Title

Core-packing constraints, hydrophobicity and protein design

Permalink

https://escholarship.org/uc/item/6k25c588

Journal

Current Opinion in Biotechnology, 5(4)

ISSN

0958-1669

Authors

Baldwin, Enoch P
Matthews, Brian W

Publication Date

1994-08-01

DOI

10.1016/0958-1669(94)90048-5

Peer reviewed

Core-packing constraints, hydrophobicity and protein design

Enoch P Baldwin and Brian W Matthews

Howard Hughes Medical Institute, University of Oregon, Eugene, USA

Recent crystallographic studies have shown that both backbone and side-chain adjustments occur when different core-packing arrangements are accommodated in proteins. Thus, modeling methods, which have typically considered only side-chain adjustments, must now also account for backbone movements to accurately predict the energies and structures of mutated or designed proteins. The 'plasticity' of protein cores demonstrated by random mutagenesis simplifies protein design by increasing the likelihood of identifying alternative core sequences.

Introduction

To effectively engineer proteins, we need to understand how amino acid sequences specify structure, function and stability. Patterns of hydrophobic residues in polypeptide sequences are a key determinant of the type of fold that is adopted [1,2]. In soluble globular proteins, burial of hydrophobic residues drives compaction and helps specify secondary and tertiary structure [3,4•,5]. Non-polar side chains typically pack tightly in the interior, forming the solvent-inaccessible protein 'core', whereas surface side chains are generally polar. Tight packing of buried residues has been suggested to limit the core sequences that are tolerated in a particular fold [6–9]. In computational approaches to protein design and structure analysis, packing constraints are a major consideration. Here, we briefly review the origins of these ideas and discuss recent work indicating that such constraints are, in fact, much looser than previously thought. Paradoxically, the structural basis for tolerance to different cores complicates prediction of the effects of mutation on structure and stability, but simplifies the problem of protein design.

The importance of the hydrophobic effect in determining protein structure and stability

Hydrophobicity patterns in amino acid sequences are one of the most conserved features of proteins that have the same fold [2]. Theoretical studies using highly simplified lattice models [10–12] of protein chains, as well as recent experimental work, suggest that such patterns can specify secondary and tertiary structure. For example, four-helix bundles were readily generated using polypeptide sequences based on heli-cal repeats of hydrophobic and hydrophilic residues [13,14••,15,16•,17•]. Coiled-coil dimers, trimers and tetramers have also been specified by simple patterns of hydrophobic residues [18••]. In mutants containing insertions between residues that are in α-helices [19•,20••,21•,22•], the choice between 'looping out' or translocating adjacent helical residues was apparently controlled by the relative dispositions of polar and non-polar residues [19•,20••]. Finally, burial of a non-polar surface has been correlated with helical propensity at solvent-accessible sites in T4 lysozyme [23•,24••].

The importance of the hydrophobic nature of protein cores for stability is also well established [25•,26•,27]. Truncation of core side chains to alanine [27–31, 32•–34•] is much more destabilizing (up to 5 kcal mol^{-1} for leucine residues) than alanine substitutions of most surface residues (± 1 kcal mol^{-1}) [23•,30,31,32•–34•,35, 36•]. Similarly, substituting non-polar buried side chains with smaller hydrophobic [27,28,37••,38,39] or polar residues [40•,41–43] is more destabilizing than increasing the hydrophobicity of surface-exposed side chains [23•,24••]. In some cases, substitution of buried polar residues with non-polar amino acids has led to stabilization [44•,45,46].

Theoretical evaluation of packing constraints

The constraints imposed on protein structure and stability by core packing have not yet been clearly defined [25•,26•,47•,48•,49]. As judged from protein crystal structures, core side chains are 'well packed': they are rigid, clustered at high density, and generally in low-energy conformations. This jigsaw puzzle-like fit of side chains led to the suggestion that very few combinations of hydrophobic residues might be compatible

---

Abbreviation

rmsd—root mean square deviation.

with a given fold [7]. In this view, packing arrangements, and thus the identities of core residues, contain crucial information specifying protein structure.

To test this idea and to evaluate packing constraints, Ponder and Richards [7] designed a prototypical packing algorithm, PROPAK, that evaluates all of the possible side-chain arrangements of a group of adjacent residues. Trial models were constructed by placing alternative side chains in the context of a fixed framework of surrounding side-chain and backbone atoms derived from crystal structures. The alternative side-chain orientations were restricted to a limited number of 'ideal' or 'preferred' conformers (rotamer libraries) that were determined from surveys of known structures [7,24••,50•,51,52•]. Models were evaluated by counting the number of close van der Waals contacts. Only a minute fraction of the total number of side-chain combinations, including the wild type, satisfied the contact criteria, suggesting that viable alternative core-packing arrangements are rare.

Recently, more sophisticated algorithms have been developed both to predict the mutational effects on protein stability and to build models of proteins from known homologous structures. These methods surveyed potential side-chain configurations, either by stepwise enumeration of the possible conformations for subsets of side chains [52•,53,54•–56•,57,58,59•] or by sampling via Monte Carlo approaches [60] or other search methods [51,61•,62]. The side-chain conformations were generated either from rotamer libraries [51,52•,53,55•,58,59•,62] (as with PROPAK) or by stepping through coarse increments of torsion angles (10–120°) [54•,57,60], and were often subsequently optimized using energy functions [51,53,54•–56•,57,60] that are very sensitive to interatomic distances [63,64•,65]. As is the procedure with PROPAK, the model backbone and surrounding framework atoms were held fixed, at least during side-chain enumeration. Where overall energy minimization was performed on a limited subset of culled structures, the backbones did not deviate far from their initial positions [53,54•,58]. The inherent assumption of these approaches is that movements of the backbone away from the wild-type position to avoid steric conflicts are destabilizing [66].

Mutational studies of protein cores

Early mutational studies supported the idea that core residues are also sensitive to hydrophobic substitutions [25•,26•,47•,48•]. Temperature-sensitive mutations often occur at buried sites [67,68•], which generally have a more limited range of permissible substitutions than surface sites [68•,69,70•]. Collisions with the backbone, disruption of the packing of adjacent residues, and the forcing of side chains into unfavorable conformations, are all potential sources of destabilization at tightly packed sites, especially for small to large substitutions

[58,71–74,75•,76•]. In an extreme example, the substitution Ala98→Val (three-letter amino acid code) in T4 lysozyme destabilized the protein by $5 \, \text{kcal} \, \text{mol}^{-1}$ [71]. The mutant structure showed that the increased size of the side chain perturbed the adjacent backbone. On the other hand, small to large mutants with increased stability have also been generated [43,44•,74,77•,78], demonstrating that increased hydrophobicity, improved van der Waals contacts, and greater packing efficiency can potentially overcome unfavorable factors, such as steric interference and torsional strain. Hydrophobicity and packing can have comparable contributions to stability. Interchanging side chains of similar hydrophobicity, but different shapes (e.g. phenylalanine, leucine, isoleucine and methionine), has lead to a wide range of stability changes ($+0.3 \, \text{kcal} \, \text{mol}^{-1}$ to $-4 \, \text{kcal} \, \text{mol}^{-1}$ [28,37••,38,39,58,76•]).

Paradoxically, although proteins can be somewhat sensitive to single substitutions, they are surprisingly tolerant of multiple adjacent interior substitutions, provided that the overall hydrophobicity of the core is roughly maintained. Libraries of core-packing variants of the DNA-binding domain of λ repressor and T4 lysozyme were evaluated for function and stability [79,80,81••,82••]. As many as 70% [80] of the possible combinations of hydrophobic side chains were tolerated, albeit few with stability near that of the wild type. Some sites were more sensitive to substitution than others. Stability or function did not correlate well with volume, hydrophobicity, root mean square deviation (rmsd) from wild-type structure (where structures were determined) or packing density, although mutants with larger hydrophobic cores tended to be more stable [48•,82••]. The ranges of tolerated side-chain volumes and hydrophobicities were within ± 5–6 methylenes and ± 3–$4 \, \text{kcal} \, \text{mol}^{-1}$, respectively, of the wild type, but repressor activity was more sensitive to changes in volume than T4 lysozyme. A triple mutant of λ repressor was actually $0.5 \, \text{kcal} \, \text{mol}^{-1}$ more stable than the wild type. Unlike multiple mutants in general, which usually have additive effects on stability [83•,84], all core mutants with adjacent substitutions that have been examined were more stable (by up to $3.9 \, \text{kcal} \, \text{mol}^{-1}$) than expected by additivity [58,71,76•,82••,83•]. This non-additivity indicates that well packed residues form cooperative arrangements that can be disrupted by a single substitution [47•,48•,85,86]. Further substitutions result in new interactions, although rarely as cooperative as in the wild type.

Structural studies of core mutants

A number of crystal structures of single and multiple core mutants have recently been determined [29,37••, 40•,41,42,44•,58,71,77•,82•• ,87,88•–90•], revealing why proteins are tolerant of changes in packing, and allowing evaluation of the assumptions that underlie current computational methods.

New packing arrangements were typically accommodated by overall adjustments of the backbone (up to 1.0 Å) without major distortions of secondary structure, together with small changes (10–20°) in side-chain torsion angles. In mutants in which the net volume of core side chains is reduced, structural relaxation decreases the amount of unfilled space [29,37••,87,88•–90•,91], presumably reducing destabilization [29,37••,92•]. Analogously, increasing total side-chain volume results in expansion of the cores and concerted shifts of secondary structure of up to 0.8 Å [44•,58,82••,91,93,94•,95••]. Increased packing efficiency and stability has been observed in some [44•,77•,95••], but not all, cases.

In a few mutant structures, substantial side-chain rotations toward cavities [29] or away from introduced bulky side chains have been observed [44•,82••]. Mutated side chains or side chains nearby the mutation sites occasionally adopt non-ideal torsion angles [37••,44•,58,77•,90•], even in stabilized mutants [44•,77•]. The conformations of alternative side chains in multiple mutants are generally similar to each other and are within 20° of wild-type or 'ideal' torsion angles [58,71,82••,95••]. Thus, they are determined by similar constraints, including the local secondary structure and the surrounding residues that are unchanged [52•,59•]. In homologous proteins with greater than 50% sequence identity, side-chain conformations and torsion angles are also typically preserved, supporting this idea [57,96–98]. Although changes in torsion angles contribute to the reduction of unfavorable contacts, shifts of secondary structure are largely responsible for the accommodation of altered side chains, at least in the helical proteins studied. Relatively small shifts in backbone α-carbons (0.5 Å) can be accompanied by larger changes (1–2 Å) in side-chain atom positions. These adjustments assist in redistributing side-chain bulk and optimizing new sets of contacts, so that many alternative side-chain combinations can lead to a tolerably well packed core.

For the same protein in different crystal environments, where it is subject to different crystal-packing forces, the backbone shifts are similar to those observed in the repacked mutants (0.3–0.4 Å rmsd after superposition [24••,80,96,97]), suggesting that protein backbones are inherently flexible, and adjustments of this magnitude may have little penalty [82••]. In an extreme case, one triple variant of T4 lysozyme had a backbone shift of 0.63 Å, yet was destabilized only by 1.4 kcal mol⁻¹ [81••]. Clearly, even relatively large deviations of crystallographically determined backbone positions from the wild type do not necessarily translate into large destabilizations.

The infrequency of alternative rotamers in packing mutants is likely due to the close-packed nature of cores. In an unusual T4 lysozyme core-repacking variant [82••], the subsititution of Ala→Trp (a large for a small residue) required an adjacent amino acid on the same helix to adopt a different rotamer to avoid a steric clash. Because this residue was also completely buried, it was expected

that other potential steric conflicts would be incurred. In this case, however, a potential collision was averted because the conflicting side chain was at the surface and freely rotated away. For tightly packed residues in the core, such coordinated conformational changes are likely to be more problematic.

Implications of core-repacking studies for protein design

Computational approaches to structure and stability prediction have been reasonably successful (60–90%) in reproducing core side-chain conformations from known structures [7,51,52•,53,54•–56•,57,59•,61•]. In some cases, the relative stabilities of mutants [58,61•, 66], have also been predicted. These methods have, however, been less successful either in correctly predicting the conformations of mutant or homologous structures based on a known 'parent', or in predicting the finer details of known structures. The variety of structural responses to core substitutions, particularly flexibility of the backbone, suggests that precise prediction of mutant protein stability and structure still presents an enormous challenge to computational methods. Substantial deviations both from wild-type backbone positions and from 'ideal' torsion angles are routinely observed in mutant structures. Therefore, modeling based on 'ideal' torsion angles [50•] or rigid backbones is likely to lead to incorrect structures or unreliable estimates of stability changes caused by mutations. Even if such adjustments are small, they can significantly influence interatomic distances and calculated energies. Furthermore, analysis of mutant structures suggests that both backbone shifts and changes in torsion angles can have similar energetic consequences and, therefore, cannot be evaluated independently. Although present prediction methods may be in error in excluding viable sequences and configurations, they may still be of use in more qualitative applications, such as in predicting the least perturbing interior substitutions [58,60], predicting the structures of homologous proteins [99,100•,101•], and suggesting core-residue combinations for packing *de novo* designed structures.

Do protein structures and stabilities need to be accurately predicted to effectively engineer them? A recent experiment by Hecht and coworkers [14••] suggests that *de novo* designed proteins can be obtained in the absence of strict design criteria [102]. They generated a library of potential four-helix bundles specified by a simple 'binary' pattern of polar and non-polar amino acids. Although previous approaches were based on fully defined sequences [13,15,16•], in this case, a potential 10⁴¹ combinations were specified using random assortments of five different non-polar amino acids and six different polar types. When expressed in *Escherichia coli*, 48 of 69 sequences (60%) yielded soluble products. Initial characterization of three polypeptides showed that they

were monomeric and highly helical, and two had stabilities approaching those of known proteins (3.7 kcal mol^{-1} and 4.4 kcal mol^{-1}). Thus, a large fraction of all possible sequences folded into compact soluble structures. At least 3% were reasonably stable, suggesting that packing details are a very weak constraint in this system. Even so, a cautionary note is required: many of these bundles may be 'molten globules' or may lack well defined three-dimensional structures. Characterization of native-like sequences obtained in this way should facilitate the formulation of rules for design of novel polypeptide structures.

Conclusions

The 'plasticity' of protein cores that is demonstrated by data from random mutagenesis, simplifies protein design by increasing the likelihood of identifying the alternative amino acid sequences that lead to folded functional proteins. At the same time, however, recent crystallographic studies have shown that both backbone and side-chain adjustments occur when different core-packing arrangements are accommodated in proteins. Thus, modeling methods, which have typically considered only side-chain adjustments, must now also account for backbone movements to accurately predict the energies and structures of mutated or designed proteins. This remains the challenge for the future.

Acknowledgements

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:
* of special interest
•• of outstanding interest

1. Dill KA: **Dominant Forces in Protein Folding.** *Biochemistry* 1990, **29**:7133–7155.

2. Bowie JU, Reidhaar Olson JF, Lim WA, Sauer RT: **Deciphering the Message in Protein Sequences: Tolerance to Amino Acid Substitutions.** *Science* 1990, **247**:1306–1310.

3. Kauzman W: **Some Factors in the Interpretation of Protein Denaturation.** *Advan Prot Chem* 1959, **14**:1–63.

4. Ponnuswamy PK: **Hydrophobic Characteristics of Folded Pro-**
•   **teins.** *Prog Biophys Mol Biol* 1993, **59**:57–103.
Comprehensive review of the role of hydrophobicity in protein folding and stability.

5. Hopp TP: **Use of Hydrophilicity Plotting Procedures to Identify Protein. Antigenic Segments and Other Interaction Sites.** *Methods Enzymol* 1989, **178**:571–585.

6. Richards FM: **Areas, Volumes, Packing and Protein Structure.** *Annu Rev Biophys Bioeng* 1977, **6**:151–176.

7. Ponder JW, Richards FM: **Tertiary Templates for Proteins. Use of Packing Criteria in the Enumeration of Allowed Sequences for Different Structural Classes.** *J Mol Biol* 1987, **193**:775–791.

8. Janin J, Wodak S: **Conformation of Amino Acid Side Chains in Proteins.** *J Mol Biol* 1978, **125**:357–386.

9. Chothia C: **Structural Invariants in Protein Folding.** *Nature* 1975, **254**:304–308.

10. Chan HS, Dill KA: **Origins of Structure in Globular Proteins.** *Proc Natl Acad Sci USA* 1990, **87**:6388–6392.

11. Yue K, Dill KA: **Inverse Protein Folding Problem: Designing Polymer Sequences.** *Proc Natl Acad Sci USA* 1992, **89**:4163–4167.

12. Sikorski A, Skolnick J: **Monte Carlo Studies on Equilibrium Globular Protein Folding. III. The Four Helix Bundle.** *Biopolymers* 1989, **28**:1097–1113.

13. DeGrado WF, Wasserman ZR, Lear JD: **Protein Design, a Minimalist Approach.** *Science* 1989, **243**:622–628.

14. Kamtekar S, Schiffer JM, Xiong H, Babik JM, Hecht MH: **Protein**
•• **Design by Binary Patterning of Polar and Nonpolar Amino Acids.** *Science* 1993, **262**:1680–1685.
This paper describes a combinatorial genetic approach to obtaining four-helix bundles. A degenerate gene library, which encodes polypeptides with the same pattern of residue polarity but different distinct sequences, is screened for products accumulating in soluble form in *Escherichia coli* that are judged to be 'folded'. Preliminary physical characterization of three proteins shows that they are highly helical and compact. This paper will likely inspire many similar experiments with other folds.

15. Hecht MH, Richardson JS, Richardson DC, Ogden RC: **De Novo Design, Expression, and Characterization of Felix: A Four-Helix Bundle Protein of Native-Like Sequence.** *Science* 1990, **249**:884–891.

16. Johnsson K, Allemann RK, Widmer H, Benner SA: **Synthesis,**
•   **Structure and Activity of Artificial, Rationally Designed Catalytic Polypeptides.** *Nature* 1993, **365**:530–532.
A description of a designed lysine rich four-helix bundle that catalyzes the decarboxylation of oxaloacetate with a rate acceleration of 10^2–10^3. Data show that a relatively non-specific surface with appropriate functionality can significantly catalyze a reaction.

17. DeGrado WF: **Peptide Engineering. Catalytic Molten Globules.**
•   *Nature* 1993, **365**:488–489.
An authoritative review on designed four-helix bundles.

18. Harbury PB, Zhang T, Kim PS, Alber T: **A Switch between**
•• **Two-, Three-, and Four-Stranded Coiled Coils in GCN4 Leucine Zipper Mutants.** *Science* 1993, **262**:1401–1407.
This paper describes simple rules for specifying different coiled-coil oligomers on the basis of sequences of residues at the helix interfaces.

19. Heinz DW, Baase WA, Zhang XJ, Blaber M, Dalquist FW,
•   Matthews BW: **Accommodation of Amino Acid Insertions in a Helix of T4 Lysozyme.** *J Mol Biol* 1994, **236**:869–886.
A complete structural and thermodynamic analysis of T4 lysozyme mutants with one to four residues inserted into the beginning, middle, and ends of an α-helix. The pattern of hydrophobic residues in sequences adjacent to the insertion site are shown to influence the choice between 'looping out' or translocation of the adjacent residues toward the end of the helix.

20. Heinz DW, Baase WA, Dahlquist FW, Matthews BW: **How**
•• **Amino Acid Insertions are Allowed in an α-Helix of T4 Lysozyme.** *Nature* 1993, **361**:561–564.
Describes the structures of insertion mutants, in which either 'looping out' of inserted residues or translocation of neighboring helical residues occurs. When amino acids are inserted, the helix length is maintained, but the adjacent loops increase in size. See also [21•,22•].

21. Keefe LJ, Sondek J, Shortle D, Lattman EE: **The α-Aneurism:**
•   **A Structural Motif Revealed in an Insertion Mutant of Staphylococcal Nuclease.** *Proc Natl Acad Sci USA* 1993, **90**:3275–3279.
See [22•].

22. Kavanaugh JS, Moo Penn WF, Arnone A: **Accommodation of**
• **Insertions in Helices: The Mutation in Hemoglobin Catonsville**
(Pro 37 α-Glu-Thr 38α) Generates a 3(10)-α. Bulge. *Biochemistry* 1993, **32**:2509–2513.
This paper and [21•] describe structural studies of insertion mutants of hemoglobin and staphylococcal nuclease, in which the inserted residue interrupts the local helical structure.

23. Blaber M, Zhang XJ, Matthews BW: **Structural Basis of Amino**
• **Acid α-Helix Propensity.** *Science* 1993, **260**:1637–1640.
This paper proposes that helical propensity of a particular residue type is influenced by the amount of surface area buried upon folding.

24. Blaber M, Zhang XJ, Lindstrom JD, Pepiot SD, Baase WA,
•• Matthews BW: **Determination of Helix Propensity Within the**
Context of a Folded Protein. *J Mol Biol* 1994, **235**:600–624.
A comprehensive structural and thermodynamic analysis of substitutions at a solvent-exposed site in the middle of a helix. It discusses the contributions of hydrophobicity and side-chain entropy to helix propensity and includes a comprehensive survey of torsion angles in highly refined X-ray structures.

25. Matthews BW: **Structural and Genetic Analysis of Protein Fold-**
• **ing and Stability.** *Curr Opin Struct Biol* 1993, **3**:589–593.
See [26•].

26. Rose GD, Wolfenden R: **Hydrogen Bonding, Hydrophobicity,**
• **Packing, and Protein Folding.** *Annu Rev Biophys Biomol Struct* 1993, **22**:381–415.
This review and [25•,27] summarize recent thermodynamic and structural analyses of mutant proteins.

27. Pace CN: **Contribution of the Hydrophobic Effect to Globular Protein Stability.** *J Mol Biol* 1992, **226**:29–35.

28. Matsumura M, Becktel WJ, Matthews BW: **Hydrophobic Stabilization in T4 Lysozyme Determined Directly by Multiple Substitutions of Ile 3.** *Nature* 1988, **334**:406–410.

29. Eriksson AE, Baase WA, Zhang XJ, Heinz DW, Blaber M, Baldwin EP, Matthews BW: **Response of a Protein Structure to Cavity Creating Mutations and its Relation to the Hydrophobic Effect.** *Science* 1992, **255**:178–183.

30. Heinz DW, Baase WA, Matthews BW: **Folding and Function of a T4 Lysozyme Containing 10 Consecutive Alanines Illustrate the Redundancy of Information in an Amino Acid Sequence.** *Proc Natl Acad Sci USA* 1992, **89**:3751–3755.

31. Zhang XJ, Baase WA, Matthews BW: **Multiple Alanine Replacements Within α-Helix 126–134 of T4 Lysozyme Have Independent, Additive Effects on Both Structure and Stability.** *Protein Sci* 1992, **1**:761—776.

32. Matthews BW: **Structural and Genetic Analysis of Protein Sta-**
• **bility.** *Annu Rev Biochem* 1993, **62**:139–160.
Review on structure and stability studies of mutant proteins.

33. Lin L, Pinker RJ, Kallenbach NR: **Helix Stability and the Native**
• **State of Myoglobin.** *Biochemistry* 1993, **32**:12638–12643.
See [34•].

34. Pinker RJ, Lin L, Rose GD, Kallenbach NR: **Effects of Alanine**
• **Substitutions in Helices of Sperm Whale Myoglobin on Protein Stability.** *Protein Sci* 1993, **2**:1099–1105.
This paper and [33•] describe recent studies using alanine mutagenesis to probe the role in stability of buried residues in myoglobin.

35. Pjura P, McIntosh LP, Wozniak JA, Matthews BW: **Perturbation of Trp 138 in T4 Lysozyme by Mutations at Gln 105 Used to Correlate Changes in Structure, Stability, Solvation, and Spectroscopic Properties.** *Proteins* 1993, **15**:401–412.

36. Gregoret LM, Sauer RT: **Additivity of Mutant Effects Assessed**
• **by Binomial Mutagenesis.** *Proc Natl Acad Sci USA* 1993, **90**:4246–4250.
Proposes a statistical method for identifying interacting pairs of residues from the frequency of corresponding pairs of substitutions in a functional population of random multiple alanine mutants.

37. Eriksson AE, Baase WA, Matthews BW: **Similar Hydrophobic**
•• **Replacements of Leu 99 and Phe 153 Within the Core of T4 Lysozyme Have Different Structural and Thermodynamic Consequences.** *J Mol Biol* 1993, **229**:747–769.

A complete thermodynamic and crystallographic analysis of site-directed core mutants that illustrates how accommodation of substitutions is affected by context.

38. Kellis JT Jr, Nyberg K, Sali D, Fersht AR: **Contribution of Hydrophobic Interactions to Protein Stability.** *Nature* 1988, **333**:784–786.

39. Varadarajan R, Connelly PR, Sturtevant JM, Richards FM: **Heat Capacity Changes for Protein–Peptide Interactions in the Ribonuclease S System.** *Biochemistry* 1992, **31**:1421–1426.

40. Blaber M, Lindstrom JD, Gassner N, Xu J, Heinz DW, Matthews
• BW: **Energetic Cost and Structural Consequences of Burying a Hydroxyl Group Within the Core of a Protein Determined from Ala→Ser and Val→Thr Substitutions in T4 Lysozyme.** *Biochemistry* 1993, **32**:11363–11373.
Describes a structural and thermodynamic survey of the responses to isosteric polar replacements of core residues. Evidence from this study shows that buried polar groups always seem to satisfy their hydrogen bonds, if necessary by sequestration of a solvent molecule.

41. Daopin S, Anderson DE, Baase WA, Dahlquist FW, Matthews BW: **Structural and Thermodynamic Consequences of Burying a Charged Residue Within the Hydrophobic Core of T4 Lysozyme.** *Biochemistry* 1991, **30**:11521–11529.

42. Stites WE, Gittis AG, Lattman EE, Shortle D: **In a Staphylococcal Nuclease Mutant the Side-Chain of a Lysine Replacing Valine 66 is Fully Buried in the Hydrophobic Core.** *J Mol Biol* 1991, **221**:7–14.

43. Lim WA, Farruggio DC, Sauer RT: **Structural and Energetic Consequences of Disruptive Mutations in a Protein Core.** *Biochemistry* 1992, **31**:4324–4333.

44. Anderson DE, Hurley JH, Nicholson H, Baase WA, Matthews
• BW: **Hydrophobic Core Repacking and Aromatic–Aromatic Interaction in the Thermostable Mutant of T4 Lysozyme Ser 117 Phe.** *Protein Sci* 1993, **2**:1285–1290.
Describes the structure of a packing mutant, containing phenylalanine in place of a buried serine, that has enhanced stability and a modest increase in packing efficiency. Backbone adjustments and side-chain conformation changes are also observed.

45. Vriend G, Berendsen HJ, van der Zee JR, van den Burg B, Venema G, Eijsink VG: **Stabilization of the Neutral Protease of *Bacillus stearothermophilus* by Removal of a Buried Water Molecule.** *Protein Eng* 1991, **4**:941–945.

46. Eijsink VG, van der Zee JR, van den Burg B, Vriend G, Venema G: **Improving the Thermostability of the Neutral Protease of *Bacillus stearothermophilus* by Replacing a Buried Asparagine by Leucine.** *FEBS Lett* 1991, **282**:13–16.

47. Hurley JH: **The Role of Interior Side-Chain Packing in Pro-**
• **tein Folding and Stability.** In *The Protein Folding Problem and Tertiary Structure Prediction.* Edited by Merz K, Le Grand S. Boston: Birkhauser; 1994:in press.
A review of the packing problem, with emphasis on mutational and structural analysis. Computational methods and the non-additivity of multiple-core mutants are also discussed.

48. Lim WA, Richards FM: **An Analysis of Packing in the Protein**
• **Folding Problem.** *Q Rev Biophys* 1993, **26**:423–498.
A comprehensive review of theoretical and experimental studies of geometric packing in proteins. A compendium of stability data for packing mutants is also included.

49. Behe MJ, Lattman EE, Rose GD: **The Protein-Folding Problem: The Native Fold Determines Packing, but Does Packing Determine the Native Fold?** *Proc Natl Acad Sci USA* 1991, **88**:4195–4199.

50. Schrauber H, Eisenhaber F, Argos P: **Rotamers: To Be or Not**
• **To Be? An Analysis of Amino Acid Side-Chain Conformations in Globular Proteins.** *J Mol Biol* 1993, **230**:592–612.
This paper analyzes side-chain torsion angles and their distributions in known structures, questioning the validity of the 'average rotamer' concept.

51. Tuffery P, Etchebest C, Hazout S, Lavery R: **A New Approach to the Rapid Determination of Protein Side-Chain Conformations.** *J Biomol Struct Dyn* 1991, **8**:1267–1289.

52. Dunbrack RL Jr, Karplus M: **Backbone-Dependent Rotamer Li-**
 • **brary for Proteins. Application to Side Chain Prediction.** *J Mol Biol* 1993, **230**:543–574.
The relationship between side-chain torsion angles and backbone conformation is investigated. This analysis shows that rotamer matching is improved if backbone torsional information is included.

53. Schiffer CA, Caldwell JW, Kollman PA, Stroud RM: **Prediction of Homologous Protein Structures Based on Conformational Searches and Energetics.** *Proteins* 1990, **8**:30–43.

54. Eisenmenger F, Argos P, Abagyan R: **A Method to Configure**
 • **Protein Side-Chains from the Main-Chain Trace in Homology Modelling.** *J Mol Biol* 1993, **231**:849–860.
See [55•].

55. Wilson C, Gregoret LM, Agard DA: **Modeling Side-Chain Con-**
 • **formation for Homologous Proteins Using an Energy-Based Rotamer Search.** *J Mol Biol* 1993, **229**:996–1006.
This paper and [54•,60] contain excellent discussions of the limitations of homology modeling as well as what can be done about them.

56. Laughton CA: **Prediction of Protein Side-Chain Conformation**
 • **from Local Three-Dimensional Homology Relationships.** *J Mol Biol* 1993, **235**:1088–1097.
This paper describes a strategy to predict packing based on arrangements observed in three-dimensional structures. A rigid backbone is assumed. This procedure is fundamentally different from the rotamer approach to packing because it incorporates context information from the database.

57. Summers NL, Karplus M: **Construction of Side-Chains in Homology Modelling. Application to the C-Terminal Lobe of Rhizopuspepsin.** *J Mol Biol* 1989, **210**:785–811.

58. Hurley JH, Baase WA, Matthews BW: **Design and Structural Analysis of Alternative Hydrophobic Core Packing Arrangements in Bacteriophage T4 Lysozyme.** *J Mol Biol* 1992, **224**:1143–1159.

59. Tuff'ery P, Lavery R: **Packing and Recognition of Protein Struc-**
 • **tural Elements: A New Approach Applied to the 4-Helix Bundle of Myohemerythrin.** *Proteins* 1993, **15**:413–425.
Describes the first method to predict side-chain configuration that incorporates rigid helical motions in energy minimization of a trial structure. These minimization steps are performed between side-chain enumeration steps. The method is used to correctly predict the orientations of three of four helices in myohemerythrin.

60. Lee C, Subbiah S: **Prediction of Protein Side-Chain Conformation by Packing Optimization.** *J Mol Biol* 1991, **217**:373–388.

61. Lee C: **Predicting Protein Mutant Energetics by Self-Consistent**
 • **Ensemble Optimization.** *J Mol Biol* 1994, **236**:918–939.
Presents qualitatively correct predictions of λ repressor mutant stabilities (see also [66,79,80]) using a rigid backbone model and a new method to sample side-chain configurations.

62. Desmet J, De Maeyer M, Hazes B, Lasters I: **The Dead-End Elimination Theorem and its Use in Protein Side-Chain Positioning.** *Nature* 1992, **356**:539–542.

63. van Gunsteren WF, Mark AE: **Prediction of the Activity and Stability Effects of Site-Directed Mutagenesis on a Protein Core.** *J Mol Biol* 1992, **227**:389–395.

64. Shi YY, Mark AE, Wang CX, Huang F, Berendsen HJ, van Gun-
 • steren WF: **Can the Stability of Protein Mutants be Predicted by Free Energy Calculations?** *Protein Eng* 1993, **6**:289–295.
This paper and [63,65] contain discussions of the accuracy and/or feasibility of current energy calculations on proteins.

65. Abagyan R, Argos P: **Optimal Protocol and Trajectory Visualization for Conformational Searches of Peptides and Proteins.** *J Mol Biol* 1992, **225**:519–532.

66. Lee C, Levitt M: **Accurate Prediction of the Stability and Activity Effects of Site-Directed Mutagenesis on a Protein Core.** *Nature* 1991, **352**:448–451.

67. Alber T, Dao-pin S, Nye JA, Muchmore DC, Matthews BW: **Temperature-Sensitive Mutations of Bacteriophage T4 Lysozyme Occur at Sites with Low Mobility and Low Solvent Accessibility in the Folded Protein.** *Biochemistry* 1987, **26**:3754–3758.

68. Terwilliger TC, Zabin HB, Horvath MP, Sandberg WS, Shlunk
 • PM: **In Vivo Characterization of Mutants of the Bacteriophage f1 Gene V Protein.** *J Mol Biol* 1994, **236**:556–571.
Complete segment survey of the relative mutability of residues in gene V protein by segment combinatorial mutagenesis.

69. Reidhaar Olson JF, Sauer RT: **Combinatorial Cassette Mutagenesis as a Probe of the Informational Content of Protein Sequences.** *Science* 1988, **241**:53–57.

70. Brunet AP, Huang ES, Huffine ME, Loeb JE, Weltman RJ, Hecht
 • MH: **The Role of Turns in the Structure of an α-Helical Protein.** *Nature* 1993, **364**:355–358.
Describes segment combinatorial mutagenesis of three residues in an interhelical loop of the four-helix bundle protein cytochrome b562. Of the 31 clones analyzed, all folded into native-like structures, as judged by their heme-binding activity *in vivo*.

71. Dao-pin S, Alber T, Baase WA, Wozniak JA, Matthews BW: **Structural and Thermodynamic Analysis of the Packing of Two α-Helices in Bacteriophage T4 Lysozyme.** *J Mol Biol* 1991, **221**:647–667.

72. Sandberg WS, Terwilliger TC: **Influence of Interior Packing and Hydrophobicity on the Stability of a Protein.** *Science* 1989, **245**:54–57.

73. Khorasanizadeh S, Peters ID, Butt TR, Roder H: **Folding and Stability of a Tryptophan containing Mutant of Ubiquitin.** *Biochemistry* 1993, **32**:7054–7063.

74. Mendel D, Ellman JA, Chang Z, Veenstra DL, Kollman PA, Schultz PG: **Probing Protein Stability with Unnatural Amino Acids.** *Science* 1992, **256**:1798–1802.

75. Wynn R, Richards FM: **Unnatural Amino Acid Packing Mutants**
 • **of Escherichia coli Thioredoxin Produced by Combined Mutagenesis/Chemical Modification Techniques.** *Protein Sci* 1993, **2**:395–403.
Describes the use of unnatural side chains to probe protein core packing and stability (see also [74]). A homologous series of alkyl side chains was introduced at the same position in thioredoxin by disulfide exchange.

76. Tsuji T, Chrunyk BA, Chen X, Matthews CR: **Mutagenic Analysis**
 • **of the Interior Packing of an α/β Barrel Protein. Effects on the Stabilities and Rates of the Subunit of Tryptophan Synthase.** *Biochemistry* 1993, **32**:5566–5575.
A study of thermodynamic and kinetic effects of single and double core mutants on protein folding.

77. Ishikawa K, Nakamura H, Morikawa K, Kanaya S: **Stabiliza-**
 • **tion of Escherichia coli Ribonuclease HI by Cavity-Filling Mutations Within a Hydrophobic Core.** *Biochemistry* 1993, **32**:6171–6178.
Reports the structures of two packing mutants with enhanced stability and increased packing efficiency. Substitutions (valine to isoleucine or leucine) are accommodated by only small adjustments, The replacement side chain in the more stable leucine variant is in a strained conformation.

78. Eijsink VG, Dijkstra BW, Vriend G, van der Zee JR, Veltman OR, van der Vinne B, van den Burg B, Kempe S: **The Effect of Cavity-Filling Mutations on the Thermostability of Bacillus stearothermophilus Neutral Protease.** *Protein Eng* 1992, **5**:421–426.

79. Lim WA, Sauer RT: **Alternative Packing Arrangements in the Hydrophobic Core of Lambda Repressor.** *Nature* 1989, **339**:31–36.

80. Lim WA, Sauer RT: **The Role of Internal Packing Interactions in Determining the Structure and Stability of a Protein.** *J Mol Biol* 1991, **219**:359–376.

81. Baldwin EP, Xu J, Hajiseyedjavadi O: **Construction and Func-**
 •• **tional Selection of a T4 Lysozyme Gene Library Randomly Mutagenized at Five Specific Sites.** In *Techniques in Protein Chemistry IV.* Edited by Angeletti R. New York: Academic Press; 1993:495–507.
See [82••].

82. Baldwin EP, Hajiseyedjavadi O, Baase WA, Matthews BW: **The**
 •• **Role of Backbone Flexibility in the Accommodation of Variants that Repack the Core of T4 Lysozyme.** *Science* 1993, **262**:1715–1718.

Structural and thermodynamic characterizations of eight repacking variants containing three to five substitutions (see also [81••]). The structures show that much of the accommodation of the repacked cores is due to shifts in the backbone. In contrast, relatively few changes are observed in the rotational angles of the substituted side chains (see [94•]).

83. Sandberg WS, Terwilliger TC: **Engineering Multiple Properties**
• **of a Protein by Combinatorial Mutagenesis.** Proc Natl Acad Sci USA 1993, 90:8367–8371.
A survey of single and double mutants in gene V protein showing that stability and activity changes in most mutants are additive. In the one exception, the two substituted residues are in direct contact, and the resulting double mutant is 3.0 kcal mol⁻¹ more stable than expected.

84. Wells JA: **Additivity of Mutational Effects in Proteins.** Biochemistry 1990, 29:8509–8517.

85. Horovitz A, Fersht AR: **Co-Operative Interactions during Protein Folding.** J Mol Biol 1992, 224:733–740.

86. Serrano L, Horovitz A, Avron B, Bycroft M, Fersht AR: **Estimating the Contribution of Engineered Surface Electrostatic Interactions to Protein Stability by Using Double-Mutant Cycles.** Biochemistry 1990, 29:9343–9352.

87. Varadarajan R, Richards FM: **Crystallographic Structures of Ribonuclease S Variants with Nonpolar Substitution at Position 13: Packing and Cavities.** Biochemistry 1992, 31:12315–12327.

88. Danishefsky AT, Housset D, Kim KS, Tao F, Fuchs J, Woodward
• C, Wlodawer A: **Crevice-Forming Mutants in the Rigid Core of Bovine Pancreatic Trypsin Inhibitor: Crystal Structures of F22A, Y23A, N43G, and F45A.** Protein Sci 1993, 2:577–587.
See [90•].

89. Kim KS, Tao F, Fuchs J, Danishefsky AT, Housset D, Wlodawer
• A, Woodward C: **Crevice-Forming Mutants of Bovine Pancreatic Trypsin Inhibitor: Stability Changes and New Hydrophobic Surface.** Protein Sci 1993, 2:588–596.
See [90•].

90. Buckle AM, Henrick K, Fersht AR: **Crystal Structural Analysis**
• **of Mutations in the Hydrophobic Cores of Barnase.** J Mol Biol 1993, 234:847–860.
This paper and [87,88•,89•,91] describe the structures of variants containing single-residue replacements in their cores.

91. Matsumura M, Wozniak JA, Dao-pin S, Matthews BW: **Structural Studies of Mutants of T4 Lysozyme that Alter Hydrophobic Stabilization.** J Biol Chem 1989, 264:16059–16066.

92. Lee B: **Estimation of the Maximum Change in Stability of Glob-**
• **ular Proteins Upon Mutation of a Hydrophobic Residue to Another of Smaller Size.** Protein Sci 1993, 2:733–738.
A rationalization of the stability of cavity-creating mutants based on differences in excluded volume of the folded and unfolded state.

93. Karpusas M, Baase WA, Matsumura M, Matthews BW: **Hydrophobic Packing in T4 Lysozyme Probed by Cavity-Filling Mutants.** Proc Natl Acad Sci USA 1989, 86:8237–8241.

94. Murphy ME, Fetrow JS, Burton RE, Brayer GD: **The Structure**
• **and Function of Omega Loop A Replacements in Cytochrome c.** Protein Sci 1993, 2:1429–1440.
Presents crystal structures of packing mutants from homologous replacements of a loop in cytochrome c. Substitutions that replace the core-packing residues in the mutant loop with those of the wild type result in a structure that more closely resembles wild-type cytochrome c.

95. Lim WA, Hodel A, Sauer RT, Richards FM: **Crystal Structure**
•• **of a Mutant Protein with Altered but Improved Hydrophobic Core Packing.** Proc Natl Acad Sci USA 1994, 91:423–427.
Reports the crystal structure of a stabilized triple mutant of λ repressor. Both an increase in packing density and preservation of wild-type torsion angles accompanied by backbone adjustments are observed (see also [82••]). Torsion angles differ from the predicted values [66].

96. Lesk AM, Chothia C: **How Different Amino Acid Sequences Determine Similar Protein Structures: The Structure and Evolutionary Dynamics of the Globins.** J Mol Biol 1980, 136:225–270.

97. Lesk AM, Chothia C: **The Response of Protein Structures to Amino Acid Sequence Changes.** Philos Trans R Soc London [A] 1986, 317:345–356.

98. Hilbert M, Bohm G, Jaenicke R: **Structural Relationships of Homologous Proteins as a Fundamental Principle in Homology Modeling.** Proteins 1993, 17:138–151.

99. Siezen RJ, de Vos WM, Leunissen JA, Dijkstra BW: **Homology Modelling and Protein Engineering Strategy of Subtilases, the Family of Subtilisin-Like Serine Proteinases.** Protein Eng 1991, 4:719–737.

100. Ring CS, Cohen FE: **Modeling Protein Structures: Construction**
• **and their Applications.** FASEB J 1993, 7:783–790.
See [101•].

101. Vriend G, Eijsink V: **Prediction and Analysis of Structure, Sta-**
• **bility and Unfolding of Thermolysin-Like Proteases.** J Comput Aided Mol Design 1993, 7:367–396.
This paper and [99,100•] review the applications of homology modeling to the prediction of structures within protein families.

102. Benner SA: **Catalysis: Design Versus Selection.** Science 1993, 261:1402–1403.

EP Baldwin and BW Matthews, Institute of Molecular Biology, Howard Hughes Medical Institute and Department of Physics, University of Oregon, Eugene, Oregon 97403, USA.