



Published in final edited form as:

*Biopolymers*. 2008 May ; 89(5): 479–485. doi:10.1002/bip.20890.

## From Helix–Coil Transitions to Protein Folding

**Harold A. Scheraga**

Baker Laboratory of Chemistry, Cornell University, Ithaca, NY 14853-1301

**Lila Gierasch**

### Abstract

An evolution of procedures to simulate protein structure and folding pathways is described. From an initial focus on the helix–coil transition and on hydrogen-bonding and hydrophobic interactions, our original attempts to determine protein structure and folding pathways were based on an experimental approach. Experiments on the oxidative folding of reduced bovine pancreatic ribonuclease A (RNase A) led to a mechanism by which the molecule folded to the native structure by a minimum of four different pathways. The experiments with RNase A were followed by development of a molecular mechanics approach, first, making use of global optimization procedures and then with molecular dynamics (MD), evolving from an all-atom to a united-residue model. This hierarchical MD approach facilitated probing of the folding trajectory to longer time scales than with all-atom MD, and hence led to the determination of complete folding trajectories, thus far for a protein containing as many as 75 amino acid residues. With increasing refinement of the computational procedures, the computed results are coming closer to experimental observations, providing an understanding as to how physics directs the folding process.

### Keywords

helix–coil transition; polyamino acids; collagen

## INTRODUCTION

After Pauling discovered the  $\alpha$ -helix and the  $\beta$ -sheet,<sup>1</sup> it was natural to assume that an understanding of the interactions that lead to these conformations could be gained by studying the changes in conformation between these regular structures and the unfolded, statistical-coil form. This led to a flurry of both experimental<sup>2–4</sup> and theoretical<sup>5–8</sup> activity to consider, primarily, the helix–coil transition, with a later similar theoretical treatment of the transition between an antiparallel  $\beta$ -sheet and the statistical coil.<sup>9</sup>

Elkan Blout played a very important role in the experimental studies of the helix–coil transition by developing methods to synthesize high-molecular-weight poly- $\alpha$ -amino acids,<sup>2,10–16</sup> to study their properties, including the helix–coil transition,<sup>17–21</sup> with optical rotation,<sup>22–29</sup> circular dichroism,<sup>30,31</sup> and infrared spectroscopy,<sup>32–38</sup> and by demonstrating the biological properties of some of these homopolymers, e.g., their antigenicity.<sup>39</sup> He also developed

© 2007 Wiley Periodicals, Inc.

Correspondence to: Harold A. Scheraga; has5@cornell.edu.

This article is dedicated to the memory of Elkan Blout, whose friendship and seminal experimental studies of biopolymers, including the helix–coil transition in polyamino acids, and models of collagen, are fondly remembered.

This article was originally published online as an accepted preprint. The “Published Online” date corresponds to the preprint version. You can request a copy of the preprint by emailing the Biopolymers editorial office at biopolymers@wiley.com

synthetic methods to study polyproline and models of the triple helix of collagen,<sup>31,40-48</sup> including the triple helix-to-coil transition,<sup>49</sup> and demonstrated the presence of cis peptide bonds preceding proline.<sup>50</sup> Elkan was thus a leader among several other eminent scientists in studies of the conformational properties of poly- $\alpha$ -amino acids.

The theoretical work on the helix-coil transition was based on the one-dimensional Ising model, for which a matrix treatment was used to formulate the partition function to calculate the thermodynamic properties of the transitions in homopolyamino acids. This theoretical approach was extended to treat copolymers, in addition to the original focus on homopolymers of amino acids, and thereby to acquire information about the tendencies of each of the 20 naturally occurring amino acids to take on the helix conformation at a given temperature.<sup>51</sup> Aside from this application, however, it was soon realized that the helix-coil transition is not a good model for conformational changes in globular proteins, because the one-dimensional Ising model does not capture the cooperative features, embodied in the interplay between short- and long-range interactions, of the folding/unfolding transition of globular proteins.<sup>52-57</sup>

The parallel development of molecular mechanics (MM) was designed to overcome the failure of the one-dimensional Ising model to treat the protein folding problem. During the development of MM, the helix-coil transition was treated as one example<sup>58-62</sup> of the valid application of the developing MM methodology. This article describes our early experimental approach, which then evolved into a theoretical approach to determine protein structure and protein-folding pathways. The experimental approach focused on bovine pancreatic ribonuclease A (RNase A), and the theoretical approach evolved from some early applications of MM to treat the properties of the  $\alpha$ -helix, and the extension of MM methodology to treat conformational changes in globular proteins, i.e., the protein-folding problem.

## EARLY STUDIES OF HELIX-COIL TRANSITIONS

The early statistical mechanical treatment of the helix-coil transition is summarized in a book that provided the various theoretical approaches and also reprinted the seminal papers in which these approaches were described.<sup>63</sup> The factors relating to the question of a phase transition in polyamino acids<sup>64</sup> and polynucleotides,<sup>65</sup> the kinetics of the helix-coil transition in polyamino acids,<sup>66-68</sup> and the role of hydrophobic interactions between nonpolar side chains in stabilizing  $\alpha$ -helices<sup>69,70</sup> received extensive early consideration.

Early MM studies were carried out to assess the factors stabilizing  $\alpha$ -helices,<sup>71,72</sup> and determining the relative preferences for left- and right-handedness<sup>73</sup> in the  $\alpha$ -helical forms of polyamino acids. With verification by experiment,<sup>74</sup> the MM method was extended to treat both the helical and coil forms, and hence the helix-coil transition in polyamino acids.<sup>58-62</sup> These studies were the forerunner for the application of MM to nonregular linear<sup>75</sup> and cyclic<sup>76</sup> oligopeptides, and fibrous<sup>77,78</sup> and globular<sup>79-82</sup> proteins.

## EARLY CONSIDERATION OF INTERRESIDUE INTERACTIONS, AND MOTIVATION FOR DEVELOPMENT AND APPLICATIONS OF MM

While proteins were known to be copolymers of amino acids linked by peptide bonds, it was the determination of the amino acid sequence of insulin by Sanger<sup>83</sup> that established that such copolymers are unbranched, and consist of unique amino acid sequences. At about the same time, Pauling proposed the  $\alpha$ -helix and  $\beta$ -sheet conformations by relying heavily on the presence of hydrogen bonds between the NH and CO groups of the backbone of the polypeptide chain.<sup>1</sup>

Recognizing that the unique properties of proteins also depend on the interactions among their side chains,<sup>84</sup> we investigated the effect of hydrogen bonds on the acid-base equilibria ( $pK_s$ ) of ionizable side-chain groups<sup>85</sup> and on limited proteolysis.<sup>86</sup> For example, if a tyrosyl-OH group is hydrogen-bonded to a glutamyl-COO<sup>-</sup> group, then the  $pK$  of the tyrosyl group will be raised, and that of the COOH group will be lowered by the free energy of formation of the hydrogen bond.<sup>85</sup> Therefore, identification of such “abnormal”  $pK$ 's provides a clue as to the existence of such a hydrogen bond. In addition, the free energy for disrupting such a hydrogen bond will have to be paid in hydrolyzing the peptide group of the backbone whose side-chain hydrogen bonds must be broken to liberate the peptide fragment.<sup>86</sup> Therefore, in contrast to the hydrolysis of, say, a simple dipeptide, which thermodynamically goes to completion, the hydrolysis of a peptide group when the side chain is hydrogen-bonded will not go to completion; i.e., as demonstrated in the thrombin-catalyzed hydrolysis of peptides from fibrinogen,<sup>87,88</sup> such a limited hydrolytic reaction catalyzed by a proteolytic enzyme can be involved in a reversible equilibrium. In other words, a proteolytic enzyme can catalyze the formation of a peptide bond, because the hydrogen bonds between the fragment and the rest of the protein provide favorable thermodynamics for peptide-bond synthesis.<sup>86</sup> Supporting experimental evidence for this view of limited proteolysis<sup>86</sup> was also provided by Seifter et al.<sup>89</sup>

This model of peptide-bond synthesis takes on added significance in considering the role of the ribosome in the synthesis of peptide bonds.<sup>90</sup> Gindulyte et al.<sup>90</sup> carried out quantum mechanical calculations on a group of 50 atoms chosen to represent the formation of the transition state (TS) during peptide-bond formation on the ribosome. They computed the activation energy for the process, and found that the hydrogen bonds that are formed between the 3' end of the A-site tRNA and atoms of the ribosomal environment, while translocating into the P-site, lowered the activation energy for formation of the TS. We may regard such hydrogen bonds as also providing the free energy of equilibrium stabilization of the newly formed peptide bond.

Hydrophobic interactions<sup>91,92</sup> can provide further stabilization of hydrogen bonds. For example, hydrophobic interactions, in which, say, a leucine side chain interacts with the nonpolar neck of a hydrogen-bonded glutamyl residue (as illustrated in Fig. 1 of Ref. 93) can strengthen such a lysyl ··· glutamyl hydrogen bond. The consequence of the earlier considerations is that hydrogen bonds can be located by potentiometric titrations that identify ionizable groups with abnormal  $pK_s$ . With associated physical-chemical and biochemical studies<sup>94</sup> to identify hydrogen-bonded pairs, such hydrogen bonds provide distance constraints on how the backbone may fold.

The foregoing provided the rationale to obtain distance constraints<sup>94</sup> to determine the folded structure of a protein before the advent of X-ray or NMR methods to determine protein structures. In particular, 3 out of 6 tyrosyl side chains, and 3 out of 11 side-chain carboxyl groups, of RNase A were found to have abnormal  $pK_s$ , which suggested the possible existence of three hydrogen bonds. In a series of chemical modifications (before the advent of recombinant DNA techniques) and physical-chemical and biochemical studies, three tyrosyl-aspartyl interactions (out of 19,800 ways to specifically pair 3 out of 6 tyrosyl side chains with 3 out of 11 side-chain carboxyl groups) were identified,<sup>94</sup> namely Tyr25 ··· Asp14, Tyr 92 ··· Asp 38, and Tyr 97 ··· Asp83. These interactions were confirmed by the subsequent determination of the crystal structure of RNase A.<sup>95</sup>

These three Tyr ··· Asp interactions, plus the location of four disulfide bonds in RNase A, and the suggestion<sup>96,97</sup> that His12, His119, and Lys41 are in the active site of this enzyme, constituted distance constraints that motivated the development of computational methodology<sup>98</sup> to calculate the three-dimensional structure of the protein, by making use of

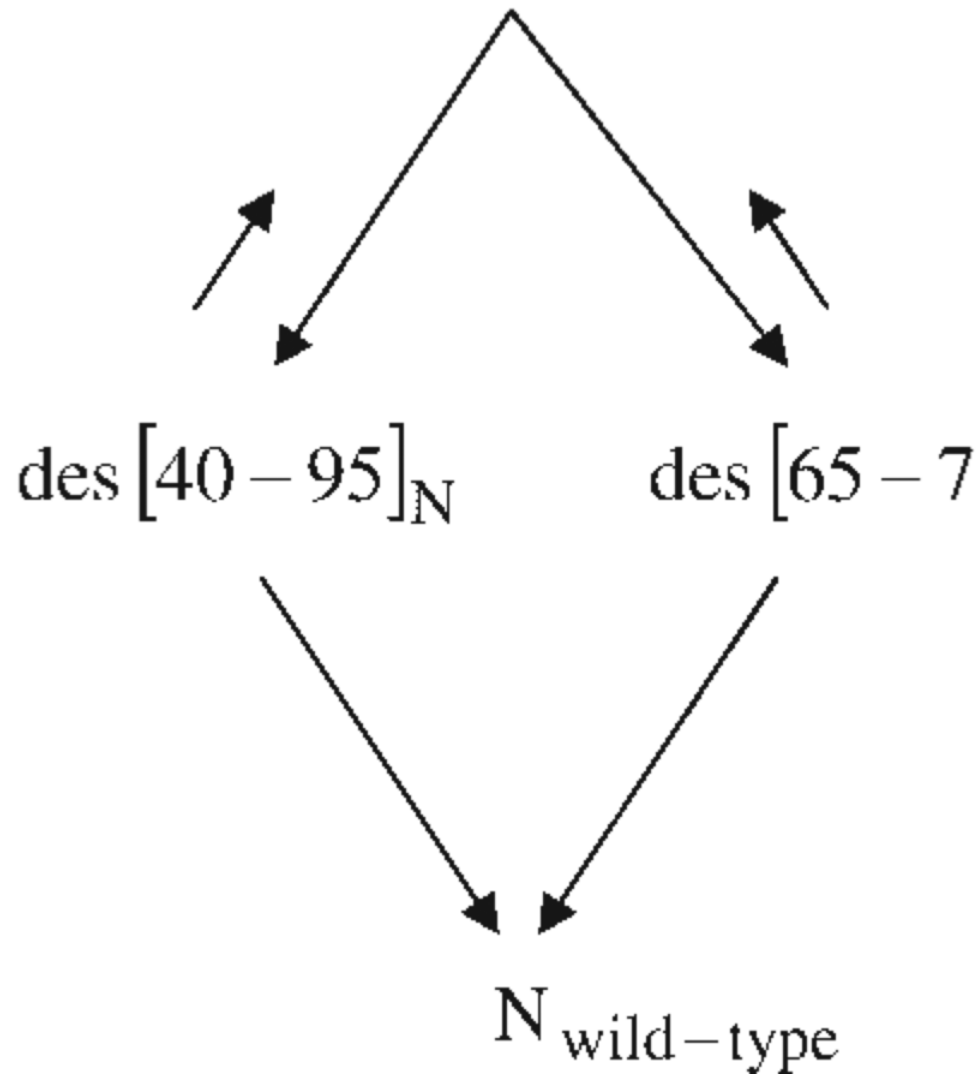
these distance constraints in the minimization of the empirical conformational energy of the protein. Without introducing the conformational energy and minimizing its energy, this limited number of distance constraints is not sufficient to determine the backbone structure. In fact, ~80 distance constraints, with half of them between residues  $i$  and  $j$  in the range of  $5 \leq |i - j| \leq 20$  and the other half with  $|i - j| \geq 21$ , would be required<sup>99</sup> in order to obtain the structure of a 57-residue protein with an RMSD  $< 2 \text{ \AA}$  with respect to the native structure if these distances are known exactly. If the distances in these two ranges are known only within an error no greater than  $\sim 2 \text{ \AA}$ , then more than  $\sim 150$  such distances are required. Additional information is available<sup>99</sup> about how the distributions of  $|i - j|$  distances and through-space distances  $d_{ij}$  affect the RMSD. Hence, the evolution of computational methodology in our laboratory proceeded in two different directions making use of (1) knowledge-based information (distance constraints, homology modeling, etc.) together with an empirical potential-energy function and (2) a physics-based potential-energy function without reliance on knowledge-based information.

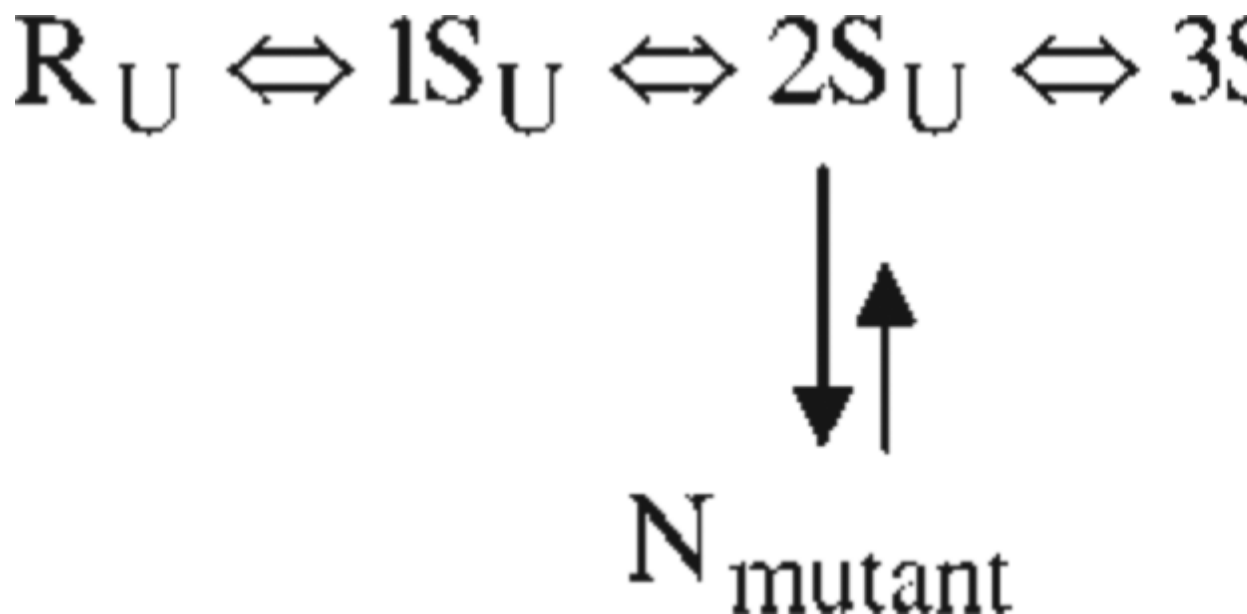
## EXPERIMENTAL DETERMINATION OF FOLDING PATHWAYS

When Anfinsen identified spontaneous protein folding and introduced the thermodynamic hypothesis for a theoretical approach,<sup>100</sup> we expanded our interest in determining protein structure to also determine folding pathways (first by experiment and later by theoretical methods).

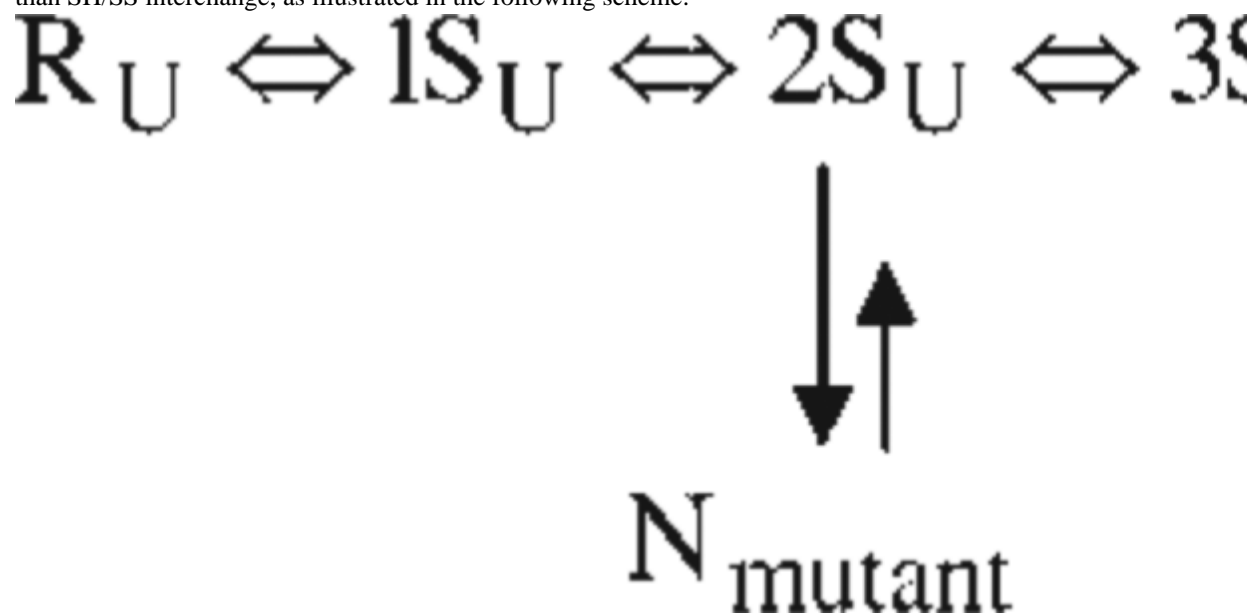
Our initial experimental work was carried out with insulin, the only protein whose amino acid sequence was known<sup>83</sup> at the time. However, experimental difficulties arose in the attempt to locate hydrogen bonds by potentiometric titration to determine pKs, because insulin is largely insoluble at neutral pH. But, as soon as the amino sequence of lysozyme was determined, our efforts turned to this protein which is soluble over the pH range of interest. However, lysozyme also presented experimental problems because of its relatively high tryptophan content. To locate hydrogen bonds by observing abnormal pKs of ionizable side-chain groups, it was necessary to carry out chemical modification and determine the amino acid composition of proteolytically obtained fragments. But, analysis for tryptophan was very difficult before the later development of improved methods for determining amino acid composition. Finally, with the availability of the amino acid sequence of RNase A, which is water-soluble and contains no tryptophan, it was possible to determine the kinetics and thermodynamics in the oxidative folding of RNase A, using either glutathione<sup>101</sup> or dithiothreitol<sup>102–104</sup> as the oxidizing agent.

With dithiothreitol as the redox reagent, the following mechanism was deduced for the oxidative folding of RNase A:



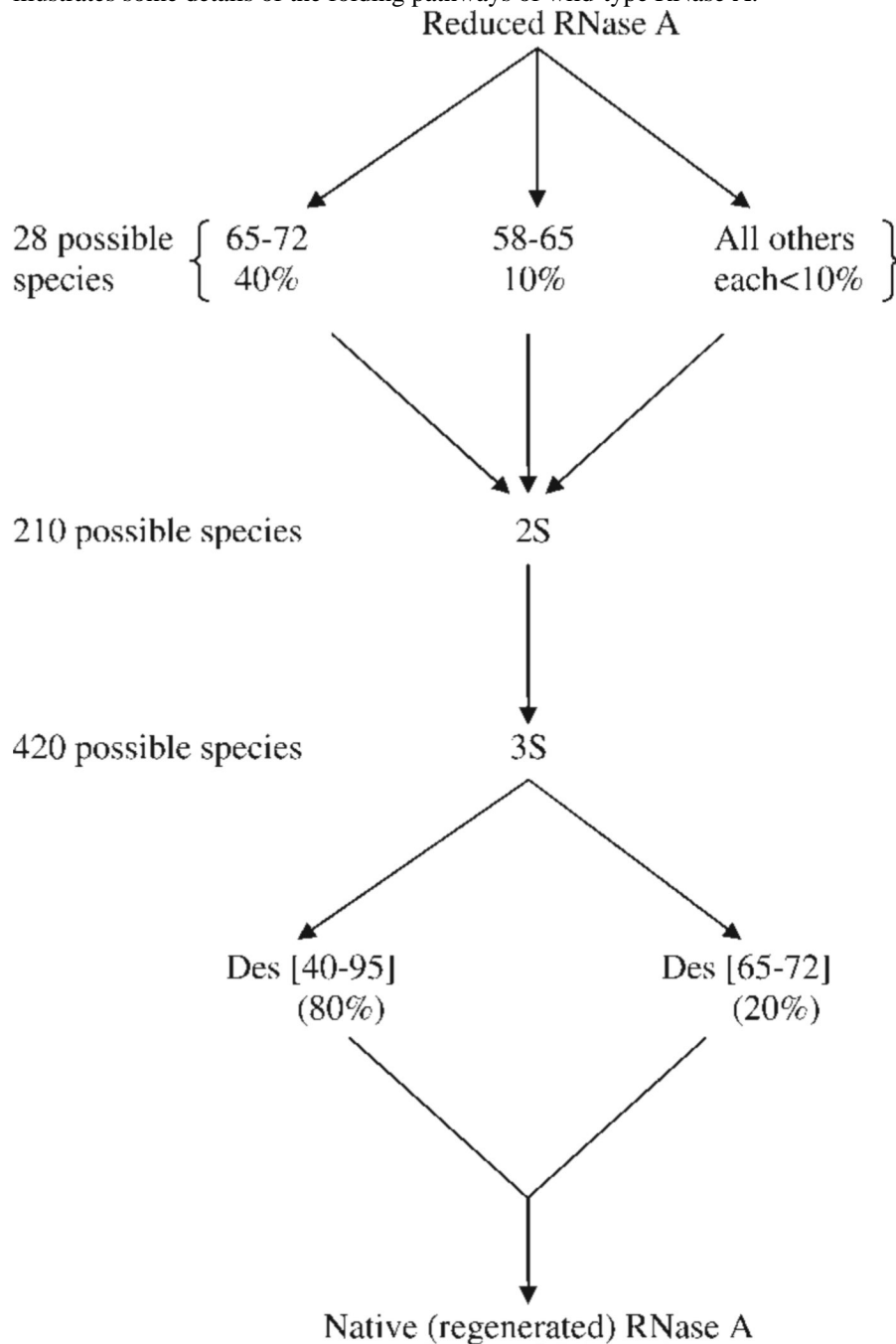


where  $R$  is the reduced protein, and  $1S$ – $4S$  are ensembles of one-, two-, three- and four-disulfide species, respectively (with the subscript  $U$  indicating that they are essentially unfolded species). The  $1S$ – $4S$  species are involved in a pre-equilibrium (quasi-steady-state condition). In the rate-determining step, the  $3S_U$  ensemble reshuffles by  $SH/SS$  interchange along two different pathways (in the approximate ratio 78:18) to form native-like des  $[40 - 95]_N$  and des  $[65 - 72]_N$ , respectively, both of which go rapidly to the native form. The remaining ~5% of the protein goes through two separate pathways, which could be detected only by mutating the Cys40, Cys95 and Cys65, Cys72 residues, respectively, to prevent formation of the corresponding disulfide bonds.<sup>105,106</sup> The rate-determining step in the oxidative folding of each of these two mutants involves oxidation from the  $2S_U$  ensemble to the native form, rather than  $SH/SS$  interchange, as illustrated in the following scheme:



Thus, these experiments identified a minimum of four different pathways for the oxidative folding of wild-type RNase A.

In the initial folding stages, i.e., up to formation of the 1S ensemble of wild-type RNase A, 40% of the 28 theoretically possible 1S forms have the 65–72 disulfide bond,<sup>107</sup> which persists to an even greater extent in the formation of the 2S ensemble.<sup>108</sup> The following scheme illustrates some details of the folding pathways of wild-type RNase A:



The preference of the 4:1 ratio for the native 65–72 disulfide bond relative to the non-native 58–65 disulfide bond is the same in the 1S folding stage of both the native protein and of an isolated 58–72 fragment.<sup>109,110</sup> Since both possible loops, 58–65 and 65–72, have the same number of residues and, hence, involve the same entropy of loop closure, this 4:1 ratio indicates that specific local interactions lead to the preference for formation of the 65–72 loop; such

preferential interactions in this loop have been identified by NMR<sup>111</sup> and MM<sup>112</sup> computations.

## COMPUTATIONS OF POLYPEPTIDE AND PROTEIN STRUCTURE WITH AN ALL-ATOM FORCE FIELD

The experimental work described earlier led to the development of MM<sup>98,113</sup> to compute protein structure. Using an all-atom representation of the polypeptide chain and an empirical potential-energy function, Empirical Conformational Energy Program for Peptides (ECEPP),<sup>114</sup> together with a variety of procedures for global optimization of the potential energy,<sup>115</sup> the low-energy (global-minimum) structures of various  $\alpha$ -helices,<sup>71-73</sup> several linear<sup>75</sup> and cyclic<sup>76</sup> peptides, and fibrous proteins,<sup>77,78</sup> were computed. As an example, one of these global optimization procedures, Electrostatically Driven Monte Carlo (EDMC),<sup>116</sup> together with an implicit hydration model, was used<sup>79</sup> in four simulations of the 46-residue three-helix bundle, protein A, providing structures within an RMSD of  $\sim 4$  Å from the known native structure. This was the largest protein whose all-atom structure had been computed at the time, and the calculation was carried out in 11 days with 60 processors of a Beowulf-type cluster. By contrast, a single trajectory on a smaller, 36-residue protein, the villin headpiece, simulated by molecular dynamics (MD),<sup>117</sup> required 4 months on a Cray computer. The villin headpiece was subsequently simulated<sup>80</sup> with the EDMC procedure.

## COMPUTATION OF FOLDING PATHWAYS WITH AN ALL-ATOM FORCE FIELD

The folding pathway of protein A<sup>118</sup> from an initial unfolded state  $i$  to a final folded state  $f$  was computed by use of a procedure, involving a stochastic difference equation (SDE), developed by Elber et al.<sup>119</sup> Both states  $i$  and  $f$  must be known in order to compute the pathways from  $i$  to  $f$  with the SDE procedure. The folded state  $f$ , of course, is either an X-ray or NMR structure of the protein. But the initial state  $i$  is not a single structure, but rather an ensemble of unfolded structures. Therefore, 130 representative structures were generated by MD; each of them was treated as an initial state  $i$ , and then all 130 trajectories were averaged.

In the SDE procedure, the action,  $S$ , defined as  $\int_i^f [2(E - U)]^{1/2} dl$ , is required to remain stationary along the trajectory length  $l$  from  $i$  to  $f$ , where  $E$  is the total energy, and  $U$  is the potential energy. From the computed results, it appeared that the C-terminal helix folded first, followed by the N-terminal helix, and then the middle helix. Contrary to expectations, it was found<sup>118</sup> that the radius of gyration did not decrease significantly until about halfway through the trajectory, i.e., an initial “hydrophobic collapse” was not observed. We have recently found<sup>120</sup> that the sequence of folding events is sensitive to the external conditions (temperature, friction, and random forces).

## COMPUTATION OF STRUCTURE AND FOLDING PATHWAYS BY A HIERARCHIAL PROCEDURE

Although an all-atom representation of the polypeptide chain could be used to compute the structure<sup>79</sup> and folding pathway<sup>118</sup> of the 46-residue protein A, it has not yet been possible to compute protein structures and folding pathways for proteins containing 100–200 residues with the all-atom representation. We have therefore developed a hierarchical approach in which the all-atom representation is initially replaced by a united-residue (UNRES) model,<sup>121</sup> and the search of the UNRES potential energy is carried out with a conformational space annealing (CSA) method.<sup>122</sup> The motivation of this search is to locate the *region* of the global minimum;



then the results of the UNRES search are converted to an all-atom representation,<sup>123,124</sup> and the search is continued in this *region* with the EDMC procedure.

The UNRES model is obtained by averaging over all “unimportant” fast degrees of freedom, thereby reducing the number of force centers and the number of degrees of freedom, and rendering the search procedure feasible for large proteins. The CSA search procedure starts with a representative set of local minima in the multidimensional conformation energy space, and reduces their distribution (to the *region* of the global minimum) by selected moves in what is essentially a genetic algorithm. The results have been subjected to blind tests<sup>82</sup> in successive (critical assessment of protein structure prediction procedures, CASP).

The predicted results have been sufficiently encouraging to motivate carrying out MD calculations with the UNRES model rather than with an all-atom one. Because all-atom MD trajectories require femtosecond moves to change conformation and, hence, do not reach the millisecond-to-second time-scale for folding all but the very fast-folding proteins, MD has not been successful in folding all-atom structures. However, because the fast degrees of freedom are averaged out in UNRES, it seemed possible that MD with UNRES might be able to simulate protein folding. Therefore, the mathematics for a Langevin-equation approach was formulated, and solved with a stochastic velocity Verlet algorithm.<sup>125,126</sup> This approach was successful in folding largely  $\alpha$ -helical single-chain<sup>81</sup> and two-chain<sup>127</sup> proteins, one containing as many as 75 residues, and in simulating folding kinetics.<sup>128</sup> To be able to treat  $\alpha+\beta$  and  $\beta$  proteins, it was found to be necessary to introduce entropic effects.<sup>129</sup>

## CONCLUSIONS

Motivated originally by the experimental acquisition of distance constraints in RNase A by physical–chemical and biochemical procedures, a MM approach was developed to compute the three-dimensional structures of proteins. Later, in order to focus on the physics of inter-residue interactions in determining protein structures, reliance on knowledge-based information, such as experimental distance constraints, was abandoned in favor of the use of only the potential energy. Early investigations were carried out by global optimization of an all-atom representation of the potential energy but, in order to treat large protein structures, this approach evolved into a hierarchy procedure based on a UNRES representation of the potential energy. Subsequently, UNRES was used to compute not only protein structure but also folding pathways by means of the Langevin equation. Further, improvement<sup>129</sup> of the force field by including entropic effects and focusing on reproduction of thermodynamic properties, such as heat capacity, has led to improved predictions of protein structure.

## Acknowledgments

Contract grant sponsors: NIH and NSF

## REFERENCES

1. Pauling L, Corey RB, Branson HR. Proc Natl Acad Sci USA 1951;37:205–211. [PubMed: 14816373]
2. Blout ER, Karlson RH. J Am Chem Soc 1956;78:941–946.
3. Doty P, Bradbury JH, Holtzer AM. J Am Chem Soc 1956;78:947–954.
4. Katchalski E, Sela M. Adv Protein Chem 1958;13:243–492. [PubMed: 14404711]
5. Schellman JA. Compt Rend Trav Lab Carlsberg Sér Chim 1955;29:223–259.
6. Zimm BH, Bragg JK. J Chem Phys 1959;31:526–535.
7. Lifson S, Roig A. J Chem Phys 1961;34:1963–1971.
8. Poland DC, Scheraga HA. J Chem Phys 1965;43:2071–2074.2074 [PubMed: 5845857] erratum: J Chem Phys 1965, 43, 3774.

9. Mattice WL, Scheraga HA. *Biopolymers* 1984;23:1701–1724. [PubMed: 6478039]
10. Blout ER, Karlson RH, Doty P, Hargitay B. *J Am Chem Soc* 1954;76:4492–4493.
11. Blout ER, Idelson M. *J Am Chem Soc* 1956;78:3857–3858.
12. Blout ER, Idelson M. *J Am Chem Soc* 1958;80:4909–4913.
13. Idelson M, Blout ER. *J Am Chem Soc* 1958;80:2387–2393.
14. Blout ER, DesRoches ME. *J Am Chem Soc* 1959;81:370–372.
15. Fasman GD, Blout ER. *J Am Chem Soc* 1960;82:2262–2267.
16. Fasman GD, Idelson M, Blout ER. *J Am Chem Soc* 1961;83:7099–712.
17. Doty P, Holtzer AM, Bradbury JH, Blout ER. *J Am Chem Soc* 1954;76:4492–4493.
18. Blout ER, Idelson M. *J Am Chem Soc* 1956;78:497–498.
19. Doty P, Wada A, Yang JT, Blout ER. *J Polym Sci* 1957;23:851–861.
20. Blout ER, Lenormant H. *Nature* 1957;179:960–963. [PubMed: 13430748]
21. Tomita K, Rich A, de Lozé C, Blout ER. *J Mol Biol* 1962;4:83–92. [PubMed: 13921759]
22. Blout ER, Doty P, Yang JT. *J Am Chem Soc* 1957;79:749–750.
23. Blout ER, Karlson RH. *J Am Chem Soc* 1958;80:1259–1260.
24. Idelson M, Blout ER. *J Am Chem Soc* 1958;80:4631–4634.
25. Blout, ER. *Proceedings of the fourth International Congress of Biochemistry*; Pergamon: London. 1959. p. 37-41.
26. Karlson RH, Norland KS, Fasman GD, Blout ER. *J Am Chem Soc* 1960;82:2268–2275.
27. Shechter E, Blout ER. *Proc Natl Acad Sci USA* 1964;51:695–702. [PubMed: 14166777]
28. Shechter E, Blout ER. *Proc Natl Acad Sci USA* 1964;51:794–800. [PubMed: 14172993]
29. Shechter E, Carver JP, Blout ER. *Proc Natl Acad Sci USA* 1964;51:1029–1036. [PubMed: 14215623]
30. Simmons NS, Cohen C, Szent-Gyorgyi AG, Wetlaufer DB, Blout ER. *J Am Chem Soc* 1961;83:4766–4769.
31. Blout ER, Carver JP, Gross J. *J Am Chem Soc* 1963;85:644–645.
32. Blout ER, Asadourian A. *J Am Chem Soc* 1956;78:955–961.
33. Idelson M, Blout ER. *J Am Chem Soc* 1957;79:3948–3955.
34. Blout ER. *Ann NY Acad Sci* 1957;69:84–93. [PubMed: 13478993]
35. Bird GR, Parrish M Jr. Blout ER. *Rev Sci Instr* 1958;29:305–309.
36. Bird GR, Blout ER. *J Am Chem Soc* 1959;81:2499–2503.
37. Miyazawa T, Blout ER. *J Am Chem Soc* 1961;83:712–719.
38. Blout ER, de Lozé C, Asadourian A. *J Am Chem Soc* 1961;83:1895–1900.
39. Maurer PH, Subrahmanyam D, Kathchalski E, Blout ER. *J Immunol* 1959;83:193–197.
40. Blout, ER.; Fasman, GD. *Recent Advances in Gelatin and Glue Research*. Pergamon; London: 1957. p. 122-130.
41. Fasman GD, Blout ER. *Biopolymers* 1963;1:3–14.
42. Oriol PJ, Blout ER. *J Am Chem Soc* 1966;88:2041–2045.
43. Carver, JP.; Blout, ER. *Treatise on Collagen*. Vol. 1. Academic Press; New York: 1967. p. 441-526.
44. Ramachandran GN, Doyle BB, Blout ER. *Biopolymers* 1968;6:1771–1775. [PubMed: 5704347]
45. Brown FR III, Carver JP, Blout ER. *J Mol Biol* 1969;39:307–313. [PubMed: 5391435]
46. Kivirikko KI, Prockop DJ, Lorenzi GP, Blout ER. *J Biol Chem* 1969;244:2755–2760. [PubMed: 5769999]
47. Lorenzi GP, Doyle BB, Blout ER. *Biochemistry* 1971;10:3046–3051. [PubMed: 5126922]
48. Doyle BB, Traub W, Lorenzi GP, Blout ER. *Biochemistry* 1971;10:3052–3060. [PubMed: 5126923]
49. Brown FR III, Hopfinger AJ, Blout ER. *J Mol Biol* 1972;63:101–115. [PubMed: 5016967]
50. Deber CM, Bovey FA, Carver JP, Blout ER. *J Am Chem Soc* 1970;92:6191–6198. [PubMed: 5505620]
51. Scheraga HA, Vila JA, Ripoll DR. *Biophys Chem* 2002;101–102:255–265.
52. Hao MH, Scheraga HA. *J Phys Chem* 1994;98:4940–4948.

53. Hao MH, Scheraga HA. *J Phys Chem* 1994;98:9882–9893.
54. Hao MH, Scheraga HA. *J Chem Phys* 1995;102:1334–1348.
55. Hao MH, Scheraga HA. *J Chem Phys* 1997;107:8089–8102.
56. Hao MH, Scheraga HA. *J Mol Biol* 1998;277:973–983. [PubMed: 9545385]
57. Hao MH, Scheraga HA. *Acc Chem Res* 1998;31:433–440.
58. Go N, Go M, Scheraga HA. *Proc Natl Acad Sci USA* 1968;59:1030–1037. [PubMed: 5240015]
59. Go M, Go N, Scheraga HA. *J Chem Phys* 1970;52:2060–2079. [PubMed: 5445420]
60. Go M, Go N, Scheraga HA. *J Chem Phys* 1971;54:4489–4503.
61. Go M, Hesselink FT, Go N, Scheraga HA. *Macromolecules* 1974;7:459–467. [PubMed: 4418767]
62. Go M, Scheraga HA. *Biopolymers* 1984;23:1961–1977. [PubMed: 6498288]
63. Poland, D.; Scheraga, HA. *Theory of Helix-Coil Transitions in Biopolymers*. Academic Press; New York: 1970.
64. Poland D, Scheraga HA. *J Chem Phys* 1966;45:1456–1463. [PubMed: 5920200]
65. Poland D, Scheraga HA. *J Chem Phys* 1966;45:1464–1469. [PubMed: 5920201]
66. Schwarz G. *J Mol Biol* 1965;11:64–77. [PubMed: 14255761]
67. McQuarrie DA, McTague JP, Reiss H. *Biopolymers* 1965;3:657–663. [PubMed: 5852527]
68. Poland D, Scheraga HA. *J Chem Phys* 1966;45:2071–2090. [PubMed: 5975749]
69. Poland DC, Scheraga HA. *Biopolymers* 1965;3:315–334.
70. Poland DC, Scheraga HA. *Biopolymers* 1965;3:335–355.
71. Ooi T, Scott RA, Vanderkooi G, Scheraga HA. *J Chem Phys* 1967;46:4410–4426. [PubMed: 6043847]
72. Yan JF, Vanderkooi G, Scheraga HA. *J Chem Phys* 1968;49:2713–2726. [PubMed: 5682471]
73. Yan JF, Momany FA, Scheraga HA. *J Am Chem Soc* 1970;92:1109–1115.
74. Erenrich EH, Andreatta RH, Scheraga HA. *J Am Chem Soc* 1970;92:1116–1119.
75. Li Z, Scheraga HA. *J Mol Struct (Theochem)* 1988;179:333–352.
76. Dygert M, Go N, Scheraga HA. *Macromolecules* 1975;8:750–761. [PubMed: 53324]
77. Miller MH, Scheraga HA. *J Polym Sci: Polym Symp* 1976;54:171–200.
78. Fossey SA, Némethy G, Gibson KD, Scheraga HA. *Biopolymers* 1991;31:1529–1541. [PubMed: 1814502]
79. Vila JA, Ripoll DR, Scheraga HA. *Proc Natl Acad Sci USA* 2003;100:14812–14816. [PubMed: 14638943]
80. Ripoll DR, Vila JA, Scheraga HA. *J Mol Biol* 2004;339:915–925. [PubMed: 15165859]
81. Liwo A, Khalili M, Scheraga HA. *Proc Natl Acad Sci USA* 2005;102:2362–2367. [PubMed: 15677316]
82. Oldziej S, Czaplowski C, Liwo A, Chinchio M, Nancias M, Vila JA, Khalili M, Arnautova YA, Jagielska A, Makowski M, Schafroth HD, Kazmierkiewicz R, Ripoll DR, Pillardy J, Saunders JA, Kang YK, Gibson KD, Scheraga HA. *Proc Natl Acad Sci USA* 2005;102:7547–7552. [PubMed: 15894609]
83. Sanger, F. *Currents in Biochemical Research*. Green, DE., editor. Interscience; New York: 1956. p. 434–459.
84. Neurath H, Greenstein JP, Putnam FW, Erickson JO. *Chem Rev* 1944;34:157–265.
85. Laskowski M Jr. Scheraga HA. *J Am Chem Soc* 1954;76:6305–6319.
86. Laskowski M Jr. Scheraga HA. *J Am Chem Soc* 1956;78:5793–5798.
87. Laskowski M Jr. Donnelly TH, Van Tijn BA, Scheraga HA. *J Biol Chem* 1956;222:815–821. [PubMed: 13367049]
88. Laskowski M Jr. Ehrenpreis S, Donnelly TH, Scheraga HA. *J Am Chem Soc* 1960;82:1340–1348.
89. Seifter S, Gallop PM, Klein L, Meilman E. *J Biol Chem* 1959;234:285–293. [PubMed: 13630896]
90. Gindulyte A, Bashan A, Agmon I, Massa L, Yonath A, Karle J. *Proc Natl Acad Sci USA* 2006;130:13327–13332. [PubMed: 16938893]
91. Némethy G, Scheraga HA. *J Phys Chem* 1962;66:1773–1789.
92. Griffith JH, Scheraga HA. *J Mol Struct* 2004;682:97–113.

93. Némethy G, Steinberg IZ, Scheraga HA. *Biopolymers* 1963;1:43–69.
94. Scheraga HA. *Fed Proc* 1967;26:1380–1387. [PubMed: 6051318]
95. Wlodawer A, Svensson LA, Sjölin L, Gilliland GL. *Biochemistry* 1988;27:2705–2717. [PubMed: 3401445]
96. Hirs, CHW.; Halmann, M.; Kycia, JH. *Biological Structure and Function*. Goodwin, TW.; Lindberg, O., editors. Vol. 1. Academic Press; New York: 1961. p. 41-57.
97. Heinrikson L. *J Biol Chem* 1966;241:1393–1405. [PubMed: 5935351]
98. Némethy G, Scheraga HA. *Biopolymers* 1965;3:155–184.
99. Wako H, Scheraga HA. *Macromolecules* 1981;14:961–969.
100. Anfinsen CB. *Science* 1973;181:223–230. [PubMed: 4124164]
101. Scheraga HA, Konishi Y, Ooi T. *Adv Biophys* 1984;18:21–41. [PubMed: 6399820]
102. Rothwarf DM, Li Y-J, Scheraga HA. *Biochemistry* 1998;37:3760–3766. [PubMed: 9521695]
103. Rothwarf DM, Li Y-J, Scheraga HA. *Biochemistry* 1998;37:3767–3776. [PubMed: 9521696]
104. Scheraga, HA.; Wedemeyer, WJ.; Welker, E. *Methods in Enzymology*. Nicholson, AW., editor. Vol. 341. Academic Press; San Diego: 2001. p. 189-221.
105. Iwaoka M, Juminaga D, Scheraga HA. *Biochemistry* 1998;37:4490–4501. [PubMed: 9521769]
106. Xu X, Scheraga HA. *Biochemistry* 1998;37:7561–7571. [PubMed: 9585571]
107. Xu X, Rothwarf DM, Scheraga HA. *Biochemistry* 1996;35:6406–6417. [PubMed: 8639587]
108. Volles MJ, Xu X, Scheraga HA. *Biochemistry* 1999;38:7284–7293. [PubMed: 10353840]
109. Milburn PJ, Scheraga HA. *J Protein Chem* 1988;7:377–398. [PubMed: 3255371]
110. Altmann KH, Scheraga HA. *J Am Chem Soc* 1990;112:4926–4931.
111. Talluri S, Falcomer CM, Scheraga HA. *J Am Chem Soc* 1993;115:3041–3047.
112. Carty RP, Pincus MR, Scheraga HA. *Biochemistry* 2002;41:14815–14819. [PubMed: 12475229]
113. Scheraga HA. *Adv Phys Org Chem* 1968;6:103–184.
114. Némethy G, Gibson KD, Palmer KA, Yoon CN, Paterlini G, Zagari A, Rumsey S, Scheraga HA. *J Phys Chem* 1992;96:6472–6484.
115. Scheraga HA, Lee J, Pillardy J, Ye Y-J, Liwo A, Ripoll J. *Glob Optimiz* 1999;15:235–260.
116. Ripoll DR, Liwo A, Scheraga HA. *Biopolymers* 1998;46:117–126. [PubMed: 9664845]
117. Duan Y, Kollman PA. *Science* 1998;283:740–744. [PubMed: 9784131]
118. Ghosh A, Elber R, Scheraga HA. *Proc Natl Acad Sci USA* 2002;99:10394–10398. [PubMed: 12140363]
119. Elber R, Ghosh A, Cárdenas AA. *Acc Chem Res* 2002;35:396–403. [PubMed: 12069624]
120. Jagielska A, Scheraga HA. *J Comput Chem* 2007;28:1068–1082. [PubMed: 17279497]
121. Liwo A, Czaplewski C, Pillardy J, Scheraga HA. *J Chem Phys* 2001;115:2323–2347.
122. Lee J, Scheraga HA, Rackovsky S. *J Comput Chem* 1997;18:1222–1232.
123. Kazmierkiewicz R, Liwo A, Scheraga HA. *J Comput Chem* 2002;23:715–723. [PubMed: 11948589]
124. Kazmierkiewicz R, Liwo A, Scheraga HA. *Biophys Chem* 2003;100:261–280.280 [PubMed: 12646370] Erratum: *Biophys Chem* 2003, 106, 91.
125. Khalili M, Liwo A, Rakowski F, Grochowski P, Scheraga HA. *J Phys Chem B* 2005;109:13785–13797. [PubMed: 16852727]
126. Khalili M, Liwo A, Jagielska A, Scheraga HA. *J Phys Chem B* 2005;109:13798–13810. [PubMed: 16852728]
127. Rojas AV, Liwo A, Scheraga HA. *J Phys Chem B* 2007;111:293–309. [PubMed: 17201452]
128. Khalili M, Liwo A, Scheraga HA. *J Mol Biol* 2006;355:536–547. [PubMed: 16324712]
129. Liwo A, Khalili M, Czaplewski C, Kalinowski S, Oldziej S, Wachucik K, Scheraga HA. *J Phys Chem B* 2007;111:260–285. [PubMed: 17201450]