UCSF
UC San Francisco Electronic Theses and Dissertations

Title
Simple models of protein and small molecule systems

Permalink
https://escholarship.org/uc/item/05b5v48h

Author
Brem, Rachel Beth

Publication Date
2000

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

Simple Models of Protein and Small Molecule Systems

by

Rachel Beth Brem

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Biophysics

in the

GRADUATE DIVISION

of the

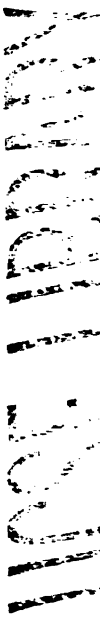
UNIVERSITY OF CALIFORNIA SAN FRANCISCO

UCSF LIBRARY

[redacted]

Date University Librarian

Degree Conferred: .....



Copyright 2000
by
Rachel Beth Brem

Acknowledgements

I think I owe the most to the people who helped me out in the roughest times—the first three years of grad school. The main pit crew back then was my parents, Amy and Jerry Brem; David Tobin; and Jen Goree. Also Danny Heap and Shi-Jie Chen, who treated me with respect in lab from the very beginning and were good friends to me throughout. I kept my head on straight thanks in large part to my work with Fred Gardner, who made *Synapse* what it was, and with the people in the Science and Education Partnership office.

After my orals, life and science got much better. I'm indebted to Hue Sun Chan for directing my solvation work via email from Toronto—Hue Sun has been exceptionally generous with his time and insight. And I learned everything I know about Monte Carlo simulations from long discussions with Kevin Silverstein over email. Thank you also to Patsy Babbitt, an important role model, career advisor, and friend; to Dick Shafer and David Agard for continuing to check in with me over the years; to Patricia Monohon for formatting this dissertation; and to my advisor Ken Dill.

Authorship

Chapter 1 is a reprint of material that was originally published in *Protein Engineering.* The paper includes Shaojian Sun, Hue Sun Chan, and Ken A. Dill as co-authors. Rachel Brem did all the off-lattice calculations that form the basis for this part of the dissertation.

Chapter 2 is a reprint of material that was originally published in *Protein Science.* The paper includes Ken A. Dill as a co-author. Rachel Brem was the principal researcher for this study; the co-author directed and supervised the research.

Chapter 3 is a reprint of material that was originally published in *The Journal of Physical Chemistry.* The paper includes Hue Sun Chan and Ken A. Dill as co-authors. Rachel Brem was the principal researcher for this study; the co-authors directed and supervised the research.

Chapter 4 includes Hue Sun Chan and Ken A. Dill as co-authors. Rachel Brem was the principal researcher for this study; the co-authors directed and supervised the research.

Chapter 5 includes Hue Sun Chan and Ken A. Dill as co-authors. Rachel Brem was the principal researcher for this study; the co-authors directed and supervised the research.

Ken Dill

Department of Pharmaceutical Chemistry

Abstract

Simple Models of Protein and Small Molecule Systems

by

Rachel Beth Brem

Doctor of Philosophy in Biophysics

University of California at San Francisco

Ken A Dill, Chair

This thesis uses simple models to address several questions in protein biophysics and liquid theory.

The work of the first chapter finds protein sequences of hydrophobic (H) and polar (P) residues that fold to a given target. We search H-P sequence space with a genetic algorithm, rewarding H-H contacts and penalizing exposed H's. In a lattice model, ~70% of designed sequences fold uniquely to their target. In biological proteins, designed and native sequences agree in ~70% of residue positions.

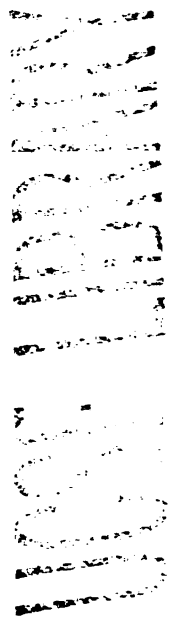
In the second chapter we ask, how severe are errors in protein-ligand scoring functions which disregard non-native ligand binding modes? We design such functions in a 2D lattice model of ligand binding, then train them against a "training set" of ligands. Errors in fitted parameters and free energies predicted by the scoring functions are as high as 25%. But if all ligands occupy the ground-state binding mode more than 50% of the time, errors are only ~10%.

The thesis's third chapter proposes a new way to estimate γ, the strength of the hydrophobic effect, from oil-water partitioning data. Our protocol harnesses oil-phase enthalpies *via* the van Laar model. In lattice tests, van Laar theory recoups model oil-phase energy

parameters to within 14%. The van Laar expression fits experimental enthalpies well ($r \geq 0.9$); the resulting estimate for γ is ~40 cal/(mol Å^2).

We address in the fourth chapter the conventional wisdom that Flory-Huggins (FH) theory should not be used on sphere-like molecules. We find evidence against this hypothesis: FH predicts trends in cyclic, branched, and linear alkanes, in the iodine solvation data of Hildebrand, and in disk models of liquids.

In the fifth chapter we study the physical basis for the success of mole-fraction theory (MF) in Hildebrand's iodine solvation entropies. We re-formulate MF in terms of excess quantities; within ~15%, this version of MF agrees with entropies of experimental liquids, Monte-Carlo-simulated hard disks, and scaled particle theory. Results suggest that MF may describe the solution structure of sphere-like molecules.



Contents

List of Figures x

List of Tables xii

1 Designing amino acid sequences to fold with good hydrophobic cores 1
1.1 Abstract. . . . 2
1.2 Introduction. . . . 2
1.3 Materials and Methods. . . . 3
1.3.1 The Burial Algorithm. . . . 4
1.3.2 The Sequence Evolution Algorithm. . . . 6
1.4 Results and discussion. . . . 10
1.4.1 Designing protein sequences. . . . 10
1.4.2 Comparing designed and natural sequences. . . . 12
1.4.3 Dimeric proteins. . . . 14
1.4.4 A lattice test of the GCSE method. . . . 15
1.4.5 Lattice tests of the sHP potential. . . . 16
1.4.6 Uniqueness. . . . 18
1.4.7 Hydrophobic core positions are more conserved. . . . 20
1.5 Conclusions. . . . 20
1.6 Acknowledgements. . . . 21

Bibliography 22

2 The effect of multiple binding modes on empirical modeling of ligand docking to proteins. 24
2.1 Abstract. . . . 25
2.2 Introduction. . . . 25
2.3 The model of ligand-receptor binding. . . . 27
2.4 Results and Discussion. . . . 30
2.4.1 Two-state errors during energy function parameterization. . . . 31
2.4.2 Two-state errors during free energy prediction. . . . 35
2.4.3 Two-state errors during the rank ordering of complexes. . . . 42
2.4.4 Can two-state error be avoided by predicting $p(gs)$? . . . 44
2.5 Conclusions. . . . 48
2.6 Acknowledgements. . . . 48

Bibliography 49

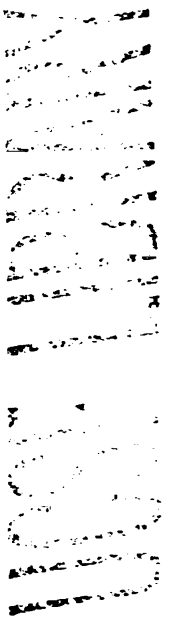
3 Extracting microscopic energies from oil-phase solvation experiments. 53
3.1 Abstract. . . . 54
3.2 Introduction. . . . 54
3.3 Computational and analytic methods. . . . 57
3.3.1 A lattice model of hydrocarbon liquids. . . . 57
3.3.2 Monte Carlo simulations. . . . 58
3.3.3 Properties of the pure lattice liquid. . . . 60
3.3.4 Analytic methods for extracting γ quantities. . . . 60
3.3.5 Choosing molecular volumes for alkanes. . . . 64
3.4 Results and discussion. . . . 64
3.4.1 The oil-phase enthalpy premise. . . . 64
3.4.2 Application of van Laar theory to solvation. . . . 67
3.4.3 The van Laar energy model extracts more accurate energy parameters than traditional methods. . . . 67
3.4.4 Application to experimental hydrocarbon data. . . . 70
3.4.5 Oil-water partitioning and the strength of the hydrophobic effect. . . . 72
3.5 Conclusions. . . . 77
3.6 Acknowledgements. . . . 78
3.7 Appendix. . . . 78

Bibliography 82

4 Tests of Flory-Huggins Theory on Fluids of Sphere-like Molecules 85
4.1 Abstract. . . . 86
4.2 Introduction. . . . 86
4.3 Computational and analytic methods. . . . 87
4.3.1 A 2D lattice model of liquids. . . . 87
4.3.2 Continuous models of liquids in 2D. . . . 87
4.3.3 FH theory. . . . 89
4.4 Results. . . . 90
4.4.1 Applying FH to a constant-V lattice model. . . . 90
4.4.2 Applying FH to simple continuum models. . . . 91
4.4.3 Applying FH to experimental data. . . . 97
4.5 Conclusions. . . . 100
4.6 Acknowledgements. . . . 100

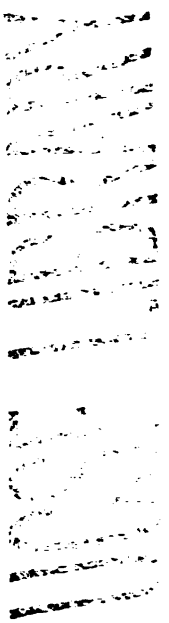
Bibliography 102

5 The Physical Basis for the Success of Mole Fraction Theory in Sphere-like Molecules 105
5.1 Abstract. . . . 106
5.2 Introduction. . . . 106
5.3 Computational and analytic methods. . . . 107
5.3.1 The toy lattice model. . . . 107
5.3.2 The hard disk model. . . . 108
5.3.3 Scaled particle theory. . . . 108
5.4 Results. . . . 109
5.4.1 Re-expressing the Hildebrand relation. . . . 109
5.4.2 Fluids of different sized molecules at the same pressure. . . . 112
5.4.3 Hildebrand's relation in hard disks. . . . 115
5.4.4 Applying scaled particle theory to iodine data. . . . 119



5.5 Conclusions. 121
5.6 Acknowledgements. 122

Bibliography **123**



List of Figures

1.1 Natural sequences and the two best (lowest energy) GCSE-designed sequences . . . 11
1.2 Native and GCSE-designed sequences superimposed on the native structures. . . . 13
1.3 Designing dimers. 15
1.4 Comparing sHP and HP potentials. 18

2.1 Lattice protein model and a sample lattice ligand model. 28
2.2 Ignoring multiple binding modes during fitting causes errors in energy parameters. 32
2.3 Regressions on stronger binders produce less negative parameter values. 33
2.4 Including multiple modes during fitting improves parameters. 34
2.5 Ignoring multiple binding modes in test sets causes errors in free energies. 36
2.6 Energy parameters fitted under the two-state assumption predict free energies inaccurately. 37
2.7 Accounting for multiple binding modes improves free energy prediction. 38
2.8 Ignoring higher-energy modes during regression can cause consistently nonphysical free energy predictions. 40
2.9 Accounting for multiple binding modes improves rank orders of test set complexes. 43
2.10 Rank-ordering errors occur only if differences in affinity errors are larger than differences in true affinities. 45
2.11 Occupancy of the ground state does not correlate with strength of binding. 46
2.12 Occupancy of the ground state correlates with enthalpy and entropy of binding. . 47

3.1 Modeling protein folding with oil-water partitioning. 55
3.2 Phase-like behavior of the 2-D lattice model. 61
3.3 Chemical potentials don't correlate with surface area. 68
3.4 FH extracts solvation energies from chemical potentials with moderate accuracy. . 70
3.5 Surface area alone doesn't extract an accurate energy parameter from the energy of insertion. 71
3.6 van Laar theory extracts accurate solvation energies from enthalpies. 72
3.7 Extracting solvation energies from experimental enthalpies with surface area. . . . 73
3.8 Extracting solvation energies from experimental enthalpies with van Laar theory. . 74
3.9 Extracting γ_{wat} from experimental data. 76
3.10 Experimental energies follow van Laar theory. 80

4.1 Correlating FH theory with simulation data. 88
4.2 Density fluctuations change with molecular shape and size. 92
4.3 Flory-Huggins predictions at constant V. 94
4.4 Snapshots of 2D fluids. 95

4.5 Flory-Huggins at constant P. 96
4.6 Flory-Huggins predictions correlate linearly with entropies of alkanes. 99
4.7 With knowledge of partial molar volumes, Flory-Huggins predictions correlate linearly with entropies of iodine. 101

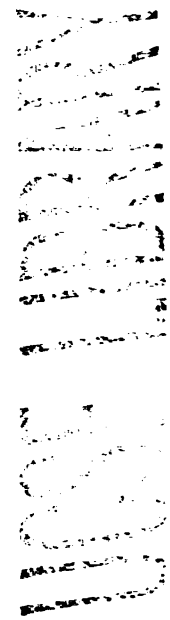
5.1 The mole-fraction expression fits iodine solvation data. 109
5.2 An equivalent treatment of Hildebrand's mole fraction data. 112
5.3 Constant pressure, density, and packing fractions in lattice fluids. 114
5.4 Snapshots of Monte Carlo simulations of hard disks. 116
5.5 Hard disks obey the Hildebrand relation. 117
5.6 Insertion of solutes orders large solvents more. 119
5.7 2-D SPT entropies obey the Hildebrand relation. 120
5.8 Applying SPT to iodine data. 122

List of Tables

1.1 PDB protein names. 5
1.2 Breakdown of average sequence assignment by degree of exposure and by residue type 12
1.3 Number of encodable $n = 16$ conformations on square lattices by exact enumeration 17
1.4 Numbers of $n = 16$ sHP sequences that fold uniquely to HP native structures on square lattices 17

3.1 Variation in solvation energy parameters due to SASA calculation and molecular shape. 75

5.1 Parameters used in SPT applied to experimental data. 121



Chapter 1

Designing amino acid sequences to fold with good hydrophobic cores

Shaojian Sun, Rachel Brem, Hue Sun Chan and Ken A. Dill

This chapter was originally published as "Designing amino acid sequences to fold with good hydrophobic cores," *Protein Engineering* **8**(12), 1205–1213 (1995).

1.1 Abstract.

We present two methods to design amino acid sequences of proteins that will fold to have good hydrophobic cores. Given the coordinates of the desired target protein or polymer structure, the methods generate sequences of hydrophobic (H) and polar (P) monomers that are intended to fold to those structures. One method designs hydrophobic inside, polar outside; the other minimizes an energy function in a sequence evolution process. The sequences generated by these methods agree at the level of 60–80% of the sequence positions in 20 proteins in the Protein Data Bank. A major challenge in protein design is to create sequences that can fold uniquely, *i.e.*, to a single conformation rather than to many. While an earlier lattice-based sequence evolution method was shown not to design unique folders, the present method generates unique folders in lattice model tests. These methods may also be useful in designing other types of foldable polymers not based on amino acids.

1.2 Introduction.

We develop here inverse folding algorithms that aim to design amino acid sequences to fold to desired "target" structures. Recently Kamtekar *et al.* [1993] have shown that a simple binary sequence code of hydrophobic (H) and polar (P) amino acids may define native chain topologies in proteins. Kamtekar *et al.* started with the native coordinates of a known biological 4-helix bundle and then designed sequences to fold to it. Their design strategy was simply to bury H monomers and expose P monomers. In the present paper, we give two automated procedures for performing such hydrophobic/polar designs. The first is a direct transcription of the procedure of Kamtekar *et al.*: it assigns an H monomer to any solvent-inaccessible position in the given structure, and P otherwise. We call it the Burial algorithm. It simply codifies the standard lore that hydrophobics should be buried, and just serves as a reference for our second method. The second method is a variant of the "evolutionary design" procedure of

Shakhnovich and Gutin [1993]. Following statistical mechanical terminology, we call ours the Grand Canonical Sequence Evolution (GCSE) method, in contrast to the canonical method of Shakhnovich and Gutin. We perform both types of designs on the x-ray structures of 20 known proteins from the Protein Data Bank (PDB). We then test the GCSE method in a simplified lattice representation of proteins, to demonstrate why we believe our designed sequences will fold to target structures and no other structures. When we compare the sequences designed by these algorithms to the known native sequences of biological proteins, agreement ranges from 57-81%. These methods may be useful tools for designing monomer sequences that will fold to specific target conformations of proteins and other polymers.

Inverse folding [Drexler 1981, Ponder and Richards 1987, Bowie *et al.* 1991, Lee and Levitt 1991, Yue and Dill 1992, Shakhnovich and Gutin 1993, Shakhnovich 1994, Gutin *et al.* 1995] can be thought of as a "bead coloring" problem. Start with the desired coordinates of a chain of "colorless beads," *i.e.*, generic amino acids that do not yet have sidechain structural identities. The design process then "paints" each bead a "color" representing each of the 20 amino acids. Any such design must solve two problems [Yue and Dill, 1992]: 1) *Positive design*: it must generate sequences that will fold to the given structure, and 2) *Negative design*: it must design sequences that will not fold to other low energy structures. We show here that one of our algorithms provides good negative design in lattice model tests.

1.3 Materials and Methods.

Our algorithms are given the coordinates of a desired target structure. Since our aim here is to test our methods, the coordinates we use are those of known proteins from the PDB. Thus we "invent" sequences for some known protein structures and compare them to the natural sequences. Of course it would defeat our purpose to input the all-atom sidechain coordinates for the target structure, because then any good sequence design method would just recreate the

natural sequence. In the present paper we ask instead, can we design just the hydrophobic and polar parts of the sequence? For such a problem, when we input the coordinates of the target structure we must represent every sidechain identically, so that they are "colorless." Hence we input coordinates for the target structure in the following way. We retain all backbone atoms but replace side chains with a centroid, or spherical virtual atom, at the C^{β} position. Sidechain centroids are placed along the native C^{α}-C^{β} bond vector at a uniform distance of 3.0Å from the C^{α}, and have a radius of 2.0Å. For residue positions occupied by glycine in the target structure, no native C^{α}-C^{β} bond vector is available, so we use a standard rotamer direction (bond angle 107.63° and dihedral angle 239.11°) for placing our centroid. Our aim is to design HP sequences that should fold to roughly these target conformations.

1.3.1 The Burial Algorithm.

In this method, we use the given coordinates of the target structure to determine the degree of burial of each amino acid, then assign H to buried positions and P to exposed sites. Each amino acid in the target structure is defined as exposed or buried according to whether its sidechain bead has more or less than 10Å^2 of exposed surface area, respectively. A residue is "painted" H if it is buried by this definition, and is painted P otherwise. This algorithm is fast because it depends only linearly on the chain length of the target structure. We designed sequences for 20 proteins using the Burial algorithm. Table 1.1 shows the fraction of sequence positions for which the Burial algorithm assignment of H or P agrees with the polar/ nonpolar identity of the residue in the true native structure, where hydrophobic = A, C, I, L, M, F, W, Y, and V, and polar = R, N, D, E, Q, G, H, K, P, S, and T in the one-letter code of amino acids. The Burial algorithm designs sequences that are identical to the natural sequences (in terms of hydrophobic and polar assignment) at about 60-80% of the residue positions. This is consistent with the many earlier studies indicating that the cores of globular proteins are predominantly hydrophobic

Table 1.1: PDB protein names, (a) chain length, (b) percent identity of HP sequences designed by the Burial algorithm compared with natural sequences, (c) percent identity of HP sequences designed by the GCSE algorithm compared with the natural sequences. "Monomer" and "dimer" refer to sequence design calculations performed using the monomer structure alone as the target and the full dimer structure as the target, respectively.

Protein	(a) Length	(b) Correct, Burial (cutoff=10Å)	(c) Correct, GCSE
1aaj	105	66%	66%
1aba	87	81%	81%
1aps	98	76%	72%
1arr monomer	53	60%	62%
1arr dimer	106	70%	73%
1bba	36	61%	58%
1bbl	37	62%	68%
1bov	69	67%	74%
1brq	174	71%	68%
1cis	66	67%	64%
1cmb monomer	104	59%	62%
1cmb dimer	208	68%	70%
1hel	129	78%	74%
1ifb	131	69%	76%
1kba monomer	66	71%	72%
1kba dimer	132	80%	73%
2gb1	56	77%	80%
2hpr	87	78%	78%
2il8	71	73%	77%
256b	106	79%	81%
3cln	143	68%	62%
3rn3	124	71%	81%
3trx	105	77%	80%
average		72%	73%

and the surfaces are extensively polar [Chothia 1976, Janin 1979, Wertz and Scheraga 1978, Meirovitch and Scheraga 1980, Guy 1985].

1.3.2 The Sequence Evolution Algorithm.

We also develop another bead painting algorithm. Following earlier work on the inverse folding problem [Ponder and Richards 1987, Lee and Levitt 1991, Shakhnovich and Gutin 1993, Shakhnovich 1994, Gutin *et al.* 1995], our second method performs positive design by the minimization of a conformational energy function through sequence mutations. Our approach differs from earlier works in the following respects: (i) Rather than using 20 amino acid types [Ponder and Richards 1987, Lee and Levitt 1991, Shakhnovich 1994] we consider just the two monomer types, H and P. In this way we can explore the design of other possible foldable polymers, in addition to proteins. We can test the method on lattice model examples for which the true global minimum in free energy is known in both sequence and conformational spaces. (ii) We do not fix the amino acid composition. In analogy with statistical mechanics, we call our approach a "grand canonical" method, which allows variable composition, in contrast to the "canonical" method of Shakhnovich and Gutin [1993], which requires fixed composition. Shakhnovich and Gutin recognized that when using the HP potential alone, sequence evolution methods will converge to the homopolymer of all hydrophobic units, since that gives the lowest achievable energy. To avoid that problem, they fixed the sequence composition. Theirs is a "bead swapping" process which iteratively reassigns H and P monomers to different positions. The total number of H monomers and the total number of P monomers remain unchanged in their method. In the canonical method, it is necessary to know in advance the correct H and P composition. The use of the wrong HP composition can lead to sequence designs that do not fold uniquely [Yue *et al.* 1995]. In contrast to their canonical bead swapping method, ours involves bead painting, and explores all possible HP compositions. It does not converge to homopolymer sequences because of a term we add to the energy function, justified below. We show below that our method often gives sequences that fold uniquely to the desired native structure, when tested in lattice models. (iii) Rather than a single sequence evolution Monte Carlo search method of the type used by Shakhnovich and coworkers [Shakhnovich and Gutin

1993, Shakhnovich 1994, Gutin *et al.* 1995], we use a genetic algorithm (GA) to search sequence space which creates many different sequences simultaneously. This provides useful statistics. (iv) Unlike lattice model methods [Yue and Dill 1992, Shakhnovich 1994, Gutin *et al.* 1995], our method allows real-space coordinates, and so can be designed and tested on real molecules.

We assumed a very simple fitness function. The twenty amino acids were divided into two classes: H, for hydrophobic, and P for polar or other. The sequence fitness energy function is

$$E = \varepsilon \sum_{i<j}^{N} h_{ij} + \sigma \sum_{j}^{N} s_j, \tag{1.1}$$

where $\varepsilon = -2$ and $\sigma = +1$. Here i and j are indices for the positions in the sequence of two monomers. The first term is negative (favorable) if both the ith residue and the jth ($\neq i$) are hydrophobic and are in nonlocal contact (here nonlocal means $| i-j | > 2$). In full-space studies, we define the distance-dependent function

$$h_{ij} = h_{ij}(r_{ij}) = \frac{1.0}{1.0 + e^{(r_{ij}-6.5)/1.0}}, \tag{1.2}$$

with the inter-residue distance r_{ij} between hydrophobics in Å. This function is sigmoidal, representing more softness in the potential than a step function. The cutoff distance of the sigmoidal function is 6.5 Å. In the lattice model test of our methods, we set $h_{ij} = 1$ when two contacting lattice residues are hydrophobic, and $h_{ij} = 0$ otherwise. The second term in Eq. (1) is the hydrophobic residue-solvent interaction, representing an additional tendency for avoidance of contacts between solvent and the hydrophobic monomers. It is positive (unfavorable) if residue j is hydrophobic and contacts a solvent site, in which case $s_j = 1$; otherwise $s_j = 0$. In the full-space model, the sum in the second term of Eq. (1) is over solvent contacts. We used the program ACCESS [Richmond, 1984] to define solvent accessibilities of the amino acids. We assumed one solvent contact for every 3Å^2 of surface area exposed of a sidechain bead, and we

used the default 1.4Å probe. In the lattice representation, the sum in the second term of Eq. (1) was over all hydrophobic residue faces not involved in contacts.

We explore sequence space with a genetic algorithm similar to one used in folding algorithms [Sun 1993, Dandekar and Argos 1992]. It contains several steps [Sun 1992]. 1) Randomly create a population of HP sequences of a given length, 2) Apply genetic operations to this initial sequence population to generate about 10-fold more sequences, 3) Use the fitness function to select the fittest (lowest energy) among these new sequences, 4) Iterate through steps (2) and (3) until the energy function reaches a minimum and sequences stop improving. This differs from the Monte Carlo sequence design method [Shakhnovich and Gutin 1993, Shakhnovich 1994, Gutin *et al.* 1995] in which the search operates only on a single evolving sequence. The genetic operations used are the following:

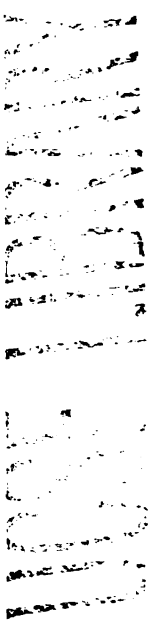
- Replication. Copy all p fittest sequences of the last generation into the the next generation. Thus p is the replication population.

- Mutation. Each of the p fittest sequences of the last generation is mutated and then copied into the next generation. The mutation sites are chosen randomly within a sequence, and we allow a maximum of three point mutations, *i.e.*, H to P or P to H, per sequence. In the mutation operation we also allow two monomers to exchange their sequence positions (permutation). The mutation population is $2p$.

- Crossover. There are different ways to perform crossover. Here we use random monogamy: the crossover operation randomly selects two sequences and cuts them at a random site, then crosses over and pastes in the parts to generate two new sequences for the next generation. The crossover sequences are created randomly between the replication sequences and the mutation sequences. Therefore the total number of new sequences generated by the crossover operation is $2p$. Although crossover could be allowed at multiple sites, we use only one crossover site for any pair of sequences.

- Shuffle. The shuffle operation interchanges two pieces of the chain from one place in the sequence to another. It conserves composition. We randomly choose the lengths (up to a maximum of 8 residues) and starting positions of the shuffled pieces. To shuffle segments at the ends of sequences, we used a periodic boundary condition. Hence if a segment four residues long is to be shuffled, and it begins at the protein's last residue, the final segment to be shuffled will include residues 1, 2, and 3. The shuffled sequences are generated from the p fittest sequences of the last generation. The shuffle population is p.

- Reverse shuffle. Reverse shuffle is identical to the shuffle operation, except that the sequence directionality is changed. Sequence piece $ABCD...$ becomes $...DCBA$ at a different sequence position. The reverse shuffle population is p. Reverse genetic shuffles are not biological, but for computer algorithms like the present one, this is not important.

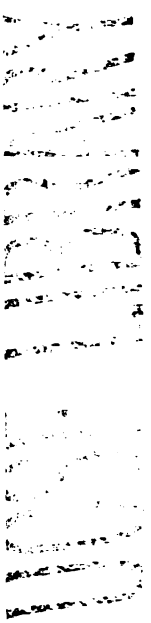


We use an initial sequence population size of $p = 500$. All the initial sequences are created randomly: at each position, H or P is initially chosen with equal probability. After the replication, mutation, crossover, shuffle and reverse shuffle operations, $7p$ sequences are created from the fittest p sequences of the last generation, and their fitness values are computed according to the energy function. In order to avoid possible premature convergence in the sequence space, a *share* mechanism [Goldberg 1989, Sun 1993] is applied in the sequence selection process. The number of identical sequences in the selected population was restricted to no more than 2. All newly created sequences are sorted according to their energy values. The energies of the sequences usually converge in about 120-180 generations. Converged sequences are about 95% identical among themselves. For targets longer than ~ 60 residues, we often find that more than 50% of the final sequences have fitness values within 1-5% of the lowest energy.

1.4 Results and discussion.

1.4.1 Designing protein sequences.

We used 20 proteins from the PDB as target structures. We designed an ensemble of HP sequences for each target structure and then compared our invented sequences with the natural sequences, translated to the HP code. Table 1.1 compares our designs from the GCSE and Burial algorithms with the natural sequences. Figure 1 gives the lowest-energy sequences from the GCSE algorithm for several example proteins. Table 1.2 correlates average success and failure at a given residue with its exposure to solvent. In sequences designed by either the Burial or the GCSE algorithm, when P's are assigned to exposed positions, 75–80% agree with the native sequence at that position, and 67–69% of H's assigned to buried positions agree with the native. These data illustrate the extent to which the "hydrophobic inside, polar outside" principle is borne out in our test set of proteins. Table 1.2 also shows that while the GCSE algorithm seldom assigns P's to buried positions (only 2.8% of all assigned P's are buried), it often assigns H's to exposed positions (25.4% of all assigned H's are exposed). Hence the physics described by GCSE is different from that of the Burial algorithm. Using GCSE, the success rates vary from 58 to 81%, averaging about 73% among the biologically relevant forms of the proteins we tested. The highest success rates are, not surprisingly, for the most globular proteins having the most buried hydrophobic cores.



Should we expect better? It is not clear. Presumably nature designs sequences not just for structure and stability, but also for function and perhaps to optimize folding kinetics. Hence even if we designed proteins using nature's true potential function, we might not achieve much greater similarities to natural proteins. Moreover, natural proteins may not be optimal. For example, stability can be improved when buried polar groups are replaced by hydrophobic residues [Schildbach *et al.*, 1995] or when exposed hydrophobic residues are made polar [Pakula and Sauer, 1990]. Hence some buried polar groups in proteins, and some exposed

(A) 1hel

```
1                                                                                80
KVFGRCELAAAMKRHGLDNYRGYSLGNWVCAAKFESNFNTQATNRNTDGSTDYGILQINSRWWCNDGRTPGSRNLCNIPC
HPHPHHPHHPHHHPHPHHPHPPHHHHHHHHHHHPHHHPHPHPHPPPHPHPHHHHHHHHHHHHHPHHPHPPHPPHPHPHPH
HPHPHPPHHPHHPPHPHPPHPPHHHHHHHHHHHPHHHPHPHPHPPPHPHPHHHHHHHHHHHHHPHHPHPPHPPHPHPHPH

81
SALLSSDITASVNCAKKIVSDGNGMNAWVAWRNRCKGTDVQAWIRGCRL
PHHHPPPHPHHHPHHHPHHPPPPPHPHHPHHPPPHPPHPHPPHHPPHPH
PHHHPPPHPHHHPHHHPHHPPPPPHPHHPHHPPPHPPHPHPPHHPPHPH
```

(B) 3rn3

```
1                                                                                80
KETAAAKFERQHMDSSTSAASSSNYCNQMMKSRNLTKDRCKPVNTFVHESLADVQAVCSQKNVACKNGQTNCYQSYSTMS
PPPPHPPHPPHHHHPPHPPHPPPPHHHPHHPPHPHHPPPHHPPHHHHHHPHPPHHPHHPHPPHPHPHPHPHHHHHPPPHH
PPPPHPPHPPHHHHPPHPPHPPPPHHHPHHPPHPHHPPPHHPPHHHHHHPHPPHHPHHPHPPHPHPHPHPHHHHHPPPHH

81
ITDCRETGSSKYPNCAYKTTQANKHIIVACEGNPYVPVHFDASV
HHHHPHHPPHPPPPHHHPHPHHPHHHHHHHHPPPHHHHPHHHHH
HHHHPHHPPHPPPPHHHPHPHHPHHHHHHHPPPPPHHHPHHHHH
```

(C) 256b

```
1                                                                                80
ADLEDNMETLNDNLKVIEKADNAAQVKDALTKMRAAALDAQKATPPKLEDKSPDSPEMKDFRHGFDILVGQIDDALKLAN
HPHPPHHPHHPPHHPPHPPHPPHPPHHPHHPPHHPHHPHHHPHPHPPHPPPPPPHPPHPPHPPHHPPHHPHPPHHPHHP
HPHPPHHPHHPPHHPPHPPHPPHPPHHPHHPPHHPHHPPHPPHPHPHPPPPPPHPPHPPHPPHHPPHHPHPPHHPHHP

81
EGKVKEAQAAAEQLKTTRNAYHQKYR
PPPHPPHHPHHPPHPPHPPPHPPPHP
PPPHPPHHPHHPPHPPHPPPHPPPHP
```

(D) 1aba

```
1                                                                                80
MFKVYGYDSNIHKCGPCDNAKRLLTVKKQPFEFINIMPEKGVFDDEKIAELLTKLGRDTQIGLTMPQVFAPDGSHIGGFD
PHPHHHHPHPHPPHPPHPPHHPHHPPPPPPHPHPHHHPPPPPPPPPPPPPHPPHPPPPPPPHPHHHHHHPPPPHHPPHP
PHPHHHHPHPHPPHPPHHPHHPHHPPPPPPHPHPHHHPPPPPPPPPPPPPHPPHPPPPPPPHPHHHHHHPPPPHHPPHP

81
QLREYFK
PHPPPHP
PHPPHHP
```

Figure 1.1: Natural sequences and the two best (lowest energy) GCSE-designed sequences, for (A) lysozyme (PDB identifier 1hel), (B) RNAse A (3rn3), (C) cytochrome b652 (256b), and (D) a glutaredoxin mutant (1aba). H = hydrophobic; P = polar.

nonpolar groups, may be accidental. It is worth noting the possibility that the GCSE method may cover sequence space more broadly, and design sequences that are more stable in their native structures, than nature does.

Table 1.2: Breakdown of average sequence assignment by degree of exposure and by residue type

	Burried position		Exposed position	
	assignment correct (%)	assignment incorrect (%)	assignment correct (%)	assignment incorrect (%)
GCSE-designed				
Total assigned Hs	51.6	23.0	12.8	12.6
Total assigned Ps	2.1	0.7	77.9	19.3
Burial-designed				
Total assigned Hs	67.4	32.6	0.0	0.0
Total assigned Ps	0.0	0.0	75.1	24.9

Results were averaged over all 20 naturally occurring proteins used in this study (see Table 1.1). Here, 'Buried position' indicates that either an H or a P residue is assigned at a position in the structure with $< 10\,\text{Å}^2$ of exposed surface area. Likewise, exposed residues are defined here to have $> 10\,\text{Å}^2$ of exposed surface area. An assignment is 'correct' if it agrees with the native sequence, i.e. if the assigned residue is H, the native sequence has an H at that position.

1.4.2 Comparing designed and natural sequences.

Figure 2 compares the native and designed sequences for six proteins, threaded onto their corresponding structures. The GCSE algorithm designs good hydrophobic cores and polar surfaces. Figure 2a-f shows that the discrepancies with natural sequences are due to exposed hydrophobics and buried polar residues. Figure 2g-h shows two examples of more global errors in the GCSE assignments, in a glutaredoxin mutant (1aba): the algorithm creates a hydrophobic "core" even at a protrusion near the surface of a protein, where there are multiple residue contacts. Despite these errors, the designed sequence for 1aba was 81% identical to the native (Table 1.1).

The method has a few other notable failures. It assigned all P residues to BPTI (PDB identifier 1bba), finding no position buried enough to be assigned an H. Functionally, inhibitors such as BPTI need exposed hydrophobic surface area, so they can bind active sites of their respective enzymes. The GCSE algorithm does not account for such functional constraints. For the hybrid protein constructed by Osmark *et al.* [1993], with secondary structural elements from chymotrypsin inhibitor-2 and subtilisin (1cis), the algorithm succeeded at only 63% of the monomer positions (Figure 2i-j and Table 1.1). Here buried polar residues like proline 82

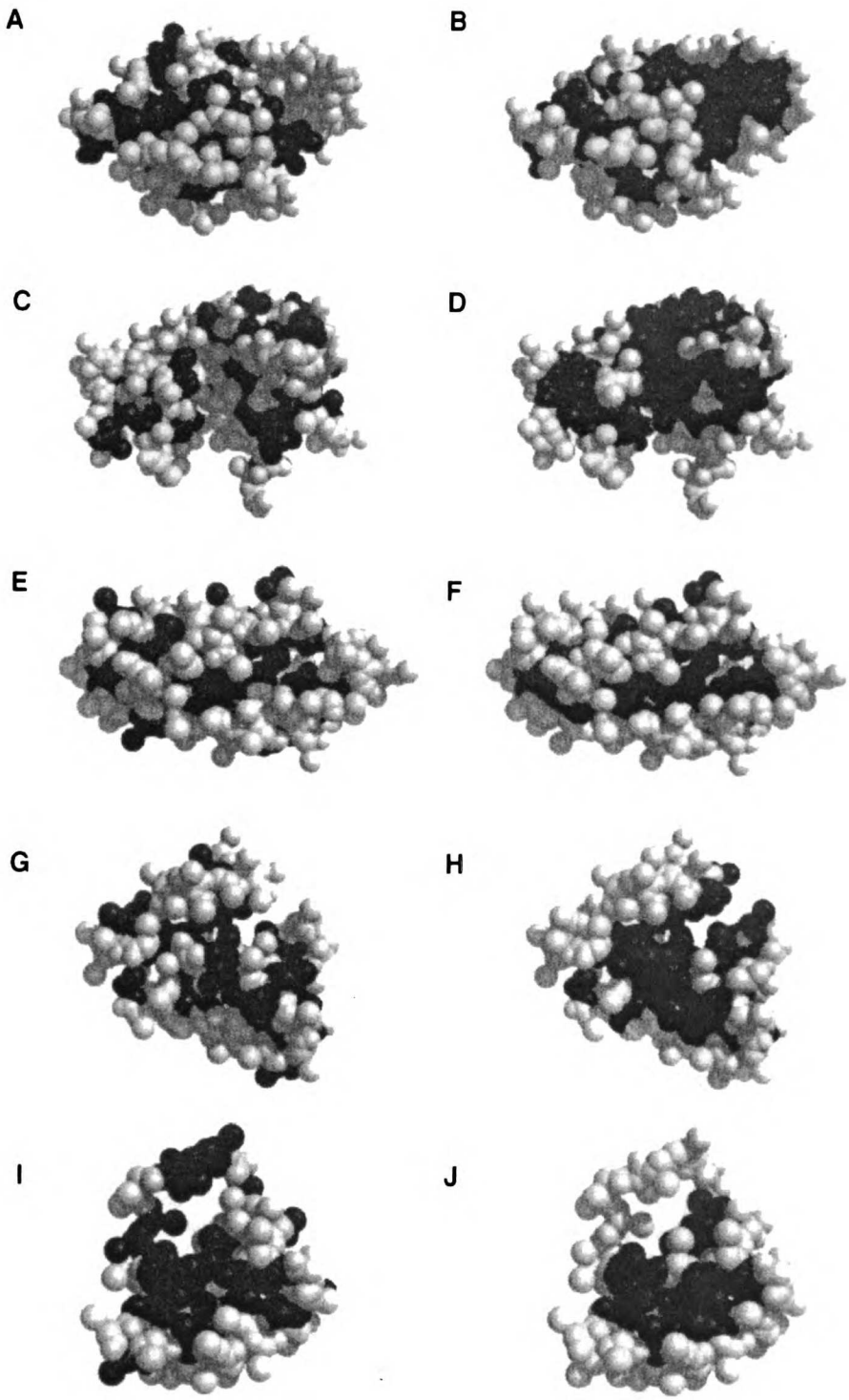
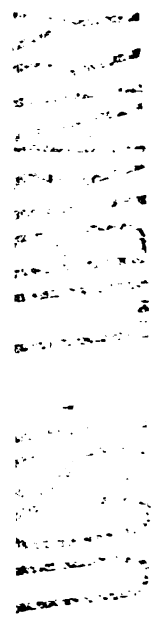


Figure 1.2: Native (left), and GCSE-designed sequences (right) superimposed on the native structures. Black = hydrophobic, white = polar. (A–B) lysozyme (1hel), (C–D), RNAse A (3rn3), (E–F) cytochrome b652 (256b), (G–H) glutaredoxin mutant (1aba), (I–J) subtilisin-chymotrypsin inhibitor (1cis).

and glycine 85 were mis-assigned. Such residues play a role in structure as well as stability. Also, as with BPTI, the native molecule has exposed hydrophobic surface in an inhibiting loop for functional reasons, which GCSE designs to be polar. As expected, GCSE does a relatively poor job (62%, Table 1.1) on calmodulin (3cln), because of its unusual shape. Hence the GCSE algorithm performs best on the most globular proteins.

1.4.3 Dimeric proteins.

We also designed three dimers: the ARC repressor (1arr), the MET repressor (1cmb), and kappa-bungarotoxin (1kba). Comparisons in Figure 3 and Table 1.1 show that when the full dimer structure of 1arr was the target, sequence designs were much closer to the native than when only the monomer was the target (73% versus 62% of residues correct). The strips of hydrophobic residues down the center of the horseshoe-shaped monomer contact each another in the dimer, but in the monomer they appear exposed to solvent. Hence the program judged these positions to be suitable for P residues in the monomer but assigned them H when given the dimer coordinates. Results with 1cmb were similar.



Sequences designed for the 1kba monomer were almost identical to those designed when the full dimer was the target; both designs are ~ 72% successful. The explanation is simple: the dimer interface is composed of backbone atoms; sidechains are mainly tucked into monomer interiors. Hence, to the inverse folding algorithm, these residues look buried in each monomer, so dimerization has little effect on HP assignment in such a case.

While the GCSE method involves considerably more computer time than the Burial algorithm, and only slightly higher success rates, it offers one significant advantage. Because GCSE is based on an energy function, rather than on a structural property like burial, it is readily generalizable and can be used to explore and compare other energy functions. This may offer an opportunity to learn about the relative importance of other types of energies beyond simple hydrophobic and polar energy functions.

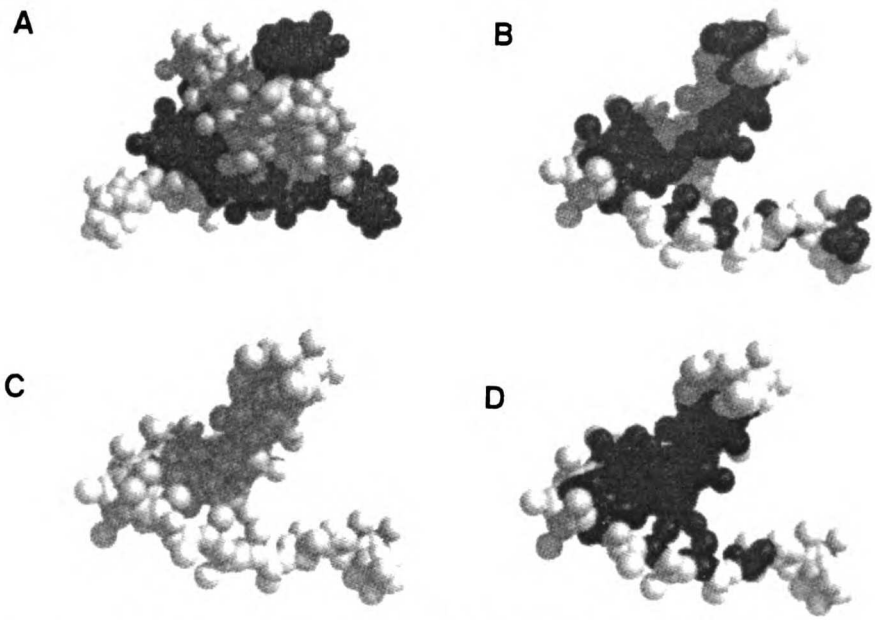


Figure 1.3: Designing dimers. (A) Illustration of the dimer interface of the ARC repressor as a dimer, with one monomer in black and the other in white. (B) Color-coded hydrophobic and polar illustration (black = hydrophobic, white = polar) of the native sequence and (C) the GCSE-designed lowest-energy sequence when the ARC monomer was used as the target. (D) The GCSE-designed sequence when the ARC dimer was used as the target. In (C) all residues appeared exposed to the algorithm, and were assigned as polar; whereas with the full dimer as the target in (D), the residues now buried are assigned to be hydrophobic.

1.4.4 A lattice test of the GCSE method.

Will the designed sequences fold to unique native conformations? The only computational way to determine uniqueness of folding at the present time is through the use of lattice models, for which the full conformational space can be explored. Accordingly, we used GCSE to design lattice model HP sequences [Lau and Dill 1989, Chan and Dill 1991], then studied whether those sequences folded uniquely. In the lattice model, a protein is represented by a linear self-avoiding path of a chain of "beads" on a two-dimensional square lattice or a three-dimensional simple cubic lattice. Side chains are not considered in this model, and H or P monomers along the chain are centered at their sites.

Before describing the design results in the lattice system, we first address a technical issue involving the potential function in Eq. (1), which we call the sHP potential because of its

additional solvent term relative to the well-studied HP model [Lau and Dill, 1989, Chan and Dill, 1991, Yue and Dill, 1992, Chan and Dill 1994, Dill *et al.*, 1995]:

$$E = \varepsilon \sum_{i<j}^{N} h_{ij}. \tag{1.3}$$

The surface term σ in Eq. (1) is empirical and serves to prevent evolution toward the all-H sequence. Does the introduction of the extra solvent term change the model? We show in the section below that while the extra solvent term in the sHP potential imparts the ability to do grand canonical design, where we do not have to know the HP composition in advance, it nevertheless does not seriously alter the HP model.

1.4.5 Lattice tests of the sHP potential.

We performed exhaustive folding simulations on the 2D lattice for the 2^{16} different sequences of chain length $n = 16$, using either the HP or the sHP potential. We use folding enumerations to address questions about the two energy functions described above. Does a given HP sequence fold to a different native structure, or with a different degeneracy, when the potential is sHP rather than HP? Our exact enumeration results show that 98.8% (1520 out of 1539) of those sequences that fold uniquely under HP also fold uniquely under sHP. However, a randomly selected sequence is more likely to fold uniquely under the sHP potential than under HP: 13.9% (9107 sequences) of sequences fold uniquely under sHP, while 2.3% (1539 sequences) fold uniquely under HP (Table 4). In the sHP potential, those structures with the lowest energy for a given sequence must not only maximize HH contacts but also minimize H-solvent contacts. As fewer configurations are likely to meet both criteria, the lowest-energy state is more likely to fold *uniquely* under the sHP potential than under HP. Suppose we are given a particular lattice native structure, and a set of sequences that fold uniquely to it (*i.e.*, its convergence set [Chan and Dill, 1991]). If a sequence is the lowest-energy sequence for that structure when evaluated with the sHP potential, we call it a lowest-sHP-energy sequence. As demonstrated in Table 4, the

Table 1.3: Number of encodable $n = 16$ conformations on square lattices by exact enumeration

Topological contacts t	Unique in sHP[a]	Unique in HP[b]	Unique in sHP and HP[c]	Lowest energy sHP[d]
6	1	0	—	—
7	337	163	163	159 (97.5%)
8	884	273	271	261 (95.6%)
9	53	20	20	20 (100%)
Total	1275	456	454	440 (96.5%)

[a]Number of conformations with t intrachain contacts that are unique native structures for the sHP potential.
[b]Number of conformations with t intrachain contacts that are unique native structures for the HP potential.
[c]Number of conformations that are encodable by at least one identical sequence for both potentials.
[d]Number of conformations among those in c with at least one sequence that encodes the given conformation uniquely for both potentials and has the lowest native energy among all converging sHP sequences.

Table 1.4: Numbers of $n = 16$ sHP sequences that fold uniquely to HP native structures on square lattices

Topological contacts t	sHP[a]	Lowest energy sHP[b]	Lowest energy sHP and HP[c]
7	887	223	167 (74.9%)
8	3265	367	268 (73.0%)
9	1200	34	26 (76.5%)
Total	5352	624	461 (73.9%)

The number of interchain contacts is t.
[a]Total number of sHP sequences that encode HP native structures.
[b] Number of sHP sequences that have the lowest energy among all converging sHP sequences that fold to the same HP native structure.
[c]Number of lowest sHP energy sequences among those in b that would also fold uniquely to the same native structure for the HP potential.

majority (73.9%) of these lowest-sHP-energy sequences uniquely encode their target structures when folded not only under the sHP potential but also under HP. Minimizing sHP energy in sequence design thus guarantees a unique folder in these lattice model tests.

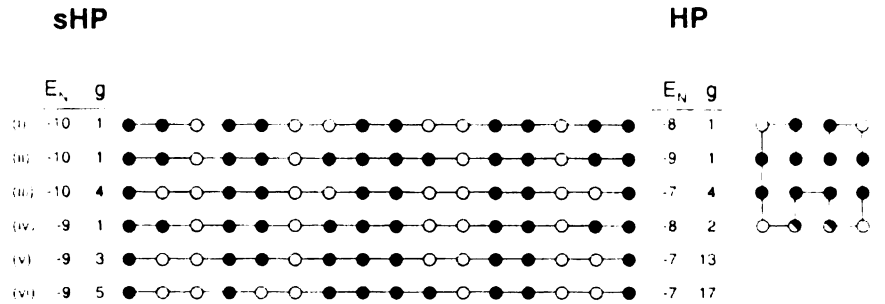


Figure 1.4: Comparing sHP and HP potentials. Six sequences are generated by the inverse folding genetic algorithm described in the text to fold to the two-dimensional target structure on the right *with the sHP potential* (H: black bead, P: white bead). The native energy E_N and degeneracy g of the sequences are determined for both potentials by exact enumeration. The target structure is either the unique native structure ($g = 1$) or one of many lowest-energy conformations ($g > 1$) of all 6 sequences for both the sHP and HP potentials. Two of the sequences with the lowest sHP energy of -10 fold uniquely to the target structure with *both* potentials (i, ii). The target structure is shown in these two sequences, and the half-filled beads are P's for sequence (i) and are H's for sequence (ii). However, sequences (iii) with the lowest sHP energy do not fold uniquely, while sequence (iv) with higher sHP energy folds uniquely with the sHP potential. Sequences with lowest sHP energies tend to fold uniquely to the target structure while those with higher sHP energies tend not to fold uniquely with the HP potential.

1.4.6 Uniqueness.

Enumeration provided us with the lowest-energy sequence for every structure possible with 16-mers. When we performed GCSE design on the structures, the algorithm nearly always produced these lowest-energy sHP sequences. Hence GCSE converges on the correct answer in these lattice model tests. Since these lowest-sHP-energy sequences also usually fold uniquely (see above), GCSE-designed sequences are often, but not always, unique folders. Figure 4 shows that GCSE is successful for the given target structure. Two of the three lowest-sHP-energy sequences identified by the GCSE method uniquely encode for the target structure when folded under both the sHP and the HP potentials. In contrast, the simpler Burial algorithm does not guarantee unique folders. For instance, assigning P's to all exposed sites of the target structure in Figure 4 results in a sequence that folds uniquely to the target under the sHP potential, but the same sequence has $g = 144$ different lowest-energy conformations under the HP potential.

Hence GCSE designs unique folding sequences for short 2D chains. Does GCSE also design unique or near-unique sequences (*i.e.*, those having minimal degeneracy) for longer chains in 3D? It was successful in the one case we tested here, a 3D cubic lattice HP model sequence of 88 residues. The target native conformation is shown in Figure 3a of [Yue and Dill 1995]. There are $2^{88} \approx 3 \times 10^{26}$ different possible sequences for this chain length. Although the total number of sequences computed by GCSE is only a small fraction of this, around $100 \times 500 \times 7 = 3.5 \times 10^5$, the GCSE algorithm converges. Obviously, extensive sequence space enumeration is impossible in this case, so checking the GCSE results against the known unique folders is less exact. However, one sequence designed by the present algorithm is the one that was designed in [Yue and Dill (1995)] by the constrained hydrophobic core construction method with the HP potential, which has been shown to have minimal degeneracy. Thus GCSE includes negative design in this case as well as in the 2D cases above.

The lattice tests indicate that the GCSE energy function is effective in designing sequences that fold to the lattice target structures. Table 4 shows that the strategy of obtaining sequences with lowest sHP energies that fold to a target structure is likely to yield sequences that would fold uniquely to the same structure in the HP potential. This observation is suggestive of a deeper underlying physical connection between the two potentials. However, while successful in providing sequences that fold to a given structure, minimizing sHP energy may not identify *all* HP sequences that fold to the target. For instance, the sequence HPHPHPPHPHPPHHHH folds uniquely on the 2D lattice under both the HP and sHP potentials. When we fold it under the HP potential, we generate its native structure (HP energy = -8). Using this structure in GCSE inverse folding, we get the original sequence back as the designed sequence. As explained above, the GCSE output sequence minmizes sHP energy, yielding in this case an sHP energy of -12. Hence when we minimize sHP energy in inverse folding for this sequence, we obtain a unique folder in HP energy. But other sHP sequences also fold uniquely to this structure, without being the very lowest in sHP energy. For example, the sequence HPPPHPPHPHPPHHHH also folds uniquely to

this target in both potentials. Its sHP and HP energies are, respectively, -11 and -7. Biological systems may be satisfied with higher-sHP-energy sequences which also fold uniquely under HP. The GCSE strategy is to find a sequence that minimizes an energy. Why should this work? While folding is a process of energy minimization, inverse folding is not. We believe that by including the extra surface term, GCSE tends to find sequences with well-separated native and denatured energies, that are thus stable in the target fold.

1.4.7 Hydrophobic core positions are more conserved.

Because GCSE produces, as output, an ensemble of 500 sequences very similar in energy (see above), we can observe the variability at each sequence position across this final "population." This gives some measure of how conserved each residue should be, according to the design strategy. Out of 42 native H residue positions, 38 H residue positions are conserved in the 3D lattice 88-mer in the computed final 500 sequences. The other 4 native H residue positions are only slightly mutable (more than 95% sequence homology at these four positions). Interestingly, P residue positions are also conserved. Out of 46 native P residue positions, 13 of them are conserved. The sHP energy function discourages burying P residues in the core, but is less sensitive to changes on the surface. This is consistent with experiments showing higher mutabilities at surface than core sites [Matthews 1993].

1.5 Conclusions.

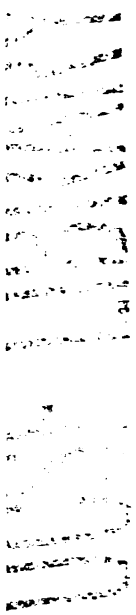
We have described a computer algorithm that takes coordinates of a known native protein structure, either from the PDB or from a lattice model, and designs a sequence of polar and nonpolar groups compatible with that structure. It is a grand canonical method, based on a genetic algorithm bead coloring process that iteratively minimizes a simple energy function. Our algorithm predicts to within a few percent accuracy the hydrophobic and hydrophilic composi-

tions of 20 different proteins we tested, and the sequence identities with the native sequences range from about 60 to 80%.

In lattice model tests, our algorithm designs sequences that fold uniquely to the target structure, *i.e.*, they fold to no other structures of the same or lower energy. Perhaps the GCSE algorithm will be useful in designing sequences that will fold to have desired hydrophobic cores in proteins and other polymers.

1.6 Acknowledgements.

We thank Kaizhi Yue for providing three-dimensional lattice conformations used in this study, and Paul Thomas for helpful discussions. We thank the NIH for financial support.



Bibliography

[1] Bowie JU, Luthy R, Eisenberg D. (1991) *Science,* **253,** 164-70.

[2] Bryngelson JD, Onuchic JN, Socci ND, Wolynes PG. (1995) *Proteins* **21(3)**:167-95.

[3] Chan HS, Dill KA. (1991) *Journal of Chemical Physics,* **95,** 3775-3787.

[4] Chan HS, Dill KA. (1994) *Journal of Chemical Physics,* **100,** 9238-9257.

[5] Chothia C. (1976) Journal of Molecular Biology, **105,** 1-12.

[6] Dandekar T, Argos P. (1992) *Protein Engineering,* **5,** 637-45.

[7] Dill KA, Bromberg S, Yue K, Fiebig K, Yee DP, Thomas PD, Chan HS. (1995) *Protein Science,* **4,** 561-602.

[8] Goldberg, D. (1989) *Genetic Algorithms in Search, Optimization, and Machine Learning,* Addison-Wesley Publishing Company, Inc.

[9] Gutin AM, Abkevich VI, Shakhnovich EI. (1995) *Proc. Nat. Acad. Sci. (USA),* **92,** 1282-1286.

[10] Guy HR. (1985) *Biophysical Journal,* **47,** 61-70.

[11] Kamtekar S, Schiffer JM, Xiong H, Babik JM, Hecht MH. (1993) *Science,* **262,** 1680-5.

[12] Janin J. (1979) *Nature (London),* **277,** 491-492.

[13] Lau KF, Dill KA. (1989) *Macromolecules,* **22,** 3986-3997.

[14] Lee C, Levitt M. (1991) *Nature (London),* **362,** 448-51.

[15] Meirovitch H, Scheraga HA. (1980) *Macromolecules,* **13,** 1406.

[16] Matthews BW. (1993) *Annu. Rev. Biochemistry,* **62,** 139-60.

[17] Osmark P, Sorensen P, Poulsen FM. (1993) *Biochemistry,* **32,** 11007-14.

[18] Pakula AA, Sauer RT. (1990) *Nature (London),* **344,** 363-4.

[19] Ponder JW, Richards FM. (1987) *Journal of Molecular Biology,* **193** 775-91.

[20] Richmond JT. (1984) *Journal of Molecular Biology,* **178,** 63-89.

[21] Šali A, Shakhnovich E, Karplus M. (1994) *Nature,* **369,** 248-251.

[22] Schildbach JF, Milla ME, Jeffrey PD, Raumann BE, and Sauer RT. (1995) *Biochemistry,* **34,** 1405-12.

[23] Shakhnovich EI, Gutin AM. (1993) *Protein Engineering,* **6,** 793-800.

[24] Shakhnovich EI. (1994) *Physics Review Letters,* **72,** 3907-3910.

[25] Socci ND, Onuchic JN. (1994) *Journal of Chemical Physics,* **101,** 1519-1528.

[26] Sun S. (1993) *Protein Science,* **2,** 762-785.

[27] Sun S. (1992) Thesis, Department of Biophysical Sciences, SUNY at Buffalo.

[28] Vasquez M, Scheraga HA. (1988) *J. Biomol. Struct. Dyn.,* **5,** 705-55.

[29] Wertz DH, Scheraga HA. (1978) *Macromolecules,* **11,** 9-15.

[30] Yue K, Dill KA. (1992) *Proc. Natl. Acad. Sci. (USA),* **89,** 4163-4167.

[31] Yue K, Dill KA (1995) *Proc. Natl. Acad. Sci. (USA),* **92,** 146-150.

[32] Yue K, Fiebig KM, Thomas PD, Chan HS, Shakhnovich EI, Dill KA (1995) *Proc. Natl. Acad. Sci. (USA),* **92,** 325-329.

Chapter 2

The effect of multiple binding modes on empirical modeling of ligand docking to proteins.

Rachel Brem and Ken A. Dill

This chapter was originally published as "The effect of multiple binding modes on empirical modeling of ligand docking to proteins," *Protein Science* **8**, 1134-1143 (1999).

2.1 Abstract.

A popular approach to the computational modeling of ligand/receptor interactions is to use an empirical free energy-like model with adjustable parameters. Parameters are learned from one set of complexes, then used to predict another set. To improve these empirical methods requires an independent way to study their inherent errors. We introduce a toy model of ligand/receptor binding as a workbench for testing such errors. We study the errors incurred from the two-state binding assumption—the assumption that a ligand is either bound in one orientation, or unbound. We find that multiple binding modes are not rarer in tight binders than in weak binders. But we find that fitted values of energy parameters, and errors in free energy predictions, do depend on binding strengths. We show that errors in two-state binding can be diagnosed by systematically adding new binding modes when predicting binding affinities: if predictions worsen as the modes are added, then the two-state assumption in the fitting step may be at fault.

2.2 Introduction.

How can we compute the binding affinities of ligands for receptors from knowledge of the atomic structures? One approach is all-atom molecular dynamics, but it is computationally expensive, conformational sampling is limited, and the energy functions are not perfect. A substantial niche has arisen for methods that are faster but more *ad hoc*. Often it is valuable just to have a method that can crudely rank order the binding affinities. A popular strategy for doing this is to assume an empirical model of free energies as a function of chemical characteristics of the ligand and receptor. Some models use adjustable energy parameters. Values of these parameters are "learned" on a training set via regression, and then used to predict a test set of ligand/protein complexes (Bohm, 1994; Jain, 1996; Head et al., 1996; Ortiz et al., 1995; Horton and Lewis, 1992; Perez et al., 1998; Wilson et al., 1991; Wallqvist et al., 1995). We call this class

of models the "Learn Then Predict" (LTP) models. Other models use fixed values for each energy type (Checa et al., 1997; Krystek et al., 1993; Vajda et al., 1994; Verkhivker et al., 1995; Jackson and Sternberg, 1995; Weng et al., 1996).

Our interest here is in how LTP methods can be improved. Their predicted free energies can err by 2-3 kcal/mol or more, and different training sets can lead to different parameters (Ajay and Murcko, 1995). One way to improve empirical methods is to focus on statistical and sampling strategies, irrespective of the underlying physical model. The other approach is to improve the degree to which the model represents the true underlying physics. We take the latter approach. To make systematic improvements in empirical models, it is necessary to have an independent way to study the various types of errors. Here, through use of a simplified model, we study a specific class of errors: those that arise from what we call *the two-state binding assumption.* It is usually assumed that a ligand is either unbound, or bound in a single dominant configuration. But real binding may sometimes involve multiple modes of binding (Weber et al., 1994; Raag and Poulos, 1991; Birdsall et al., 1989; Murthy et al., 1992; Ladbury et al., 1996; Wojtczak et al., 1992; Montfort et al., 1990; Bohm and Klebe, 1996, and references therein).

Recently, multiple modes have been modelled explicitly during high-resolution free energy calculations (Wilson et al., 1991; Head et al., 1997). Also, it has been argued that multiple modes must be accounted for during free energy perturbation (Miller and Dill, submitted) and when computing binding equilibria when competing ligands are present (Janin, 1996). In the present paper, with the help of a simplified but rigorous model, we take three new steps: (i) we quantitate the effects of multiple modes on LTP modeling; (ii) we propose a way to predict when the two-state assumption fails; and (iii) we study how the predictive power of an LTP model changes when the magnitudes of training and test set complexes' free energies change.

2.3 The model of ligand-receptor binding.

We study a toy model of ligand binding. Toy models are the only methods currently available that can test the principles we are interested in here. Despite their artificialities, they have several advantages. We can know with rigor and exactness the underlying physics that drives the binding in the model. We invent the "true" energy function-equivalent in the toy model to the unknown energy function that nature uses to bind real ligands to proteins. So, when we fit parameters, we can always compare our results to the "right answers." In addition, we know the binding partition function exactly, with no errors due to incomplete sampling. Our approach involves no uncertainties about the structural representation, about the dominant binding modes, or about any of the binding thermodynamics, so we can study the way that multiple binding modes might confound modeling efforts. Our simple model serves to generate hypotheses that might then be testable through experiments or refined simulations.

Ours is a 2-dimensional lattice model of receptors and ligands. In this model, a pairwise contact energy function is used that can evaluate the energy E of every possible ligand/receptor configuration. In some cases, our ligands have hundreds of different configurations in which they bind, to different degrees. We call these configurations *binding modes.* Figure 2.1 shows an example lattice ligand/receptor pair.

In the lattice model, receptor sites and ligands are rigid. Ligand/receptor interactions involve only nearest-neighbor pairwise contact energies, $\varepsilon_A, \varepsilon_B, \ldots$, which apply when a ligand molecule sits in a lattice site that is adjacent to a unit surface area of a receptor site. There are no interactions of longer range. A binding mode has an energy $E = n_A\varepsilon_A + n_B\varepsilon_B + \cdots$, where n_α is the number of contacts of α-type surfaces. These energy terms are meant to mimic the short-ranged energies in all-atom functions. To keep the model simple, all cross-terms $\varepsilon_{A-B}, \varepsilon_{B-C}, \ldots = 0$. All the lattice model data reported in this paper were generated with three energy types: ε_A = 0.53 kT, ε_B = 0.88 kT, ε_C = 0.65 kT.

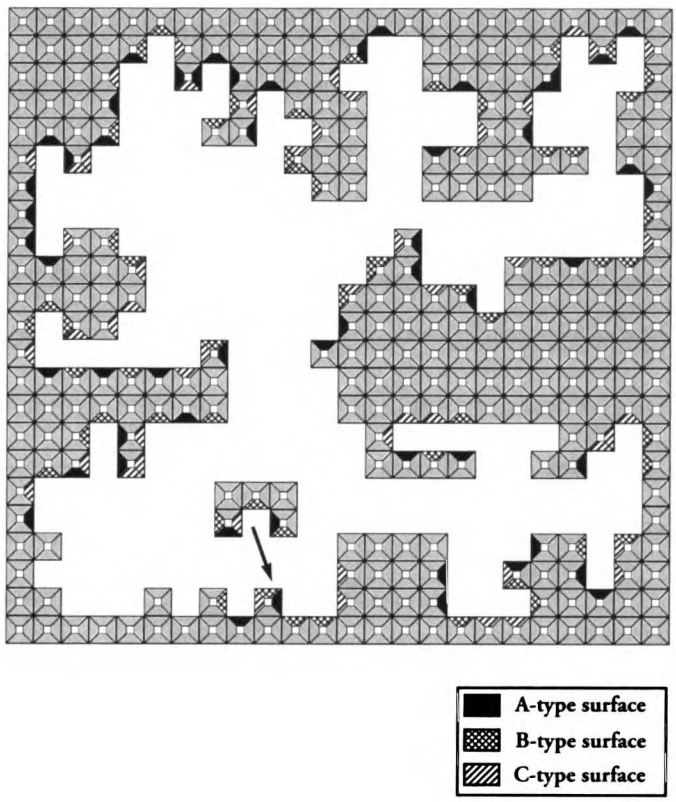


Figure 2.1: **Lattice protein model and a sample lattice ligand model.** In the lattice model, a ligand (small object) binds to its lowest-energy binding site (arrow) within the binding cavity of its receptor (large object). In this paper, the receptor shown is paired with many ligands of varying shapes and surface "chemistries." Only one ligand is illustrated here.

We study 106 ligands which vary widely in chemistry and size, but only one model receptor, shown in Figure 2.1. The need for a large, complicated receptor reflects the need to have a large variety of binding sites for different ligands, and does not otherwise indicate any attempt to mimic any known biomolecule receptor. Our aim is to keep the model as simple as possible. Nevertheless, in order to study a range of complexes, sometimes over a wide range of binding affinities, our ligands must range from small and simple to large and multi-contacting with the receptor. Rather than developing both types of ligand, we simplify the calculations without loss of rigor by multiplying contact counts $n_A, n_B, \ldots$, and/or configuration counts Ω_E, by constants. These constants are chosen arbitrarily to give rise to desired binding affinities.

To calculate the exact free energy of binding of the ligand to the receptor, we enumerate all the ways the ligand and receptor can associate in a box of volume V. The partition function Q is then the sum of the Boltzmann factors over all arrangements $i = 0, 1, 2, \ldots$: $Q \equiv \sum_i e^{-E_i/kT}$ for energy E, temperature T, and Boltzmann's constant k. From our exact partition function, we obtain an exact Helmholtz free energy F through the relation $F = -kT \ln Q$. Because we are interested in binding processes at some standard state concentration, and not interested in concentration dependences, it is simplest to leave out the translational entropy and set $F_{free} = 0$. The binding free energy is $\Delta F = F - F_{free}$. For each complex, given the energy of its ground state E_{gs} and its partition function Q, we can also calculate the Boltzmann population of the dominant (ground-state) mode, $p(gs) \equiv e^{-E_{gs}/kT}/Q$.

To mimic the LTP modeling process, we need to adjust energy parameters on a training set of exact free energies. First, the exact binding free energies of training set ligands are calculated from the given physical model we invent. Those free energies play the role that experimental free energies of binding would play in a real LTP method. Next we need another energy function, which serves as an empirical predictor model. Even though we know the true energy function and partition function by construction, our aim here is to make predictor models that we design to be flawed in various ways. We train the empirical model using linear or nonlinear regression in Mathematica (Wolfram Research, Inc., 1993). This regression assigns energy parameters ε, which make the predicted empirical free energy match the exact free energies of the training set ligands (the "data") as closely as possible. The output of this *fitting step* of the LTP process is a set of learned ε's. Then in the *prediction step*, we substitute the learned ε's back into the empirical function, and predict free energies of more ligands, which make up the test set.

To quantitate the errors in LTP modeling, we report how wrong the fitted values of parameters are, relative to the exact right answers. When predicting binding affinities, we can report errors in the predictions, relative to true exact affinities. We also are interested in how

well the flawed LTP models can distinguish strong from weak binders. As a simple measure of these errors, we report the differences in affinity between pairs of ligand/receptor complexes whose rank order, when predicted by the LTP model, is of the wrong sign. To calculate this quantity we first parameterize an energy function on a training set. We then use the fitted energy function to predict affinities of test set complexes. For any pair of complexes 1 and 2, if $\Delta F_1 < \Delta F_2$, but the model predicts $\Delta F_1 > \Delta F_2$, we call this a "mis-ranking." We make a list of all the mis-ranked pairs in a test set, and report the largest difference in affinities, $\Delta\Delta F$, that we observe among these pairs.

2.4 Results and Discussion.

For both training set and test set complexes, it is common to assume that the ligand is either bound in one particular mode, or unbound. We call this the two-state assumption. In reality, a ligand may have other important binding modes, or may bind nonspecifically, in addition to its dominant mode. It is common to assume that such higher-energy modes are negligible in real binding to biomolecules because NMR or crystallography often finds each ligand at a single site, in a single conformation, hence in a single dominant binding mode. Since binding free energies are usually several kilocalories per mole, it is usually assumed that when binding is tight it is also specific, and that a single mode is sufficient for modeling. While this may often be true, we believe that structural evidence alone is not convincing. Thermodynamically important binding modes need not always be apparent in structures. If binding modes are weak but numerous, they can contribute substantially to the thermodynamics of binding, and yet be invisible in structural studies. It is not clear how important these kinds of errors might be (Janin, 1996; Miller and Dill, submitted). A common approach is to assume the errors arising from the two-state binding assumption are small, and that the dominant errors are in the energy functions. The philosophy underlying this paper is different. Our view is that since

LTP models are not yet sufficiently predictive, it is important to understand the physical basis for any errors, large or small, because then we can hope that strategies will emerge to repair them and allow us to focus attention on other sources of error. Below, we use our lattice model of ligand/receptor binding to study the errors that multiple binding modes may cause in LTP modeling. We suggest diagnostics and remedies for these errors.

2.4.1 Two-state errors during energy function parameterization.

If a true binding process involves one dominant mode and other weaker ones, what errors are incurred if fitting procedures are based on the two-state assumption? In Figure 2.2 we train energy parameters using the two-state assumption, then compare them to the exactly correct values. The magnitudes of the errors are dependent on the dominant mode population. Not surprisingly, when the weak modes have very small populations, errors are small. Interestingly, we find that fitted energies are always more negative than the correct answers. The reason for this is that when we make the two-state assumption, partition functions are truncated, from several contributing Boltzmann terms to just one. The regression compensates by fitting all (favorable) energy parameters to overly negative values.

Figure 2.2 shows that fitting errors depend not only on the Boltzmann population of the ground state, $p(gs)$, but also on the tightness of binding. Why? It turns out that irrespective of $p(gs)$, *the magnitude of the ground-state energy affects fitting error*, as illustrated by Figure 2.3. To understand this, first recall how the fitting step of LTP modeling works. We model the true "experimental" binding free energy ΔF_{exact} with an approximate partition function. In Figures 2.2 and 2.3, we have used the two-state assumption, so our approximation is just the Boltzmann factor of the ground-state binding mode. This means we assume that ΔF_{exact} can be modelled by the expression $n_A\varepsilon_A + n_B\varepsilon_B + n_C\varepsilon_C$. The ε's are the adjustable parameters that will be assigned to make $\Delta F_{exact} \approx n_A\varepsilon_A + n_B\varepsilon_B + n_C\varepsilon_C$. The coefficients n are given; they are the numbers of contacts made in the ground state (mimicking the structural

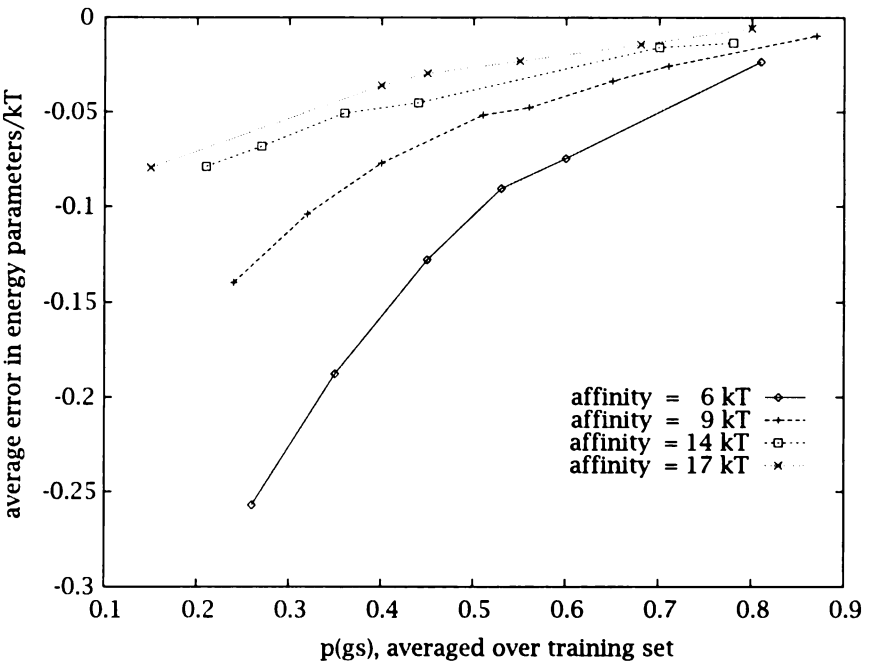


Figure 2.2: **Ignoring multiple binding modes during fitting causes errors in energy parameters.** Every point represents one regression on one training set. Every line style represents a different average affinity among training set complexes. The two-state assumption was used: partition functions in all regression models contained only a ground-state Boltzmann term. The x axis gives $p(gs)$, a complex's occupancy of its ground-state mode, averaged over all the complexes in a training set. The y axis plots the error in each fitted energy parameter output by a regression on the training set, averaged over the 3 parameters in the lattice energy function. Following a curve from left (low-$p(gs)$) to right (high-$p(gs)$) shows that errors in parameters decrease with increasing $p(gs)$ of training set complexes. Comparing among curves from top (high-affinity) to bottom (low-affinity) shows that training sets of low-affinity binders produce parameters with larger errors.

information known from crystallography or NMR). Now consider this fitting process using either of two ligand/protein complexes 1 and 2, where $\Delta F_{exact,1} < \Delta F_{exact,2}$ but $p(gs)$ is the same for both complexes. Under these circumstances, complex 1 will have a more negative ground-state energy than complex 2. Since ε's are identical for the two complexes, this means that there are higher contact counts n_A, n_B, and/or n_C for complex 1. This difference in contact counts explains Figure 2.3. When n values are large, the fitting process compensates by assigning ε's to be small in magnitude. That is, when trying to assign ε's that make $\Delta F_{exact} \approx n_A\varepsilon_A + n_B\varepsilon_B + n_C\varepsilon_C$, there is no reason to assign strong ε's, because large n's already make the right-hand

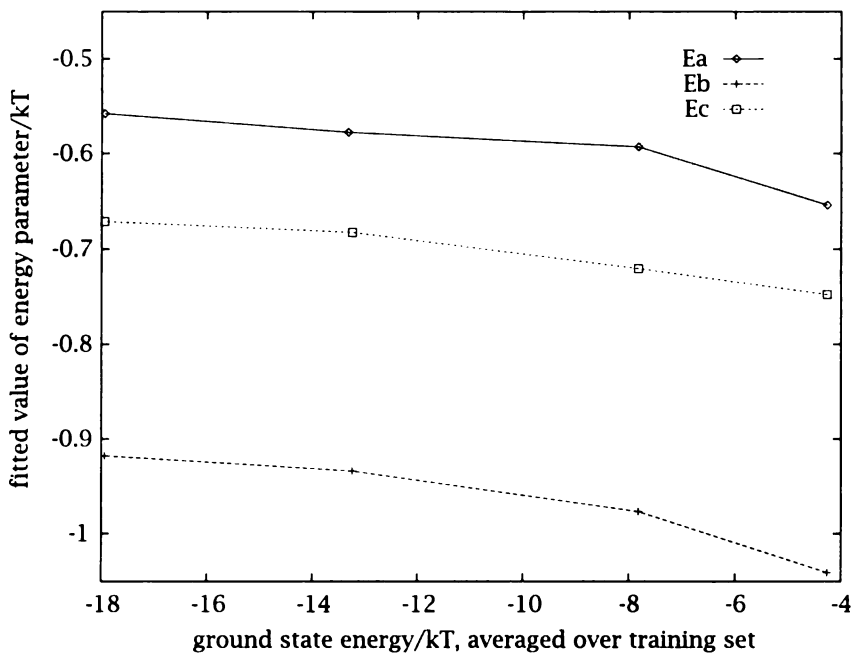


Figure 2.3: **Regressions on stronger binders produce less negative parameter values.** Every point represents one regression. Every line style represents one energy parameter. The two-state assumption was used: partition functions in all regression models contained only a ground-state Boltzmann term. All training sets contained complexes of $p(gs) \approx 0.45$. The x axis gives the energy of the ground-state binding mode, averaged over all the complexes in a training set. The y axis plots the value of each energy parameter output by a regression on the training set. Correct values are ε_A = 0.53 kT, ε_B = 0.88 kT, ε_C = 0.65 kT. Training sets of complexes which bind with stronger ground-state energies produce parameters that are less negative.

expression large enough. By the same token, n's are small among weak binders, so fitted parameters must be larger in magnitude to make the right-hand expression large enough. If we did not use the two-state assumption, this argument would still hold. Stronger binders would still make more contacts overall, which would drive fitted parameter magnitudes down. It follows that the affinity dependence of fitted parameters is not an effect of the two-state assumption. Rather, it comes from a bias in the composition of strong and weak binders in the training set.

The curves in Figure 2.2 are the result of these two opposing trends. Two-state errors make parameters more negative, but strong affinities make parameters less negative.

Adding more binding modes into a model can improve errors arising from the two-state assumption. Figure 2.4 shows the result of learning energy parameters using a partition

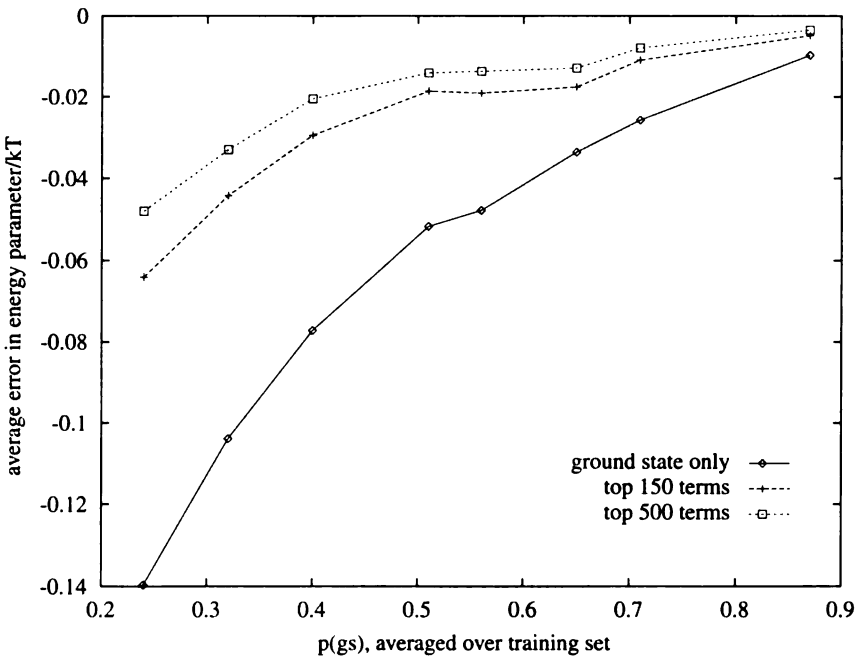


Figure 2.4: **Including multiple modes during fitting improves parameters.** Every point represents one regression. Every line style represents a different number of lowest-energy Boltzmann terms used in the partition function of the regression model. The x axis gives $p(gs)$, a complex's occupancy of its ground-state mode, averaged over all the complexes in a training set. The y axis plots the error in each fitted energy parameter output by a regression on the training set, averaged over the 3 parameters in the lattice energy function. Correct values are ε_A = 0.53 kT, ε_B = 0.88 kT, ε_C = 0.65 kT. Complexes had affinities of around 9kT. Comparing among curves from top (many terms included) to bottom (few terms included) shows that including Boltzmann terms of higher-energy binding modes in partition functions of regression models improves the accuracy of fitted parameters.

function having many Boltzmann terms. Errors are smaller than when the two-state approximation is made. Improving the accuracy of the partition function model improves the accuracy of energy parameters. *How many* higher-energy modes are needed to make this improvement? Below, we develop a diagnostic that indicates when enough modes have been added.

There are two implications for more realistic LTP modeling. First, training sets should contain as many strong binders as weak binders, in order to minimize the bias in fitted energy parameters. This bias may be especially strong when the training set involves very similar complexes, *e.g.* a congeneric series of ligands that bind to the same receptor. Second, the most physically accurate energy parameters can be obtained by including higher-energy

binding modes in partition functions. All-atom methods are being developed to make ensemble calculations of this type quicker and easier (Wilson et al., 1991; Knegtel et al., 1997; Gilson et al., 1997; Head et al., 1997).

2.4.2 Two-state errors during free energy prediction.

So far, we have examined errors from the two-state assumption in the fitting step of LTP modeling. In this section, we examine errors in the prediction step.

Using the two-state assumption in test set partition functions.

Before we put errors into the fitting step, we do the reference experiment where the prediction model is exactly correct. Now when we keep only the ground-state term in the partition function of test set ligands, this two-state assumption in the prediction step is the only source of error. In this case, Figure 2.5 shows that main determinant of prediction error is the degree to which the main binding mode dominates the binding. The more dominant is the principal binding mode, the smaller the errors.

In Figure 2.5 the error in free energy prediction is exactly equal to $-kT\ln[p(gs)]$, no matter what the strength of binding. The reason is as follows. Call a the Boltzmann factor for the ground-state binding mode. We can write the partition function as $Q = a/p(gs)$. (For example, if $p(gs) = 0.5$, $Q = 2a$.) If we now modelled the partition function with only the ground-state binding mode, we would be approximating the partition function as $Q' = a$. Then the error in the free energy we predict is $\Delta F_{predict} - \Delta F_{exact} = -kT\ln Q' + kT\ln Q = kT\ln(1/p(gs)) = -kT\ln[p(gs)]$. Figure 2.5 shows that this error is fairly small whenever $p(gs) > 0.1$.

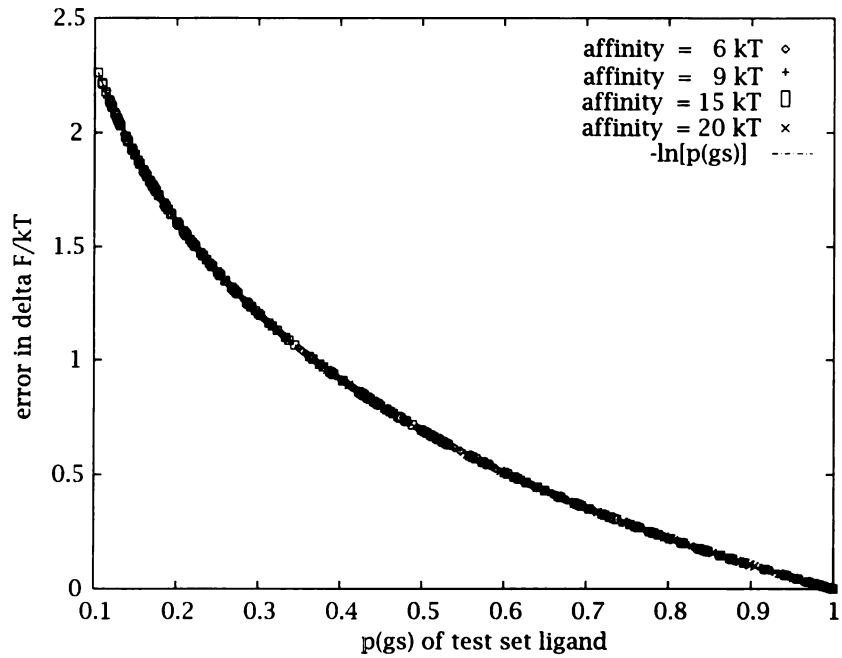


Figure 2.5: **Ignoring multiple binding modes in test sets causes errors in free energies.** Every point represents one test set complex. Every point style represents a different average affinity among test set complexes. The two-state assumption was used: on each complex, we predicted ΔF using a partition function composed of only the ground-state term, and using exact values for energy parameters. The x axis gives $p(gs)$ of a test set complex. The y axis shows the error in the predicted ΔF for the complex, $(\Delta F_{\text{predicted}} - \Delta F_{\text{exact}})/kT$. The dashed line shows the curve of $-\ln[p(gs)]$.

Using the two-state assumption in training and test sets.

Now we return to the full LTP paradigm. Here our prediction energy function uses fitted values of energy parameters, which are imperfect. We showed in the previous section that using the two-state assumption when fitting energy functions produces energy parameters that are in error. How poorly do these parameters predict free energies of test set complexes? Figure 2.6 shows that flawed parameterized energy functions can make sizable mistakes in predicting free energies of complexes.

Note in Figure 2.6 that the LTP paradigm makes larger errors on stronger binders. This is because, at a given value of $p(gs)$, stronger binders make more contacts than weaker ones. For stronger binders, there are more ligand/receptor interactions, so flawed energy terms are used more often, and their errors add up.

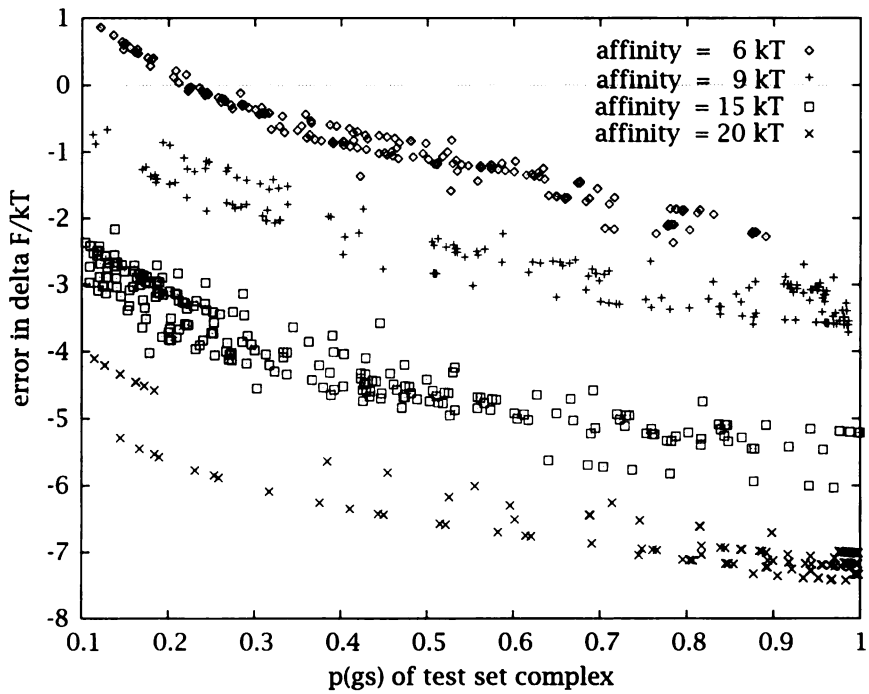


Figure 2.6: **Energy parameters fitted under the two-state assumption predict free energies inaccurately.** Every point represents one test set complex. Every point style represents a different average affinity among test set complexes. The two-state assumption was used during fitting and prediction: partition functions in all regression models and all test set free energy expressions contained only a ground-state Boltzmann term. Energy parameters from regressions on complexes with $p(gs) \approx 0.2$ and affinities around 6 kT were used. The x axis plots $p(gs)$, a test set complex's occupancy of its ground-state mode. The y axis plots the error in the predicted free energy, $(\Delta F_{predicted} - \Delta F_{exact})/kT$. Following a curve from left (low-$p(gs)$) to right (high-$p(gs)$) shows that errors in predicted free energies increase with increasing $p(gs)$ of test set complexes, because energy parameters were optimized for low-$p(gs)$ binders. Comparing among curves from top (low-affinity) to bottom (high-affinity) shows that low-affinity test set complexes are predicted with smaller errors.

As parameters get worse, this accumulation of errors gets worse. The difference between errors in weak binders and strong binders becomes more pronounced. To see this, compare Figure 2.5 with Figure 2.6. In Figure 2.5, exactly correct values of energy parameters are used to predict affinities. In this case, complexes that bind with ΔF = 6 kT have the same errors as those that bind with ΔF = 20 kT. This is because there are no errors that can sum up differently in tighter *versus* weaker binders. The curves all superimpose, for all binding affinities. But Figure 2.6 shows that when the model is flawed, errors are more severe among stronger binders. If we use even worse parameter sets, the differences between errors in weak binders

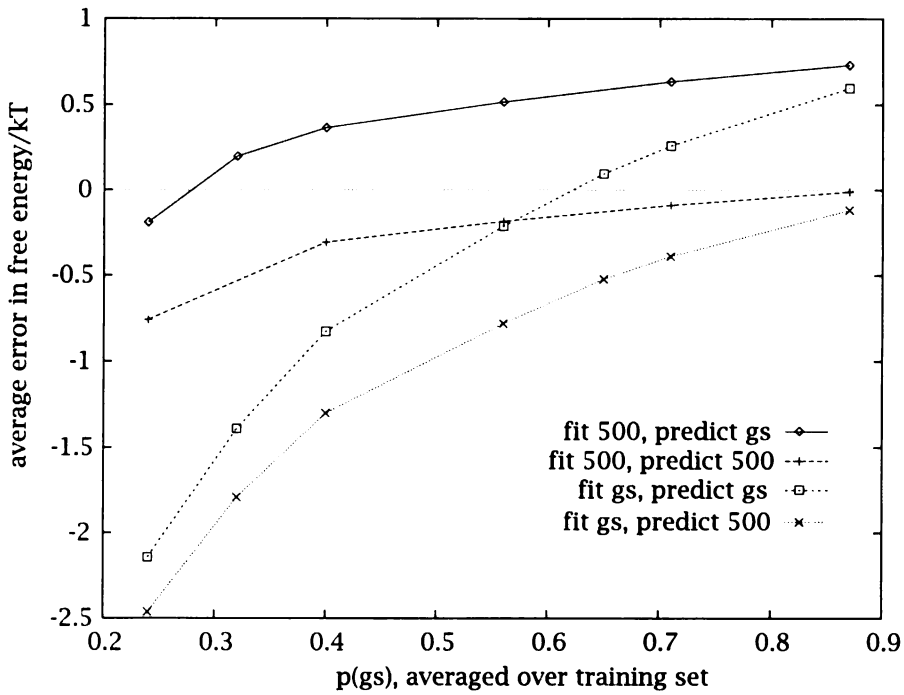


Figure 2.7: **Accounting for multiple binding modes improves free energy prediction.** Every point represents one regression. Every line style represents a different number of lowest-energy binding modes used in training and test set partition functions. In the legend, "fit" indicates the way the partition function was written during the fitting step; "predict" indicates the partition function during the prediction step; "gs" indicates that only the ground-state Boltzmann term was used; "500" indicates that the top 500 lowest-energy Boltzmann terms were used. The x axis plots $p(gs)$, a complex's occupancy of its ground-state mode, averaged over all the complexes in a training set. The y axis shows the error in the predicted ΔF for a complex, $(\Delta F_{predicted} - \Delta F_{exact})/kT$, averaged over all 106 lattice ligands. Following a curve from left (low-$p(gs)$) to right (high-$p(gs)$) shows that errors in predicted free energies decrease with increasing $p(gs)$ of training set complexes. Comparing among curves shows that using many Boltzmann terms during the fitting and prediction steps (crosses) gives better results than using the two-state assumption in both steps (squares) or in just one (diamonds and x's). Using the two-state assumption during fitting, but multiple Boltzmann terms during prediction (x's), gives the worst results. Energy parameters from regressions on complexes with $p(gs) \approx 0.2$ and affinities around 9 kT were used. Predictions were done on complexes with affinities around 14 kT.

and strong binders get even larger (data not shown). This spreading out of errors is important because it leads to wrong rank-orderings of binding affinities (see below).

Figure 2.6 also shows that errors in predicted affinities depend strongly on $p(gs)$ of the test set ligand. This happens because we are using the two-state assumption twice: once during energy function fitting on the training set, and once during free energy prediction on the

test set. When we make the two-state assumption during *the fitting step,* energies ε tend to be more negative than they should be (see Figure 2.2). So predicting with them makes Boltzmann factors too big, and predicted binding free energies too negative. However, when we make the two-state assumption during *the prediction step,* we truncate the partition functions of test set ligands, making the absolute value of the free energy smaller. This tends to err in the opposite direction: predicted free energies are too positive.

Thus, the trends from the two LTP steps oppose each other and sometimes cancel, which is why some errors are small in Figure 2.6. When test set ligands have very low $p(gs)$, errors from the test set partition function dominate, and free energies tend to be too positive. When test set ligands have high $p(gs)$, the test set partition function is fairly accurate. Then errors from energy parameters dominate, and free energies tend to be too negative.

Using multiple modes in training and test set partition functions.

As noted above, we can improve the accuracy of parameters by including multiple binding modes in the training step. If we also include multiple binding modes in the prediction step, free energies are predicted quite accurately by LTP energy functions, as seen in Figure 2.7 (second curve). But Figure 2.7 shows an interesting puzzle: *improving* the modeling in the fitting step (by upgrading from the two-state to multiple binding assumption) can *degrade* affinity predictions. We illustrate this more fully in Figure 2.8.

This non-physical behavior arises from the tendency, described above, for the two-state assumption to produce parameters that are too negative. These parameters then produce Boltzmann factors and partition functions of the test set that are too big. This causes the consistent trend that all free energies of test set complexes are less accurate, because they are too negative, when predicted with multiple modes. As illustrated in Figure 2.8(d), we find that almost any set of energy parameters which is bad enough can make less accurate predictions as more binding modes are used in the test set. But in the general case, it does not happen consistently,

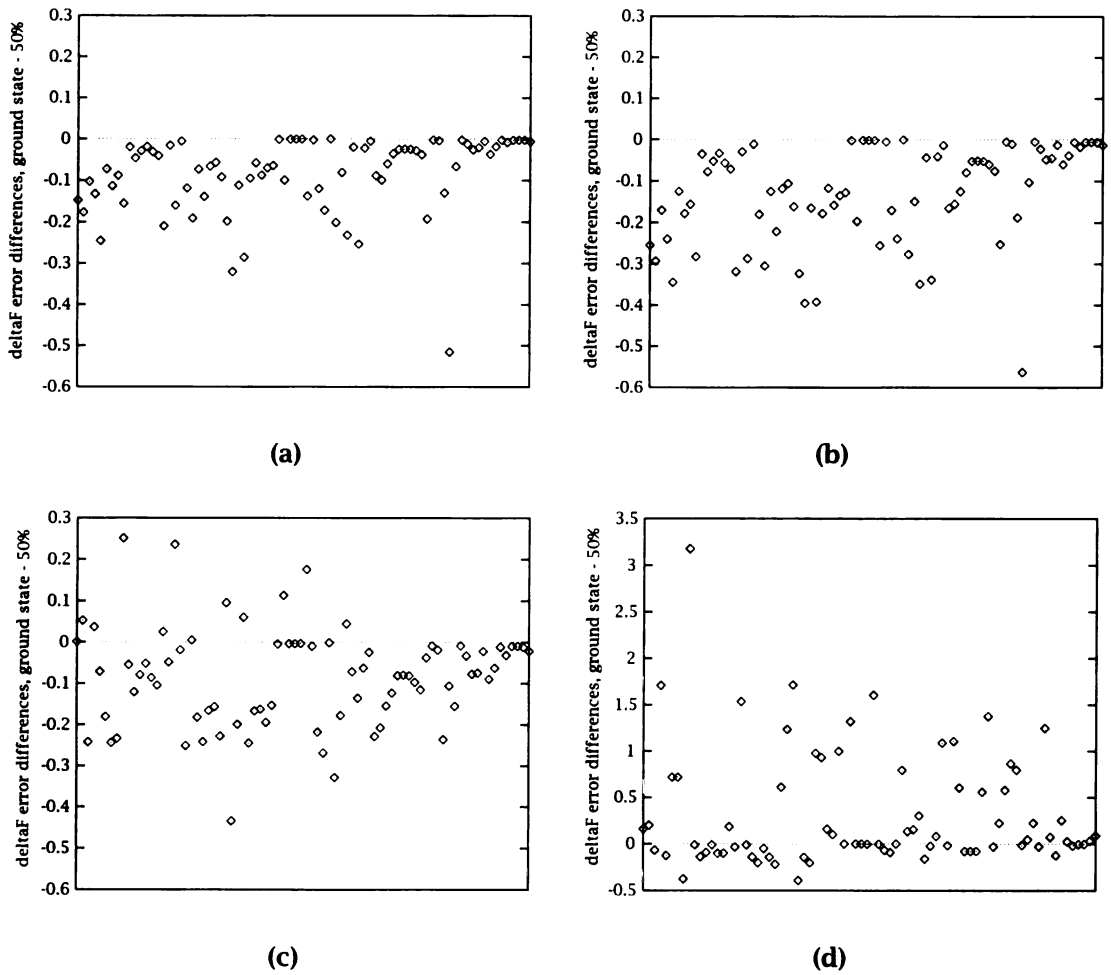


Figure 2.8: **Ignoring higher-energy modes during regression can cause consistently nonphysical free energy predictions.** Each panel shows data produced with energy parameters from a single regression. On a given panel, each point corresponds to the difference in the errors e from two free energy predictions on a single test set complex. One prediction used only the Boltzmann term of the ground-state binding mode (gs) to model the test set complex's partition function; the other used the Boltzmann terms of the top 50% of all lowest-energy binding modes. Each data point shows the quantity $|e(gs)| - |e(50\%)|$ for a single complex. If this quantity is negative, using *more* binding modes during free energy prediction gives *worse* errors. (a), Energy parameters came from a regression on training set complexes with an average $p(gs)$ around 0.25; (b), $p(gs) = 0.4$; (c), $p(gs) = 0.65$; (d), $p(gs) > 0.7$, but the energy function was fit with the C-type contact term omitted. The partition function in the regression model contained only the Boltzmann term of the ground state. Energy parameters from regressions on complexes of affinities around 9 kT were used. Predictions were done on complexes with affinities around 9 kT. Free energy predictions are shown for all 106 lattice ligands used in this study.

across a group of many free energy predictions. The artifact only appears consistenly when all energy parameters are more negative than they should be.

The results in Figure 2.8 suggest a way to weed out poor models or training sets. If predictions worsen upon increasing the number of binding modes assumed in the test set, it implies that the fitted values of parameters are flawed. And if predictions of many test set complexes consistently show this behavior, it suggests that all energy parameters are overly negative, possibly due to the omission of higher-energy binding modes in training set models. Thus, this behavior can be used as a reporter for two-state error in fitted parameters: multiple modes should be added to the training set partition functions until the artifact no longer occurs.

In conclusion, errors in free energies predicted by LTP models depend on $p(gs)$ and binding affinities for both test set and training set ligands. Our model doesn't allow us to make quantitative estimates of these errors in all-atom systems. But our modeling suggests principles that may be of general value. (i) When the two-state assumption is made during the free energy prediction step, if a "perfect" or very accurate energy function is used, two-state error in predicted free energies is equal to $-kT\ln[p(gs)]$, which is probably small for most drug design cases of interest. (ii) Poor energy parameters cause less error in the predictions of binding free energies of weak binders than of tight binders. (iii) In some cases, two-state error during the free energy prediction step cancels two-state error from the fitting step. (iv) The most physical and accurate predictions are those done with accurate partition functions, that account for multiple binding modes, in both steps. The number of binding modes assumed in modeling should be the same in learning and prediction steps. But herein, we believe, lies a useful diagnostic. If adding multiple binding modes worsens the predictions, it may indicate two-state errors in the learned parameters.

2.4.3 Two-state errors during the rank ordering of complexes.

In LTP modeling, often the key prediction of interest is just a correct rank-ordering of binding affinities, not necessarily the affinities themselves. In this section, we pay attention to the pairs of binding complexes that are ranked incorrectly—the mis-ranks. If the ΔF's of complexes in a mis-ranked pair are very far apart, it implies that the LTP model confuses weak binders with strong ones. Then if the model is used to screen a database to output a short list of putative strong binders, some weak binders are likely to be on the list. Such errors are a useful way to characterize a database screening tool. A good screening tool will not mis-rank a weak and a strong binder.

Can the two-state assumption cause LTP models to mis-rank very disparate complexes? In Figure 2.9, we use the two-state assumption in the fitting and/or the prediction step of the LTP process. We report the largest $\Delta\Delta F$ of all the mis-ranked pairs in a large test set. Figure 2.9 shows that the mis-ranking errors are small, even with the most severely flawed energy parameters (those from a training set of low-$p(gs)$ ligands of weak affinity; see Figure 2.2). By contrast, Figure 2.6 shows large errors in ΔF's.

Thus, LTP models using the two-state assumption make poor predictions of affinities, but reasonable predictions of rank orders of affinities. Why? We find that the two-state assumption does not cause the error in predicted affinity to change much as the magnitude of the true affinity changes. The latter behavior is what causes large ranking errors. To understand this, imagine two complexes 1 and 2, where $\Delta F_1 < \Delta F_2$. If a computational method is used to predict the two free energies, the errors in the predictions will be e_1 and e_2, respectively. To predict the wrong rank order of the two complexes, it must be true that $|e_1 - e_2| \geq |\Delta F_1 - \Delta F_2|$. We illustrate this *ranking error criterion* in Figure 2.10. In Figure 2.10(b) the error difference is at the critical value, $|e_1 - e_2| = |\Delta F_1 - \Delta F_2|$, and the two affinities are predicted to be equal. Any greater error difference, and the predicted rank will be wrong, as in Figure 2.10(c); any lesser error difference, and the predicted rank will be accurate, as in Figure 2.10(a). When we use the

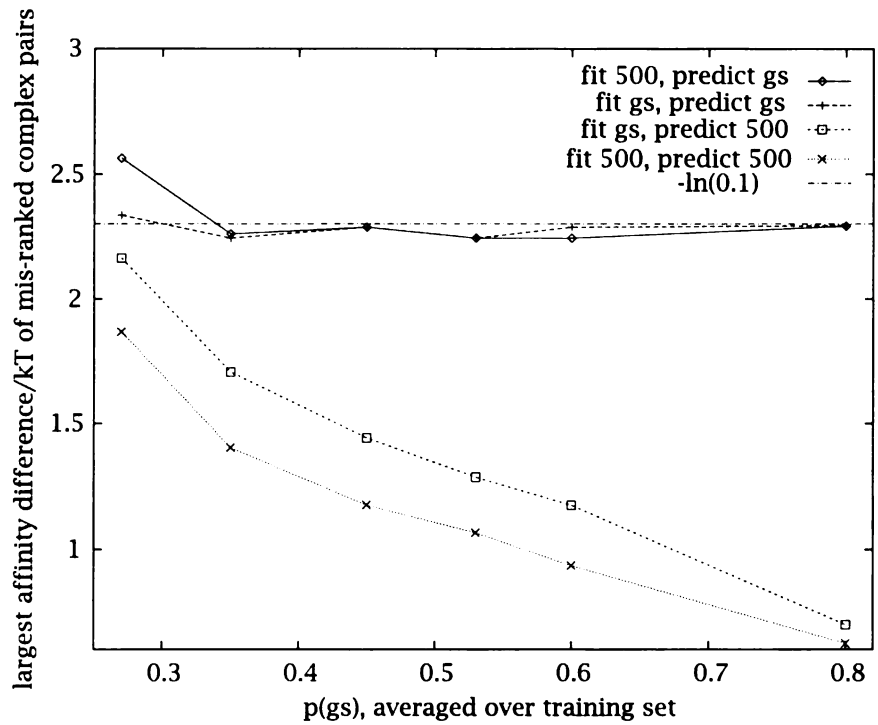


Figure 2.9: **Accounting for multiple binding modes improves rank orders of test set complexes.** Every point represents one regression. Every line style represents a different number of binding modes used in training and test set partition functions. In the legend, "fit" indicates the way the partition function was written during the fitting step; "predict" indicates the partition function during the prediction step; "gs" indicates that only the ground-state Boltzmann term was used; "500" indicates that the top 500 lowest-energy Boltzmann terms were used. The x axis plots $p(gs)$, a complex's occupancy of its ground-state mode, averaged over all the complexes in a training set. The y axis plots the largest affinity difference between complex pairs whose affinities were mis-ranked, from all pairs of complexes in a large test set with affinities from 6–100 kT. Following a curve from left (low-$p(gs)$) to right (high-$p(gs)$) shows that errors in ranking decrease with increasing $p(gs)$ of training set complexes. Comparing among curves shows that using many Boltzmann terms during the fitting and prediction steps (x's) gives better results than using the two-state assumption in both steps (diamonds) or in just one (crosses and squares). The prediction step is the most important in reducing ranking errors. Energy parameters from regressions on complexes with $p(gs) \approx 0.2$ and affinities around 6 kT were used.

two-state assumption in an LTP model, and predict free energies on test set ligands, predictions do not satisfy the ranking error criterion for very large $|\Delta F_1 - \Delta F_2|$. In other words, the error in a predicted affinity does not change much with the magnitude of the affinity. All errors are about the same, and they mostly cancel. As we argued above (see Figure 2.6), the errors in flawed parameters add up when test set complexes make more contacts, and this behavior is

more pronounced when more flawed energy parameters are used. But even our worst energy parameters are not "bad enough" to cause large error differences.

Also interesting in Figure 2.9 is that the prediction step dominates ranking errors. The top two curves show that, if we use the two-state assumption when we write partition functions of test set ligands, it doesn't matter whether we have used multiple modes or not during the fitting step. Parameters from both kinds of fit give almost identical ranking errors, which are around 2.3 kT. This is noteworthy because 2.3 kT $\approx -kT \ln 0.1$ is the maximum ranking error that arises when *perfectly accurate* parameters are plugged in to the energy function. Under the latter conditions, $|e_1 - e_2|$ is largest when complex 1 has $p(gs) \approx 1$ and complex 2 has very low $p(gs)$, here 0.1. (To see this graphically, compare the two extremes of Figure 2.5.) Then $e_1 = kT \ln 1 = 0$ and $e_2 = -kT \ln 0.1 = 2.3kT$. Of course, we haven't used exactly accurate ε's in Figure 2.9, but we get almost the same results as if we did: the largest $|e_1 - e_2| \approx 2.3$ kT, so the largest $|\Delta F_1 - \Delta F_2|$ of a mis-ranked pair is approximately 2.3 kT. Figure 2.9 shows that only the very worst energy parameters (the far left regions of the top curves) cause ranking errors greater than 2.3 kT. In all the other cases, errors from the two-state assumption during fitting are negligible.

We conclude that in the absence of other errors, for predicting rank orders correctly, it is more important to use multiple binding modes in the prediction step than in the fitting step. Nevertheless, ranking errors are not very severe even if the two-state assumption is used in both steps.

2.4.4 Can two-state error be avoided by predicting $p(gs)$?

Tight binding does not guarantee that the two-state assumption is valid, as we noted in Figure 2.5. The validity of the two-state assumption depends on the relative strength of the weak and strong modes. Figure 2.11 confirms that even strong binders with large negative free energies of binding can populate many higher-energy binding modes. There is no correlation

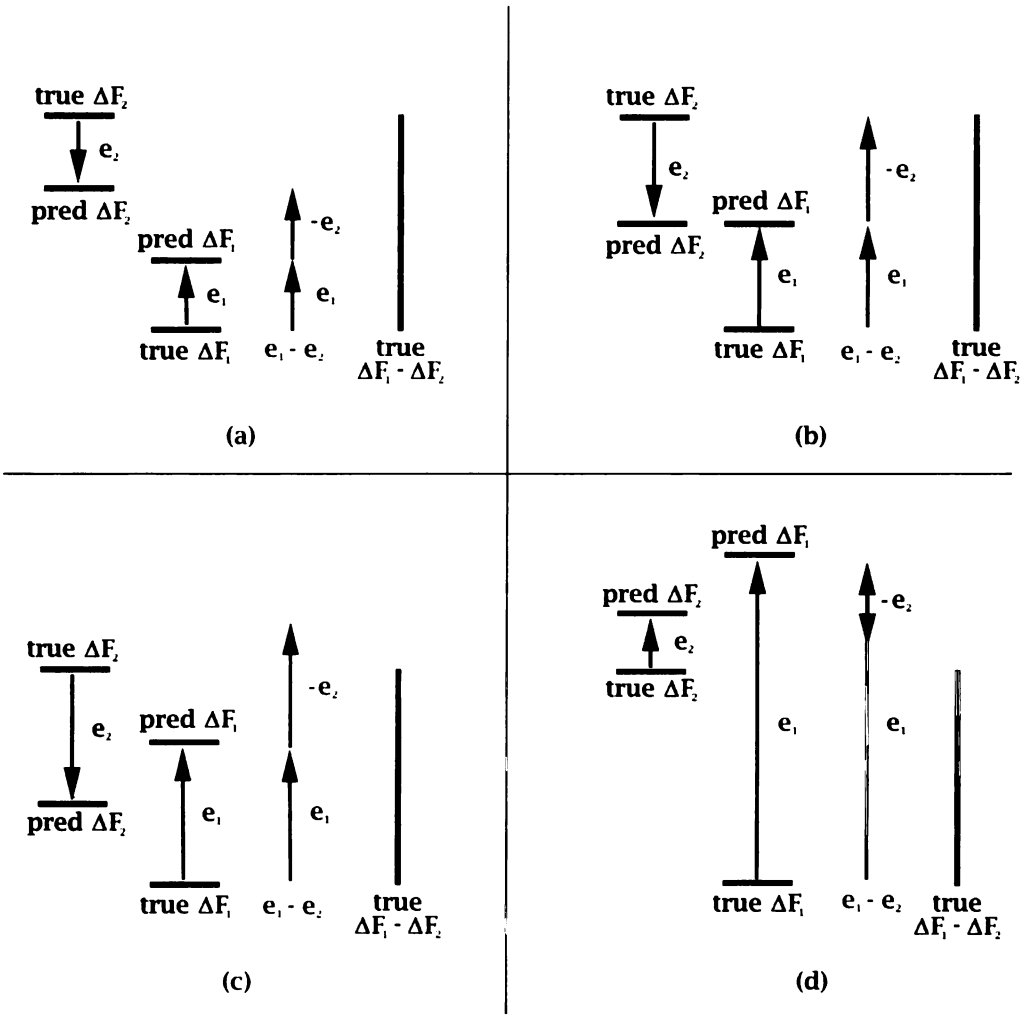


Figure 2.10: **Rank-ordering errors occur only if differences in affinity errors are larger than differences in true affinities.** Horizontal bars represent free energy values of imaginary test set complexes. Distance from the top of each panel represents the magnitude of the free energy. Each pair of stacked horizontal bars represents one test set complex; there are two test set complexes, 1 and 2, in each panel. True free energies are marked "true ΔF," and free energies predicted by an LTP method are marked "pred ΔF." Arrows represent errors in predicted free energies, marked "e." Stacked arrows represent the differences between errors from two test set complexes. An inverted arrow represents the negative of an error value. Vertical bars represent the true $\Delta\Delta F$, the difference between the true free energies of two test set complexes. The rank of two test set complexes is the order with respect to affinity. In each panel, complex 1 is a stronger binder than complex 2, but in some panels, complex 2 is predicted to be stronger by the LTP method. The *ranking error criterion* for a rank to be predicted incorrectly is that the difference in prediction errors (height of stacked arrows) must be greater than or equal to the difference in true free energies (height of vertical bar). (a), The predicted rank has the correct sign; (b), the predicted rank is 0; (c), the predicted rank has the incorrect sign; (d), errors in ΔF are of the same sign, but the predicted rank is still incorrect, showing that ranking errors are independent of the signs of errors in ΔF.

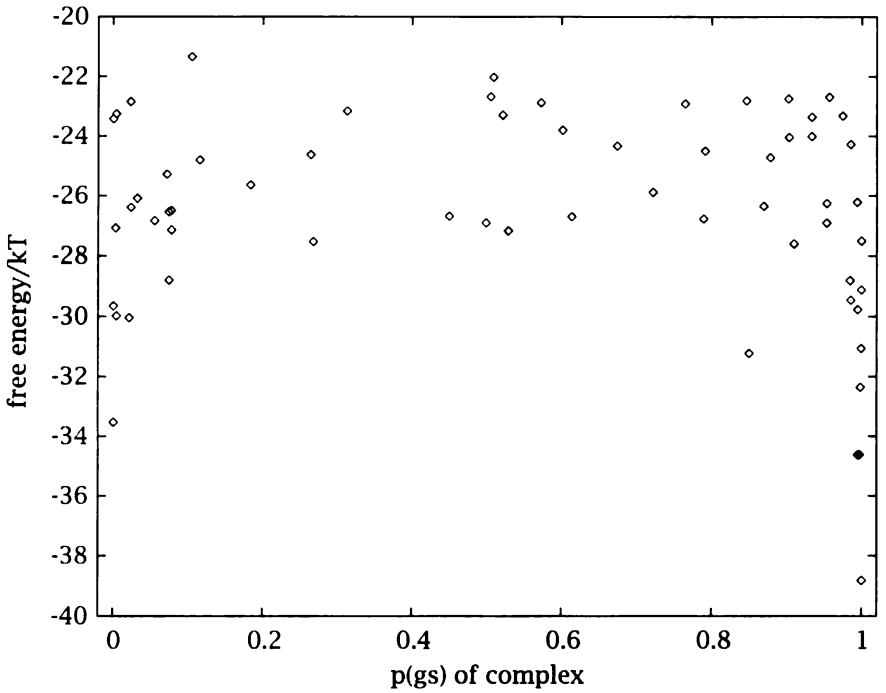


Figure 2.11: **Occupancy of the ground state does not correlate with strength of binding.** Every point represents one complex. Data in this graph are exact values for each complex. The x axis plots the occupancy of the ground-state mode for the complex. The y axis gives the exact free energy of binding of the complex. Data are shown for all 106 lattice ligands used in this study, with affinities around 25 kT.

between the strength of binding of a ligand and the probability that it will occupy its ground-state mode. Some ligands are strong binders not because they bind tightly to the ground-state mode, but because of their favorable entropies of the bound state. Figure 2.12 illustrates this for our lattice complexes. Configurational entropy may play an important role even under conditions when enthalpies dominate the free energy (Weber et al., 1994; Raag and Poulos, 1991; Ladbury et al., 1996). It cannot be known whether or not a complex has high $p(gs)$ just from its free energy of binding.

We do not know of any foolproof predictor of $p(gs)$ in more realistic modeling. Structural experiments are not unambiguous either. Even if a parent ligand/receptor complex is crystallized and well-ordered, it does not follow that a *derivative* will bind specifically too: small changes in ligand chemistry or experimental conditions can produce large changes in binding

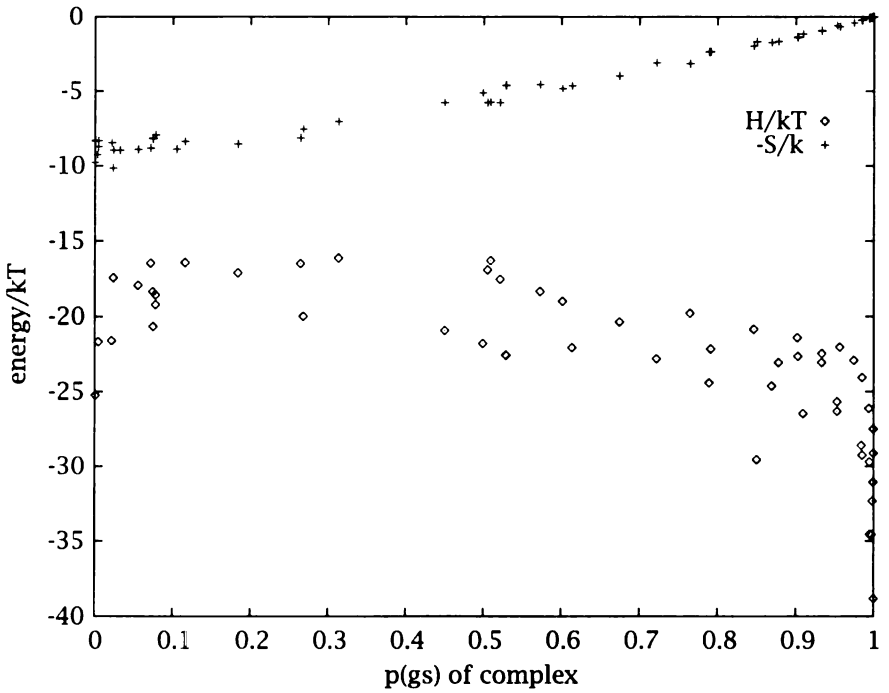


Figure 2.12: **Occupancy of the ground state correlates with enthalpy and entropy of binding.** Every point represents one complex. Data in this graph are exact values for each complex, not predictions of any kind. The x axis shows the occupancy of the ground-state mode for the complex. The y axis gives the exact enthalpy (diamonds) or entropy (crosses) of binding of the complex. Data are shown for all 106 lattice ligands used in this study, with affinities around 25 kT.

behavior (Montfort et al., 1990; Mattos et al., 1994). If a training set could be assembled of only "good" complexes which occupy their ground states most of the time, two-state errors might be avoided in the fitting step. But such a training set will suffer from other problems, since it is useful to choose training sets that are as broad and representative as possible, including both bad and good binders (Bohm, 1994; Jain, 1996). As for the prediction step, restricting LTP calculations to "good" binders is not a good strategy, because unoptimized small molecules are the most interesting targets of LTP analysis. For both training sets and test sets, one way to repair multiple-mode binding errors is to include additional binding modes in the empirical model.

2.5 Conclusions.

We present results from a simple model of ligand/receptor binding. We study the magnitudes and sources of errors in LTP modeling that come from the two-state approximation. We show that two-state errors may arise even among strong binders, and we show that they can be remedied by including additional binding modes. We show that fitted values of energy parameters, and errors in free energy predictions, depend on binding strengths. We observe that if two-state error causes flawed parameters, the parameters consistently predict free energies of binding of test set ligand/protein complexes to be too negative. And, we find that in the absence of other errors, it is more important to use multiple modes during free energy prediction than during parameter fitting, if accurate rank orders of complexes are needed. The principles learned from this simple model may apply to more realistic models of binding.

2.6 Acknowledgements.

We thank Drs. David Agard, Irwin Kuntz, and David Miller for helpful discussions, and Chuck Sindelar for graphic art expertise. We also thank members of the lab of Robert Fletterick for the use of their computers. R.B. was an NSF predoctoral fellow while this work was in progress.

Bibliography

[1] Ajay and Murcko MA. 1995. Computational methods to predict binding free energy in ligand-receptor complexes. *J Med Chem 38(26)*:4953–4967.

[2] Birdsall B, Feeney J, Tendler SJB, Hammond SJ, and Roberts GCK. 1989. Dihydrofolate reductase: multiple conformations and alternative modes of substrate binding. *Biochemistry 28(5)*:2297–2305.

[3] Bohm HJ. 1994. The development of a simple empirical scoring function to estimate the binding constant for a protein-ligand complex of known three-dimensional structure. *J Comp-Aid Mol Des 8*:243–256.

[4] Bohm HJ and Klebe G. 1996. What can we learn from molecular recognition in protein-ligand complexes for the design of new drugs? *Angew Chem Int Ed Engl 35*:2588–2614.

[5] Checa A, Ortiz AR, de Pascual-Teresa B, and Gago F. 1997. Assessment of solvation effects on calculated binding affinity differences: trypsin inhibition by flavonoids as a model system for congeneric series. *J Med Chem 40(25)*:4136–4145.

[6] Gilson MK, Given JA, and Head MS. 1997. A new class of models for computing receptor-ligand binding affinities. *Chemistry & Biology 4(2)*:87–92.

[7] Head MS, Given JA, and Gilson MK. 1997. "Mining minima": direct computation of conformational free energy. *J Phys Chem 101*:1609–1618.

[8] Head RD, Smythe ML, Oprea TI, Waller CL, Green SM, and Marshall GR. 1996. VALIDATE: A new method for the receptor-based prediction of binding affinities of novel liganeds. *J Am Chem Soc 118*:3959-3969.

[9] Horton N and Lewis M. 1992. Calculation of the free energy of association for protein complexes. *Protein Science 1(1)*:169-181.

[10] Jackson RM and Sternberg MJE. 1995. A continuum model for protein-protein interactions: application to the docking problem. *J Mol Biol 250*:258-275.

[11] Jain AN. 1996. Scoring noncovalent protein-ligand interactions: a continuous differentiable function tuned to compute binding affinities. *J Comp-Aid Mol Des 10*:427-440.

[12] Janin J. 1996. Quantifying biological specificity: the statistical mechanics of molecular recognition. *Proteins 25(4)*:438-445.

[13] Knegtel RMA, Kuntz ID, and Oshiro CM. 1997. Molecular docking to ensembles of protein structures. *J Mol Biol 266(2)*:424-440.

[14] Krystek S, Stouch T, and Novotny J. 1993. Affinity and specificity of serine endopeptidase-protein inhibitor interactions. Empirical free energy calculations based on X-ray crystallographic structures. *J Mol Biol 234(3)*:661-679.

[15] Ladbury JE, Hensmann M, Panayotou G, and Campbell ID. 1996. Alternative modes of tyrosyl phosphopeptide binding to a Src family SH2 domain: implications for regulation of tyrosine kinase activity. *Biochemistry 35(34)*:11062-11069.

[16] *Mathematica 2.2*, Wolfram Research, Inc., Champaign, Illinois, 1993.

[17] Mattos C, Rasmussen B, Xiachun D, Petsko GA, and Ringe D. 1994. Analogous inhibitors of elastase do not always bind analogously. *Nat Struct Biol 1(1)*:55-58.

[18] Miller D and Dill KA. 1998. Computer simulations of ligand binding: are non-dominant binding modes important? Submitted.

[19] Montfort WR, Perry KM, Fauman EB, Finer-Moore JS, Maley GF, Hardy L, Maley F, and Stroud RM. 1990. Structure, multiple site binding, and segmental accommodation in thymidylate synthase on binding dUMP and an anti-folate. *Biochemistry 29(30)*: 6964-77.

[20] Murthy KHM, Winborne EL, Minnich MD, Culp JS, and Debouck C. 1992. The crystal structures at 2.2—A resolution of hydroxyethylene-based inhibitors bound to human immunodeficiency virus type 1 protease show that the inhibitors are present in two distinct orientations. *J Biol Chem 267(32)*:22770-22778.

[21] Ortiz AR, Pisabarro MT, Gago F, and Wade RC. 1995. Prediction of drug binding affinities by comparative binding energy analysis. *J Med Chem 38(14)*:2681-2691.

[22] Perez C, Pastor M, Ortiz AR, and Gago F. 1998. Comparative binding energy analysis of HIV-1 protease inhibitors: incorporation of solvent effects and validation as a powerful tool in receptor-based drug design. *J Med Chem 41(6)*:836-852.

[23] Raag R, Poulos TL. 1991. Crystal structures of cytochrome P-450CAM complexed with camphane, thiocamphor, and adamantane: factors controlling P-450 substrate hydroxylation. *Biochemistry 30(10)*:2674-2684.

[24] Vajda S, Weng Z, Rosenfeld R, and DeLisi C. 1994. Effect of conformational flexibility and solvation on receptor-ligand binding free energies. *Biochemistry 33(47)*:13977-13988.

[25] Verkhivker G, Appelt K, Freer ST, and Villafranca JE. 1995. Empirical free energy calculations of ligand-protein crystallographic complexes. I. Knowledge-based ligand-protein interaction potentials applied to the prediction of human immunodeficiency virus 1 protease binding affinity. *Protein Engineering 8(7)*:677-691.

[26] Wallqvist A, Jernigan RL, and Covell DG. 1995. A preference-based free-energy parameterization of enzyme-inhibitor binding. Applications to HIV-1-protease inhibitor design. *Protein Science 4(8)*:1881-1903.

[27] Weber PC, Pantoliano MW, Simons DM, and Salemme FR. 1994. Structure-based design of synthetic azobenzene ligands for streptavidin. *J Am Chem Soc 116(7)*:2717-2724.

[28] Weng Z, Vajda S, and DeLisi C. 1996. Prediction of protein complexes using empirical free energy functions. *Protein Science 5(4)*:614-626.

[29] Wilson C, Mace JE, and Agard DA. 1991. Computational method for the design of enzymes with altered substrate specificity. *J Mol Biol 220(2)*:495-506.

[30] Wojtczak A, Luft J, and Cody V. 1992. Mechanism of molecular recognition. Structural aspects of 3,3'-diiodo-L-thyronine binding to human serum transthyretin. *J Biol Chem 267(1)*, 353-357.

Chapter 3

Extracting microscopic energies from oil-phase solvation experiments.

Rachel Brem, Hue Sun Chan, and Ken A. Dill

This chapter was originally published as "Extracting microscopic energies from oil-phase solvation experiments," *J. Phys. Chem. B* **104**(31) (2000)

3.1 Abstract.

There is a large literature that reports oil-water partition coefficients of small molecule compounds. A goal of much of this work is to obtain hydrophobicity parameters for models of folding, docking, binding, conformational changes, partitioning and flow, mainly of biomolecules and drugs. The quantity obtained from experiments is the difference between a solute's oil and water-phase chemical potentials $\Delta\mu \equiv \mu_{wat} - \mu_{oil}$, but the quantity needed for models is a contact free energy per unit area. Usually, solvation modelers find this quantity in two steps: (i) subtract off some assumed "irrelevant" solute entropies from the experimental $\Delta\mu$'s, then (ii) divide the result by the solute area. Here we show that such a procedure can lead to large errors. We use a 2-dimensional lattice model of hydrocarbon liquids in NVT Monte Carlo simulations, where the underlying contact energy is rigorously known in advance, to test strategies for converting partitioning data to contact energies. We find that rather than using $\mu_{wat} - \mu_{oil}$ and subtracting assumed entropies, it is better to use $\mu_{wat} - h_{oil}$, where h_{oil} is the solute's enthalpy of solvation in the oil phase. We apply this strategy to experimental thermodynamic data from liquid alkanes and propose that the best estimate for the microscopic free energy of transfer of a short-chain hydrocarbon from oil to water is around 40 cal/mol$Å^2$.

3.2 Introduction.

A large body of literature describes the oil-water partitioning equilibria of solutes. The principal motivation for such measurements has been to model the way more complex molecules, particularly biomolecules, fold, dock, bind, and partition from one medium into another. We use the term BIPSE to refer to models that rely on oil-water partitioning data for their parameters. BIPSE stands for Break Into Pieces, Sum the Energies. For example, Figure 3.1 shows how to model protein folding as a transfer from water to oil of small molecule mimics of core hydrophobic amino acids. This parallels the true folding process in which the hydrophobic

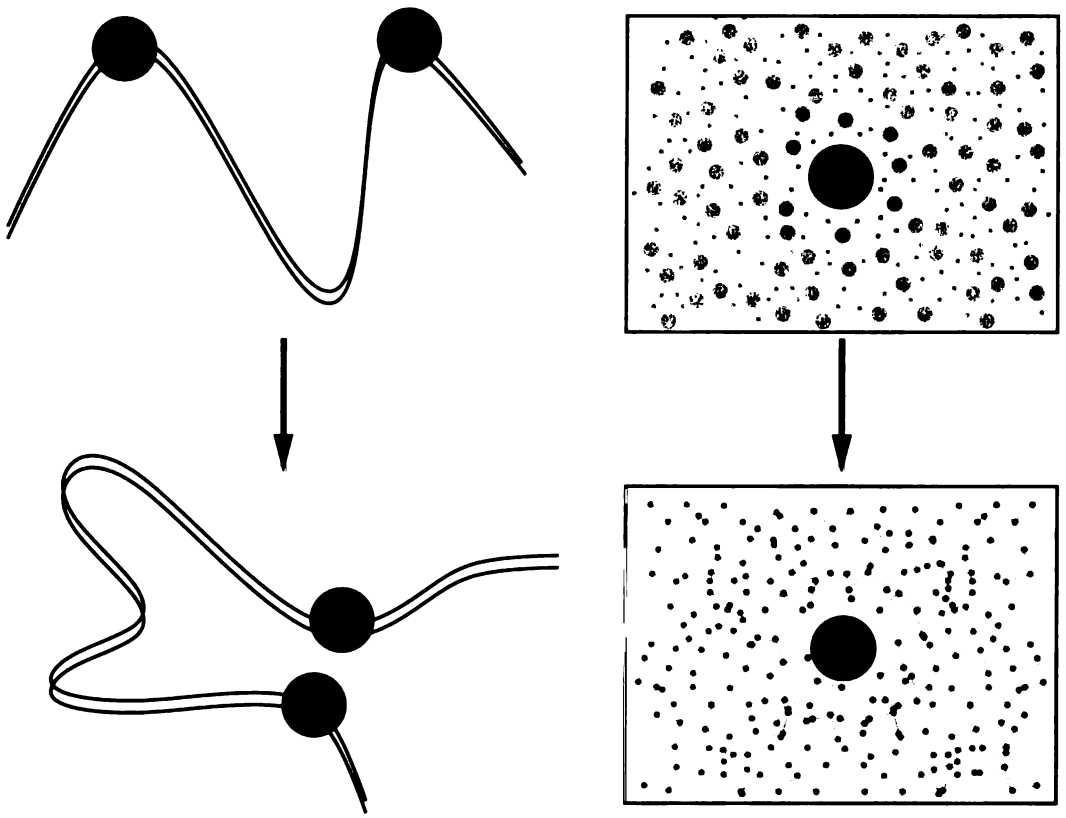


Figure 3.1: **Modeling protein folding with oil-water partitioning.** Large black balls represent amino acids. The fully solvated state of an unfolded protein (top left) can be modeled by small molecules in water (top right); interactions between core amino acids in a folded protein (bottom left) can be modeled by small molecules in alkane liquids (bottom right).

amino acids start by being solvated in water and end by being "solvated" in a hydrophobic core.

To use a BIPSE model, one needs to know the free energy change of each amino acid's transfer between media; an estimate comes from oil-water partitioning data on model compounds in solution. Then in the BIPSE model, transfer free energies of all the amino acids are summed to get the total free energy of folding. Thus, for a BIPSE model of protein folding,

$$\Delta G = \sum_i \gamma_i \sigma_i - T\Delta S_{\text{conf}} + \Delta E. \tag{3.1}$$

In Equation 3.1, ΔE represents all energies except those involved in solvation; ΔS_{conf} represents the entropy of the relevant conformational change; and the first term represents all effects of solvation. σ_i is the monomer surface area whose exposure changes upon folding, and

γ_i is the free energy cost of that change in surface exposure. Since γ contains information from two separate solute insertions, one into oil and one into water, we can also write $\gamma = \gamma_{wat} - \gamma_{oil}$. These parameters may be derived from experimental data for use in the BIPSE model; they are not experimental observables themselves. The simple BIPSE models we study here assume that all monomers have the same transfer free energy, $\gamma_i = \gamma$ for all i; we make this simplification because we are primarily interested here in the statistical mechanical basis of these analyses. But the method can readily be generalized to treat the transfer of different amino acids, as has been used for the determination of hydrophobicity scales [1].

In the second term of Equation 3.1, ΔS_{conf} includes all the relevant conformational entropies that are not accounted for in the model compound solvation and desolvation process. For example, for folding processes, ΔS_{conf} would include the conformational entropies due to the loss of backbone and sidechain degrees of freedom when the chain becomes compact and folded. Since it is usually impossible to get ΔS_{conf} from experiments, that term in a BIPSE model usually comes either from theory or computer simulation. Perhaps the first major BIPSE model was the Flory-Huggins theory of polymer solutions [2], which uses a lattice model to obtain ΔS_{conf}, and the γ_i quantities multiplied by $-z/(2kT)$ (where z is the lattice coordination number) are called Flory χ parameters.

Our focus here is not on BIPSE models. Rather, our focus is on ways to extract from experimental partitioning data a proper value of a microscopically meaningful contact free energy γ that can be used in BIPSE models. The problem is that the partitioning of a small molecule compound into oil is not exactly the same as when protein folding sequesters an amino acid into a core, as when a ligand is binding to a protein, or as when a solute partitions into a lipid bilayer membrane. There are differences. Entropies and enthalpies for oil-water partitioning can differ from those of biomolecule conformational changes. For example, the density inside a protein interior can be different than the density within bulk oils, so some attempt must be made to treat these differences if oil-water energies are to be inserted into

models of protein folding, binding, *etc.* It is necessary to decide which aspects of partitioning experiments are relevant to BIPSE models, and which aspects are irrelevant.

In the oil phase, several types of interaction have been proposed to be irrelevant: free volume changes [3], polymer entanglement and steric interference [2, 8, 2], translational entropy of the solute [13], and pressure-volume effects [11, 12, 7]. For each such property, solvation theories have been developed to subtract off its contribution from raw experimental data, presumably leaving only the relevant physics for BIPSE modeling.

To study solvation in the oil phase, we develop a statistical mechanical lattice model of liquid oils. The advantage of our model is that we know exactly in advance the true underlying energy parameter that γ_{oil} should report. Given this knowledge, we ask what mathematical process correctly extracts γ_{oil} from the "data," which in our case come from a Monte Carlo simulation. We suggest that virtually all the entropies of interaction of a nonpolar solute solvated by an oil phase are irrelevant to BIPSE models. Therefore, we propose that the experimental solvation free energy quantity that gives the most useful microscopic interaction information is a water-phase solvation free energy minus an oil-phase solvation enthalpy. We show that our procedure recoups γ_{oil} accurately in lattice model tests. Then we apply the procedure to experimental oil-water data to estimate γ. This γ can be used in folding, docking, binding, and conformational change models.

3.3 Computational and analytic methods.

3.3.1 A lattice model of hydrocarbon liquids.

We consider several different model hydrocarbon liquids, each in a 2-dimensional lattice model. Each liquid is comprised of N identically shaped molecules. By "shape" we mean that each molecule is not necessarily a single lattice site square, as in many lattice models, but rather a covalently linked collection of lattice sites in some particular arrangement [8]. Such

shapes are intended to capture several physical features: that real molecules are not all perfect spheres; that there are different packing arrangements of real molecules in liquids; and that real molecules have internal partition functions. A given lattice molecule in our model is not meant to mimic any specific alkane. Rather, we want a pantheon of molecular species to explore effects of shape in general.

Molecules in our model interact *via* a pairwise contact energy. When one lattice site unit of surface area is in contact with another lattice site unit of surface area not covalently linked to the first, they make a favorable interaction of ε energy units. All results in this paper were generated with $\varepsilon = -1kT$. There is only one energy type: all lattice sites have the same "chemistry." There are only contact interactions among immediate spatial neighbors, and no longer-ranged interactions. Excluded volume is enforced by the requirement that no lattice site can be occupied by more than one molecular segment. We focus on the insertion of molecule $N + 1$ of the given shape, called the test particle, into a solvent of N other particles of the same shape.

3.3.2 Monte Carlo simulations.

The simulation data we use is obtained from Monte Carlo simulations in the canonical (NVT) ensemble. See the Appendix for a discussion of our choice of constant-volume conditions. All simulation data presented here were generated using periodic boundary conditions, in systems of about 50,000 molecules and simulation boxes of 100,000 to 400,000 lattice sites. Our move set contains only one type of move, which is as follows. We randomly choose a single molecule m from all the N molecules in the simulation box. That molecule is picked up from its current position a, then put down into a new position b elsewhere in an arbitrary orientation. We accept or reject the move based on the Metropolis criterion. To reduce computer time, we do not try moves which we know will always be rejected, *i.e.*, those that violate excluded volume. For this algorithm, after picking up the molecule from position a, we make a list of

all *unobstructed* molecular-shaped holes in the box. Say there are H such holes. We pick the molecule's final position, b, randomly from this list of H entries. The probability of picking molecule m and attempting the $a \rightarrow b$ transition is then $(1/N) \times (1/H)$. Detailed balance is satisfied because we only make the list of empty holes after picking up the molecule. Imagine the $b \rightarrow a$ transition instead, with b as the original position; there are still H available holes once the molecule is picked up. So the probability of attempting the $b \rightarrow a$ transition is still $(1/N) \times (1/H)$, as is required by detailed balance. This means our simulations will converge to the Boltzmann distribution.

To calculate the chemical potential, energy, and entropy of insertion μ, e, and s, we use the Widom procedure [9]. It gives the thermodynamics of inserting a test particle $N + 1$ into a solvent of N other molecules. In the Widom method, the Monte Carlo trajectory is simulated using a Hamiltonian H_N which takes account of the interactions between the N solvent molecules. In each Monte Carlo frame, however, the test particle is placed at random in the simulation box, and its energy of interaction U with the solvent is tabulated separately. The final thermodynamic calculations require averages over the ensemble of N solvent molecules [9, 11]:

$$\begin{aligned} \mu &= -kT \ln\langle \exp(-\beta U)\rangle_N \\ e &= \frac{\langle H_{N+1} \exp(-\beta U)\rangle_N}{\langle \exp(-\beta U)\rangle_N} - \langle H_N\rangle_N \\ Ts &= \mu - e. \end{aligned}$$

Here $\beta = 1/kT$ where k is Boltzmann's constant (here, units are chosen so that $k = 1$) and T is temperature, and $H_{N+1} = H_N + U$. All averages $\langle \cdots \rangle_N$ are performed over a converged trajectory of the pure reference solvent.

The Widom process averages over all solvent configurations in the presence of the inserted test particle. e, s, and μ don't contain translational or rotational entropy of the test particle; they describe only the insertion of test particles at fixed positions. We call these excess

quantities [12], μ_{ex}, e_{ex}, and s_{ex} respectively, because they characterize the thermodynamics beyond that for inserting an "ideal gas molecule" into an otherwise empty simulation box. See Equation 3.2 below for the mathematical definition of μ_{ex}; $s_{ex} = \partial\mu_{ex}/\partial T$, and $h_{ex} = \mu_{ex} + Ts_{ex}$. Uncertainties in our calculated μ_{ex}, e_{ex}, and s_{ex} are $0.05kT$ or less.

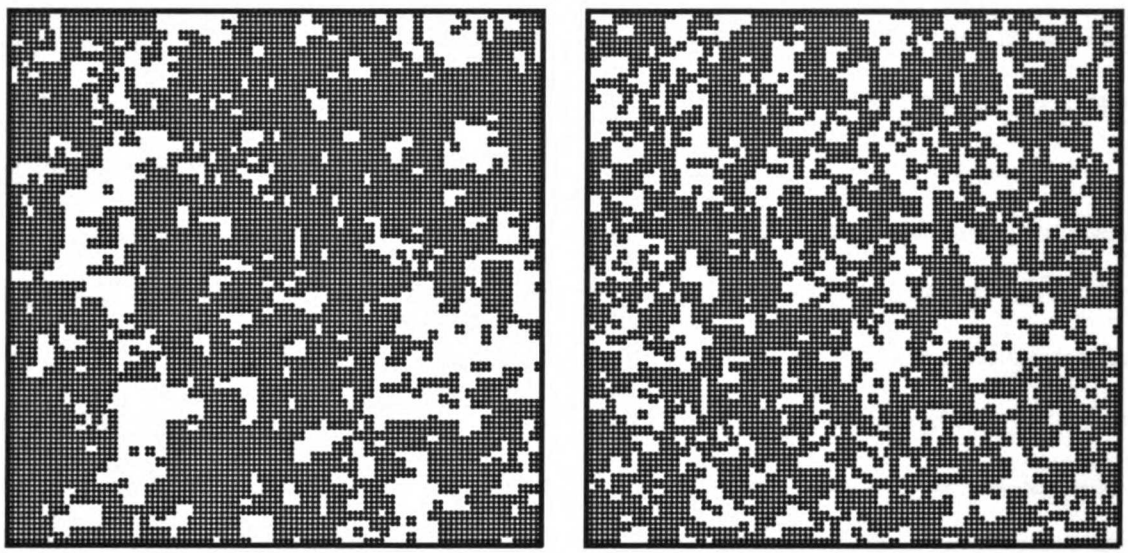
3.3.3 Properties of the pure lattice liquid.

Our lattice model involves substantial simplifications compared to real hydrocarbons. But we believe it contains the relevant physics we need to study the problem of extracting γ_{oil}, for the following reasons: (i) We study many different sizes and shapes of molecules, to mimic the diversity of alkanes. (ii) We study large systems, so we eliminate artifacts due to small system size. (iii) We study systems having about the same space-filling fraction, $\phi \approx 0.7$, as in hydrocarbon liquids [19]. (iv) We assign the temperature so that energies dominate the thermodynamics of insertion, as is true in hydrocarbon liquids [13]. (v) At appropriate temperatures, the particles are mobile but local density fluctuations are small; see Figure 5.4.

Our lattice fluids undergo no sharp phase transitions. Rather, with increasing temperature, these systems pass gradually from a well-packed, solid-like phase to a high-pressure-gas-like phase. (They cannot condense to a true solid or expand to a true gas, because of the constant-V restriction.) The state we call "liquid-like" is in between these two extremes. All thermodynamics shown below are from this state, at $T = 1.5|\varepsilon|/k$. At this temperature, energies dominate the thermodynamics of insertion, and local density fluctuations are small. Because our interest here is only in the liquid-like range, and not in the behavior at phase boundaries, we believe the present model is adequate for our purposes.

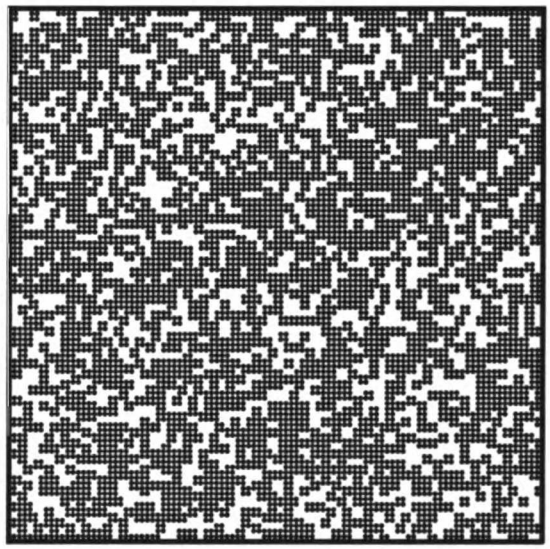
3.3.4 Analytic methods for extracting γ quantities.

This paper describes ways to find γ_{oil}, the solvation parameter in BIPSE models that comes from the oil phase of the partitioning experiment. In our calculations below, we mimic



(a)

(b)



(c)

Figure 3.2: **Phase-like behavior of the 2-D lattice model.** Every dot on the figure represents a filled lattice site. Each panel shows a randomly chosen snapshot from the top left 100×100 lattice sites of 250,000-site Monte Carlo simulation boxes, during simulations of homogeneous fluids of 2×2 lattice squares. (a) $T = 1.0|\varepsilon|/k$; (b) $T = 1.5|\varepsilon|/k$; (c), $T = 10.0|\varepsilon|/k$.

the oil-phase half of the experiment and the theory applied to it. First we run Monte Carlo simulations on lattice model "oil" fluids and calculate thermodynamics over the trajectories. These measurements represent "experimental data" of test particles inserting into oil solvents. To the data we apply approximate theoretical treatments which are designed to derive a value for the parameter γ_{oil}.

The two standard approaches that have been used are density-based theory and Flory-Huggins theory. Density-based theory assumes that the free energy of insertion can be split into two contributions: (i) the entropy of the test particle s_t, and (ii) all other enthalpy and entropy changes that happen upon test particle insertion, lumped together and denoted μ^*. In the simplest density-based theory, the test particle is assumed to have no internal degrees of freedom [13]. In this case s_t, the test particle entropy term, is entirely translational. Assuming that the insertion happens in a box of volume V, the test particle translational component is simply $k \ln V$. A factor of $\ln(1/N)$ also appears in entropies of insertion to account for indistinguishability of particles, so $s_t/k = -\ln \rho$, where $\rho = N/V$. Our case is slightly more complicated than this, since our lattice molecules also have rotational degrees of freedom. The test particle entropy factor in our case is $s_t/k = \ln q_{\text{int}} + \ln V + \ln(1/N)$, where q_{int} is the internal (rotational) partition function of the test particle. Thus, for density-based theory,

$$\begin{aligned} \mu^* &= \mu - (-Ts_t) \\ &= \mu - (-kT \ln q_{\text{int}} + kT \ln \rho) \\ &= \mu_{\text{ex}} \end{aligned} \tag{3.2}$$

where the last line follows from the definition of excess quantities. μ_{ex} is directly computable from our simulations. Density-based theory assumes that all the physics in μ^* are relevant to protein folding, drug binding, and other biomolecule problems. If so, the microscopic contact free energy quantity we seek, γ_{oil}, could be obtained from the experimental quantity μ^* using

$$\gamma_{\text{oil}} = \frac{\mu^*}{f(\sigma)}. \tag{3.3}$$

where $f(\sigma)$ is some function of surface area σ, usually $f(\sigma) = \sigma$ [3].

An alternative way to assign a value for the parameter γ_{oil} is with Flory-Huggins (FH) theory, which was originally developed [2] for long polymers. FH theory treats the translational entropy and the conformational entropy of the chain, which are assumed to be coupled through excluded volume. The Flory-Huggins chemical potential at constant V, for a test particle inserted into its own pure liquid, can be written [2]

$$\mu = \sigma\phi\varepsilon + kT\ln\rho - mkT\ln(1-\phi) - kT\ln q_{\text{int}}. \tag{3.4}$$

The first term on the right-hand side of Equation 4.1 is the energy of insertion; the second accounts for test particle translational entropy; the third is the excluded volume term; and in the fourth, q_{int} is a single polymer's internal partition function. Here m is the polymer chain length (or molecular volume), σ is the total number of lattice units of surface area in a polymer molecule, ϕ is the fraction of space filled in the solution ($\phi = m\rho$), and ε is the contact energy between two unit lattice site contacts of surface area. In terms of excess quantities, the FH expression is

$$\mu_{\text{ex}} = \sigma\phi\varepsilon - mkT\ln(1-\phi). \tag{3.5}$$

The Flory-Huggins solvation strategy is based on the idea that when a solute partitions into an oil phase of chain molecules, the dilution of the oil leads to a gain in conformational entropy of the oil-phase chains; dilution reduces the constraints imposed by neighboring chains. Because this dilution entropy of the oil chains is not relevant for BIPSE models, it should be subtracted from experimental partitioning free energies to give a proper microscopic contact free energy γ_{oil}, as follows:

$$\gamma_{\text{oil}} = \varepsilon = \frac{\mu_{\text{ex}} + mkT\ln(1-\phi)}{\sigma\phi}. \tag{3.6}$$

3.3.5 Choosing molecular volumes for alkanes.

For some of our calculations below we need to know ϕ, the fraction of filled space, in alkane liquids. For real molecules $\phi = \rho M v$, where M is the molecular weight and v is the volume taken up by one alkane molecule. We take ρ (at 20° C) and m from published experimental data [20], but v must be calculated. We use the v values calculated by Abraham *et al.* [24]. When we substitute those values into $\rho M v$, we get $\phi = 0.65$ to 0.8 for all the alkanes we study, which is the general liquid range [19].

3.4 Results and discussion.

3.4.1 The oil-phase enthalpy premise.

Here is the traditional way to extract a microscopic solvation free energy γ. First measure the partition coefficient K of a solute, which can be defined in two ways: (i) in terms of x_{wat} and x_{oil}, the mole fraction of the solute in the oil and water phase respectively, in which case $K \equiv x_{wat}/x_{oil}$ [17]; or (ii) in terms of molarity of the solute in each phase ρ_{wat} and ρ_{oil}, in which case $K \equiv \rho_{wat}/\rho_{oil}$ [18, 19]. The molarity-based definition is more directly related to a simple general statistical mechanical formulation of solvation [13, 3]. Next, given an experimental K, take its logarithm and multiply by $-kT$ to get $\Delta\mu^0 = \mu^0_{wat} - \mu^0_{oil} = -kT \ln K$, where μ^0_{oil} and μ^0_{wat} are the excess chemical potentials of the solute into oil and water phases, respectively. (If the molarity-based definition is used in the traditional method, these excess chemical potentials μ^0 correspond to the μ_{ex} defined above. However, if mole fractions are used, they are different from μ_{ex} [3].) Then assume that $\gamma = \Delta\mu^0/\sigma$. This simple scheme yields the familiar 25-33 cal/molÅ^2 estimate for γ [20]; because the $\ln K$ procedure involves excess chemical potentials, entropic effects from translation of the solute are assumed not to contribute to the final γ.

This procedure has been criticized [2, 13, 11, 12, 7, 3, 21, 10, 6] on the grounds that

excess chemical potentials contain entropies that are irrelevant to BIPSE models. Solvation theories deal with the problem by first estimating the numerical contribution of the irrelevant entropies, and then subtracting it from the experimental μ. Such a subtraction scheme yields a corrected value μ^{corr}. For Flory-Huggins theory, $\mu^{corr} = \mu - (-kT \ln q_{\text{int}} + kT \ln \rho) + mkT \ln(1 - \phi)$ (see Computational Methods section for details). The difference between corrected chemical potentials is the resultant unit solvation parameter for use in BIPSE models:

$$\begin{aligned} \gamma &= \gamma_{\text{wat}} - \gamma_{\text{oil}} \\ &= \frac{\mu_{\text{wat}}^{corr}}{f(\sigma)} - \frac{\mu_{\text{oil}}^{corr}}{g(\sigma)}. \end{aligned}$$

Here $f(\sigma)$ and $g(\sigma)$ are functions of the surface area σ; in Flory-Huggins theory, $g(\sigma) = \sigma\phi$ (see Equation 3.6).

Subtraction schemes are useful because all their ingredients come from experiments. However, they rely on estimates and models of the putative irrelevant entropies which can't be checked independently.

In this paper we propose an improved strategy for the oil-phase half of the partitioning experiment. Our goal is to avoid subtracting entropies that depend on model-based assumptions. We develop a protocol which uses experimental data without resting on modeled entropies. Our protocol starts with the assumption that *all* of the complicated entropies of oil molecules are irrelevant to BIPSE models. No oil-phase entropy—chain effects, translation, *etc.*—belongs in γ quantities. The reason is that, within the philosophy of BIPSE modeling, the first term in Equation 3.1 aims to capture only the drive for the solute to replace its water-like surroundings by oil-like surroundings. Any entropies of restriction, conformational change, density change, or organization that are specifically relevant to folding, docking, or other BIPSE processes, are to be captured in the conformational entropy term.

This idea, which we call the enthalpy premise, means that γ_{oil} should be an interaction energy and not a free energy. It follows that when we estimate γ_{oil}, we shouldn't use the chemical potential μ_{oil} itself, because we aim to leave out s_{oil}. Instead, we use only the oil-phase enthalpy of insertion, h_{oil}. A similar proposal was made by Makhatadze and Privalov [24]. Oil-phase enthalpies can be obtained directly from experiments [13].

The logic of the enthalpy premise holds only for the oil phase. In the water phase, solvation and desolvation entropies of nonpolar solutes are essential in γ_{wat}, since these are the driving forces that γ aims to capture in BIPSE models. For example, a nonpolar solute may order surrounding waters; this involves an entropic free energy cost that drives partitioning [11, 25, 26]. That same entropy is relevant to folding, binding, *etc.* in BIPSE models. So we want water-phase entropies in our solvation parameter. It's only in the oil phase that we use enthalpies exclusively.

Thus, the enthalpy premise is

$$\gamma = \frac{\mu_{wat}^{corr}}{f(\sigma)} - \frac{h_{oil}}{g(\sigma)}. \tag{3.7}$$

In other words, we take the difference between the chemical potential of a solute into the water phase, and its enthalpy of insertion into the oil phase, as the basis for the solvation parameter γ.

The purpose of this enterprise has sometimes been misinterpreted as having to do with associating some thermodynamic quantity with the experimental data. But that is not our goal here. We aim to extract from the experimental data a quantity that will have microscopic meaning in BIPSE models. In general γ is not equal to the quantity μ^* which Ben-Naim refers to as the "solvation free energy" [13, 27].

The Flory-Huggins strategy embodied in Equation 3.6 can also be regarded as a statement of the enthalpy premise, except that it applies under more limited circumstances than the approach we propose here. FH theory explicitly subtracts a mean-field estimate of the dilution

entropy. If there are other polymeric entropies, or flaws in the FH mean-field approximation, our approach should handle them more accurately because our strategy is to let the experimental data speak for itself, rather than to make model approximations and assumptions about it.

Note that a key assumption in BIPSE models is that any entropies in the γ and ΔS_{conf} terms (see Equation 3.1) are additive. Additivity assumptions such as these have been criticized [28]. The enthalpy premise we propose here helps circumvent at least one additivity assumption, since the model no longer needs to add the entropies from liquid oils to the biomolecule entropies in ΔS_{conf}.

3.4.2 Application of van Laar theory to solvation.

We now show that there is a third approach-based on the van Laar theory of solvation [2]—that extracts a microscopic energy quantity from partitioning data. It is very simple, and yet we find that when used in conjunction with the enthalpy premise, it more accurately captures the true underlying energy quantities we seek than the two methods described above. The van Laar model estimates the observable, macroscopic energy e as $e = \sigma\phi\varepsilon$, where σ is the surface area of the inserted molecule, $\phi = v\rho$ is the fraction of space filled in the solution, v is the molecular volume, and ρ is the density. The strategy is to measure e and predict γ_{oil} using known values of surface area and the fraction of filled space in the liquid:

$$\gamma_{\text{oil}} = \varepsilon = \frac{e}{\sigma\phi}. \tag{3.8}$$

3.4.3 The van Laar energy model extracts more accurate energy parameters than traditional methods.

In this section, we test several strategies for parameter extraction to see which strategy most accurately extracts the true underlying solvation energetics from data. We mimic experimental solvation data with those from Monte Carlo simulations of a known underlying model.

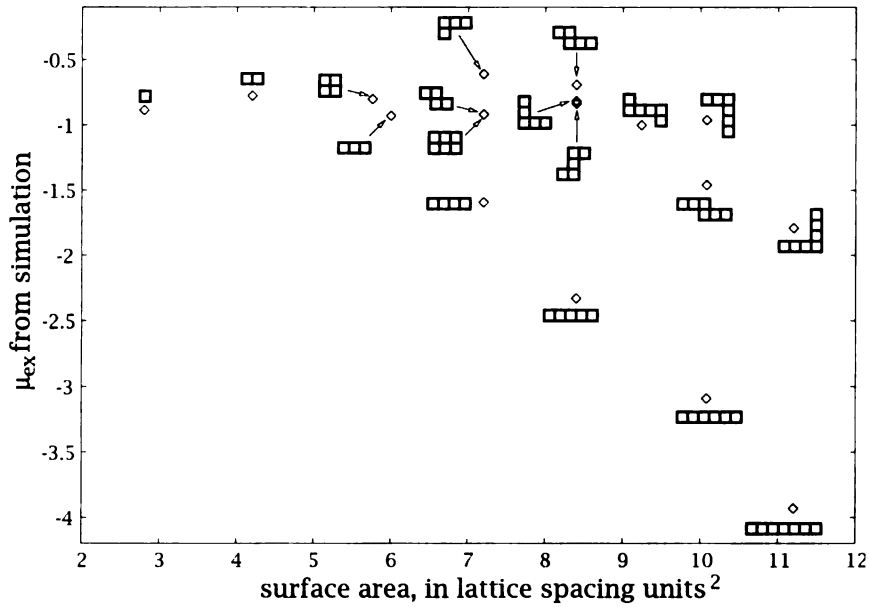


Figure 3.3: **Chemical potentials don't correlate with surface area.** Every data point represents one lattice molecule type. The x axis plots the test particle surface area. The y axis plots the chemical potential of insertion of the molecule into its own pure fluid, calculated from Monte Carlo simulations at $T = 1.5|\varepsilon|/k$.

For each of the several different model hydrocarbons we study, we compute the excess chemical potential μ_{ex} or excess energy e_{ex} of insertion of a particle into a fluid of identical molecules. Then we use Equations 3.3, 3.6, or 3.8 to extract from the "experimental data" an estimate of γ_{oil}. The ideal value of γ_{oil} should be the sticking energy ε in our model, which we assign in advance to be $\varepsilon = -1kT$. Knowing this, we can determine which extraction method most accurately captures the underlying energetics of our fluids.

The most traditional way to extract energy parameters is simply to plot the chemical potential of insertion against the surface area of the test particle, Equation 3.3. The slope of such a line is usually taken as the fundamental quantity of solvation, an energy per unit area. But Figure 3.3 shows that in our lattice model, there isn't a significant correlation in such a plot. Thus, in general it's not possible to extract accurate contact energies from data using the traditional approach. The reason is that chemical potentials contain more information than just contact energies; this is why we need more detailed solvation models.

Next we test the Flory-Huggins strategy, Equation 3.6. In this approach, we extract energies from chemical potentials as follows. First we use the FH entropy model to calculate the entropy of insertion analytically. Then we multiply the latter by $-kT$, and subtract it from an "experimental" chemical potential. If the model works well, the resulting quantity is an accurate estimate of the energy of insertion. Finally we analyze these estimated energies with the van Laar model to find the solvation interaction energy. The results of the FH protocol are shown in Figure 3.4. In the van Laar model $\sigma\phi$, which is given on the x-axis of Figure 3.4, is a measure of the number of intermolecular contacts between the test molecule and its surroundings. Thus, the slope of this plot—the number of contacts made upon insertion *versus* the total energy of insertion—should report the true underlying interaction energy, which we want for our solvation parameter γ_{oil} (see Equation 3.8). However, the predicted γ_{oil} in Figure 3.4 is about 30% in error. This is due to shortcomings of the FH entropy model and van Laar theory.

Next we apply the enthalpy premise: we use "experimental" enthalpies instead of chemical potentials to estimate microscopic energy parameters. If we take the traditional approach and plot e_{ex} *versus* surface area, we get a straight line, as can be seen in Figure 3.5. But the slope, the putative energy per unit surface area, is only 60% of the true interaction energy of the model. The reason is that surface area alone doesn't adequately describe the thermodynamics of insertion.

Finally we test the enthalpy premise in conjunction with van Laar theory. The results are shown in Figure 3.6. The estimated energy parameter is 86% of the true interaction energy of the model. This indicates that the van Laar method is a substantial improvement relative to the strategies described above. Moreover, the slope approaches 100% accuracy as we increase the temperature of the Monte Carlo simulations (data not shown). We believe this indicates that the lattice model becomes more realistic at higher temperatures because of reduced density fluctuations (see Figure 5.4).

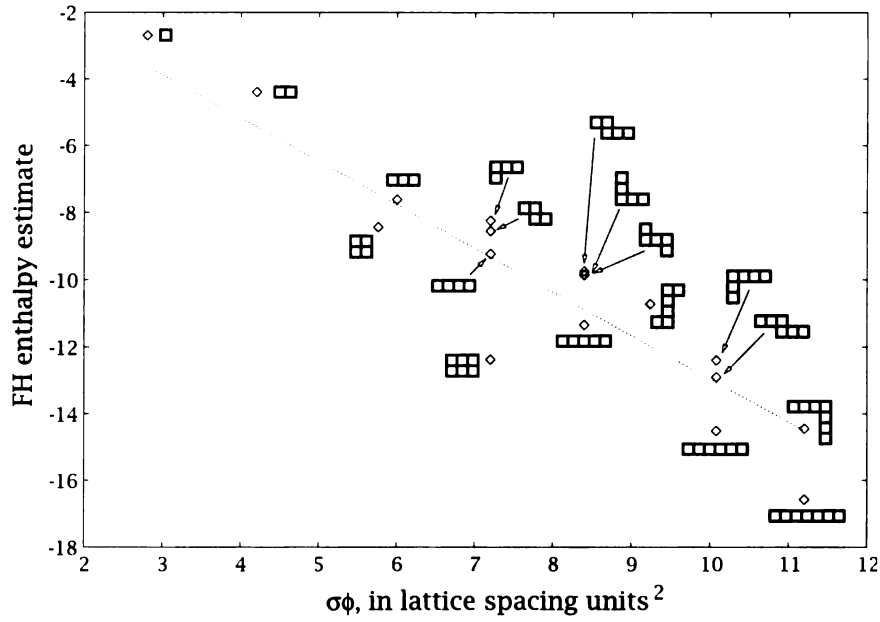


Figure 3.4: **FH extracts solvation energies from chemical potentials with moderate accuracy.** Every data point represents one lattice molecule type. The x axis plots the van Laar estimate of the number of contacts made by the test particle upon insertion: σ is the test particle surface area, and ϕ is the fraction of space filled in the model alkane fluid. The y axis plots the Flory-Huggins estimate of energy: $\mu_{ex} + kTm\ln(1-\phi)$, where μ_{ex} is the chemical potential of insertion of the molecule into its own pure fluid, calculated from Monte Carlo simulations at $T = 1.5|\varepsilon|/k$; m is the number of lattice sites in the molecule. The fit to a line of intercept 0 (dashed line) is $y = -1.29535x$. If Flory-Huggins and van Laar theory made perfect predictions, the data would fall along a line of slope $\varepsilon = -1kT$ and intercept 0 (see Equation 3.8).

This last result, Figure 3.6, is the cornerstone of this paper. It illustrates an accurate way to extract microscopic energies from models of liquid oils. Our protocol, which analyzes energies of insertion with van Laar theory, is more accurate than traditional methods because it captures the physics of energies, but it doesn't require modeling of entropies.

3.4.4 Application to experimental hydrocarbon data.

In this section, we apply the same strategies as above to experimental data. Figure 3.7 shows the correlation of experimental enthalpies of insertion of various hydrocarbons into their own pure liquids, *versus* surface area of the molecule. As in the lattice model (see Figures 3.5), the plot is linear; the slope of the curve here is -27.0 cal/molÅ^2. Figure 3.8 shows the

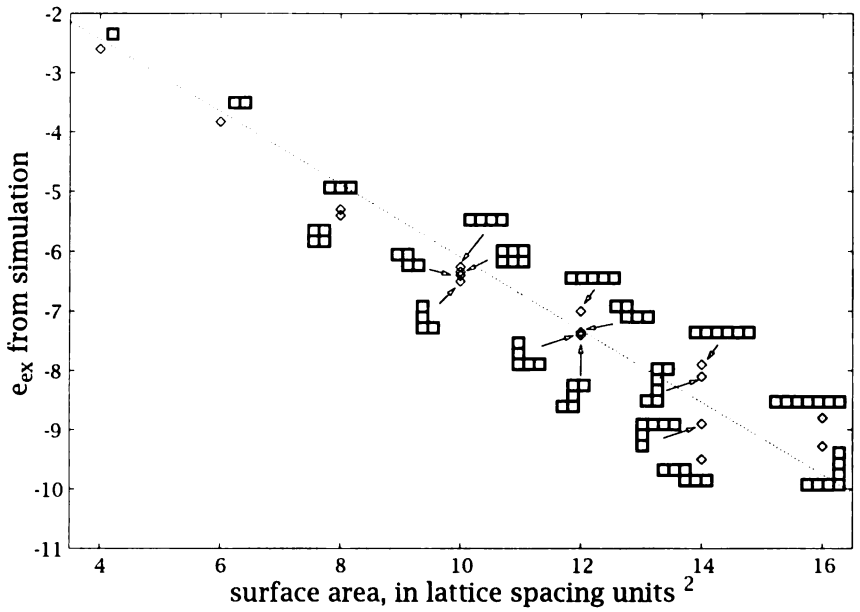


Figure 3.5: **Surface area alone doesn't extract an accurate energy parameter from the energy of insertion.** Every data point represents one lattice molecule type. Surface area of the molecule is on the x axis. The y axis plots the energy of insertion of the molecule into its own pure fluid, calculated from Monte Carlo simulations at $T = 1.5|\varepsilon|/k$. The fit to a line of intercept 0 (dashed line) is $y = -0.609568x$. $r = 0.975$. If surface area alone described energies well, the data would fall along a line of slope $\varepsilon = -1kT$ and intercept 0.

corresponding plot for the van Laar strategy. The slope of this line is -36.3 cal/molÅ^2. Because we have shown that the latter strategy is more reliable in the lattice model, for which the true underlying energy is known, we believe the latter value is also therefore the more accurate value for solvation energies of real hydrocarbons in their own liquids, for which the true energies are not known.

Column (v) of Table 1 lists the slopes from Figure 3.8 and other applications of van Laar theory to experimental enthalpies, all using solvent-accessible surface area (SASA). The slopes in Table 1 give $\gamma_{oil} = -31$ to -36 cal/molÅ^2. The spread depends on which alkanes we use in the data sets and on the method used to calculate surface area.

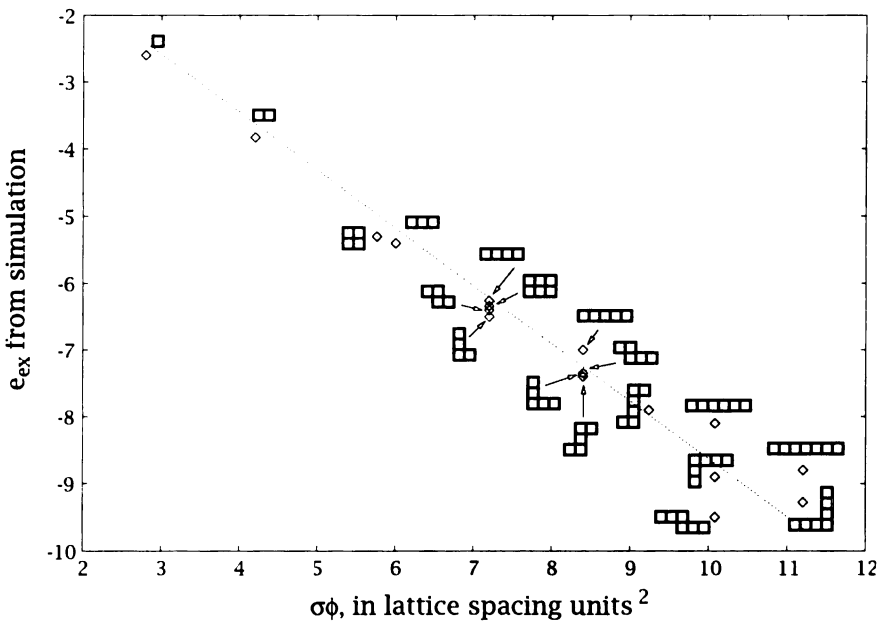


Figure 3.6: **van Laar theory extracts accurate solvation energies from enthalpies.** Every data point represents one lattice molecule type. The x axis plots the van Laar estimate of the number of contacts made by the test particle upon insertion: σ is the test particle surface area, and ϕ is the fraction of space filled in the model alkane fluid. The y axis plots the energy of insertion of the molecule into its own pure fluid, calculated from Monte Carlo simulations at $T = 1.5|\varepsilon|/k$. The fit to a line of intercept 0 (dashed line) is $y = -0.86292x$. If fit to a two-parameter curve, $y = -0.788703x - 0.63577$. $r = -0.983$. If van Laar theory made perfect predictions, the data would fall along a line of slope $\varepsilon = -1.0$ and intercept 0 (see Equation 3.8).

3.4.5 Oil-water partitioning and the strength of the hydrophobic effect.

In the sections above, we have described the solvation of a nonpolar molecule in its own liquid phase. The hydrophobic effect is often characterized by a free energy of transfer of such a molecule between an oil and water phase. To get the free energy of transfer, we now subtract the oil phase free energy of solvation obtained above, from the water phase free energy of solvation,

$$\gamma = \gamma_{wat} - \gamma_{oil} = \frac{\mu_{wat}^{corr}}{f(\sigma)} - \frac{h_{oil}}{\sigma\phi}.$$

In the previous section, we found that γ_{oil} was -32 to -36 cal/molÅ^2. Now we need γ_{wat}. This is the surface-area-dependent free energy of insertion into the water phase, which we can get directly from experiments. If we take $f(\sigma) = \sigma$, then to find γ_{wat} we need to plot

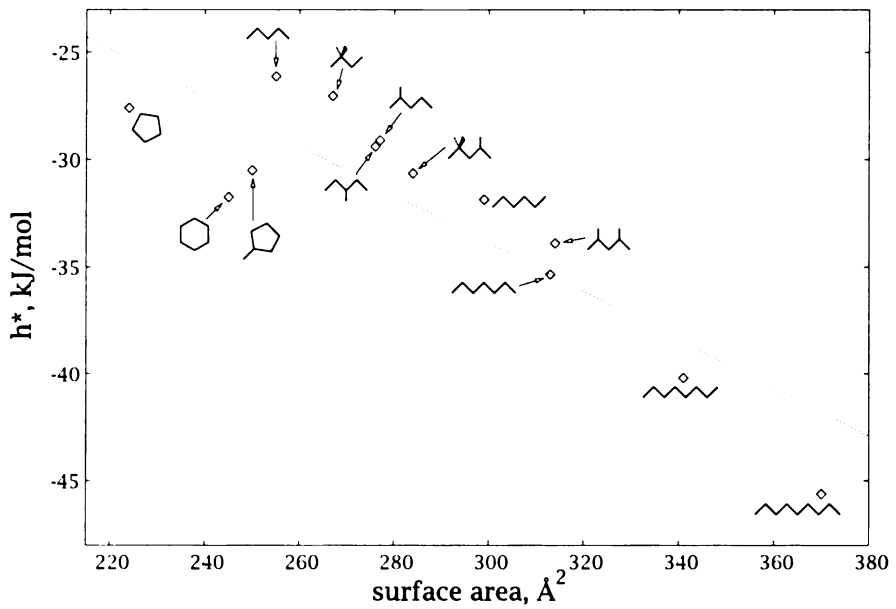


Figure 3.7: **Extracting solvation energies from experimental enthalpies with surface area.** Every data point represents one alkane species. SASA of the alkane molecule, taken from Sitkoff [37], is on the x axis. The y axis plots the enthalpy of insertion of the molecule into its own pure liquid, from Ben-Naim [13]. The fit to a line of intercept 0 (dashed line) is $y = -0.112946x$ (slope = -27.02 cal/molÅ^2). $r = -0.888999$.

chemical potentials against the surface area of the test particle and take the slope. A plot of this type is shown in Figure 3.9. These experimental data are μ^* of hydrocarbons into water from Ben-Naim [13].

Figure 3.9 shows that cyclic and branched alkanes don't have a simple dependence on surface area. However, most γ estimates to date [11, 20] have been from linear alkanes alone, which do fall on a straight line with respect to surface area. In the absence of more sophisticated treatments of the water phase, we follow this convention, and use linear alkanes alone to estimate γ_{wat}. The slopes of the linear alkanes, from this plot and others like it, are listed in column (a) of Table 1: slopes are 5 to 6 cal/molÅ^2 using SASA.

Now we combine the oil phase and water phase results. We compute a total γ estimate from the difference between γ_{wat} and γ_{oil}. When we do this with the data in Table 1 for alkanes with fewer than 10 carbons, using all alkanes for the oil phase (analyzed with a one-parameter

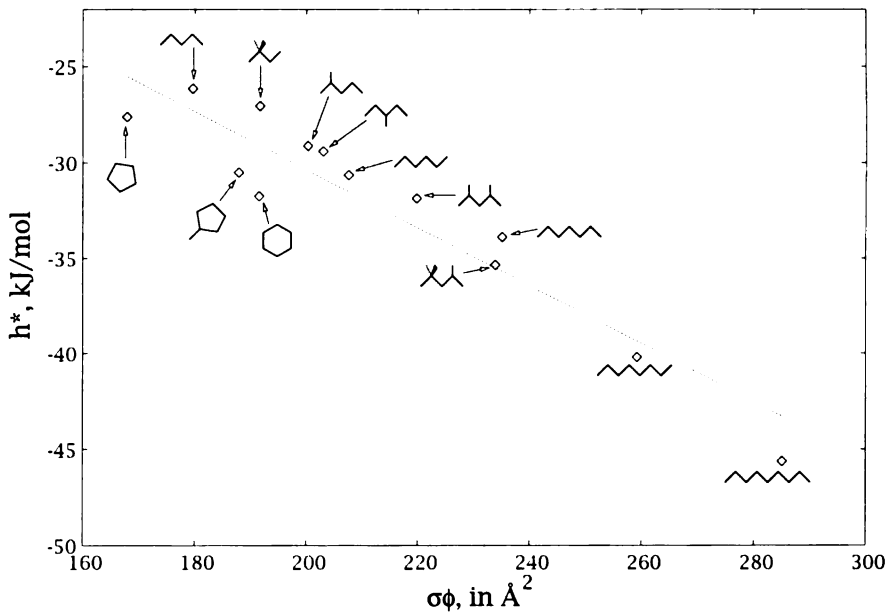


Figure 3.8: **Extracting solvation energies from experimental enthalpies with van Laar theory.** Every data point represents one hydrocarbon species. The x axis plots the van Laar estimate of the number of contacts made by the test particle upon insertion: σ is the test particle surface area, and ϕ is the fraction of space filled in the liquid. The y axis plots the enthalpy of insertion of the molecule into its own pure liquid in kJ/mol, from Ben-Naim [13]. The fit to a line of intercept 0 (dashed line) is $y = -0.151832x$ (slope = -36.32 cal/molÅ^2). If van Laar theory made perfect predictions, the data would fall along a line whose slope is the contact energy parameter (see Equation 3.8). Surface areas are from Sitkoff [37]. Results using other surface area data sets look similar.

linear fit) and straight-chain alkanes only for the water phase, we get 38.5 to 42.6 cal/molÅ^2 using SASA. This is $\gamma = \gamma_{wat} - \gamma_{oil}$ for alkanes.

Note that this value of γ is essentially independent of density or molecular shape. Such a result apparently contradicts the recent arguments of Siepmann and co-workers [21]. They simulate all-atom alkane chain models with Monte Carlo, calculate free energies of insertion, and then apply the FH entropy correction to their data to find a γ_{oil}. The protocol of Siepmann *et al.* is analogous to our Equation 3.6, except that these authors use a denominator of unity instead of $\sigma\phi$. To calculate FH entropy in their models, Siepmann *et al.* convert molecular size from FH lattice units to a unit that's relevant in all-atom systems. They denote such a unit volume V_a. Rather than choose its value arbitrarily, they leave V_a as a free parameter. So they

Table 3.1: **Variation in solvation energy parameters due to SASA calculation and molecular shape.** Numbers are solvation energy parameters, taken from the slopes of linear fits to data, in cal/molÅ^2. Comparing values down a column shows that results depend on the shapes of molecules used in the fits and on the literature reference from which we take SASA. The spread over all this variation is about 6 cal/molÅ^2, which serves as an error bar for our γ_{oil} estimate. (i) Reference for SASA. (ii) Molecular species in data set. a indicates pentane; b, cyclopentane; c, 2,2-dimethylpropane; d, 2-methylbutane; e, hexane; f, cyclohexane; g, methylcyclopentane; h, 2,2-dimethylbutane; i, 2-methylpentane; j, 3-methylpentane; k, heptane; l, 2,4-dimethylpentane; m, octane; n, 2,2,4-trimethylpentane; o, nonane. (iii) Slopes of plots of μ^* of alkanes into water *versus* SASA. (iv) Slopes of plots of μ^* of alkanes into their own pure liquids *versus* SASA. (v) Slopes of plots of h^* of alkanes into their own pure liquids *versus* $\sigma\phi$. Traditional hydrophobicity parameters would be calculated by taking the difference between (iii) and (iv); in this work, we take the difference between (iii) and (v). The number in parentheses is the linear correlation coefficient r of each plot. All alkane thermodynamics were taken from [13].

(i) SASA reference	**(ii)** Molecular species used	**(iii)** γ_{wat}	**(iv)** γ_{oil} **from** μ	**(v)** γ_{oil} **from** h
	straight-chain alkanes			
[37]	a,e,k,m,o	6.0 (0.981)	−22.4(−1.000)	−36.7(−0.999)
[38]	a,e,k,m	5.5 (0.985)	−20.4(−0.994)	−33.8(−0.999)
[39]	a,e,k,m	5.4 (0.982)	−20.0(−1.000)	−32.1(−1.000)
	straight-chain and branched alkanes			
[20]	a,c,d,h,i,j,l,n		−23.2(−0.976)	−31.5(−0.986)
[36]	a,c,h,l,n		−24.8(−0.991)	−33.2(−0.992)
[37]	a,c,d,e,h,i,j,k,l,m,n,o		−22.4(−0.996)	−35.7(−0.988)
[38]	a,c,d,e,h,i,j,k,l,m,n		−21.7(−0.959)	−32.5(−0.979)
[39]	a,c,d,e,h,j,k,l,n,m		−16.6(−0.928)	−31.1(−0.970)
	straight-chain, branched, and cyclic alkanes			
[20]	a,b,c,d,f,g,h,i,j,l,n		−12.5(−0.570)	−32.3(−0.896)
[36]	a,b,c,f,h,l,n		−20.5(−0.815)	−34.3(−0.891)
[37]	a through o		−14.6(−0.840)	−36.3(−0.910)
[38]	a through n		−13.1(−0.709)	−33.0(−0.921)

estimate γ_{oil} multiple times, each time using a different V_a value. They then plot the resulting γ_{oil} against the V_a values that produced them, from calculations using data on hexane, octane, decane, and dodecane. Curves from the different alkanes don't intersect at a single point. Thus, Siepmann *et al.* are unable to find a single value of V_a to make Flory-Huggins fit the data from alkanes of different lengths. Siepmann *et al.* conclude that it's impossible to find a chain-length-independent γ_{oil}.

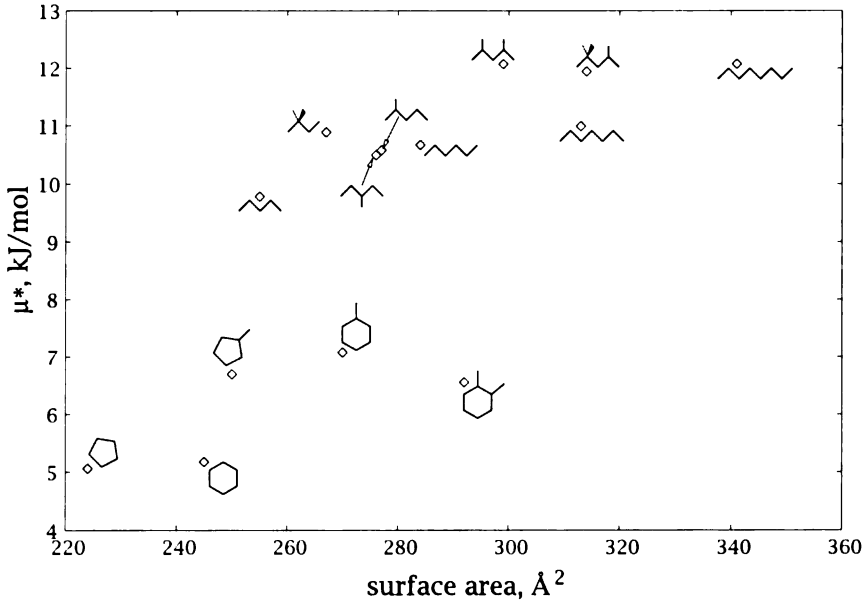


Figure 3.9: **Extracting γ_{wat} from experimental data.** Every data point represents one hydrocarbon species. The x axis plots SASA. The y axis plots the chemical potential of insertion of the molecule into water in kJ/mol, from Ben-Naim [13]. Surface areas are from Simonson [38]. Results using other surface area data sets look similar.

We agree that subtracting the conventional FH term from chemical potentials doesn't account for all oil-phase entropies. But in this paper, we extract γ_{oil} without a subtraction scheme or chemical potentials; we harness oil-phase enthalpies directly. For this reason we are able to extract an essentially universal γ_{oil} from experimental data. Siepmann *et al.* might have gotten closer to a universal γ_{oil} using subtraction schemes if they had used a full FH form like our Equation 3.6—*i.e.*, if they had written μ (their Equation 2.b) with the FH entropy term plus a van Laar term that accounts for the $\sigma\phi$ dependence of enthalpy (our Equation 3.5). The van Laar term is part of classical Flory derivations [2, 2] but doesn't appear in most modern applications of FH theory to solvation [3, 6]. In summary, it is most likely that the results of Siepmann *et al.* arise from inadequacies in the conventional form of FH, which does not model the $\sigma\phi$ dependence of enthalpies with a van Laar term. When we apply van Laar theory to experimental data to account for packing fraction effects, we can extract an essentially chain-length-independent γ_{oil}.

Our γ estimate of $\gamma \approx 40$ cal/molÅ^2 is close to the value $\gamma = 43.1$ cal/molÅ^2 from Makhatadze and Privalov [24, 30]. Makhatadze and Privalov also used the enthalpy premise—that is, they assigned γ to be the difference between a solute's free energy in the water phase and enthalpy in the alkane phase. However, for the latter they used the enthalpy of sublimation from alkane crystals, for which direct experimental data are lacking. Makhatadze and Privalov estimated the enthalpy of sublimation from heat capacity of fusion and vaporization data [30]. By contrast, we use oil-phase enthalpies that come from temperature derivatives of chemical potentials [13]. Makhatadze and Privalov did not apply a van Laar-like model to alkane enthalpy data.

It is also useful to compare our new γ estimate to those derived previously from oil-water partitioning data. Density-based theory, when applied to partition coefficients of transfer between water and chain alkanes with fewer than 10 carbons, yields 31 cal/molÅ^2 for SASA [10]. Using FH in the oil and water phases gives 47 cal/molÅ^2 for SASA [10, 6], using chain alkanes as oils. Our γ estimates lie between those from these two methods, just as they did in the lattice model. Analyzing with density-based theory on cyclohexane instead of chain alkanes, to minimize effects of steric interference [8, 2], Chan and Dill [3] made a tentative estimate for γ of 34 cal/Å^2 using SASA. Our current ESE estimate differs from that of Chan and Dill by ~ 6 cal/molÅ^2. We conclude that ~ 6 cal/molÅ^2 indicates the degree to which density-based theory fails to capture all the entropies in cyclohexane oil phases.

3.5 Conclusions.

Oil-water partitioning experiments are often performed for the purpose of extracting energy differences of the solute with oil, relative to the solute with water. Such quantities can then be used in folding, binding, docking, and transport models. Our focus here is on ways to obtain microscopically meaningful measures of the interaction of a molecule with the oil phase.

To do this, we have performed Monte Carlo simulations of lattice model fluids that are complex (*i.e.* those involving molecules of different shapes). We propose that among the methods we tested, the most meaningful way to extract proper energies is to use experimental enthalpies of solvation in the oil phase, rather than experimental free energies, and then to use the van Laar model to convert the data to contact quantities.

3.6 Acknowledgements.

The authors thank Kevin Silverstein for allowing us to use his unpublished computer code, and for his generosity with advice and help. We thank Seishi Shimizu for very helpful discussions on the relationship between constant-pressure and constant-volume enthalpies and entropies. We also thank the following colleagues for their input: Shi-Jie Chen, Josh Deutsch, Danny Heap, Anton Krukowski, Albert Leo, Jed Pitera, Jack Schonbrun, Noel Southall, and Karen Tang. R.B. was a National Science Foundation predoctoral fellow and a University of California Regents' Fellow while this work was in progress. H.S.C. is supported by a grant (MT-15323) from the Medical Research Council of Canada, and K.A.D. by NIH grant GM34993.

3.7 Appendix.

In this Appendix we describe the relationship between enthalpies of insertion at constant pressure, h_P^*, and energies of insertion at constant volume, e_V^*, defined by:

$$h_P^* \equiv -T^2 \left[\frac{\partial}{\partial T} \left(\frac{\mu^*}{T} \right) \right]_P$$
$$e_V^* \equiv -T^2 \left[\frac{\partial}{\partial T} \left(\frac{\mu^*}{T} \right) \right]_V ,$$

where μ^* is defined in Equation 3.2 in the main text. In general, e_V^* is expected to contain non-local energy contributions beyond those near the inserted solute. During a constant-volume insertion, say when an alkane molecule is added to a solution of N identical molecules at

volume V, the density of the system increases from N/V to $(N+1)/V$. Solvents are crowded together, and the solvent-solvent interaction energy may change significantly. Thus, at constant volume the experimental insertion energy can report much more than just solute-solvent interactions [31, 32, 33]. In this physical picture, it is intuitive that constant-P data should mitigate the problem. At constant pressure, the solution can expand when a molecule is inserted, so crowding and solvent-solvent interaction energy changes should be minimal. This is especially likely in the homogeneous systems we study, since the density at constant pressure of a test tube of N identical molecules is the same as that of $N+1$ molecules. Therefore, constant-P data better reflect the solute-solvent interactions one aims to model by contact energies. A previous general analytical treatment [32] supports this picture.

In our lattice model calculations in the main text, we used constant-volume conditions. However, the properties of the model are such that non-local energy differences caused by test particle insertion at constant V are small in our simulation data. We know this because we can measure such differences directly in Monte Carlo simulations. We calculate the average difference $\langle W\rangle_{N+1} - \langle W\rangle_N$ in solvent-solvent interaction energies W before and after test particle insertion (when the system contains N and $N+1$ shapes, respectively). We find this difference to be 10% or less of the magnitude of the total e_V^* (data not shown). Thus, energies of insertion in our simulations contain only small non-local components, so it is not unreasonable to use these data as a model of experimental constant-P insertion enthalpies.

What is the quantitative difference between e_V^* and h_P^*? From classical thermodynamics, we have:

$$h_P^* - e_V^* = kT^2\alpha + T\Delta V\frac{\alpha}{\kappa_T} = kT^2\alpha + T\frac{\alpha}{\kappa_T\rho}. \tag{3.9}$$

where α is the coefficient of thermal expansion, κ_T is the isothermal compressibility, ΔV is the partial molar volume (the change in volume of the solution upon insertion of the solute) at constant pressure, and ρ is the density of the solvent. We have used $\Delta V = 1/\rho$, since in the systems we study, the solute and solvent are the same species and $1/\rho$ is the partial molar

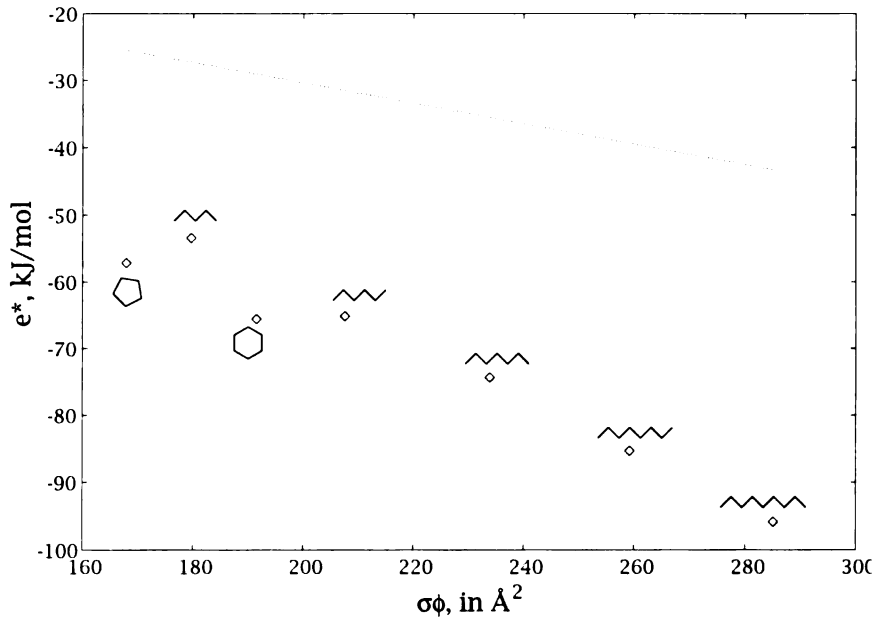


Figure 3.10: **Experimental energies follow van Laar theory.** Every data point represents one hydrocarbon species. The x axis plots the test particle surface area times the fraction of space filled in the solvent, which has units of Å^2. The y axis plots e_V^*, the constant-volume energy of insertion of the molecule into its own pure liquid in kJ/mol. This quantity was computed *via* Equation 3.9 with enthalpies from Ben-Naim [13]. If van Laar theory made perfect predictions, the data would fall along a line whose slope is the contact energy parameter (see text). The dotted line shows the best linear fit to the data in Figure 3.8, in which $\sigma\phi$ is plotted against constant-P enthalpies of insertion h_P^*. The slopes of constant-P and constant-V data are different. Surface areas are from Sitkoff [37]. Isothermal compressibilities and cubic thermal expansion coefficients are from the CRC Handbook [20].

volume of a solvent molecule in solution. Equation 3.9 is very similar to relations derived previously [31, 34].

When we apply Equation 3.9 to experimental enthalpies of solvation of alkanes into their own pure liquids taken at constant pressure [13], we can convert them to e_V^* quantities. We present this e_V^* data as the y values in Figure 3.10. These are about a factor of 2 more favorable than the corresponding h_P^* values (see Figure 3.8). This is consistent with other reports of sizable differences between solvation thermodynamic quantities at constant P *versus* those obtained at constant V [31, 33, 34, 14]. Presumably, the difference arises from significant non-local contributions, *i.e.* changes in interaction energy between solvents upon insertion.

Next we apply van Laar theory to e_V^*, and plot the results in Figure 3.10. Interestingly, e_V^* of alkanes into their own pure liquids are linear in $\sigma\phi$. This is consistent with results from our constant-V lattice model simulation (see Figure 3.6). Thus, the van Laar expression scales with the sum of non-local and local insertion energies in e_V^* as well as it does the local ones alone in h_P^* (compare Figure 3.10 and Figure 3.8).

What is the significance of the slope of Figure 3.10? Our lattice model shows that the slope of a plot of $\sigma\phi$ *versus* an experimental energy quantity gives the underlying solute-solvent contact energy strength. But it will give this value accurately only when the experimental quantity doesn't contain non-local contributions, since van Laar theory doesn't model the latter. For this reason, we believe that the slope of a plot like the e_V^* part of Figure 3.10 isn't a good estimate of γ_{oil}. The slope in Figure 3.10 is about -80 cal/molÅ^2 for alkanes with fewer than 10 carbons. This is twice as strong as the analogous γ_{oil} estimate from constant-P data (see Table 1).

In summary, constant-V insertion energies can be very different from constant-P enthalpies. We attribute the difference to non-local crowding of solvents upon insertion of a solute at constant V, and we argue that this crowding is minimized at constant P. Because van Laar theory doesn't account for non-local effects, we believe it is most appropriate to apply the theory to constant-P data. Based on this reasoning, in Table 1 we have used values of h_P^* rather than e_V^* to estimate γ_{oil}.

Bibliography

[1] DeVido DR, Dorsey JG, Chan HS, and Dill KA. *J. Chem. Phys. B* **1998**, *102*:7272-7279.

[2] Flory PJ. *Principles of Polymer Chemistry,* Cornell University Press, Ithaca, New York, 1953.

[3] Hildebrand JH. *J. Chem. Phys.* **1947**, *15*:225-28.

[4] Krukowski AE, Chan HS, and Dill KA. *J. Chem. Phys.* bf 1995,*103*:10675-10688.

[5] Chan HS and Dill KA. *J. Chem. Phys.* **1994**, *101*:7007-7026.

[6] Ben-Naim A. *Solvation Thermodynamics.* Plenum Press, New York, 1987.

[7] Sharp KA, Nicholls A, Friedman R, and Honig B. *Biochemistry* **1991**, *30*:9686-97.

[8] Kumar SK, Szleifer I, Sharp K, Rossky PJ, Friedman R, and Honig B. *J. Phys. Chem.* **1995**, *99*:8382-8391.

[9] Sharp KA, Kumar S, Rossky PJ, Friedman RA, and Honig B. *J. Phys. Chem.* **1996**, *100*:14166-14177.

[10] Widom B. *J. Chem. Phys.* **1963**, *39*:2808-2812.

[11] Silverstein KAT, Haymet ADJ, and Dill KA. *J. Am. Chem. Soc.* **1998**, *120*:3166-3175.

[12] Allen MP and Tildesley DJ. *Computer Simulation of Liquids.* Oxford University Press, New York, 1996.

[13] Bader RFW, Carroll MT, Cheeseman JR, and Chang C. *J. Am. Chem. Soc.* **1987**, *109*:7968-7979.

[14] Chan HS and Dill KA. *Annu. Rev. Biophys. Biomol. Struct.* **1997**, *26*:425-59.

[15] Lide DR, Editor. *CRC Handbook of Chemistry and Physics.* CRC Press, Boca Raton, FL, 2000.

[16] Abraham MH, Chadha HS, Whiting GS, and Mitchell RC. *J. Pharm. Sci.* **1994**, *83*:1085-1100.

[17] Tanford C. *The Hydrophobic Effect.* Wiley, New York, 1980.

[18] Ben-Naim A. *J. Phys. Chem.* **1979**, *83*:1803.

[19] Tanford C. *J. Phys. Chem.* **1979**, *83*:1802-3.

[20] Hermann RB. *J. Phys. Chem.* **1972**, *76*:2754-2759.

[21] Shimizu S, Ikeguchi M, Nakamura S and Shimizu K. *J. Chem. Phys.* **1999**, *110*:2971-2982.

[22] Sharp KA, Nicholls A, Fine RF, and Honig B. *Science* **1991**, *252*:106-109.

[23] De Young LR and Dill KA. *J. Phys. Chem.* **1990**, *94*:801-809.

[24] Makhatadze GI and Privalov PL. *Adv. Protein Chem.* **1995**, *47*:307-425.

[25] Geiger A, Rahman A, Stillinger FH. *Journal of Chemical Physics* **1979**, *70*:263-76.

[26] Mancera RL and Buckingham AD. *J. Phys. Chem.* **1995**, *99*:14632-14640.

[27] Ben-Naim A, Lovett R. *J. Phys. Chem. B* **1997**, *101*:10535-10541.

[28] Dill KA. *J. Biol. Chem.* **1997**, *272*:701-4.

[29] Martin MG, Zhuravlev ND, Chen B, Carr PW, and Siepmann JI. *J. Phys. Chem. B* **1999**, *103*:2977-2980.

[30] Makhatadze GI and Privalov PL. *J. Mol. Biol.* **1993**, *232*:639-659.

[31] Matubayasi N, Reed LH, and Levy RM. *J. Phys. Chem.* **1994**, *98*:10640-10649.

[32] Matubayasi N, Gallicchio E, and Levy RM. *J. Chem. Phys.* **1998**, *109*:4864-4872.

[33] Rashin AA and Bukatin MA. *J. Phys. Chem.* **1994**, *98*:386-389.

[34] Yu H-A, Roux B, and Karplus M. *J. Chem. Phys.* **1990**, *92*:5020-33.

[35] Lazaridis T and Paulaitis ME. *J. Phys. Chem.* **1993**, *97*:5789-5790.

[36] Hermann RB. *J. Comp. Chem.* **1993**, *14*:741-750.

[37] Sitkoff D, Sharp KA, and Honig B. *Biophysical Chemistry* **1994**, *51*:397-409.

[38] Simonson T and Brunger AT. *J. Phys. Chem.* **1994**, *98*:4683-4694.

[39] Amidon GL, Yalkowsky SH, and Leung S. *J. Pharm. Sci.* **1974**, *63*:1858-1866.

Chapter 4

Tests of Flory-Huggins Theory on Fluids of Sphere-like Molecules

Rachel Brem, Hue Sun Chan, and Ken A. Dill

4.1 Abstract.

The Flory-Huggins (FH) polymer theory can be used to predict solvation entropies in liquids. However, conventional wisdom states that FH's mean-field model is less accurate when applied to compact sphere-shaped molecules (*e.g.* cycloalkanes) than extended molecules (*e.g.* n-alkanes). Here we test this notion in simple models and in experimental data. We find that FH predicts qualitative trends in the experimental entropies of cyclic alkanes as well as in those of branched and linear shape, contrary to intuition from lattice models at constant volume. We also find, despite previous arguments to the contrary, that FH predicts trends in the iodine solvation data of Hildebrand and in disk models of liquids. For these predictions, we require an accurate estimate of the system volume change dV upon insertion of a solute at constant pressure. We conclude that FH can capture the physics of "compact" solvents under liquid-like conditions. In some cases, FH requires a more accurate estimate of dV than is traditionally used.

4.2 Introduction.

The Flory-Huggins (FH) polymer theory [1] is a classic theory of the liquid state. Traditionally, theorists have believed that FH theory is most accurate on extended polymer chains, and is not appropriate for "compact" sphere-like molecules [2, 3, 4, 5, 6, 7, 8, 9].

This dichotomy is of current interest in the biophysical literature that deals with oil-water solvation experiments. Such experiments measure the partitioning of a small molecule solute between liquid oil and liquid water. This is taken as a model of the partitioning of chemical groups on proteins or drugs, between water and the oily environments in proteins or biomembranes. Theories of the liquid state are sometimes needed to convert raw data from the small molecule experiment into a form useful to biomodelers. In particular, modelers often need to account for entropic differences between liquid oil and the oily environment of the biomolecule. FH theory has been proposed as a treatment of liquid oils for this purpose [6, 7,

8, 9, 10, 11, 12, 13]. However, conventional wisdom states that FH theory shouldn't be applied to some of these liquid oils-cycloalkanes or other compact shapes.

In this paper, we study the dependence of errors in FH theory on molecular size and shape, in several simple model systems as well as previously published experimental data. Our goal is to guide the use of FH theory for solvation modeling and other studies of the liquid state.

4.3 Computational and analytic methods.

4.3.1 A 2D lattice model of liquids.

For some analysis below, we use a 2D lattice model of liquids. In this model we perform Monte Carlo simulations in the NVT ensemble, on fluids of 2D lattice molecules interacting via a pairwise contact energy potential, as described previously [14]. Briefly, lattice segments of the molecules interact with an energy of $\varepsilon_{\text{lat}} = -1$ units when they sit in adjacent lattice sites, and otherwise their interaction energy is 0. See Figure 4.1 for cartoons of the multi-node lattice molecules we use here. We calculate thermodynamics of insertion with the Widom method [9] from simulations under liquid-like conditions [14], in which $T = 1.5\varepsilon_{\text{lat}}/k$ and the filling fraction $\phi = 0.7$.

4.3.2 Continuous models of liquids in 2D.

We also use a continuous-space liquid model, composed of 2D disk-shaped molecules interacting *via* the Lennard-Jones potential:

$$U(r_{ij}) = 4\varepsilon_{LJ}\left[\left(\frac{\sigma}{r_{ij}}\right)^{12} - \left(\frac{\sigma}{r_{ij}}\right)^{6}\right]$$

Here U is the interaction energy between two molecules i and j; r_{ij} is the distance between the centers of the molecules; ε_{LJ} is the energy of interaction between two molecules at the most favorable contact distance $r_{ij} = \sigma$. In all our studies here, $\varepsilon_{LJ} = -1.0$ unit.

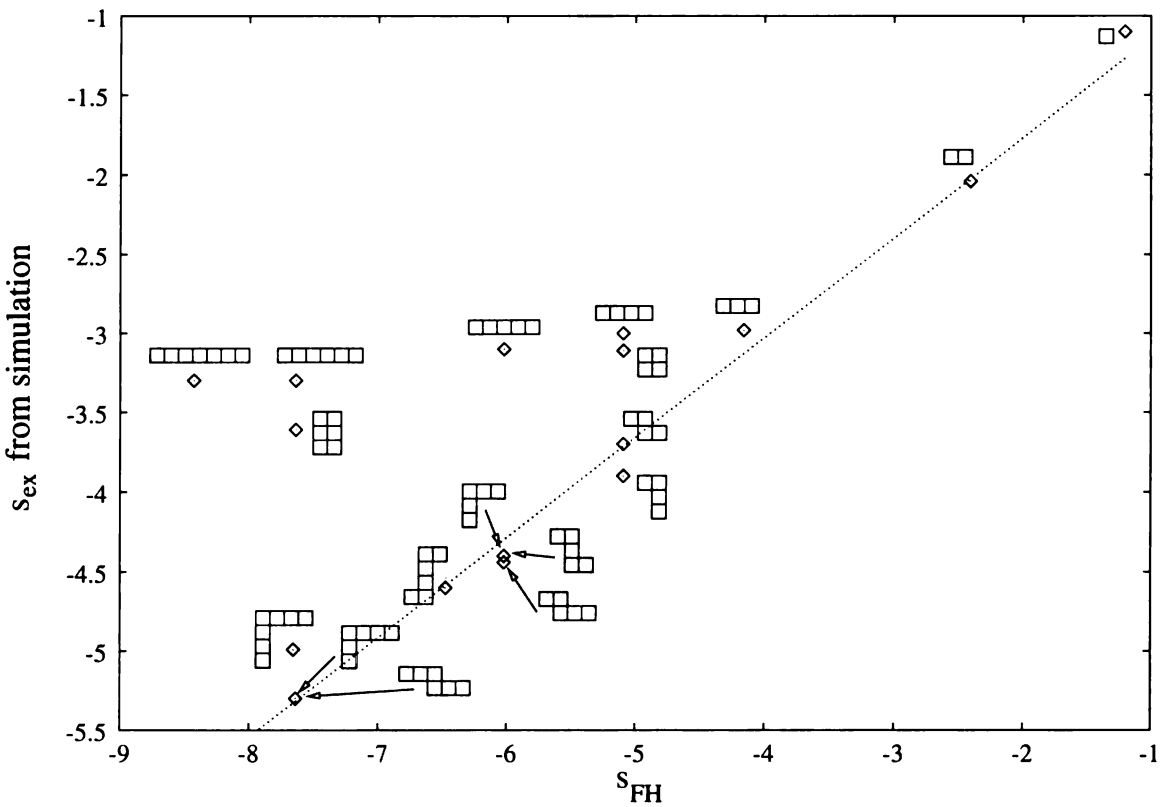


Figure 4.1: **Correlating FH theory with simulation data.** Every data point represents one molecular species. The x axis plots the predicted excess entropy of insertion of each solute into its own pure fluid at constant volume, calculated with Flory-Huggins theory, Equation 4.2. The y axis plots the true entropy of insertion from constant-volume Monte Carlo simulations at $T = 1.5\varepsilon_{lat}/k$. Dashed line indicates the best linear fit to all data except those from squares and rods, $y = 0.63x - 0.51$. Error bars on entropies are less than $0.05kT$.

We ran NVT and NPT Monte Carlo simulations in the liquid-like region of the 2D Lennard-Jones model [17, 18]: reduced temperature $T = 0.55\varepsilon_{LJ}$ and reduced pressure $p\sigma^2/\varepsilon_{LJ} = 0.6$ or reduced molar density $\rho = 0.755\sigma^2$. We also ran NVT and NPT simulations on hard disks in 2D at various densities and pressures. In all simulations, each system contained 100 solvent molecules, and we used periodic boundary conditions. We calculated thermodynamics of insertion of solutes with the Widom method [9].

4.3.3 FH theory.

In this paper we study Flory-Huggins (FH) theory, which was originally developed [1] for long polymers. FH theory treats the translational entropy and the conformational entropy of the chain, which are assumed to be coupled through excluded volume. The FH expression for the excess entropy of a solute (indicated by subscript 2) inserting into a solvent (indicated by subscript 1) at infinite dilution can be written [8]

$$s_{ex}/k = -\frac{1}{V}\frac{\partial V}{\partial N_2}N_1(m_1 - 1) - \frac{\partial v}{\partial N_2}\ln\frac{v}{V}. \tag{4.1}$$

Here V is the volume of the fluid before insertion of the solute, N_2 is the number of solute molecules, N_1 is the number of solvent molecules, m_1 is the solvent chain length (or molecular volume), and v is the amount of empty volume in the fluid, where $V = m_1N_1 + v$.

The form of FH one uses depends on the conditions of the experiment. If the solute is inserted at constant volume, then $\partial V/\partial N_2 = 0$ and $\partial v/\partial N_2 = -m_2$; this yields

$$s_{ex}/k = m_2\ln(1 - \phi) \tag{4.2}$$

where the filling fraction $\phi = m_1N_1/V$.

One can also approximate constant-pressure conditions within FH theory. At constant pressure, upon insertion of a solute, the solution changes volume. To use FH predictively, we must know how big the volume change is. If we assume the expansion is exactly m_2 volume units, there is no change in the free volume of the solution. Then $\partial V/\partial N_2 = m_2$ and $\partial v/\partial N_2 = 0$, which gives

$$s_{ex}/k = -\rho_1 m_2(m_1 - 1) \tag{4.3}$$

where ρ_1 is the molar density of the solvent.

In some calculations below we don't need to make any assumptions regarding the volume change of a solution upon insertion of a solute at constant pressure. Rather, we can measure this volume change directly from a constant-P experiment. Call the measured volume change

dV. In the notation of Equation 4.1 for this general case, $\partial V/\partial N_2 = dV$ and $\partial v/\partial N_2 = (dV - m_2)$. This gives

$$s_{ex}/k = -\rho_1(dV)(m_1 - 1) - (dV - m_2)\ln(1 - \phi). \tag{4.4}$$

4.4 Results.

4.4.1 Applying FH to a constant-V lattice model.

Our goal in this paper is to evaluate the conventional wisdom that FH theory predicts entropies of extended molecules with more accuracy than those of compact sphere-like shapes. In this section we address the question in a constant-volume 2-D lattice model of liquids.

For comparison with experimental data below, we study the entropies of solvation of lattice molecules into their own pure fluids. We estimate these entropies using Equation 4.2, then compare them to true entropies from Monte Carlo simulations of our lattice model. Figure 4.1 shows the result. FH predicted entropies are about a factor of 1.5 too negative, relative to the true values. Qualitatively speaking, FH estimates reproduce the correct trend among lattice shapes, except in the case of squares and liquid-crystal-like rod molecules, as has been seen in a previous lattice model [7]. Errors are more severe in fluids of larger molecules.

Why do rods and squares behave differently from other shapes? Molecules in a fluid of rods or squares can pack flush against one another. We illustrate this in the first four panels of Figure 4.2. Here efficient packing of rods and squares causes more microscopic density fluctuations than in other fluids, such as the ones in Figure 4.2(e) and (f). In rods and squares, instantaneous large voids form, allowing easier insertion of a solute. FH's mean-field model doesn't account for these voids, so FH predicts entropies to be less favorable than the true values. Figure 4.2 also gives an explanation for the dependence of FH errors on solvent size. Comparing across rows in Figure 4.2 shows that larger voids open in fluids of larger molecules,

because more of the matter is "bonded together" in the latter system. FH's mean-field model is better suited to fluids with smaller density fluctuations, so FH's predictions are more accurate on fluids of small molecules.

To see the effect of solute size and shape, we next do a simple control. We measure the insertion entropies of a 1-mer lattice square into each of many lattice solvents, then predict these entropies with FH (Equation 4.2) and evaluate the prediction. The prediction is always exactly correct within simulation error (data not shown). The excess entropy of insertion of a 1-mer in the NVT lattice model is always exactly $\ln(1-\phi)$. This is true by construction: on average over the entire simulation box, the fraction of open lattice sites is exactly $(1-\phi)$, no matter what the structure of the fluid. As we have seen, error in FH predictions arises when solutes are larger than one lattice unit, because FH neglects the fact that the occupancy of neighboring sites is correlated [1].

We conclude that at constant volume, FH predicts entropies less accurately in solvents of compact and liquid-crystal-like shapes. This is in keeping with conventional wisdom in the liquid theory literature [2, 3, 4, 5, 6, 7, 8, 9]. We also find that FH errors depend on solvent and solute size.

4.4.2 Applying FH to simple continuum models.

Next we apply FH to two more simple liquid-like models: 2-D hard disks and 2-D Lennard-Jones disks, both of which are described in the Computational and Analytic Methods section above. We observe entropies of insertion of hard disk solutes into each of these solvents, *via* Monte Carlo simulations. Then, as in the previous section, we compute predicted entropies with FH theory and ask how well they agree with true simulated values.

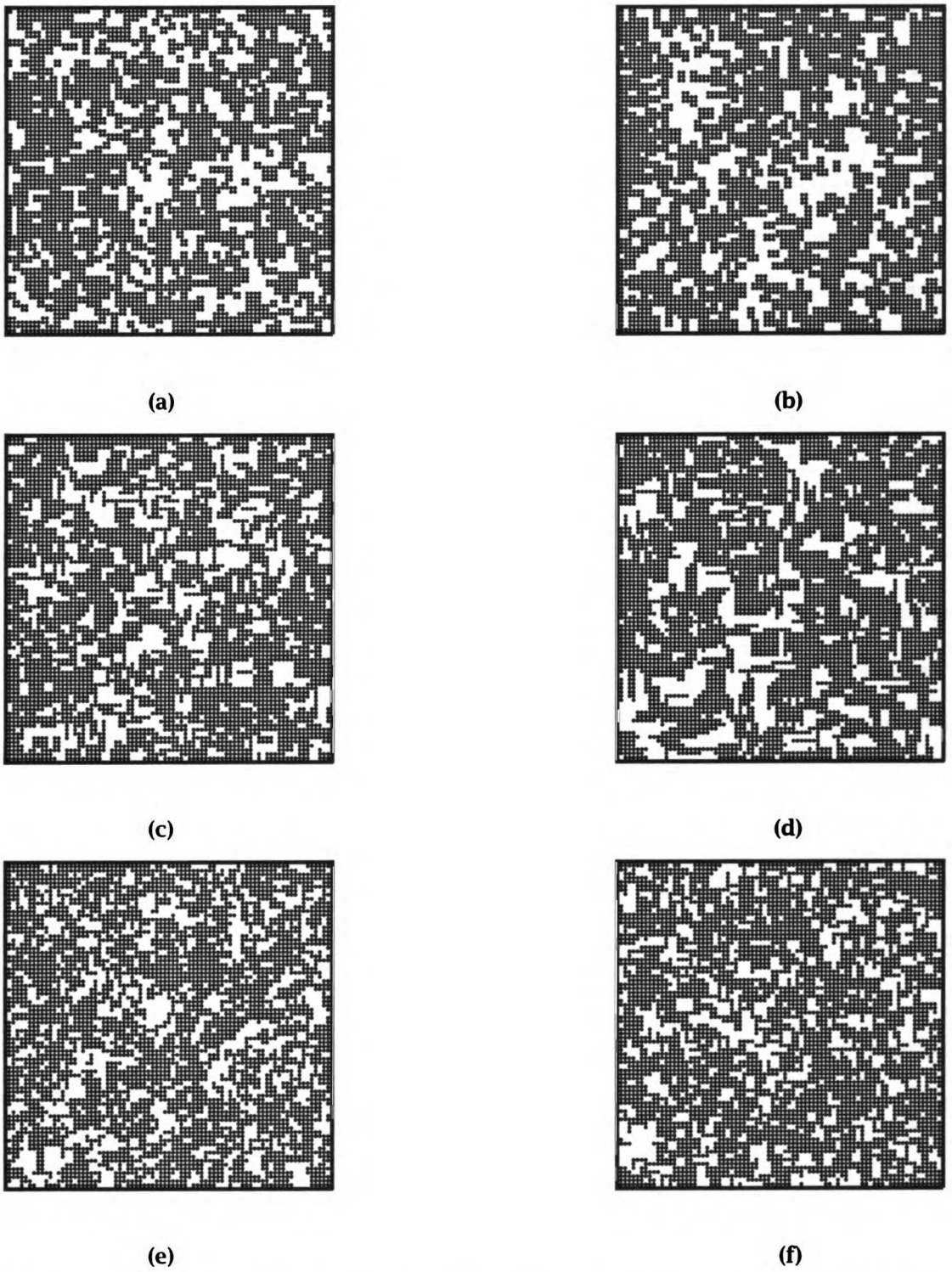


Figure 4.2: **Density fluctuations change with molecular shape and size.** Every dot on the figure represents a filled lattice site. Each panel shows a sample snapshot from the top left 80×80 lattice sites of ~ 250,000-site Monte Carlo simulation boxes, during simulations of homogeneous fluids. (a), molecules are 2×2 lattice squares; (b), 3×2 rectangles; (c), 4×1 rods; (d), 6×1 rods; (e), 4-mer L shapes; (f), 6-mer L shapes.

Applying FH to constant-volume continuum models.

Figure 4.3 compares predicted and simulated entropies under constant volume conditions. We use Equation 4.2 for FH predictions here. The bottom curve in Figure 4.3 gives results from hard disks. FH predicted entropies are about a factor of 10 too positive, relative to true simulation values. This is the opposite trend from the one we saw in the lattice model (see Figure 4.1), and we can understand it using the inverse of our reasoning above. Disk molecules don't pack efficiently. Much of the empty space in a fluid of disks isn't available for insertion of a solute. See Figure 4.4(a) for a snapshot of our hard disk fluid, and note that there isn't much room for a solute to insert, even though the simulation box is 40% empty. The FH model doesn't account for this "wasted space," so it predicts insertion to be more favorable than is actually the case.

In the top curve in Figure 4.3 we find that FH predictions are more accurate in solvents of Lennard-Jones disks. We believe this to be a result of a cancellation of errors. In fluids governed by distance-dependent potentials, solvents may order themselves at regular intervals to minimize energy. Introduction of a hard solute disrupts this order—a favorable entropic effect. As evidence for this idea, we show a snapshot of the Lennard-Jones fluid in Figure 4.4(b), in which the solvent ordering can be observed (compare to the clumps of solvents in Figure 4.4(a), especially at the top left). Thus, entropies are more favorable in Lennard-Jones systems than in hard disks, which cancels some of the FH error.

We concude that although FH accuracy suffers due to shape effects, accuracy improves due to solvent sticking energies. In our most liquid-like models at constant V, FH predicts entropies within about a factor of 3 and reproduces size trends correctly.

Applying FH to constant-pressure continuum models.

Next we measure solvation entropies in disk models at constant pressure, and compare them to FH predictions. Recall that to use FH at constant pressure we need an estimate of dV,

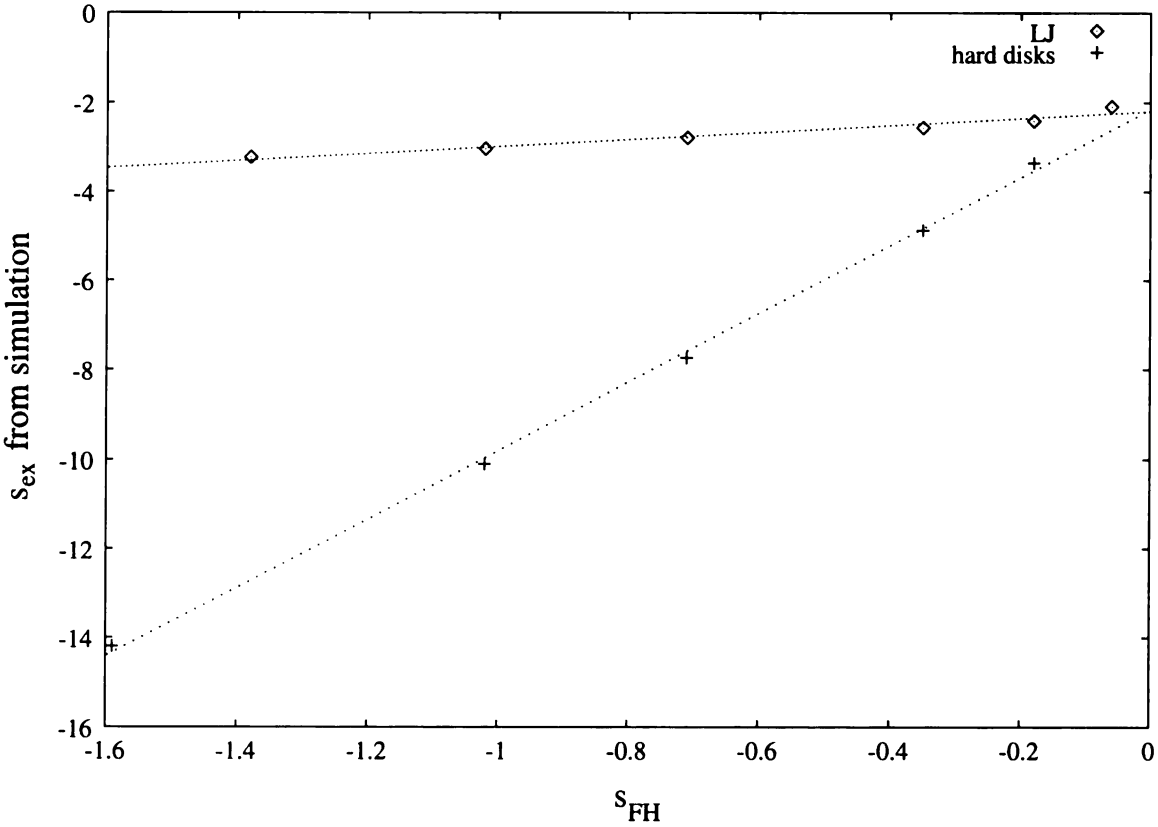
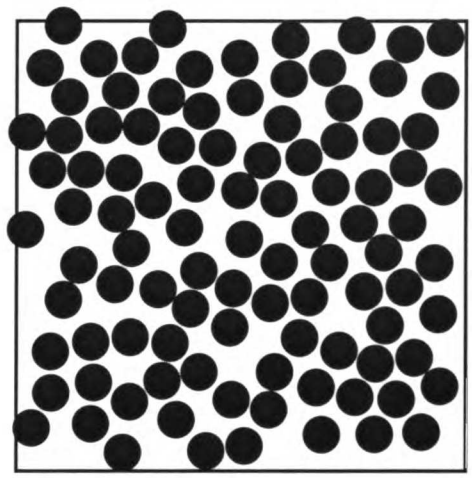


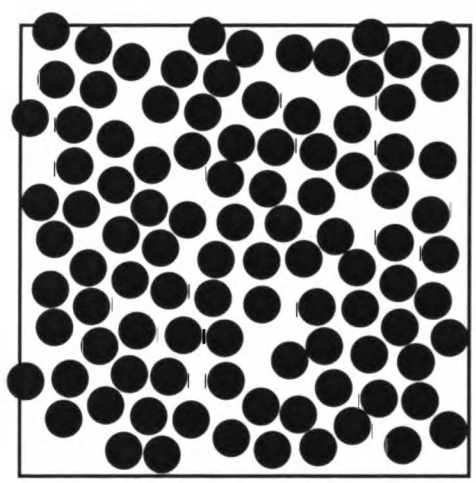
Figure 4.3: **Flory-Huggins predictions at constant V.** Each curve represents the solvation entropy of several hard disk solutes into the same solvent. Each data style represents a different liquid model. Dashed lines are best linear fits. "LJ" indicates Lennard-Jones solvent; "hard disks" indicates hard disk solvent. The x axis plots the Flory-Huggins constant-V excess entropy prediction, Equation 4.2. The y axis plots entropies of insertion from Monte Carlo simulations. Fit in top curve is $y = 0.80x - 2.20$ and in bottom curve, $y = 7.67x - 2.15$. Solute diameters are 0.5 to 1.5 units, and all solvent molecules have a diameter of 1.0 unit. Error bars on entropies are 0.2 or less. Hard disks at are at the phase point $\rho = 0.755\sigma^2$. Lennard-Jones disks are under liquid-like conditions (see text).

the volume change of the solvent upon insertion of the solute. We test two estimates here: (i) We assume dV is equal to the molecular volume of the solute, in which case FH takes the form in Equation 4.3. (ii) We measure the true dV from our simulations and substitute it into FH. Then the FH form we use is Equation 4.4.

Figure 4.5(a) shows the results of our applications of FH theory to solvation entropies in hard disk solvents at constant pressure. As in the constant-V case (see Figure 4.3), FH predictions are more positive than true simulation values. Comparing the right- and left-hand curves



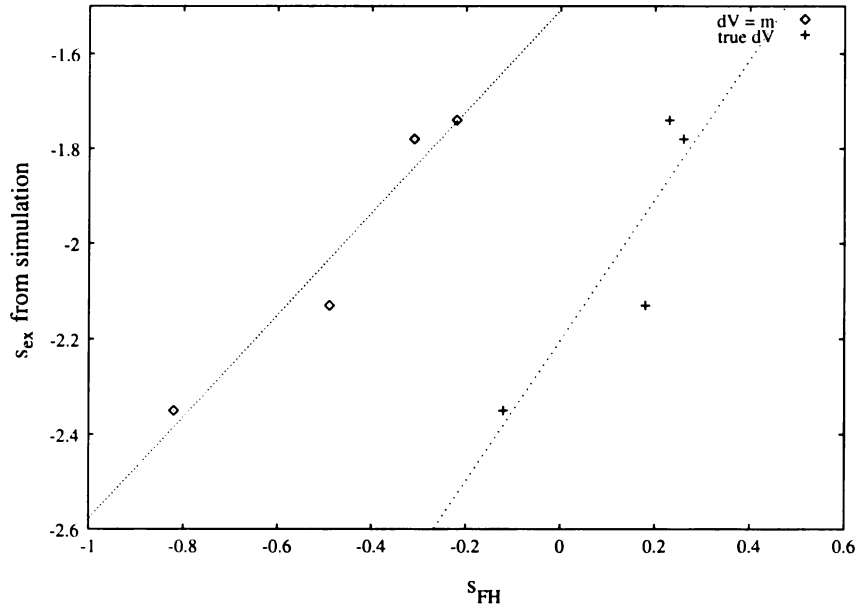
(a)



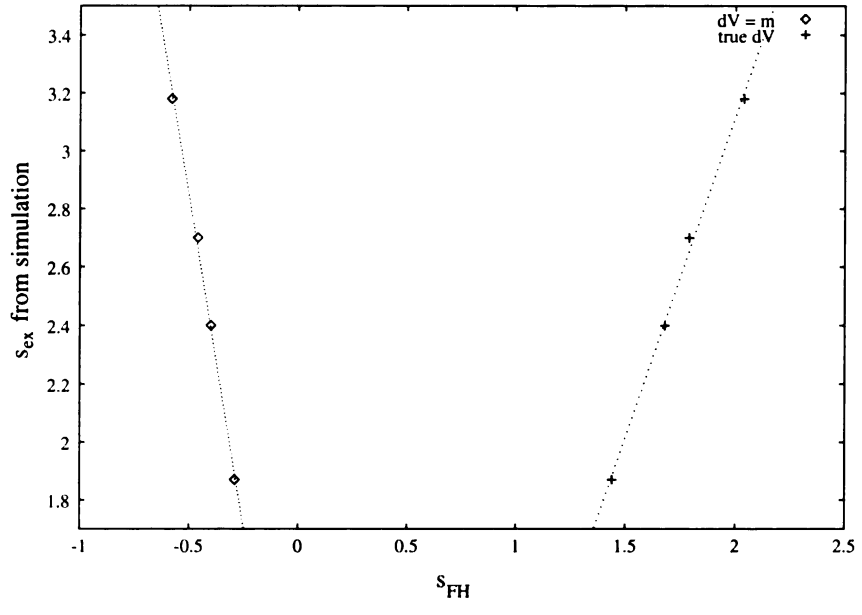
(b)

Figure 4.4: **Snapshots of 2D fluids.** Each panel represents a configuration randomly chosen from a converged constant-volume Monte Carlo trajectory. (a) Hard disks at $\rho = 0.755\sigma^2$. (b) Lennard-Jones disks under liquid-like conditions (see text).

of Figure 4.5(a) shows that the error is worse when we model dV exactly. We conclude, as above, that FH underestimates excluded volume effects in hard disks at constant pressure; disks pack more inefficiently than the mean-field model assumes, so true solute insertion is less favorable than the mean-field estimate.



(a)



(b)

Figure 4.5: **Flory-Huggins at constant P.** Each curve represents entropies of several hard disk solutes in the same solvent. Each panel gives data from one solvent model. x axes plot FH predictions, and y axes plot simulation data. "$dV = m$" indicates FH calculated with Equation 4.3; "true dV" indicates Equation 4.4 with dV's taken from simulations. Dashed lines are best linear fits. Solutes are σ = 1.2 to 1.7, and solvents are σ = 1.5. (a) Solvent is hard disks at an average filling fraction of 0.64. Fits are $y = 1.48x - 2.20$ (right-hand curve) and $y = 1.06x - 1.51$ (left-hand curve). Error bars on entropies are 0.2 or less. (b) Solvent is Lennard-Jones disks under liquid-like conditions (see text). Fits are $y = 2.20x - 1.29$ (right-hand curve) and $y = -4.53x + 0.58$ (left-hand curve). Error bars on entropies are 0.1 or less.

The story is different in Lennard-Jones solvents, as can be seen in Figure 4.5(b). Assuming that dV is equal to the solute molecular volume (the left-hand curve in Figure 4.5(b)) predicts entropies that have the wrong sign, and the wrong trend with solute size. But when we model dV exactly, FH predictions are correct within a factor of 1.5—more accurate than in the hard disk case. Why? Upon insertion of a solute, the system volume expands more in Lennard-Jones solvents than in hard disk solvents, and the expansion depends more on solute size in the former (data not shown). This large dV in Lennard-Jones solvents causes a large free volume contribution (the second term in Equation 4.4), which dominates the FH estimate. In other words, free volume effects dominate the magnitudes of entropies and their trends with solute size in Lennard-Jones solvents. When we model free volume exactly in FH, predictions are reasonably good, and other errors in FH contribute only secondarily.

We conclude that FH theory predicts entropies most accurately in our most liquid-like models. The traditional form of FH is not sufficient for this purpose; a good model of the system volume change dV, used as in Equation 4.4, is necessary.

4.4.3 Applying FH to experimental data.

In this section we compare FH predicted entropies to those measured in experimental systems. Since all experiments we study were done at constant pressure, we use Equations 4.3 and 4.4 to compute FH predictions. Where the molecular size m is called for in these equations we use molecular volumes in cm^3/mol, calculated as in [26]. These yield FH predictions with magnitudes about a factor of 10 too large. So we are unable to draw quantitative conclusions about FH in these systems; we can, however, study the trends of FH with size and shape.

Applying FH to alkane solvation.

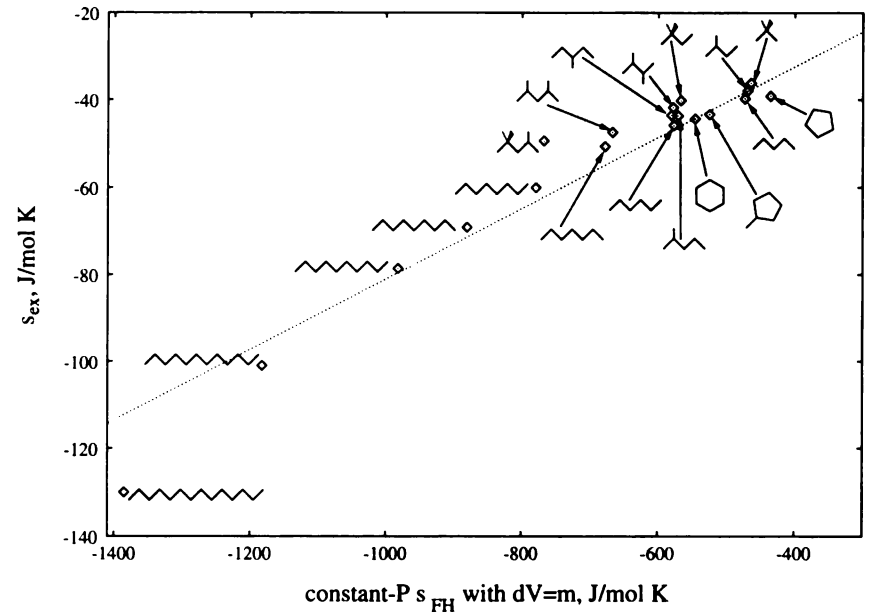
First we compare FH theory against the entropies of short-chain alkanes solvating in their own pure liquids, taken from [13]. The results are shown in Figure 4.6. For a given form

of FH the data from cyclic, branched, and linear alkanes fall near one line, with a y-intercept of zero. This runs counter to predictions from our lattice model (see Figure 4.1) and a previous lattice model [7], in which the accuracy of FH predictions depends on shape. Our alkane results also contradict the work of Chan and Dill [8, 9], who argued that FH should not be used on cyclic solvents. We conclude that with respect to shape, intuition from lattice models may not apply to the experimental case. We don't observe the accuracy of FH predicted entropies to depend on shape in alkane solutions.

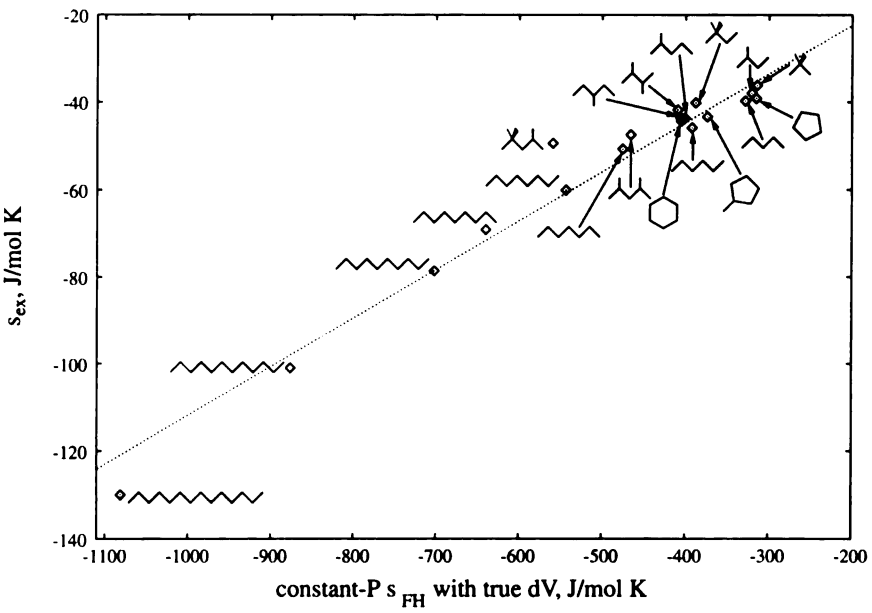
Note in Figure 4.6 that correlations are about as strong whether we use true exact system volume changes, or approximate volume changes, in our FH computation. This suggests that free volume effects don't dominate entropies in the alkane data set. Indeed, if we plot only FH's free volume term (the second term in Equation 4.4) against experimental entropies, the correlation is weak (data not shown). Thus, the correlations in Figure 4.6 are due to FH's treatment of excluded volume, which works equally well on all alkane shapes we study here.

Applying FH to iodine solvation.

Next we compare FH predictions to the entropies of iodine into "compact" solvents, taken from [2, 3]. Much of the conventional wisdom about FH's performance on sphere-like molecules has come from studies on this data set. Our results are shown in Figure 4.7. First we compute the FH estimate for entropy assuming that the system volume change dV is equal to the solute molecular volume (Equation 4.3). Figure 4.7(a) shows no correlation between theory and experiment, in keeping with previous observations [2, 3, 9]. Next we make FH predictions with Equation 4.4, using the true experimental dV's from [23]. Now we see a striking correlation (with a single outlier), as Figure 4.7(b) shows. This is reminiscent of our results from Lennard-Jones solvents in Figure 4.5(b): FH predictions using an assumed dV don't follow the right trends, but those using the true correct dV have the correct qualitative behavior. We conclude that in



(a)



(b)

Figure 4.6: **Flory-Huggins predictions correlate linearly with entropies of alkanes.** Each data point represents the entropy of one alkane in its own pure liquid. Each panel represents predictions from one form of FH. In both panels the y axis plots experimental entropies [13] corrected for solute translation in solvent. Dashed lines are best linear fits with zero intercept. (a) The x axis plots the FH prediction assuming that upon solute insertion, the system volume change dV is equal to the solute volume m (Equation 4.3). The linear fit is $y = 0.08x$. (b) The x axis plots the FH prediction using the true dV (Equation 4.4 with $dV = 1/\rho_1$). The linear fit is $y = 0.11x$. Molecular volumes are from [26] and alkane densities at 20°C are from [20].

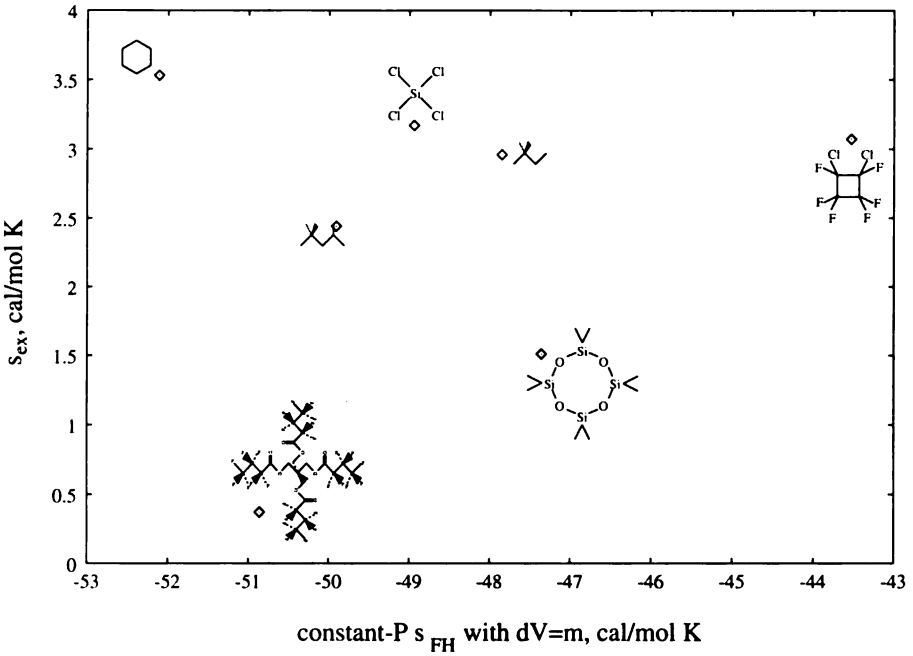
liquid-like models and experiments, FH can predict qualitative trends in entropies of compact molecules. For this successful prediction, FH requires an accurate estimate of dV.

4.5 Conclusions.

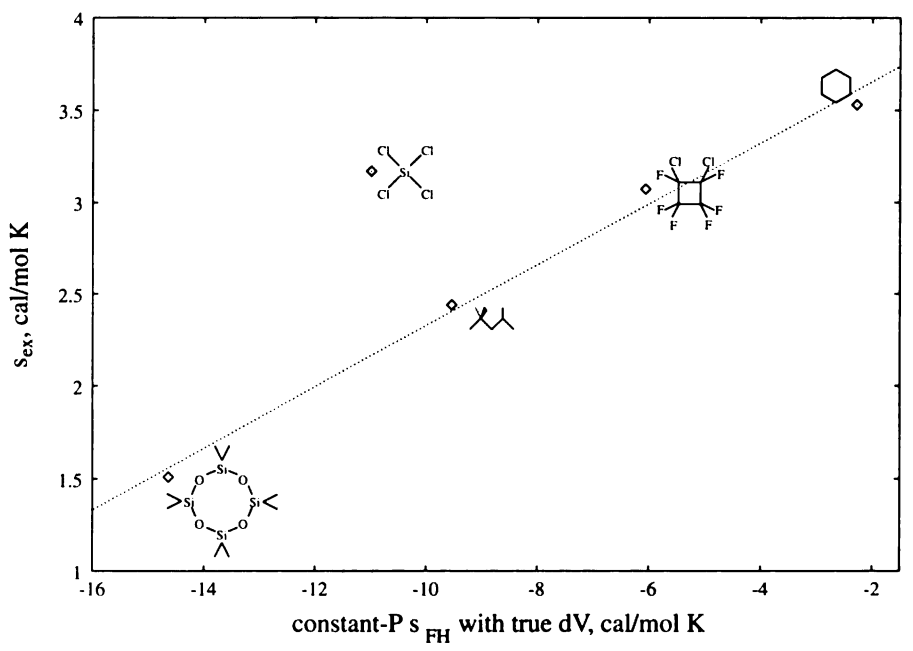
Our goal in this work has been to test the conventional wisdom that Flory-Huggins theory is not appropriate for "compact" sphere-shaped solvents. We have shown that in a 2-D lattice model of liquids of complex shape, the accuracy of FH entropy predictions depends on shape. But in an experimental data set of alkanes, FH performs equally well on cyclic molecules as on those of branched and linear shape. In 2-D disk models, we find that FH predictions are more accurate in Lennard-Jones solvents than hard solvents. FH predictions in Lennard-Jones solvents at constant pressure agree qualitatively with simulation data, if FH is given an accurate estimate of the system volume change dV upon insertion of a solute. This is borne out in an experimental data set of iodine solvation entropies into sphere-like solvents: FH predictions correlate with experimental entropies, if we use the true dV in FH computations. We conclude that FH can capture the physics of "compact" solvents under liquid-like conditions. In some cases, FH requires a more accurate estimate of dV than is traditionally used.

4.6 Acknowledgements.

We thank Kevin Silverstein for the use of his computer code and for technical advice. We also thank Jack Schonbrun, Noel Southall, and Karen Tang for helpful discussions. R.B. was a National Science Foundation predoctoral fellow and a University of California Regents' Fellow while this work was in progress.



(a)



(b)

Figure 4.7: **With knowledge of partial molar volumes, Flory-Huggins predictions correlate linearly with entropies of iodine.** Each data point represents the solvation entropy of iodine into one solvent. Each panel represents predictions from one form of FH. y axes plots experimental entropies [2, 3, 1] corrected for solute translation in solvent. (a) The x axis plots the FH prediction assuming that upon solute insertion, the system volume change dV is equal to the solute volume m (Equation 4.3). (b) The x axis plots the FH prediction using the true dV (Equation 4.4 with dV's taken from [23]). Dashed line is the best linear fit to all data except $SiCl_4$, $y = 0.17x + 3.99$. Molecular volumes are from [26] and densities are from [1]. Solvent structures are not to scale.

Bibliography

[1] Flory PJ. *Principles of Polymer Chemistry,* Cornell University Press, Ithaca, New York, 1953.

[2] Shinoda K and Hildebrand JH. The solubility and entropy of solution of iodine in *n*-C_7F_16, *c*-$C_6F_{11}CF_3$, $(C_3F_7COOCH_2)_4C$, *c*-$C_4Cl_2F_6$, CCl_2FCClF_2 and $CHBr_3$. 1958. *J. Phys. Chem. 62*:292-294.

[3] Shinoda K and Hildebrand JH. The solubility and entropy of solution of iodine in octamethylcyclotetrasiloxane and tetraethoxysilane. 1957. *J. Phys. Chem. 61*:789-791.

[4] Chiou CT, Porter PE, and Schmedding DW. Partition equilibria of nonionic organic compounds between soil organic matter and water. 1983. *Environ. Sci. Technol. 17(4)*:227-231.

[5] Chiou CT. Parition coefficients of organic compounds in lipid-water systems and correlations with fish bioconcentration factors. 1985. *Environ. Sci. Technol. 19(1)*:57-62.

[6] De Young LR and Dill KA. Partitioning of nonpolar solutes into bilayers and amorphous *n*-alkanes. 1990. *J. Phys. Chem. 94*:801-809.

[7] Krukowski AE, Chan HS, and Dill KA. An exact lattice model of complex solutions: chemical potentials depend on solute and solvent shape. 1995. *J. Chem. Phys. 103(24)*:10675-10688.

[8] Chan HS and Dill KA. Solvation: effects of molecular size and shape. 1994. *J. Chem. Phys. 101(8)*:7007-7026.

[9] Chan HS and Dill KA. Solvation: how to obtain microscopic energies from partitioning and solvation epxeriments. 1997. *Annu. Rev. Biophys. Biomol. Struct. 26*:425-59.

[10] Sharp KA, Nicholls A, Fine RF, and Honig B. Reconciling the magnitude of the microscopic and macroscopic hydrophobic effects. 1991. *Science 252*:106-109.

[11] Sharp KA, Nicholls A, Friedman R, and Honig B. Extracting hydrophobic free energies from experimental data: relationship to protein folding and theoretical models. *Biochemistry 30*:9686-97.

[12] Kumar SK, Szleifer I, Sharp K, Rossky PJ, Friedman R, and Honig B. Size dependence of transfer free energies. 1. A Flory-Huggins approach. 1995. *J. Phys. Chem. 99*:8382-8391.

[13] Sharp KA, Kumar S, Rossky PJ, Friedman RA, and Honig B. Size dependence of transfer free energies. 2. Hard sphere models. 1996. *J. Phys. Chem. 100*:14166-14177.

[14] Brem R, Chan HS, and Dill KA. Extracting microscopic energies from oil-phase solvation experiments. 2000. *J. Phys. Chem. B 104(31)*:7471-7482.

[15] Widom B. Some topics in the theory of fluids. 1963. *J. Chem. Phys. 39*:2808-2812.

[16] Tang KES and Bloomfield VA. Excluded volume in solvation: sensitivity of scaled-particle theory to solvent size and density. 2000. *Biophys. J.*, in press.

[17] Phillips JM, Bruch LW, and Murphy RD. The two-dimensional Lennard-Jones system: Sublimation, vaporization, and melting. 1981. *J. Chem. Phys. 75(10)*:5097-5109.

[18] Mitus AC, Patashinskii AZ, and Sokolowski S. Local structures in a computer-generated liquid: two-dimensional Lennard-Jones liquid. 1991. *Physica A 174*:244-271.

[19] Bader RFW, Carroll MT, Cheeseman JR, and Chang C. Properties of atoms in molecules: atomic volumes. 1987. *J. Am. Chem. Soc. 109*:7968-7979.

[20] Ben-Naim A. 1987. *Solvation Thermodynamics.* Plenum Press, New York.

[21] Martin MG, Zhuravlev ND, Chen B, Carr PW, and Siepmann JI. Origins of the solvent chain-length dependence of Gibbs free energies of transfer. 1999. *J. Phys. Chem. B 103*:2977-2980.

[22] Hildebrand JH, Scott RL. *Regular Solutions.* Prentice-Hall, Englewood Cliffs, NJ, 1962.

[23] Shinoda K and Hildebrand JH. Partial molal volumes of iodine in various complexing and non-complexing solvents. 1958. *J. Phys. Chem. 62*:295-296.

[24] Hirschfelder JO, Curtiss CF, Bird RB. *Molecular Theory of Gases and Liquids.* Wiley and Sons, New York, 1964.

[25] Jorgensen WL, Madura JD, and Swenson, CJ. Optimized intermolecular potential functions for liquid alkanes. 1984. *J. Am. Chem. Soc. 106(22)*:6638-46.

[26] Abraham MH and McGowan JC. The use of characteristic volumes to measure cavity terms in reversed phase liquid chromatography. 1987. *Chromatographia 23(4)*:243-246.

[27] Abraham MH, Chadha HS, Whiting GS, and Mitchell RC. Hydrogen bonding. 32. An analysis of water-octanol and water-alkane partitioning and the $\Delta \log P$ parameter of Seiler. 1994. *J. Pharm. Sci. 83(8)*:1085-1100.

[28] Lide DR, Editor. *CRC Handbook of Chemistry and Physics.* CRC Press, Boca Raton, FL, 2000.

Chapter 5

The Physical Basis for the Success of Mole Fraction Theory in Sphere-like Molecules

Rachel Brem, Hue Sun Chan, and Ken A. Dill

5.1 Abstract.

How do solvation entropies depend on solvent size? The question is of current interest in the water theory community; it was also addressed in the last century by Hildebrand and co-workers using mole-fraction theory. Modern liquid state theorists don't often use mole-fraction theory, because of its assumption that solute and solvent be the same size. But its success in describing Hildebrand's series of data, in which the solute remains fixed while the solvent changes, has never been fully explained. Here we ask why and how mole-fraction theory applies to sphere-like molecules. We re-formulate the theory in terms of excess quantities, then test it in a hard disk model. Like Hildebrand's entropies, hard disk solvation entropies follow our mole-fraction-like expression; we believe the reason relates to ordering effects among solvents close to the solute. Predictions from scaled particle theory also are in approximate agreement with the Hildebrand expression. We conclude that the mole-fraction expression may describe a general property of entropies of sphere-like molecules.

5.2 Introduction.

The classic experiments of Hildebrand and co-workers [1] are one of the cornerstones of modern physical chemistry. These experiments measured partial molar entropies of solvation of iodine into "compact" sphere-like solvents of different size. Hildebrand *et al.* were especially interested in solvents in which the iodine had a violet color, *i.e.* those solvents which did not interact specifically with the iodine. Entropies of these violet solutions were seen to be equal to the quantity $-R \ln x + C$, where R is the gas constant, x is the mole fraction of the solute in solution, and C is a constant.

Hildebrand and co-workers interpreted this success as support for the simple mole-fraction theory of solvation [1]. The latter is based on a lattice model of solute and solvent species that are the same size. Precisely because of this assumption, mole-fraction theory has

fallen out of favor among liquid state theorists [2, 3]. But its original success on the iodine solvation experiments has never been fully rationalized. This question may be of broader interest to modern hydrophobicity modelers [4, 5, 6] who, like Hildebrand, study the effects of solvent size on solvation thermodynamics. In an attempt to explain the physical basis for Hildebrand's success, Sharp *et al.* applied a hard sphere model to the iodine data [7]. However, the accuracy of their conclusions was questioned by Chan and Dill [2].

In this work, we ask whether Hildebrand's results are due to general properties of compact molecules. First we re-formulate Hildebrand's relation in terms of excess quantities. Then we perform a Hildebrand analysis on the entropies of 2-D hard disks, in Monte Carlo simulations and in scaled particle theory (SPT). We also analyze Hildebrand's original experimental data in terms of SPT.

5.3 Computational and analytic methods.

5.3.1 The toy lattice model.

For some of our work below we use a toy 2-D lattice model of molecular fluids, developed by Krukowski *et al.* [8]. In this model each molecule is represented by a collection of fused lattice nodes, and allowed to occupy positions across a grid of fixed volume, the "simulation box." No two lattice nodes may share a grid position, ensuring that excluded volume is not violated. There are no other energetic interactions between molecules. Several lattice molecules placed on the same grid composes a "fluid" in the model. The density of states of the fluid is calculated by exactly enumerating all possible configurations of the molecules in the simulation box, *via* the algorithm in [8].

5.3.2 The hard disk model.

We also make use of a 2-D model of hard disks. We run Monte Carlo simulations on systems of 100 solvent disks whose only energetic interaction is that of excluded volume. All simulations are under periodic boundary conditions, at constant hard-disk pressures as specified below. From equilibrated simulations we calculate thermodynamics of insertion of hard disk solutes *via* the Widom method [9].

5.3.3 Scaled particle theory.

As an analytic treatment for our hard sphere and disk models, we use scaled particle theory (SPT). This theory predicts the excess chemical potential of insertion of a cavity of diameter d_2 in a solvent of spheres of diameter d_1 to be [10]

$$\mu_{ex} = -RT\ln(1-y) + \frac{9}{2}RT\left[q^2 + \frac{2q}{3}\right]f_R^2 + 3RTqf_R + \frac{p\pi d_1^3}{6},$$

where R is the gas constant, T the temperature, p the pressure, $y \equiv \pi d_1^3\rho_1/6$ is the packing fraction, ρ_1 is the molar density of the solvent, $f_R \equiv d_2/d_1$, and $q \equiv y/(1-y)$. Using the fundamental thermodynamic identity $s_{ex} = (\partial\mu/\partial T)_P$ to define the excess partial molar entropy of insertion s_{ex}, SPT gives the following prediction:

$$\begin{aligned} s_{ex} &= -R\ln(1-y) - RTq\alpha + \frac{9}{2}Rf_R^2q^2 + 9RTf_R^2\left[-q^2\alpha - q^3\alpha\right] \\ &+ 3Rf_R^2q - 3RTf_R^2\left[q^2\alpha + q\alpha\right] + 3Rf_Rq + 3RTf_R\left[-q\alpha - q^2\alpha\right], \end{aligned} \tag{5.1}$$

where $\alpha \equiv (1/V)(\partial V/\partial T)_P$, the coefficient of thermal expansion of the solvent, for a system volume V.

In two dimensions, the SPT chemical potential is [11, 12]

$$\mu_{ex} = -RT\ln(1-y) + 2RTf_Rq + \frac{p\pi d_1^2}{4}$$

for $y = \pi d_1^2\rho_1/4$. This gives for the excess entropy

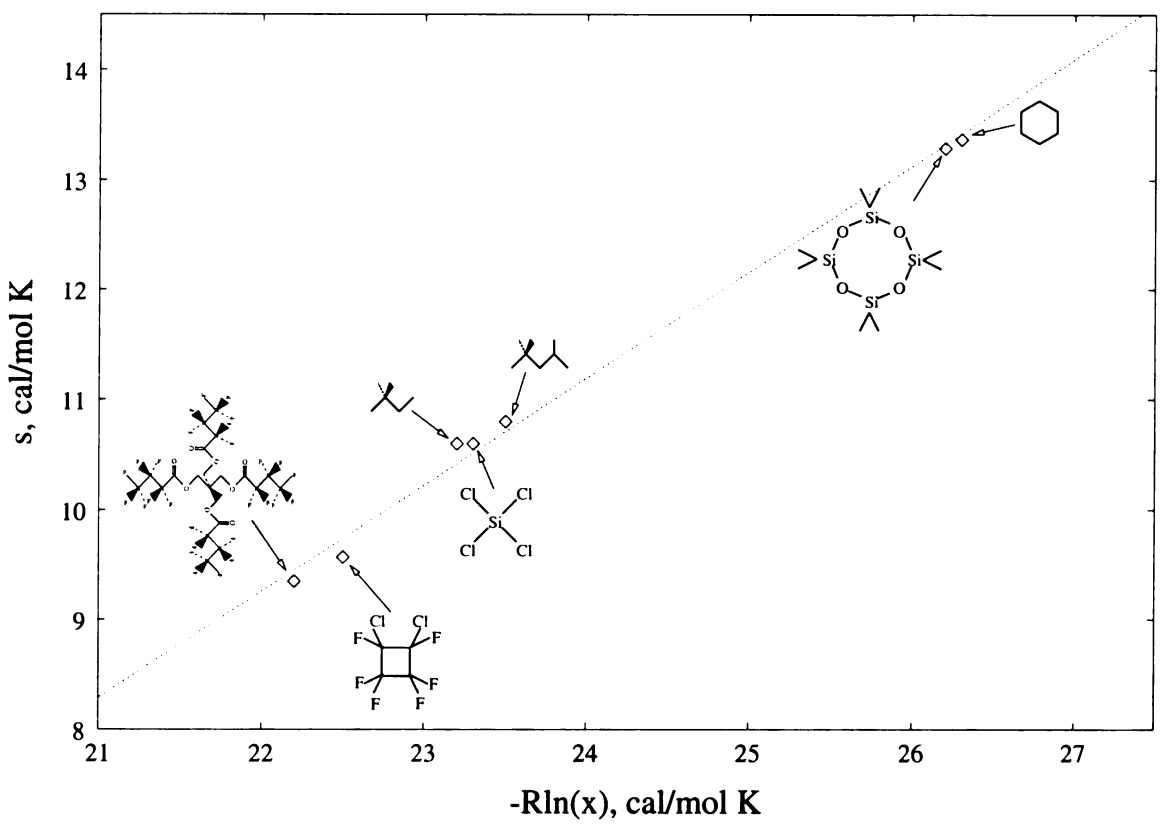


Figure 5.1: **The mole-fraction expression fits iodine solvation data.** Every data point represents the entropy of insertion of iodine into one solvent. The x axis plots $R \ln x_2$, where x_2 is the mole fraction of the solute, $x_2 \equiv N_2/(N_1 + N_2)$ for N_1 moles of solvent and N_2 moles of solute. The y axis plots experimental entropies of insertion. The dashed line is the best linear fit to the data, $y = 0.97x - 12.04$. Data are from [18, 19]. Solvent structures are not to scale.

$$s_{ex} = -R\ln(1-y) - RTq\alpha + Rf_R\left[2q - 2Tq\alpha - 2Tq^2\alpha\right] \tag{5.2}$$

given $\alpha \equiv (1/A)(\partial A/\partial T)_P$ for a system area A.

5.4 Results.

5.4.1 Re-expressing the Hildebrand relation.

Figure 5.1 shows Hildebrand's original entropies of insertion of iodine into "compact" solvents, as these data were published in 1947. Since the slope of Figure 5.1 is approximately

unity, we can express the relation between the ordinate and abcissa as

$$s = -R \ln x_2 + C$$

where x_2 is the mole fraction of the solute in solution and C is a constant; when fit to Hildebrand's data, $C = 12.4$ e.u. If we define ρ_2 to be the density of the solute in solution, then adding $R \ln \rho_2$ to both sides yields

$$s + R \ln \rho_2 = -R \ln(x_2/\rho_2) + C. \tag{5.3}$$

Now by the definition of the molar fraction and density,

$$x_2/\rho_2 = \frac{V}{N_1 + N_2}$$

where V is the volume of the solution, N_1 is the number of moles of solvent, and N_2 is the number of moles of solute. If we assume very low density of the solute ($N_2 \ll N_1$), as is applicable in Hildebrand's data, we have

$$x_2/\rho_2 = 1/\rho_1$$

for a solvent density ρ_1. When substituted into Equation 5.3, this gives

$$s + R \ln \rho_2 = R \ln \rho_1 + C. \tag{5.4}$$

Here's how to interpret the left-hand side of Equation 5.4 for Hildebrand's iodine experiments. First, define the entropy of the initial state s_{init}, which in Hildebrand's case is the iodine solute in solid iodine. Also define the entropy of the final state s_{fin}, which is the iodine solute in liquid solvent. The experimentally determined partial molar entropy s is then $s = s_{fin} - s_{init}$. Now add $R \ln \rho_2$ to the liquid state entropy to give s_{ex}^{fin}, the excess entropy in the liquid phase:

$$s_{ex}^{fin} = s^{fin} + R \ln \rho_2$$

s_{ex}^{fin} is believed to give the entropy of solvent reorganization upon insertion of a solute at a fixed position in space [13]. Ordinarily one would next do a similar correction for the initial

state; we choose to leave Hildebrand's initial state entropies uncorrected. Then the entropy we ultimately use for analysis is the difference between the final, corrected entropy of the iodine in the liquid phase and the intitial, uncorrected iodine in the solid phase. For convenience we call this quantity s_{ex}:

$$\begin{aligned} s_{ex} &= (s^{fin} + R\ln\rho_2) - s^{init} \\ &= (s^{fin} - s^{init}) + R\ln\rho_2 \\ &= s + R\ln\rho_2. \end{aligned}$$

s_{ex} contains solvent reorganization entropies in the liquid (final) state, plus an invariant additive term from the solid (initial) state. Substituting the definition of s_{ex} into Equation 5.4, we find that for Hildebrand's iodine experiments,

$$s_{ex} = R\ln\rho_1 + C. \tag{5.5}$$

Equation 5.5 expresses the mole fraction relation in terms of excess entropies, and is the major result of this section.

From the preceding derivation we predict that a plot of Equation 5.5, when computed with Hildebrand's raw data, should have the same properties as Hildebrand's original plot. Figure 5.2 bears out this prediction: its slope and intercept are roughly the same as in Figure 5.1. Note in Figure 5.2 that s_{ex} is more favorable for insertion of iodine into smaller solvents. This trend, from analysis of s_{ex}, contradicts Hildebrand's original observation [1] based on s, that "the wide range in molal volumes has no effect whatever" on iodine solvation in these data. Thus studying the sum of solute and solvent entropies in the total s, as Hildebrand did, obscures the dependence of solvation behavior on solvent size.

Also, note in Figure 5.2 that s_{ex} of Hildebrand's data are positive, which suggests that the iodine solute causes ordering of solvent molecules in the liquid less than it orders other iodine molecules in the solid. The mechanism for this is not understood.

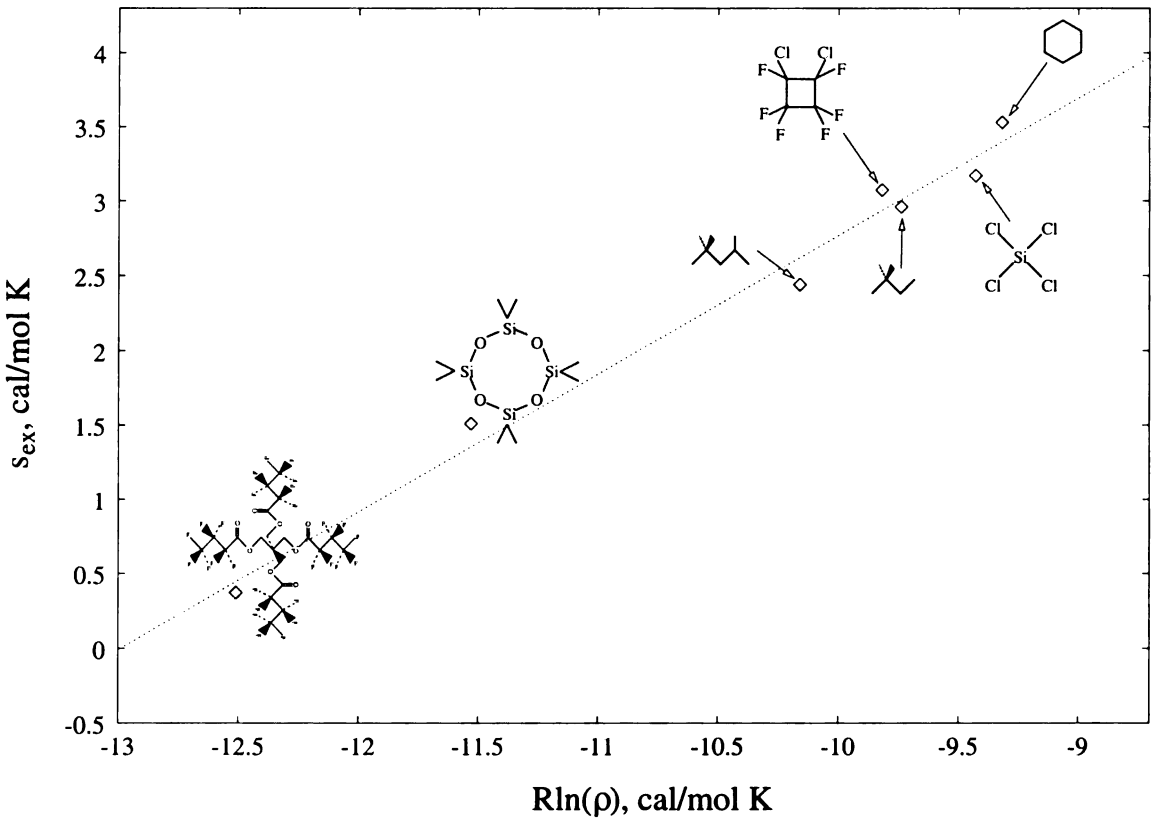


Figure 5.2: **An equivalent treatment of Hildebrand's mole fraction data.** Each data point represents the entropy of insertion of iodine into one solvent. The x axis plots $R \ln \rho$, where ρ is the molar density of the solvent at 1 atm. The y axis plots the difference between the experimental entropy of insertion of the iodine and the quantity $R \ln \rho_2$, where ρ_2 is the molar density of the iodine in solution. The dashed line is the best linear fit to the data, $y = 0.93x - 12.03$. Entropies are from [18, 19] and densities are from [1, 20]. Solvent structures are not to scale.

5.4.2 Fluids of different sized molecules at the same pressure.

In this work we aim to study Hildebrand's experiments in simple models. For this we need to compare several fluids of different-sized molecules. Recent papers have done this sort of comparison by fixing packing fractions [14] or volumes [5] to be the same across a series of solvent sizes. We are interested here in comparing fluids at the same pressure, as Hildebrand did. What is the difference?

We first address this question in a toy lattice model of molecular liquids [8]. See the Computational and Analytic Methods section for a brief description of the model. The advantage of this model is that we can calculate a fluid's entire density of states Ω exactly. Given Ω of a 2-D

lattice liquid at constant volume, we then study the pressure P through the following relation:

$$\begin{aligned} P/kT &= -\frac{1}{kT}\left(\frac{\partial F}{\partial V}\right)_T \\ &= \frac{\partial \ln \Omega}{\partial V}. \end{aligned} \tag{5.6}$$

Here k is Boltzmann's constant, T is the temperature of the lattice fluid, F is the Helmholtz free energy F, and V is the box volume occupied by the lattice solvent molecules. Equation 5.6 tells us that the slope of a plot of $\ln \Omega$ *versus* V will give an estimate of the pressure in the toy lattice model. We show such a plot in Figure 5.3 for fluids of several "sphere-like" lattice molecules.

To visualize the condition of constant packing fraction in Figure 5.3, note the arrows in this figure; they indicate the location of the phase point where the filling fraction is 0.3. At each arrow, the slope of the curve—that is, P/kT—is different. Pressures are higher in fluids of small molecules. In these fluids, changing the box volume by a small amount, say dV, causes a steep increase in the number of possible configurations of the system. We believe this is because the length scale for molecular displacement is small: centers of the solvent molecules can "squeeze by" each other to adopt many new positions given a small dV. Centers of the larger solvent molecules shift within a smaller range given a small dV, so the increase in the density of states with dV is not so steep.

To visualize the condition of constant volume (or density) in Figure 5.3, compare the curves from top to bottom at one position along the x axis. Slopes of fluids of large molecules are steeper than those of small molecules at the same density. We believe this is because at a given density, the packing fraction is higher in fluids of large molecules. More configurations violate excluded volume in these fluids, so more new configurations are added to Ω as we increase the system volume by dV.

Now visualize the condition of constant pressure in Figure 5.3 by finding phase points at which the slope of the curve is about the same. At any such phase points, fluids of small molecules will have low filling fractions, and fluids of large molecules will have high filling

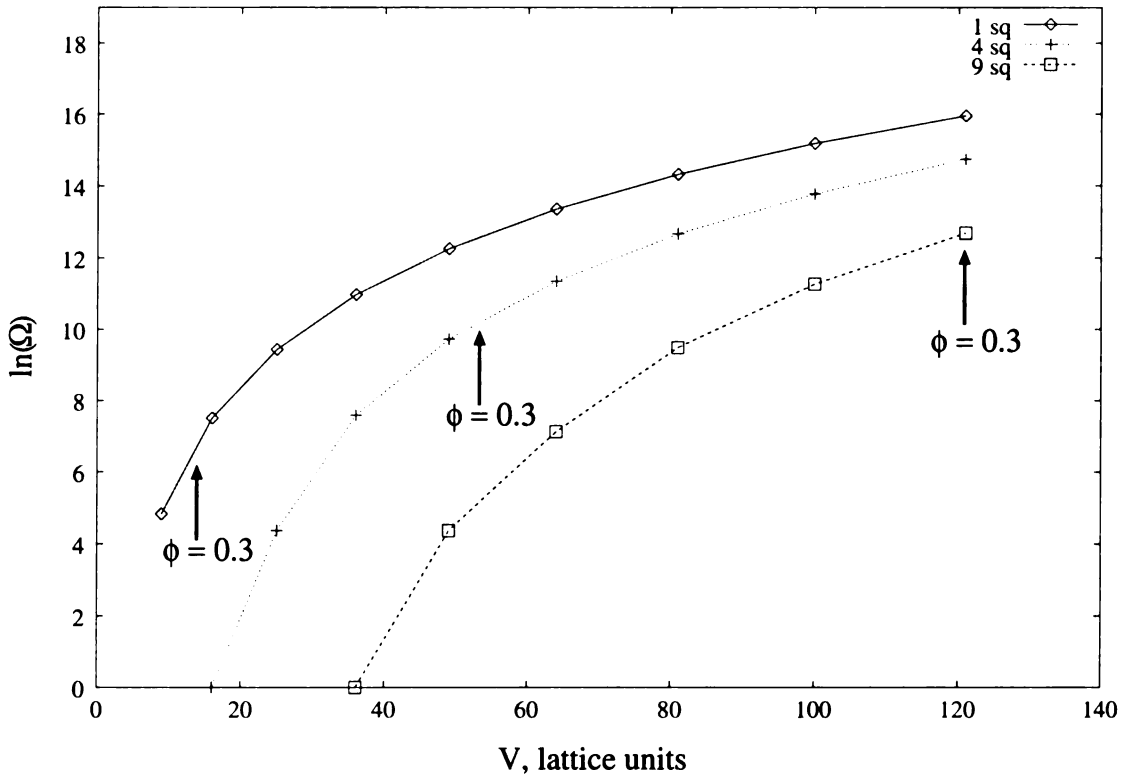


Figure 5.3: **Constant pressure, density, and packing fractions in lattice fluids.** Each data point represents one phase point in a 2-D lattice fluid. Each data style represents one lattice molecular species. In the legend, "1 sq" indicates a square composed of 1 lattice unit, "4 sq" indicates a square composed of 4 linked lattice units, and "9sq" indicates a square of 9 linked lattice units; for details, see [8]. The x axis plots the simulation box volume in lattice units. The y axis plots the natural logarithm of the density of states of four solvent molecules in the volume V, calculated as in [8]. Arrows indicate the volume at which the packing fraction $\phi = 0.3$.

fractions. This trend is approximately borne out in organic liquids [4, 6] and in Hildebrand's solvents (data not shown).

We conclude that in a series of fluids of different sized molecules, the condition of constant pressure is qualitatively different from those of constant packing fraction [14] or volume [5]. At the same pressure, fluids of larger molecules are more tightly packed than fluids of small molecules.

5.4.3 Hildebrand's relation in hard disks.

Next we introduce a continuum model—that of hard disks. Though we can't calculate $\ln\Omega$ directly in this model, we can simulate large systems and calculate thermodynamic quantities with the Monte Carlo method. The advantage here is that we can simulate a series of solvents at constant pressure. Intuition from the lattice model above suggests that at the same pressure, fluids of large disks should have a higher packing fraction than fluids of small disks. Figure 5.4 confirms this prediction. Thus, to mimic Hildebrand's iodine solvation experiments we insert the same solute disk into different disk solvents at the same pressure, as shown in Figure 5.4; these fluids have different packing fractions, dependent on solvent size.

With the hard disk model we aim to study the Hildebrand relation, Equation 5.5. To test this equation in our model, we first calculate excess entropies of solute insertion from Monte Carlo simulations. Then we plot these entropies against the natural logarithm of the densities attained by the simulations at constant pressure, and measure the slope. The result is a Hildebrand plot for the hard disk model, shown in Figure 5.5(a). Under the conditions we have examined so far, error bars are sizable, but the data appear to follow Equation 5.5: as in the original iodine experiments, our slopes are all within simulation error of 1.0. Also, as in the iodine data, entropies are less favorable in fluids of large molecules than in small molecules. Thus, 2-D hard disks reproduce the qualitative behavior of Hildebrand's iodine solutions.

Why does Equation 5.5 fit our model entropies? To find out whether free volume effects are part of the answer, we did the following test. We ran Monte Carlo simulations of fluids of different sized solvent disks at constant volume. We fixed densities of the solvents to be those the fluids would have attained at constant pressure. Into each of these fluids we inserted the same solute, without letting the simulation box expand, and measured the entropy of solvation. Our Hildebrand analysis of these data is shown in Figure 5.5(b). Entropies of solvation at constant volume are in qualitative agreement with Equation 5.5. Recall that the same is true at constant pressure (see Figure 5.5(a)), where the simulation box can expand as the solute enters.

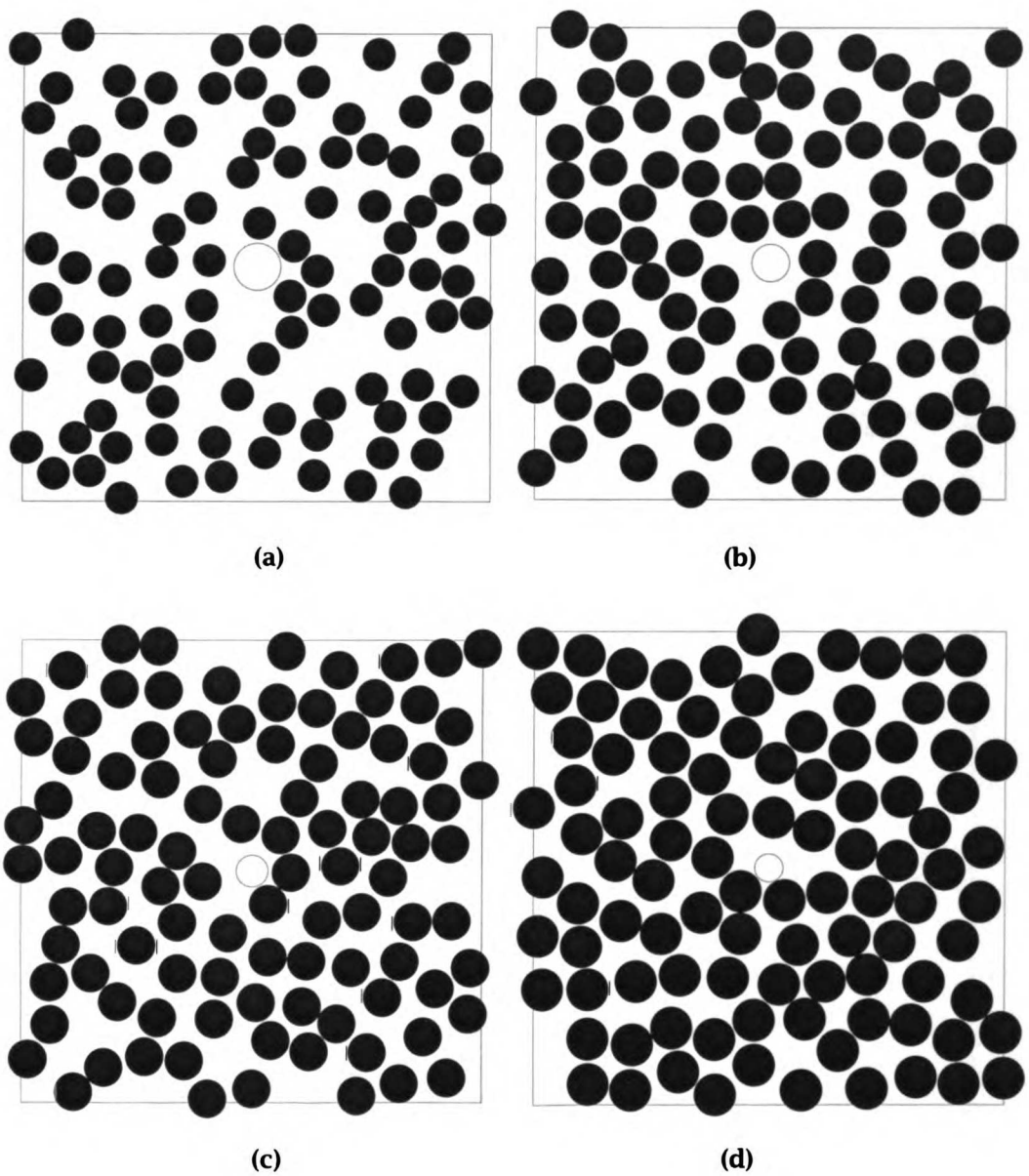
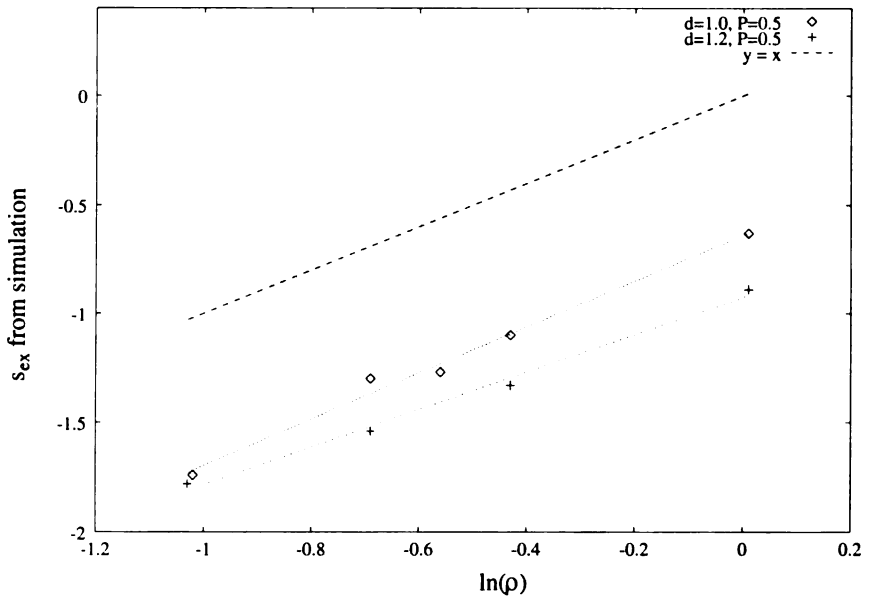
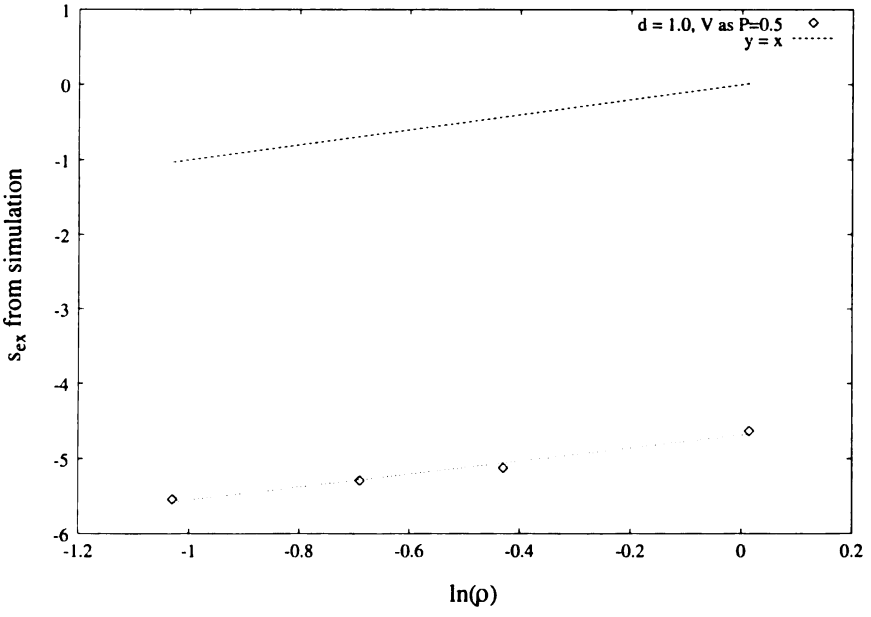


Figure 5.4: **Snapshots of Monte Carlo simulations of hard disks.** Black dots are solvents and white dots are solutes. All solutes are $d = 1.0$, and all simulations are at $P = 0.5$. Configurations were randomly chosen from converged simulations at constant pressure. Snapshots are scaled so that simulation box volumes appear the same. (a) Solvent $d = 0.7$; (b) solvent $d = 1.0$; (c) solvent $d = 1.2$; (d) solvent $d = 1.5$.



(a)



(b)

Figure 5.5: **Hard disks obey the Hildebrand relation.** Each data style represents the entropies of insertion of one solute of diameter d into solvents of diameter $0.5d$ to $1.5d$. Each panel represents one pressure-volume condition. Dashed lines between points are best linear fits. x axes plot the natural logarithm of the solvent density and y axes plot entropies; both quantities are from Monte Carlo simulations. (a) Entropies at constant pressure. Fitted lines are, from the top curve down, $y = 1.06x - 0.64$, $y = 0.86x - 0.93$. Error bars on entropies are 0.4 or less. (b) Entropies at constant volume, with volumes as in simulations at $P = 0.5$. Linear fit is $y = 0.87x - 4.68$. Error bars on entropies are less than 0.05%.

Thus, Equation 5.5 appears to describe hard disk entropies regardless of the sign of free volume changes upon insertion. This can only be true if entropic effects from free volume changes aren't solvent-dependent. In this case, free volume contributions would only appear in the constant term of Equation 5.5. Indeed, in hard disks at constant pressure, system volume changes upon insertion of a given solute don't change significantly from one solvent to another (data not shown). We conclude that in hard disks, the $\ln \rho_1$ term of Equation 5.5 doesn't capture free volume effects on bulk solvent. The $\ln \rho_1$ term may describe microscopic physics of interaction close to the solute upon insertion.

To study the structure of our hard disk solvents near a solute, we next calculate solute-solvent pair correlation functions. These report the density of solvent centers, relative to bulk density, at a given distance from the solute. We show several pair correlation functions in Figure 5.6. This figure shows that in fluids of larger molecules, peaks of normalized density near the solute are higher and troughs lower than in fluids of smaller molecules. We believe this is related to the higher packing fraction in fluids of large disks, as seen in Figure 5.4. Because there is so little empty space in these fluids, large disks must adopt a tightly ordered structure to accommodate the solute. The link between this structuring force and the $\ln \rho_1$ term in Equation 5.5 is not yet clear.

Next we ask whether scaled particle theory (SPT) predicts Equation 5.5. Here we use the 2-D SPT formulation, Equation 5.2, for comparison with our 2-D Monte Carlo data. Taking pressures, densities, and expansion coefficients from our simulations, we use Equation 5.2 to compute the insertion entropies s_{ex} of one solute in a series of solvents. With these entropies we make a Hildebrand plot, to determine whether these SPT predictions follow Equation 5.5. Figure 5.7(b) shows the result. Slopes of the Hildebrand curves are within about 20% of 1.0.

What does this study of SPT tell us? SPT reproduces our constant-pressure hard disk entropies from Monte Carlo to within about 15% (data not shown). Since we have already seen that Equation 5.5 applies to our simulation data (see Figure 5.5), its success on SPT predictions

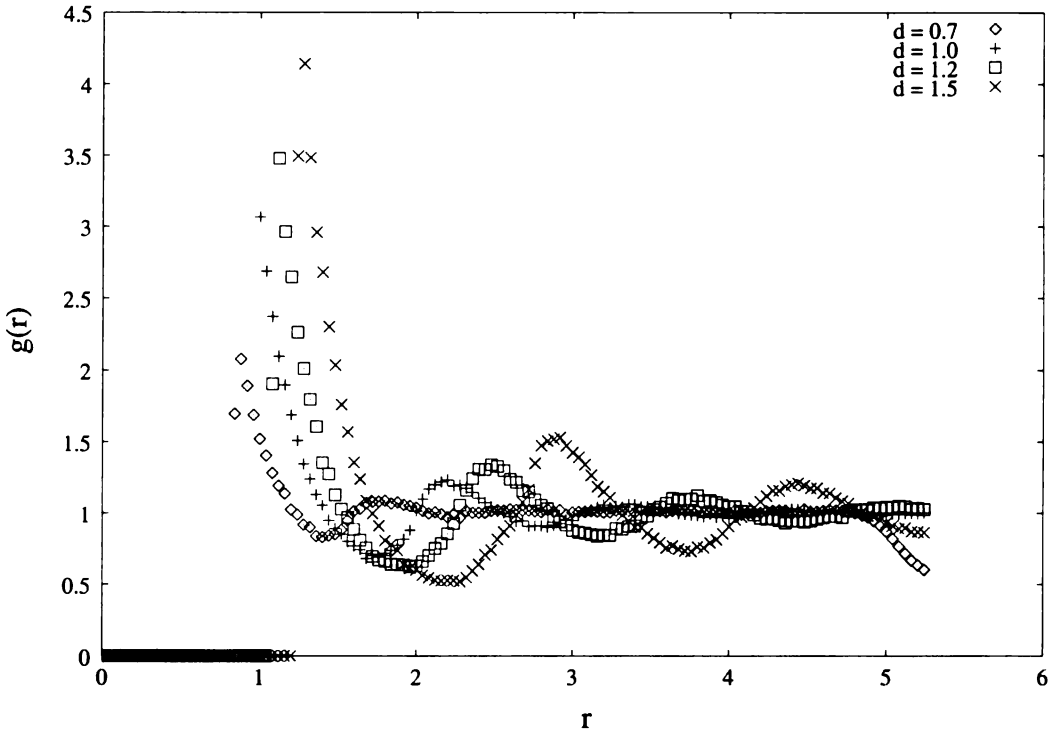


Figure 5.6: **Insertion of solutes orders large solvents more.** Each data point represents one solute-solvent interaction. Each data style represents one solvent size d. The x axis plots a solute-solvent separation r. The y axis plots the solute-solvent pair correlation function $g(r) = \frac{V}{N^2} \langle \sum_i \delta(\mathbf{r} - \mathbf{r}_{\mathbf{solute,i}}) \rangle$. Here $\langle \cdots \rangle$ indicates an ensemble average, $\mathbf{r}_{\mathbf{solute,i}}$ is the separation between solvent i and the fixed solute, $\mathbf{r}$ is a vector with length r, δ specifies a window function for numerical integration, V is the system volume and N is the number of solvents. $g(r)$ expresses the density of solvents within r of the solute, relative to the bulk density N/V.

in Figure 5.7(b) isn't surprising. But these results do suggest that SPT might be used to find the analytic basis for Equation 5.5.

5.4.4 Applying scaled particle theory to iodine data.

In the last section we saw that SPT predicted the Hildebrand relation, Equation 5.5, to within about 20% in 2-D systems of hard disks. This implies that SPT entropies calculated for the original iodine experimental data should also follow Equation 5.5 within some error. To test this idea, first we compute the 3-D SPT prediction for entropy with Equation 5.1, using parameters listed in Table 5.1. Then we plot this prediction against true experimental entropies,

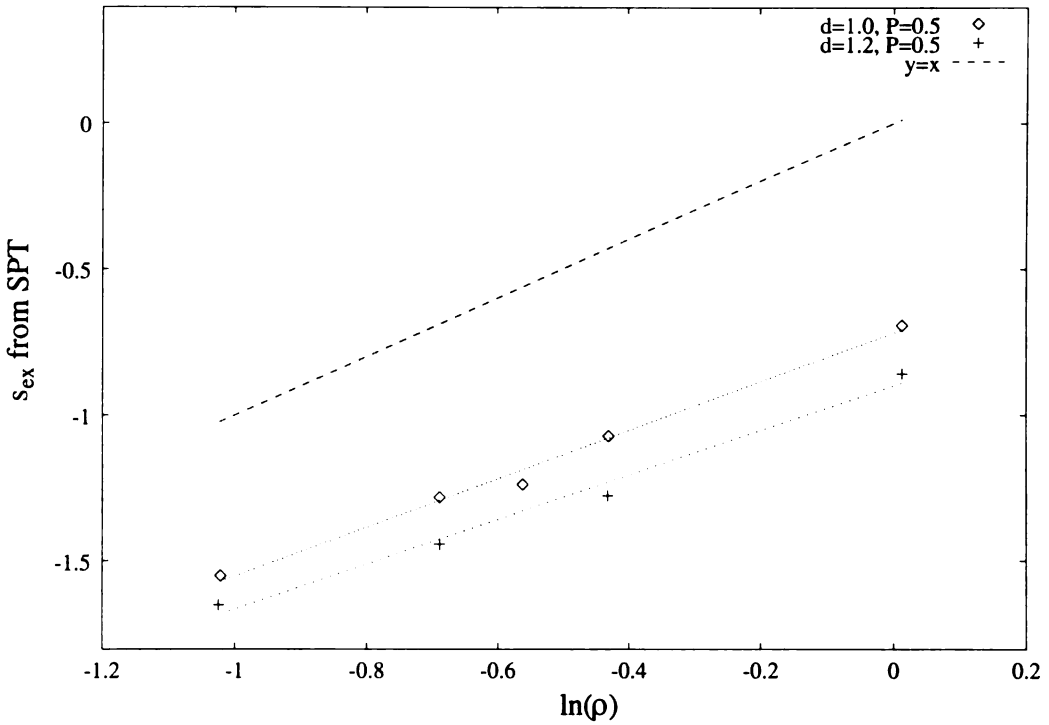


Figure 5.7: **2-D SPT entropies obey the Hildebrand relation.** Each data style represents the entropies of insertion of one solute of diameter d into solvents of diameter $0.5d$ to $1.5d$ at the same pressure P. Dashed lines between points are best linear fits. The x axis plots the natural logarithm of the solvent density from Monte Carlo simulations; the y axis plots entropies predicted by 2-D SPT (Equation 5.2) with parameter values taken from Monte Carlo simulations. Fitted lines are, from the top down, $y = 0.83x - 0.72$, $y = 0.76x - 0.90$. Error bars on SPT parameters are 3% or less.

or against the Hildebrand expression, $R \ln \rho_1$. The results are shown in Figure 5.8. All SPT predictions correlate linearly with experimental data. SPT follows Equation 5.5 most closely if we scale effective hard sphere diameters d of all the Hildebrand solvents by a single adjustable parameter.

Thus, our prediction from 2-D hard disks is only borne out when we adjust the effective diameters that describe real solvents. Is this scaling justified? It is not clear. Effective diameters come from fits of energy functions or equations of state to experimental data [15, 16]; they are semi-empirical estimates of the best way to describe solvent molecules as hard spheres. Also, SPT is known to be very sensitive to diameter values [17]. For these reasons, we conclude that we don't know the effective diameters at high enough resolution to learn anything quantitative

Table 5.1: **Parameters used in SPT applied to experimental data.** (a) Molecular species. 1 indicates 2,2-dimethylbutane; 2, silicon tetrachloride; 3, 2,2,4-trimethylpentane; 4, cyclohexane; 5, octamethylcyclotetrasiloxane; 6, 1,2-dichloro-1,2,3,3,4,4-hexafluorocyclobutane. (b) Molar density. (c) Coefficient of thermal expansion. Where no reference is cited, the value for α was assigned empirically. (d) Effective diameter of a hard sphere representing the solvent molecule. Where [15] is cited, d values were taken straight from Table I of that paper when available, and otherwise calculated from Table III of that paper.

(a) species	(b) ρ, mol/cm^3	(c) α, K^{-1}	(d) d, cm/mol
1	0.00747 [20]	0.00147 [21]	2.50 [15]
2	0.00873 [20]	0.00143 [22]	2.47 [16]
3	0.00606 [20]	0.0012 [23]	2.76 [15]
4	0.00926 [20]	0.0015 [20]	2.38 [15]
5	0.00303 [1]	0.0006	3.51 [15]
6	0.00714 [1]	0.0013	2.65 [15]
iodine			2.02 [24]

from our test. Qualitatively, however, Hildebrand's iodine solvation data appear to follow scaled particle theory, suggesting that solvents in these experiments behave like hard spheres.

5.5 Conclusions.

In this paper, we have asked whether the agreement of Hildebrand *et al*'s iodine solvation data with mole-fraction theory is due to a general property of "compact" molecules. We first re-formulated the mole fraction expression in terms of the excess entropy s_{ex}: the result is $s_{ex} = R \ln \rho_1 + C$, where ρ_1 is the molar density of the solvent, and C is a solute-dependent constant. We found that entropies of a simple hard disk model of liquids obey this expression. The reason for this success may be linked to ordering effects among solvents close to the solute. We have also compared the $R \ln \rho_1 + C$ formula to predictions from scaled particle theory (SPT). In the hard disk model, SPT excess entropies follow the formula reasonably well, suggesting that SPT can be used to find the analytic basis for the success of $R \ln \rho_1 + C$. However, when applied to experimental data, SPT results are ambiguous because of uncertainty in the parameters needed for the calculation.

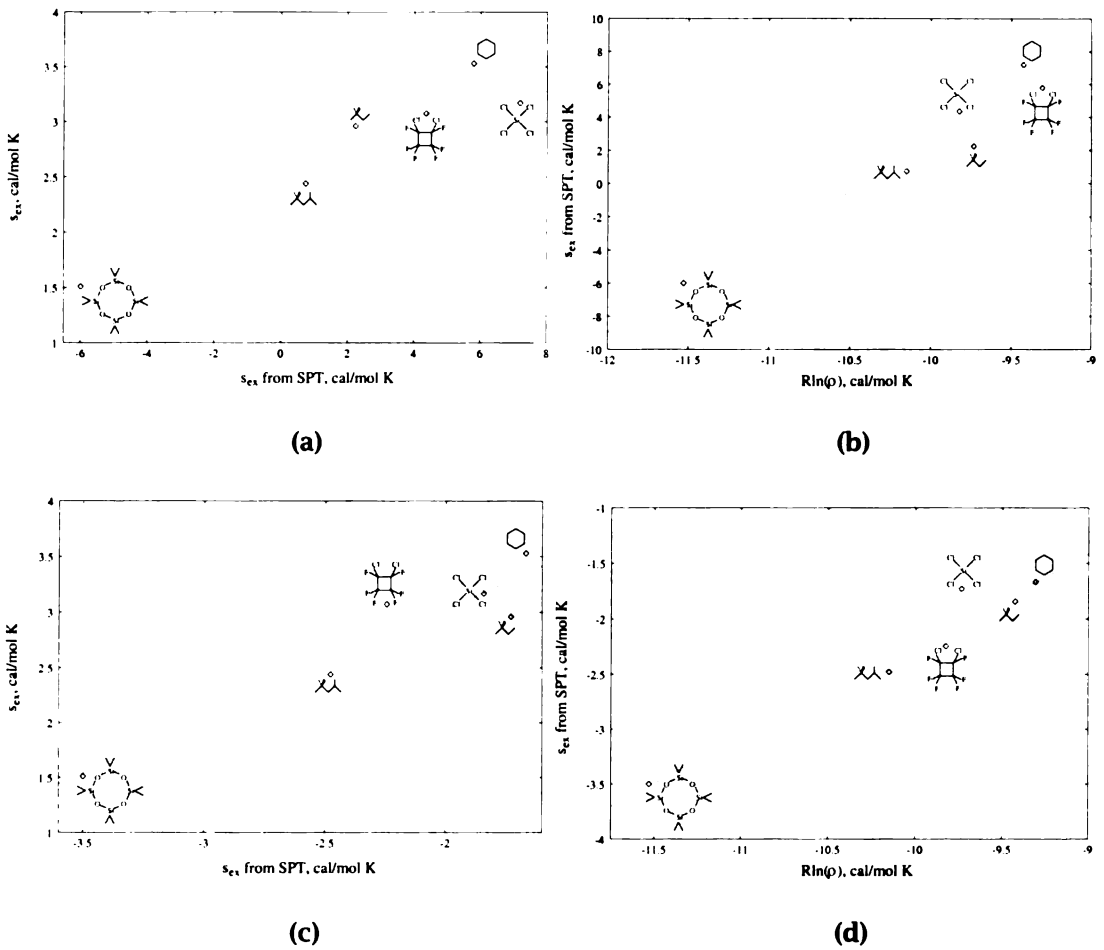


Figure 5.8: **Applying SPT to iodine data.** All experimental entropies of iodine solvation are from [18, 19], densities are from [1, 20], and parameters for SPT are as in Table 5.1. Dashed lines are best linear fits. (a) SPT predicted entropy *versus* true experimental entropy. Fitted line is $y = 0.15x + 2.43$. (b) Hildebrand expression, $R \ln \rho_1$, *versus* SPT prediction. Fitted line is $y = 5.65x + 58.8$. (c) SPT predicted entropy *versus* true experimental entropy, when molecular diameters are scaled by 0.84. Fitted line is $y = 0.99x + 4.99$. (d) Hildebrand expression, $R \ln \rho_1$, *versus* SPT prediction, when molecular diameters are scaled by 0.84. Fitted line is $y = 0.83x + 6.07$.

5.6 Acknowledgements.

We thank Kevin Silverstein for the use of his computer code and for technical advice. We also thank Noel Southall, Jack Schonbrun, and Karen Tang for helpful discussions. R.B. was a National Science Foundation predoctoral fellow and a University of California Regents' Fellow while this work was in progress.

Bibliography

[1] Hildebrand JH, Scott RL. *Regular Solutions.* Prentice-Hall, Englewood Cliffs, NJ, 1962.

[2] Chan HS and Dill KA. Solvation: effects of molecular size and shape. 1994. *J. Chem. Phys. 101(8):*7007-7026.

[3] Chan HS and Dill KA. Solvation: how to obtain microscopic energies from partitioning and solvation epxeriments. 1997. *Annu. Rev. Biophys. Biomol. Struct. 26:*425-59.

[4] Lee B. The physical origin of the low solubility of nonpolar solutes in water. 1985. *Biopolymers 24:*813-823.

[5] Madan B and Lee B. Role of hydrogen bonds in hydrophobicity: the free energy of cavity formation in water models with and without the hydrogen bonds. 1994. *Biophys. Chem. 51:*279-289.

[6] Pratt LR and Pohorille A. Theory of hydrophobicity: transient cavities in molecular liquids. 1992. *Proc. Natl. Acad. Sci. USA 89:*2995-2999.

[7] Sharp KA, Kumar S, Rossky PJ, Friedman RA, and Honig B. Size dependence of transfer free energies. 2. Hard sphere models. 1996. *J. Phys. Chem. 100:*14166-14177.

[8] Krukowski AE, Chan HS and Dill KA. An exact lattice model of complex solutions: chemical potentials depend on solute size and shape. 1995. *J. Chem. Phys. 103(24):*10675-10688.

[9] Widom B. Some topics in the theory of fluids. 1963. *J. Chem. Phys. 39*:2808-2812.

[10] Reiss H, Frisch HL, Helfand E, and Lebowitz JL. Aspects of the statistical thermodynamics of real fluids. 1960. *J. Chem. Phys. 32(1)*:119-124.

[11] Helfand E, Frisch HL, and Lebowitz JL. Theory of the two- and one-dimensional rigid sphere fluids. 1961. *J. Chem. Phys. 34(3)*:1037-1042.

[12] Chatelier RC and Minton AP. Adsorption of globular proteins on locally planar surfaces: models for the effect of excluded surface area and aggregation of adsorbed protein on adsorption equilibria. 1996. *Biophys. J. 71*:2367-2374.

[13] Ben-Naim A. 1987. *Solvation Thermodynamics.* Plenum Press, New York.

[14] Lazaridis T. The origin of hydrophobicity. 2000. *Proc. Nat. Acad. Sci.*, submitted.

[15] Ben-Amotz D and Willis KG. Molecular hard-sphere volume increments. 1993. *J. Phys. Chem. 97*:7736-7742.

[16] Zarkova L, Petkov I and Pirgov P. An approach to the calculation of self-consistent thermophysical properties of scarcely examined heavy gaseous halides. 1998. *J. Phys. B 31*:805-813.

[17] Tang KES and Bloomfield VA. Excluded volume in solvation: sensitivity of scaled-particle theory to solvent size and density. 2000. *Biophys. J.*, in press.

[18] Shinoda K and Hildebrand JH. The solubility and entropy of solution of iodine in octamethylcyclotetrasiloxane and tetraethoxysilane. 1957. *J. Phys. Chem. 61*:789-791.

[19] Shinoda K and Hildebrand JH. The solubility and entropy of solution of iodine in *n*-C_7F_16, *c*-$C_6F_{11}CF_3$, $(C_3F_7COOCH_2)_4C$, *c*-$C_4Cl_2F_6$, CCl_2FCClF_2 and $CHBr_3$. 1958. *J. Phys. Chem. 62*:292-294.

[20] Lide DR, Editor. *CRC Handbook of Chemistry and Physics.* 2000. CRC Press, Boca Raton, FL.

[21] Riddick JA, Burger WB, Sakano TK. *Organic Solvents: Physical Properties and Methods of Purification.* 1986. Wiley, New York.

[22] Dean JA, Editor. *Lange's Handbook of Chemistry.* 1979. McGraw-Hill, New York.

[23] Vroom EM, Noordegraaf MA, and Jacobs HJC. Contraction of isooctane, diethyl ether and ethanol between +25 and −80° C. 1991. *Spectrochimica Acta 47A(12)*:1781-1782.

[24] Abraham MH and McGowan JC. The use of characteristic volumes to measure cavity terms in reversed phase liquid chromatography. 1987. *Chromatographia 23(4)*:243-246.

The San Francisco Library

For reference

Not to be taken from the room.

628377

3 1378 00628 3777