptor interaction, are broadly described. In addition, for each element of the process, several representative algorithms, along with their source references, are provided. The chapter will appear in Drug

Discovery Research: New Frontiers in the Post-Genomic Era (John Wiley and Sons), slated for publication in 2007.

Chapter 2, "Evaluating the High-Throughput Screening Computations," describes the results of the McMaster University High-Throughput Data-Mining and Docking Competition. The purpose of the competition was to perform a blind evaluation of the ability of computational algorithms to reproduce experimental results. In total, 32 separate groups participated using a wide variety of computational methods, of which four were identified as enriching the test set—a library of small molecules screened against dihydrofolate reductase—by 15% or better and three were identified as making useful comments about the nature of the test set. The chapter also included lessons learned about the nature of this type of blind evaluation from both the computational and experimental perspectives. The manuscript was published, along with the experimental data and several representative computational summaries, in the October 2005 issue of the Journal of Biomolecular Screening.

Finally, chapter 3, "Development and Validation of a Modular, Extensible Docking Program: DOCK 5," presents the summary of several graduate students' work. Version 5 of the DOCK suite of programs was written by Demetri Moustakas, with help from Scott Pegg, Scott Brozell and me. The program was optimized to predict binding poses for experimentally determined protein-ligand complexes using a test compiled by Natasja Brooijmans. In general, it was found that the sampling algorithm could recreate the correct binding pose with the optimized parameters. However, in many cases, the scoring function did not

rank the correct pose at the top of the list, suggesting that DOCK is sampling **adequately** but that the scoring needs to be improved. The manuscript was **accepted** and will be published in the Journal of Computer-Aided Molecular **Design**.

Armed with experience in both code development and experimental **validation** of protein-ligand interactions, I then refocused my efforts on RNA **targets**, particularly to interrupting the Tat-TAR interaction described above. **According** to structural studies, the binding interaction of the Tat-TAR complex is **thought** to be dominated by a single Tat arginine, which interacts with the bulge **region** of the TAR, and is stabilized by additional contacts between the protein **and** the phosphates on the RNA backbone (Aboul-ela F et al., 1995; Puglisi JD et **al.**, 1992). Because TAR is located at the beginning of each viral transcript, an **inhibitor** that targets the Tat-TAR interaction should prevent HIV transcription at **the** earliest stages.

In a series of studies by the James lab, the small molecule **acetylpromazine**, was shown to disrupt the Tat-TAR complex. In the first set of **experiments**, a subset of compounds from the Available Chemicals Directory, **filtered** for “drug-like” molecules, was screened against the TAR bulge region **with** DOCK using a modified scoring function. This resulted in several promising **scaffolds** that were then tested experimentally for binding. The acetylpromazine **was** shown to interfere with the binding of Tat to TAR *in vitro* (Lind KE et al., 2002). In a follow-up study, an NMR structure confirmed that acteylpromazine **interacted** specifically with the bulge region of TAR (Du Z et al., 2002). However,

when acetylpromazine was tested for specific interaction with TAR over other RNA molecules, it was found that acetylpromazine was a promiscuous RNA binder (Mayer M and James TL, 2004). It was therefore hypothesized that derivatives of the acetylpromazine scaffold would serve to both improve binding with the active site and to increase specificity for TAR over other RNA molecules.

In Chapter 4, "Synthesis and Testing of a Focused Phenothiazine Library for Binding to HIV-1 TAR RNA," Peter Madrid and Sabina Gerber synthesized a library of phenothiazine—the base scaffold of acetylpromazine—derivatives. This library was then screened against TAR using saturation transfer NMR, which monitors the shift of ligand peaks upon binding, by Moriz Mayer and Irene Gomez-Pinto. It was found that binding could be nontrivially enhanced by several-fold through modifications of the substituents. In addition, several areas of the molecule were identified that could be optimized in later studies. While binding was not enhanced enough for a useful specificity study, it was determined that some areas of the active site were amenable to larger substituents while others were not, suggesting that specificity based on the shape of the inhibitor could potentially be obtained. The study was published in the September 2006 issue of the journal of Chemistry & Biology.

In response to the acetylpromazine study and anticipating the need to develop small molecules for other RNA systems, there is a need for computational tools that will allow scaffold hopping were required. As stated above, RNA presents a difficult computational challenge because of its high charge density and its flexibility. In the remainder of my thesis, I present

methodologies that attempt to address these issues in the hopes of providing **tools** for RNA structure-based drug design.

Chapter 5, "Optimization of DOCK for RNA Targets" discusses **optimization** of the DOCK suite of programs for the ability of DOCK to reproduce **experimental** binding modes for a set of RNA-ligand complexes. As for the **Protein** test set in Chapter 3, the sampling parameters were optimized to **reproduce** experimentally determined binding poses of a test set of RNA-ligand **complexes**. In addition, several modifications were made to the flexible ligand **algorithm** in response to problems identified in the studies in Chapter 3. With the **optimized** parameters, DOCK is able to recreate the experimental binding poses **for** 60% of the ligands in the test set with less than seven rotatable bonds, 41% **of** the test set with less than ten rotatable bonds, and 37% of the test set with **less** than twelve rotatable bonds. In the next stage, we will explore the effect of **more** advanced methods for modeling solvation, various charge models, and **va**riation protonation states on the docking success rates. I anticipate completing **this** study and submitting a manuscript in the next few months.

Although the DOCK algorithm is fast, one of its limitations is that only the **ligand** is permitted to be flexible while the receptor remains rigid during the **docking** process. This limitation is applied for time constraints, but, for RNA **systems**, restraining the target to one conformation may significantly influence ranking a library of small molecules. In Chapter 6, entitled "Steps Toward Fully Flexible Docking to RNA Targets: Strengths and Weaknesses of Various Force Fields in Generating Ensembles of RNA Structures," several sampling

algorithms—replica exchange, single temperature molecular dynamics, **miniCarlo**, and Path Exploration with Distance Constraints—were explored for **their** ability to reproduce the experimentally determined ensemble of TAR RNA. **This** survey points out the strengths in the tested algorithms and also identified **areas** of potential improvement, particularly for the molecular dynamics **simulations**. At this point in the analysis, we have found that all the simulations **generate** ensembles that are on par with the diversity of experimental structures **using** RMSD as the criterion. However, more work needs to be done to **validate** **the** sampling techniques using experimental NMR data. Once we are confident **in** the generated ensembles, we will explore the effect of cross-docking to the **co**nformations using a library of known inhibitors and decoys using Saturation **T**ransfer Difference NMR. I expect to continue work on this project during the **be**ginning of my post doctoral research and to submit the manuscript early in **2007**.

The final chapter of my thesis, "A Fragment-Based Screening Method **Designed** for RNA Targets," presents preliminary results for a slightly new spin **on** the concept of fragment-based libraries, a technique that screens libraries of **structurally** minimal functional groups against targets and then connects the hits **into** larger scaffolds. In this case, the fragments were selected to have the **characteristics** of known RNA binders, including hydrogen bond donors and acceptors, aromatic rings, and electron driving groups. Each fragment was then **filtered** for reasonable NMR Saturation Transfer Difference spectra and **connected** to a chemically reactive group for easy linking into pieces. Preliminary

results indicated that several of the bits interacted with TAR RNA. At this point, **t**he chemistry to connect the fragments into scaffolds is being optimized and will **b**e screened by NMR. Once the screening is complete, both docking and **m**olecular dynamics simulations will be used to create structures of the bound **c**omplexes. Because the synthesis is being optimized, it is not clear when this **p**roject will be completed.

Taken as an entity, this thesis reflects the current status of the field of **d**eveloping drugs to target RNA based on the RNA structure. The majority of the **s**cience, and thus basic guiding knowledge, has been derived from protein **t**argets. In theory, if we completely understand and have correctly modeled the **p**hysical and chemical principles behind protein-ligand binding, then transferring **t**hese models to RNA should be trivial. However, the reality is that our models for **p**roteins are based on basic assumptions about the underlying character of **p**roteins that are simply not transferable to RNA. Rather, in my opinion, it has **b**ecome increasingly apparent that RNA is an entity unto itself with its own new **d**iscoveries, its own set of pitfalls. It is my hope that the pages of this thesis—**t**hese naïve fumbblings in the dark—will provide more specific insight into areas of **i**mprovement for the computational studies of RNA drug design.

Aboul-ela, F, Karn, J and Varani, G. *J Mol Biol.* **253** (1995) 313-32.

Baltimore, D. *Nature.* **226** (1970) 1209-11.

Ban, N, Nissen, P, Hansen, J, Moore, PB and Steitz, TA. *Science.* **289** (2000) 905-20.

- Calnan, BJ, Tidor, B, Biancalana, S, Hudson, D and Frankel, AD. *Science*. 252 (1991) 1167-71.**
- Crick, FHC. "On protein synthesis." Presented at *Symp Soc Exp Biol* (1958).**
- Downing, PT. "Notes from Junior Year Biology." Generated at *Mount Saint Joseph's Academy* (1996).**
- Du, Z, Lind, KE and James, TL. *Chem Biol*. 9 (2002) 707-12.**
- Frankel, AD and Young, JA. *Annu Rev Biochem*. 67 (1998) 1-25.**
- Gilbert, W. *Nature*. 319 (1986) 618-618.**
- Harms, J, Schluenzen, F, Zarivach, R, Bashan, A, Gat, S, Agmon, I, Bartels, H, Franceschi, F and Yonath, A. *Cell*. 107 (2001) 679-88.**
- Hsu, MC, Schutt, AD, Holly, M, Slice, LW, Sherman, MI, Richman, DD, Potash, MJ and Volsky, DJ. *Science*. 254 (1991) 1799-802.**
- Karn, J. *J Mol Biol*. 293 (1999) 235-54.**
- Lind, KE, Du, Z, Fujinaga, K, Peterlin, BM and James, TL. *Chem Biol*. 9 (2002) 185-93.**
- Mayer, M and James, TL. *J Am Chem Soc*. 126 (2004) 4453-60.**
- Puglisi, JD, Tan, R, Calnan, BJ, Frankel, AD and Williamson, JR. *Science*. 257 (1992) 76-80.**
- Schluenzen, F, Tocilj, A, Zarivach, R, Harms, J, Gluehmann, M, Janell, D, Bashan, A, Bartels, H, Agmon, I, Franceschi, F and Yonath, A. *Cell*. 102 (2000) 615-23.**
- Temin, HM and Mizutani, S. *Nature*. 226 (1970) 1211-3.**
- UNAIDS/WHO. "2006 Report on the global AIDS epidemic."
<http://www.unaids.org/en/HIV%5Fdata/2006GlobalReport/>. (Sept 2006).**
- Wimberly, BT, Brodersen, DE, Clemons, WM, Jr., Morgan-Warren, RJ, Carter, AP, Vornrhein, C, Hartsch, T and Ramakrishnan, V. *Nature*. 407 (2000) 327-39.**
- Yusupov, MM, Yusupova, GZ, Baucom, A, Lieberman, K, Earnest, TN, Cate, JH and Noller, HF. *Science*. 292 (2001) 883-96.**

"I'm still not convinced we will ever get docking to work well."

--Tack Kuntz, Retirement Symposium

Chapter 1

Molecular Docking and Structure-Based Design

P. Therese Lang¹, Tiba Aynечи², Demetri Moustakas³, Brian Shoichet², Irwin D. Kuntz^{2*}, Natasja Brooijmans¹, Connie M. Oshiro⁴

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

²Department of Pharmaceutical Chemistry
University of California, San Francisco

³Joint Graduate Program in Bioengineering
University of California, Berkeley/University of California, San Francisco

⁴Roche Pharmaceuticals, Palo Alto, CA

**To whom all correspondence should be addressed (kuntz@cgl.ucsf.edu)*

ABSTRACT

The discovery of new drugs is a complex process. Computational methods have proved useful in many aspects. This chapter focuses on “structure-based drug design” describing strategies which use the receptor structure to identify or design ligands. These protocols include virtual screening, compound optimization, and fragment-based design. We describe in a particular aspect of structure-based design which makes use of molecular docking. We present the underlying assumptions in this approach and summarize recent tests that document pragmatic successes of the methodology. Finally, we point to the need for better treatment of entropic terms and conformational sampling, especially sampling receptor flexibility.

Keywords: Structure-based; Drug design; Ligand design; Molecular docking; Drug discovery; Molecular mechanics; Virtual screening

Abbreviations: DHFR, dihydrofolate reductase; GB, Generalized Born; K_i , experimental binding affinity; MCSS, Multiple Copy Simultaneous Search; MD, molecular dynamics; PB, Poisson-Boltzmann; QSAR, quantitative structure-activity relationships; SA, surface area; vdW, van der Waals.

INTRODUCTION

The discovery of new drugs is a complex process. It generally starts with the identification of compounds that bind to a target or show efficacy in a simple screen. Molecules that show good affinity are called “hits.” The next step is to find compounds that have attractive pharmaceutical properties, for example, low toxicity and sufficient aqueous solubility to be orally active. Such compounds are often called “leads.” Traditionally, “hits” have been found by screening, while “leads” are developed from “hits” through chemical synthesis. Screening normally involves large numbers of compounds from natural products, corporate databases, or organic chemistry companies that can be examined for biological activity in high-throughput assays. Commercial systems can process a million tests per day for enzyme targets. The best compounds are moved forward in a process aimed at modifying their chemical structure to improve potency, specificity, and in vivo activity while lowering toxicity and side effects. Synthetic methods include combinatorial chemistry and library synthesis (Figure 1).

Computational methods have proved useful in many aspects of the discovery process (Alvarez JC, 2004). A variety of strategies are available. If an active “lead” is known, it is straightforward to query a database for molecules with similar properties using pharmacophore searches or quantitative structure-activity relationships (QSAR). These methods typically use information available for the ligand, inhibitor, or substrate rather than the receptor. However, there is growing interest in mapping out receptor properties—either from known family relationships with other members of the receptor family or through

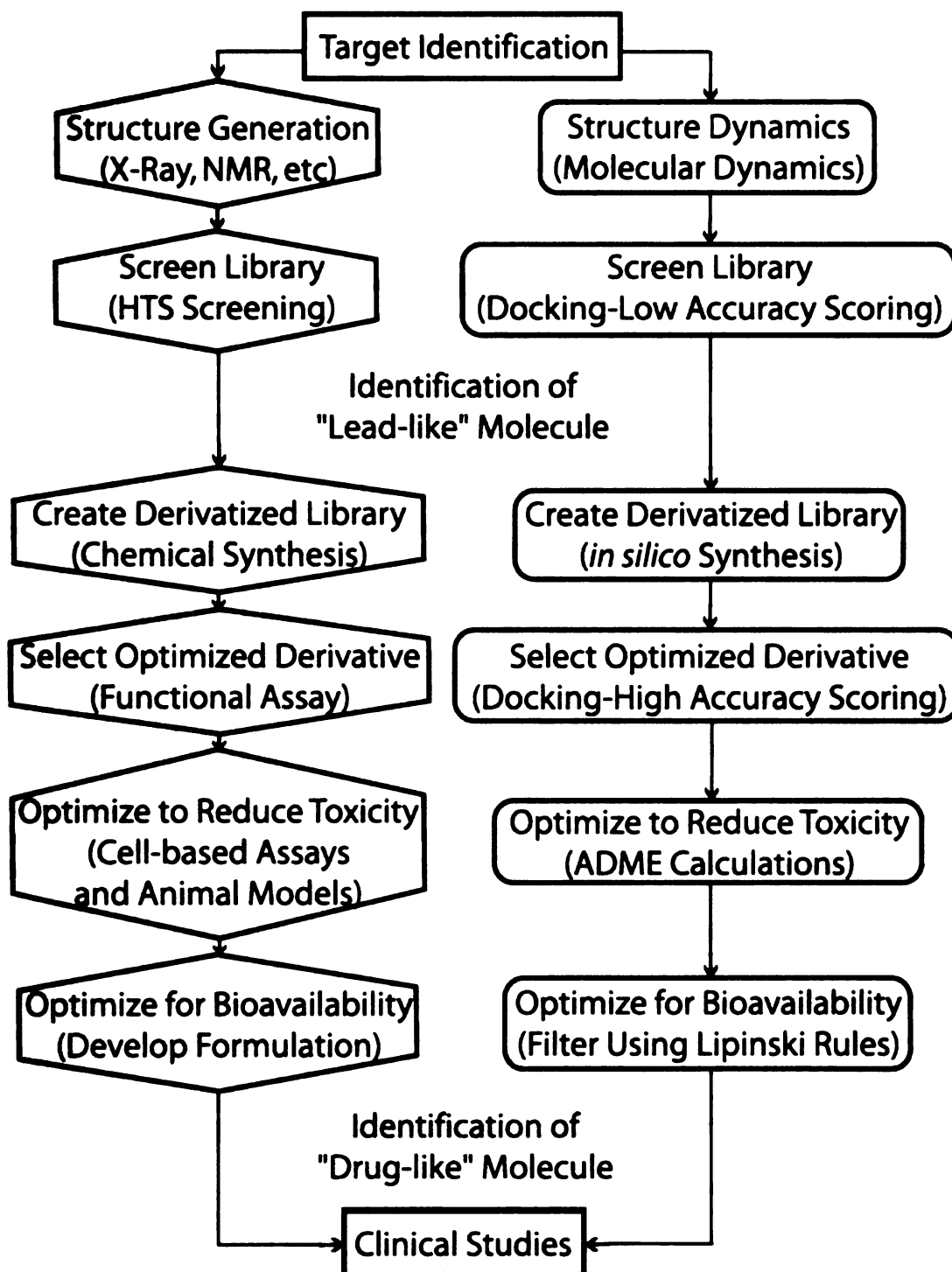


Figure 1: Example flow of drug design process from both experimental and computational perspectives

pharmacophore strategies applied directly to the receptor structures (Arnold JR et al., 2004; Hajduk PJ et al., 2005). This chapter focuses on a set of strategies in which direct knowledge of the receptor structure is used to identify or design ligands that possess good steric and chemical complementarity to specific sites on the target macromolecule. This process is referred to as “structure-based drug design” (Brooijmans N and Kuntz ID, 2003).

The structure-based drug design paradigm is analogous to experimental screening. Structures for the receptor or target are obtained either from the literature or in-house operations. These structures come from crystallography or NMR experiments, but there is increasing interest in high-quality homology-modeled structures (Chance MR et al., 2004). Computer analogs of ligands are generated. These families are often called “virtual libraries” and may consist of compounds from corporate, academic, or commercial holdings (Laird ER and Blake JF, 2004; Webb TR, 2005). A virtual library might also include molecules that are not physically available but might be obtained through chemical synthesis, perhaps using combinatorial chemistry (Jorgensen WL, 2004; Kick EK et al., 1997). Screening of the virtual library against the target structure involves some form of positioning the putative ligand in three-dimensional space and evaluating the intermolecular interactions for that particular geometry. Typically, the process is an iterative one: the ligand is moved, and the new geometry is evaluated. This cycle is repeated until some “best scoring” geometry is identified for the particular ligand under test. Then, the next ligand in the list is chosen, and the whole procedure begins again. The goal of virtual screening is to identify the

best binding candidates from the library for experimental testing. Because the virtual libraries can be huge—upwards of a billion compounds—this triage procedure is a critical step.

Once a binding candidate has been found, structure-based design can be used to optimize binding affinity. For this operation, one starts with a “hit” or “lead” with a known activity. Often, a structure for the ligand-target complex is available. There are many computational methods available for evaluating chemical variants of the “lead” that offer suggestions about the direction for the next round of synthesis. Ideally, a number of such variants are prepared, and their properties and structures obtained so that a selection of molecules are available to take into further biological testing. Optimization methods are typically much more computationally intensive than other virtual screening approaches.

There has recently been a merging of these two ideas. Virtual libraries containing 1,000–10,000 molecular fragments (sometimes called “anchors” or “scaffolds”) are used in the initial screening. The most promising are then expanded using computer synthesis in a combinatorial fashion (Jorgensen WL, 2004; Kick EK et al., 1997; Miranker A and Karplus M, 1991).

In this chapter, we will focus on a particular subset of molecular design strategies called “docking,” in which candidate molecules are matched to receptor structures and evaluated for chemical and geometric complementarity. We will not discuss the broad field of quantitative structure-activity relationships (QSAR) that focus on the chemical structures of ligands alone (Bender A and Glen RC, 2005).

MOLECULAR DOCKING

Overview

The basis of molecular docking is the calculation or estimation of the free energy of binding of a ligand to a specific receptor site in a fixed environment. The free energy of binding yields, directly, an equilibrium binding constant and, indirectly, the preferred binding mode of the ligand-receptor complex. There are important scientific and mathematical issues involved. For example, it is currently much easier to calculate the energy/enthalpy of interaction than to obtain the free energy because we lack efficient ways to obtain the entropic contributions. Second, the interactions of the ligand and receptor with the solvent (water, salts, and other components) are not easy to quantitate. Third, searching through the large number of conformations of the receptor and the ligand and their relative positions are difficult computer science problems. The need to repeat these calculations for large numbers of putative ligands and many possible targets requires serious attention to the algorithms. Docking protocols have adopted a variety of heuristics to make useful calculations with the knowledge that a complete high-level calculation is not feasible for the systems of biological or therapeutic interest.

What can we expect from current approaches? The best-case calculations are accurate to within approximately 0.5 kcal/mol of experimental results, but these are generally free energy differences obtained using perturbation techniques on a related family of ligands (Jorgensen WL, 2004). Routine results

are rarely within 1 kcal/mol of experimental results, and library searches of diverse chemical types have larger inaccuracies. Work continues on improving the force fields that model the enthalpic terms (Bernacki K et al., 2005).

Estimates of entropic contributions are empirical, and adequate sampling of configurations and conformations search is a complex combinatorial problem.

Consequently, searching large databases for new leads requires protocols that deal with three specific tasks:

- (i) receptor site identification;
- (ii) receptor site characterization;
- (iii) orientation of the ligand within the site; and
- (iv) evaluation of the ligand.

These steps are described in turn in the succeeding sections, and examples are given.

Receptor site identification

With the sequencing of the human genome and recent advances in structural techniques, the number of publicly available biomolecular structures has exploded over the last few years with over 30,000 in the Protein Data Bank (Berman HM et al., 2000). The two main experimental sources for 3D structures of biomolecules are X-ray crystallography and high-resolution NMR spectroscopy. X-ray crystallography provides structures of biomolecules in the crystalline states, and with the crystallization of the ribosome, the upper limit of the experiment has been extended to the 1,000 kDa range. NMR spectroscopy

is limited to approximately 50 kDa, but the method has the advantage of providing additional information about the dynamics of the structure. As an alternative, if the structure has not been solved experimentally, computational techniques such as homology modeling can be used to predict 3D structures. We next discuss some pros and cons of each of these sources for structure-based drug design.

For X-ray crystal structures to be sufficiently accurate for drug designing purposes, a resolution of approximately 2 Å, an R-factor below 20%, and an R_{free} factor below 30% are preferred. It is important to note that the majority of 3D crystal structures of biomolecules do not have hydrogens or highly flexible residues included in the file. The missing atoms must be considered before structure-based drug design can begin. In addition, crystal packing forces may locally influence protein conformation, particularly for nucleic acids and surface active sites.

The result of structure determination with NMR is an ensemble of structures that agree equally well with experimental data. Although an averaged structure can be derived, it has been shown that the entire ensemble provides a more complete description of the system from an experimental perspective (Staunton D et al., 2003). For structure-based drug design purposes, though, there are several methods for choosing the appropriate structure, including selecting the member of the ensemble closest to the average as measured by some distance metric or cross-docking to all members of the ensemble. Unfortunately, there is no generally accepted standard of accuracy for NMR-

generated structures. As a rough rule of thumb, a high-resolution NMR structure should preferably have approximately 20 (distance or dihedral) restraints per residue (Berman HM and Westbrook JD, 2004).

If no experimental structural information is available for the target biomolecule, homology modeling can provide structures to guide the search for novel lead compounds. It should be noted that, depending on the method, homology modeling yields average errors of 3 Å root mean square deviation of proteins with greater than 50% sequence similarity, with larger errors for increasing sequence dissimilarity (Nissen P et al., 2000). Nevertheless, homology modeling has proven to be successful in several cases, including discoveries of a highly potent DNA methylation inhibitor and a compound that discriminates between two voltage-gated K⁺ channels with 20-fold accuracy (Staunton D et al., 2003).

A number of more general issues associated with the selection and preparation of a receptor structure should be noted. In many structures, ions are required for structural or functional purposes. However, modeling this type of chemistry is often difficult because the formal charge and associated desolvation energy of ions are extremely complicated to compute accurately. The protonation states of residues such as histidine, lysine, glutamic acid, or aspartic acid are highly dependent on the environment in the active site and may even change in response to the ligand (Hensen C et al., 2004). In particular for crystal structures, critical water molecules may be present in the active site, and it is often difficult to predict whether the water can be replaced or should be included

in the model of the receptor. All these issues can affect the quality of the model and should be carefully considered.

Receptor site characterization

Once an accurate structure has been determined for the target, ligands are typically restricted to lie within one geometric region of the macromolecule, generally known as the “binding site.” This region is generally selected because, upon ligand binding, normal function is altered. A given receptor can have one or more binding sites, such as the active sites of enzymes, allosteric sites, the binding or recognition sites of receptors, or even a dimer interface. The exact location of the binding site may be well known through experiment. However, if the binding site is not known, automated methods exist to identify potentially interesting regions.

Experimental data that indicate the binding site is the best source, if available. Experimentally derived structures of the biomolecule complexed with the natural substrate or a known inhibitor directly indicate the binding site. The Protein Data Bank and the Nucleic Acid Data Bank contain a large number of these types of structures (Berman HM et al., 1992; Berman HM et al., 2000). The Cambridge Crystallographic Data Centre/Astex has compiled a subset of biologically relevant protein-ligand complexes identified as being reliable for structure-based drug design purposes (Nissink JW et al., 2002). If direct observations are not available, binding regions can be identified through mutational experiments, such as alanine screening, or other biochemical assays.

There are cases where binding site information is not available. In these instances, computational tools can be used to indicate probable binding areas. We will describe two methods: the first is based upon geometric features, and the second uses chemical functionality of the receptor surface. For illustration, we consider SPHGEN from the DOCK suite of programs and the Multiple Copy Simultaneous Search (MCSS) approach (Ewing TJ et al., 2001; Kuntz ID, 1982; Meng EC et al., 1992; Miranker A and Karplus M, 1991). In addition, an interesting statistical characterization has been recently described (Hajduk PJ et al., 2005).

SPHGEN automatically identifies a target site by computing a set of site points or sphere centers, which serves to create a negative image of the surface. The algorithm begins by mapping the geometric features of the receptor surface, as defined by Lee and Richards, using the dms program (Richards FM, 1977). Then, spheres of varying radii are analytically generated to touch the molecular surface at two points, with the sphere center lying along the surface normal and with no portion of the sphere intersecting with a receptor atom. These overlapped spheres indicate various surface features, including invaginations and clefts. A clustering protocol, using radial overlap as a metric, is then used to indicate potential areas for ligand binding. The largest cluster is generally used as the binding site, and, once generated, the cluster is used as a template for possible ligand atom positions.

MCSS identifies binding sites by mapping the chemical properties of the biomolecular surface. Thousands of copies of specific molecular fragments are

distributed in the target region of the protein. Then, energy minimization is performed on the ensemble, creating distinct local minima for each fragment. This process is repeated on a variety of chemical functionalities until the surface is adequately described. Although this method does not capture every geometric detail of the binding site, it does provide a basic pharmacophore that can be used in later studies (Arnold JR et al., 2004; Miranker A and Karplus M, 1991).

Orientation of the ligand in the target site

There are two basic strategies for exploring the orientational degrees of freedom for putative ligands. The first uses a search grid for both the translational space and the euler angle space. This brute-force method is feasible if one is studying a few ligands in a restricted site. It can be extended to larger libraries if parallel processing is available (Jorgensen WL, 2004). Alternatively, protocols have been developed to prescreen orientation space. For example, the DOCK program uses a geometric pairing algorithm, matching the sphere centers (described above) with ligand centers (usually ligand atoms). The match criterion is based on a comparison of inter-sphere and inter-atom distances. Exhaustive or selective searches can be done over the match matrix. Careful placement of the spheres is an important step in getting good-quality results. The second type of method selectively samples all of orientation and conformation space using search engines. For example, the Metropolis algorithm and simulated annealing are used in QXP, and the genetic algorithm has been implemented an Autodock (McMartin C and Bohacek RS, 1997; Morris GM et al., 1998). These algorithms

have been studied extensively for many applications. Their strengths and limitations are well understood.

Evaluation of ligand orientations

The many configurations (orientations and conformations) of the ligand need to be evaluated with a scoring function to identify the energetically most favorable ligand binding pose. Ideally, the scoring function would calculate the ligand free energy of binding in aqueous solution (Beveridge DL and Dicapua FM, 1989; Kollman P, 1993). However, the large computational expense of these calculations leads to the introduction of scoring functions that calculate a range of simplifications of the ligand binding free energy.

Scoring functions can be broadly classified into two categories: those based on first-principles derived molecular mechanics force fields, and those based on functions fit to empirically derived binding data. For the purposes of this review, scoring functions that employ quantum mechanics are not considered as the extreme computational cost of these calculations make them prohibitive for use during small molecule docking.

Of the first-principles derived scoring functions, the most computationally efficient are those that approximate ΔG_{bind} as the molecular mechanics protein-ligand interaction energy. Molecular mechanics treats the molecule as a collection of atoms governed by a set of classical mechanical potential functions (Weiner PK and Kollman PA, 1981). Parameters for these potentials are derived from small molecule experiments and refined to yield correct structural and

thermodynamic quantities such as bond stretching frequencies or heats of formation. The primary DOCK energy scoring function approximates the ligand-receptor binding energy using the AMBER molecular mechanics intermolecular interaction energy, a sum of the Lennard-Jones 6-12 van der Waals (vdW) potential, and the Coulombic potential, given in equation 1:

$$E = \sum_i^{Lig} \sum_j^{Rec} \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} + K \frac{q_i q_j}{D r_{ij}} \right) \quad (1)$$

where i indexes the ligand atoms and j indexes the receptor atoms; A and B are the vdW attractive and dispersive parameters, respectively; q is the partial charge on the atom; K is the scaling constant that converts electrostatic energy into kcal/mol; D is the dielectric constant of the medium; and r_{ij} is the distance between ligand atom i and receptor atom j (Pearlman DA et al., 1995). This scoring function is limited by its use of a distance-dependent dielectric screening function to mediate all charge-charge interactions. This dielectric treatment assumes that the dielectric value of the solvent is uniform between all charge pairs.

A class of scoring functions has been developed that utilizes implicit models of solvation to calculate the electrostatic component of the molecular mechanics intermolecular interaction energy in a more sophisticated fashion than the simple Coulombic approach previously described. Both the Generalized Born (GB) and Poisson-Boltzmann (PB) terms have been combined with an empirically derived surface area (SA) term to include the energy of desolvating non-polar atoms, and the resulting GB/SA and PB/SA methods have been

implemented into molecular dynamics and docking methods (Feig M and Brooks CL, 2004; Feig M et al., 2004; Honig B and Nicholls A, 1995). While these functions are more computationally intensive than the Coulombic electrostatic energy functions, their proper treatment of solvation effects yields more accurate energy scores, and they are therefore frequently used in a hierarchical fashion to rescore docked ligand poses. DOCK 5 implements a GB/SA scoring function that is recommended for use in a rescoring capacity (Zou XQ et al., 1999).

The most computationally intensive class of first-principles derived scoring functions combine molecular dynamics (Lingham RB et al.) simulations with implicit or explicit solvation to average the interaction energies from a Boltzmann-weighted ensemble of complex structures, yielding accurate estimates of the binding free energy that takes into account protein flexibility. The implicit solvation methods, MM-PB/SA and MM-GB/SA, perform a short, explicit solvent MD simulation from which a set of snapshots of the protein-ligand complex structure are saved (Gohlke H and Case DA, 2004; Wang JM et al., 2001; Zhou ZG and Madura JD, 2004). These snapshots, representing a Boltzmann-weighted ensemble of complex structures, are rescored with either the PB/SA or GB/SA scoring functions, and the average interaction score of the snapshots is taken as the free energy of binding for the ligand. Several research groups in academia and industry use these methods in a hierarchical fashion to rescore ligands identified as potential binders (Kollman PA et al., 2000).

The second major class of scoring functions models the binding free energy as a weighted sum of several different types of interaction energies, with

and without explicit vdW and electrostatic terms. Many of these functions are based upon a comparison of receptor-ligand complexes and experimental binding data (K_i). Programs such as AutoDock, FlexX, GOLD, and Glide implement a variety of empirically derived energy score functions, including the well-known ChemScore and PLP functions (Eldridge MD et al., 1997; Friesner RA et al., 2004; Gehlhaar DK et al., 1995; Halgren TA et al., 2004; Hoffmann D et al., 1999; Morris GM et al., 1998; Verdonk ML et al., 2003).

LIGAND STRUCTURE GENERATION

To identify virtual screening “hits,” databases consisting of three-dimensional structures of putative ligands are searched using the methodology described above. How are these structures obtained? In some cases, structures can be taken directly from a database of experimentally-determined structures, e.g., the Cambridge Crystallographic Database contains 325,709 small molecule crystallographically determined structures as of November 2004. In cases where only the two-dimensional information is known, one can create computer representations of the covalent structure using conformation generation programs such as Concord (Tripos), ChemX (Accelrys), Rubicon (Daylight), and Omega (OpenEye). These programs provide one or more conformers consistent with the chemical connectivity and general rules of physical organic chemistry. Often, these conformers will be ranked by some energy formula. Finally, some vendors make libraries of three-dimensional structures directly available, including the Advanced Chemical Development (www.acdlabs.com), MDL Drug

Data Report, National Cancer Institute Open Database Compounds (cactus.nci.nih.gov), Tripos Discovery Research Screening Libraries (www.tripos.com), InfoChem GmbH database (www.infochem.de), Thomson Index Chemicus database (scientific.thomson.com), and ZINC (blaster.docking.org/zinc).

Two strategies have emerged to study flexible ligands. The first, generally referred to as incremental construction, breaks down the ligand into smaller pieces and then rebuilds it during the docking calculation. One example of this technique starts with a fragment of the compound (anchor) and then adds atoms in layers during a docking or an optimization cycle. This approach has been called "anchor and grow." In the alternate method, conformers for each compound can be pregenerated, stored in a database and then rigidly docked. Molecular dynamics and Monte Carlo techniques offer a combination strategy where the starting point is a single conformation of the ligand that then explores alternatives during the dynamics phase. There have been some tests of the two strategies, but there is no strong consensus of which is better (Lorber DM and Shoichet BK, 2005; Moustakas D et al., in press).

Other important issues that influence ligand structure and docking are choices of partial charges, tautomer preferences, and pKa values. A new database (ZINC) that deals with many of these concerns is now available through the Shoichet group at UCSF (Irwin JJ and Shoichet BK, 2005).

DESCRIPTION OF DOCKING PROGRAMS

While we do not have the space to provide descriptions of the many different approaches to molecular docking, Table 1 summarizes features of several of the most frequently used programs (Anderson AC, 2003; Murcko MA, 1997). We also present a brief synopsis of DOCK 5, developed at UCSF.

DOCK 5, the current version of the DOCK program, is written in C++ and provides an object-oriented implementation in which each major component of the DOCK algorithm is a class with a documented interface, allowing these DOCK functions to be modified or replaced easily. As a result, it has been possible to independently validate and optimize the rigid body sampling, the flexible sampling, the energy scoring functions, and our minimizers. DOCK 5 features an energy scoring function based on a molecular mechanics force field, solvation corrections using implicit solvent models, integration with the complete AMBER force field score, rigid body docking, ligand conformational searching, binding pose cluster analysis, and local minimization methods and also includes support for parallel computing using the MPI standard.

TESTS OF DOCKING AND STRUCTURE-BASED DESIGN

Despite well-known methodological weaknesses, structure-based screening using molecular docking has had important successes. Pragmatically, docking has suggested new, non-obvious ligands for multiple targets; these have been subsequently tested and shown to bind experimentally. Hugo Kubinyi, in a

Method	Ligand Sampling Method ^a	Receptor Sampling Method ^{a,b}	Scoring Function ^{b,c}	Solvation Scoring ^{b,d}	Reference(s)
AutoDock	GA	SE	MM + ED	DDD, DS	(Morris GM et al., 1998; Osterberg F et al., 2002)
DOCK 3	CE	SE	MM	PBE, DS	(Lorber DM and Shoichet BK, 1998; Wei BQ et al., 2004)
DOCK 4/5	IC	SE	MM	DDD, GB, PB	(Ewing TJ et al., 2001; Moustakas D et al., in press; Zou XQ et al., 1999)
EUDOCK	CE	CE	MM	DDD	(Pang YP et al., 2001a)
FlexX/FlexE	IC	SE	ED	NA	(Claussen H et al., 2001)
Glide	CE + MC	TS	MM + ED	DS	(Eldridge MD et al., 1997; Friesner RA et al., 2004; Halgren TA et al., 2004; Sherman W et al., 2006)
GOLD	GA	NA	MM + ED	NA	(Jones G et al., 1997; Verdonk ML et al., 2003)
CM-Dock	MC	MC	MM + ED	DDD, PBE, DS	(Abagyan R et al., 1994; Totrov M and Abagyan R, 1997)
MM-PBSA	MD	MD	MM	GB, PB	(Kollman PA et al., 2000)
QXP	TS + MC	MD	MM + ED	DDD	(McMartin C and Bohacek RS, 1997)

Table 1. Examples of Commonly Used Structure-based Drug Design Packages

- a)** Sampling methods are defined as Genetic Algorithm (GA), Conformational Expansion (CE), Monte Carlo (MC), Molecular Dynamics (Lingham RB et al.), Incremental Construction (IC), Merged Target Structure Ensemble (SE), Torsional Search (TS)—see *Ligand Structure Generation* for more information.
- b)** If the package does not accommodate this option, the symbol NA (Not Available) is used.
- c)** Scoring functions are defined as either empirically derived (ED) or based on molecule mechanics (MM)—see *Evaluation of Ligand Orientation* for more information.
- d)** Additional accuracy can be added to the scoring function using implicit solvent models. The most commonly used options are Distance Dependent Dielectric (DDD), Poisson Boltzmann Dielectric (PBE), a parameterized desolvation term (DS), Generalized Born (GB) and linearized Poisson Boltzmann (PB)—see *Evaluation of Ligand Orientation* for more information.

recent review, describes over 50 macromolecular targets for which ligands have been discovered using docking-based approaches (see Table 2 for a partial list) (Kubinyi H, in press). Most of these projects used experimental X-ray structures to represent the protein. In several cases, homology-modeled structures were employed (Evers A and Klebe G, 2004a; Schapira M et al., 2003).

In recent work, the structures of known ligands in complex with their receptors have been predicted by docking, beginning with the structures of the independent molecules (Figure 2) (Rizzo RC et al., 2000; Rosenfeld RJ et al., 2003). In these studies, where the binding affinity is known but the structure of the complex is not, the docking predictions have been relatively accurate. The caveat to this is that there are many cases where docking mis-predicts geometries in retrospective tests. Still, in published cases where the goal was genuine prediction, the docked geometry has often turned out to correspond closely to the subsequent experimental result.

A more difficult test is comparing the predicted geometries of novel ligands that emerge from the docking screens themselves. There are many examples of such predictions of ligand and geometry from docking screens against simple model cavity sites. These sites are small, completely enclosed by the protein, and dominated by one particular type of interaction, such as hydrophobicity, a single hydrogen bond acceptor or a single electrostatic interaction. These features have allowed for multiple predictions of new ligands that are tested experimentally, often including structure determination (Figure 3)

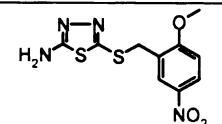
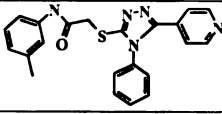
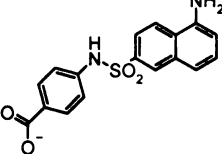
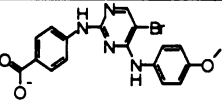
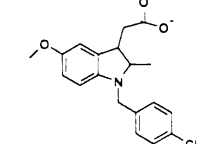
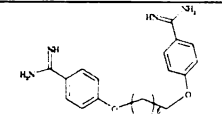
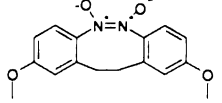
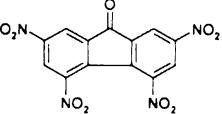
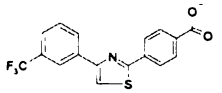
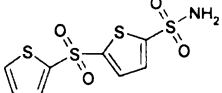
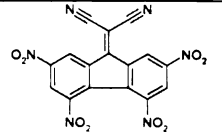
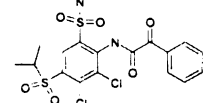
Target	Representative Hit	Lead Inhibitor IC ₅₀	Follow-Up Inhibitor IC ₅₀	Complex Structure
p56 Lck SH2 domain (Huang N et al., 2004)		10 μM	NR	No
neurokinin-1 receptor (Evers A and Klebe G, 2004b)		0.25 μM	NR	No
AICAR Transformylase (Li C et al., 2004)		0.15 μM	NR	No
Checkpoint Kinase 1 (Lyne PD et al., 2004)		0.11 μM	NR	No
aldose reductase (Iwata Y et al., 2001)		4.3 μM	0 μM	No
matriptase (Eenyedy IJ et al., 2001)		0.92 μM	0 μM	No
Bcl-2 (Eenyedy IJ et al., 2001)		10.4 μM	NR	No
adenovirus protease (Pang YP et al., 2001b)		3.1 μM	NR	No
retinoic acid receptor (Schapira M et al., 2001)		2 μM	NR	No
Carbonic anhydrase II (Gruneberg S et al., 2001)		0.0008 μM	NR	Yes
HPRTase (Freymann DM et al., 2000)		2.2 μM	NR	No
hydro-dipicolinate (Iwata AM et al., 2001)		7.2 μM	NR	No

Table 2. Recent Examples of Novel Inhibitor Discovery Using Molecular Docking

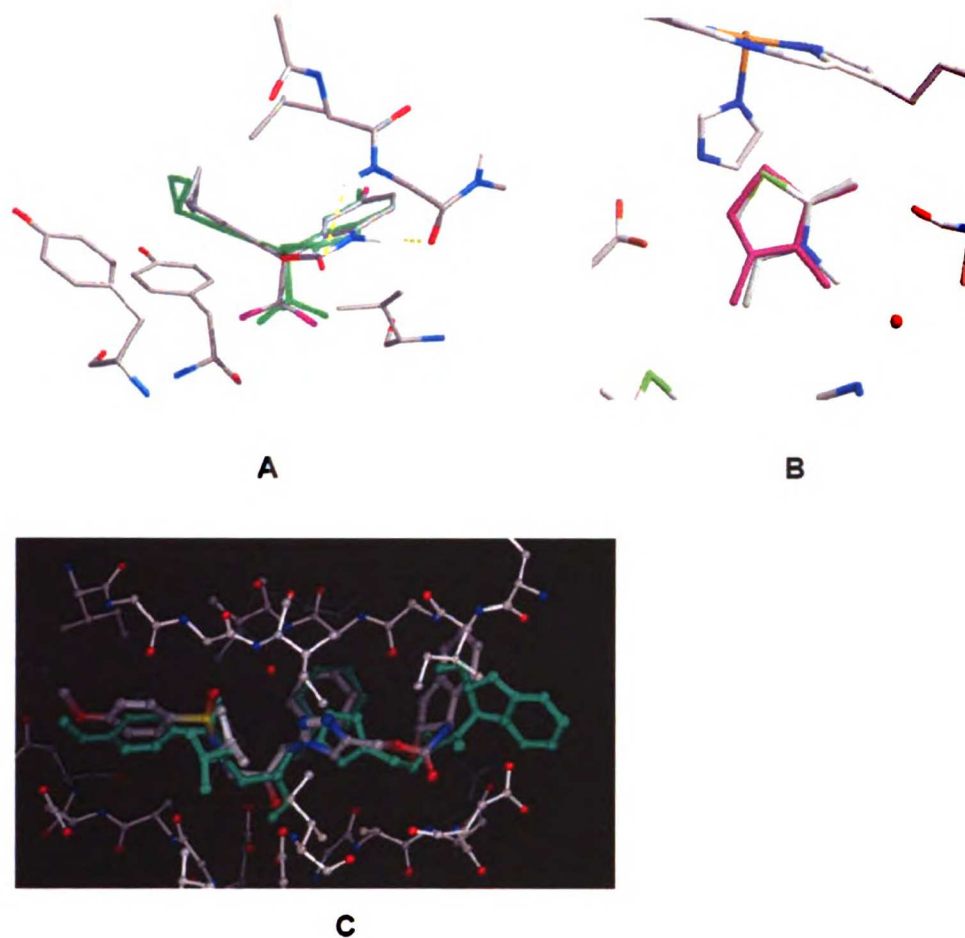


Figure 2: Predicted complexes versus X-ray crystallographic structures that were subsequently determined

A) Predicted (carbons in gray) and experimental (green) structures for sustiva in HIV reverse transcriptase (Rizzo RC et al., 2000; Rosenfeld RJ et al., 2003)

B) Predicted (magenta) and experimental (white carbon atoms) structures of 2,3,4-trimethylthizole in the W191G cavity of cytochrome c peroxidase. (Rizzo RC et al., 2000; Rosenfeld RJ et al., 2003)

C) Predicted (green) and experimental structure (carbons in gray) of an HIV protease inhibitor (ligands with thick bonds, enzyme residues with thin bonds) (Brik A et al., 2005; Brik A et al., 2003); courtesy of Art Olson, TSRI

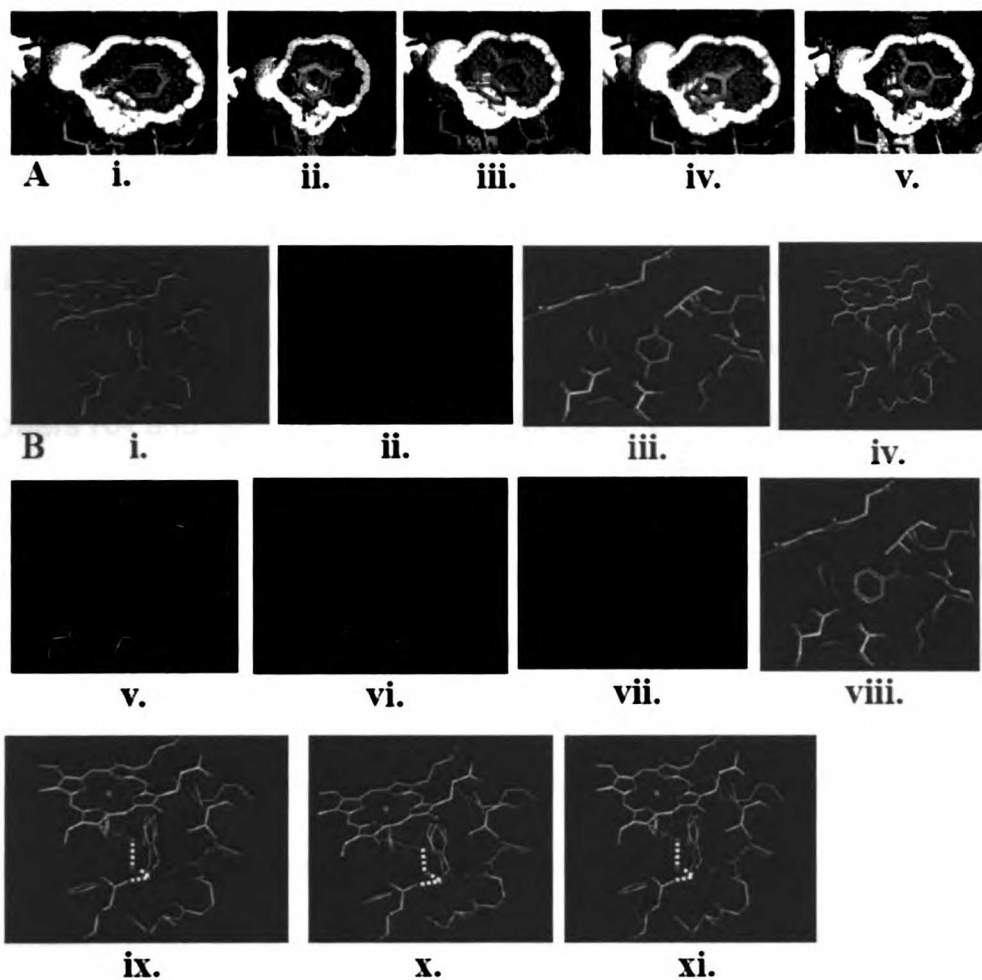


Figure 3: Docking predicted ligands from virtual screening against simple cavity sites.

A. The docked prediction (carbons in green) superposed on the crystallographic result (carbons in cyan). The surface of the L99A/M102Q cavity of T4 lysozyme (yellow) is cut away to reveal the complex. i. phenol; ii. chlorophenol; iii. fluoroaniline; iv. methylpyrrole; v. difluorophenol

B. The docked prediction (carbons in green) superposed on the crystallographic result (carbons in yellow) in the W191G cavity of cytochrome c peroxidase. i. thiophene-amidine; ii. diaminopyridine; iii. 2-amino-5-methylpyridine; iv. 2-amino-4-methylpyridine; v. diaminopyrimidine; vi. hydroxymethyl-imidazole; vii. 3-methyl-n-methylpyridine; viii. 4-hydroxymethyl-pyridine; ix. aminomethyl-cyclopentane; x. aminomethyl-benzene; xi. aminomethyl-furan

(Brenk R et al., 2006; Graves AP et al., 2005; Wei BQ et al., 2002; Wei BQ et al., 2004). Thus, X-ray crystal structures have been determined for about 25 ligands bound to three different cavities; in every case, the docking prediction corresponds closely to the X-ray crystallographic result. These results suggest that current docking algorithms are adequate to capture first-order determinants of binding fidelity (Figure 4) (Gradler U et al., 2001; Gruneberg S et al., 2002; Powers RA and Shoichet BK, 2002; Wei BQ et al., 2002).

How does performance in simple sites translate into larger, more drug-like sites? The consensus of many retrospective and prospective docking screens is that the ability to predict ligands and their geometries diminishes considerably in biology-relevant targets. In most cases, this failure reflects the increased complexity of the binding sites and the greater opportunities to find decoy ligand geometries. Nevertheless, there are examples of successful ligand prediction followed by structural determination, and in these cases the docking prediction is often close to the experimental result (Figure 4) (Gradler U et al., 2001; Gruneberg S et al., 2002; Powers RA and Shoichet BK, 2002; Wei BQ et al., 2002). These studies suggest that when the method does correctly predict a new ligand, even for a complicated, drug-like binding site, it does so for the right reasons.

An important question is whether structure-based screening is worth the effort, assuming groups have access to high-throughput screening for ligand discovery. The two types of screens have been compared only a few times

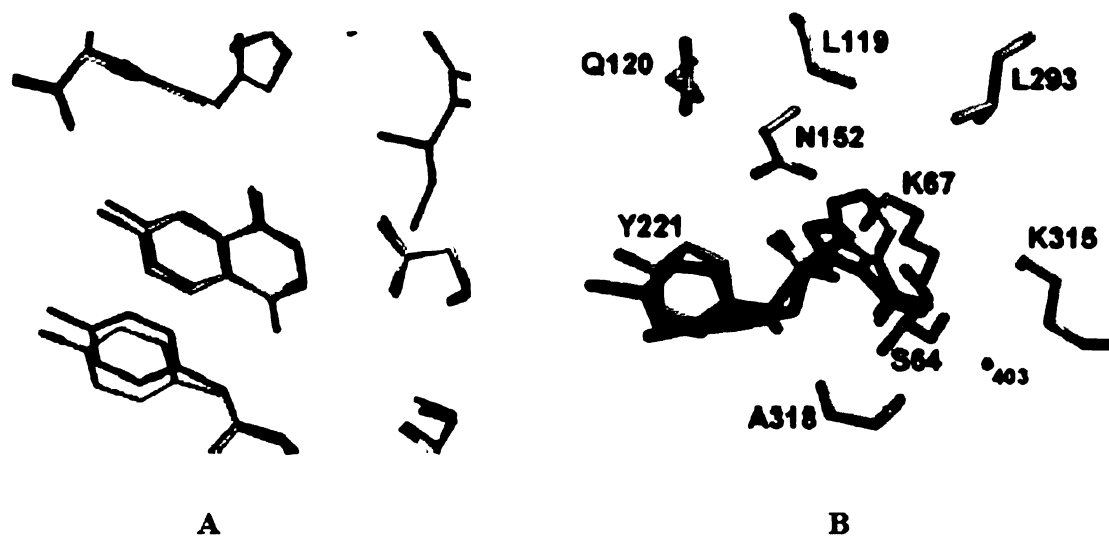


Figure 4: Predicted vs. experimental structures from virtual screening.

A) The docked (carbons in orange) vs. the crystallographic structure of the 4-aminophthalhydrazide bound to tRNA guanine transglycosylase (Gradler U et al., 2001)

B) The docked (carbons in green) versus the crystallographic structure (carbons in orange) of 3-((4-chloroanilino)-sulfonyl)-thiophene-2-carboxylate bound to β -lactamase (enzyme carbons in gray) (Powers RA et al., 2002)

Technique	Number compounds tested	Hits with $IC_{50} < 100 \mu\text{M}$	Hits with $IC_{50} < 10 \mu\text{M}$	Rule of five compliant hits ^a	Hit Rate ^b
HTS	400,000	85	6	23	0.021%
Docking	365	127	18	73	34.8%

Table 3. Hit rates and drug-like properties for inhibitors discovered with high-throughput and virtual screening against the enzyme PTP-1B (Doman TN et al., 2002)

a) Number of 100 μM or better inhibitors that passed all four “rule of five” criteria (Doman TN et al., 1996; Paiva AM et al., 2001).

b) The number of compounds experimentally tested divided by the number with IC_{50} values of 100 μM or less.

publicly, though rather more in unpublished industrial work. In the few published studies, the virtual screens had “hit rates” 10- to 1,700-fold higher than the empirical screens (Table 3) (Doman TN et al., 2002; Kick EK et al., 1997; Oshiro C et al., 2004; Paiva AM et al., 2001; Wyss PC et al., 2003). In the case with the best hit-rate enhancement, that of the diabetes-associated enzyme PTP-1B, the comparison was an imperfect one. Here, different libraries were targeted by the virtual and high-throughput screen, and a slightly different assay was used. In very recent work, Eric Brown and colleagues at McMaster University challenged the virtual screening community to predict the affinities of 50,000 molecules, none of which had been tested before but which were about to be tested in a high-throughput screen against dihydrofolate reductase (DHFR). In this experiment, the docking and HTS libraries were precisely the same, as were the experimental conditions. One of the startling results of this experiment was the very small number of hits to emerge from the screen. Indeed, whereas several groups were able to enrich putative inhibitors among their high-scoring molecules, the experimental group eventually concluded that they had no reliable hits at all (Elowe NH et al., 2005; Lang PT et al., 2005). Intriguingly, several of the computational groups were able to indicate the lack of binders as part of their predictions (Brenk R et al., 2005). Whereas the lack of experimental hits prevents definitive conclusions from this study, what does seem clear is that there is room for more of these comparative studies and “competitions.”

CONCLUSIONS AND FUTURE DIRECTIONS

We have described the general steps involved in docking as well as typical algorithms for addressing each stage of the methodology. We have also listed examples in which prepackaged programs have been successfully used to discover novel inhibitors of a wide range of medically relevant applications. In the future, these methods will become progressively more integrated in the drug design process.

However, there are still several components of docking that need improvement. The two most-debated open questions in the field involve improving the scoring functions and developing algorithms for receptor flexibility. For scoring functions, research focuses on improving the treatment of solvent and the effects of entropy loss upon binding. Most docking approaches currently include drastic approximations of both of these properties, which have shown improvements over older methods. However, it is necessary to develop new schemes that treat these issues more accurately while preserving the speed of the calculation. Configurational entropy contributions are also difficult to calculate. Techniques that use molecular dynamics simulations to generate ensembles of ligand positions that are then rescored with high-accuracy scoring functions generate very accurate free energies of binding; however, they are computationally expensive (Kollman PA et al., 2000). It will be necessary to develop sampling techniques able to generate Boltzmann-weighted ensembles of ligand poses without requiring the use of expensive molecular dynamics calculations. Similarly, for receptor flexibility, many prepackaged programs have

recently begun to develop new algorithms that allow for some measure of induced fit in the binding site. Several of these methods show great promise and will be further perfected with time, allowing for more and more structural rearrangement.

ACKNOWLEDGMENTS

We thank the sponsors of this work—the National Institute of General Medical Sciences of the National Institutes of Health through grants GM 56531 (P. O. Montellano, PI to IDK) and GM 59957 (BKS) as well as the AFPE fellowship to PTL.

REFERENCES

- Abagyan, R, Totrov, M and Kuznetsov, D. *J Comput Chem.* **15** (1994) 488-506.
- Alvarez, JC. *Curr Opin Chem Biol.* **8** (2004) 365-370.
- Anderson, AC. *Chem Biol.* **10** (2003) 787-97.
- Arnold, JR, Burdick, KW, Pegg, SCH, Toba, S, Lamb, ML and Kuntz, ID. *J Chem Inf Comput Sci.* **44** (2004) 2190-2198.
- Bender, A and Glen, RC. *J Chem Inf Model.* **45** (2005) 1369-1375.
- Berman, HM, Olson, WK, Beveridge, DL, Westbrook, J, Gelbin, A, Demeny, T, Hsieh, SH, Srinivasan, AR and Schneider, B. *Biophys J.* **63** (1992) 751-9.
- Berman, HM, Westbrook, J, Feng, Z, Gilliland, G, Bhat, TN, Weissig, H, Shindyalov, IN and Bourne, PE. *Nucleic Acids Res.* **28** (2000) 235-242.
- Berman, HM and Westbrook, JD. *Am J Pharmacogenomics.* **4** (2004) 247-52.
- Bernacki, K, Kalyanaraman, C and Jacobson, MP. *J Biomol Screen.* **10** (2005) 675-681.

- Beveridge, DL and Dicapua, FM. *Annu Rev Biophys Biophys Chem.* **18** (1989) 431-492.
- Brenk, R, Irwin, JJ and Shoichet, BK. *J Biomol Screen.* **10** (2005) 667-674.
- Brenk, R, Vetter, SW, Boyce, SE, Goodin, DB and Shoichet, BK. *J Mol Biol.* **357** (2006) 1449-1470.
- Brik, A, Alexandratos, J, Lin, YC, Elder, JH, Olson, AJ, Wlodawer, A, Goodsell, DS and Wong, CH. *ChemBioChem.* **6** (2005) 1167-9.
- Brik, A, Muldoon, J, Lin, YC, Elder, JH, Goodsell, DS, Olson, AJ, Fokin, VV, Sharpless, KB and Wong, CH. *ChemBioChem.* **4** (2003) 1246-1248.
- Brooijmans, N and Kuntz, ID. *Annu Rev Biophys Biomol Struct.* **32** (2003) 335-373.
- Chance, MR, Fiser, A, Sali, A, Pieper, U, Eswar, N, Xu, GP, Fajardo, JE, Radhakannan, T and Marinkovic, N. *Genome Res.* **14** (2004) 2145-2154.
- Claussen, H, Buning, C, Rarey, M and Lengauer, T. *J Mol Biol.* **308** (2001) 377-395.
- Doman, TN, Cibulskis, JM, Cibulskis, MJ, McCray, PD and Spangler, DP. *J Chem Inf Comput Sci.* **36** (1996) 1195-1204.
- Doman, TN, McGovern, SL, Witherbee, BJ, Kasten, TP, Kurumbail, R, Stallings, WC, Connolly, DT and Shoichet, BK. *J Med Chem.* **45** (2002) 2213-2221.
- Eldridge, MD, Murray, CW, Auton, TR, Paolini, GV and Mee, RP. *J Comput Aided Mol Des.* **11** (1997) 425-45.
- Elowe, NH, Blanchard, JE, Cechetto, JD and Brown, ED. *J Biomol Screen.* **10** (2005) 653-657.
- Enyedy, IJ, Ling, Y, Nacro, K, Tomita, Y, Wu, X, Cao, Y, Guo, R, Li, B, Zhu, X, Huang, Y, Long, YQ, Roller, PP, Yang, D and Wang, S. *J Med Chem.* **44** (2001) 4313-24.
- Evers, A and Klebe, G. *Angew Chem, Int Ed.* **43** (2004a) 248-251.
- Evers, A and Klebe, G. *J Med Chem.* **47** (2004b) 5381-92.
- Ewing, TJ, Makino, S, Skillman, AG and Kuntz, ID. *J Comput Aided Mol Des.* **15** (2001) 411-28.

- Feig, M and Brooks, CL. *Curr Opin Struct Biol.* **14** (2004) 217-224.
- Feig, M, Onufriev, A, Lee, MS, Im, W, Case, DA and Brooks, CL. *J Comput Chem.* **25** (2004) 265-284.
- Freymann, DM, Wenck, MA, Engel, JC, Feng, J, Focia, PJ, Eakin, AE and Craig, SP. *Chem Biol.* **7** (2000) 957-68.
- Friesner, RA, Banks, JL, Murphy, RB, Halgren, TA, Klicic, JJ, Mainz, DT, Repasky, MP, Knoll, EH, Shelley, M, Perry, JK, Shaw, DE, Francis, P and Shenkin, PS. *J Med Chem.* **47** (2004) 1739-1749.
- Gehlhaar, DK, Verkhivker, GM, Rejto, PA, Sherman, CJ, Fogel, DB, Fogel, LJ and Freer, ST. *Chem Biol.* **2** (1995) 317-24.
- Gohlke, H and Case, DA. *J Comput Chem.* **25** (2004) 238-250.
- Gradler, U, Gerber, HD, Goodenough-Lashua, DM, Garcia, GA, Ficner, R, Reuter, K, Stubbs, MT and Klebe, G. *J Mol Biol.* **306** (2001) 455-467.
- Graves, AP, Brenk, R and Shoichet, BK. *J Med Chem.* **48** (2005) 3714-3728.
- Gruneberg, S, Stubbs, MT and Klebe, G. *J Med Chem.* **45** (2002) 3588-3602.
- Gruneberg, S, Wendt, B and Klebe, G. *Angew Chem Int Ed Engl.* **40** (2001) 389-393.
- Hajduk, PJ, Huth, JR and Fesik, SW. *J Med Chem.* **48** (2005) 2518-25.
- Halgren, TA, Murphy, RB, Friesner, RA, Beard, HS, Frye, LL, Pollard, WT and Banks, JL. *J Med Chem.* **47** (2004) 1750-1759.
- Hensen, C, Hermann, JC, Nam, KH, Ma, SH, Gao, JL and Holtje, HD. *J Med Chem.* **47** (2004) 6673-6680.
- Hoffmann, D, Kramer, B, Washio, T, Steinmetzer, T, Rarey, M and Lengauer, T. *J Med Chem.* **42** (1999) 4422-4433.
- Honig, B and Nicholls, A. *Science.* **268** (1995) 1144-1149.
- Huang, N, Nagarsekar, A, Xia, G, Hayashi, J and MacKerell, AD, Jr. *J Med Chem.* **47** (2004) 3502-11.
- Irwin, JJ and Shoichet, BK. *J Chem Inf Model.* **45** (2005) 177-182.

- Iwata, Y, Arisawa, M, Hamada, R, Kita, Y, Mizutani, MY, Tomioka, N, Itai, A and Miyamoto, S. *J Med Chem.* **44** (2001) 1718-28.
- Jones, G, Willett, P, Glen, RC, Leach, AR and Taylor, R. *J Mol Biol.* **267** (1997) 727-748.
- Jorgensen, WL. *Science.* **303** (2004) 1813-1818.
- Kick, EK, Roe, DC, Skillman, AG, Liu, G, Ewing, TJ, Sun, Y, Kuntz, ID and Ellman, JA. *Chem Biol.* **4** (1997) 297-307.
- Kollman, P. *Chemical Reviews.* **93** (1993) 2395-2417.
- Kollman, PA, Massova, I, Reyes, C, Kuhn, B, Huo, S, Chong, L, Lee, M, Lee, T, Duan, Y, Wang, W, Donini, O, Cieplak, P, Srinivasan, J, Case, DA and Cheatham, TE, 3rd. *Acc Chem Res.* **33** (2000) 889-97.
- Kubinyi, H Ed. Success Stories of Computer-Aided Design. Wiley: New York (in press).
- Kuntz, ID. "Stability and Dynamics of Globular Proteins." Presented at AAAS (1982).
- Laird, ER and Blake, JF. *Curr Opin Drug Discovery Dev.* **7** (2004) 354-359.
- Lang, PT, Kuntz, ID, Maggiora, GM and Bajorath, J. *J Biomol Screen.* **10** (2005) 649-652.
- Li, C, Xu, L, Wolan, DW, Wilson, IA and Olson, AJ. *J Med Chem.* **47** (2004) 6681-90.
- Lingham, RB, Silverman, KC, Jayasuriya, H, Kim, BM, Amo, SE, Wilson, FR, Rew, DJ, MD, MDS, Bergstrom, JD, Koblan, KS, Graham, SL, Kohl, NE, Gibbs, JB and Singh, SB. *J Med Chem.* **41** (1998) 4492-501.
- Lorber, DM and Shoichet, BK. *Protein Sci.* **7** (1998) 938-950.
- Lorber, DM and Shoichet, BK. *Curr Top Med Chem.* **5** (2005) 739-749.
- Lyne, PD, Kenny, PW, Cosgrove, DA, Deng, C, Zabudoff, S, Wendoloski, JJ and Ashwell, S. *J Med Chem.* **47** (2004) 1962-8.
- McMartin, C and Bohacek, RS. *J Comput-Aided Mol Des.* **11** (1997) 333-344.
- Meng, EC, Shoichet, BK and Kuntz, ID. *J Comput Chem.* **13** (1992) 505-524.

- Miranker, A and Karplus, M. *Proteins*. **11** (1991) 29-34.
- Morris, GM, Goodsell, DS, Halliday, RS, Huey, R, Hart, WE, Belew, RK and Olson, AJ. *J Comput Chem*. **19** (1998) 1639-1662.
- Moustakas, D, Lang, P, Pegg, S, Pettersen, E, Kuntz, I, Brooijmans, N and Rizzo, R. *J Comput-Aided Mol Des*. (in press)
- Murcko, MA. "An Introduction to De Novo Ligand Design " In Practical Application of Computer-Aided Drug Design. P. S. Charifson Ed. Marcel Dekker, Inc: New York (1997) 355-395.
- Nissen, P, Hansen, J, Ban, N, Moore, PB and Steitz, TA. *Science*. **289** (2000) 920-30.
- Nissink, JW, Murray, C, Hartshorn, M, Verdonk, ML, Cole, JC and Taylor, R. *Proteins*. **49** (2002) 457-71.
- Oshiro, C, Bradley, EK, Eksterowicz, J, Evensen, E, Lamb, ML, Lanctot, JK, Putta, S, Stanton, R and Grootenhuys, PD. *J Med Chem*. **47** (2004) 764-7.
- Osterberg, F, Morris, GM, Sanner, MF, Olson, AJ and Goodsell, DS. *Proteins: Struct, Funct, Genet*. **46** (2002) 34-40.
- Paiva, AM, Vanderwall, DE, Blanchard, JS, Kozarich, JW, Williamson, JM and Kelly, TM. *Biochim Biophys Acta Protein Struct Mol Enzymol*. **1545** (2001) 67-77.
- Pang, YP, Perola, E, Xu, K and Prendergast, FG. *J Comput Chem*. **22** (2001a) 1750-1771.
- Pang, YP, Xu, K, Kollmeyer, TM, Perola, E, McGrath, WJ, Green, DT and Mangel, WF. *FEBS Lett*. **502** (2001b) 93-7.
- Pearlman, DA, Case, DA, Caldwell, JW, Ross, WS, Cheatham, TE, Debolt, S, Ferguson, D, Seibel, G and Kollman, P. *Comput Phys Commun*. **91** (1995) 1-41.
- Powers, RA, Morandi, F and Shoichet, BK. *Structure*. **10** (2002) 1013-1023.
- Powers, RA and Shoichet, BK. *J Med Chem*. **45** (2002) 3222-34.
- Richards, FM. *Annu Rev Biophys Bioeng*. **6** (1977) 151-176.
- Rizzo, RC, Wang, DP, Tirado-Rives, J and Jorgensen, WL. *J Am Chem Soc*. **122** (2000) 12898-12900.

- Rosenfeld, RJ, Goodsell, DS, Musah, RA, Morris, GM, Goodin, DB and Olson, AJ. *J Comput-Aided Mol Des.* **17** (2003) 525-536.
- Schapira, M, Raaka, BM, Das, S, Fan, L, Totrov, M, Zhou, ZU, Wilson, SR, Abagyan, R and Samuels, HH. *Proc Natl Acad Sci U S A.* **100** (2003) 7354-7359.
- Schapira, M, Raaka, BM, Samuels, HH and Abagyan, R. *BMC Struct Biol.* **1** (2001) 1.
- Sherman, W, Day, T, Jacobson, MP, Friesner, RA and Farid, R. *J Med Chem.* **49** (2006) 534-553.
- Staunton, D, Owen, J and Campbell, ID. *Acc Chem Res.* **36** (2003) 207-14.
- Totrov, M and Abagyan, R. *Proteins: Struct, Funct, Genet.* (1997) 215-220.
- Verdonk, ML, Cole, JC, Hartshorn, MJ, Murray, CW and Taylor, RD. *Proteins: Struct, Funct, Genet.* **52** (2003) 609-623.
- Wang, JM, Morin, P, Wang, W and Kollman, PA. *J Am Chem Soc.* **123** (2001) 5221-5230.
- Webb, TR. *Curr Opin Drug Discovery Dev.* **8** (2005) 303-308.
- Wei, BQ, Baase, WA, Weaver, LH, Matthews, BW and Shoichet, BK. *J Mol Biol.* **322** (2002) 339-355.
- Wei, BQ, Weaver, LH, Ferrari, AM, Matthews, BW and Shoichet, BK. *J Mol Biol.* **337** (2004) 1161-1182.
- Weiner, PK and Kollman, PA. *J Comp Chem.* **2** (1981) 287.
- Wyss, PC, Gerber, P, Hartman, PG, Hubschwerlen, C, Locher, H, Marty, HP and Stahl, M. *J Med Chem.* **46** (2003) 2304-12.
- Zhou, ZG and Madura, JD. *Proteins: Struct, Funct, Bioinf.* **57** (2004) 493-503.
- Zou, XQ, Sun, YX and Kuntz, ID. *J Am Chem Soc.* **121** (1999) 8033-8043.

“The strongest arguments prove nothing so long as the conclusions are not verified by experience. Experimental science is the queen of sciences and the goal of all speculation.”

--Roger Bacon

Chapter 2

Evaluating the High-Throughput Screening Computations

P. Therese Lang¹, Irwin D. Kuntz^{2*}, Gerald M. Maggiora³, and Jürgen Bajorath⁴

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

²Department of Pharmaceutical Chemistry
University of California, San Francisco

³College of Pharmacy
University of Arizona, Tucson

⁴Department of Life Science Informatics, B-IT
Rheinische Friedrich-Wilhelms-University Bonn, Germany

**To whom all correspondence should be addressed (kuntz@cgl.ucsf.edu)*

ABSTRACT

The judges evaluated the submissions for the McMaster University High-Throughput Data-Mining and Docking Competition based on 3 criteria: identification of active compounds, percent enrichment, and overview of the competition. Using these metrics, 4 of the participating groups found meaningful enrichment, and 3 groups made perceptive comments about the general nature of the competition.

Keywords: high-throughput screening, docking, QSAR/MS, DHFR, judges' review

INTRODUCTION

The stated objective of the McMaster University High-Throughput Data-Mining and Docking Competition was for computational chemists and data analysts to predict the results of a high-throughput screening (HTS) experiment. The collected predictions were to then enable assessment of the utility of computational approaches to identify compounds for follow-on screening. The protocol called for a single target, *Escherichia coli* dihydrofolate reductase (DHFR), and 2 small-molecule libraries to be screened: a training set whose results would be given to the investigators and a test set that would be predicted blind. The computational groups were asked to rank-order test set compounds or otherwise indicate those compounds thought to be active.

Blind tests, in which investigators are asked to predict an unknown outcome, have a long history in science and play an important role in the advancement of methods. They are more powerful than a posteriori rationalization because of their potential to expose random and systematic errors in both computations and experiments. Such initiatives are especially useful if they provide a forum for the evaluation and comparison of different types of computational approaches.

DHFR catalyzes the NADPH-dependent reduction of 7,8-dihydrofolate (DHF) into 5,6,7,8 tetrahydrofolate (THF). Because THF is essential for the biosynthesis of purines, pyrimidines, and several amino acids, DHFR is an established drug target for the treatment of bacterial infections, cancer, and malaria (Anderson AC, 2005; Hitchings GH and Smith SL, 1980; Huennekens

FM, 1994). In previous work, Zolli-Juran et al.(Zolli-Juran M et al., 2003) identified 32 hit compounds—12 of which were found to be competitive inhibitors—out of a diverse library of 50,000. For the contest, the participants were given a 3-dimensional structure for each compound in this library (training set) along with the measured levels of inhibition. They were free to use these data to validate their methods or to develop predictive models of DHFR inhibition. Each group was also given coordinates for a second, chemically diverse library (test set), but the experimental results were withheld. Once the competitors submitted their results and the judges completed their evaluation, the measured levels of inhibition for the test set were released (Figure 1) (Elowe NH et al., 2005).

The design of this competition has some similarity to the Critical Assessment of Techniques for Protein Structure Prediction (CASP) competition, a biennial effort in which competitors predict the 3-dimensional structures of proteins from their amino acid sequences (*Proteins*, 1995; *Proteins*, 1997; *Proteins*, 1999; *Proteins*, 2001; *Proteins*, 2003a). Here also, the experimental structures are withheld until after the entries are submitted, and predictions are judged using a variety of structure-based metrics. A protein-protein docking competition, Critical Assessment of Predicted Interactions (CAPRI), in which participants are asked to predict the mode of association of 2 proteins based on their 3-dimensional structures, has also been established (*Proteins*, 2002; *Proteins*, 2003b). Competitions such as CASP and CAPRI have had significant impacts on their fields by setting standards that allow direct comparison of

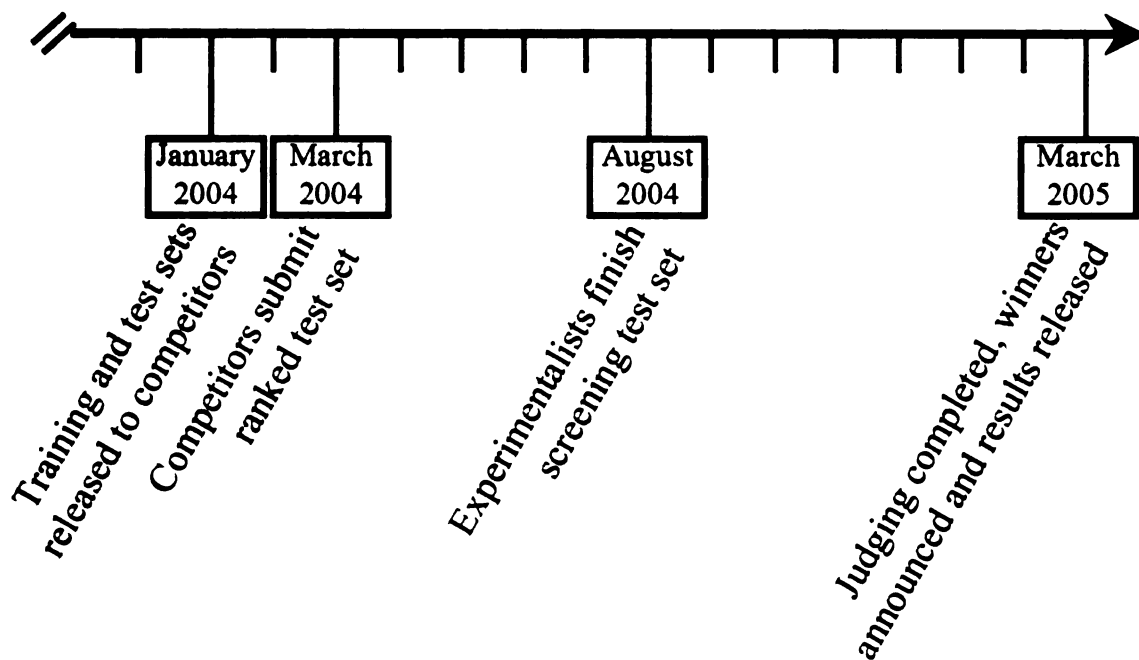


Figure 1. Timeline and major events in the McMaster University High-Throughput Data-Mining and Docking Competition.

different computational methodologies. In the spirit of these efforts, it was our hope that a well designed screening contest would illuminate and perhaps improve the computational and experimental aspects of HTS technology

EVALUATING THE ENTRIES

There was considerable interest in the computational community, yielding a total of 32 independent submissions spanning a wide variety of approaches. The methods can generally be classified as follows: quantitative structure-activity relationship-based (QSAR), molecular similarity-based (MS), or target structure-based (docking) approaches. QSAR and MS methods compare physicochemical properties and structural features of known active and inactive compounds to predict the activities of novel molecules. Docking methods predict 3-dimensional structures of target-ligand complexes and employ scoring functions that capture various physicochemical interactions such as electrostatics to rank the docked compounds and identify potentially active ones. Of the submissions, about 50% used only QSAR or MS techniques, 10% used only docking methods, and 40% used some combination of both.

The basic task of the judges was to evaluate the performance of each of the computational predictions on the test set compounds. However, after obtaining the test set screening data and the submissions, 3 complications emerged. First, although 96 test set compounds were considered active in the primary screen (defined as 75% or less residual enzymatic activity for the average of the replicates), including several compounds that showed partial-dose

response, follow-up experiments failed to confirm that any of these compounds were competitive inhibitors, suggesting that all these putative hits might function by different mechanisms such as allosteric binding. Also, as noted in the accompanying experimental article, preliminary data indicated that at least some of the test set inhibitors were nonspecific aggregators (Brenk R et al., 2005). However, because none of the training set and only a small subset of the test set compounds have been evaluated, we could not incorporate this information into the judging. Finally, some of the computational groups ranked the entire library, whereas others only submitted their predicted actives, making sophisticated data analysis unproductive.

The judges decided to use a simple enrichment test based on the predicted and reported inhibition levels. The 96 “active” test set compounds were used to determine relative enrichment values. Each group was evaluated for the ability to correctly identify actives in the top 1% (500 compounds) and the top 5% (2500 compounds) of the ranked list (Table 1). Although more sophisticated evaluation schemes could have been applied, the absence of confirmed competitive inhibitors in the test set, as well as the variability of the submissions, made such refinement impractical.

Interestingly, in addition to ranking compounds, several of the entrants commented on what they perceived as major differences between the training set and the test set. Specifically, these groups noted either that the 2 libraries showed significant chemical differences, or they predicted that the test set would

Group ^a	# Submitted ^b	Top 1% ^c	Top 5% ^d
1	50,000	1	6
2	495	0	0
3	22	0	0
4	50	1	1
5	2000	1	3
6	127	0	0
7	50,000	1	7
8	150	0	0
9	20	0	0
10	200	0	0
11	30	0	0
12	77	0	0
13	59	0	0
14	294	0	0
15	46,901	0	4
16	344	0	0
17	10	0	0
18	21	0	0
19	105	0	0
20	50,000	0	1
21	59	0	0
22	44	0	0
23	6	1	1
24	40	0	0
25	28	0	0
26	21	0	0
27	121	0	0
28	601	2	2
29	46,720	2	13
30	439	1	1
31	26	0	0
32	1000	0	0

Table 1: Number of Active Compounds Identified in Each Group's Ranked List
 Actives are defined as 75% or less residual enzymatic activity for the average of the replicates. Using this criterion, the hit rate for the test set of compounds was 0.19%.

a) The judges were blinded throughout the competition. These numbers served only to identify the groups and have no relation to either the performance in the competition or the classification of the methods.

b) If group submitted a larger list, only the top 2500 ranked compounds were used in the evaluation.

c) Actives for the top 1% (500 compounds) of the submitted ranked test set.

d) Actives for the top 5% (2500 compounds) of the submitted ranked test set.

perform worse in the experimental screen than the training set. We feel these groups deserve special recognition because they considered the overall experiment and reached conclusions that would have had a substantial impact in a drug discovery setting

IMPLICATIONS FOR EXPERIMENTAL DESIGN

Because this competition was aimed at assessing the utility of various computational methods, the optimal experimental data would include concentration data (ideally, K_{is}) and experimental structures of representative protein-ligand complexes. Prediction of actual binding affinities and, for docking, of the binding geometries would be more stringent measures of computational methods. Although this level of effort is often expended for structure-based design projects in industry, it is impractical for early stage screening involving large libraries. Nevertheless, because it is desirable that experimental screening data are as accurate as possible, the basic assays need to show statistical reproducibility with low standard deviations. This level of data would also facilitate analysis of the false-positive and false-negative rates in the computational predictions. Finally, in analyzing the experimental data for this particular screen, the definition of “active” compounds was statistically different for the test set as compared to the training set because the experimental values had a different distribution. To be fair to the competitors, the cutoff between “active” and “inactive” should have remained constant regardless of the composition of the library.

IMPLICATIONS FOR COMPUTATIONAL METHODS

In general, no group predicted more than 15% of the apparent inhibitors in the test set, and the best group showed a nominal enrichment of only 2-fold. A contributing factor to such low success rates may have been the decision to design the test set library to be structurally divergent from the training set library. In particular, the difference between the two may have presented a significant problem for QSAR and MS methods because these approaches rely on building predictive models from training set data that are transferable. An analogy would be if CASP offered a training set of proteins from a limited number of folding classes and a test set consisting only of proteins not in those classes.

Of course, another major complication was that the test set contained very few, if any, competitive DHFR inhibitors. However, in this case, a desirable computational result would have been the prediction of “no inhibitors” in the test set. In addition, we noted that many of the computational groups either did not consider or had no way to judge the absolute level of their scores, and thus they reported only the rank order. Extra credit was given to the 3 groups that noted the qualitative difference between the training and test sets. However, we suspect that an implicit assumption was made that there would be well-characterized inhibitors in the test set, which may also have influenced the reporting process. It is important to note that a screen that finds no useful inhibitors is often encountered in drug discovery programs, and the ability to make reliable negative predictions would have real utility in the design of primary

screens. Regardless, in the context of blind predictions, the screening data situation was less than ideal.

RECOMMENDATIONS FOR FUTURE COMPETITIONS

We note that the choice of any specific target, library, or experimental design will affect the outcome of the overall contest. However, for competitions of this type to be useful to the community, it should address recognized needs. As stated above, this first experience brought to light several important issues. For example, the lack of competitive inhibitors made it extremely difficult to evaluate the predictive ability of the different methods. Prescreening a library of 100,000 compounds and then using it to create training and test sets based on equal distributions of experimental data and chemical properties would resolve this issue and also remove the bias against QSAR- and MS-type methods. This observation also points to the need for methods to identify those compounds for which a QSAR or MS model can be applied and for those that lie beyond the scope of the original training set, potentially through similarity-based methods, as suggested by Sheridan et al (Sheridan RP et al., 2004).

In addition to the small-molecule libraries, another area of improvement would be expanding the number and variety of targets screened. By presenting at least 2 unrelated targets and evaluating predictions based on both sets of rankings, any bias introduced by the characteristics of a particular binding site (i.e., hydrophobic, polar, highly charged, critical hydrogen bond, etc.) can be reduced. To further challenge the capabilities of the prediction methods, it would

also be useful to present 2 targets from the same protein family. This type of test would make it possible to evaluate the ability of the computational methods to discriminate between selective inhibitors. Conversely, it may be useful to present systems in which the target structure is unknown, but could perhaps be modeled, to test the ability of docking algorithms to deal with this equally real type of screening situation.

CONCLUSIONS

Although unintentional, this event has revealed 2 major areas for improvements in the field of computational prediction. First, the aggregation phenomenon manifests itself as false positives in the experimental screen.(Brenk R et al.) Computational procedures that could predict this property de novo would be extremely useful in culling out these types of molecules from databases before screening begins. Second, the experimental result that fails to find competitive inhibitors is one that occurs frequently in real-life screens. Computational methods that can move beyond ranking actives to predicting whether actives even exist would be of great utility.

ACKNOWLEDGEMENTS

We thank the American Foundation for Pharmaceutical Chemistry and the National Institutes of Health for support for P.T.L. We also thank Eric Brown, McMaster University HTS Lab, the Ontario Research and Development

Challenge Fund, the ChemBridge Corporation for developing this competition, and Christian Parker for helpful conversations and insights.

REFERENCES

- Anderson, AC. *Drug Discov Today*. 10 (2005) 121-8.
- Brenk, R, Irwin, JJ and Shoichet, BK. *J Biomol Screen*. 10 (2005) 667-674.
- Elowe, NH, Blanchard, JE, Cechetto, JD and Brown, ED. *J Biomol Screen*. 10 (2005) 653-657.
- Hitchings, GH and Smith, SL. *Adv Enzyme Regul*. 18 (1980) 349-71.
- Huennekens, FM. *Adv Enzyme Regul*. 34 (1994) 397-419.
- Proteins*. 23 (1995) 295-460.
- Proteins*. S1 (1997) 1-230.
- Proteins*. S3 (1999) 1-237.
- Proteins*. S5 (2001) 1-199.
- Proteins*. 47 (2002) 257-407.
- Proteins*. S6 (2003a) 333-595.
- Proteins*. 52 (2003b) 1-122.
- Sheridan, RP, Feuston, BP, Maiorov, VN and Kearsley, SK. *J Chem Inf Comput Sci*. 44 (2004) 1912-28.
- Zolli-Juran, M, Cechetto, JD, Hartlen, R, Daigle, DM and Brown, ED. *Bioorg Med Chem Lett*. 13 (2003) 2493-6.

“In this house, we OBEY the laws of thermodynamics!”

--Homer Simpson

Chapter 3

Development and Validation of a Modular, Extensible Docking Program: DOCK 5

Demetri T. Moustakas^{1*}, P. Therese Lang^{2*}, Scott Pegg³, Eric Pettersen³, Irwin D. Kuntz^{3*}, Natasja Broojimans², Robert C. Rizzo⁴

¹Joint Graduate Program in Bioengineering
University of California, Berkeley/University of California, San Francisco

²Graduate Program in Chemistry and Chemical Biology
University of California, San Francisco

³Department of Pharmaceutical Chemistry
University of California, San Francisco

⁴Department of Applied Mathematics and Statistics
Stony Brook University, New York

** Authors contributed equally to this work*

**To whom all correspondence should be addressed (kuntz@cgl.ucsf.edu)*

ABSTRACT

We report on the development and validation of a new version of DOCK. The algorithm has been rewritten in a modular format, which allows for easy implementation of new scoring functions, sampling methods and analysis tools. We validated the sampling algorithm with a test set of 114 protein-ligand complexes. Using an optimized parameter set, we are able to reproduce the crystal ligand pose to within 2 Å of the crystal structure for 79% of the test cases using our rigid ligand docking algorithm with an average run time of 1 minute per complex and for 72% of the test cases using our flexible ligand docking algorithm with an average run time of 5 minutes per complex. Finally, we perform an analysis of the docking failures in the test set and determine that the sampling algorithm is generally sufficient for the binding pose prediction problem for up to 7 rotatable bonds; i.e. 99% of the rigid ligand docking cases and 95% of the flexible ligand docking cases are sampled successfully. We point out that success rates could be improved through more advanced modeling of the receptor prior to docking and through improvement of the force field parameters, particularly for metal-based cofactors.

Keywords: automated docking, scoring functions, structure-based drug design, flexible docking, binding mode prediction, incremental construction, validation

INTRODUCTION

Transient non-covalent interactions are critical for biological processes. The sequencing of a variety of genomes and the development of proteomics techniques have enabled scientists to study these interactions on the widest scales (Kopec KK et al., 2005). Advances in X-ray crystallography, nuclear magnetic resonance spectroscopy, and other experimental structure techniques provide the ability to study these interactions at an atomic level of detail (Congreve M et al., 2005). One important application of these advances is the design of small molecules that interact with cellular processes to modify biological activity and treat disease.

The drug discovery process typically requires between 10-15 years from early discovery until FDA approval (Kraljevic S et al., 2004). Computational tools—such as virtual screening, homology modeling and cheminformatics—are applied both to facilitate various stages of research and to create models that explain experimental data (Hillisch A et al., 2004; Posner BA, 2005; Schneck V and Bostrom J, 2006). Molecular docking, which can broadly be defined as the prediction of the orientation of two molecules with respect to one another, is a computational technique that has been successfully used in both of these capacities (Alvarez JC, 2004). In drug design applications, one molecule is typically a protein or nucleic acid drug target—the receptor—and the other is a potential ligand. In these applications, docking is used to identify novel ligands that interact with a biomolecular target and to predict the geometric position (binding mode) of ligands with respect to the target of interest.

DOCK Background

DOCK is one example of a family of molecular docking packages available, which includes Glide, FlexX, and GOLD (Table 1) (Friesner RA et al., 2004; Halgren TA et al., 2004; Kramer B et al., 1999; Verdonk ML et al., 2003). Each of these programs consists of two key parts: a search algorithm and a scoring function. The search algorithm samples both the relative orientations of the two objects as well as their conformations. It must be thorough enough to ensure adequate coverage of the binding free energy landscape in order to find the global minimum of the scoring function. The scoring function ranks the various geometries generated by the search algorithm, proposing the top-scoring pose as the global minimum. It must rapidly evaluate receptor-ligand complex stability with sufficient accuracy such that the global minimum of the scoring function agrees with experimental data.

The number of degrees of freedom in receptor-ligand interactions is very large, and several approximations must be made to ensure that the docking problem is tractable. Many different approaches, ranging from freezing non-essential motions to the use of preferred conformations, have been developed to reduce the number of degrees of freedom sampled (Kitchen DB et al., 2004). In the DOCK algorithm, for example, the receptor is considered to be conformationally rigid, requiring only the ligand conformational, translational and rotational degrees of freedom to be sampled during complex formation. This assumption is reasonable in docking applications in which either the receptor conformation does not change dramatically upon ligand binding or in which the aim is to

Method	Ligand Sampling Method ^a	Receptor Sampling Method ^a	Scoring Function ^b	Solvation Scoring ^{c,d}
DOCK 4/5	IC	SE	MM	DDD, GB, PB
FlexX/FlexE	IC	SE	ED	NA
Glide	CE + MC	TS	MM + ED	DS
GOLD	GA	GA	MM + ED	NA

Table 1: Summary of scoring functions and sampling algorithms for commonly used docking programs. (a) Sampling methods are defined as Genetic Algorithm (GA), Conformational Expansion (CE), Monte Carlo (Smith GM et al.), Incremental Construction (IC), Merged Target Structure Ensemble (SE), Torsional Search (TS). (b) Scoring functions are defined as either empirically derived (Hillisch A et al.) or based on molecule mechanics (Case DA et al.). (c) If the package does not accommodate this option, the symbol NA (Not Available) is used. (d) Additional accuracy can be added to the scoring function using implicit solvent models. The most commonly used options are Distance Dependent Dielectric (DDD), a parameterized desolvation term (Richards FM), Generalized Born (GB) and linearized Poisson Boltzmann (PB).

stabilize a particular receptor conformation.

In order to guide the search for ligand orientations with respect to the receptor, a negative image of the active site volume is created by placing spheres on the solvent accessible surface area of the receptor, thus restricting the ligand orientational sampling to the most relevant region on the surface of the receptor (Shoichet BK et al., 1992). To sample the internal degrees of freedom of the ligand, DOCK uses the incremental construction algorithm, anchor-and-grow, which separates the ligand flexibility into two steps (Ewing TJA and Kuntz ID, 1997; Leach AR and Kuntz ID, 1992), (Figure 1). First, the largest rigid substructure of the ligand (anchor) is identified and rigidly oriented in the active site by matching its heavy atoms centers to the receptor sphere centers (orientation). The anchor orientations are evaluated and optimized using the scoring function and the energy minimizer. The orientations are then ranked according to their score, spatially clustered by heavy atom root mean squared deviation (RMSD), and prioritized (pruning). Next, the remaining flexible portion of the ligand is built onto the best anchor orientations within the context of the receptor (grow). It is assumed that the shape of the binding site will help restrict the sampling of ligand conformations to those that are most relevant for the receptor geometry.

In order to evaluate a large number of ligand poses in a reasonable amount of time, approximate scoring functions must be used. Once again, numerous solutions to this problem have been proposed, including a variety of empirical and physics-based terms (Kitchen DB et al., 2004). DOCK uses an

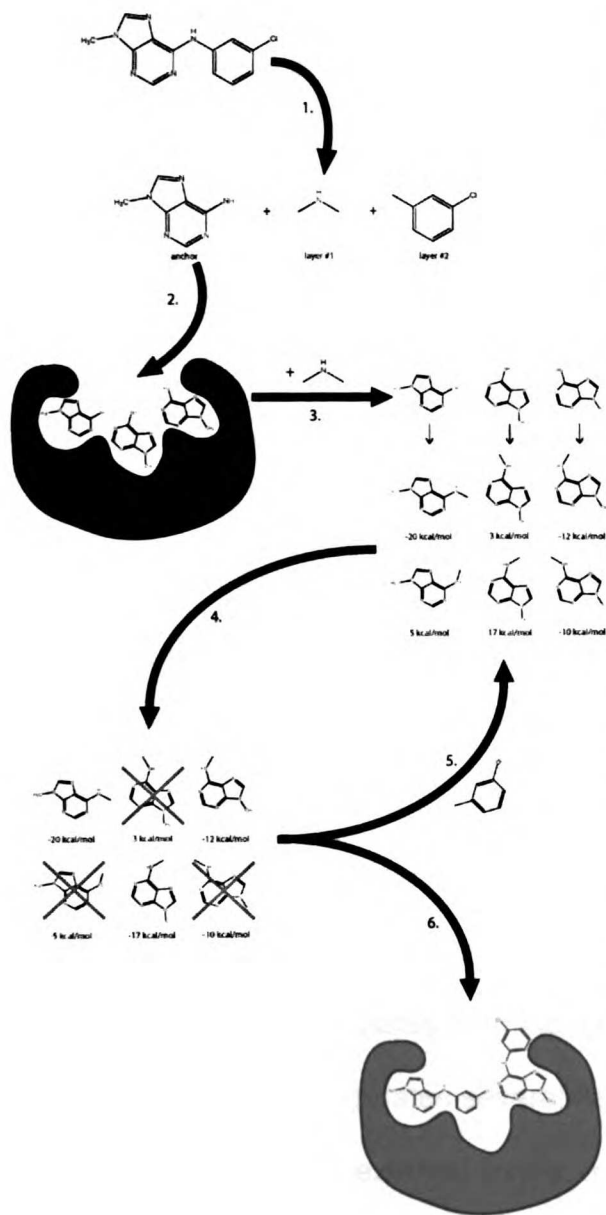


Figure 1. The “anchor and grow” conformational search algorithm. The algorithm performs the following steps: (1) DOCK perceives the molecule’s rotatable bonds, which it uses to identify an anchor segment and overlapping rigid layer segments. (2) Rigid docking is used to generate multiple poses of the anchor within the receptor. (3) The first layer atoms are added to each anchor pose, and multiple conformations of the layer 1 atoms are generated. An energy score within the context of the receptor is computed for each conformation. (4) The partially grown conformations are ranked by their score and are spatially clustered. The least energetically favorable and spatially diverse conformations are discarded. (5) The next rigid layer is added to each remaining conformation, generating a new set of conformations. (6) If all layers have been added, the set of completely grown conformations and orientations is returned

energy scoring function based on the AMBER molecular mechanics force field (Ewing TJA and Kuntz ID, 1997). Only the interactions between the ligand and protein are considered, leaving only intermolecular van der Waals (VDW) and electrostatic components in the function. Since the receptor is considered to be rigid, the receptor contribution to the potential energy can be pre-calculated and stored on a grid (Meng EC et al., 1992). These approximations enable the program to evaluate large libraries of small molecules against a receptor in a reasonable period of time.

This paper describes a new version of the DOCK program and explores the critical variables that control its ability to find correct binding modes in a suite of test problems. Our motivation is to provide a modular docking package that permits the easy development of new scoring functions, search algorithms, and analysis tools. Thus, each functional unit of the DOCK algorithm was implemented as a self-contained and portable module that interacts with the user through a well-defined interface (Figure 2). The object-oriented language C++ was chosen to allow each component of the DOCK algorithm to be implemented as a class, which encapsulates both the data structures and functions (Lischner R, 2003). DOCK 5 incorporates several new routines, including parallelization of the algorithm through an external library, modification of the ligand structural class to enable greater user control over sampling, and clustering of the final results by root mean square deviation. The implications of these additions will be discussed in this paper. Additional scoring functions and alternate sampling

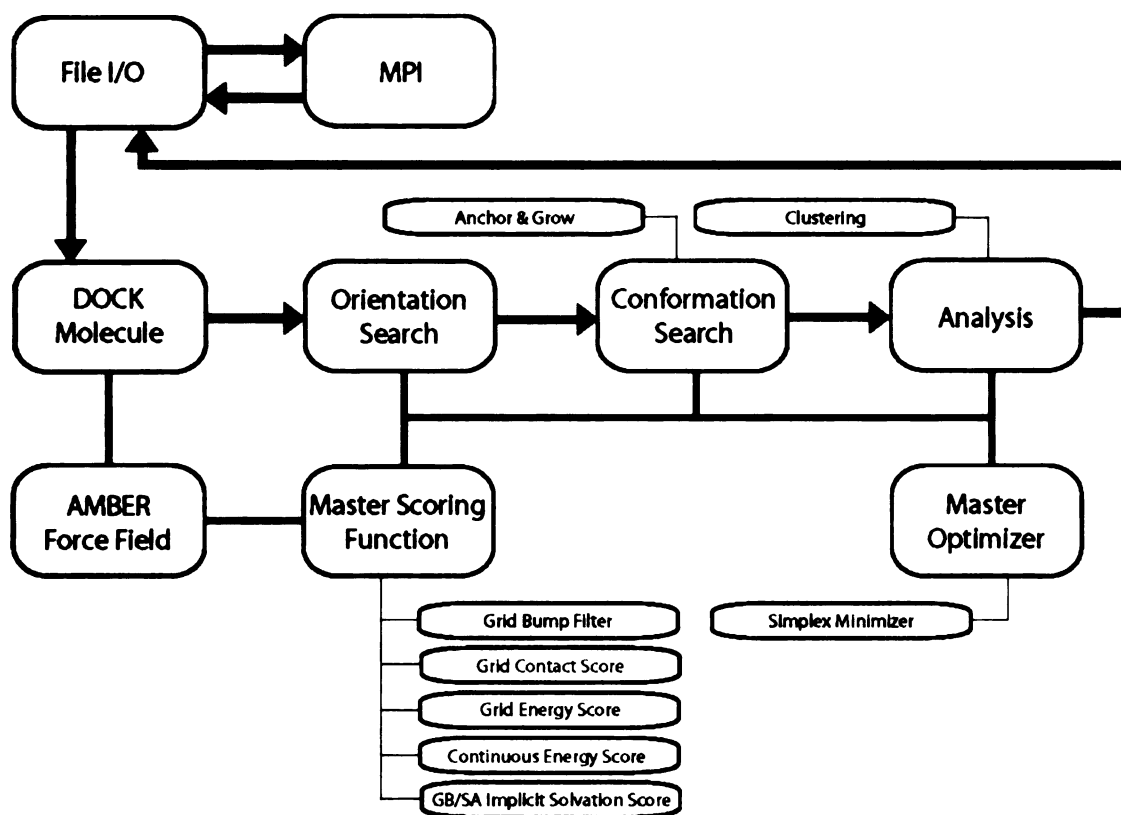


Figure 2. The major DOCK 5 classes and their interconnections. The bold arrows denote the connections between the classes that implement the DOCK sampling algorithm. The path traced by the arrows illustrates the sequence of operations performed upon a ligand molecule during docking. The bold lines (without arrowheads) denote functional connections between classes. These connections allow one class to call functions implemented in another. This diagram demonstrates that the classes implementing the DOCK sampling methods are heavily connected to a layer of classes that implement the physics engine: the force field, the scoring functions, and the energy minimizers. The thin lines denote hierarchical relationships between a master class and modular subclasses. These hierarchical arrangements allow new functional classes (scoring functions, energy minimizers, etc...) to be plugged into the existing DOCK algorithm in a modular fashion.

techniques have been implemented as well and will be discussed in future papers (<http://dock.compbio.ucsf.edu>).

Previous studies have examined the scoring function and the matching algorithm of DOCK in detail ((Ewing TJA and Kuntz ID, 1997; Meng EC et al., 1992) and equations 1-6 in (Meng EC et al., 1992)). In this paper, we pay particular attention to the robustness of the anchor-and-grow portion of the DOCK algorithm. We seek to maximize the success of complex structure prediction by independently optimizing the various steps in the anchor-and-grow algorithm. In the process, we also quantify and bound the errors for cases in which flexible docking fails and provide direction for potential areas of improvement.

Overview of Test Set

The validation of any software program requires careful testing of all aspects of the algorithm and assessment of its utility in all anticipated applications of the software. Molecular docking is commonly used in several modes, namely ligand binding mode prediction, virtual screening, and prioritization of a set of related compounds based on their affinity. However, predicting the correct binding mode of a ligand-receptor complex is a requisite step for the successful comparison of different ligands and therefore will be the focus of this paper. It is important to note, however, that predicting binding orientations is not the only metric for the accuracy and utility of docking algorithms. Optimizing DOCK for applications, including ranking libraries of small

molecules and calculating absolute free energies of binding, will be addressed in other papers (<http://dock.compbio.ucsf.edu>).

Large-scale validation of docking algorithms was long hampered by the lack of a large number of high quality protein-ligand complex crystal structures. Thanks to advances in automation in molecular biology and crystallography, the number of structures in the Protein Data Bank (PDB) continues to grow at a rapid pace (Berman HM et al., 2000). The developers of GOLD were first to test their program on a large number of available structures (Jones G et al., 1997). Their test set was compiled using a number of criteria to select candidate protein-ligand complex structures. The protein must be of pharmacological interest and the ligands must be drug-like. In addition, complexes were chosen that exhibited interesting and unusual interactions between the ligand and the protein. The final set of 100 (more recently expanded to 134) protein-ligand complexes has served as the basis for other, larger test sets (Kramer B et al., 1999; Nissink JW et al., 2002; Pang YP et al., 2001; Perola E et al., 2004).

More recently, the CCDC/Astex set compiled 305 protein-ligand complex structures by expanding the original GOLD test set (Nissink JW et al., 2002). However, the authors note that many of the new entries contain larger ligands that have more rotatable bonds, making this set less drug-like. The crystal structures in the CCDC/Astex set were evaluated for crystallographic errors and inconsistencies, yielding a “clean” set of 224 protein-ligand complexes. To create the test set for the DOCK validation studies, we filtered out 84 complexes

Protein Data Bank Identifier					
1A28	1COM	1FLR	1OKL	1TYL	2MCP
1A6W	1COY	1HAK	1PBD	1UKZ	2PCP
1A9U	1CPS	1HDC	1PDZ	1ULB	2PHH
1ABE	1D3H	1HSL	1PHD	1WAP	2PK4
1ABF	1D4P	1HYT	1PHG	1XID	2TMN
1ACJ	1DBB	1IMB	1PTV	1XIE	2YPI
1ACM	1DBJ	1IVB	1QCF	1YDR	3CPA
1ACO	1DG5	1LAH	1QPE	2AAD	3ERD
1AI5	1DID	1LCP	1QPQ	2ACK	3GPB
1AOE	1DOG	1LDM	1RNT	2ADA	3HVT
1AQW	1DR1	1LST	1ROB	2AK3	4AAH
1AZM	1DWB	1LYL	1RT2	2CHT	4COX
1BYG	1EBG	1MDR	1SNC	2CMD	4CTS
1C5C	1ETT	1MLD	1SRJ	2CPP	4FBP
1C5X	1F0R	1MRG	1TDB	2CTC	4LBD
1C83	1F0S	1MRK	1TNG	2DBL	5ABP
1CBX	1F3D	1MUP	1TNH	2GBP	5CPP
1CIL	1FGI	1NGP	1TNI	2H4N	6RNT
1CKP	1FKI	1NIS	1TNL	2LGS	7TIM

Table 2: Complexes used in the test set (total of 114 complexes)

Tetra-coordinated Zinc^a

Radius 1.700 Å

Well depth 0.067 kcal/mol

Penta-coordinated Zinc^b

Radius 1.100 Å

Well depth 0.0125 kcal/mol

Table 3: Zinc VDW parameters used to generate grids (a) Parameters used for receptors with tetra coordinated zinc ions (Aqvist J and Warshel A, 1990) (b) Parameters used for receptors with penta coordinated zinc ions (Merz KM et al., 1991)

with 8 or more rotatable ligand bonds. In addition, several of the complexes had properties that we felt made them inappropriate for a validation set. These issues included ligands that were covalently bound to the receptor (PDB code 1ASE), ligands with missing electron density (PDB code 1EED), and known sequence misregistry in the receptor (PDB code 3HVT). Ligands with vanadium that required VDW types in which we were not completely confident were also removed. The final test set contained 114 drug-like complexes (see Methods, Table 2).

METHODS

DOCK 4 to DOCK 5 Conversion

The new DOCK rigid body orienting code was written as a direct implementation of the isomorphous subgraph matching method of Crippen and Kuhl (Kuhl FS et al., 1984). All receptor sphere pairs and atom center pairs are considered for inclusion in a matching clique. This is more computationally demanding than the clique matching algorithm implemented in previous versions of DOCK that used a distance binning algorithm to restrict the clique search, in which pairs of spheres and atom centers were binned by distance. Only sphere pairs and center pairs that were within the same distance bin were considered as potential matches (Ewing TJA and Kuntz ID, 1997). The new DOCK clique matching implementation avoids bin boundaries that prevent some receptor sphere and ligand atom pairs from matching, and, as a result, it can find good matches missed by previous versions of DOCK. The rigid body rotation code was

also corrected to avoid a singularity that occurred if the spheres in the match lay within the same plane. Both of these changes improved orientational sampling.

The anchor-and-grow algorithm in the new version of DOCK was also modified to prevent premature pruning of the growth tree. The DOCK 5 anchor-and-grow code was completely rewritten with several differences in the implementations between DOCK 4 and 5. The anchor-and-grow implementation in DOCK 5 fixed a series of bugs that caused some branches of the search to be pruned when they should have been preserved for the next round of growth. The mechanism of minimization of partially grown conformers was also changed to allow the entire partial conformer to move, instead of just the latest layer, enabling more accurate ranking and pruning of the partially grown conformers.

In addition, the simplex minimizer was re-coded based on the original Nelder and Mead algorithm (Nelder JA and Mead R, 1965). The new minimizer implementation consistently found lower energy minima when using the same set of 1000 ligand orientations in a receptor, indicating that it was performing better than the previous version (data not shown). In addition, we changed the mechanism of minimization of partially grown ligand conformers to allow all atoms in the partial conformer to be minimized, rather than only the outermost layer of atoms. These changes may explain why DOCK 4 performs more poorly when run with the DOCK 5 optimized parameters (see below).

The final version of the new DOCK code, including all functions described below and all bug fixes, was posted to the DOCK web site as version 5.4.0 (<http://dock.compbio.ucsf.edu>). All experiments performed with the new

implementation of DOCK used this version and will be referred to as DOCK 5 for convenience. All experiments performed with the previous version of DOCK used version 4.0.1 and will be referred to as DOCK 4.

Conversion of the DOCK Codebase from C to C++

The design of the new DOCK 5 architecture balances the speed of the code, or computational performance, against its modularity and extensibility. The code was developed using ANSI C++ to ensure portability across multiple platforms (Lischner R, 2003). The only external library used by DOCK 5 is MPICH for parallel processing (Gropp W et al., 1996). To enable easy modification or replacement of DOCK 5 algorithm components, the DOCK 5 class structure was designed so that there are classes for each major DOCK algorithm function, and these classes interface with each other by passing instances of the DOCK 5 molecule class. Within the major functions, there are two layers of classes: those that implement the ligand sampling functions--rigid orienting, conformational searching, and minimizing--and those that implement the underlying physics engine--the force field definitions and the scoring functions. The sampling classes are applied sequentially to the ligand molecule; the physics engine classes are utilized by the sampling classes to score the ligand-receptor interaction after each step.

As a specific example of modularity, the DOCK 5 scoring functions are implemented as a master score class with five scoring function subclasses. The master score class acts as an interface to the scoring subclasses, enabling the

user to designate primary and secondary scoring functions at runtime. This design was chosen because the individual scoring functions were best implemented as individual classes; they each require different input and use different internal data structures. While they could have been implemented into one large scoring class, the result would have been quite large and disjoint. This solution was also applied to the ligand conformational search, energy minimization and post-docking analysis classes.

The DOCK 5 molecule class was designed to contain the minimum information required to specify a three-dimensional ligand conformation (atom coordinates, bond connectivity, atom partial charges, atom types and bond types) to minimize the memory required to store a molecule, allowing large arrays of molecules to be stored in RAM. Standard C-style arrays were used to store the molecular data to maximize the speed of accessing this information.

Test Set Preparation

The proteins and ligands were extracted from the PDB files, which were downloaded from the PDB website (<http://www.rcsb.org>, Table 2). The ligands were assigned atom types and bond types manually, and hydrogens were added using Sybyl. Subsequently, AM1-BCC partial electrostatic charges were calculated using the Antechamber package distributed with Amber 8 (Case DA et al., 2004; Jakalian A et al., 2000). The number of rotatable bonds of each of the ligands was measured using DOCK, and ligands with >7 rotatable bonds were eliminated from the test set. We choose 7 or fewer bonds to give a reasonable

representation of DOCK's performance using compounds similar to those of most interest in drug discovery (Hann MM and Oprea TI, 2004; Oprea TI, 2002; Oprea TI et al., 2001). The final test set that was used consisted of 114 non-covalent protein-ligand complexes (Brooijmans N, 2003) (Table 2).

For the proteins, we removed all waters, covalently linked sugars, sulfates, and halogens that were not part of the ligand. Co-factors, such as heme, ATP, and NADPH, were kept, atom and bond types were assigned manually, and Gasteiger-Hückel partial electrostatic charges were calculated using the "Compute" module in Sybyl (Gasteiger J and Marsili M, 1980; Lorber DM and Shoichet BK, 2005; Purcell WP and Singer JA, 1967). Ions, such as calcium and zinc, were considered to be part of the protein and the correct charge was assigned manually. Different VDW parameters for zinc were used depending on the coordination state of the zinc atom in the protein-ligand complex (Table 3). Hydrogens were added to the protein residues using the "Biopolymer" module in Sybyl, as were AMBER partial charges and VDW parameters (Cornell WD et al., 1995; Lorber DM and Shoichet BK, 2005). No additional optimization of the protein structure was carried out at this point.

The GRID accessory program of DOCK was used to pre-calculate scoring function potential grids (Meng EC et al., 1992). All parameters were set to default parameters, except for the "energy_cutoff_distance," which was set to 9999, resulting in the inclusion of all protein atoms in the energy calculation. For matching, the dms program was used to generate a molecular surface for each receptor (Richards FM, 1977). The SPHGEN accessory program of DOCK was

used to create a negative image of the surface using spheres (DesJarlais RL et al., 1988; Kuntz ID et al., 1982). For the purpose of this validation study, a general procedure was established to generate a sphere cluster for every protein in the test set. In this procedure, we select all the spheres found within 10 Å of any ligand atom. The receptor box delimiting the active side was calculated with the accessory program SHOWBOX using the sphere set with an additional 5 Å boundary. We have explored additional box sizes ranging from 1 Å – 9 Å padding and found that there is little sensitivity to the exact padding amount (ie success rate for rigid ligand docking of $80\% \pm 1\%$, time increase 10% with padding size increase, and an average test set energy of -50 ± 0.1 DOCK units). The final procedure creates sphere sets with an average of 101 docking spheres and boxes of $\sim 20 \text{ \AA}^3$. These receptor sphere sets are larger than what one would typically use in most docking applications. This adds stringency to our testing of DOCK 5 by increasing the orientational and translational space that it must search.

Optimized Hydrogen Locations for Test Set Receptors

To assess the effect of hydrogen placements on docking outcomes, we also optimized the hydrogen atom placement and hydrogen-bonding network for the receptor using the “Dock Prep” module in Chimera (Pettersen EF et al., 2004). In this module, the hybridization states of the nonhydrogen atoms of a PDB structure are determined by an enhanced version of the IDATM atom-typing algorithm (Meng EC and Lewis RA, 1991). Then, all hydrogens that can be

unambiguously positioned are added to the file. To assist in positioning ambiguous hydrogens, hydrogen-bonding interactions are examined. The definitions of hydrogen-bonding donors and acceptors as well as hydrogen-bonding angle and distance criteria are based on the values found in Mills et al (Mills JEJ and Dean PM, 1996). Relevant hydrogen bonds (H-bonds) are examined from shortest to longest, with satisfaction of shorter bonds having priority. For H-bonds where it is unclear which end is acting as the donor (e.g. water-water), use of that bond is postponed until either end is resolved further, though any lower-priority bonds that conflict geometrically with the postponed bond are eliminated from consideration at that time. If neither end is resolved by other interactions, the ambiguity is decided arbitrarily. Should examination of H-bond interactions not completely determine the positions of all of the hydrogens bound to a heavy atom, they are positioned to first satisfy potential H-bond interactions, then any remaining hydrogens are positioned to avoid steric clashes with other atoms. For histidine residues, normally one nitrogen will be protonated (chosen based on H-bond/ steric considerations); however if both ring nitrogens are H-bond donors, they will both be protonated.

Selection of Active Site Waters

All waters within 3 Å RMSD of any ligand heavy atom were selected. These waters were included as part of the receptor. The new receptor-water complexes were then subjected to the same hydrogen bonding optimization as above.

DOCK Parameter Optimization

To characterize the performance of DOCK 5 in regenerating known complex structures, we explored the optimum parameters for use with rigid and flexible ligand docking strategies (see Appendix). Unless otherwise stated, all docking experiments were carried out on 2.2 GHz dual processor Opteron 828s running Linux Fedora Core 3. The code was compiled using open-source GNU compilers (<http://www.gnu.org>). The optimized parameters have been implemented as the defaults. We note that our primary criterion for optimization was success in finding the proper ligand geometry and not the CPU time required per compound. Unless otherwise stated, these parameters were used for all experiments in this paper.

Greedy Clustering of Conformational Ensemble

The greedy clustering algorithm is designed to eliminate redundant ligand orientations from consideration. DOCK generates a set of ligand orientations that are ranked by the scoring function. The RMSD between each ligand orientation in the list is calculated. If the RMSD between two ligand orientations falls within the clustering threshold, the second orientation is assigned to a cluster with the first. The first ligand orientation is selected and compared to all subsequent unclustered orientations in the list; this process is repeated until the last unclustered orientation has been selected. Once the entire list has been processed, only the best scoring ligand pose in each cluster, designated as the cluster head, is retained.

Evaluation of MPI Functionality

Parallel Processing is fully integrated into the DOCK calculation. The DOCK program starts a single master node and a set of processing nodes. The master node performs file processing and molecule input/output, whereas the processing nodes perform the actual docking calculations. If the number of processors is set to 1, the code defaults to non-MPI behavior. As a result of this configuration, there will be minimal difference in performance between 1 and 2 processors. Improved performance will only become evident with more than 2 nodes. It should be emphasized that the primary benefit in using DOCK 5 in parallel mode is to reduce bookkeeping tasks associated with manually splitting up a database into multiple chunks, which then must be submitted to different processors individually. DOCK 5 automatically partitions out subsets of a database to various nodes, collates and ranks the final results, and takes care of all intermediate bookkeeping.

To gauge the performance of parallelization of the DOCK 5 algorithm, two small subsets of the NCI database from the ZINC database were constructed (Gropp W et al., 1996; Irwin JJ and Shoichet BK, 2005). The two subsets, one containing 500 and the other 1000 small molecules, were filtered to have ≤ 5 and ≤ 14 rotatable bonds, respectively. The receptor used as a target for this study was HIV-1 reverse transcriptase in complex with nevirapine (PDB code 1VRT). Because the receptor was not part of the test set, nevirapine was flexibly redocked using the optimized parameters, which yielded a ligand orientation 0.28 Å RMSD from the crystal structure orientation. In addition, a library consisting of

1000 copies of nevirapine was generated to remove dependence on the order and size of the compound library. All parallelization study calculations were executed at the Computational Science Center at Brookhaven National Laboratory (<http://www.bnl.gov/csc>) on a cluster consisting of 34 nodes with dual 3.2GHz Xeon processors running Linux. Tests were performed using between 2 and 68 nodes. The code was compiled using open-source GNU compilers and MPI software mpich version 1.2.7 from Argonne National Laboratory (<http://www-unix.mcs.anl.gov/mpi/mpich>).

RESULTS

We first consider the results of rigidly docking ligands, which used a conformation taken directly from the complex crystal structure, to the complex crystal structure conformation of the receptor. We then present the results of flexible ligand docking tests. In each case, we consider a) the overall performance of each sampling algorithm, b) the ability of each algorithm to reproduce the crystal ligand orientation as the top-scoring pose, c) the effect of the initial ligand conformation on the performance of the algorithm, d) any additional information contained in the set of all sampled ligand orientations, and e) the ability extract additional information by clustering docking results. We also compare the performance of DOCK 5 to equivalent DOCK 4 experiments. Finally, we analyze the cases in which DOCK 5 fails to reproduce the crystal structure and propose some directions for improvement of both the DOCK algorithm and our test set preparation method.

Rigid Ligand DOCKing

Overall Performance

Unless otherwise noted, all experiments described in this section involved rigid docking of the complex crystal structure ligand conformation to the receptor complex crystal structure. For each case in the test set, the heavy atom RMSD between the top-scoring docked ligand pose and the complex crystal structure ligand pose was evaluated. A DOCK 5 run was considered to be successful for cases in which the RMSD between for the top-scoring ligand orientation and the crystal ligand orientation was less than 2.0 Å. DOCK 5 selects the correct pose as the lowest energy structure for 79% (90/114) of the test cases using the rigid docking protocol with an average time of 55 seconds per complex.

Dependence on Ligand Conformation

An ensemble of ligand conformations was generated using the anchor-and-grow algorithm to apply changes of each of the ligand's rotatable bonds. This expansion generated a conformation ensemble for each ligand that covered all torsional parameters that DOCK samples. Each generated conformation was rigidly docked to the receptor, and the results from all the dockings were binned according to the magnitude of the ligand's conformational perturbation (Figure 3a). The curve shows dramatic and continual decrease in the success rate as the perturbation magnitude increases with little success for any ligand conformations greater than 0.5 Å heavy atom RMSD away from the crystal conformation. Therefore, any conformation generation method must generate

ligand conformations within 0.5 Å heavy atom RMSD of the crystal conformation for rigid docking to have a reasonable chance to succeed.

Analysis of Total Orientational Ensemble

To this point, we have disregarded “near misses,” which we define as any generated orientations within 2 Å RMSD from the crystal structure that are close to the top of the ranked conformation list, but are not the best scoring poses. We can examine the remaining poses either by including all poses that differ by a fixed energy unit from the most favorable geometry or by including those that differ by a fixed number of ranked poses from the most favorable energy. In order to quantify the extent of these partial successes, all generated ligand poses for each test case were preserved and sorted by their energy scores.

An energy gap is defined as the difference between the DOCK score of the top scoring ligand orientation and the score of a ligand ranked further down the list. Considering all docked ligand orientations with an energy gap of 2.5 DOCK units—an average of 5 ligand orientations—increases the rigid ligand docking success rate to 90% for the entire test set, while an average of 50 orientations increase the rigid docking success rate to 99% (Figure 4a and b). These results indicate that the orienting method samples near-crystal ligand orientations well, but the current energy scoring function cannot discriminate well between the top-ranked orientations.

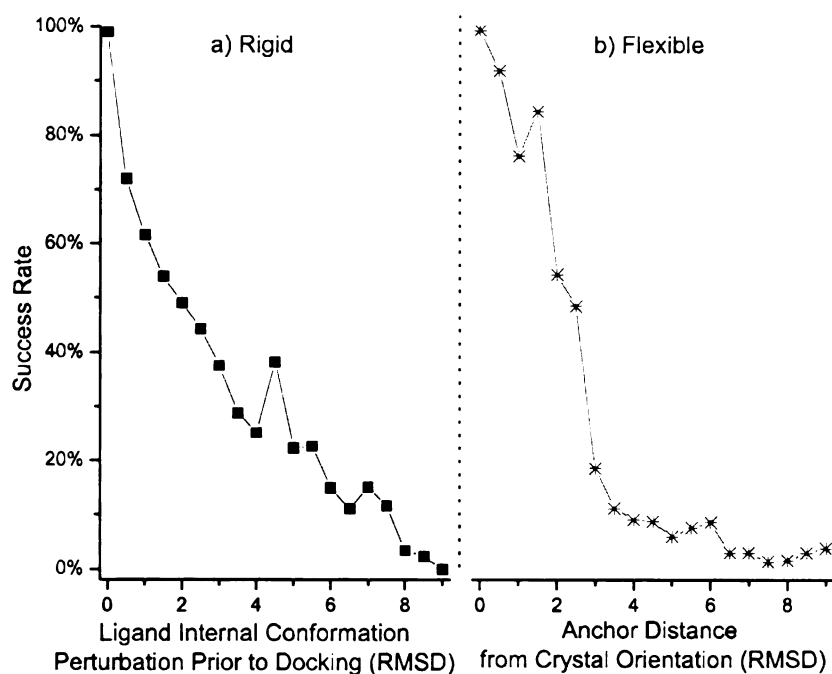


Figure 3. (a) Rigid docking success rates (■)--as calculated by any conformation being within 2 Å heavy atom RMSD of the complex crystal orientation--shown as a function of the ligand internal conformation perturbation magnitude (RMSD) (b) Flexible growth success rates (*)--as calculated by any conformation being within 2 Å heavy atom RMSD of the complex crystal orientation--shown as a function of the magnitude of the anchor perturbation (RMSD)

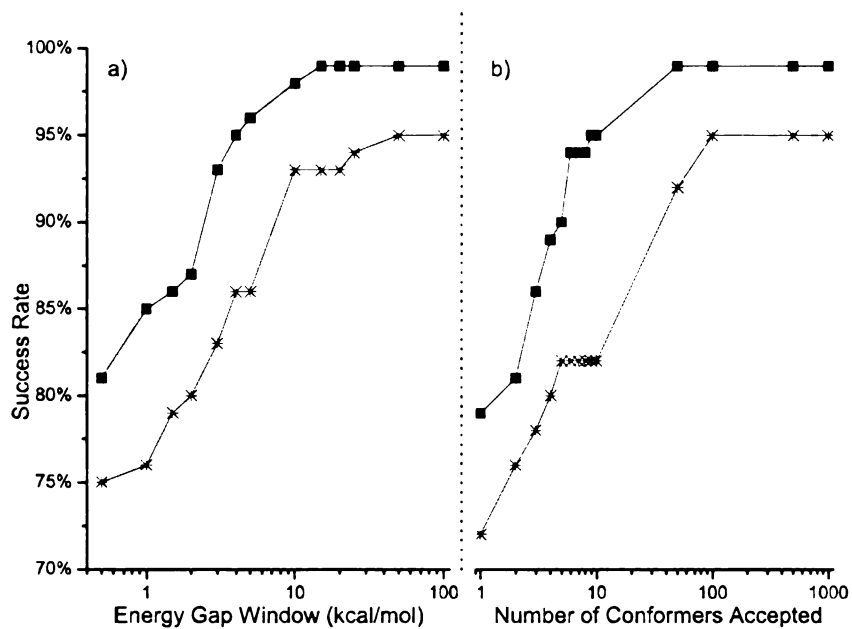


Figure 4. (a) The rigid (■) and flexible (*) docking success rate as a function of the DOCK score energy gap (kcal/mol) for all conformers generated. (b) The rigid and flexible docking success rate as a function of the number of ranked conformers examined

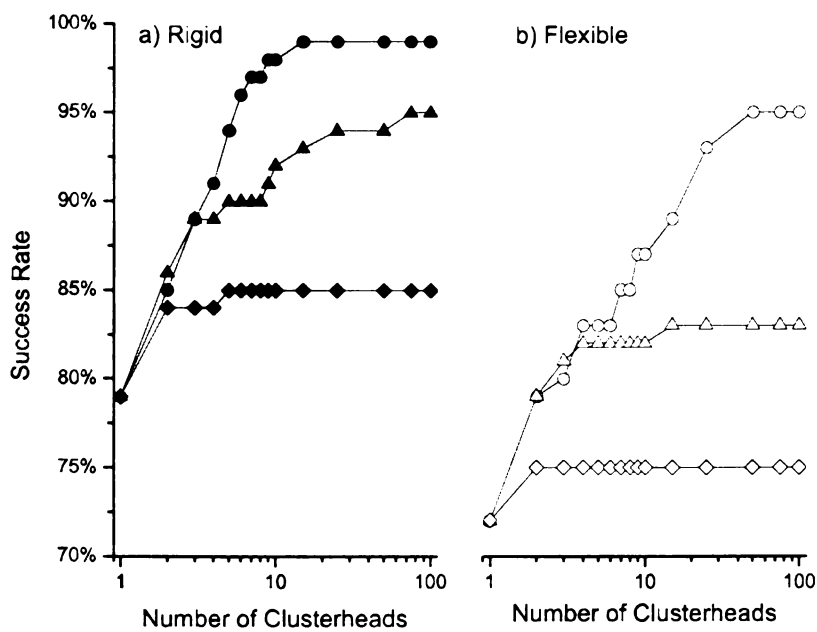


Figure 5. The rigid (filled) and flexible (open) docking success rate as a function of the number of cluster heads examined. Clusters with heavy atom RMSD cutoffs of 1.0 Å (●), 3.0 Å (▲), and 5.0 Å (◆) were compared.

Geometric Clustering of Poses

Each ligand conformational ensemble was spatially clustered according to inter-pose RMSD values (see Methods section for algorithm details). After examining a range of potential cut-offs, an optimal value of 1.0 Å was chosen (Figure 5). Using this clustering threshold, only 15 clusterheads are required to achieve a success rate of 99%, compared with the top 50 ranked unclustered orientations. This result is encouraging, suggesting that the clustering helps sort through the conformers efficiently.

Flexible Ligand DOCKing

Overall Performance

Unless otherwise noted, all experiments described in this section involved flexible docking of the ligand to the receptor complex crystal structure. As with the rigid docking tests, the heavy atom RMSD between the top-scoring docked ligand pose and the complex crystal structure ligand pose was evaluated for each complex in the test set. The success rate over the entire test set using the optimized flexible ligand anchor-and-grow protocol was 72% (82/114) with an average time of 314 seconds per complex.

Dependence on Anchor Position

The anchor-and-grow algorithm belongs to the set of incremental construction algorithms for searching ligand conformational space (Ewing TJA and Kuntz ID, 1997; Leach AR and Kuntz ID, 1992). It uses a rigid docking step

for the “anchors” to identify likely anchor positions (anchor orienting), and a torsion angle search step to generate ligand conformations rooted at the previously identified anchor positions (flexible growth). In order for flexible docking to succeed, both of these individual steps must be successful.

To measure the dependence of success rate on the precision of the anchor location, the crystal position of the anchor for each complex in the test set was perturbed randomly from 0 Å to more than 10 Å. Each perturbed anchor position was then considered as the starting point for flexible growth (Figure 3b). With the anchor starting less than 0.5 Å heavy atom RMSD from the crystal orientation, the growth algorithm can find the experimental orientation 99% of the time. However, the results demonstrate a rapid decrease in success rate as the anchor is moved further away from its crystal structure position, decreasing to 76% at 1.0 Å perturbation down to 54% at 2.0 Å. These data imply that if the flexible ligand docking algorithm can place the anchor within 0.5 Å heavy atom RMSD of the crystal anchor position, DOCK 5 has a very high probability of successfully predicting the full binding mode correctly.

Analysis of Total Conformational Ensemble

We examined the entire ensemble of conformers generated by flexible docking, as we described previously in the rigid ligand docking analysis. Considering all docked ligand conformations with a 2.5 DOCK unit energy gap—an average of 5 ligand orientations—increases the success rate to 82%, while an average of 100 orientations increasing the success rate to 95% (Figure 4a and

b). Again, these results indicate that the sampling density produced by the optimized parameters is quite high, but there is little discrimination between very similar poses by the current scoring function.

Geometric Clustering of Poses

As with the rigid ligand docking tests, each conformational ensemble was spatially clustered according to interpose RMSD (see Methods section for algorithm details). A clustering threshold of 1.0 Å, as determined in the rigid docking section, was used (Figure 5). Using this clustering threshold, only 50 clusterheads must be examined to reach a success rate of 95% as compared to 100 purely ranked orientations. Once again, this result is encouraging, as it requires a small number of ligand poses to be retained for rescoring with more advanced scoring functions that are better at discriminating between very similar ligand poses.

Comparison to DOCK 4

Using the optimized DOCK 5 parameters, we performed the same rigid and flexible ligand docking experiments on the entire test set using the last available version of DOCK 4. The performance of the current implementation of DOCK 5 compared favorably with the DOCK 4 performance (Table 4). We attribute the improved accuracy in performance to improvements outlined in the Methods Section. However, when comparing the speed of docking experiments between DOCK 4 and DOCK 5, DOCK 4 is five-fold faster for rigid docking and

thirty-fold faster for flexible ligand docking than DOCK 5 (Table 5). We attribute this increased calculation time to extra stages of minimization and sampling in DOCK 5, as well as additional overhead necessary to preserve the modularity of the code (see Methods).

Comparison to Other Docking Methods

Developers of Glide, GOLD and FlexX have also evaluated their methods using similar test sets and made some of their analyses available (Friesner RA et al., 2004; Hoffmann D et al., 1999; Wagoner J and Baker NA, 2004). Based on this data, we note that DOCK's flexible docking success rate of 70% is comparable to Glide's and FlexX's success rates of 82% and 61%, respectively (Table 6). Unfortunately, GOLD has not posted the results for the entire CCDC/Astex test set, so a complete comparison could not be made. However, for the subset of the test set they did report, DOCK's success rate of 67% is once again reasonable as compared to the success rate of 77% for GOLD, considering that the DOCK scoring function does not use either empirically weighted parameters or adjustable parameters.

Analysis of Successes and Failures of DOCKing Protocols

Docking failures can be categorized into two categories: sampling (soft) and scoring (Richards FM) failures (Verkhivker GM et al., 2000). For scoring failures, an orientation near the crystal structure was sampled in the course of the DOCK run, but the scoring function failed to rank it at the top of the

DOCK Version	Rigid Ligand	Flexible Ligand
4.0.1	71.9%	42.1%
5.4.0	79.0%	71.9%

Table 4: Success based on DOCK version (see Methods)

	Average (sec)	Minimum (sec)	Maximum (sec)
DOCK 4 Rigid Lig	10.9 ± 12.1	0.99	66.8
DOCK 4 Flexible Lig	7.1 ± 6.04	0.44	33.5
DOCK 5 Rigid Lig	55.4 ± 37.5	6.0	198.0
DOCK 5 Flexible Lig	314.7 ± 449.8	2.0	2638.0

Table 5: Average length of time in seconds for docking calculation using the optimized parameter set (see Appendix)

Program	# Complexes	Success	DOCK Success
GOLD	43	77%	67%
Glide	71	82%	70%
FlexX	71	61%	70%

Table 6: Comparison of DOCK success rates to other docking programs for flexible ligand docking

list. A sampling failure indicates that the DOCK run failed to sample any orientations within 2 Å RMSD of the crystal structure. The major caveat of this classification scheme is the assumption that the model of both the receptor and the ligand, including the VDW parameters, electrostatics, and hydrogen orientations and protonation states, reflect those that occur in the experimental structure (Kuntz ID and Agard DA, 2003). Here, we analyze the flexible docking ligand failures within the sampling-scoring classification scheme.

Failures Resulting from Receptor Modeling/Structural Problems

The original CCDC/Astex test set was filtered for experimental errors using a variety of metrics (Nissink JW et al., 2002). We plotted the flexible ligand success rate as a function of various metrics of the quality of the x-ray structures to determine if the selection criteria were appropriate for testing the DOCK algorithm (Figure 6). There appears to be at best a weak correlation between the RMSD of the best scoring DOCK pose and either crystal resolution or b-factor of active site or backbone atoms, indicating that the cut-offs chosen for the original set were reasonable for docking purposes.

We next explored whether specific atom types caused problems with the DOCK force field terms by correlating the test set success rate with the presence and type of active site cofactor (Table 7). The only clear problem involved metal ions in the receptor. These structures showed a much lower success rate, accounting for nearly half of both the rigid and flexible ligand docking failures. However, there still are a number of failures in the portion of the test set without

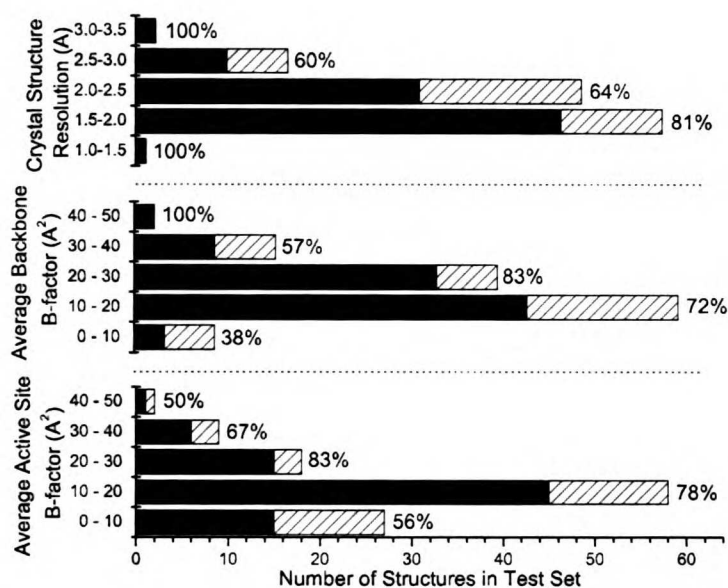


Figure 6. Correlation of flexible ligand success (filled) and failure (striped) rates with crystallographic resolution (Å) and experimental B-factor (Å²). For active site B-factors, the active site was defined as any atom within 9 Å of the experimental ligand orientation.

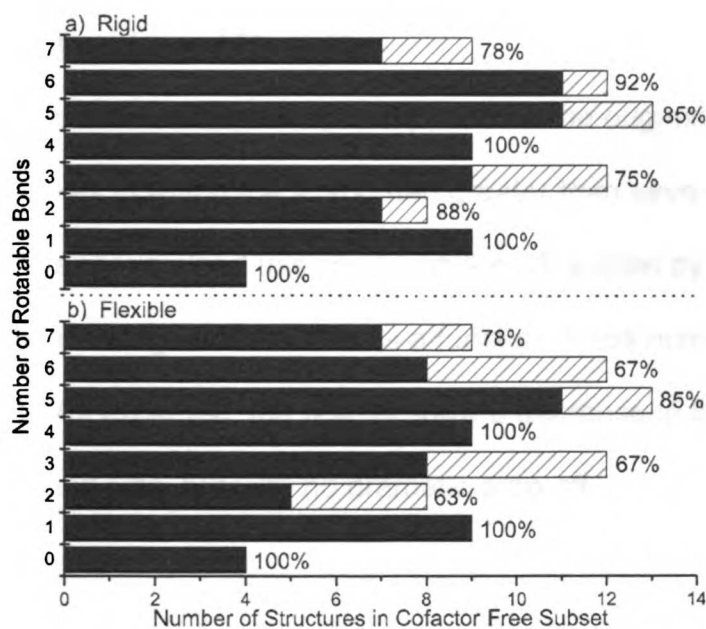


Figure 7. Rigid and flexible docking success (filled) and failure (striped) rates as a function of the number of rotatable bonds in each ligand in the CF test set

cofactors in the active site that require further characterization. Unless otherwise mentioned, all studies below were performed on this subset, referred to as the Cofactor Free (CF) subset.

For all members of the test set, the experimental resolution of the crystal structures was too poor to identify hydrogen atom locations. We originally modeled the hydrogen atom positions using a rule-based method. To test this scheme, we applied a more advanced hydrogen addition procedure that accounted for steric clashes and hydrogen-bonding networks to the CF subset (see Methods). As a follow-up, we assumed all crystallographically bound waters found within 3 Å of any ligand heavy atom were critical for binding and included them in the receptor model as well. We found that both of these procedures improved the flexible ligand docking success rate (Table 8).

Failures Resulting from Ligand Flexibility

In addition to the selection criteria imposed on the original test set, we also filtered out complexes in which the ligand had greater than seven rotatable bonds (see Methods). We reexamined this choice on the CF subset by plotting the rigid and flexible ligand docking success rate as a function of the number of flexible bonds (Figure 7). As expected, the results show a decrease in the success rate with increasing ligand size, but with no dramatic drop-off.

	Total Count	Rigid Lig Success	Flexible Lig Success
Entire Test Set	114	79.0%	71.9%
CF Subset	76	81.6%	76.3%
Active Site Cofactor	38	73.7%	63.2%
Active Site Metal Cofactor	28	64.3%	50.0%

Table 7: Flexible ligand success as function of active site cofactor

Test Set Preparation Technique	Success
Standard	76.3%
Hydrogen Optimization	78.9%
Active Site Waters + Hydrogen Optimization	80.3%

Table 8: Flexible ligand success as function of CF test set preparation (total of 76 complexes)

	Rigid Sampling Failure	Rigid Scoring Failure	Rigid Success
Flexible Sampling Failure	0	0	2
Flexible Scoring Failure	0	9	3
Flexible Success	0	1	23

Table 9: Comparison of success and failure cases of both rigid and flexible docking for complexes in test set with cofactors in active site (total of 36 complexes)

	Rigid Sampling Failure	Rigid Scoring Failure	Rigid Success
Flexible Sampling Failure	1	1	2
Flexible Scoring Failure	0	7	7
Flexible Success	0	5	53

Table 10: Comparison of success and failure cases of both rigid and flexible docking for complexes in CF subset (total of 76 complexes)

Sampling Versus Scoring Failures

We now return to classification of DOCK failures based on scoring and sampling classifications (Verkhivker GM et al., 2000). First, we examined the test set failure cases with active site cofactors (Table 9). Within this set, nine examples were scoring failures for both rigid and flexible ligand docking, indicating that new VDW and electrostatic parameters need to be developed for magnesium, heme groups, and some coordination states of zinc. In addition, there were three flexible ligand scoring failures that were rigid successes, thus suggesting that the flexible algorithm was able to identify additional orientations with better scores than the experimental ligand orientation. Only two flexible ligand docking cases were sampling failures. We expected flexible ligand docking sampling failures due to the increased ligand degrees of freedom compared with rigid ligand docking, but it does not appear to be a severe problem in this test set containing ligands with less than eight rotatable bonds. Finally, one of the rigid ligand docking scoring failures was a flexible ligand success. In this case, there was a large VDW clash between one of the ligand atoms and the receptor. The anchor-and-grow algorithm was able to build the ligand in the active site to avoid this clash which the rigid ligand docking algorithm could not accommodate.

We repeated this analysis with the CF subset (Table 10). Here, there was one rigid ligand docking sampling failure, which also failed for flexible ligand docking. Upon closer examination of the receptor site, a residue making critical interactions with the ligand was not resolved in the experimental complex

structure (PDB code 1A6W). We anticipate that there may not be enough contacts to correctly place the molecule. Seven examples were scoring failures for both rigid and flexible ligand docking. In this subset, though, we cannot attribute the failure to unusual atom types, indicating that the scoring function is incorrectly modeling some portion of the energy landscape. There were also seven scoring failures for flexible ligand docking that were successes for rigid ligand docking, once again suggesting that the flexible docking algorithm identified additional orientations that scored better than the experimental orientation.

As in the cofactor set above, there were only three additional flexible ligand docking sampling failures. One of these was also a scoring failure in rigid ligand docking, implying that this failure case may actually be due to a combination of both sampling and scoring factors. The remaining two flexible ligand docking sampling failures once again indicate that the flexible algorithm was able to identify alternative orientations that scored better than the crystal complex orientation. Finally, five rigid ligand docking scoring failures were flexible ligand dockings successes, signifying that the flexible ligand docking algorithm is able to compensate for intermolecular clashes in the active site of the experimental structure that the rigid ligand algorithm simple cannot accommodate (data not shown).

Analysis of DOCK Score for DOCKing Protocols

To analyze the ability of DOCK to reproduce the ligand-receptor interaction energy as measured by the DOCK scoring function, we plotted the score from the top-ranking pose for both rigid and flexible ligand docking that were successful against the DOCK score of the complex crystal structure (Figure 8 a-b). Each crystal structure ligand was minimized with 1000 steps of the DOCK simplex minimizer. The significant feature of both plots is that the docked pose generally scores more favorably than the minimized crystal structure. When rigid ligand docking is compared with flexible ligand docking, the flexibly docked ligand conformations almost always have a lower score (Figure 8c). These results indicate that increasing the amount of ligand orientational and conformational sampling increasingly identifies deeper wells in the binding energy landscape. When we plotted the flexible ligand success rate against the minimized crystal score, there was little correlation, though DOCK was observed to perform better using crystal structures with scores more negative than -20 DOCK units (Figure 8d). This lack of correlation indicates that, while having a negative interaction energy for the crystal structure will increase the probability of DOCK finding the correct binding orientation, this metric is not a good predictive indicator of DOCKing success.

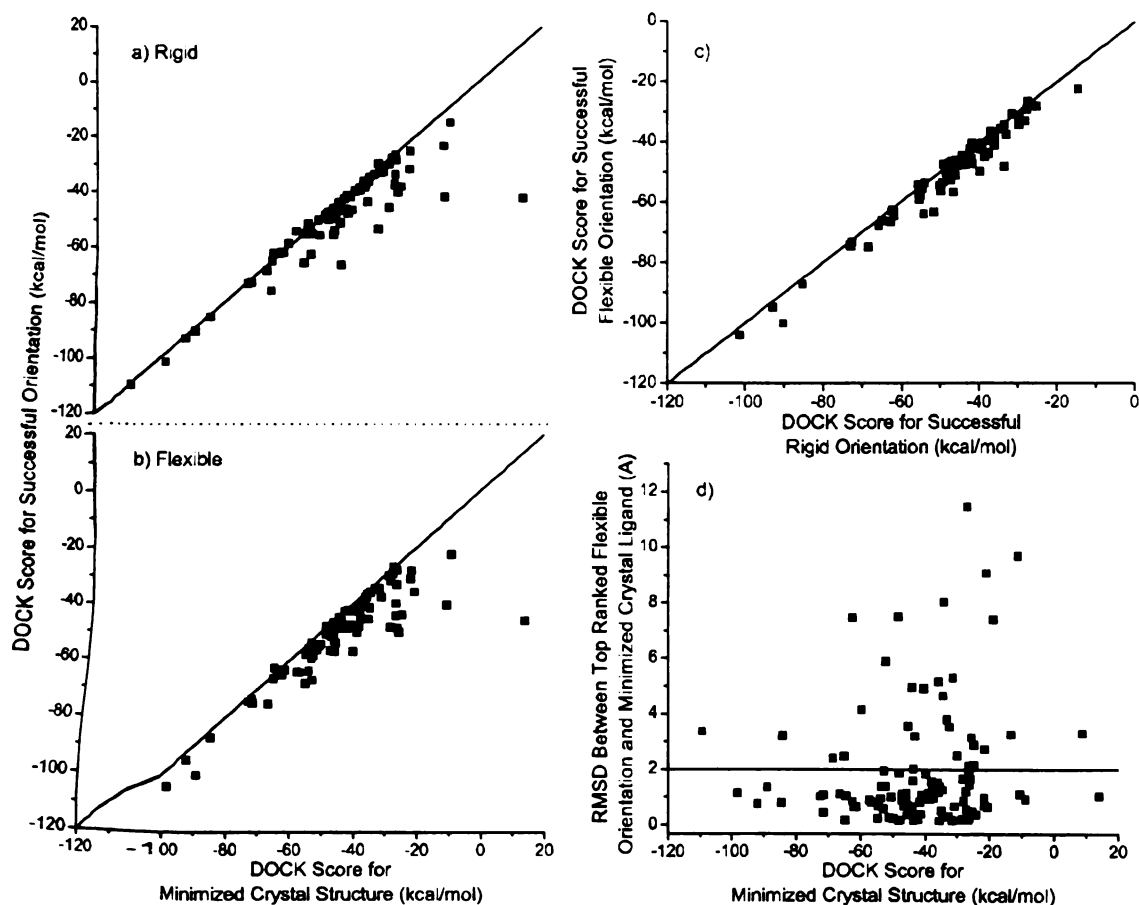


Figure 8. (a) Successful rigid ligand docking scores (kcal/mol) as a function of minimized crystal structure ligand scores (kcal/mol) (b) Successful flexible ligand docking scores (kcal/mol) as a function of minimized crystal structure ligand scores (kcal/mol) (c) Successful flexible ligand docking energy scores (kcal/mol) as a function of successful rigid ligand docking energy scores (kcal/mol) (d) Comparison of the RMSD between all top ranked flexible ligand orientations and the minimized crystal ligand orientations to the minimized crystal interaction energy as measured by the DOCK score (kcal/mol)

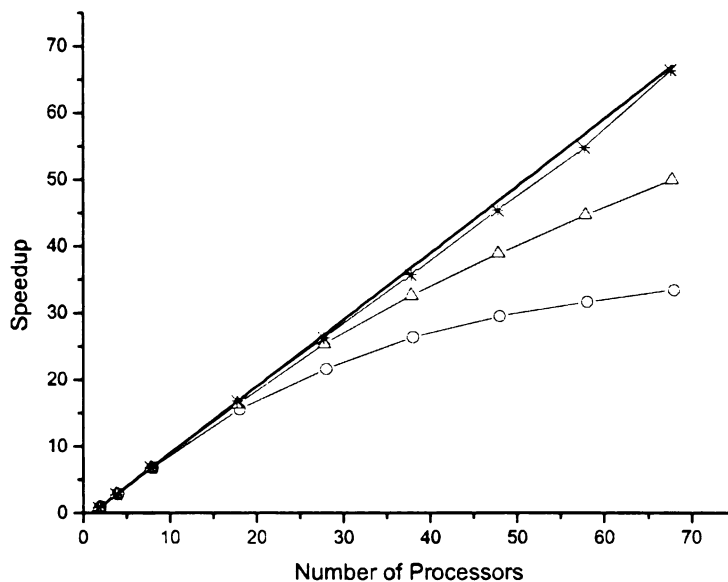


Figure 9. Speedup (calculated as length of time for calculation on a single processor / length of time for calculation on n processors) for docking a library of 500 different small molecules (○), 1000 different small molecules (Δ), and 1000 copies of nevaripine (*) using flexible ligand docking as a function of the number of processors in MPI mode. A perfectly parallel calculation (-) is plotted for comparison.

Database DOCKing using MPI

Substantial speedup is observed for up to about 14 processors for the 500 compound library and 18 processors for the 1000 compound library (Figure 9). Interestingly, the library with 1000 copies of neviripine shows almost perfectly parallel behavior up to 68 processors. We hypothesize that the speedup for the heterogeneous libraries will continue to approach ideal as larger libraries with increased numbers of rotatable bonds are used, but will never be completely linear due to overhead from input and output and lag resulting from communication between the nodes.

DISCUSSION

In this paper we have described a new version of the DOCK program. Our main purpose was to develop modular code that was straightforward to modify and which showed improved performance over the old version. By using an object-oriented language for DOCK 5, we were able to accomplish this goal, and we demonstrate, here, how routines such as the simplex minimizer and the clustering algorithm can be added or replaced without changes in other parts of the program. The successful parallelization of the calculation and the addition of post-processing clustering were simple but useful modifications to the algorithm, which encourages further investigations and algorithm experimentation.

The performance of DOCK 5 on a curated test set of 114 protein-ligand complexes proved to be superior to DOCK 4, with an over-all success rate of 79% for rigid ligand docking and 72% for flexible ligand docking, compared with

72% and 42% respectively for DOCK 4. We ascribe the improvements to significant changes in the flexible search sampling and pruning procedures and to code corrections. The difference in performance of DOCK 5 for rigid and flexible docking is relatively modest (79% vs 72%) even though the search for flexible ligands includes both configurational and conformational spaces. Using the receptor structure to prune the conformational search tree is clearly a reasonably efficient procedure. Although, the DOCK 5 code takes longer on average to run a calculation than DOCK 4, we feel this drawback is balanced by the improved results and the modularity of DOCK 5. Efforts to increase throughput are underway.

We also wish to stress the importance of having a high quality test set for *evaluation* of docking programs. Xray crystallography typically provides essential *but incomplete* data for the calculations we wish to carry out. For example, in the majority of cases, hydrogen positions must be determined. In other cases, *critical water* molecules must be placed and some residues need to be modeled where *experimental* data is lacking. The ligand conformations may also contain *significant* uncertainties. Finally, we must be aware of the inherent assumptions *underlying* the force field parameters used in the molecular modeling steps. All of these *considerations* speak to the need for careful inspection of test set *complexes*. Our results demonstrate this issue: the success rate for *reconstitution* of the complex geometries was shown to depend on the nature of the *cofactors*, the optimization of hydrogen placements, and the inclusion of *critical waters*.

The primary result that emerges from the analysis of the docking failures is that the current force field requires improvement, particularly in the treatment of metal-containing cofactors. We also note that binding conformations and configurations are determined by the free energy of the system while we are only, at best, estimating the enthalpy. Finally, we do identify a few situations in flexible ligand docking where the conformational sampling is insufficient. A test set with ligands containing more than 7 rotatable bonds would, presumably, show an increase in these sampling failures. We hypothesize that the key weakness is the pruning algorithm, which we will explore in future studies.

What are the routes to improvement? An obvious starting point is the use of more accurate methods for preparing experimental structures, including tools for accurate pKa prediction and *de novo* identification of critical waters. For the docking calculation itself, it would be helpful to improve VDW and electrostatic parameters for all atoms heavier than oxygen, particularly for metal atoms. Ideally, one would directly include charge polarization and ligation geometry in the force field. In addition, modifications to the force field to better approximate the free energy – e.g. generalized Born or Poisson Boltzmann implicit solvation electrostatics with surface area corrections to account for the hydrophobic effect—would also improve modeling accuracy. The DOCK 5 platform is positioned to enable future developments and work is underway to incorporate them into future releases.

CONCLUSIONS

In this study, we have evaluated a new version of DOCK. We have found that it predicts binding geometries of a structurally diverse test set comparably to similar algorithms and better than the previous version of DOCK.

Simultaneously, we have thoroughly explored the sampling portions of the algorithm and found that the majority of binding pose prediction failures is a result of scoring function deficiencies. In further exploration of these failures, we have determined that the docking success seems to be a function of whether there are alternative orientations that score well—as defined by the scoring function—rather than the interaction energy of the experimental structure itself. Finally, we have implemented new functionalities and shown that they improve the success rates of both rigid and flexible ligand docking. In general, we have a new tool that not only performs well on a typical test set but is an ideal tool to explore any number of new algorithms in the context of the molecular docking problem.

ACKNOWLEDGMENTS

Gratitude is expressed to Dr. Bentley Strockbine and Sudipto Mukherjee for computational assistance with MPI calculations. Demetri Moustakas, Natasja Brooijmans, P. Therese Lang and Irwin D. Kuntz would like to thank the NIH grant GM 56531 (Paul Ortiz de Montellano, PI) for support. P. Therese Lang would also like to thank the Burroughs Welcome Foundation and the American Foundation for Pharmaceutical Education for additional support. The authors

would like to thank Scott Brozell, Mathew Jacobson, and Brian Shoichet and members of his group for helpful conversations.

APPENDIX

Rigid DOCKing Parameter Optimization

The parameters listed in Table 11 control the sampling of ligand poses within the receptor active site during rigid ligand docking. The parameters that control the step sizes for the simplex minimizer (`simplex_trans_step`, `simplex_rot_step`, and `simplex_tors_step`) were optimized in a previous study and were held at those values (Ewing TJA and Kuntz ID, 1997; Gschwend DA and Kuntz ID, 1996). For the remaining parameters—the number of orientations (`max_orientations`) and the number of minimization steps (`simplex_final_max_iterations`)—a series of rigid ligand docking experiments were performed to optimize the DOCK score for the top ranking pose averaged over the entire test set and the success rate, defined as the orientation of the top ranking pose being within 2 Å heavy atom RMSD from the crystal ligand. The success rate and DOCK scores initially improved as the number of orientations and the amount of minimization increased and then converged (Figure 10). We selected the lowest converged values—1000 orientations and 1000 minimization steps—as optimal.

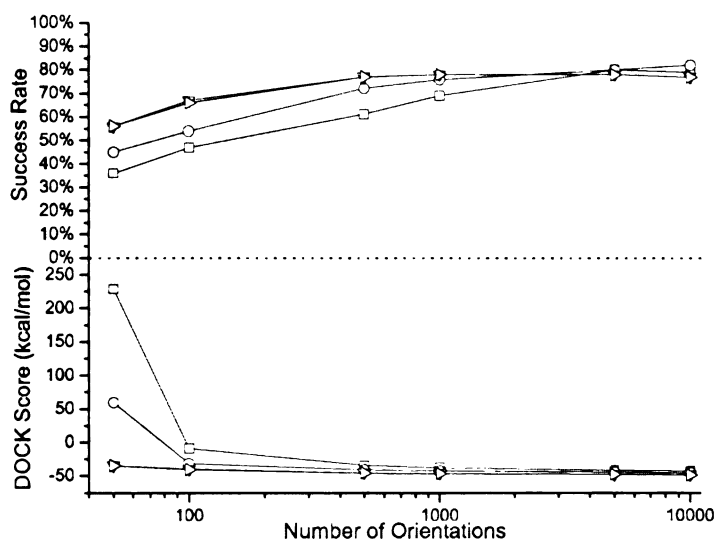


Figure 10. Optimization of parameters for rigid ligand docking. Parameters of 50 (\square), 100 (\circ), 1000 (∇), and 10000 (\triangleright) minimization steps (*simplex_final_max_iterations*) are examined as a function of the number of orientations (*max_orientations*).

Parameter Name	Parameter Description	Value
<i>max_orientations</i>	The number of ligand poses sampled by the rigid orienting algorithm	1000
<i>simplex_score_converge</i>	The score threshold used to determine simplex convergence	0.1
<i>simplex_trans_step</i>	The maximum initial translation step size for the simplex minimizer	1.0 Å
<i>simplex_rot_step</i>	The maximum initial rotational euler angle step size for the simplex minimizer	0.1 radian
<i>simplex_tors_step</i>	The maximum initial dihedral angle step size	10°
<i>simplex_final_max_iterations</i>	The maximum number of simplex iterations	1000

Table 11: Description of and optimized default values for parameters that affect rigid ligand docking

Flexible DOCKing Parameter Optimization

For the more complex flexible ligand algorithm, the parameter optimization was performed first on the anchor docking, and the best parameters were then used for optimizing the growth. The parameters that control the sampling in both these steps are listed in Table 12. As for rigid ligand docking, the parameters that control step sizes for the simplex minimizer were set to the previously defined optimal values.

The first step in the anchor-and-grow algorithm is ring identification or anchor segmentation. All bonds within molecular rings are treated as rigid. This classification scheme is a first-order approximation of molecular flexibility, since some amount of flexibility can exist in non-aromatic rings. To treat such phenomena as sugar puckering and chair-boat hexane conformations, the user needs to supply each ring conformation as a separate input molecule. If the molecule does not have a ring, the largest rigid segment is specified as the anchor. Additional bonds may be specified as rigid by the user. For simplicity, all runs in this study used the default of largest anchor only. If the molecule had multiple anchors of the same size, the first anchor on the anchor list was used. Once the anchor had been identified, the parameters that control the number of anchor orientations (`max_orientations`), the number of anchor minimization steps (`simplex_anchor_max_iterations`), and the cutoff for the anchor pruning (`num_confs_for_next_growth`) were explored. Because the anchors are substructures of the ligand, the parameter convergence was monitored as a function of the RMSD between the anchor orientation and the corresponding

substructure of the crystal ligand averaged over all generated orientations before the pruning function. When the number of anchor orientations and minimization steps were varied systematically, the number of minimization steps converged at 500 (Figure 11). We expected this optimized value to be lower than rigid docking because anchors are typically smaller than the final ligand.

Because the anchor orientations are pruned before the growth step, we used the optimized number of minimization steps while exploring the number of anchor orientations and the pruning cutoff. The optimal anchor pruning cutoff of 100 was chosen as a balance between convergence and the length of the calculation, which remained fixed for the final exploration of the number of orientations. The optimal number of orientations was selected to be 500 because the combination of these three variables generated the highest number of anchors near the crystal structure (Figure 11a). Note that if the number of orientations was increased beyond the selected value, the number of anchors near the crystal structure dropped dramatically. We hypothesized that this resulted from a combination of increased sampling and pruning. The pruning function was designed to identify a representative orientation from each energy well that the matching algorithm finds (see Introduction: DOCK Background). As sampling increased, the ranked orientations began to converge toward the bottom of the deepest energy wells, sampling less of the alternative high energy wells. Because the pruning function is designed to supply the most diverse ligands, fewer orientations made it through the pruning step as the sampling is

increased. We felt that this effect was reducing the potential sampling for the algorithm and plan to explore alternatives in future studies.

The next step in the anchor-and-grow algorithm is flexible bond identification. Each flexible bond is associated with a label defined in an editable file. The parameter file is identified with the `flex_definition_file` parameter. Each label in the file contains a definition based on the atom types and chemical environment of the bonded atoms. Typically, bonds with some degree of double bond character are excluded from minimization so that planarity is preserved. Each label is also associated with a set of preferred torsion positions. The location of each flexible bond is used to partition the molecule into rigid segments. A segment is the largest local set of atoms that contains only non-flexible bonds.

Using the optimal anchor parameters, we varied number of minimization steps for each layer of growth (`simplex_grow_max_iterations`) and the cutoff of number of conformers for the growth pruning function (`num_confs_for_next_growth`). Because the dock run now creates a complete pose, we return to using a combination of the score for the top ranking pose averaged over the entire test set and the success rate to monitor convergence. As with rigid ligand docking, the success rate improves modestly with improved sampling and eventually converges (Figure 11). However, although DOCK scores improved as the number of orientations and the amount of minimization increased, the values do not converge. We once again attribute this phenomenon to the pruning function. Therefore, we used the success rate to

select the lowest converged values—500 minimization steps and the cutoff for the number of conformers for the growth section as 100—as optimal.

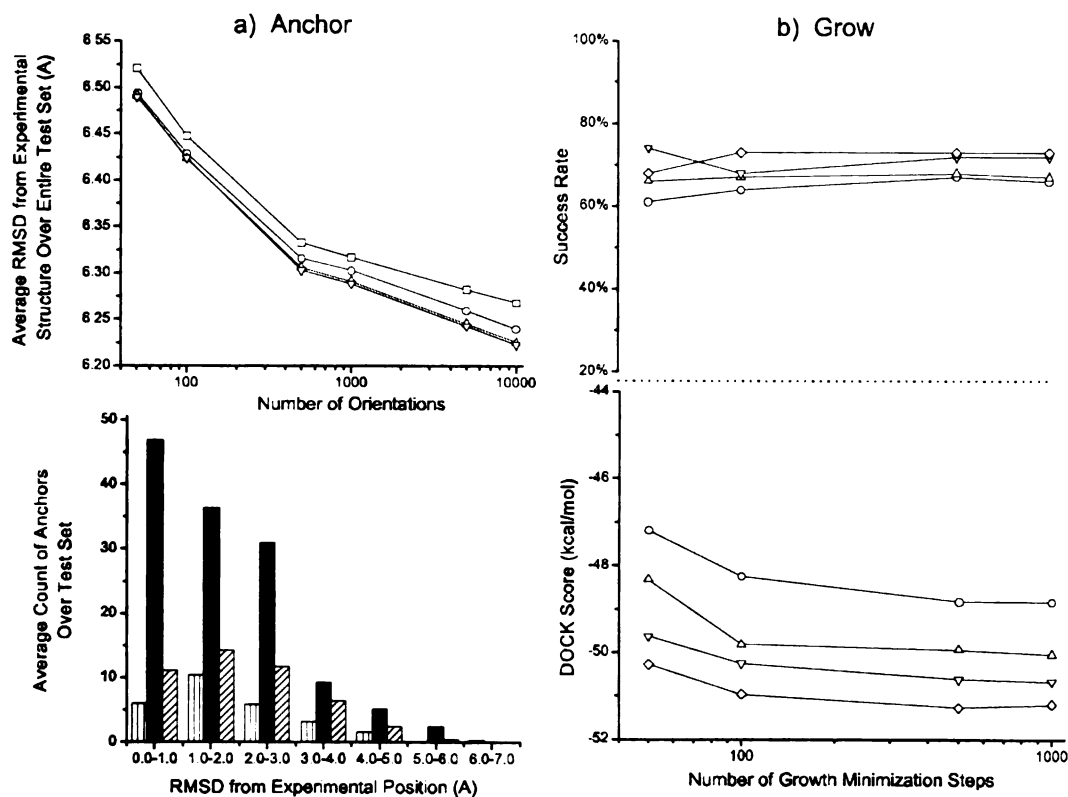


Figure 11. Optimization of parameters for flexible ligand docking. (a) Parameter optimization for anchor sampling portion of flexible ligand docking. TOP: Parameters of 0 (□), 50 (○), 100 (△), and 500 (▽) anchor minimization steps (simplex_anchor_max_iterations) are plotted as a function of the number of orientations (max_orientations). BOTTOM: Parameters of 50 (vertical stripes), 500 (filled), and 5000 (diagonal stripes) anchor orientations (max_orientations) are compared using an anchor pruning cutoff (num_confs_for_next_growth) of 100. (b) Parameter optimization for growth sampling portion of flexible ligand docking. Growth pruning cutoffs (num_confs_for_next_growth) of 25 (○), 50 (△), 100 (▽), and 200 (◇) are plotted as a function of the number of growth minimization steps (simplex_grow_max_iterations).

Parameter Name	Parameter Description	Value
max_orientations	The number of anchor poses sampled by the rigid orienting algorithm	500
num_anchor_orients_for_growth	The maximum number of anchor orientations promoted to the conformational search	100
num_confs_for_next_growth	The number of partially grown ligand conformers stored at each stage of the flexible growth procedure	100
simplex_anchor_max_iterations	The maximum number of simplex iterations applied to the ligand anchor during anchor docking	500
simplex_grow_max_iterations	The maximum number of simplex iterations applied to the ligand during the flexible growth procedure	500
simplex_final_max_iterations	The maximum number of simplex iterations applied to the complete ligand by the secondary scoring function	0

Table 12: Description of and optimized default values for parameters that affect flexible ligand docking

REFERENCES

- Alvarez, JC. *Curr Opin Chem Biol.* **8** (2004) 365-70.
- Aqvist, J and Warshel, A. *J Am Chem Soc.* **112** (1990) 2860-2868.
- Berman, HM, Westbrook, J, Feng, Z, Gilliland, G, Bhat, TN, Weissig, H, Shindyalov, IN and Bourne, PE. *Nucleic Acids Res.* **28** (2000) 235-42.
- Brooijmans, N. *Theoretical studies of molecular recognition.* (2003) Graduate Department of Chemistry and Chemical Biology, University of California, San Francisco, San Francisco, CA.
- Case, DA, Darden, TA, Cheatham, I, T.E. , Simmerling, CL, Want, J, Duke, RE, Luo, R, Merz, KM, Wang, B, Pearlman, DA, Crowley, M, Brozell, S, Tsui, V, Gohlke, H, Mongan, J, Hornak, V, Cui, G, Beroza, P, Schafmeister, C, Caldwell, JW, Ross, WS and Kollman, PA. "AMBER 8." (2004). University of California, San Francisco.
- Congreve, M, Murray, CW and Blundell, TL. *Drug Discovery Today.* **10** (2005) 895-907.
- Cornell, WD, Cieplak, P, Bayly, CI, Gould, IR, Merz, KM, Ferguson, DM, Spellmeyer, DC, Fox, T, Caldwell, JW and Kollman, PA. *J Am Chem Soc.* **117** (1995) 5179-5197.
- DesJarlais, RL, Sheridan, RP, Seibel, GL, Dixon, JS, Kuntz, ID and Venkataraghavan, R. *J Med Chem.* **31** (1988) 722-9.
- Ewing, TJA and Kuntz, ID. *J Comput Chem.* **18** (1997) 1175-1189.
- Friesner, RA, Banks, JL, Murphy, RB, Halgren, TA, Klicic, JJ, Mainz, DT, Repasky, MP, Knoll, EH, Shelley, M, Perry, JK, Shaw, DE, Francis, P and Shenkin, PS. *J Med Chem.* **47** (2004) 1739-1749.
- Gasteiger, J and Marsili, M. *Tetrahedron.* **36** (1980) 3219-3228.
- Gropp, W, Lusk, E, Doss, N and Skjellum, A. *Parallel Computing.* **22** (1996) 789-828.
- Gschwend, DA and Kuntz, ID. *J Comput-Aided Mol Des.* **10** (1996) 123-132.
- Halgren, TA, Murphy, RB, Friesner, RA, Beard, HS, Frye, LL, Pollard, WT and Banks, JL. *J Med Chem.* **47** (2004) 1750-1759.
- Hann, MM and Oprea, TI. *Curr Opin Chem Biol.* **8** (2004) 255-63.

- Hillisch, A, Pineda, LF and Hilgenfeld, R. *Drug Discovery Today*. **9** (2004) 659-669.
- Hoffmann, D, Kramer, B, Washio, T, Steinmetzer, T, Rarey, M and Lengauer, T. *J Med Chem*. **42** (1999) 4422-4433.
- Irwin, JJ and Shoichet, BK. *J Chem Inf Model*. **45** (2005) 177-182.
- Jakalian, A, Bush, BL, Jack, DB and Bayly, CI. *J Comp Chem*. **21** (2000) 132-146.
- Jones, G, Willett, P, Glen, RC, Leach, AR and Taylor, R. *J Mol Biol*. **267** (1997) 727-48.
- Kitchen, DB, Decornez, H, Furr, JR and Bajorath, J. *Nat Rev Drug Discovery*. **3** (2004) 935-949.
- Kopec, KK, Bozyczko-Coyne, D and Williams, M. *Biochem Pharmacol*. **69** (2005) 1133-1139.
- Kraljevic, S, Stambrook, PJ and Pavelic, K. *EMBO Rep*. **5** (2004) 837-42.
- Kramer, B, Rarey, M and Lengauer, T. *Proteins*. **37** (1999) 228-41.
- Kuhl, FS, Crippen, GM and Friesen, DK. *J Comput Chem*. **5** (1984) 24-34.
- Kuntz, ID and Agard, DA. *Protein Simulations*. **66** (2003) 1-25.
- Kuntz, ID, Blaney, JM, Oatley, SJ, Langridge, R and Ferrin, TE. *J Mol Biol*. **161** (1982) 269-288.
- Leach, AR and Kuntz, ID. *J Comput Chem*. **13** (1992) 730-748.
- Lischner, R. C++ in a nutshell. 1st, Sebastopol, CA.: O'Reilly Media, Inc. (2003)
- Lorber, DM and Shoichet, BK. *Curr Top Med Chem*. **5** (2005) 739-749.
- Meng, EC and Lewis, RA. *J Comput Chem*. **12** (1991) 891-898.
- Meng, EC, Shoichet, BK and Kuntz, ID. *J Comput Chem*. **13** (1992) 505-524.
- Merz, KM, Murcko, MA and Kollman, PA. *J Am Chem Soc*. **113** (1991) 4484-4490.
- Mills, JEJ and Dean, PM. *J Comput-Aided Mol Des*. **10** (1996) 607-622.

- Nelder, JA and Mead, R. *Computer Journal*. **7** (1965) 308-313.
- Nissink, JW, Murray, C, Hartshorn, M, Verdonk, ML, Cole, JC and Taylor, R. *Proteins*. **49** (2002) 457-71.
- Oprea, TI. *J Comput-Aided Mol Des*. **16** (2002) 325-334.
- Oprea, TI, Davis, AM, Teague, SJ and Leeson, PD. *J Chem Inf Model*. **41** (2001) 1308-1315.
- Pang, YP, Perola, E, Xu, K and Prendergast, FG. *J Comput Chem*. **22** (2001) 1750-1771.
- Perola, E, Walters, WP and Charifson, PS. *Proteins*. **56** (2004) 235-49.
- Petterson, EF, Goddard, TD, Huang, CC, Couch, GS, Greenblatt, DM, Meng, EC and Ferrin, TE. *J Comput Chem*. **25** (2004) 1605-1612.
- Posner, BA. *Curr Opin Drug Discovery Dev*. **8** (2005) 487-494.
- Purcell, WP and Singer, JA. *J Chem Eng Data*. **12** (1967) 235-246.
- Richards, FM. *Ann Rev of Biophys Bioeng*. **6** (1977) 151-176.
- Schnecke, V and Bostrom, J. *Drug Discovery Today*. **11** (2006) 43-50.
- Shoichet, BK, Bodian, DL and Kuntz, ID. *J Comput Chem*. **13** (1992) 380-397.
- Smith, GM, Alexander, RS, Christianson, DW, McKeever, BM, Ponticello, GS, Springer, JP, Randall, WC, Baldwin, JJ and Habecker, CN. *Protein Sci*. **3** (1994) 118-25.
- Verdonk, ML, Cole, JC, Hartshorn, MJ, Murray, CW and Taylor, RD. *Proteins*. **52** (2003) 609-623.
- Verkhivker, GM, Bouzida, D, Gehlhaar, DK, Rejto, PA, Arthurs, S, Colson, AB, Freer, ST, Larson, V, Luty, BA, Marrone, T and Rose, PW. *J Comput-Aided Mol Des*. **14** (2000) 731-751.
- Wagoner, J and Baker, NA. *J Comput Chem*. **25** (2004) 1623-1629.

“We are in the ordinary position of scientists of having to be content with piecemeal improvements: we can make several things clearer, but we cannot make anything clear.”

--Frank Plumpton Ramsay

Chapter 4

Synthesis and Testing of a Focused Phenothiazine Library for Binding to HIV-1 TAR RNA

Moriz Mayer¹, P. Therese Lang^{1,2}, Sabina Gerber¹, Peter B. Madrid^{1,2}, Irene Gomez Pinto¹, R. Kiplin Guy¹, and Thomas L. James^{1,*}

¹Department of Pharmaceutical Chemistry
MC 2280

²Graduate Group in Chemistry and Chemical Biology
University of California, San Francisco
600 16th Street
San Francisco, California 94158

**To whom all correspondence should be addressed (james@picasso.ucsf.edu)*

SUMMARY

We have synthesized a series of phenothiazine derivatives, which were used to test the structure-activity relationship of binding to HIV-1 TAR RNA. Variations from our initial compound, 2-acetylphenothiazine, focused on two moieties: ring substitutions and n-alkyl substitutions. Binding characteristics were ascertained via NMR, principally by saturation transfer difference spectra of the ligand and imino proton resonance shifts of the RNA. Both ring and alkyl substitutions manifested NMR changes upon binding. In general, the active site, while somewhat flexible, has regions that can be capitalized for increased binding through van der Waals interactions and others that can be optimized for solubility in subsequent stages of development. However, binding can be nontrivially enhanced several-fold through optimization of van der Waals and hydrophilic sites of the scaffold.

INTRODUCTION

Synthesis and evaluation of focused libraries of small molecules for binding to specific targets are crucial for generating leads in the drug discovery process. During the knowledge-based drug design process, repeated cycles of synthesis and testing of smaller libraries while optimizing for a desired specificity or affinity is an efficient way to generate high-quality lead compounds (Hajduk PJ et al., 2000; Rees DC et al., 2004). Although RNA is a potentially valuable drug target, few compounds specifically targeting RNA are currently on the market; some antibiotics in current clinical use provide a precedent for targeting RNA since they bind to conserved microbial rRNA structural motifs. One advantage of targeting RNA compared to developing drugs against protein receptors is that it can open new fronts in the fight against diseases (Hermann T, 2000).

According to World Health Organization estimates, 42 million people are living with HIV/AIDS worldwide. With growing resistance of the retrovirus HIV-1 (human immunodeficiency virus, type 1) to current drugs, there is need for research on other HIV-1 targets. The HIV-1 genomic RNA itself presents possibilities. One of the more prominent RNA targets is the HIV-1 TAR RNA motif, which plays an important role in the life cycle of HIV (Aboul-ela F et al., 1995; Cheng AC et al., 2001). Use of a selective inhibitor that blocks the interaction of TAR and the virus-encoded protein Tat, which regulates RNA transcriptase processivity, is one possible way of keeping the virus from proliferating (Hamy F et al., 1997). A number of recent studies have identified small-molecule ligands that bind to TAR RNA, thus inhibiting Tat binding or Tat

transactivation (Davis B et al., 2004; Hwang S et al., 2003; Lind KE et al., 2002; Renner S et al., 2005).

We have previously identified a series of “hits,” i.e., compounds that specifically bind HIV-1 TAR RNA, by using a sequential combination of virtual screening followed by limited experimental screening (Lind KE et al., 2002). Some of these hits belong to a class of compounds called phenothiazines, which have been used as antipsychotic agents for many years. Besides its proven high bioavailability and low toxicity, relative to other RNA binding drugs, phenothiazine is an attractive scaffold for potential lead development due to its low molecular weight and single positive charge, which afford ample room to build in enhanced binding affinity.

By testing a small set of commercially available phenothiazines against the TAR RNA structure, we were able to establish that this class of compounds is an interesting scaffold for further derivatization (Mayer M and James TL, 2004). NMR is especially useful at this stage of the optimization process because the tested compounds do not yet have a high affinity, thus requiring methods that are able to identify and classify weak interactions. NMR is also quite robust in that no false positives or negatives are found in NMR screening, where a hit is defined as a small molecule interacting specifically with the active site of the biomolecule. NMR also allows us to determine the affinity and specificity of the small molecules tested by monitoring imino proton chemical shifts of the RNA target (Mayer M and James TL, 2005; Yu L et al., 2003). Differential line broadening (DLB) has been used for many years to identify ligand moieties involved in

binding interactions (Fejzo J et al., 1999). Similarly, saturation transfer difference (STD) NMR has proven to be efficient in identifying and characterizing weakly binding compounds even to small RNA constructs (Mayer M and James TL, 2002).

The initial testing of commercial phenothiazines showed that substitutions around the phenothiazine ring and the presence of a basic amine functionality were important structural features for binding activity (Mayer M and James TL, 2004). An NMR structure of acetylpromazine complexed with the HIV-1 TAR RNA construct has also been solved in our lab (Du Z et al., 2002), and we used this structure to illuminate the SAR data obtained (Figure 1). The same 27 nt RNA construct used in structural elucidation of the acetylpromazine TAR RNA complex is utilized for the binding studies reported here. The objective was to generate derivatives of the known ligand basic phenothiazine scaffold in order to ascertain some aspects of a structure-activity relationship. To identify the crucial binding interactions of the original scaffold, two libraries were designed: the first one to probe substitutions of the phenothiazine ring system, and the second one to probe aliphatic side chain substitutions.

RESULTS AND DISCUSSION

Chemistry

We set out to evaluate structure-activity relationships for binding of a series of compounds with the phenothiazine scaffold to TAR RNA. A concomitant goal was to explore the ability to moderate the potency of the commercially.

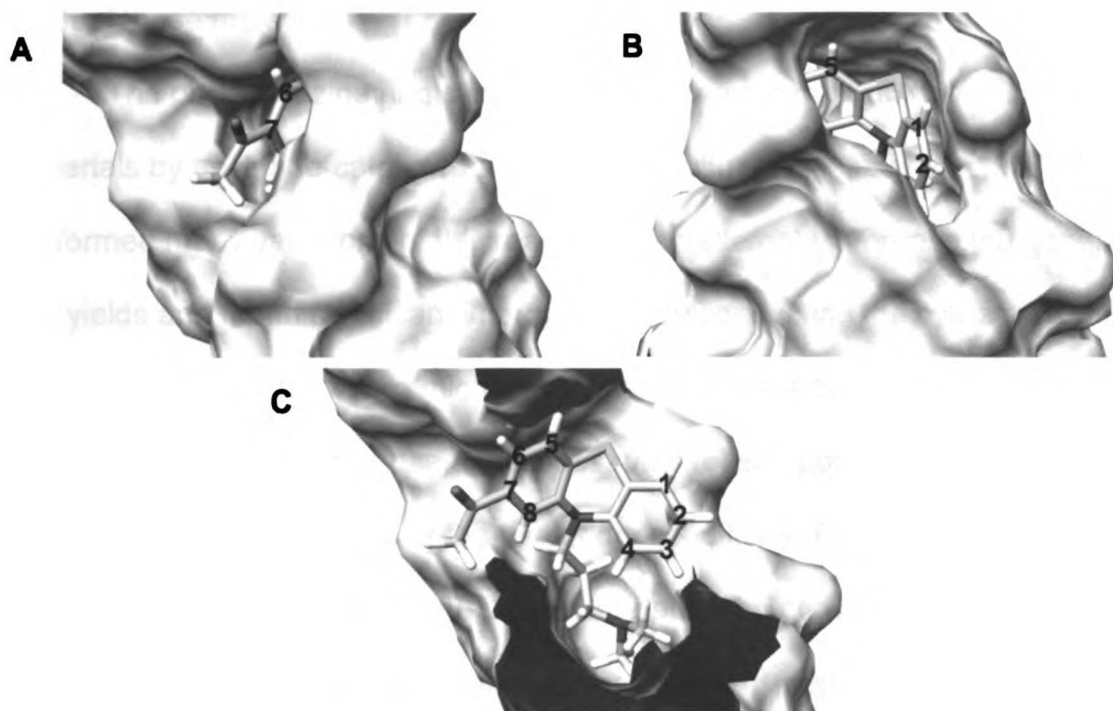


Figure 1. Binding Site of TAR RNA with Acetylpromazine Bound.

See Compound 6 in Figure 3. The NMR structure has PDB code 1LVJ(Du Z et al.).

A) Perspective of the phenothiazine scaffold from the minor groove.

B) Perspective of the phenothiazine scaffold from the major groove.

C) Perspective of the entire compound. The second half of RNA was removed for visualization purposes.

available compounds already tested. We dissected the scaffold into two parts: substitutions of the phenothiazine rings and substitution of the n-alkyl linker variations. We proceeded to synthesize chemical libraries with modifications in each of those moieties

The phenothiazine ring system was constructed from diarylamine starting materials by an iodine-catalyzed reaction with sulfur (Yale HL et al., 1957). We performed many variations of this reaction in an attempt to improve the generally low yields and found significant improvement by conducting the cyclization with microwave irradiation. While microwave conditions have been briefly described before for the synthesis of phenothiazines, they relied upon the use of strong acid or Lewis acid catalysts or the use of dry conditions (Filip SV et al., 1998; Villemin D and Vlieghe X, 1998). Under our optimized conditions, the reaction is carried out in water in a sealed reaction vessel and is heated to 190°C for 20 min under microwave irradiation (Figure 2A, Compounds 1 and 2). This method was very convenient, since the hydrophobic product immediately precipitated upon cooling and could easily be purified by filtration. This new, to our knowledge, method for the synthesis of phenothiazines was successfully used to synthesize a set of new substituted 10H-phenothiazines (Figure 2A, Compound 2). The 10H-phenothiazines were then alkylated by using sodium hydride and 1-chloro-3-iodopropane (or appropriate chloriodoalkane) to give the chloroalkylphenothiazine (Figure 2A, Compound 3). This intermediate could then be aminated with dimethylamine (or another secondary amine) in a sealed vessel with microwave irradiation to afford the ring-substituted phenothiazine library

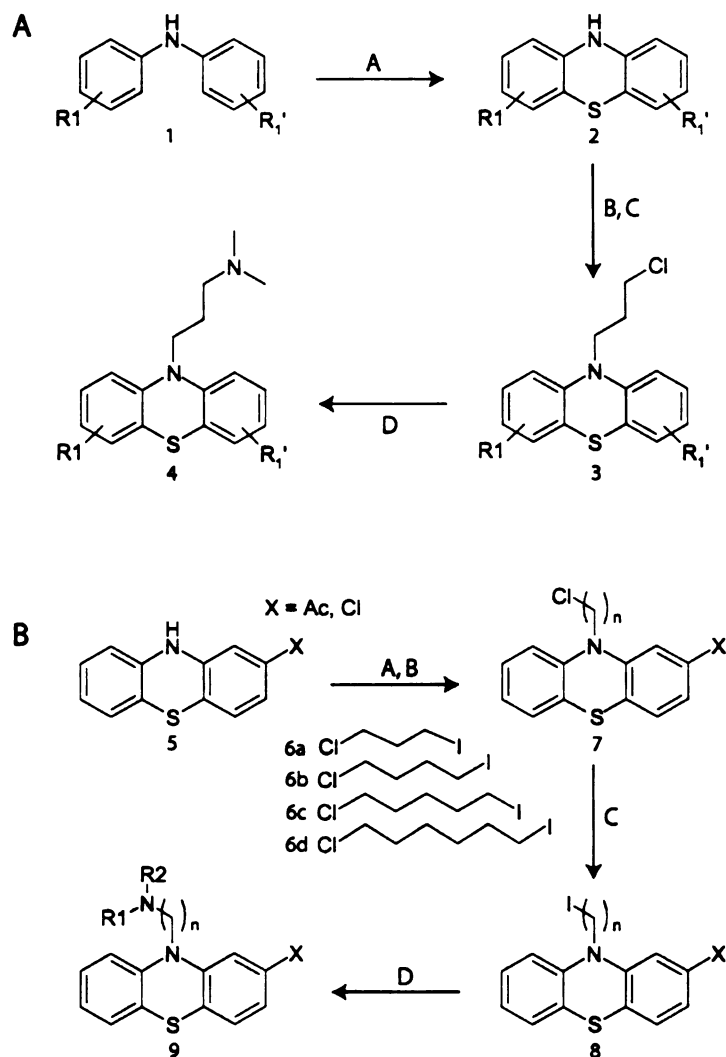


Figure 2. Synthesis of a 10H-Phenothiazine Library

For more information, see Experimental Procedures (“Synthesis of Compounds”) and Results and Discussion (“Chemistry”).

A) Ring variant syntheses. The conditions for syntheses are: (A) 2 eq. S₈, catalyst I₂, water, 50 W microwave, 190°C, 20 min; (B) 1.2 eq. NaH, DMF, 0°C, 1 hr; (C) 5 eq. Cl(CH₂)₃I (slow reverse addition), RT, 1 hr; (D) 4 eq. dimethylamine (8{1}), 2M THF, DMF, 30 W microwave, 100°C, 40 min.

B) Side chain variations. The conditions for syntheses are: (A) 2.1 eq. NaH, THF, DMSO; (B) 3 eq. Cl(CH₂)_nI, (6a–6d), reverse add., THF, DMSO, RT, 2 hr; (C) 5 eq. KI, 2-methyl-4-propanone, 130°C, reflux, 40 hr; (D) 2 eq. HNR₁R₂, DMF, 60°C, 20 hr.

(Figure 2A, Compound 4). The subset of these compounds that were soluble and stable enough to be tested for binding by NMR is reported in Figure 3.

For evaluation of side chain substituents, we decided to fix the ring scaffold as 2-acetylphenothiazine and 2-chlorophenothiazine and to vary the length of the alkyl linkers and the substituents on the basic amine group (Figure 2B). We again alkylated the 10H-phenothiazine ring by using sodium hydride and four chloriodo alkyl chains, the length of which varied from 3 to 6 methylene units (Figure 2B, Compounds 5–7). We found that the resulting alkyl chloride intermediate was not sufficiently reactive to yield product with a diverse set of amines, so we used the Finkelstein reaction to convert the chloroalkane to an iodoalkane (Figure 2B, Compound 8). The iodoalkane intermediate then could be reacted with a diverse set of primary and secondary amines to afford the completed side chain variation library (Figure 2B, Compound 9). In general, these reactions proceeded quite cleanly, with the major side reaction being the elimination of the iodide in longer linkers. For this set, those compounds soluble and stable enough to be tested by NMR are reported in Figure 4.

Library Screening

STD NMR experiments and differential line broadening (DLB) of proton resonances on the ligand enabled assessment of ligand binding to the RNA target (Fejzo J et al., 1999; Mayer M and James TL, 2002; Mayer M and James TL, 2005). RNA imino proton chemical shifts were monitored to identify and

characterize the location of small-molecule binding to the 27 nt construct of HIV-1 TAR RNA. Only stable, water-soluble compounds were utilized for these studies.

The affinities of the compounds for the RNA construct were divided into three categories by a combination of DLB and STD effects. For binders deemed more promising by ligand NMR results, RNA imino proton chemical shift measurements were also carried out. We were able to classify 24 compounds according to their RNA binding interactions manifest in resonance line broadening and STD spectral intensities. Figure 5 shows reference and STD spectra of two selected compounds from the libraries. Compound 17 exhibited strong line broadening effects and substantial STD enhancements (Figure 5A). Compounds with these characteristics were categorized as strong binders. On the other hand, Compound 23 shows only weak broadening of the aromatic protons, at similar ligand and RNA concentrations as for Compound 17 (Figure 5B). Compounds showing these characteristics were categorized as weak binders.

Nine compounds were examined to determine the effects of aromatic ring substitutions on binding affinity (Figure 3). For analysis, we divided the substituents by location on the ring system (see Figure 1 for numbering). Note, however, that the symmetry of the phenothiazine ring system requires us to consider whether a substituent is exerting its influence on the minor groove or major groove side of the complex. Bond rotation around the linker-N-to-alkyl-C bond will enable almost the same conformation in the complex. As shown in Figure 1, we have found that a hydrophilic group at position 7 (or 3) will be.

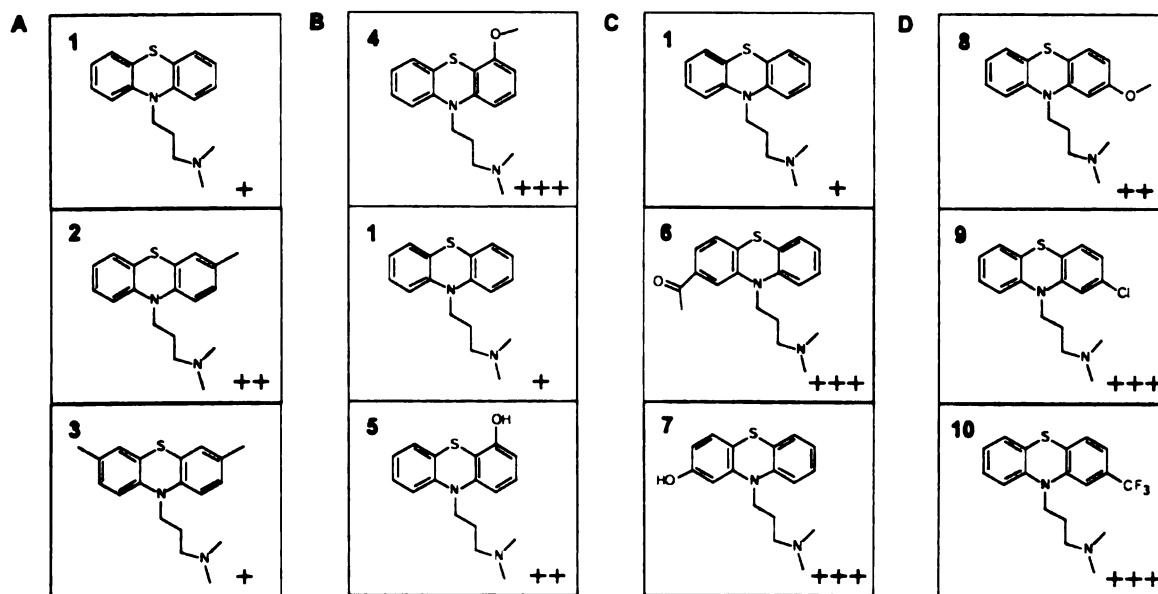


Figure 3. Binding between Phenothiazine Ring Derivatives and TAR RNA

A–D) Binding was classified into three binding categories: affinities of ~ 0.1 – 1 mM (+++), affinities of ~ 1 – 5 mM (++) , and affinities weaker than ~ 5 mM (+). For all panels, the compounds are at a concentration of $500 \mu\text{M}$ and the RNA is at a concentration of $50 \mu\text{M}$. Compounds are separated from left to right by the location of substituents on the rings and are roughly ordered from top to bottom by increasing hydrophilicity.

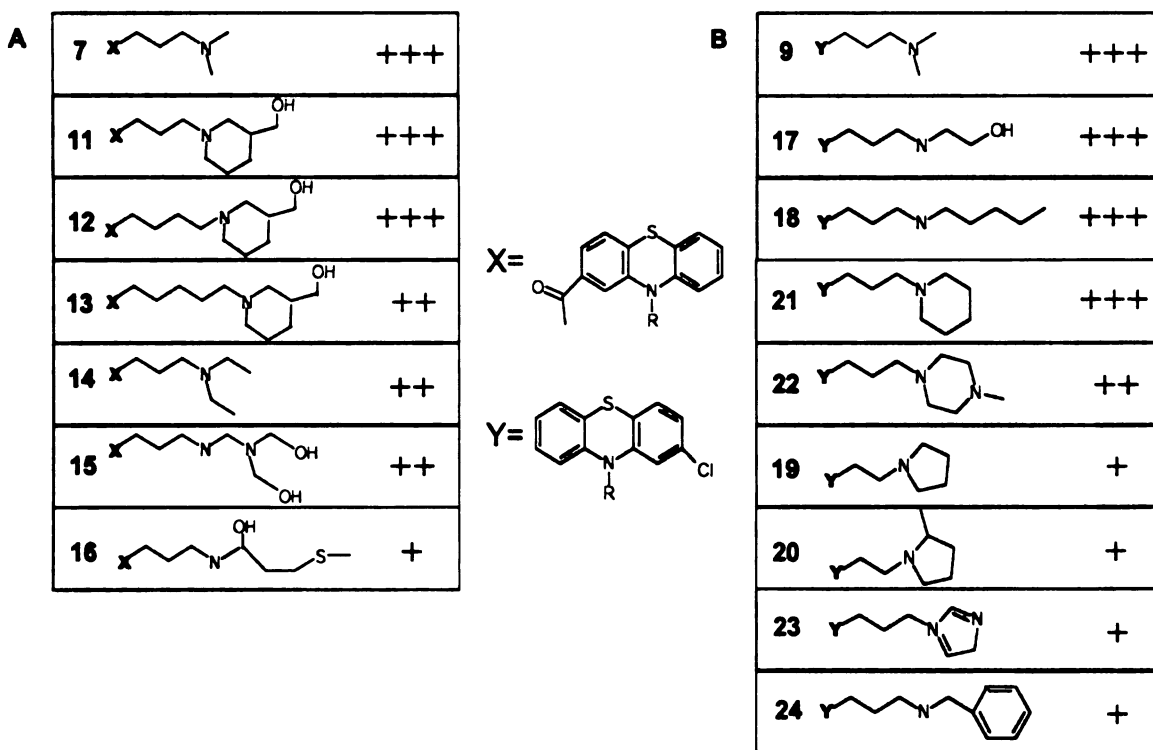


Figure 4. Phenothiazine Scaffolds with Side Chain Substitutions Tested for Binding to TAR RNA

A and B) As explained in the Results and Discussion (“Library Screening”), the chlorine substituent causes the ring system to flip in orientation from the structure in Figure 1. As a result, affinity comparisons can only be made within a particular column. Binding categories and experimental conditions same as for Figure 3.

exposed to solvent on the minor groove side of the complex. That being said, TAR RNA has been shown to be extremely flexible in solution, and portions of the active site can potentially distort to accommodate particular substituents (Bayer P et al., 1995; Davis B et al., 2004; Du Z et al., 2002; Ippolito JA and Steitz TA, 1998; Murchie AIH et al., 2004)

In Figure 3A, we examine the effect of adding methyl groups. In Compound 2, the methyl group is in position 2 on the ring scaffold. If a substituent in this position is oriented toward the major groove side of the binding pocket, as in position 6, it will sustain some amount of van der Waals (VDW) clash. Therefore, we expect that the methyl orientation is toward the minor groove, making room for the single methyl substituent. The double methyl substituent supports this hypothesis with reduced binding affinity, which we attribute to VDW clash from the second methyl group, which must be on the major groove side.

In Figures 3B–3D, we have ranked the compounds by hydrophobicity. In Figure 3B, substituents in position 1 are solvent exposed but could also pack with the top portion of the binding site. This orientation rationalizes the improved binding for the more hydrophilic hydroxyl substituent as well as the additional increase in affinity from VDW interactions with the methoxy substituent. The carbon at position 7 in Figure 3C is completely solvent exposed. Therefore, we propose that increasing hydrophobicity of the substituent in this position improves binding in the series. Finally, the carbon 3 position is completely buried. Therefore, the more hydrophobic compounds in Figure 3D still have reasonable

binding. The methoxyl substituent has reduced binding in the series due to VDW clashes in the major groove orientation.

For the side chain-substituted library, the ratio of the STD intensities of the aromatic region resonances versus the aliphatic region resonances and line broadening were used to rank the small library of compounds. From analysis of the STD spectra of the originally discovered ligand acetylpromazine, we knew that the aromatic rings of the heterotricyclic structure have a closer interaction with the RNA protons than the aliphatic side chain (Mayer M and James TL, 2004). This is not apparent in the NMR structure of acetylpromazine in complex with TAR RNA, where the side chain and ring system of the ligand appear to be equally interactive with the RNA (Du Z et al., 2002). As discussed previously (Mayer M and James TL, 2004), we believe that this apparent discrepancy may be due to the fact that we are pinning down the relatively flexible side chain in the structure calculation by intermolecular NOEs, whereas that particular conformation with the side chain pinned down in the minor groove may represent less than half of the population of the ligand in the bound form. The weaker STD intensities of the side chain implied that optimization of that region of the ligand would be beneficial. Phenothiazine derivatives with stronger STD intensities in the aliphatic side chain may reflect a larger population of bound ligand with the side chain immobilized in the minor groove.

For the side chain variation series, the relative STD enhancements of the aromatic protons were larger than those of the side chains, providing strong evidence that it is the phenothiazine ring that provides much of the binding

energy. However, in the substituent comparison, the phenothiazine ring system remained the same within each series. Therefore, it was possible to correlate stronger aliphatic STD enhancements with higher affinity of the ligands because the phenothiazine remained largely unchanged. In addition, stronger relative STD signals of the aliphatic groups were reflected in stronger DLB effects. Detailed epitope mapping of the ligands was not performed due to significant overlap in some cases, especially of the aromatic protons.

The qualitative binding affinities for the side chain variations can be found in Figure 4. In the solved structure, the minor groove portion of the lower half of the binding site is largely comprised of an unpaired cytosine from the loop region of the RNA. This cytosine closes over the ligand situated at the binding site; in fact, the ligand cannot get into the structure determined without the bulge cytosine being moved aside. While this cytosine is within hydrogen bonding distance of the ribose of the following nucleotide, we hypothesize that this portion of the binding site is somewhat malleable and thus able to accommodate larger groups on the side chain region. For the acetylphenothiazine scaffold derivatives in Figure 4A, we suggest that the cyclohexyl structure is intrinsically stable enough to deform the active site when combined with additional VDW interactions. Compound 14, however, has a bit too much bulk for the active site as well as significant entropy loss upon binding, resulting in a decrease in binding. For the remaining compounds in the list, the substituents have enough bulk to cause little VDW clash, but too much internal flexibility to distort the active



Figure 5. Representative NMR Spectra of Phenothiazine Compounds from Figure 4

For all spectra, the compound is at a concentration of 500 μM and the RNA is at a concentration of 50 μM .

(A) Compound 17: (1) Reference spectrum of compound alone; an asterisk indicates impurities. (2) Reference spectrum of compound in the presence of TAR RNA showing considerable line broadening induced by the oligonucleotide. (3) STD NMR spectrum of the compound in the presence of TAR RNA. The RNA is irradiated at 5.8 ppm, as indicated in the figure by the arrow. Signals of Compound 17 indicating the binding interaction are clearly visible, while those of the impurities indicated by asterisks are subtracted. RNA signals are also visible because the molecular size is not large enough to eliminate the background signals by a relaxation filter.

(B) Compound 23: (1) Reference spectrum of compound alone. (2) Reference spectrum of the compound in the presence of TAR RNA showing a small amount of line broadening localized to the aromatic ring protons. (3) STD NMR spectrum of the compound in the presence of TAR RNA. The RNA is irradiated at 5.8 ppm, as indicated in the figure by the arrow. In this case, signals for the aromatic protons manifest an effect, while the side chain and the imidazole group are subtracted.

site. In addition, because these compounds are so flexible, the entropic penalty to bind these molecules must be higher than the enthalpic gain.

For the choro-substituted phenothiazines in Figure 4B, the side chain will be in a slightly different orientation in response to the flipped scaffold (relative to the acetylphenothiazine series); thus, this series cannot be directly compared to that in Figure 4A. However, many of the same conclusions can be drawn. For Compounds 17 and 18, the single substituents have a thinner diameter than the original dimethyl substituents, which we hypothesize enables these flexible chains to adapt to the minor groove. The two cyclohexyl compounds most likely have the same effect as what is shown in Figure 4A; they are intrinsically stable enough to deform the active site while compensating with additional VDW packing. The cyclopentyl rings have a shorter chain length than any other substituent in the series. Because the active site does not widen until a bit further down, these substituents are likely disrupting the upper portion of the active site. Finally, for both of the aromatic compounds, the stacking interactions with bases are not possible without significant rearrangement of the TAR structure, resulting in entropic loss of binding. In this case, the substituents would be solvent exposed, thus explaining their reduced binding.

For compounds that showed large STD intensities and exchange broadening effects, the RNA imino proton resonances were monitored. Imino proton chemical shift changes upon addition of ligand provide information about the binding site on the RNA and also permit K_D values to be determined (Figure 6). For Compound 11, the most potent of the tested molecules, the calculated K_D

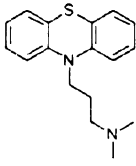
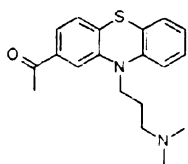
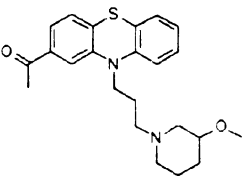
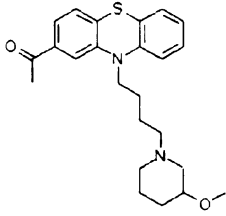
Compound	Dissociation Constants
	5 mM
	0.27 mM
	0.18 mM
	0.14 mM

Figure 6. Binding Affinities as Determined by Chemical Shift Monitoring of TAR RNA Imino Proton Resonances for Selected Compounds

value was 140 μM , or about twice that of the original ligand, acetylpromazine (Mayer M and James TL, 2004). As shown for Compound 11 with the 3-hydroxymethylpiperidine ring (see Figure 7), the imino resonance that shifts most upon addition of the ligand is that of G26 of the HIV-1 TAR RNA. The imino proton resonance of G26 exhibits a large upfield shift due to ring current effects from the stacked ligand benzene I. As with Compound 11, the other ligands exhibited imino chemical shifts only for residues in the vicinity of the TAR bulge. It should also be noted that there was no evidence that more than one of the tested ligands binds to the RNA at the same time, as no significant imino chemical shifts were observed outside of the binding site for any of the tested compounds. From previous studies, which entailed several different kinds of multinuclear and multidimensional NMR spectra, we also found that only residues in the vicinity of the bulge experienced changes upon ligand binding, and that there were no changes evident for other residues upon addition of a large excess of ligand (Du Z et al., 2002; Lind KE et al., 2002).

We would like to emphasize that a well-defined SAR cannot be expected from compounds with binding affinities in this range. However, the SAR performed for the focused library of phenothiazines suggests a few avenues for future developments. A side chain with 3–4 methylenes between the ring and the heteroatom seems optimal (Figure 4). Consistent with the NMR structure, a flexible linear chain will snake nicely along the minor groove. In addition to VDW interactions, heteroatoms or functional groups with hydrogen bonding capabilities may improve binding. Some rigidity in the side chain will decrease entropic loss

upon binding, but the orientation and nature of further derivitization will become more critical. Ring substituents may further enhance binding, and extension beyond the methoxyl group in position 1 should gain additional VDW interactions. The results of Figure 3 also imply that it might be of value to have a phenothiazine with a hydrophilic substituent in position 7 and a small hydrophobic substituent in position 3. These implications from the SAR are entirely consonant with the structure.

SIGNIFICANCE

The synthesis of the focused library of phenothiazine derivatives—a class of nontoxic, bioavailable RNA binders—and analysis of the SAR for binding to HIV-1 TAR RNA reported here are important extensions of earlier work. The study explores structural features affecting phenothiazine interactions with TAR RNA. The observation that both moieties varied here—ring substituents and the n-alkyl substitutions—affect binding agrees with our published structure of TAR with acetylpromazine. In addition, the results described above suggest that the binding site, while flexible, has regions that restrict growth. There are also several areas of the binding site that are solvent exposed, which can be exploited in later stages of development.

The most significant factor in the data is that compounds possessing a four-atom linker between the tricycle and the distal base are among the tightest binders. In the pharmacological literature, neuroleptic activity is strongly associated with a three-atom linker, and antihistamine activity is associated with

a two- or three-carbon linker. Obviously, any RNA binding compound would require the ability to dissociate these undesirable activities from therapeutic effects. Compounds with longer linkers provide an attractive potential route toward that goal.

In future work, knowledge gained in this study can guide further enhancement of specificity. The phenothiazines studied have a single positive charge and relatively low molecular weight. In addition to modifications studied here, increased affinity may be gained by combining the appropriate substituted phenothiazine with other groups known to enhance RNA binding (e.g., amino sugars, guanidino groups, and intercalating groups). However, as shown here, binding affinity can be improved through optimization of VDW and hydrophilic sites of the scaffold. This hypothesis is supported by the increase in binding from the naked promazine scaffold to the final product from this study, which have binding affinities of 5 mM and 140 μ M, respectively.

EXPERIMENTAL PROCEDURES

NMR Experiments

Unlabeled RNA samples were prepared and purified as previously described (Mayer M and James TL, 2004). NMR experiments were performed on a Varian Inova 600 MHz spectrometer. Typically, a spectrum of the compound alone at a 250 or 500 mM concentration was acquired, and then RNA was added from a concentrated stock solution to produce a 25 or 50 μ M final RNA concentration. Reference and STD spectra were compared. STD NMR spectra

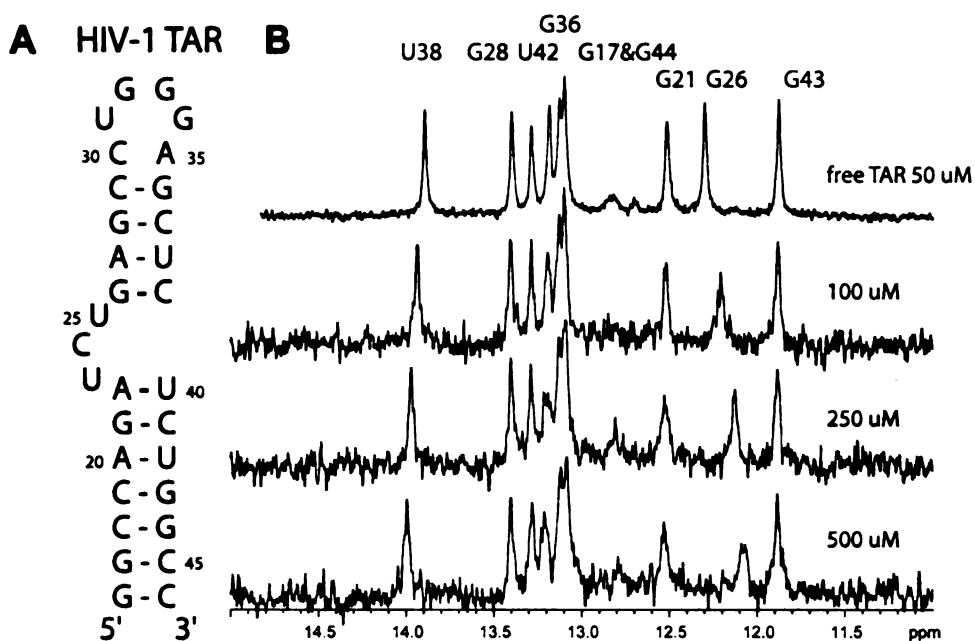


Figure 7. Representative Imino Proton NMR Spectra of RNA in the Presence of Compound 11 from Figure 4

For all spectra, the RNA is at a concentration of 50 mM.

A) Sequence of the TAR RNA construct used for testing the phenothiazine derivatives.

B) Chemical shift monitoring of TAR RNA imino proton resonances upon addition of Compound 11 in increasing concentrations, from 0 to 500 mM. The chemical shift difference is plotted against the compound concentration to determine the dissociation constant of the TAR-ligand complex.

were acquired by internal subtraction via phase cycling. On-resonance irradiation was set to 5.8 ppm, and off-resonance irradiation was set to ~30 ppm where no RNA resonances are present. Typically, 256 scans were acquired for the STD experiments. Presaturation of RNA resonances was achieved by an appropriate number of band-selective G4 Gaussian cascade pulses to give a saturation time of 2 s. The temperature was set to 15°C in all STD experiments. 1D-NMR imino proton spectra were acquired by using hard-pulse WATERGATE water suppression, and the excitation profile was optimized for maximum intensity at 13 ppm

For the qualitative ranking of compound binding affinities, spectra of resonances from the ligand alone were compared with the same resonances in STD spectra acquired with ligand in the presence of RNA (see Figure 5 for representative examples). A classification of binding was made according to observed STD intensities and line broadening effects (Figures 3 and 4).

For a select subset of derivatives, dissociation constant (K_D) values were determined by titrating the RNA with ligand and monitoring the RNA imino proton chemical shift differences (Figure 6). Typically, five titration points were taken for each compound for which we determined the K_D . In a previous study, the first compound of this series, acetylpromazine, was found to have a binding affinity of 270 μM (Mayer M and James TL, 2004). The qualitative study described above found the remaining compounds to be in the same binding range of ~0.1–1.0 mM. Substantial line broadening observed for some of the tested compounds stems mostly from the broadening of resonances due to chemical exchange of

the ligand either between free and bound species or, more likely, from multiple conformations on the RNA; the latter may result from the symmetry of the ligand binding in a similar fashion, but with the ring flipping 180°. Except for the parent compound listed for comparison, for the compounds whose K_D values were determined (Figure 6), the symmetry was broken such that one orientation of binding was strongly preferred; this was clearly evident in the quality of the various spectra (primarily lineshapes) acquired for these complexes compared with the others. For these three compounds, it appears that fast-exchange conditions were obtained. As noted already, our NMR results indicate that these ligands bind to only a single site on our TAR RNA construct. The equation for single-site binding under the fast-exchange regime,

$$K_D = C_L \left(\frac{\delta - \delta_F}{\delta_B - \delta} \right) \quad (1)$$

was used to obtain K_D s as chemical shifts were observed for the complex (δ), as a function of the concentration of the ligand (C_L), relative to the chemical shifts for free RNA (δ_F) and ligand bound RNA (δ_B) (Peng JW et al., 2001).

Synthesis of Compounds

All reagents and starting materials were purchased from commercial sources and were used without further purification; solvents were anhydrous HPLC grade. All parallel synthesis steps were carried out in polypropylene fritted reaction tubes with the Bohdan Miniblock 48-position reaction blocks. Microwave

reactions were done in a CEM Discover microwave reactor system. The reaction schemes described below can be seen in Figure 2.

General Procedure for the Preparation of Substituted 10*H*-Phenothiazines

Each diphenylamine (Figure 2A, Compound 1) (2.0 mmol) was sealed in a microwave reaction vessel with sulfur (128 mg, 4.0 mmol), a crystal of iodine, and 2.5 ml doubly distilled water. The vessels were heated to 190°C for 20 min and then cooled to room temperature. The crude mixture was extracted with 10 ml ethyl acetate and then purified by column chromatography (silica gel; 90% ethyl acetate, 10% hexanes) to give the substituted 10*H*-phenothiazine product (Figure 2A, Compound 2) (5%–45% yield).

General Procedure for the Alkylation of 10*H*-Phenothiazines

Sodium hydride (0.06 mmol) was added slowly at 0°C to each stirring solution of 10*H*-phenothiazine (Figure 2A, Compound 2) (0.05 mmol) dissolved in DMF (10 ml). The reaction was stirred for 30 min, and was then allowed to warm to room temperature. This solution was then added dropwise into a stirring solution of 1-chloro-3-iodopropane (27 μ l, 0.25 mmol) in DMF and was stirred for 1 hr. The reaction mixture was extracted from a brine solution with ethyl acetate (25ml, 3x) and was then evaporated under reduced pressure. The crude product was purified by column chromatography (silica gel; 90% ethyl acetate, 10% hexanes) to give the substituted 10*H*-phenothiazine product (Figure 2A, Compound 3) (35%–85% yield).

General Procedure for the Amination of 10-(3-Iodopropyl)-Phenothiazines

Dimethylamine (4 eq., 40 μ l of a 2 M solution in THF) was added to each member of the substituted 10*H*-phenothiazine product (Figure 2A, Compound 3) (0.02 mmol) dissolved in 1 ml DMF. The mixture was sealed into a microwave reaction vessel and was heated at 100°C for 40 min. The crude reaction mixture was then purified by silica gel prep-TLC by using a 94.9:5:0.1 ratio of dichloromethane:methanol:triethylamine. Pure products (Figure 2A, Compound 4) were isolated in 60%–85% yields.

General Procedure for the Alkylation of (2-Chloro or 2-Acetyl)-10*H*-Phenothiazines

Sodium hydride (21 mmol, 60% dispersion in mineral oil) was slowly added to a mixture of 2-chloro- or 2-acetyl-10*H*-phenothiazine (Figure 2B, Compound 5) (10 mmol) stirring in a 1:1 solution of DMSO:THF at 0°C. The mixture was allowed to warm to 25°C and was then added to a stirring solution of iodochloroalkanes (Figure 2B, Compounds 6a–6d) (30 mmol) in *N,N*-dimethylformamide (20 ml). The reaction mixture was stirred at room temperature for 2 hr, and 100 ml brine was added to stop the reaction. The crude product was extracted into ethyl acetate (3x, 20 ml) and then purified by column chromatography (silica gel; 90% ethyl acetate, 10% hexanes) to give the title compounds (Figure 2B, Compound 7) in 40%–80% yield.

General Procedure for the Iodo Halogen Exchange Reaction

A mixture of potassium iodide (5 re, 5 mmol) and each member of the alkylated (2-chloro or 2-acetyl)-10*H*-phenothiazines (Figure 2B, Compound 7) (1

mmol) was heated until reflux in 2-methyl-4-propanone until the reaction had reached completion as determined by NMR (about 3 days). The solvent was removed under vacuum, and the product (Figure 2B, Compound 8) (95% yield) was used without any further purification.

General Procedure for the Amination Diversity Step

Each iodoalkane intermediate (Figure 2B, Compound 8) (0.33 mmol) was dissolved with N,N-dimethylformamide (5 ml) and divided into five portions (0.06 mmol each). Each portion was mixed with 0.13 mmol *n*-butylamine, benzylamine, 2-aminoethanol, piperidine, and imidazole. The crude reaction mixtures were worked up by extracting into methylene chloride (40 ml) from a 10% NaOH solution (40 ml). Solvent was removed under vacuum, and each compound was then purified by silica gel prep-TLC by using a 94.9:5:0.1 ratio of dichloromethane:methanol:triethylamine. Pure products (Figure 2B, Compound 9) were isolated in 60%–85% yields.

ACKNOWLEDGEMENTS

This work was supported by grant AI46967 from the National Institutes of Health and grant 02881-31-RG from the American Foundation for AIDS Research. P.T.L. wishes to acknowledge support from the Burroughs Wellcome Fund and the American Foundation for Pharmaceutical Education.

REFERENCES

- Aboul-ela, F, Karn, J and Varani, G. *J Mol Biol.* **253** (1995) 313-32.
- Bayer, P, Kraft, M, Ejchart, A, Westendorp, M, Frank, R and Rösch, P. *J Mol Biol.* **247** (1995) 529-35.
- Cheng, AC, Calabro, V and Frankel, AD. *Curr Opin Struct Biol.* **11** (2001) 478-484.
- Davis, B, Afshar, M, Varani, G, Murchie, AIH, Karn, J, Lentzen, G, Drysdale, MJ, Bower, J, Potter, AJ, Starkey, ID, Swarbrick, TM and Aboul-Ela, F. *J Mol Biol.* **336** (2004) 343-356.
- Du, Z, Lind, KE and James, TL. *Chem Biol.* **9** (2002) 707-12.
- Fejzo, J, Lepre, CA, Peng, JW, Bemis, GW, Ajay, Murcko, MA and Moore, JM. *Chem Biol.* **6** (1999) 755-69.
- Filip, SV, Silberg, IA, Surducun, E, Vlassa, M and Surducun, V. *Synth Commun.* **28** (1998) 337-345.
- Hajduk, PJ, Gomtsyan, A, Didomenico, S, Cowart, M, Bayburt, EK, Solomon, L, Severin, J, Smith, R, Walter, K, Holzman, TF, Stewart, A, McGaraughty, S, Jarvis, MF, Kowaluk, EA and Fesik, SW. *J Med Chem.* **43** (2000) 4781-6.
- Hamy, F, Felder, ER, Heizmann, G, Lazdins, J, Aboul-ela, F, Varani, G, Karn, J and Klimkait, T. *Proc Nat Acad Sci USA.* **94** (1997) 3548-3553.
- Hermann, T. *Angew Chem Int Ed.* **39** (2000) 1891-1905.
- Hwang, S, Tamilarasu, N, Kibler, K, Cao, H, Ali, A, Ping, YH, Jeang, KT and Rana, TM. *J Biol Chem.* **278** (2003) 39092-103.
- Ippolito, JA and Steitz, TA. *Proc Natl Acad Sci USA.* **95** (1998) 9819-24.
- Lind, KE, Du, Z, Fujinaga, K, Peterlin, BM and James, TL. *Chem Biol.* **9** (2002) 185-93.
- Mayer, M and James, TL. *J Am Chem Soc.* **124** (2002) 13376-7.
- Mayer, M and James, TL. *J Am Chem Soc.* **126** (2004) 4453-60.
- Mayer, M and James, TL. *Methods Enzymol.* **394** (2005) 571-87.

Murchie, AIH, Davis, B, Isel, C, Afshar, M, Drysdale, MJ, Bower, J, Potter, AJ, Starkey, ID, Swarbrick, TM, Mirza, S, Prescott, CD, Vaglio, P, Aboul-ela, F and Karn, J. *J Mol Biol.* **336** (2004) 625-638.

Peng, JW, Lepre, CA, Fejzo, J, Abdul-Manan, N and Moore, JM. "Nuclear Magnetic Resonance-Based Approaches for Lead Generation in Drug Discovery." In *Methods Enzymol.* T. L. James, V. Dotsch and U. Schmitz Ed. Academic Press: San Diego (2001) 202-230.

Rees, DC, Congreve, M, Murray, CW and Carr, R. *Nat Rev Drug Discov.* **3** (2004) 660-72.

Renner, S, Ludwig, V, Boden, O, Scheffer, U, Gobel, M and Schneider, G. *ChemBioChem.* **6** (2005) 1119-25.

Villemin, D and Vlieghe, X. *Sulfur Lett.* **21** (1998) 191-198.

Yale, HL, Sowinski, F and Bernstein, J. *J Am Chem Soc.* **79** (1957) 4375-4379.

Yu, L, Oost, TK, Schkeryantz, JM, Yang, J, Janowick, D and Fesik, SW. *J Am Chem Soc.* **125** (2003) 4444-50.

“The essence of science: ask an impertinent question, and you are on the way to a pertinent answer.”

- Jacob Bronowski

Chapter 5

Optimization of DOCK for RNA Targets

P. Therese Lang¹, Scott Brozell², Alan Graves³, Kaushik Raha⁴, Sudipto Mukherjee⁵, Veena Thomas⁶, Robert Rizzo⁵, Brian Shoichet⁴, David Case⁷, Thomas L. James⁴, and Irwin D. Kuntz^{4*}

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

²Ohio Supercomputer Center
Columbus, OH

³Graduate Program in Biophysics
University of California, San Francisco

⁴Department of Pharmaceutical Chemistry
University of California, San Francisco

⁵Department of Applied Mathematics and Statistics
Stony Brook University, New York

⁶Pharmaceutical Sciences and Pharmacogenomics
University of California, San Francisco

⁷Department of Molecular Biology
The Scripps Research Institute, San Diego

**To whom all correspondence should be addressed (kuntz@cgl.ucsf.edu)*

ABSTRACT

With a continually increasing interest in using RNA for therapy and for potentially targeting RNA to treat disease, there is a need for the tools used for protein-based drug design, particularly docking algorithms, to be extended or adapted for nucleic acids. Here, we have compiled a test set of RNA-ligand complexes to validate the ability of the DOCK suite of programs to successfully recreate experimentally determined binding poses. At this point, we have characterized the ligand sampling algorithm and made a few improvements in both success rate and length of calculation. With the optimized parameters, 60% of a subset of the test set with less than 7 rotatable bonds in the ligand, 41% of the test set with less than 10 rotatable bonds, and 37% of the test set with less than 12 rotatable bonds can be successfully recreated. In the next stage, we will apply the new scoring functions available in the latest version of DOCK, which we hypothesize will further increase our success rates.

INTRODUCTION

In the past few years, the role of RNA in cellular processes has been greatly expanded. No longer is RNA simply for transporting genetic information from the nucleus to the cytoplasm for translation. Rather, it has been shown to be an integral part of biological processes. For example, in the ribosome, RNA has been shown to be responsible for a wide range of function including catalyzing the formation of the nascent peptide bonds (Frank J and Spahn CMT, 2006; Polacek N and Mankin AS, 2005). Other RNA molecules, like HIV TAR and bacterial riboswitches, recruit and bind proteins to regulate reproduction of the HIV genome and the production of various processes, respectively (Bannwarth S and Gagnol A, 2005; Frankel AD and Young JA, 1998; Tucker BJ and Breaker RR, 2005). These and other RNA-protein interactions are critical for cellular function and thus make potential drug targets.

Several drug design efforts for RNA targets have already been attempted with various levels of success, and, with the increasing evidence of the importance of RNA in regulation of the cell, these efforts will increase as well (see (Johansson D et al., 2005; Mayer M and James TL, 2005; Mayer M et al., 2006; Nakatani K et al., 2006; Renner S et al., 2005; Yu XL et al., 2005) for example). As a result, there is a need for the same tools that are used in drug design, in particular docking algorithms, for protein targets to be adapted and extended for nucleic acids. Previous studies suggest that poor modeling of the high localized charge in RNA targets through both the scoring function and the estimation of charge may be limiting the success of docking algorithms (Lind KE

et al., 2002). In addition, in a prior investigation, we explored the ability of the docking algorithm DOCK to recreate experimentally determined binding poses for a protein test set and also identified the limiting factor for success was the scoring function (Moustakas D et al., in press). The newest release of the DOCK suite of programs, version 6, is ideally suited to address the scoring function issues for both RNA and protein targets.

New Features of DOCK

The release of DOCK version 6 is an important extension of previous versions of the code. Version 5 was a reimplementaion of version 4 algorithm in a modular, extendable format. The purpose of this rewrite was to create the basic scaffold of a docking algorithm, in which the algorithm for each major component could be easily replaced by an alternate method without affecting the overall workflow. This newest release is a direct application of the extensibility with the addition of a number of new features, including DOCK 3.5 scoring, Hawkins-Cramer-Truhlar Generalized Born with Solvent Accessible Surface Area (GB/SA) solvation scoring with optional salt screening, Poisson Boltzmann with Solvent Accessible Surface Area (PB/SA) solvation scoring, and AMBER scoring with optional receptor flexibility. All of these new features, which are described below in greater detail, were plugged into the basic core of the original docking code.

The first new scoring function, originally implemented in DOCK version 3.5, is a variant of the grid-based scoring function implemented in version 5 (Meng EC et al., 1992; Moustakas D et al., in press). While the van der Waals

(VDW) values used for both methods are the same, instead of using the AMBER parm99 partial charges for the biomolecular target, the DOCK 3.5 charges are calculated using finite difference Poisson-Boltzmann equation as implemented in the program DelPhi (Nicholls A and Honig B, 1991). The scoring function also has the option to incorporate both ligand and receptor desolvation. For the ligand alone, the atomic desolvation for each atom is precomputed, typically by the program AMSOL, and is stored as an additional field at the end of the standard ligand file (G.D. Hawkins et al., 2004). The desolvation of each atom is then normalized by the extent to which the ligand is buried in the active site by the SOLVMAP accessory distributed with DOCK. Alternately, for the receptor and ligand together, the desolvation can alternatively be calculated using the occupancy desolvation method, in which the desolvation energy is calculated as the affinity of the ligand for solvent weighted by the volume of the solvent displaced from the receptor atom due to binding and vice versa (Luty BA et al., 1995). The desolvation energy for both the receptor and ligand are precomputed using the SOLVGRID accessory also distributed with DOCK.

The next two scoring methods, GB/SA and PB/SA, are implicit solvent models that account for the effect of solvent on the electrostatics of the complex. In the GB/SA scoring function, the interaction between the ligand and receptor are calculated using Coulomb's Law without distance dependent dielectric for electrostatics and the Lennard-Jones potential for VDW energies as the grid-based scoring, but also adds the change in solvation for the formation of the complex. The polar portion of the solvation term (GB) is calculated using the

formalism developed by Hawkins et al using parameters developed by Tsui et al (Hawkins GD et al., 1995; Hawkins GD et al., 1996; Tsui V and Case DA, 2000a). The nonpolar solvent accessible surface area (SA) is calculated using the icosahedra algorithm (Weiser J et al., 1999). This method also offers the option of incorporating the effect of salt screening into the polar term based on the Debye-Huckel limiting law for ion screening described in Srinivasan et al (J. Srinivasan et al., 1999). The PB/SA scoring function is an implementation of the Zap Tool Kit from OpenEye (Gilson MK et al., 1985; Jean-Charles A et al., 1991; Prabhu NV et al., 2004). Similar to the GB/SA function, this method calculates the VDW interaction between the ligand and receptor using the Lennard-Jones potential. The shielded electrostatics, on the other hand, are calculated using a dielectric function based on atom-centered Gaussians, which avoids the pitfalls of discrete dielectrics (Grant JA et al., 2001).

The last new scoring function, AMBER score, implements Molecular Mechanics GB/SA (MM GB/SA) simulations with traditional all-atom AMBER force field for protein atoms, and general AMBER force field for ligand molecules (Pearlman DA et al., 1995; Wang J et al., 2004). Unlike the methods above, this method calculates the energy terms for the entire AMBER force field, including bond, angle and dihedral terms for both the ligand and receptor, as well as Coulomb's Law and the Lennard-Jones potential as described above. The solvation energy is calculated using the Generalized Born (GB) solvation models: (1) Hawkins, Cramer and Truhlar model with parameters described by Tsui and Case, (2) Onufriev, Bashford and Case model, and (3) a modified version of

model (2) (Feig M et al., 2004; Hawkins GD et al., 1995; Hawkins GD et al., 1996; Onufriev A et al., 2000; Onufriev A et al., 2004; Tsui V and Case DA, 2000b). The surface area term is derived using a fast LCPO algorithm (Weiser J et al., 1999). Because the entire scoring function is calculated, a full thermodynamic cycle is employed (i.e., $\text{Score} = E_{\text{complex}} - (E_{\text{receptor}} + E_{\text{ligand}})$). Because the entire AMBER force field is available, minimization using conjugate gradient method is available in lieu of the simplex minimizer for the other scoring functions. In addition, Langevin molecular dynamics (MD) simulations at constant temperature can also be performed. As a result of both the new minimizer and the molecular dynamics capabilities, the AMBER score function now allows both ligand and receptor flexibility during minimization and MD simulations.

Development of Test Set

As stated above, there have been several docking studies on RNA-small molecule complexes (Barbault F et al., 2006; Detering C and Varani G, 2004; Lind KE et al., 2002; Moitessier N et al., 2006). Since these studies, though, a number of additional structures have been solved through both NMR and crystallographic techniques. To generate the test set used for this paper, we collected all structures that were labeled as containing both an RNA molecule and a ligand from the protein data bank (PDB). Crystal structures with resolution worse than 3.0 Å were removed to reduce the occurrence of experimental error. All structures solved by NMR were kept. This set of structures was then filtered

to remove all complexes where the ligands were either ions or artifacts of the structural determination method (eg, ethanol). To bias toward biologically relevant structures, all receptors with less than 15 bases were also removed. This procedure left a total of 68 structures.

In the next stage, the complexes were prepared for docking. The complexes with ligands containing cobalt and receptors with modified bases were removed as these chemistries were not available in our parameter set. Of the remaining complexes, ten had multiple active sites with multiple ligands bound. One complex (PDB code 2ET5) actually had four ligands bound to the same receptor, which led us to remove it from the test set due to nonspecific binding (Francois B et al., 2005). For all other complexes with multiple ligands bound, each site was treated as unique and prepared separately. Finally, once the test set was prepared for docking (see Methods: Test Set Preparation), each ligand was minimized while keeping the receptor rigid to detect complexes that were not stable in our scoring function. The ligands that moved more than 2 Å heavy-atom RMSD from the starting structure were removed from the set. The final set had a total of 46 structures (36 structures without the multiple binding sites) (Table 1).

PDB Codes for Test Set			
1AKX	1LVJ	1QD3	2BE0
1AM0	1MWL	1TOB	2BEE
1BYJ	1NEM	<i>1U8D</i>	2ESJ
1EHT	<i>1NTA</i>	1UTS	2ET3
<i>1F27</i>	<i>1NTB</i>	1UUD	2ET4
1FMN	<i>1NYI</i>	1UUI	2ET8
1FYP	1O15	<i>1XPF</i>	2F4S
1J7T	<i>1O9M</i>	<i>1Y26</i>	2F4U
1LC4	1PBR	1YRJ	2TOB

Table 1: List of PDB codes for all RNA-ligand complexes in test set. Codes in bold indicate complexes with multiple active sites, which were modeled independently. Codes in italics are structures determined by X-ray crystallography; others were determined using NMR.

METHODS APPLIED TO DATE

Test Set Preparation

A little less than half of the test set structures were determined using NMR, and thus deposited to the PDB as an ensemble of structures. To select a structure to use for docking, we first calculated the average structure from the coordinates of the entire ensemble. Because this structure was likely to be physically unrealistic, we then selected the structure in the ensemble that was closest to the average by RMSD. For the crystal structures, all solvent and counter ions were removed, and, if there was more than one location for a base, the highest occupancy was selected. Hydrogens were added and optimized in the presence of the bound ligand using the algorithm described in Moustakas et al. (Moustakas DT, 2004). AMBER parm99 partial charges were used for the bases (Cornell WD et al., 1995). Partial charges for receptor cofactors were calculated using the AM1BCC model, and formal charges were modified by hand (Jakalian A et al., 2000). The entire receptor preparation was completed using the Dock Prep procedure in Chimera (Moustakas D et al., in press; Pettersen EF et al., 2004).

To identify the active site for docking, all hydrogens were removed from the receptor, and the molecular surface was calculated using the dms program (Richards FM, 1977). Spheres were then created along the molecular surface to create a negative image of the features. All spheres within 10 Å of any atom of the bound ligand were selected, resulting in an average of 123 ± 22 spheres per active site. Different distances from 1 – 10 Å from the ligand were explored. We

found that there was no change in the success rate using anything between 1 and 10 Å and only minimal changes in the energy (-98.9 ± 2.5 kcal/mol) and length of calculation (1207 ± 3 seconds). We therefore selected the 10 Å radius from the ligand for historical purposes to compare with the results from the protein test set (Moustakas D et al., in press).

Next, to account for the receptor contribution to the score during docking, grids that store the VDW and electrostatic values for the receptor were calculated. The dimensions of the grids were calculated by padding the selected spheres by 5 Å using the accessory SHOWBOX. Default parameters, except for the energy_cutoff_distance set to 9999, were used in the GRID accessory program to calculate the data for the grids themselves. The final grids averaged $\sim 40,000$ Å³ in volume.

The ligands were prepared using two different techniques. For the sampling parameter optimization, the ligands were prepared once again using the Dock Prep module in Chimera using the AM1BCC charge model (Jakalian A et al., 2000; Pettersen EF et al., 2004). For the comparison and evaluation of the scoring functions, only ligands with 12 or less rotatable bonds were prepared (see Results and Discussion of Current Status: Flexible Ligand Sampling Optimization). The ligands were first run through the standard ZINC database preparation (Irwin JJ and Shoichet BK, 2005). In this method, the three-dimensional files are converted to canonical SMILE strings. Then, an energy-minimized three-dimensional structure is generated, and the stereochemical centers are expanded. Because it is not clear what the local pH of the RNA

active sites might be, the protonation states for each ligand were also expanded between pH 4.5 - 9.5. The current version of ZINC does not allow for monoanionic protonation states for phosphate groups, so these compounds were protonated by hand. Once the ensembles of stereochemistry and protonation states were generated for each ligand, Gasteiger-Hückel and AM1-BCC partial charges were calculated using the ANTECHAMBER version in DOCK, AMSOL partial charges were calculated using the AMSOL program, and RESP partial charges were calculated using the ANTECHAMBER accessory in AMBER 8 (Bayly CI et al., 1993; Gasteiger J and Marsili M, 1980; Hawkins GD et al., 2004; Jakalian A et al., 2000; Wang J et al., 2006; Wang J et al., 2004).

Modification of Pruning Algorithm

The ligand flexibility sampling algorithm is an incremental construction method called anchor-and-grow. In this method, the ligand is first divided into the largest rigid portion and layers of flexible regions (see Figure 1 for an example). The largest rigid portion of the ligand, or anchor, is identified and then oriented in the active site and minimized. All orientations with scores above 1000 kcal/mol were removed and the remaining ranked by score and then clustered by RMSD using a greedy algorithm (pruning) (Moustakas D et al., in press). One layer of flexible bonds is then grown from each cluster, minimized, ranked, and clustered again. The growth phase is repeated until the molecule is fully built. In a previous study, we had shown that the pruning portion of the algorithm was limiting sampling during growth and preventing energy convergence

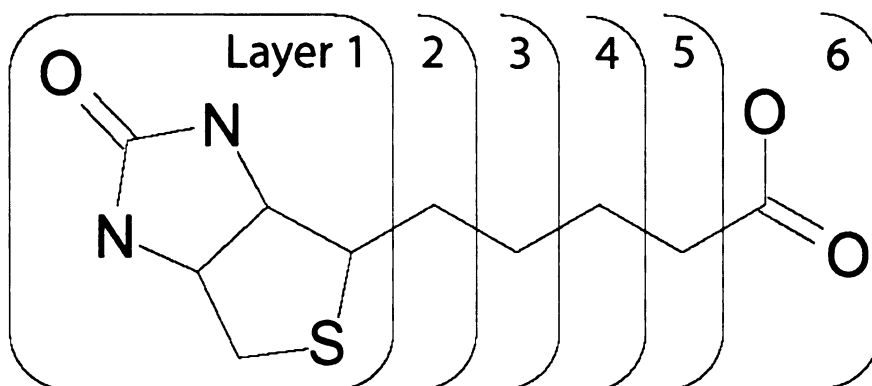


Figure 1. Diagram of identification of rigid anchor (layer 1) and flexible layers for growth.

(Moustakas D et al., in press). Therefore, while we were optimizing the sampling parameters for flexible ligands for the RNA test set, we also reworked the algorithm.

After comparing several clustering algorithms, it was determined that the clustering itself was limiting the sampling. In an ideal pruning algorithm for anchor orientation, the anchors that are close to the location of the experimentally determined ligand should be moved on to growth. With clustering by root mean squared deviation (RMSD), however, all anchors that are close to the correct position will fall into the same cluster. Thus, in the worst case, only one of these anchors would be propagated to the next stage even though several were generated by the sampling algorithm. Instead of clustering, we found that using a simple scoring cutoff of 25 kcal/mol and a hard limit of 100 orientations increased the number of orientations near the active site that passed pruning for 56% of the test set (Table 2). Applying the new pruning algorithm and parameters to the growth cycles as well improved the success rate of flexible docking to match that of rigid docking (Figure 4 and Discussion: Comparison to Older Versions of DOCK).

Scaling of Repulsive VDW Radii

In our previous study, we found that flexible sampling failures often occurred as a result of minor clashes between the ligand and the protein receptor (Moustakas D et al., in press). Because all but one of these failed ligands could successfully be docked using rigid sampling, we hypothesized that those failures

were the result of the flexible algorithms inability to build around the clashes due to the overly coarse sampling. Due to this finding, we then hypothesized that reducing the repulsive portion of the VDW term during docking and then performing a final minimization under the standard scoring function would reduce this problem. Many physics-based force fields address this type of problem by reducing the exponent of the repulsive term in the Lennard-Jones Potential. However, modifying the equation this way actually changes the overall shape of the function, which was used to generate all the parameters in the force field (Figure 2). Since we would like to retain the integrity of the force field as much as possible, we have instead scaled the radius of each atom used for the repulsive term only. This scaling factor shifts the graph to the left but maintains the overall shape. In more practical terms, the energies are less affected, but the new function has a huge impact on the number of anchors generated within 1.5 Å heavy atom RMSD of the crystal structure (Figure 4 and Discussion: Flexible Ligand Parameter Optimization).

Modification of Bump Filter

In previous versions of DOCK, the bump filter could be applied to filter orientations that significantly overlap with receptor atoms before minimization. Because minimization is the most time-consuming portion of the calculation, this filtering helps to increase the speed of the calculation by directing sampling toward more productive routes. Here, we showed that implementation of the bump filter in DOCK 5 is significantly slower than earlier versions. To begin to

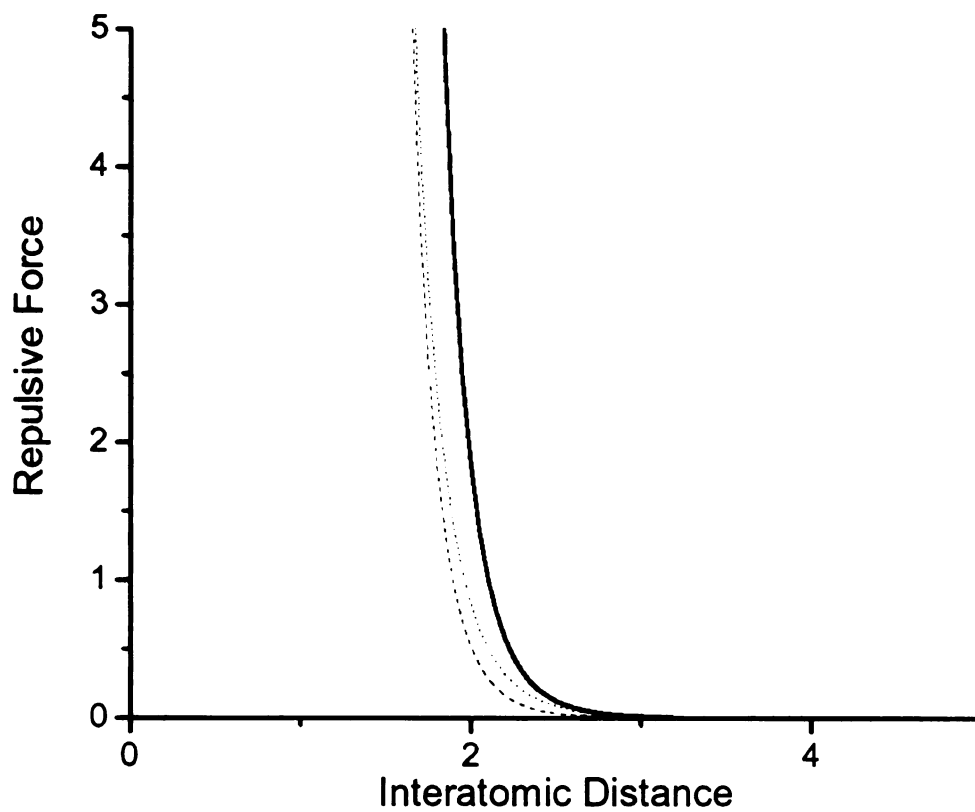


Figure 2. Influence of different methods of softening the repulsive force for the VDW term. The force is reduced either by reducing the exponent of the repulsive term to 10 (dot) or by scaling the radii for the repulsive force back 10% (dash). The graph is calculated using the radius and well depth for an aliphatic hydrogen.

address this problem in the newer version of DOCK, we have implemented the same filtering by bump during growth, in this case between torsional sampling and minimization. Because the number of atoms in the anchor is much larger than in each flexible layer, we have added user parameters to separately control the maximum number of bumps allowed for both the anchor and the growth stages.

Optimization of Parameters for DOCKing

Because the chemical characteristics of the RNA test set are different from our protein test set, both due to the differences between proteins and nucleic acids and due to the differences in the make-up of the ligand set, we reoptimized the sampling parameters for both rigid and flexible ligand docking (see Appendix). All docking experiments were carried out on 2.2 GHz dual processor Opteron 828s running Linux Fedora Core 3. The code was compiled using open-source GNU compilers (<http://www.gnu.org>). We note that our primary criterion for optimization was success in finding the proper ligand geometry and not the CPU time required per compound. Unless otherwise stated, these parameters were used for all experiments in this paper.

The final version of the new DOCK code, including all functions described in this paper, was posted to the DOCK web site as version 6.1.0 (<http://dock.compbio.ucsf.edu>). All experiments performed with the new implementation of DOCK used this version and will be referred to as DOCK 6 for

convenience. All experiments performed with the previous versions of DOCK used version 4.0.1 and version 5.4.0 and will be referred to as DOCK 4 and DOCK 5, respectively.

RESULTS AND DISCUSSION OF CURRENT STATUS

Rigid Ligand Sampling Optimization

The parameters listed in Table 2 control the sampling during rigid ligand docking. We selected these parameters based on a series of rigid ligand docking experiments in which we monitored both the success rate, where success is defined as the top-ranking orientation being within 2 Å heavy atom RMSD of the experimental structure, and DOCK scores averaged over the entire test set. When we explored a range of values for the maximum number of orientations (`max_orientations`), or matches between the ligand and spheres, in conjunction with the maximum number of minimization steps (`simplex_max_iterations`), the DOCK score decreased rapidly as the number of orientations and the amount of minimization increased and then eventually plateaued (Figure 3). The success rate did not converge as well. However, we attribute this instability to the inherent shallow, bumpy energy surface of RNA. We selected values that balanced between the length of time for each calculation and convergence of the DOCK score—500 orientations and 1000 minimization steps—as optimal. Throughout the optimization procedure, all other parameters were kept the same as those optimized for DOCK 5 (Moustakas D et al., in press).

Parameter Name	Parameter Description	Value
max_orientations	The number of ligand poses sampled by the rigid orienting algorithm	500
simplex_max_iterations	The maximum number of simplex iterations	1000
If Bump Filter is used		
max_bumps	Maximum number of clashes between ligand and receptor	12

Table 2. Description of and optimized default values for parameters that affect rigid ligand docking for RNA test set

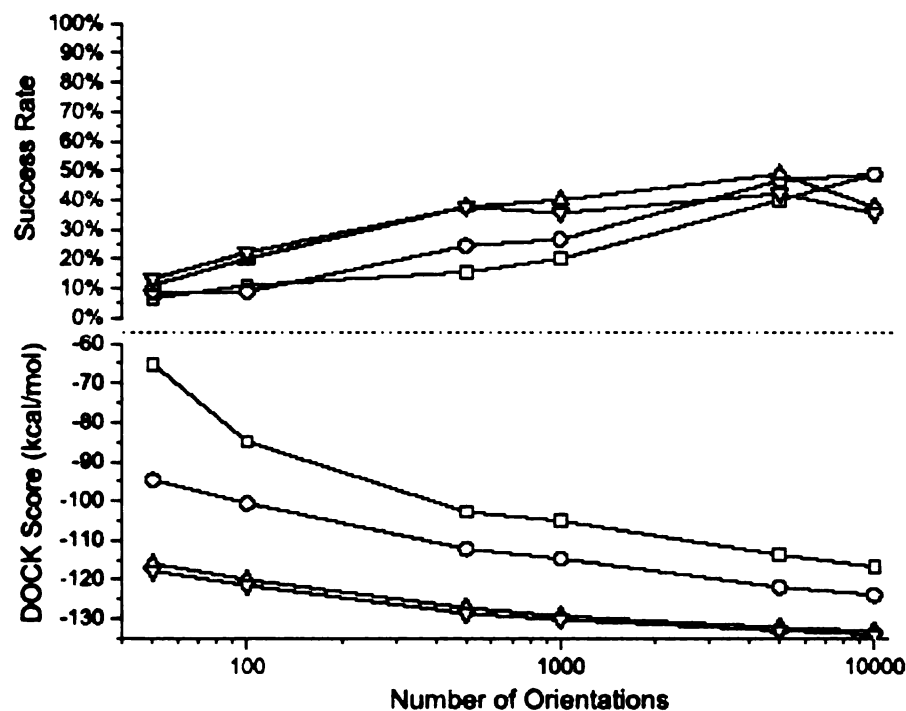


Figure 3. Optimization of parameters for rigid ligand docking. Parameters of 50 (\square), 100 (\circ), 1000 (\triangle), and 10000 (∇) minimization steps (simplex_max_iterations) are examined as a function of the number of orientations (max_orientations).

Once the sampling parameters were optimized, we then explored the use of the bump filter to speed up the dock runs. Using the sampling parameters from above, we explored a range of bump cutoffs (`max_bumps`), or the maximum number of heavy atom clashes allowed for a particular orientation, and monitored the binding pose success rates (data not shown). The optimal value — 12 bumps — was selected to be the fewest number of bumps that recreated the same success rate as without the bumps.

Flexible Ligand Sampling Optimization

For the more complex flexible ligand algorithm, the parameter optimization was performed first on the anchor docking, and the best parameters were then used for optimizing the growth. The parameters that control the sampling in both these steps are listed in Table 3. As for rigid ligand docking, the parameters that control step sizes for the simplex minimizer were set to the previously defined optimal values.

The first step in the anchor-and-grow algorithm is ring identification or anchor segmentation. All bonds within molecular rings are treated as rigid. To treat sugar puckering and chair-boat hexane conformations, we have left the rings as they were in the experimental structure. If the molecule does not have a ring, the largest rigid segment is specified as the anchor. For this stage of the optimization, all runs used the default of the largest anchor only. If the molecule had multiple anchors of the same size, the first anchor on the anchor list was used. Once the anchor had been identified, the parameters that control the

number of anchor orientations (`max_orientations`) and the number of anchor minimization steps (`simplex_anchor_max_iterations`) were explored. Because the anchors are substructures of the ligand, the parameter convergence was monitored as a function of the RMSD between the anchor orientation and the corresponding substructure of the experimental ligand averaged over all generated orientations before pruning. When the number of anchor orientations and minimization steps were varied systematically, the number of minimization steps converged to 1000 for ligands with more than twelve rotatable bonds (filled shapes) (Figure 4a). For the ligands with less than twelve rotatable bonds (open shapes), we found the same lack of convergence issue as we did for the rigid ligand docking. Once again, we attribute this problem to the roughness of the RNA energy landscape.

The anchor pruning parameters were comparatively easy to optimize. Based on the DOCK 5 studies, we chose the maximum number of anchors allowed to the growth phase (`pruning_max_orients`) to be 100 (Moustakas D et al.). As stated in the Methods Section, the perfect pruning function would maximize the number of anchors close to the experimental data while filtering out those further away. Therefore, using the parameters optimized above, we selected the scoring cutoff (`pruning_orient_score_cutoff`) for the orientations — 25 kcal/mol — to be the most negative DOCK score that did not affect the percent of anchor orientations that were within 1.5 Å of the experimental for each individual structure in the test set (data not shown).

Parameter Name	Parameter Description	Value
min_anchor_size	The minimum number of heavy atoms in a rigid segment to be considered an anchor	40 ^a 5 ^b
max_orientations	The number of ligand poses sampled by the rigid orienting algorithm	1000
pruning_max_orients	The maximum number of orientations after ranking by score to be carried to the growth phase	100
pruning_orient_score_cutoff	The DOCK score cutoff above which orientations are not carried on to the growth phase	25 kcal/mol
pruning_max_conformers	The maximum number of conformations after ranking by score to be carried to the next growth phase	75
pruning_conformer_score_cutoff	The DOCK score cutoff above which conformations are not carried on to the growth phase	25 kcal/mol ^a 1000 kcal/mol ^b
simplex_anchor_max_iterations	The maximum number of simplex iterations for each orientation	500
simplex_growth_max_iterations	The maximum number of simplex iterations for each conformation	500
simplex_final_min_max_iterations	The maximum number of simplex iterations for the fully grown set of conformations	500
If Grid Score is used		
grid_score_rep_rad_scale	Scaling value for VDW repulsive radius during docking calculation	0.9
simplex_final_min_rep_rad_scale	Scaling value for VDW repulsive radius during final minimization	1
If Bump Filter is used		
max_bumps_anchor	The maximum number of clashes between each orientation and receptor	12
max_bumps_growth	The maximum number of clashes between conformation and receptor	8

Table 3. Description of and optimized values for parameters that affect flexible ligand docking for RNA test set

(a) Sampling parameter for ligands with less than 12 rotatable bonds (b) Sampling parameters for ligands with more than 12 rotatable bonds

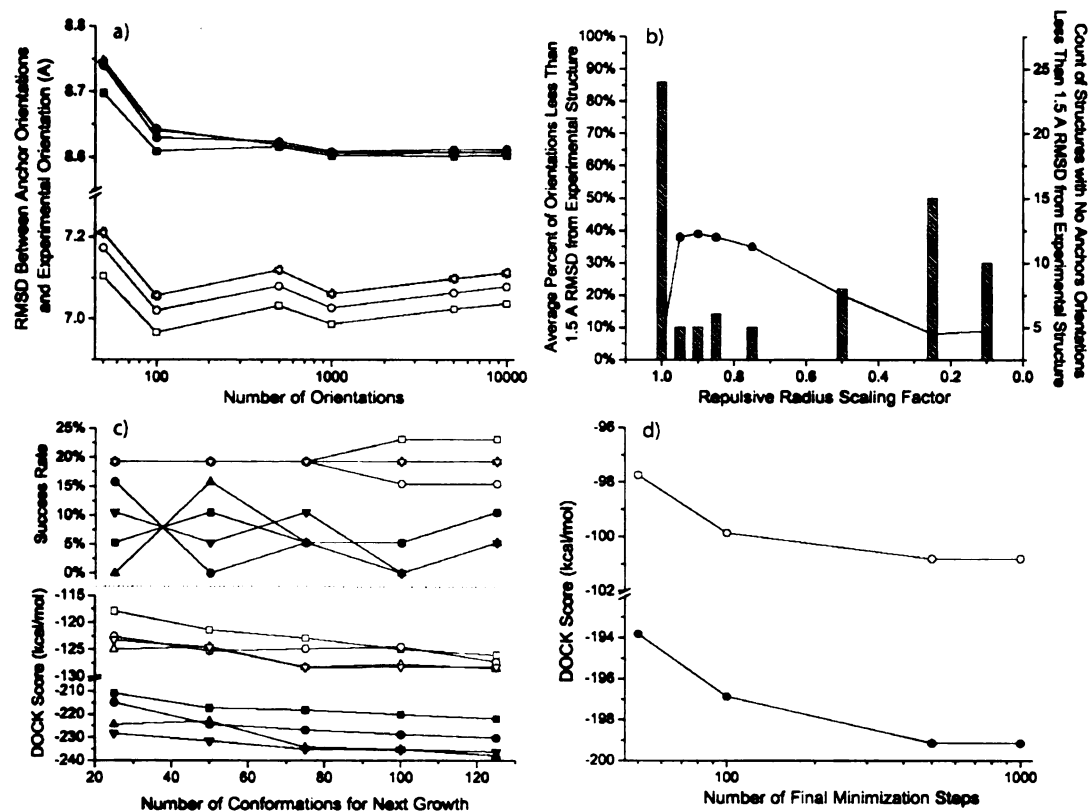


Figure 4. Optimization of parameters for flexible ligand docking. Values for ligands with less and more than twelve rotatable bonds are open and filled, respectively for a), b) and d). (a) Parameter optimization for anchor sampling portion of flexible ligand docking. Parameters of 50 (\square), 100 (\circ), 500 (\triangle), and 1000 (∇) anchor minimization steps (`simplex_anchor_max_iterations`) are plotted as a function of the number of orientations (`max_orientations`). (b) Optimization of scaling factor for atomic radii for repulsive VDW term (`grid_score_rep_rad_scale`). The average percent of orientations that pass the pruning filter that are within 1.5 Å of the experimental structure (line), as well as the count of structures with no orientations near the active site (bar), are plotted as a function of scaling factor. (c) Parameter optimization for growth sampling portion of flexible ligand docking. Maximum number of conformations allowed to the next stage of growth (`pruning_max_conformers`) of 25 (\square), 50 (\circ), 75 (\triangle), and 100 (∇) are plotted as a function of the number of growth minimization steps (`simplex_grow_max_iterations`). (d) Parameter optimization for final minimization (`simplex_final_max_iterations`) after docking is complete.

Finally, once the sampling parameters had been fully optimized, we used them to explore scaling the atomic radius used to calculate the repulsive VDW term (`grid_score_rep_rad_scale`). Once again, we explored a range of scaling factors, monitoring both the percent of orientations that pass the pruning that are within 1.5 Å heavy atom RMSD of the experimental structure averaged over the entire test set (line) as well as counts of the number of complexes in which no orientations were generated within the distance cutoff (bar) (Figure 4b). The percentage of orientations close to the experimental structure increases sharply with only minimal scaling, quickly peaks at the optimal value of 0.9 and then slowly tails off. We hypothesize that the left side of the graph accounts for differences in the various force fields used to determine the experimental structures. The less dramatic tailing on the right side of the graph simply emphasizes the loss of balance between the repulsive term and attractive term. The optimal scaling value of 0.9 was also applied for repulsive radii in the growth phase (see below).

The next step in the anchor-and-grow algorithm is flexible bond identification. Each flexible bond is associated with a label defined in the `flex_definition_file` parameter. Each label in the file contains a definition based on the atom types and chemical environment of the bonded atoms as well as a set of preferred torsion positions. Typically, bonds with some degree of double bond character are excluded from minimization so that planarity is preserved. For convenience, the scoring cutoff for the conformations (`pruning_conformer_score_cutoff`) was set to 25 kcal/mol, which is the same

cutoff for the anchor score. Using the optimal anchor and pruning parameters, we varied the number of minimization steps for each layer of growth (`simplex_growth_max_iterations`) and the cutoff of number of conformers to be carried on to the next phase of growth (`pruning_max_conformers`). We return to using a combination of the score for the top ranking pose averaged over the entire test set and the success rate to monitor convergence. As with rigid ligand docking, the DOCK scores eventually converge for both subsets of the test set (Figure 4c). We selected the converged values — 500 minimization steps and the cutoff for the number of conformers for the growth section as 75 — as optimal.

For the more flexible ligands, it was discovered in the process of optimizing the sampling parameters that, in some cases, no conformations could be generated. We determined that all molecules could be built if the scoring cutoff for the conformations (`pruning_max_conformers`) was raised to 1000 kcal/mol and if all anchors with five or more atoms were tried. However, this change in parameters greatly increased the length of the calculation time. Therefore, we split the test set into two segments: those with less than 12 rotatable bonds and those with more than 12. While the parameters optimized above were still used for the ligands with less than twelve rotatable bonds, these more permissive parameters were used for the more flexible ligands for optimization of the final parameter.

Because the radii for the repulsive term for the Lennard-Jones potential were scaled throughout the docking process, the full set of grown conformations

was re-minimized using the unscaled scoring function (`simplex_final_min_rep_rad_scale`) and then reranked to determine the final pose. We explored a range of the number of simplex steps (`simplex_final_max_iterations`), and the score for the entire test set converged at 500, which we selected as the optimal parameter (Figure 4d). This final minimization and reranking step increased the success rate from 19% to 27% for the subset with less than twelve rotatable bonds. While there was no effect on the success rate of 5% for the subset with more than twelve rotatable bonds, the location of each of the ligands changes by an average of 6.7 Å heavy atom RMSD.

For reasons explained below (see Success as a Function of the Number of Rotatable Bonds), we only optimized the bump filter parameters for flexible docking for the set of ligands with less than twelve rotatable bonds. For convenience, the maximum number of bumps for the anchor (`max_bumps_anchor`) was set to be the same as the optimized value for rigid docking — 12 bumps. The maximum number of bumps for growth was explored over a range of values. Once again, the optimal parameter was selected as the minimum number of bumps that reproduced the docking success rate without the filter (data not shown). For ligands with less than 10 rotatable bonds, the optimal number of bumps was set to 6. For ligands with between 10 and 12 rotatable bonds, the optimal parameter was 10. We hypothesize that the difference in this parameters is actually a result of the nature of the more floppy molecules. The molecules with more rotatable bonds are small aminoglycosides and many of the

flexible layers are ring systems, which result in more clashes with the active site during the coarse torsional sampling.

Success as a Function of the Number of Rotatable Bonds

In a drug design effort for targeting a protein, the number of rotatable bonds is typically limited to eight to ten (Lipinski CA, 2000; Wunberg T et al., 2006). This limit is used to reduce the loss of entropy of the ligand, as well as to reduce desolvation penalties, upon binding. It also increases the possibility that the molecule will be bioavailable as a drug. However, using aminoglycosides as an example, this rotatable bond cutoff does not appear to apply to ligands that bind to nucleic acids. Because we include aminoglycosides, our test set covers a wide range of rotatable bonds. We examined the ability of DOCK to reproduce the experimental binding pose within 2 Å heavy atom RMSD with flexible ligand docking (Figure 5). Here, the success rate drops off dramatically after zero rotatable bonds--there are no ligands in the test set with between one and four bonds--and then levels off after twelve bonds. However, when looking at the average length of the calculation, the less floppy molecules' time increases relatively linearly, whereas the additional sampling required for the ligands with more than 12 rotatable bonds results in a huge increase. Because the length of time is so much greater for the larger molecules with such little return in success rate, we will only be using the subset of the test set with less than twelve rotatable bonds for the remainder of this chapter. In addition, we will be subdividing this smaller set into a) all compounds with less than seven rotatable bonds (L7) to enable direct comparison with our protein test set (total of 10

complexes), b) all compounds with less than ten rotatable bonds (L10) to indicate the types of success possible for drug-like molecules (total of 17 complexes), and c) all compounds with less than twelve rotatable bonds (L12 — total of 26 complexes).

Comparison Between Versions of DOCK

To compare the performance with previous versions of DOCK, we repeated both the rigid and flexible docking experiments both with and without the bump filter using DOCK versions 4 and 5 (Table 4). For both types of rigid ligand docking, the success rate and energy score were identical for versions 5 and 6. The times differ very slightly, with version 6 performing a bit faster for all three test subsets. We anticipated this result, because we have not made any changes in the rigid docking algorithm in the new release. The bump filter increased the speed for the calculations by approximately 5-fold for all subsets. The calculations for version 4 have similar energies and run faster than newer versions both with and without the bump filter. In this case, the bump filter only improved the docking speed by 2- to 3-fold. However, because the length of the calculation is so fast without the bump filter, the application is a bit superfluous.

The success rates without the bump filter for version 4 were the same as versions 5 and 6. However, one of the top-scoring poses had atoms that extended beyond the edge of the grid. In version 4, if atoms extend beyond the edge of the grid, the code simply loops over the atoms, only including the atoms inside the grid. As a result, if the atoms outside the grid are clashing with the

Method	L7 Subset			L10 Subset			L12 Subset		
	Score	Success	Time	Score	Success	Time	Score	Success	Time
	Rigid Ligand Docking								
DOCK 4.0.1	-57.75	50% ^a	4.4	-74.72	41% ^a	8.4	-96.69	38% ^a	12.7
DOCK 5.4.0	-58.76	50%	25.7	-76.89	41%	38.3	-100.35	38%	50.3
DOCK 6.1	-58.76	50%	24.1	-76.89	41%	36.9	-100.35	38%	48.9
	Rigid Ligand Docking with Bump Filter								
DOCK 4.0.1	-57.20	50%	2.2	-75.46	41%	3.2	-97.11	31%	4.1
DOCK 5.4.0	-57.04	60%	5.1	-74.41	41%	7.6	-97.15	38%	10.4
DOCK 6.1	-57.04	60%	5.4	-74.41	41%	7.9	-97.15	38%	10.5
	Flexible Ligand Docking								
DOCK 4.0.1	-76.16	50% ^b	8.7	-93.13	29% ^c	13.4	-117.78 ^d	23% ^{d,e}	17.5
DOCK 5.4.0	-64.65	50%	298.2	-84.43	35%	1099.9	-113.04	27%	2437.7
DOCK 6.1	-59.82	60%	328.0	-78.78	41%	656.5	-100.82	27%	1255.0
	Flexible Ligand Docking with Bump Filter								
DOCK 4.0.1	-84.57	40% ^b	7.7	-101.00	24% ^f	12.3	-118.68 ^d	19% ^{d,g}	16.3
DOCK 5.4.0	-64.54	40%	298.1	-84.55	29%	1099.2	-113.44	23%	2336.4
DOCK 6.1	-58.71	60%	252.9	-77.84	41%	528.0	-101.36	27%	1149.5

Table 4. Average score (kcal/mol), success rate (measured as percent of complexes in test set where best scoring pose is within 2 Å heavy atom RMSD from experimental structure), and length of calculation (sec) for various subsets of test set using different versions of DOCK. (a) One complex outside the grid box. (b) Three complexes outside the grid box. (c) Five complexes outside the grid box. (d) No conformations could be generated for one compound. (e) Ten complexes outside the grid box. (f) Four complexes outside the grid box. (g) Seven complexes outside the grid box

receptor, this information is not included in the calculation. While not necessarily a problem for binding pose prediction—the pose outside the box was in this case was a failure—these poses may create difficulties when reranking with more advanced scoring functions without grids. In addition to clash checking, the bump filter performs a quick check to filter out all anchor orientations that are outside the box before minimization, so this off-the-edge issue has been resolved for version 4 with the bump filter. However, one can tell the sampling search has been limited as well for version 4, because the success rate is lower for both subset L7 and L12. For versions 5 and 6, a hard cutoff that does not score or minimize anchors that extend beyond the grid box, even without the bump filter, has been applied.

For flexible docking, the average scores for all version 4 subsets are the lowest and the length of the calculation is the shortest. However, between three and ten of the ligands in each subset extend beyond the edge of the grid, which influences these scores. In version 4, because the bump filter only checks if the anchors are off the edge of the grid, the growth and minimization phases can still result in a final compound with portions beyond the grid box. More importantly, though, the success rates for all of the subsets are lower for version 4 than for versions 5 and 6, indicating that the sampling for the oldest implementation of flexible docking is not adequate for these purposes. The average scores for version 6 are the least negative, but the success rates are the highest across all the subsets, attaining the same rates as rigid docking for both the L7 and L10 subsets. We hypothesize that this improvement in success can be attributed to

scaling of the radii for the repulsive portion of the VDW component of the score. This scaling guides the sampling algorithm to identify the correct pose, which is often modeled by experimentalists a bit too close to the active site in tight-binding ligands, while avoiding the other traps on the surface of the receptor.

In comparing the length of the calculation for version 5, the bump filter does not seem to make much of a difference except for the L12 subset. This lack of effect indicates that anchor orientation is not a time-consuming portion of the algorithm. Profiling experiments demonstrated instead that the biggest time sink in the calculation is actually in minimization (data not shown). Therefore, assuming the minimizer is as optimized as possible, the best way to reduce the length of the calculation would be to reduce the number of calls to the minimizer. For version 6, the modified pruning algorithm serves to limit the number of conformations for each stage of growth and minimization, thus making the calculation twice as fast for both the L10 and L12 subsets. We also applied the bump filter and outside-the-box check between torsional sampling and minimization, resulting in an additional increase in speed. Even with these changes, though, the version 6 calculations are still significantly slower than version 4, indicating room for additional improvement.

Examination of Ensemble of Generated Orientations

When rescored with more advanced scoring functions, we hypothesize that reranking some portion of the current ensemble will assist in the correct binding pose rising to the top of the ranking list. However, because these scoring

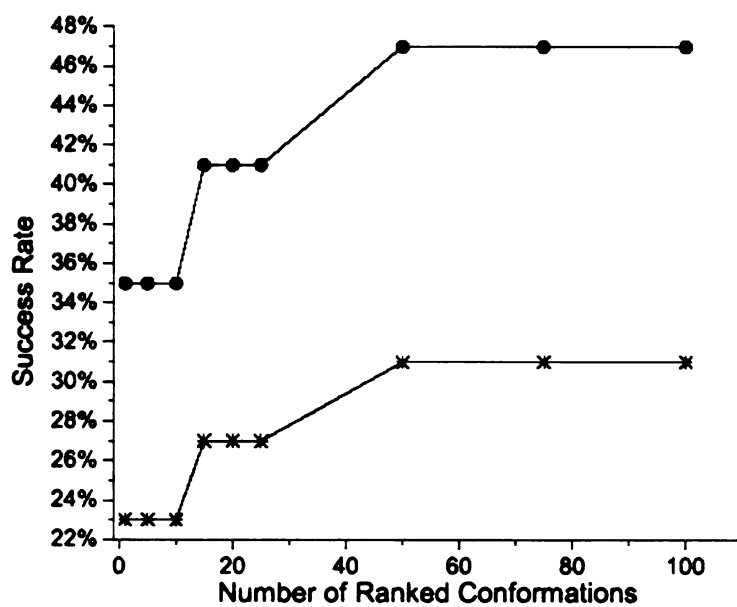


Figure 5. Success rate (closed) and average length of calculation (open) as a function of the number of rotatable bonds in the ligand for flexible ligands. Both the rate and time are calculated as a cumulative average over the test set.

functions take a nontrivial amount of time to rescore these poses, we want to reduce the number of poses for rescoring. To this end, we looked at the entire list of conformations that were generated by the sampling above to determine a reasonable cutoff for rescoring (Figure 5). For the L7 subset, there is no improvement by looking at the entire ensemble. However, for both the L10 and L12 subsets, the success rate is improved if up to fifty conformations are considered.

CONCLUSIONS

Up to this point, we have optimized the sampling parameters of the newest version of the DOCK suite of programs. Compared to the optimization procedures for DOCK 5, the sampling parameters are very similar to those for the protein test set for the rigid docking (Moustakas D et al., in press). With minor modifications to the algorithm, we have additionally improved the ability to reproduce experimental binding poses for flexible ligand docking. However, the success rates are still not as high as those for proteins. We hypothesize that this failure is at least in part due to the shallow and bumpy nature of the energy landscape of RNA molecules in comparison to proteins. We have once again fully explored the sampling algorithm and plan to explore improvements in the scoring function next. In general, we have found that DOCK can be successfully employed for binding mode prediction for RNA-ligand complexes and, even with the rudimentary grid-based scoring, should be useful in the drug design setting.

FUTURE DIRECTIONS

At this point, we have completed the optimization and characterization of the sampling portion of the DOCK algorithm for RNA. The next step will be to explore the effect of the newly added scoring functions on the success rates for binding mode prediction. In the drug design setting, ligands are not available in the ideal conformation for binding as they are in this structure-based test set. Therefore, at this point, we will switch to using the ligand sets prepared using the ZINC procedure (see Methods). The DOCK 3.5 scoring function runs at approximately the same speed as the Grid Score scoring function used for sampling parameter optimization. Therefore, the first scoring method test will be to repeat the flexible docking with the bump filter using the DOCK 3.5 and the Grid Score scoring functions. The time for calculation, success rate based on best scoring pose, and the distribution of conformations, both with and without clustering, will be analyzed. For the DOCK 3.5 scoring function, because the more advanced receptor-ligand desolvation method has not been tested on nucleic acids, we will only be adding desolvation for the ligand. For both procedures, we will also compare the Gasteiger, AM1BCC, AMSOL and RESP charge models for the ligands. Finally, throughout this process, we plan to run the same procedures on the protein test set used to validate DOCK version 5. The sampling parameters for the protein test set were very similar and should produce comparable results in DOCK 6. However, the secondary scoring functions have not yet been explored and would provide a useful frame of reference for the current state of success rate for the RNA set.

Because the more advanced GB/SA and PB/SA scoring functions take significantly more time to compute each point energy than either Grid or DOCK 3.5 score, they can only practically be used for rescoring previously generated poses. Therefore, once the comparison between the Grid and DOCK 3.5 scoring functions is complete, the top N scoring conformations or clusterheads will be rescored using the GB/SA, PB/SA, and AMBER scoring functions. The value of N will be chosen as a balance between the number of conformations to be rescored, which determines the length of the calculation, and the possibility of a successful pose being found in the list. For the GB/SA function, we will also explore the effect of various salt concentrations on the docking success rates. For the AMBER scoring function, we will use the sampling parameters and procedure developed in the Case lab (unpublished communication). Once these evaluations are complete, we will once again compare the success rate and length of time for the calculations, as well as the effect of partial charge models.

Within this test set, we have several ensembles of similar structures. The first type of ensemble — NMR structures — is the same ligand and receptor in different conformations. The second type is multiple structures of the same receptor with different ligands bound. We would like to perform cross-docking experiments to both of these types of ensembles to get an idea of the use of conformational expansion in RNA docking. In addition, there is an open question about the effect of minimization of the receptor-ligand complex before docking on both the docking results and the rescoring results. We plan to take a subset of the test set and explore if this minimization further improves the success rate by

resolving internal clashes in both the ligand and receptor, optimizing the interaction between the ligand and receptor before docking, and regularizing the structure if it was resolved using another scoring function. Finally, we have a collaborator who is screening a library of approximately 5000 small molecules against both BIV TAR RNA and HIV RRE RNA. Once the optimal procedure for binding pose prediction has been identified, we will DOCK this library and compare our ability to rank a library of small molecules to the experimental values.

Once all of the above is complete, we plan to go through and characterize where we are still failing. In this study, we have explored a wide variety of variables that improved the success rates. Here, we will reflect on what is still not working and try to make some useful suggestions for the most constructive ways to further improvement.

REFERENCES

Bannwarth, S and Gatinol, A. *Curr HIV Res.* **3** (2005) 61-71.

Barbault, F, Zhang, LR, Zhang, LH and Fan, BT. *Chemom Intell Lab Syst.* **82** (2006) 269-275.

Bayly, CI, Cieplak, P, Cornell, WD and Kollman, PA. *J Phys Chem.* **97** (1993) 10269-10280.

Cornell, WD, Cieplak, P, Bayly, CI, Gould, IR, Merz, KM, Ferguson, DM, Spellmeyer, DC, Fox, T, Caldwell, JW and Kollman, PA. *J Am Chem Soc.* **117** (1995) 5179-5197.

Detering, C and Varani, G. *J Med Chem.* **47** (2004) 4188-4201.

Feig, M, Onufriev, A, Lee, MS, Im, W, Case, DA and Brooks, CL. *J Comput Chem.* **25** (2004) 265-284.

- Francois, B, Russell, RJ, Murray, JB, Aboul-ela, F, Masquida, B, Vicens, Q and Westhof, E. *Nucleic Acids Res.* **33** (2005) 5677-90.
- Frank, J and Spahn, CMT. *Rep Prog Phys.* **69** (2006) 1383-1417.
- Frankel, AD and Young, JA. *Annu Rev Biochem.* **67** (1998) 1-25.
- G.D. Hawkins, D. J. Giesen, G. C. Lynch, C. C. Chambers, I. Rossi, J. W. Storer, J. Li, T. Zhu, J. D. Thompson, P. Winget, B. J. Lynch, D. Rinaldi, D. A. Liotard, C. J. Cramer and Truhlar, DG. "AMSOL 7.1." (2004). University of Minnesota.
- Gasteiger, J and Marsili, M. *Tetrahedron.* **36** (1980) 3219 - 3288.
- Gilson, MK, Rashin, A, Fine, R and Honig, B. *J Mol Biol.* **184** (1985) 503-16.
- Grant, JA, Pickup, BT and Nicholls, A. *J Comput Chem.* **22** (2001) 608-640.
- Hawkins, GD, Cramer, CJ and Truhlar, DG. *Chem Phys Lett.* **246** (1995) 122-129.
- Hawkins, GD, Cramer, CJ and Truhlar, DG. *J Phys Chem.* **100** (1996) 19824-19839.
- Hawkins, GD, Giesen, DJ, Lynch, GC, Chambers, CC, Rossi, I, Storer, JW, Li, J, Zhu, T, Thompson, JD, Winget, P, Lynch, BJ, Rinaldi, D, Liotard, DA, Cramer, CJ and Truhlar, DG. "AMSOL 7.1." (2004). University of Minnesota.
- Irwin, JJ and Shoichet, BK. *J Chem Inf Model.* **45** (2005) 177-182.
- J. Srinivasan, M. W. Trevathan, P. Beroza and Case, DA. *Theor Chem Acc.* **101** (1999) 426-434.
- Jakalian, A, Bush, BL, Jack, DB and Bayly, CI. *J Comp Chem.* **21** (2000) 132-146.
- Jean-Charles, A, Nichols, A, Sharp, K, Honig, B, Tempczyk, A, Hendrickson, TF and Still, WC. *J Am Chem Soc.* **113** (1991) 1454-1455.
- Johansson, D, Jessen, CH, Pohlsgaard, J, Jensen, KB, Vester, B, Pedersen, EB and Nielsen, P. *Bioorg Med Chem Lett.* **15** (2005) 2079-2083.
- Lind, KE, Du, Z, Fujinaga, K, Peterlin, BM and James, TL. *Chem Biol.* **9** (2002) 185-93.
- Lipinski, CA. *J Pharmacol Toxicol Method.* **44** (2000) 235-249.

- Luty, BA, Wasserman, ZR, Stouten, PFW, Hodge, CN, Zacharias, M and McCammon, JA. *J Comp Chem.* **16** (1995) 454-464.
- Mayer, M and James, TL. "Discovery of ligands by a combination of computational and NMR-based screening: RNA as an example target." In Nuclear Magnetic Resonance of Biological Macromolecules, Part C. Ed. (2005) 571-587.
- Mayer, M, Lang, PT, Gerber, S, Madrid, PB, Pinto, IG, Guy, RK and James, TL. *Chem Biol.* **13** (2006) 993-1000.
- Meng, EC, Shoichet, BK and Kuntz, ID. *J Comput Chem.* **13** (1992) 505-524.
- Moitessier, N, Westhof, E and Hanessian, S. *J Med Chem.* **49** (2006) 1023-1033.
- Moustakas, D, Lang, P, Pegg, S, Pettersen, E, Kuntz, I, Brooijmans, N and Rizzo, R. *J Comput-Aided Mol Des.* in press.
- Moustakas, D, Lang, P, Pegg, S, Pettersen, E, Kuntz, I, Brooijmans, N and Rizzo, R. *J Comput-Aided Mol Des.* (in press)
- Moustakas, DT. *Development and Evaluation of Structure-Based Drug Design Algorithms in the Object-Oriented Docking Program DOCK 5*. (2004) Bioengineering, University of California, Berkeley, Berkeley.
- Nakatani, K, Horie, S, Goto, Y, Kobori, A and Hagihara, S. *Bioorg Med Chem.* **14** (2006) 5384-5388.
- Nicholls, A and Honig, B. *J Comp Chem.* **12** (1991) 435-445.
- Onufriev, A, Bashford, D and Case, DA. *J Phys Chem B.* **104** (2000) 3712-3720.
- Onufriev, A, Bashford, D and Case, DA. *Proteins.* **55** (2004) 383-94.
- Pearlman, DA, Case, DA, Caldwell, JW, Ross, WS, Cheatham, TE, Debolt, S, Ferguson, D, Seibel, G and Kollman, P. *Comput Phys Commun.* **91** (1995) 1-41.
- Pettersen, EF, Goddard, TD, Huang, CC, Couch, GS, Greenblatt, DM, Meng, EC and Ferrin, TE. *J Comput Chem.* **25** (2004) 1605-1612.
- Polacek, N and Mankin, AS. *Crit Rev Biochem Mol Biol.* **40** (2005) 285-311.
- Prabhu, NV, Zhu, PJ and Sharp, KA. *J Comput Chem.* **25** (2004) 2049-2064.
- Renner, S, Ludwig, V, Boden, O, Scheffer, U, Gobel, M and Schneider, G. *ChemBioChem.* **6** (2005) 1119-1125.

- Richards, FM. *Annu Rev Biophys Bioeng.* **6** (1977) 151-176.
- Tsui, V and Case, DA. *J Amer Chem Soc.* **122** (2000a) 2489-2498.
- Tsui, V and Case, DA. *Biopolymers.* **56** (2000b) 275-91.
- Tucker, BJ and Breaker, RR. *Curr Opin Struct Biol.* **15** (2005) 342-348.
- Wang, J, Wang, W, Kollman, PA and Case, DA. *J Mol Graph Model.* **25** (2006) 247-260.
- Wang, J, Wolf, RM, Caldwell, JW, Kollman, PA and Case, DA. *J Comput Chem.* **25** (2004) 1157-74.
- Weiser, J, Shenkin, PS and Still, WC. *J Comput Chem.* **20** (1999) 217-230.
- Wunberg, T, Hendrix, M, Hillisch, A, Lobell, M, Meier, H, Schmeck, C, Wild, H and Hinzen, B. *Drug Discov Today.* **11** (2006) 175-180.
- Yu, XL, Lin, W, Pang, RF and Yang, M. *Eur J Med Chem.* **40** (2005) 831-839.

"The most exciting phrase to hear in science, the one that heralds new discoveries, is not Eureka! (I found it!) but rather, "hmm.... that's funny...."
--Isaac Asimov

Chapter 6

Steps Toward Fully Flexible Docking to RNA Targets: Strengths and Weaknesses of Various Force Fields in Generating Ensembles of RNA Structures

P. Therese Lang¹, John Chodera², Nikolai Ulyanov³, Christophe Guilbert³, Irwin D. Kuntz³ and Thomas L. James^{3*}

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

²Graduate Program in Biophysics
University of California, San Francisco

³Department of Pharmaceutical Chemistry
University of California, San Francisco

* *To whom all correspondence should be addressed (james@picasso.ucsf.edu)*

ABSTRACT

As a molecule, RNA has traditionally been very difficult to model due to its high charge density and flexibility. In a previous study, we optimized the DOCK suite of programs to model charges for RNA targets in order to best reproduce the binding mode of known small molecule binders. Here, we extend the method by incorporating multiple receptor conformations into the docking protocol. At this point, we have explore several physics-based methods—molecular dynamics, replica exchange, miniCarlo, and Path Exploration with Distance Constraints—for generating an ensemble of RNA structures. A library of small molecules will then be cross-docked to the ensemble using the optimized DOCK protocol. To explore and optimize the sampling techniques, the procedure has been applied to HIV TAR RNA, which has a number of published structures. To validate the docking protocol, we will be screening a library of known inhibitors and decoys using Saturation Transfer Difference NMR. The diversity and quality of the generated RNA ensembles has been evaluated by comparing to experimental ensembles and to the NMR data. Thus far, we have determined that the simulations generate ensembles that are on par with the structural diversity of published apo and holo structures using RMSD as a metric. However, additional work needs to be done to evaluate the ensembles as well as the docking methodology.

INTRODUCTION

Inherent in its very nature, RNA has an extremely flexible tertiary structure. In addition to the ability of the backbone helix to twist and untwist in response to the environment, secondary structure elements like bulges and loops contain unpaired bases that allow the structures to assume both kinked and more open topologies. When in solution, RNA molecules are able to access a broad range of conformations (see Figure 1a for example). However, when bound to either a native protein binding partner or a small molecule, the RNA typically adopts one conformation particularly suited for that interaction (Aboul-ela F et al., 1996; Du Z et al., 2002; Wang S et al., 1998). For drug design purposes, this property can be capitalized by stabilizing structures that are different from the conformation the RNA assumes when binding its native partner. It may also enable several slightly different scaffolds to be developed to target different conformations of the same RNA target. With these goals in mind, we combine several commonly used sampling techniques to incorporate multiple receptor conformations into the DOCK suite of programs for drug design for RNA targets.

A variety of different methods have been employed to incorporate multiple receptor conformations into docking algorithms. These techniques fall into two main categories: conformational expansion and flexibility on-the-fly. In the conformational expansion technique, reasonable conformations for the receptor are enumerated prior to the docking calculation. The ensemble of conformations is then held constant and the small molecules are cross-docked to the ensemble.

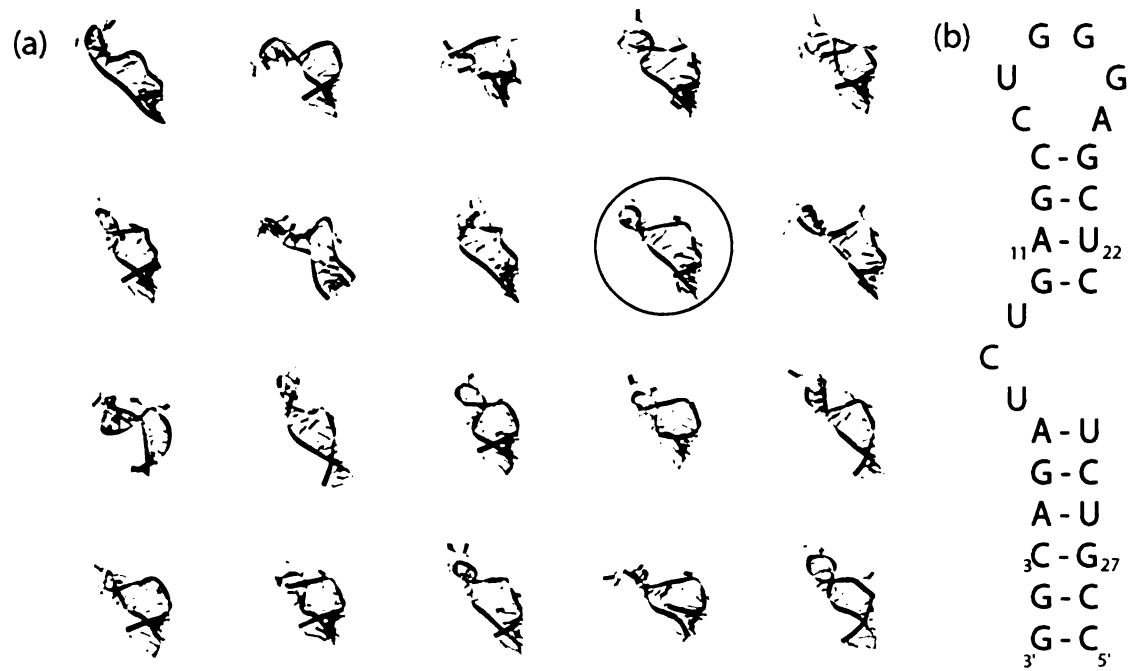


Figure 1. HIV-1 TAR RNA (a) Range of tertiary structure of TAR as determined by NMR (PDB code 1ANR). The model selected to seed the enhanced sampling techniques is circled. Image generated by Chimera (Pettersen EF et al., 2004). (b) Sequence and secondary structure of TAR

In the final stage, the ensemble score for each ligand is either combined or weighted to provide the final ranking (Knegtel RMA et al., 1997; Sherman W et al., 2006). In the flexibility on-the-fly method, the receptor conformation is sampled in conjunction with the ligand during the docking process. In most cases, this sampling is an iterative process in which the ligand is placed in the active site allowing some overlap with the receptor, and then both the receptor and the ligand are sampled to relieve the clashes (Abagyan RA et al., 1994). Currently, the only generally available algorithm that has reported successful flexible RNA docking, Autodock, used the conformational expansion technique (Moitessier N et al., 2006).

Because our docking algorithm is designed for a rigid receptor and because others have attempted it successfully before, we have decided to focus on conformational expansion (Knegtel RMA et al., 1997). In the original implementation of flexible docking in the DOCK suite of programs, the receptor ensemble came from structures that had been solved bound to a variety of different ligands. The ligand-binding events resulted in rearrangement of the active site, the new shapes of which could then be capitalized for designing novel scaffolds. Unfortunately, structural information of this diversity is only known for a handful of RNA targets, greatly reducing the transferability of this technique. To get around this problem, we have tested a variety of physics-based sampling techniques for creating ensembles of RNA structures using computational methods that still comply with experimental data. This newly created ensemble can then be used as the ensemble of receptor structures for cross-docking.

The computational sampling techniques we will be exploring include replica exchange molecular dynamics (REMD), standard molecular dynamics (MD), Monte Carlo, and Path Exploration with Distance Constraints (PEDC). Replica exchange simulation is an enhanced sampling technique, in which multiple simulations are run simultaneously at increasing temperatures, and, at set intervals, conformations are allowed to move between simulations of varying temperatures based on the Metropolis criterion (Sugita Y and Okamoto Y, 1999). If the data is analyzed by replica, or temperature, this technique has the benefit of running multiple simulations simultaneously, thus providing for the possibility of statistically less frequent events through the additional sampling. If the starting conformation of each replica is followed instead, the simulation behaves as an annealing simulation, allowing conformations to climb energy barriers very slowly while sampling at each temperature ramp. In the end, the replica exchange simulation should provide ensembles of conformations from an ensemble of energy wells. As a control, we have also performed a standard MD simulation, which will help us determine whether the enhanced sampling method is an unnecessary amount of work for our purposes.

The third method, a specialized form of Metropolis Monte Carlo called miniCarlo, is specifically designed for nucleic acids. The move sets for the method are based on internal coordinates that describe the typical movements the sugar-phosphate backbone and base pairs (Zhurkin VB et al., 1991). We anticipate this method will provide a different ensemble of structures because it will focus sampling on the predominant motions of nucleic acids. In PEDC, a

standard minimization is performed with a modified scoring function, in which a biasing restraint moves the molecule a user defined distance from the starting point (Guilbert C et al., 1995). Once again, this method should provide a different ensemble because it explores the valley of the energy landscape, typically the breathing motions of the molecule.

All of the sampling techniques described above will generate a huge number of conformations, many of which will contain redundant information. In the next stage, this large ensemble needs to be culled down to ten to fifteen structures for docking to enable reasonable lengths of calculation, but still represent the structural diversity of the entire group (see Figure 2). Because many of these techniques sample both high and low portions of the energy landscape, some of the conformations are no longer completely folded. As a first coarse processing step, then, we filter the ensemble to ensure that the minimal base pairing is still present. Those conformations that pass the filter are then clustered to reduce redundant structures. Finally, the lowest ten scoring conformations, where the score is based on the scoring scheme for the associated sampling technique, are selected for docking.

Because we are attempting to validate the ability of the sampling and culling techniques, we will be using HIV TAR RNA, a critical component of the HIV life cycle, as a test system. Once the HIV DNA has been incorporated into the host genome, it co-opts the cellular machinery and starts to replicate itself. The TAR hairpin, located at the 5' end of the nascent RNA, binds to the HIV protein Tat. This interaction then recruits a number of cellular proteins, which

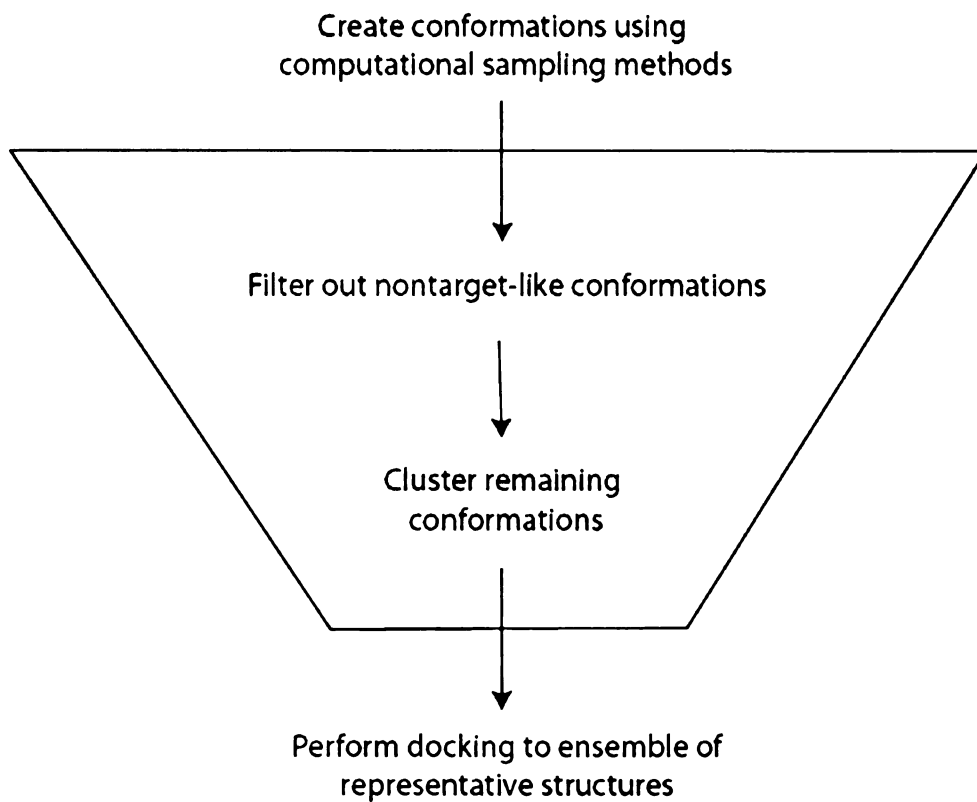


Figure 2. General scheme of receptor conformation generation and selection.

serve to facilitate production of the HIV genome (Calnan BJ et al., 1991; Frankel AD and Young JA, 1998). Because TAR is involved in such a critical portion of the life cycle, it has been the subject of numerous structural and drug design studies over the years (Davis B et al., 2004; Mayer M et al., 2006; Murchie AIH et al., 2004; Renner S et al., 2005; Yu XL et al., 2005; Zhao H et al., 2004). Thus, there is a wealth both of NMR and crystallographic data, as well as some binding data, available with which to evaluate the performance of our methods. Once the final computational ensemble has been generated, it will be compared to experimentally determined NOE restraints to determine whether the sampling and culling methodologies have produced ensembles that agree with experimental data (Aboul-ela F et al., 1995; Aboul-ela F et al., 1996; Davis B et al., 2004; Murchie AIH et al., 2004). Finally, these ensembles will be cross docked with a library of small molecules, which have already been screened against TAR and found to have both known inhibitors and known decoys (unpublished data). Once we validate our methods with the TAR test system, we will perform a prospective study on A site RNA, a ribosomal target, in which we will use the method developed on TAR to attempt to identify a novel inhibitor.

METHODS EMPLOYED TO DATE

Selection of TAR Structure

The ensemble of unbound TAR structures (PDB code 1ANR) was used as a starting point (Aboul-ela F et al., 1996). The average structure for the ensemble was calculated, and then the all-atom root mean squared deviation

(RMSD) of the entire ensemble was calculated to the average. The conformation closest to the average structure, model 14, was selected to seed all simulations (see Figure 1a).

There are currently 4 additional structures of TAR RNA with the same sequence available (PDB codes 1ARJ, 1UTS, 1UUD, and 1UUI) (Aboul-ela F et al., 1995; Aboul-ela F et al., 1996; Davis B et al., 2004; Murchie AIH et al., 2004). All of these structures were determined by NMR and are bound to a variety of drug-like molecules. The experimental data from these complexes, in addition to the apo structure, will be used to monitor and evaluate the various sampling algorithms.

Molecular Dynamics Simulations

All MD simulations were carried out with the AMBER 8.0 molecular dynamics package. The accessory program LEaP was used to prepare the selected TAR structure for simulation with the parm99 parameter set (Cornell WD et al., 1995). The Hawkins, Cramer, Truhlar Generalized Born solvation model with a salt concentration of 0.3 M and a surface area calculated using the LCPO algorithm was used to account for solvation effects (Hawkins GD et al., 1995; Hawkins GD et al., 1996; Srinivasan J et al., 1999; Tsui V and Case DA, 2000; Weiser J et al., 1999). Before simulations were performed, the starting structure was minimized using 250 steps of steepest descent energy followed by 1000 steps of conjugate gradient optimization. During MD simulations, hydrogen bond stretching was restricted using the Shake algorithm (Ryckaert JP et al., 1977).

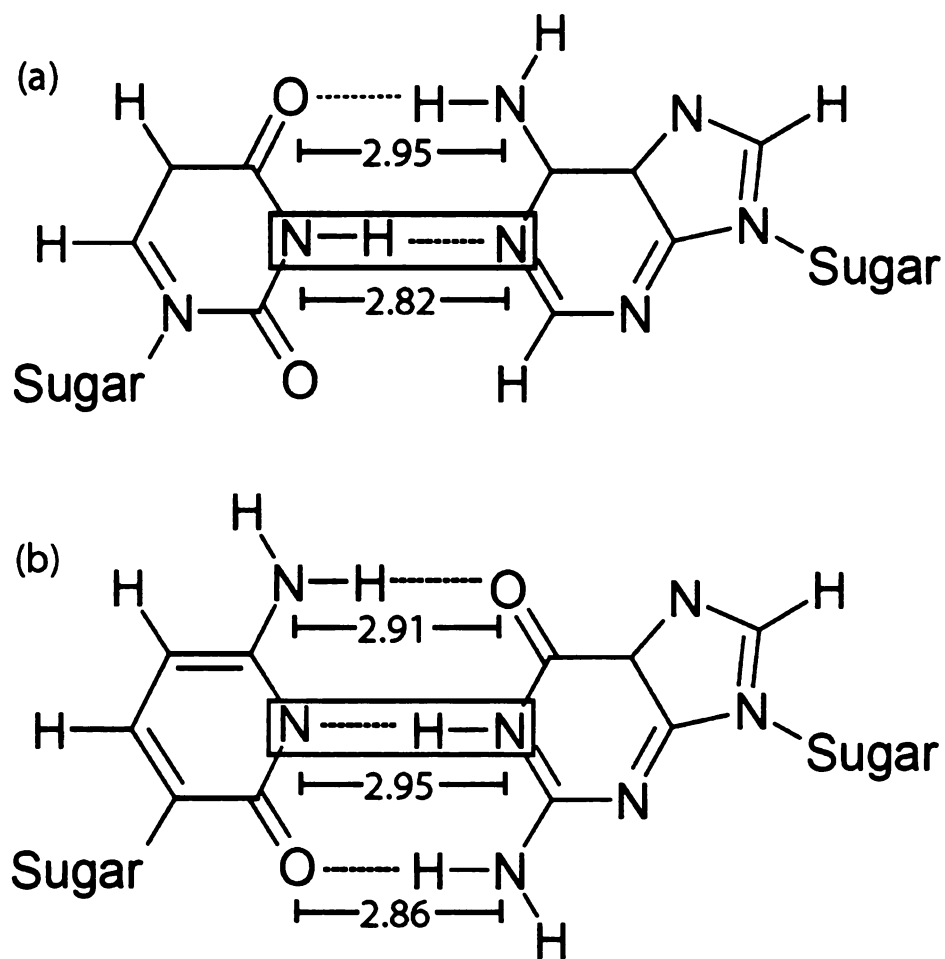


Figure 3. Hydrogen bonding distances of base pairs (Å). Distances that were restrained for MD simulations circled in red. (a) Distances of the adenine-uracil base pair. (b) Distances of the guanine-cytosine base pair.

Two single temperature MD simulations—one with restraints and one without—were run. Both simulations were started from the minimized structure and allowed to equilibrate for 2 ns at 300 K. The equilibration was followed by an 8 ns production run, and solute configurations and potential energies were saved from the production run 1 ps. The restrained simulation was repeated using the conditions above, but with square well restraints between the middle hydrogen bonds of bases 3 and 27, 4 and 26, 11 and 22, and 12 and 21 (see Figure 1b and Figure 3). If the heavy atoms of these bonds were between 1.8 and 4.0 Å, there was no penalty. If the distance was between 1.3 and 1.8 Å or 4.0 and 4.5, the scoring function was linearly penalized with a force constant of 5.0. If the distance is less than 1.3 Å or greater than 4.5 Å, the scoring function was penalized with a parabolic function with force constant of 10.0.

The REMD simulations were performed using a parallel perl wrapper for the SANDER program (Case DA et al., 2004; Chodera JD). For the first set of simulations, twenty replicas were run with temperatures ranging exponentially from 270 – 430 K. To evaluate whether conformations should be switched between replica, the algorithm starts from the highest-temperature replica and attempts to swap the configuration for the next-lowest temperature replica using a Metropolis-like criteria, and proceeds down the temperatures in this manner. On the next iteration, swapping attempts start from the lowest temperature and proceed upward; this alternation in direction is continued in subsequent pairs of iterations. All momenta were reassigned from the Maxwell-

Boltzmann distribution at the appropriate replica temperature after each exchange attempt.

As with the single temperature MD, two REMD simulations were run—with and without restraints. For the free simulation, all replicas were started from the minimized conformation and equilibrated for 1 ns at 2 fs timesteps. This equilibration run was followed by an 8 ns production run with 1 ps between exchange attempts. The exchange acceptance probability between neighboring temperatures was approximately 49.5%. Solute configurations and potential energies were saved from the production run every 1 ps. Once again, the restrained simulation was repeated using the same conditions and restraints described for the single temperature MD simulations. The exchange acceptance probability between neighboring temperatures for this simulation was approximately 50.4%.

Path Exploration with Distance Constraints

All PEDC calculations were calculated using a modified version of the CHARMM suite of programs (Foloppe N and MacKerell AD, 2000; Guilbert C et al., 1995). Before the PEDC production run, hydrogens were added to the starting structure using the Hbuild accessory and then prepared for the simulation with the param27 parameter set. Generalized Born Molecular Volume was applied to account for solvation (Lee MS et al., 2003; Lee MS et al., 2002). The conformation was then gradually minimized within CHARMM using incremental RMSD restraints on all atoms progressing from 0.1 – 1.6 Å in 0.1 Å

steps while monitoring the energy of the system. Once the energy had equilibrated, a conformation was selected from the converged set. This conformation, which was 0.55 Å all-atom RMSD from the starting structure, was used to seed the next calculation.

For the production runs, an RMSD-based biasing force was added to the scoring function to push the minimized conformation up to 0.6 Å away from the minimized conformation in steps of 0.1 Å. Once the 0.6 Å target was obtained, this conformation was reset as the reference. The biasing force was removed, and the conformation was minimized for an additional 200 steps while being restrained to stay with 0.6 Å of the new reference conformation. To additionally sample the bulge region, the PEDC simulation was repeated as above with NOE restraints between all hydrogen bonds on bases 3 and 27, 4 and 26, 11 and 22, and 12 and 21. Once again, all minimization was performed using the GBMV solvation model.

MiniCarlo

In the last method, the structures were generated using Metropolis Monte Carlo simulations with the miniCarlo program (Zhurkin VB et al., 1991). The miniCarlo program performs minimization or Metropolis Monte Carlo simulation of nucleic acid molecules using internal coordinates as independent variables. The set of internal coordinates was designed specifically for nucleic acid structures, and it includes generalized helical parameters defining the relative position of nucleic bases in space and also the internal conformation of

sugar moieties. Aromatic bases are treated as rigid bodies with idealized geometry. The conformations of the backbones connecting adjacent bases are calculated with a specialized chain closure algorithm (Zhurkin VB et al., 1978). The conformational energy is calculated *in vacuo* with a distance-dependent dielectric constant using an empirical force field optimized for nucleic acids (Poltev VI and Shulyupina NV, 1986; Zhurkin VB et al., 1980). To take the shielding effect of counterions into account, the phosphate groups are assumed charge-neutral.

The selected TAR conformation was used to generate the starting structure. Helical parameters were extracted from this structure as described in Ulyanov et al. and input to miniCarlo (Ulyanov NB and James TL, 1995). Structure was extensively energy minimized with miniCarlo and used as initial conformation for Metropolis Monte Carlo simulations. No restraints were used in any simulations. Two Monte Carlo chains were generated with 100 conformations in each chain; each conformation was stored after 10,000 Metropolis iterations. The first Monte Carlo chain was generated at 300 K. In the second Monte Carlo chain, stems were simulated at 500 K, and the bulge and the loop were simulated at 1500 K. For the high-temperature simulation, each of the 100 generated structures was additionally energy-minimized with miniCarlo.

Culling the Conformations

To ensure that the conformations were still folded properly, we checked to make sure the appropriate bases in both the upper and lower stem were still

formed. We calculated the distance between bases 3 and 27, 4 and 26, 5, and 25, 11 and 22, and 12 and 21 for all conformations generated by the sampling methods above; hydrogen-bonding angles were not considered to allow for some movement between base pairs (see Figure 1b). Those conformations with distances that violated ideal hydrogen bonding distances by more than 0.5 Å were removed from each ensemble (see Figure 3). In addition, the distance between each base pair, or rise, was calculated using the program 3DNA (Lu X and Olson WK, 2003). Those conformations with negative rise, meaning that melting occurred and resulted in shifting of the pairing, were also removed from the ensemble.

All conformations that were statistically folded were then clustered to remove redundant structures using the k-medoid clustering algorithm, a variation of k-means (Hastie T et al., 2003). In the k-medoid algorithm, a user defined number of centroids are randomly selected and then the data is divided into clusters around the centroids. The initial centroids are then randomly shifted within the cluster and the data is reassigned. If the standard deviation of the cluster members to the new set of centroids is lower than the previous set, the centroids and clusters are kept. The method continues until the standard deviation converges to a user defined amount. Because we are not sure from method to method what the optimal number of clusters should be, we monitor the average variance and its derivative for a range of centroid values (Tibshirani R et al., 2001). Once the slope of both graphs has stabilized, no more information is

being gained by adding additional clusters, and that number of centroids is selected as optimal.

TAR RNA Small Molecule Test Set

To validate both the generated structural ensemble and the docking calculation, we needed a test set with consistently derived binding affinities as well as structural information. At this time, there are four small molecules that are known to bind to TAR RNA, are available for purchase and have solved structures: Hoescht 33258, chlorpromazine, argininamide, and neomycin B (Aboul-ela F et al., 1995; Dassonneville L et al., 1997; Du Z et al., 2002; Faber C et al., 2000). We have purchased these compounds, along with 46 randomly selected, drug-like molecules selected from the drug-like subset of the ZINC database (Irwin JJ and Shoichet BK, 2005). These compounds will be screened against TAR using Saturation Transfer Difference NMR (Mayer M and James TL, 2005). This method, while often too slow to be used in screening large libraries, has the dual benefit for our purposes of providing both binding affinity and structural information for the bound structures.

RESULTS AND DISCUSSION OF CURRENT STATUS

Filtering

Once all the sampling simulations had completed, all generated conformations were subjected to filtering (Table 1). To ensure that this stage did not dramatically change the overall distribution of structures, the mean and

	Starting	Filtered
REMD	200,000	12,732
Restrained REMD	200,000	40,560
MD	10,000	5434
Restrained MD	10,000	5972
Room Temp miniCarlo	100	100
High Temp miniCarlo	100	44
PEDC	2000	2000
Restrained PEDC	2000	1963

Table 1. Total Number of Conformations at Each Stage of the Culling Process.

standard deviation of heavy atom RMSD between each experimental structure and each ensemble before and after filtering were calculated. Because these structures will eventually be used in docking calculations, the heavy atom RMSD of the active site, defined as bases 5 – 11 and bases 22 – 25 (Figure 1b), was also monitored.

Looking at the number of conformations that survive filtering already provides some useful insights (Table 1). For example, in the replica exchange simulation, over 94% of the conformations did not pass. Because the filter is designed to remove structures that are no longer properly folded, this dramatic loss of structures indicates that the majority of the ensemble is at least somewhat melted (see below for a more detailed analysis). Adding restraints to the upper and lower stems appears to have reduced this problem, with approximately 20% of structures passing the filter, but have not eliminated it completely. The same trend appears for the single temperature molecular dynamics simulation, although to a much lesser extent, indicating that the melting effect is at least partly inherent in the force field. For our purposes, we are not sure at this point what type of structural rearrangement, including partial refolding, is needed to sample the space that is covered by the experimental data. We will therefore carry the ensembles from both the standard and restrained simulations into the clustering stage and future analysis.

For both the PEDC and miniCarlo calculations, all structures passed the filter for the standard simulations. In the high temperature miniCarlo simulation, over half of the conformations were filtered out. This loss was anticipated

	Ensemble	All Holo	1ANR	1ARJ	1UTS	1UUD	1UUI	
			Standard					
Starting	9.09±2.09	9.21±2.07	8.86±2.11	9.92±1.99	8.40±1.82	9.02±2.12	8.41±2.15	
Filtered	7.17±1.39	7.21±1.26	7.09±1.60	7.59±1.43	6.89±0.77	6.31±1.25	5.80±0.98	
			Restrained					
Starting	8.42±1.87	8.54±1.84	8.20±1.90	9.23±1.72	7.76±1.63	8.18±1.78	7.55±1.88	
Filtered	7.86±1.67	8.01±1.64	7.58±1.70	8.85±1.56	7.06±1.06	7.63±1.56	6.75±1.60	

Table 2a. Heavy Atom RMSD and Standard Deviation (Å) Between Structures from REMD Simulations and All Experimental Structures

	Ensemble	All Holo	1ANR	1ARJ	1UTS	1UUD	1UUI	
			Standard					
Starting	6.45±1.32	6.80±1.16	5.78±1.36	7.00±1.25	6.56±0.96	6.82±1.26	6.75±1.26	
Filtered	5.53±1.11	5.94±0.92	4.76±1.01	5.75±0.97	6.22±0.77	5.51±0.95	5.49±0.89	
			Restrained					
Starting	6.02±1.04	6.36±0.89	5.37±0.99	6.44±1.02	6.29±0.69	6.23±1.00	6.18±0.96	
Filtered	5.85±1.05	6.22±0.89	5.14±0.96	6.32±1.01	6.11±0.67	6.08±1.00	6.01±0.97	

Table 2b. Active Site Heavy Atom RMSD and Standard Deviation (Å) Between Structures from REMD Simulations and All Experimental Structures

	Ensemble	Apo (1ANR)	All Holo	1ARJ	1UTS	1UUD	1UUI
			Standard				
Starting	6.65±1.14	6.82±1.53	6.56±0.86	6.77±0.92	6.47±0.64	5.51±0.55	5.16±0.39
Filtered	6.65±1.14	6.82±1.53	6.56±0.85	6.74±0.91	6.48±0.64	5.48±0.53	5.15±0.39
			Restrained				
Starting	6.66±1.25	6.63±1.61	6.67±1.01	6.90±1.08	6.55±0.81	5.57±0.73	5.38±0.42
Filtered	6.66±1.25	6.64±1.61	6.68±1.01	6.91±1.08	6.55±0.80	5.57±0.73	5.37±0.42

Table 3a. Heavy Atom RMSD and Standard Deviation (Å) Between Structures from MD Simulations and All Experimental Structures

	Ensemble	Apo (1ANR)	All Holo	1ARJ	1UTS	1UUD	1UUI
			Standard				
Starting	5.06±0.87	4.44±0.95	5.39±0.61	5.00±0.35	5.96±0.39	4.74±0.26	4.73±0.23
Filtered	5.06±0.87	4.44±0.95	5.38±0.61	4.99±0.35	5.96±0.39	4.73±0.25	4.73±0.23
			Restrained				
Starting	5.16±0.92	4.37±0.92	5.57±0.60	5.37±0.60	5.89±0.43	5.10±0.55	5.08±0.44
Filtered	5.16±0.92	4.37±0.92	5.57±0.60	5.37±0.60	5.89±0.43	5.09±0.55	5.08±0.44

Table 3b. Active Site Heavy Atom RMSD and Standard Deviation (Å) Between Structures from MD Simulations and All Experimental Structures

	Ensemble	Apo (1ANR)	All Holo	1ARJ	1UTS	1UUD	1UUI
	Room Temperature						
Starting	6.13±1.47	5.89±2.18	6.26±0.85	5.90±0.51	6.76±0.96	5.37±0.10	6.17±0.31
Filtered	6.13±1.47	5.89±2.18	6.26±0.85	5.90±0.51	6.76±0.96	5.37±0.10	6.17±0.31
	High Temperature						
Starting	7.74±2.10	8.19±2.37	7.50±1.91	6.12±0.83	9.30±1.41	6.28±0.64	7.63±0.71
Filtered	7.41±2.04	7.68±2.51	7.28±1.73	6.13±0.74	8.75±1.56	6.31±0.70	7.50±0.85

Table 4a. Heavy Atom RMSD and Standard Deviation (Å) Between Structures from miniCarlo Simulations and All Experimental Structures

	Ensemble	Apo (1ANR)	All Holo	1ARJ	1UTS	1UUD	1UUI
	Room Temperature						
Starting	4.43±1.29	3.25±1.10	5.05±0.88	5.76±0.31	4.08±0.36	5.60±0.09	5.62±0.09
Filtered	4.43±1.29	3.25±1.10	5.05±0.88	5.76±0.31	4.08±0.36	5.60±0.09	5.62±0.09
	High Temperature						
Starting	4.96±0.94	4.93±1.22	4.98±0.74	4.80±0.60	5.18±0.86	5.06±0.40	5.25±0.39
Filtered	4.68±0.91	4.39±1.16	4.83±0.70	5.01±0.71	4.56±0.60	5.16±0.49	5.21±0.45

Table 4b. Active Site Heavy Atom RMSD and Standard Deviation (Å) Between Structures from miniCarlo Simulations and All Experimental Structures

	Ensemble	Apo (1ANR)	All Holo	1ARJ	1UTS	1UUD	1UUI
			Standard				
Starting	7.39±2.10	6.75±2.06	7.73±2.05	9.08±1.72	6.07±1.02	7.79±1.33	6.94±1.00
Filtered	7.39±2.10	6.75±2.06	7.73±2.05	9.08±1.72	6.07±1.02	7.79±1.33	6.94±1.00
			Restrained				
Starting	7.32±1.90	6.75±1.78	7.62±1.89	9.05±1.31	5.86±0.81	7.78±0.87	7.21±0.56
Filtered	7.32±1.91	6.75±1.78	7.63±1.90	9.06±1.31	5.85±0.80	7.79±0.88	7.21±0.56

Table 5a. Heavy Atom RMSD and Standard Deviation (Å) Between Structures from PEDC Simulations and All Experimental Structures

	Ensemble	Apo (1ANR)	All Holo	1ARJ	1UTS	1UUD	1UUI
			Standard				
Starting	5.02±1.46	3.81±0.88	5.66±1.29	6.68±0.55	4.29±0.55	6.48±0.42	6.34±0.40
Filtered	5.02±1.46	3.81±0.88	5.66±1.29	6.68±0.55	4.29±0.55	6.48±0.42	6.34±0.40
			Restrained				
Starting	5.07±1.52	4.11±0.81	5.57±1.56	6.87±0.51	3.83±0.43	6.76±0.40	6.51±0.39
Filtered	5.07±1.52	4.11±0.81	5.58±1.57	6.87±0.51	3.83±0.43	6.77±0.40	6.52±0.39

Table 5b. Active Site Heavy Atom RMSD and Standard Deviation (Å) Between Structures from PEDC Simulations and All Experimental Structures

	Heavy Atom	Active Site Heavy Atom
Ensemble	6.71±2.00	4.66±1.75
Apo (1ANR)	7.10±1.87	4.89±1.31
All Holo	6.51±2.03	4.54±1.94
1ARJ	6.65±2.08	4.49±2.10
1UTS	6.39±2.03	4.60±1.70
1UUD	5.88±1.63	4.46±2.05
1UUI	6.02±1.14	4.54±1.79

Table 6. Average RMSD and Standard Deviation (Å) Between from All Experimental Structures

because the simulation was run without restraints, allowing the structure to melt. Unfortunately, it also means that fewer structures are available for analysis (e.g., there are fewer structures remaining after filtering than in the experimental ensemble). In the future, this simulation may need to be run again with restraints similar to those used in the other simulations. For the restrained PEDC calculation, about 2% of the conformations did not pass the filtering process. Because this simulation is based on minimization instead of more rigorous sampling methods, this statistic was slightly unexpected but not unreasonable. Because of the restraints, the biasing force was required to act predominantly on the free sections of the molecule. Therefore, because the base pairs that were restrained were a subset of those that were used for filtering, some of the conformations were able to expand beyond the acceptable range. As above, all of the simulations will be carried to clustering and further analysis.

The heavy atom RMSD between all the experimental structures and all the ensembles was monitored (Tables 2-5). There was no significant change in either the mean or the standard deviation for the entire structure or active site for any of the simulations with the exception of all the REMD simulations and the high temperature miniCarlo simulation, which can be attributed to the sheer number of conformations removed. In addition, the mean heavy atom values for the entire structure are a bit higher for the PEDC calculation. While we are not able to completely identify the cause currently, the primary motion of nucleic acids in this type of simulation is typically helical twisting and untwisting (data not shown). Knowing this trend, we hypothesize that this motion explains the higher

overall RMSD, while the bulge region in the active site, which does not belong to a helix, has lower values.

The values for the computational ensembles can also be loosely compared to the all-by-all RMSD values for the experimental structures (Table 6). In both the computational and experimental ensembles, the active sites have lower mean values than the entire structures, indicating that other portions of the biomolecule (eg backbone twist or the hairpin) are more varied. When looking at all the computational methods, the room temperature miniCarlo simulation produced the lowest overall mean for the apo structure and for two of the bound structures. It also produced the lowest active site mean for the apo structure. However, the single temperature molecular dynamics simulation had lower active site means for the majority of the bound structures. We hesitate to place too much meaning on this data pending a more thorough analysis using experimental NOE values. However, at this point, we suggest that the miniCarlo technique may be more appropriate for simulating overall structure variability whereas molecular dynamics may be more appropriate for bulge and loop sampling.

Replica Exchange Simulation

Upon seeing the number of conformations that were failing the filtering step for the replica exchange simulations, we wanted to understand what was occurring in more detail. After spot-checking a few of the conformations from the simulation, it was determined that, in many cases, all the base pairs for one or

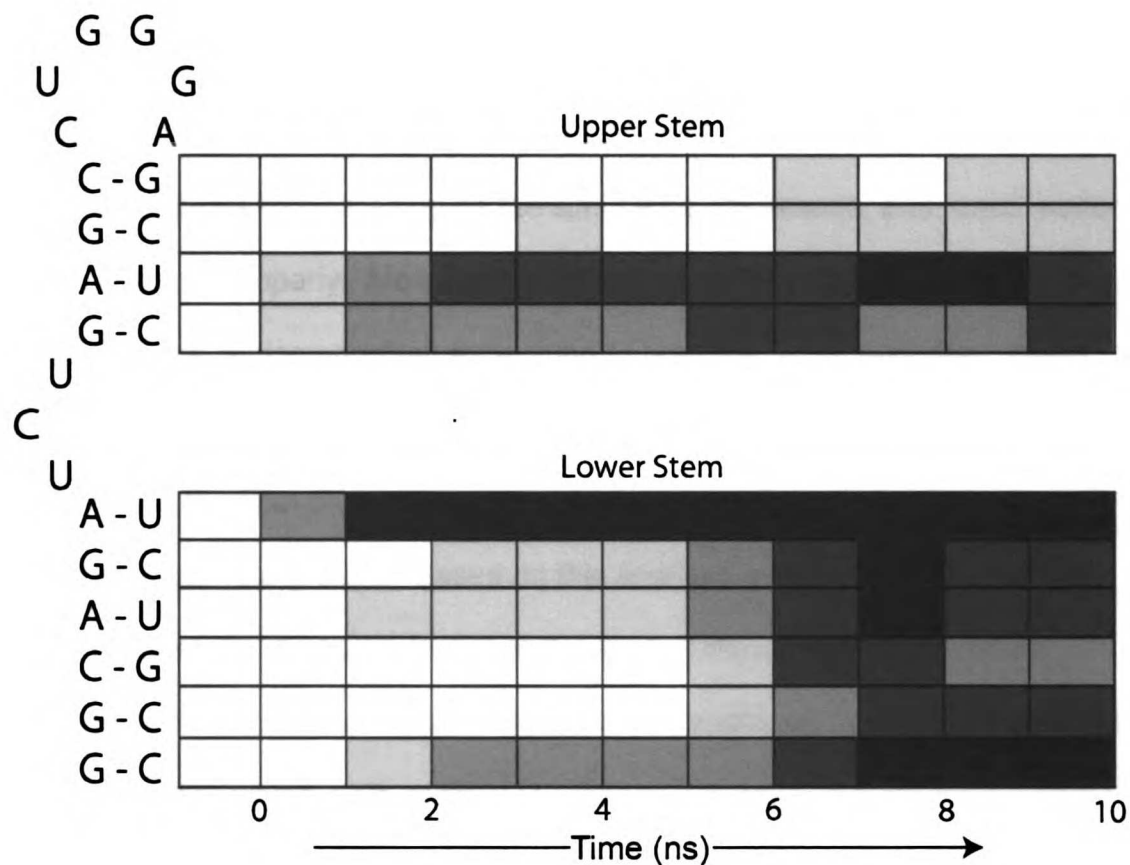


Figure 4. Average hydrogen bond distances for base pairs over course of unrestrained REMD simulation. Averages are calculated for all hydrogen bonds in each pair. Each row represents the average for the base pair of the secondary structure pictured on left. The first column contains data for the starting structure and the remaining columns contain averages over the course of the simulation in 1 ns intervals. Each block is colored by average distance, where white is less than 4 Å, light grey is 4 – 5 Å, medium grey is 5 – 6 Å, dark grey is 6 – 7 Å and black is greater than 7 Å.

both of the stems were no longer paired (data not shown). To obtain a more holistic picture, the hydrogen bonding distances for all of the base pairs were calculated over the course of the simulation (Figure 4). In this figure, the heavy atom hydrogen bond distance was calculated and averaged over 1 ns intervals. As can be seen by the increasing amount of violation over time, the RNA structure appears to be melting as the simulation progresses, and, once melted, does not refold properly. More interestingly, the melting appears to be starting from the bulge region and then expanding to both the upper and lower stems.

While we have not determined the exact cause of the unfolding, many of the bases in the spot checked structures appeared to be interacting with nearby charged phosphate groups. Based on this analysis, we hypothesize that the unpaired bases in the bulge region at the start of the simulation search for hydrogen bonding partners. Because there is no water in the simulation, these bases attach to whatever is nearby, tugging on the rest of the structure. This sampling is not inherently problematic except that, when these unpaired bases find the naked phosphate charge on the backbone, the interaction is so strong they do not detach in the course of the simulation. The base pairs in the stems, which are not strong enough to counter this tugging, are pulled apart, and then start sampling the backbone themselves. In addition to our study, there is anecdotal evidence of this type of base opening behavior (Cheng X et al., 2005). In all of these simulations, as also shown here, this problem can be overcome with extremely minimal constraints on the previously formed base pairs.

However, it also points to an inherent weakness either in the force field or solvation model of AMBER as applied to nucleic acids.

PRELIMINARY CONCLUSIONS

At this point in the project, we are still exploring the ability of the simulations to produce reasonable ensembles of RNA structures. We have shown that molecular dynamics simulations seem to have some serious issues with simulating this system in implicit solvent, primarily due either to over stabilization of unpaired base-backbone interactions or the under stabilization of base pairs. While explicit solvent may alleviate some of this problem, it does not solve it in cases where the counter-ion has traveled away from its partner phosphate group. We are not in the business of modifying or developing force fields, so we cannot make any definitive suggestions about how to address this problem. However, we hypothesize that the partial charges in the AMBER parm99 force field may not take the polarizing effect of the backbone into account and may need to be reparameterized to reduce this effect.

For all simulations, the filtering step does not seem to be negatively affecting the diversity of the ensembles using the crude metric of RMSD. It appears that the miniCarlo and single temperature molecular dynamics simulations are better able to reproduce the range of structure and active sites, respectively, of the experimental ensemble. Clustering of these, and the other computational ensembles, should serve to further reduce redundancy while maintaining the current distribution. However, it remains to be seen whether

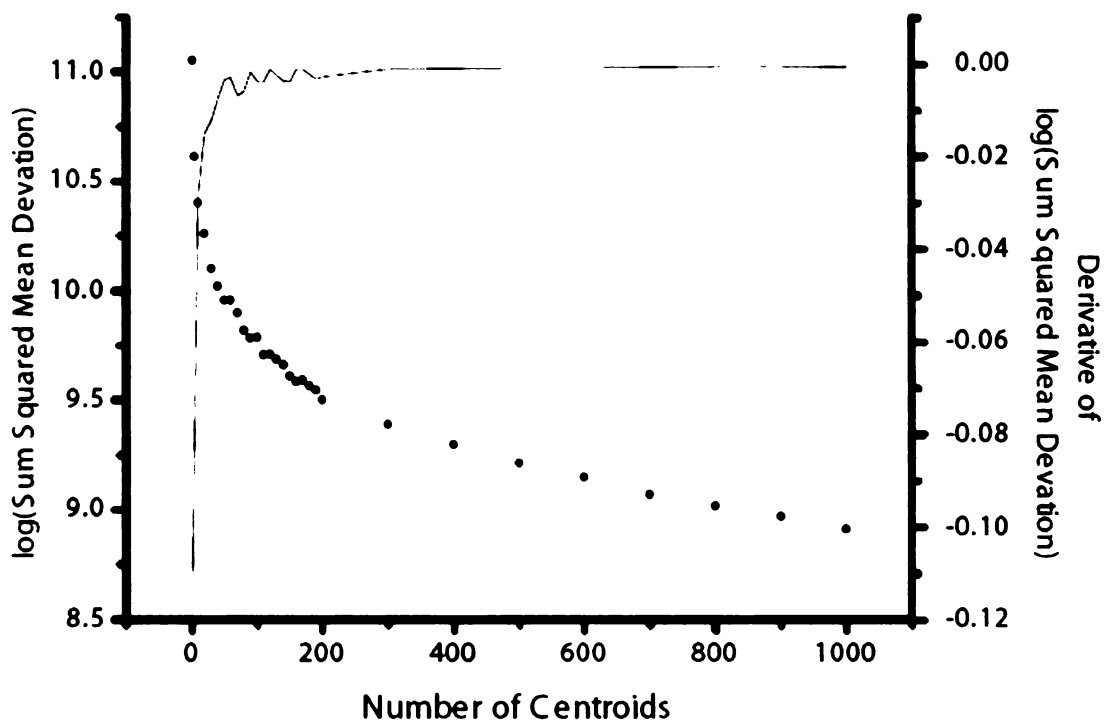


Figure 5. Example of optimization of number of clusters for the replica exchange simulation. Average variance (-) and standard deviation of average variance of clusters (•) of all clusters is monitored over a range of centroids. The number of centroids at the point in which the average variance has converged is selected for the next stage of the culling process.

these distributions contain ensembles that are reasonable according to experiment, and, more importantly, if these ensembles will improve docking success rates.

FUTURE DIRECTIONS

Clustering

At this point, the exact details of the clustering algorithm are still being finalized. We are using the structures from the MD simulations to prototype the methodology and have begun preliminary work on methods for optimizing the number of clusters that will capture the main features of the conformational ensemble while removing redundancy (Figure 5). Finally, we will explore and compare clustering based on the active site and backbone alone.

Analysis of Conformational Ensembles

There are two types of questions we can ask of the structural ensembles that we have currently generated. First, we hope to explore the structural diversity of the ensembles by calculating various metrics including active site volume, torsional sampling and nucleic acid helical parameters. For the molecular dynamics simulations, which are time dependent, we will also explore how long it takes to move between conformations. As an orthogonal question, we will explore how these structures reproduce experimental data. While we are currently comparing to structures based on RMSD, we anticipate that analyzing the ensembles based on raw NOE and J coupling data will be a significantly

more informative metric, as this type of data will provide information distributed over the entire structure. We will also break the analysis down by the ability of each ensemble to recreate the original apo structure ensemble from which the starting structure was obtained as compared to the ability to reproduce the bound structures.

Cross Docking

Once we are comfortable with the ensembles of structures, we will begin the cross docking experiments using the validation library of small molecules. As a first pass, we will be computing the score for each receptor-ligand complex using a thermodynamic cycle and then taking the best overall score. It is anticipated that this conformational expansion, followed by cross docking and minimization of the entire complex, will improve success rates in a library of small molecules.

REFERENCES

- Abagyan, RA, Totrov, MM and Kuznetsov, DN. *J Comp Chem.* **15** (1994) 488-506.
- Aboul-ela, F, Karn, J and Varani, G. *J Mol Biol.* **253** (1995) 313-32.
- Aboul-ela, F, Karn, J and Varani, G. *Nucl Acids Res.* **24** (1996) 3974-3981.
- Calnan, BJ, Biancalana, S, Hudson, D and Frankel, AD. *Genes Dev.* **5** (1991) 201-210.
- Case, DA, Darden, TA, Cheatham, I, T.E. , Simmerling, CL, Want, J, Duke, RE, Luo, R, Merz, KM, Wang, B, Pearlman, DA, Crowley, M, Brozell, S, Tsui, V, Gohlke, H, Mongan, J, Hornak, V, Cui, G, Beroza, P, Schafmeister, C, Caldwell,

JW, Ross, WS and Kollman, PA. "AMBER 8." (2004). University of California, San Francisco.

Cheng, X, Cui, G, Hornak, V and Simmerling, C. *J Phys Chem B Condens Matter Mater Surf Interfaces Biophys.* **109** (2005) 8220-30.

University of California, San Francisco. A copy of this Perl wrapper to perform replica-exchange simulations using AMBER7 and AMBER8 can be obtained from <http://www.dillgroup.ucsf.edu/~jchodera/code/rex>. San Francisco, CA.

Cornell, WD, Cieplak, P, Bayly, CI, Gould, IR, Merz, KM, Ferguson, DM, Spellmeyer, DC, Fox, T, Caldwell, JW and Kollman, PA. *J Am Chem Soc.* **117** (1995) 5179-5197.

Dassonneville, L, Hamy, F, Colson, P, Houssier, C and Bailly, C. *Nucleic Acids Res.* **25** (1997) 4487-92.

Davis, B, Afshar, M, Varani, G, Murchie, AIH, Karn, J, Lentzen, G, Drysdale, MJ, Bower, J, Potter, AJ, Starkey, ID, Swarbrick, TM and Aboul-Ela, F. *J Mol Biol.* **336** (2004) 343-356.

Du, Z, Lind, KE and James, TL. *Chem Biol.* **9** (2002) 707-12.

Faber, C, Sticht, H, Schweimer, K and Rösch, P. *J Biol Chem.* **275** (2000) 20660-20666.

Foloppe, N and MacKerell, AD. *J Comput Chem.* **21** (2000) 86-104.

Frankel, AD and Young, JA. *Annu Rev Biochem.* **67** (1998) 1-25.

Guilbert, C, Perahia, D and Mouawad, L. *Comp Phys Com.* **91** (1995) 263-273.

Hastie, T, Tibshirani, R and Friedman, J. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. New York: Springer-Verlag. (2003)

Hawkins, GD, Cramer, CJ and Truhlar, DG. *Chem Phys Lett.* **246** (1995) 122-129.

Hawkins, GD, Cramer, CJ and Truhlar, DG. *J Phys Chem.* **100** (1996) 19824-19839.

Irwin, JJ and Shoichet, BK. *J Chem Inf Model.* **45** (2005) 177-182.

Knegt, RMA, Kuntz, ID and Oshiro, CM. *J Mol Biol.* **266** (1997) 424-440.

- Lee, MS, Feig, M, Salsbury, FR and Brooks, CL. *J Comput Chem.* **24** (2003) 1348-1356.
- Lee, MS, Salsbury, FR and Brooks, CL. *J Chem Phys.* **116** (2002) 10606-10614.
- Lu, X and Olson, WK. *Nucl Acids Res.* **31** (2003) 5108-5121.
- Mayer, M and James, TL. "Discovery of ligands by a combination of computational and NMR-based screening: RNA as an example target." In Nuclear Magnetic Resonance of Biological Macromolecules, Part C. Ed. (2005) 571-587.
- Mayer, M, Lang, PT, Gerber, S, Madrid, PB, Pinto, IG, Guy, RK and James, TL. *Chem Biol.* **13** (2006) 993-1000.
- Moitessier, N, Westhof, E and Hanessian, S. *J Med Chem.* **49** (2006) 1023-1033.
- Murchie, AIH, Davis, B, Isel, C, Afshar, M, Drysdale, MJ, Bower, J, Potter, AJ, Starkey, ID, Swarbrick, TM, Mirza, S, Prescott, CD, Vaglio, P, Aboul-ela, F and Karn, J. *J Mol Biol.* **336** (2004) 625-638.
- Pettersen, EF, Goddard, TD, Huang, CC, Couch, GS, Greenblatt, DM, Meng, EC and Ferrin, TE. *J Comput Chem.* **25** (2004) 1605-1612.
- Poltev, VI and Shulyupina, NV. *J Biomol Struct Dyn.* **3** (1986) 739-65.
- Renner, S, Ludwig, V, Boden, O, Scheffer, U, Gobel, M and Schneider, G. *ChemBioChem.* **6** (2005) 1119-1125.
- Ryckaert, JP, Cicotti, G and Berendsen, HJC. *J Comput Phys.* **23** (1977) 327-341.
- Sherman, W, Day, T, Jacobson, MP, Friesner, RA and Farid, R. *J Med Chem.* **49** (2006) 534-553.
- Srinivasan, J, Trevathan, MW, Beroza, P and Case, DA. *Theor Chem Acc.* **101** (1999) 426-434.
- Sugita, Y and Okamoto, Y. *Chem Phys Lett.* **314** (1999) 141-151.
- Tibshirani, R, Walther, G and Hastie, T. *Journal of the Royal Statistical Society Series B-Statistical Methodology.* **63** (2001) 411-423.
- Tsui, V and Case, DA. *J Amer Chem Soc.* **122** (2000) 2489-2498.

Ulyanov, NB and James, TL. "Statistical Analysis of DNA Duplex Structural Features." In Meth Enzymol. T. L. James Ed. Academic Press: New York (1995) 3-44.

Wang, S, Huber, PW, Cui, M, Czarnik, AW and Mei, H-Y. *Biochemistry*. **37** (1998) 5549-5557.

Weiser, J, Shenkin, PS and Still, WC. *J Comput Chem*. **20** (1999) 217-230.

Yu, XL, Lin, W, Pang, RF and Yang, M. *Eur J Med Chem*. **40** (2005) 831-839.

Zhao, H, Li, J, Xi, F and Jiang, L. *FEBS Lett*. **563** (2004) 241-5.

Zhurkin, VB, Lysov, YP and Ivanov, VI. *Biopolymers*. **17** (1978) 377-412.

Zhurkin, VB, Poltev, VI and Florent'ev, VL. *Mol Biol (Mosk)*. **14** (1980) 1116-30.

Zhurkin, VB, Ulyanov, NB, Gorin, AA and Jernigan, RL. *Proc Nat Acad Sci USA*. **88** (1991) 7046-7050.

“Research is what I'm doing when I don't know what I'm doing.”

--Wernher Von Braun

Chapter 7

A Fragment-Based Screening Method Designed for RNA Targets

P. Therese Lang¹, David Smithson², Irene Gomez-Pinto³, Anang Shelat², R. Kip Guy², Irwin D. Kuntz³, Thomas L. James^{3*}

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

² Chemical Biology and Therapeutics
St. Jude's Children's Research Hospital, Memphis

³ Department of Pharmaceutical Chemistry
University of California, San Francisco

* To whom all correspondence should be addressed (james@picasso.ucsf.edu)

ABSTRACT

This project represents the first steps toward the design of a fragment-based screening program for RNA targets. The fragments for the library have been biased toward chemical moieties that resemble the major properties of known RNA binders. Because fragments typically bind at lower affinity and because RNA targets are so flexible, saturation transfer difference NMR is initially used to screen for binding. If STD results imply binding, simple RNA resonances, in particular, imino resonances, are monitored for line broadening indicative of ligand binding to a specific moiety of the RNA. Once the fragments are identified as interacting with the RNA target, the structure of the bound complex will be modeled using docking, minimization and molecular dynamics simulations in the DOCK 6 suite of programs. Finally, the fragments that look the most promising will be synthetically combined using “click” chemistry. To test this new methodology, the HIV target TAR has been selected as a target. Preliminary results indicate that many of the fragments bind to the target of interest and may make good lead scaffolds. However, the “click” chemistry reaction has been shown to be less general than originally anticipated, and, as a result, may have to be removed from the general scheme. As a next step, controls need to be run to fully characterize the fragments that have been shown to bind, and the three-dimensional bound complexes need to be modeled.

INTRODUCTION

In the last few years, it has become apparent that the role of RNA in cellular systems is significantly more complicated than originally assumed. Along with this new knowledge of RNA function, there is increasing amounts of evidence that RNA may be a useful target for drug design, particularly in systems like HIV and antibiotic-resistant bacteria where protein therapies have been less successful (Barker JJ, 2006; Mayer M and James TL, 2005). However, because RNA has so many physical and chemical characteristics that differ from proteins, many of the traditional drug design techniques need to be adapted or extended specifically to suite nucleic acids.

Fragment-based screening has become widely used to identify lead scaffolds as an alternative to high-throughput screening of large libraries of small molecules (Keseru GM and Makara GM, 2006; Zartler ER and Shapiro MJ, 2005). The basic principle is to screen a range of chemical moieties that represents commonly occurring motifs in successful drugs (fragments), identify the chemical moiety(ies) that bind to the target of interest, and then synthetically combine them (scaffolds) (Figure 1). Typically, the fragments have a molecular weight between 150-250 D and are selected to represent a wide range of chemical space (Carr RA et al., 2005). The screening technique is generally structure-based because the fragments tend to bind at lower affinity that would be expected from larger molecules, making them difficult to detect using other screening methods (Lepre CA, 2001).

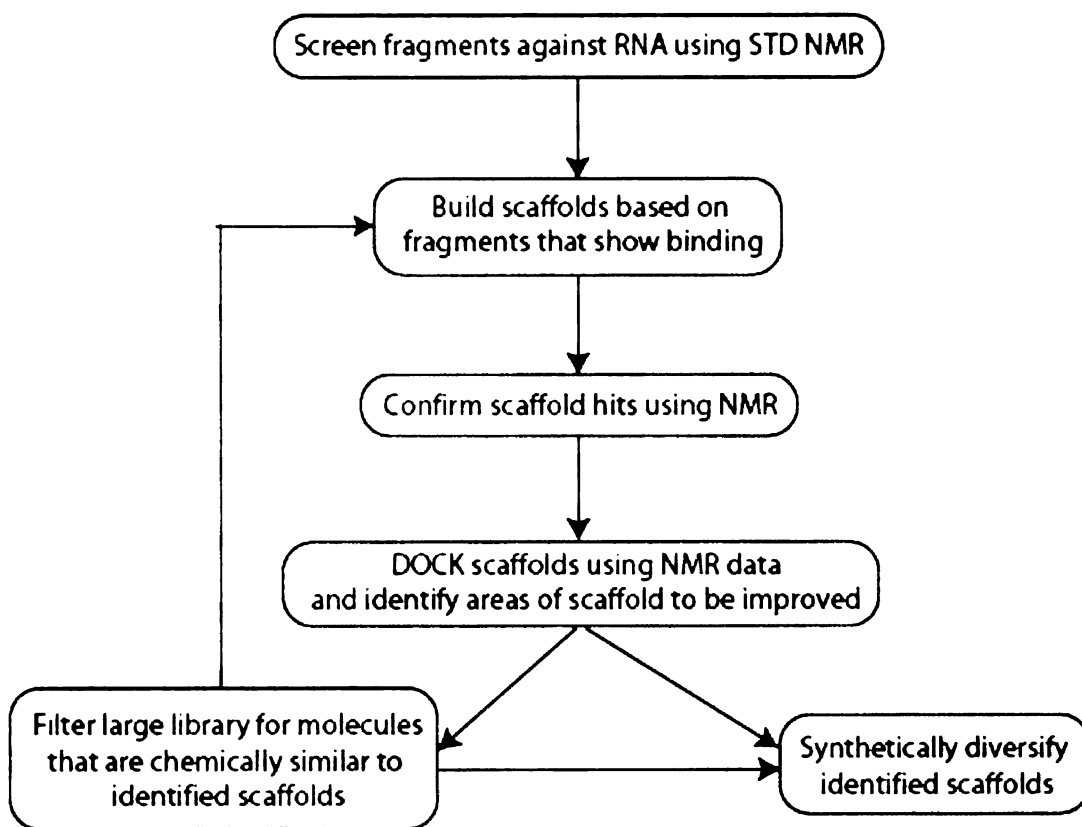


Figure 1. Scheme for RNA-biased fragment based screening method

As with other screening methods, fragment-based libraries have been used to screen against RNA targets. However, the technique was not as successful as when applied to protein systems (Fejzo J et al., 1999; Johnson EC et al., 2003). We hypothesize that the primary reason for the lack of success is because the fragments were designed to represent chemical moieties that commonly occur in small molecules that interact with protein targets. If one examines the molecules that have been shown to bind RNA, the most common interactions include hydrogen bonding, electrostatic, electron withdrawing groups that can react to the localized charge on the RNA backbone, and π -stacking (Francois B et al., 2005; Kaul M and Pilch DS, 2002; Pilch DS et al., 2003; Vicens Q and Westhof E, 2003). These properties are different from protein interactions, which are dominated by shape complementarity. Therefore, in addition to chemical diversity, we have biased the fragments of our library toward chemical moieties that encompass the properties of RNA binders.

Because RNA is so malleable, it often globally rearranges upon ligand binding. Therefore, NMR is an appropriate technique for measuring binding events because the biomolecule is in solution and free to move in reaction to its environment. Saturation transfer difference (STD) has been proven to be efficient in identifying and characterizing weakly binding compounds to small RNA constructs (Mayer M and James TL, 2002). STD signals reflect the transfer of saturation from the RNA to the ligand while in the bound state. This method allows the determination of the binding epitope of the ligand. Therefore, ligand protons which are close to the RNA surface exhibit high STD NMR signal

intensities and are more likely to be involved in substrate binding. Subtracting a spectrum, where the RNA is saturated, from one without RNA saturation produces a spectrum where only signals of the ligand(s) remain in the difference spectrum.

Once the NMR studies have identified the fragments, the structures of the bound complexes need to be solved to determine which fragments should be joined and which areas of the scaffolds would be most useful to derivatize. The fragment will first be placed in approximately the correct location using docking. In docking programs, the surface of the receptor is sampled to place a small molecule in the most energetically favorable position as scored by a molecular mechanics scoring function. Once each ligand is placed, the entire complex will be relaxed in response to the binding event using minimization followed by molecular dynamics. The newest release of DOCK, version 6, is particularly suited for this purpose because it contains the capacity for docking, minimization, and molecular dynamics in one package (see Chapter 5). If necessary, the data from the NMR experiments can be applied to restrain both the docking and the more sophisticated simulations and guide them to sample the most useful areas of space.

Finally, to make it easier to combine the fragments once they have been identified, we have made use of one of the “click” reactions. These reactions, collected by Kolb et al (Kolb HC and Sharpless KB, 2003), are termed “click” chemistry if they have simple reaction conditions with high yields, readily available starting materials and reagents, and simple stereospecific product

isolation. Since its inception, “click” chemistry has become commonly used for a wide variety of synthetic applications (Aucagne V and Leigh DA, 2006; Choi WJ et al., 2006; Montagnat OD et al., 2006). For our purposes, because we are interested in π -stacking interactions, we selected the [3+2] dipolar azido-cycloaddition reaction. Each of fragments, therefore, has also been selected to contain either an azide or acetylene moiety to facilitate the reaction.

Finally, to validate this entire proposed scheme, we have chosen TAR RNA as a test target. The Tat-TAR complex has been identified as an attractive target for the inhibition of HIV (Hsu MC et al., 1991). In the first stages of HIV replication, the Tat protein facilitates viral transcription from the promoter region of the provirus incorporated into the DNA of the host cell. In order to form this interaction, Tat binds specifically to an RNA hairpin known as trans-activating response element (TAR) at the 5' end of the newly formed viral transcripts (Calnan BJ et al., 1991). The Tat-TAR complex has been found to enhance the overall rate of viral mRNA production by as much as 100-fold (Calnan BJ et al., 1991; Frankel AD and Young JA, 1998). It has been shown that disruption of this complex prevents elongation of the RNA genome by RNA polymerase, reducing viral replication (Karn J, 1999). In addition, the TAR system has been the focus of a number of successful structure-based drug design efforts over the years (Davis B et al., 2004; Mayer M et al., 2006; Murchie AIH et al., 2004; Renner S et al., 2005; Yu XL et al., 2005). While none of the small molecules made it to the clinic, these studies suggest it is possible to target TAR with a small molecule, making it useful for exploring the design of a fragment-based library.

METHODS EMPLOYED TO DATE

Design of Fragment Library

In order to design the first fragment library, a number of libraries, including the MDL Available Chemical Directory and Screening Compounds (<http://www.mdl.com>), the ZINC database (Irwin JJ and Shoichet BK, 2005), the Bay Area Screening Center Database (<http://www.ucsf.edu/basc>), and the UCSF Inventory, were screened using MDL Isis/Base (Elsevier MDL). Each fragment was required to contain either an azide or acetylene functional group to make them amenable for synthesis. Next, any fragment without at least one non-carbon heavy atom and a ring system with at least four heavy atoms were removed to bias the compounds toward properties of known RNA binders. The fragments were also loosely filtered for drug-likeness, in this case defined as less than four rotatable bonds and a molecular weight of less than 230 D and less than 200 D for the azide- and acetylene-containing compounds, respectively.

The fragments were then subjected to a Rapid Elimination of Swill filter built in Pipeline Pilot (SciTegic) to remove chemical groups that have undesirable reactivity in biological settings and clustered for diversity (Hann M et al., 1999; Rishton GM, 1997; Walters WP and Murcko MA, 2002). Because the fragments would be evaluated by STD NMR, the H^1 spectrum of each fragment was calculated using ACD/HNMR Predictor, and those fragments whose peaks almost completely overlapped with RNA signals were removed (Golotvin SS et al., 2006). The remaining fragments were then combinatorially combined *in silico*

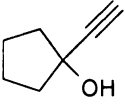
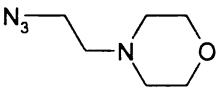
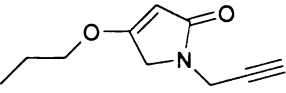
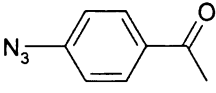
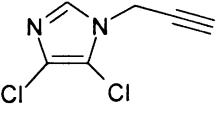
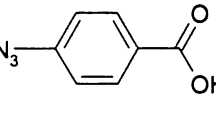
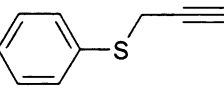
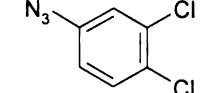
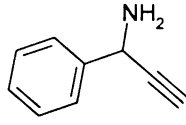
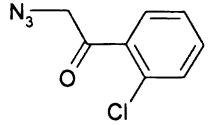
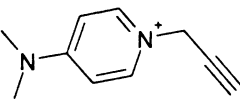
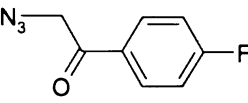
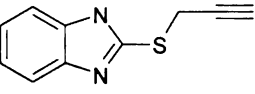
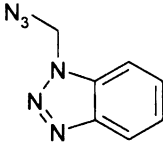
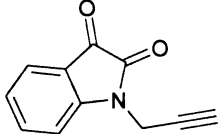
Acetylene Fragments	Binding	Azide Fragments	Binding
	---		++
	---		+
	++		---
	+++		---
	+		+++
	+++		---
	+++		+++
	+++		

Table 1. Binding of Original Fragment Library to TAR RNA. All values measured using STD NMR with 500 μM of compound and 25 μM TAR RNA. Binding values correspond to the following scheme: strong (+++), medium (++), weak (+), and no (---) change in fragment NMR upon binding. Because there was not enough compound for titration experiments, binding cannot be quantitatively evaluated at this point. For more details, see Methods Section.

in Pipeline Pilot into full scaffolds, and the scaffolds were filtered for solubility between -7 and 7 using the SlogP metric in the program MOE (Chemical Computing Group). The final set of scaffolds was broken back into fragments and those that were available for purchase were obtained (Table 1).

Detection of Binding by STD NMR

Unlabeled RNA samples were prepared and purified as previously described (Mayer M and James TL, 2004). All NMR spectra were acquired in a Bruker DRX 500 MHz spectrometer equipped with a cryoprobe and a sample changer. In all STD experiments, on-resonance irradiation was set to 5.5 ppm and off-resonance irradiation was set to 30 ppm where no RNA resonances were present. 256 scans were acquired for the STD experiments. To reduce molecular motion, all experiments were performed at 15°C. STD experiments were processed with the XWIN-NMR software (Bruker Biospin GmbH). Typically, a spectrum of the compound alone at 500 μ M concentration was acquired, and then RNA was added to produce a 25 μ M final RNA concentration. 256 scans were acquired for STD experiments. Presaturation of RNA resonances was achieved by an appropriate number of band-selective G4 Gaussian cascade pulses to give a saturation time of 2 seconds.

Initial Optimization of Synthetic Conditions

To optimize the initial reaction conditions, 1-ethynyl-4-fluorobenzene, 4-azidoaniline hydrochloride, phenylacetylene, and 4-azidophenacylbromide were

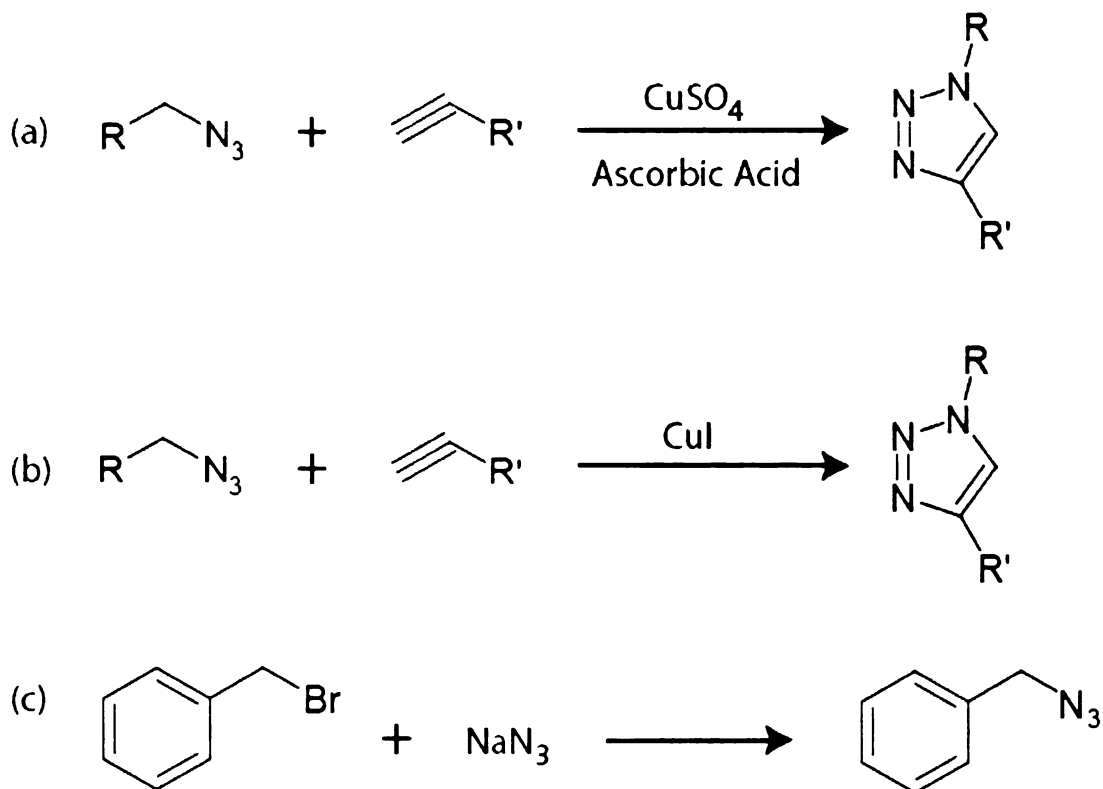


Figure 2. Synthesis of scaffolds from fragment library. For more information, see Methods Section. (a) Scheme 1 for scaffold synthesis. The conditions for synthesis are 1:1 azide: acetylene, 0.01 eq. CuSO_4 , 0.1 eq. ascorbic acid, water, t-butanol, room temperature, 5-10 days. (b) Scheme 2 for scaffold synthesis. The conditions for synthesis are 1:1 azide: acetylene, 1 eq CuI , 50 eq DIPEA, toluene, dimethylformide, room temperature, 5 days. (c) Synthesis of azide from bromide starting material. The conditions for synthesis are 0.9 eq. NaN_3 , DMSO, 50°C for 3 hours.

chosen because they were inexpensive and available in large quantities. For optimization purposes, two synthetic schemes were compared (Figure 2). In scheme 1, the acetylene compounds were combined with the azide compounds in equal portions with 0.01 eq of CuSO₄ and 0.1 eq of Na ascorbate in a 50% *tert*-butanol/H₂O buffer (Sivakumar K et al., 2004) (Figure 2a). In scheme 2, the acetylene compounds were combined with the azide compounds in equal portions with 1-2 eq. of CuI and 50 eq of DIPEA. A variety of solvents were tested during optimization, including toluene, dichloromethane, ethyl acetate, ethynol, *t*-butanol, and diisopropylethylamine (DIPEA) (Lee LV et al., 2003; Tornøe CW et al., 2002) (Figure 2b). Reactions were run for up to 5 days with approximately 75% yield by TLC for the best condition—scheme 2 in DIPEA.

Synthesis of Scaffolds from Fragment Library

For the library, all reactions were carried out on a polypropylene reaction block with solvent-resistant rubber seals. The final conditions for each well were 2.3 mg CuI, 20 μ L DIPEA, equal amounts of azide and acetylene, 2.0 mL of toluene, and 0.5 mL dimethyl formamide. Reactions were run for 4 days at room temperature. Unfortunately, the rubber seals were dissolved by the solvent, resulting in loss and contamination of product and preventing a full characterization of the reaction.

To avoid the issue with the seals, the second round of synthesis was carried out in 3 mL vials. Reactions were run for up to 14 days at room temperature. The yield of each of the reactions was determined using LCMS

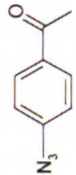
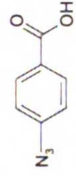
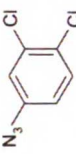
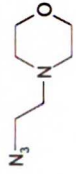
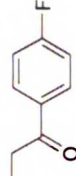

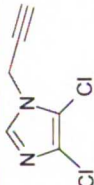
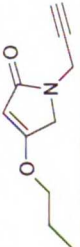
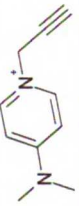
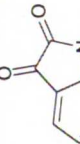
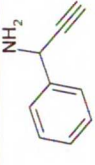
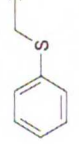
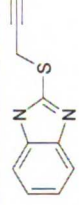
Azides	Acetylenes					
	---	---	---	---	---	43%
	---	---	NA	17%	31%	26%
	---	---	30%	8%	60%	70%
	---	---	---	NA	84%	50%
	---	---	30%	9%	100%	60%
	---	---	---	---	---	18%
	37%	70%	71%	24%	24%	35%
	---	NA	---	5%	5%	14%

Table 2. Percent Yield from Synthesis Using Optimized Conditions. Values are the isolated yield as measured by LCMS. Several reactions showed no detectable product (---), and several were not attempted under these conditions (NA). For more information on the synthetic protocol, see the Methods Section.

(Table 2). Unfortunately, many of the compounds were lost during purification due to a technical error. In the end, only seven compounds were collected, none of which were soluble at the concentration needed for testing by NMR.

DISCUSSION OF CURRENT STATUS

Synthesis

In general, we found that the “click” chemistry reaction was not as widely applicable as originally reported. From the original fragment library, we determined that the products were stereoselective and easy to isolate. However, over 40% of the reactions did not proceed, and, of those that completed, the average yield was only 41%. When looking at those reactions that did complete, some hypotheses can be made about what groups effect the reaction. For example, electron-withdrawing groups seem to reduce the yield and vice versa for electron-donating groups. Sterics seem to have an influence on the reactions as well. At this point, we feel the best course of action is to take a step back from the fragment-focused synthesis and fully characterize the reaction between substituted azido-benzene and substituted ethynyl-benzene using a Hammett series (Hammett LP, 1937).

Binding of Fragments

For the original library of fragments, 67% of the fragments were shown to have some interaction with the TAR RNA target. The fact that some of the molecules did not bind serves as an indirect control that the interactions are real

rather than the result of the interaction of the acetylene or azide groups. However, because the reciprocal experiment of the effect of the ligand on the RNA target has not yet been performed, it is difficult to determine whether the fragments are binding in the active site of the RNA. In addition, because titration experiments have not been completed, it is also inappropriate to quantify the binding outside of the qualitative ranking that is currently being employed.

PRELIMINARY CONCLUSIONS

At this point, the project is still obviously in the “proof of principle” stage. However, the results thus far are relatively encouraging. Many of the fragments in the original library were shown to bind to the TAR target, indicating that biasing the fragment library toward RNA binder-like characteristics may indeed improve the success of fragment-based screening for RNA. Unfortunately, the synthetic effort has met a number of obstacles along the way. We hope that the Hammett series synthesis will provide light and direction for the limits of the reaction. However, it is possible that the azido cyclization will prove too limiting for this application. In this case, we will simply use the fragments as the most basic lead for further derivatization.

FUTURE DIRECTIONS

Fragment Library

The fragments that resulted from the original search were from a wide range of chemistry. However, due to the specialized requirement of having either

an azide or acetylene group to facilitate synthesis, expanding the chemical space will prove to be difficult. To address this issue, particularly for fragments containing azides, the next search will have to encompass additional specialty libraries. As an alternative, other synthetic steps may need to be utilized to convert from another chemical group to the form needed for the click chemistry reaction, for example converting bromides into azides, to expand the fragments into more diverse chemical space. On the other hand, if the synthetic reaction eventually proves untenable for this project, the options for fragment should increase greatly.

Synthesis

At this point, we believe that the “click” reaction is viable as a synthetic technique, and can be useful for library synthesis. However, our preliminary results indicate that the reaction has some limits when used in combination with aromatic compounds. The Hammett series study should provide some insight into the exact limitations of the reaction in this setting (Hammett LP, 1937). Additionally, synthesizing azides in-house from their corresponding bromides and using these to test reaction conditions would probably result in an overall improvement in performance.

Modeling

At this point, there has been little computational involvement in this project beyond the design of the fragment library. However, recent advances in

computational methods as applied to RNA systems by our group and others should make modeling a useful part of this project. In addition, these experiments will provide a unique opportunity to validate the computational techniques with a set of chemically diverse fragments, some of which are binders, all of which will have binding affinities determined using the same assays and have structural data.

Binding

It is obvious that the fragments are binding to the TAR RNA. However, a number of controls need to be run. For example, as stated in the Discussion Section, there needs to be further study of where the fragments are binding on the structure of the RNA. The affinity of the fragments that are shown to bind, of course, will also need to be determined. While we are working on optimizing generic chemistry that will enable use to easily combine the fragments, it is not clear at this time that the final set of reactions will be able to generate scaffolds that will align the fragments in their originally optimal positions. Even if the chemistry does not work, though, the fragments themselves could be used to generate lead-like molecules.

The specificity of the fragments may be a much larger, and in the end much more critical issue. The good and bad news of the fragment library that we have constructed is that so many members bind to the TAR. Having so many fragments bind means that there are many chemical moieties for us to select from as a starting point for the lead optimization. On the other hand, having such

a high hit rate, even from a fragment library, may indicate that the fragment library as designed may not result in lead compounds that are specific for the target of interest. The extent of this situation can be determined by screening against other RNA hairpins, like tRNA, during the various stages of lead optimization.

REFERENCES

- Aucagne, V and Leigh, DA. *Org Lett.* **8** (2006) 4505-4507.
- Barker, JJ. *Drug Discov Today.* **11** (2006) 391-404.
- XWIN-NMR. Bruker Biospin GmbH. Silberstreifen 4, 76287 Rheinstetten.
- Calnan, BJ, Tidor, B, Biancalana, S, Hudson, D and Frankel, AD. *Science.* **252** (1991) 1167-71.
- Carr, RA, Congreve, M, Murray, CW and Rees, DC. *Drug Discov Today.* **10** (2005) 987-92.
- MOE: Molecular Operating System. Chemical Computing Group. 1255 University Street, Suite 1600, Montreal, Quebec, Canada H3B 3X3.
- Choi, WJ, Shi, ZD, Worthy, KM, Bindu, L, Karki, RG, Nicklaus, MC, Fisher, RJ and Burke, TR. *Bioorg Med Chem Lett.* **16** (2006) 5265-5269.
- Davis, B, Afshar, M, Varani, G, Murchie, AIH, Karn, J, Lentzen, G, Drysdale, MJ, Bower, J, Potter, AJ, Starkey, ID, Swarbrick, TM and Aboul-Ela, F. *J Mol Biol.* **336** (2004) 343-356.
- MDL ISIS/Base. Elsevier MDL. 2440 Camino Ramon, Suite 300, San Ramon, CA 94583.
- Fejzo, J, Lepre, CA, Peng, JW, Bemis, GW, Ajay, Murcko, MA and Moore, JM. *Chem Biol.* **6** (1999) 755-69.
- Francois, B, Russell, RJ, Murray, JB, Aboul-ela, F, Masquida, B, Vicens, Q and Westhof, E. *Nucleic Acids Res.* **33** (2005) 5677-90.
- Frankel, AD and Young, JA. *Annu Rev Biochem.* **67** (1998) 1-25.

Golotvin, SS, Vodopianov, E, Lefebvre, BA, Williams, AJ and Spitzer, TD. *Magn Reson Chem.* **44** (2006) 524-38.

Hammett, LP. *J Am Chem Soc.* **59** (1937) 96-103.

Hann, M, Hudson, B, Lewell, X, Lifely, R, Miller, L and Ramsden, N. *J Chem Inf Comput Sci.* **39** (1999) 897-902.

Hsu, MC, Schutt, AD, Holly, M, Slice, LW, Sherman, MI, Richman, DD, Potash, MJ and Volsky, DJ. *Science.* **254** (1991) 1799-802.

Irwin, JJ and Shoichet, BK. *J Chem Inf Model.* **45** (2005) 177-182.

Johnson, EC, Feher, VA, Peng, JW, Moore, JM and Williamson, JR. *J Am Chem Soc.* **125** (2003) 15724-5.

Karn, J. *J Mol Biol.* **293** (1999) 235-54.

Kaul, M and Pilch, DS. *Biochemistry.* **41** (2002) 7695-7706.

Keseru, GM and Makara, GM. *Drug Discov Today.* **11** (2006) 741-748.

Kolb, HC and Sharpless, KB. *Drug Discov Today.* **8** (2003) 1128-37.

Lee, LV, Mitchell, ML, Huang, SJ, Fokin, VV, Sharpless, KB and Wong, CH. *J Am Chem Soc.* **125** (2003) 9588-9589.

Lepre, CA. *Drug Discov Today.* **6** (2001) 133-140.

Mayer, M and James, TL. *J Am Chem Soc.* **124** (2002) 13376-7.

Mayer, M and James, TL. *J Am Chem Soc.* **126** (2004) 4453-60.

Mayer, M and James, TL. "Discovery of ligands by a combination of computational and NMR-based screening: RNA as an example target." In Nuclear Magnetic Resonance of Biological Macromolecules, Part C. Ed. (2005) 571-587.

Mayer, M, Lang, PT, Gerber, S, Madrid, PB, Pinto, IG, Guy, RK and James, TL. *Chem Biol.* **13** (2006) 993-1000.

Montagnat, OD, Lessene, G and Hughes, AB. *Tetrahedron Lett.* **47** (2006) 6971-6974.

Murchie, AIH, Davis, B, Isel, C, Afshar, M, Drysdale, MJ, Bower, J, Potter, AJ, Starkey, ID, Swarbrick, TM, Mirza, S, Prescott, CD, Vaglio, P, Aboul-ela, F and Karn, J. *J Mol Biol.* **336** (2004) 625-638.

Pilch, DS, Kaul, M, Barbieri, CM and Kerrigan, JE. *Biopolymers.* **70** (2003) 58-79.

Renner, S, Ludwig, V, Boden, O, Scheffer, U, Gobel, M and Schneider, G. *ChemBioChem.* **6** (2005) 1119-1125.

Rishton, GM. *Drug Discov Today.* **2** (1997) 382-384.

Pipeline Pilot. SciTegic. 10188 Telesis Court. Suite 100, San Diego, CA 92121-4779.

Tornøe, CW, Christensen, C and Meldal, M. *J Org Chem.* **67** (2002) 3057-3064.

Vicens, Q and Westhof, E. *Biopolymers.* **70** (2003) 42-57.

Walters, WP and Murcko, MA. *Adv Drug Deliv Rev.* **54** (2002) 255-71.

Yu, XL, Lin, W, Pang, RF and Yang, M. *Eur J Med Chem.* **40** (2005) 831-839.

Zartler, ER and Shapiro, MJ. *Curr Opin Chem Biol.* **9** (2005) 366-370.

“There are seven sins in the world: Wealth without work; pleasure without conscience; knowledge without character; commerce without morality; science without humanity; worship without sacrifice; politics without principle.”

--Mahatma Gandhi

Conclusion

At this point in my thesis and in my graduate career, both Tack and Tom thought it would be appropriate for me to take a step back to summarize the lessons I have learned and to frame some open questions for the field. I am of a few minds about the best way to address this task, so I have decided to answer in a few ways:

From the perspective of senioritis:

Theory only works to predict the data that was used to build the model and, even then, only works as a result of fortuitous cancellation of errors. Experiments take too long and often are influenced by the questions asked and the moons being in proper alignment. Give up and become an event planner; the pay and the hours are better.

From the perspective of a scientist:

When I started my thesis work, I had hoped to address and solve some of the issues pertaining to drug design in general and for RNA targets in particular. While I have made some baby steps in that direction, I have perhaps more

appropriately begun to identify where the current models fail. Many of the more technical aspects of these issues have been addressed in one form or another in the chapters of this thesis. For example, it appears as though the scoring functions we are using are good enough to guide predictions but still have room for improvement for both proteins and nucleic acids. From the various types of sampling studies I have explored here, I would attribute this issue to improper modeling of charge polarization. For example, in RNA molecules in particular, the flexible backbone can modulate the local chemical environment and allow for fairly specific regulation of changes in protonation states. Even in simple interactions like hydrogen bonding, the charges are redistributed over the interacting atoms. These phenomena are also present in proteins, particularly in active sites. Modeling these types of interactions through the new polarizable force fields will hopefully serve both to improve the stability of nucleic acid simulations. It will then have to be determined how best to incorporate this type of information into docking simulations.

For all models, continual validation using current experimental data is critical to improving predictions. With the current technology, we can generate orders of magnitude more data both experimentally and computationally than even a few years ago. More information helps to generate statistics on the reproducibility of events. With the drive for more information, important, although seemingly mundane, details can be lost. Being able to predict the general trend is useful. However, I have found that the most pertinent information often comes from tracking down why the outliers do not fit the curve. For the future of

modeling RNA in particular, a big limitation is the lack of diverse experimental data. The vast majority of data collected to date is for HIV RNA molecules and the ribosome, making models and future experiments biased towards these systems. In addition, most RNA molecules are studied in their most minimal states. Only the ribosome has been solved in its entirety, and it is still not clear, for example, how the remainder of the HIV genome will affect the folding and tertiary structure of the TAR hairpin.

Finally, I think there is an inherent trap in comparing the current status of anything in the RNA world to that in the protein world. While it is useful to use proteins as a benchmark, it is also useful to remember that the study of RNA is truly in its infancy compared to proteins. To assume that the same experimental assays will apply to both systems without significant adaptation has been shown to be incredibly naïve (ask Nick Mills in the Guy Lab). The same is true for modeling and, perhaps more subtly, for drug design. We need to adjust our expectations away from “Let us find the nanomolar inhibitor of RNA,” particularly because we do not even know at this point if it is feasible to find a drug that is specific to any one RNA target. I honestly believe that some day, the FDA will approve a small molecule inhibitor of a protein-RNA interaction. However, I also feel that a great deal of basic experimental and theoretical research must be done into the biophysics of nucleic acids in order for drug design efforts to be a true success.

Random Bits of Wisdom:

The results of both theory and experiment are continually made to be more complicated than they need to be, particularly when the simple, logical answer is usually the truth (e.g., “It’s gotta be a bug.”). In the case of two competing models, the answer is usually some combination of both ends of the spectrum rather than one over the other. A good systems administrator can prevent years of frustration. It is critically important to know how to communicate your ideas to a wide range of people. It is equally critical—the underpinning of any successful collaboration—to understand not only other people’s work but their frame of reference. Always allow at least three times the amount of time you think you need to complete something. If you cannot answer the question “Why are you doing this?” at any point about your research, then you have gotten off track.

