Proc. Natl. Acad. Sci. USA
Vol. 76, No. 12, pp. 6050–6054, December 1979
Chemistry

A possible folding pathway of bovine pancreatic RNase

(protein conformation/contact map/long-range interactions/nucleation of folding)

GEORGE NÉMETHY AND HAROLD A. SCHERAGA*

Baker Laboratory of Chemistry, Cornell University, Ithaca, New York 14853

Contributed by H. A. Scheraga, September 26, 1979

ABSTRACT **A theoretical pathway for the folding of RNase into its native conformation is derived from the contact map computed from crystallographic coordinates. The pathway is based on the hypothesis of Tanaka and Scheraga, according to which localized conformations stabilized by short- and medium-range interactions form before those conformational features that are stabilized primarily by long-range interactions. The pathway deduced from the contact map agrees with experimental information on intermediates detected in the thermal unfolding of RNase and in immunochemical studies on the formation of stable antigenic sites when reduced RNase is oxidized with glutathione. Ambiguities in the interpretation of the contact map are resolved by the combination of structural information contained in the contact map and experimental information.**

Contact maps (1–5) [and related, more quantitative representations, distance maps (6–10)] have been used to analyze the structures of native proteins (2, 6–8) and to compare the results of simulations of protein folding to the native structure (1, 3–5, 9). Some studies have dealt with the description of ordered structures, such as α-helices and parallel or antiparallel pleated sheets on the contact map (7, 8) or with the recognition of large structural domains (2, 10). Although most of these studies concerned the native structure, we recently exploited the results of Monte Carlo calculations (1) and introduced a hypothetical mechanism of protein folding, based on a qualitative visual analysis of contact maps (2). It was shown that a self-consistent pathway of folding can be postulated for several proteins on the basis of the contact map alone.

In this communication, we show that the pathway for the folding of RNase, derived from the contact map, is consistent with the one proposed on the basis of experimental information on intermediates during both thermal unfolding (with intact disulfide bonds) and glutathione oxidation of the reduced protein. Theoretical analysis and the interpretation of experimental data complement each other by narrowing the possible choices of pathways. Experimental data on the unfolding of RNase and the arguments that support the concept of a preferred pathway have been reviewed elsewhere (11–13).

METHODS

The contact map shown in Fig. 1 is based on the crystal structure of RNase S (14, 15). The coordinate listing obtained from the Protein Data Bank (Brookhaven National Laboratories, Upton, NY) was used.

Construction of the Contact Map. The contact map was constructed according to the method described by Tanaka and Scheraga (2). Two residues i and j are defined to be in contact if at least one pair of their atoms (or groups) is separated by a distance that is smaller than the diameter of a water molecule (see equation 1 of ref. 1). Contact between residues i and j is indicated by a square on the map (Fig. 1). The heavy horizontal and vertical lines in the map separate contacts between various parts of the protein molecule: (*i*) within the S-protein (in the large triangle in the lower right portion); (*ii*) within the S-peptide (in the small triangle in the upper left portion); and (*iii*) between the S-peptide and the S-protein (within the rectangle in the lower left portion). Because of the close similarity of the structures of RNase S (14, 15) and RNase A (16), essentially the same map would apply to the latter protein, with small modifications near residues 20 and 21.

Specification of Contact Regions. Following Tanaka and Scheraga (2), a contact region is defined on the contact map as either one contiguous set of contacts or several nearby contiguous sets separated from each other by only small gaps on the map, if these sets can be distinguished clearly from other contact regions. The distinction may be based on the distance or the difference in shape from neighboring contact regions. For example, region L is definitely separated from all other regions. The antiparallel structure (7, 8) represented by contact region H differs from the more irregular packing of segments of the polypeptide chain in region I. The subdivision of the contact map into contact regions has been carried out by visual inspection. In most cases, contact regions can be defined unambiguously, both in RNase S (Fig. 1) and in other proteins (e.g., see figures 2–4 of ref. 2). There are only a few locations in the contact map of RNase S where the position of a boundary line is uncertain (e.g., between regions D and H or between regions B and G). In these cases, the boundaries were drawn in such a manner that the shapes of the resulting regions would be as simple as possible.

RESULTS

Pathway of Folding Deduced from the Contact Map. A pathway of folding was deduced from Fig. 1 by using the hypothesis of Tanaka and Scheraga (2). This hypothesis assumes that short-range interactions take precedence over medium-range interactions, and medium-range interactions take precedence over long-range interactions. This criterion is satisfied by pathways in which contacts near the diagonal of the contact map are formed first, and contact regions are then formed in the order of increasing distance from the diagonal. The contact regions can be ordered by using the point with the smallest values of $|i - j|$ in each contact region (2), designated as Min$|i - j|$, as an index. According to the hypothesis, this index uniquely defines the order in which contact regions are formed, except for regions with the same value of Min$|i - j|$. The values of Min$|i - j|$ of all contact regions of Fig. 1 are listed in Table 1.

The rigorous application of the hypothesis of Tanaka and Scheraga (2) leads to the following pathway of folding of RNase. The initial steps consist of the formation of one or several of six localized conformations, A–F, mainly under the influence of

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

* To whom requests for reprints should be addressed.

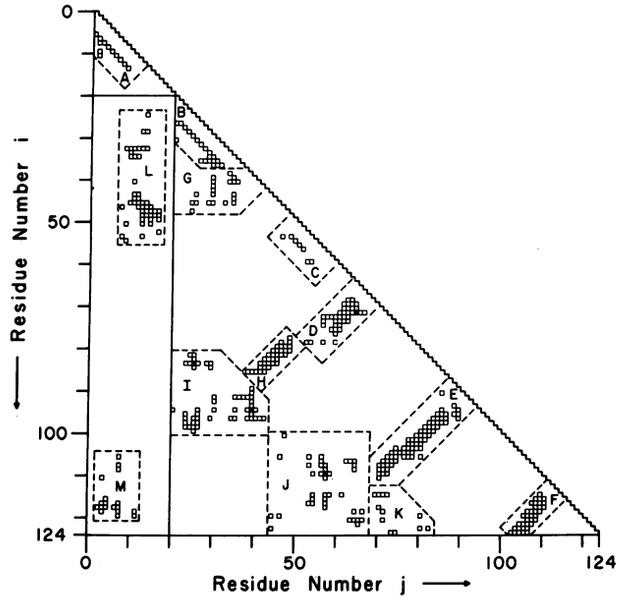

FIG. 1. Contact map of RNase S, constructed by the method of ref. 2. Each point of the map represents the presence (square) or absence (no marking) of a contact between two amino acid residues i and j. A contact is defined according to ref. 1. Contacts between residues are omitted from the figure whenever $|i-j| \leq 4$. The pairs of half-cystine residues forming the disulfide bridges are denoted by black squares. See the text for an explanation of the heavy vertical and horizontal lines between residues 20 and 21 and for the definition of contact regions. Contact regions (A–M) are bounded by dashed lines.

short- and medium-range interactions: α-helices from residues 4 to 11 (region A), 25 to 34 (region B), 51 to 57 (region C), and antiparallel structures (not necessarily pleated sheets) formed by the association of residues 53 to 67 with residues 69 to 79 (region D), residues 71 to 90 with residues 91 to 111 (region E), and residues 103 to 111 with residues 115 to 124 (region F). These nuclei do not have to form simultaneously, and the contact map does not furnish any means to decide which of these regions form first (see *Discussion*). The subsequent steps along the folding pathway consist of the growth or the coalescence of these nuclei. Region G is formed when residues 36–48 fold against helix B, in a manner indicated by computations of interactions between α-helices and neighboring unordered segments (17). Min$|i-j| = 5$ for region G. Region G, however, contains longer-range contacts than does helix B (Max$|i-j| = 10$ and 23, for regions B and G, respectively). Therefore, G is considered to form at a later stage than B.

Table 1. Width of contact regions* in RNase S, expressed in terms of the range of $|i-j|$

Region	Min $\lvert i-j\rvert$	Max $\lvert i-j\rvert$
A	5	9
B	5	10
C	5	7
D	5	26
E	5	40
F	5	21
G	5	23
H	29	47
I	50	75
J	42	78
K	39	50
L	11	46
M	94	115

* Min $|i-j|$ is the value of $|i-j|$ for the contact closest to the diagonal in each contact region (2). Max $|i-j|$ is the corresponding quantity for the contact furthest from the diagonal.

The contact region with the next-lowest value of Min$|i-j|$, 11, is region L, formed by the association of region A with B, G, and C. This is followed by: region H (Min$|i-j| = 29$), formed when regions C and D come into contact; region K (Min$|i-j| = 39$), containing contacts between regions E and F; region J (Min$|i-j| = 42$) in which regions C and D form contacts with regions E and F; region I (Min$|i-j| = 50$) containing contacts of regions E and H with G. Finally, region M (Min$|i-j| = 94$) is formed when region A comes into contact with E and F.

The pathway just outlined must be modified if one takes into account that the S-protein of RNase S can fold to a significant extent in the absence of the S-peptide (18, 19). Regions L and M represent the association between the S-protein and the S-peptide. Therefore, they occur presumably as the last step in the folding of RNase S. It is reasonable to assume that they are the last steps of folding of RNase A as well, because of the similarity of the two structures.

As a result, the following modified pathway can be proposed, based only on considerations of the contact map (and this information about the folding of S-protein). One or more of the local contact regions B, C, D, E, and F represent the nucleation site(s) for the folding of RNase. These regions do not have to form simultaneously. One of them might be the primary nucleation site. After the formation of some or all local contact regions, further contact regions form by growth or association of the local contact regions: region G forms by growth of region B. Regions H and K form by association of regions C and D and of E and F, respectively. On the basis of the contact map alone, it is not clear which of regions G, H, and K is established first. These steps, however, must precede the formation of contact regions I and J. Region I arises from the association of region G with parts of H and E, and region J arises from association of C and D with E and F. This completes the structure of the S-protein or the corresponding part (residues 21–124) of RNase A. After the α-helix (region A) of residues 4–11 forms, it folds against (or it associates with) regions C and B plus G, and E and F, forming the new contact regions L and M, respectively, and thereby completing the folding of the molecule. The sequence of events described here is summarized in a schematic form in Fig. 2.

Comparison of the Proposed Pathway with Other Experimental and Theoretical Studies. The pathway proposed here is consistent with the experimental information available on possible intermediates in the thermal unfolding of RNase A with intact disulfide bonds and in glutathione oxidation of the reduced protein.

Matheson and Scheraga (20) introduced a theoretical method for the prediction of the primary nucleation site in the renaturation of unfolded proteins. Based on the hypothesis that the nucleation site is the local hairpin-like pocket with the lowest (i.e., most favorabole) free energy of hydrophobic interactions, they proposed that residues 106–118 form the primary nucleation site in RNase. This prediction agrees with an observed bend (14, 15) at residues 113 and 114, with the very high degree of conservation of nonpolar residues in the predicted nucleation site of 23 mammalian species of this protein, (20), and with immunochemical experiments on the folding of the protein (19). In the immunochemical work, it was found that the antigenic site in segment 87–104 folded before the other antigenic sites (in segments 1–10, 40–61, and 63–75)—i.e., a segment of the protein in the sequence 87–118 folds before other parts. This observation is consistent with the contact map of Fig. 1, which

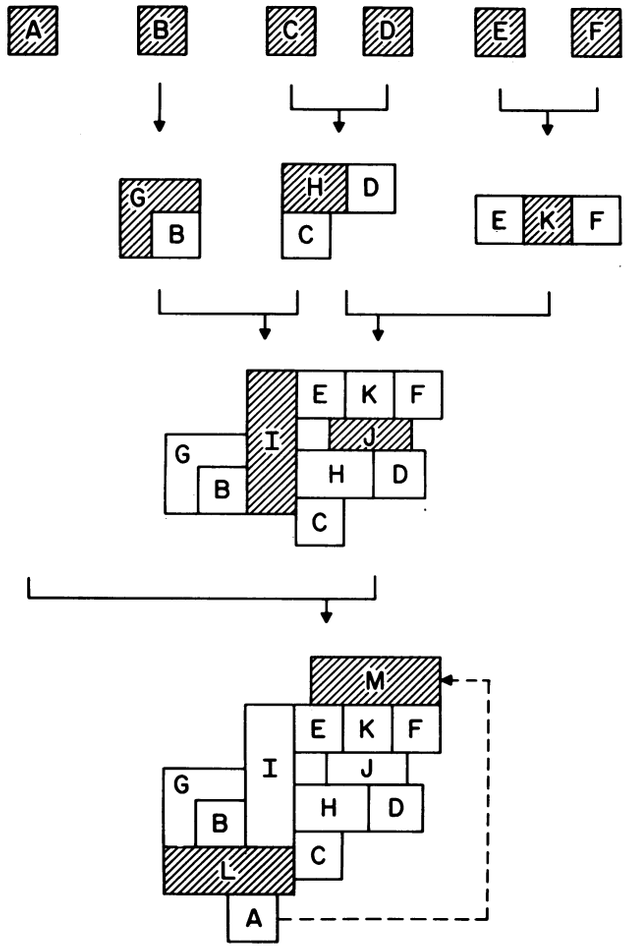

FIG. 2. Diagram (2) of the pathway of folding of RNase S. The letters correspond to the contact regions shown in Fig. 1. Folding is postulated to take place (by the association of contact regions to form new contact regions) from top to bottom of the figure. The hatched regions are the new ones formed at each stage. The formation of each contact region must precede the formation of the ones shown below it in the figure and connected to it by arrows. No information is obtained, however, from the contact map about the temporal sequence of formation of contact regions elsewhere in the diagram. The dashed line indicates a contact (in three dimensions) that cannot be represented easily in this two-dimensional diagram. The open space between regions *I* and *J* indicates the absence of contacts between these two regions and it has no significance in the three-dimensional structure.

shows that segment 70–124 of the protein can fold independently of the rest of the protein; i.e., this is an independent domain. Contacts formed *within* this sequence appear on the map in the triangular region to the right of a vertical line (not shown in Fig. 1) at $j = 69$. The proposed first nucleation site (20) is contained in contact region *F*. Nucleation, represented by the formation of all or part of region *F*, is followed by further folding of this segment of the chain, to form contact region *E*, to enlarge region *F*, and to form region *K*, thereby forming the antigenic site.

Burgess and Scheraga (21) proposed a six-stage pathway for the thermally induced unfolding of RNase (with intact disulfide bonds), based on the assessment of various experimental information and on diverse physical and chemical properties. This pathway was subsequently modified slightly (19, 22) to take into account the results of photochemical surface labeling experiments (22) and of immunochemical studies (19). The modified pathway, derived from experimental data, is summarized in Table 2. Table 2 also lists those contact regions whose formation (or dissolution) corresponds to each stage of folding (or unfolding). The ordering of the contact regions in Table 2 can be compared with that in Fig. 2. Because of the scarcity of experimental data on intermediates, it is not possible to cite an experimentally observed stage for each of the 13 contact regions. It is seen, however, that the sequence of unfolding events deduced from experiment and the pathway derived from the contact map alone—i.e., from the structure of the native molecule—are in complete agreement.

DISCUSSION

Use and Limitations of the Contact Map. The contact map of a protein, derived from the x-ray crystallographic structure, contains detailed information about the conformation of the native molecule. It does not tell anything directly about the intermediate stages during folding of the protein. Nevertheless, it was shown above (and in earlier papers 1,2) that the analysis of the distribution of contact regions can yield some information about the folding process. This utilization of the contact map to derive a pathway of folding, however, is subject to some uncertainties.

(*i*) There may be ambiguities in the manner in which the contact map is subdivided into contact regions if the regions of the map with dense distribution of contacts are not distinctly separated from each other. There are very few such uncertainties in the case of RNase, as mentioned earlier, and these do not influence the results reported here.

(*ii*) The contact regions shown in the map are those of the *native* molecule. It is likely that they are fairly stable structures because of the interactions implied by the presence of many contacts between residues within a given region. Nevertheless, it cannot be taken for granted that they also correspond to

Table 2. Comparison of pathways for folding RNase A derived from experimental studies (19–22) and from the contact map

	Experimental*	Contact map
Stage	Folding of residues	Formation of contact region†
VI	104–120	*F*
	81–102	*E, K*
V	1–12	*A*
	35–50	*B, G*
	62–74	*D*
IV	51–60	*C, (H)*
	121–124	Growth of *F*
III	27–34	Growth of *G*
	75–80	*J*‡
II	13–25	*(I), L*
	1–12§	*L, M*
I	92¶	—

* The Roman numerals refer to successive stages of *unfolding*, as proposed by Burgess and Scheraga (21), and modified by Matheson and Scheraga (22) and Chavez and Scheraga (19).

† No segment listed in the experimentally obtained stages is related directly to contact region *H* (involving contacts of residues 39–49 with residues 78–86). See the text for a possible assignment (indicated in parentheses) of region *H* to stage IV or III. Similar arguments place contact region *I* (involving contacts of residues 21–43 with residues 82–100) in stage III or II.

‡ Region *J* itself does not contain any contacts involving residues 75–80. However, the contacts between residues 45–67 and 100–123, contained in this region, can form when the molecule folds at residues 75–80.

§ The exposure of these residues changes in this stage (21).

¶ Unfolding of the side chain without alteration of the backbone. As a result, a few contacts may be altered in region *I*, but there is no gross change of the contact map.

folding nuclei or, generally, to conformational features formed early in the folding process. This would be true only if it can be assumed that contact regions, once they are formed, are not modified significantly during later stages of folding. This assumption cannot be justified rigorously at present, although it can be rendered plausible (see below). It is possible, in principle, that folding nuclei dissolve during subsequent stages, although this is unlikely. Also, they may undergo small shifts or they may grow in size. Contact region *F* (residues 103–124) of RNase is an example of this possibility. Several considerations, cited earlier, suggest that the primary nucleation site consists of the sequence 106–118—i.e., that it is part of region *F*. The folding of another part of region *F* (residues 121–124) occurs at a later stage (Table 2). The nucleation site is retained but in a slightly altered conformation. The assumption of the essential retention of structures formed early in folding seems reasonable because of several facts (2): (*a*) the ordered structures (α-helices, extended structures, and bends) that are predicted with one-dimensional short-range models correlate fairly well with those found in the native state; (*b*) many experimental observations, leading to the folding scheme shown in Table 2, are consistent with this assumption; (*c*) the assumption was valid in a Monte Carlo simulation of protein folding (1).

(*iii*) The hypothesis of Tanaka and Scheraga (2), cited at the beginning of *Results*, has not been proven independently. It can be justified only by its success in making predictions that may be compared with information about pathways derived from experiment (as in Table 2).

(*iv*) Even if the hypothesis can be assumed to be valid, uncertainties remain in the interpretation of the contact map. (*a*) It is not possible to decide, on the basis of the contact map alone, which one of several structures forms first if these structures involve interactions of similar ranges (e.g., contact regions *A* to *C*, or *D* to *F*, or *G*, *H*, and *K*). (*b*) It may happen that some structures involving long-range interactions (i.e., domains) are formed before the formation of short-range structures in other parts of the molecule. Examples will be given below. (*c*) It may even be possible that folding of part of a molecule is a prerequisite to the folding of the rest so that long-range contacts in one domain must precede the formation of shorter-range contacts in other parts of the molecule. Because of possibilities such as *a* to *c*, it cannot be determined from the contact map alone which of the branches shown in Fig. 2 represents the earlier stages of folding. The vertical positioning of the intermediate structures in Fig. 2 is therefore merely schematic. A temporal sequence of folding can be established only if there is experimental information about the identity and ordering of several intermediate structures.

Consequently, several possible pathways may be derived from the contact map even if uncertainties (*i*) and (*ii*) can be disregarded. The hypothesis of Tanaka and Scheraga can be applied to the contact map in at least three different ways, depending on the added assumptions one makes.

(*i*) Rigorous selection of a single pathway follows from the strict application of the rule, assuming that contact regions form strictly in the order of increasing distance from the diagonal—i.e., in the order of increasing values of $\mathrm{Min}|i-j|$. This leads to the pathway for RNase described in the second and third paragraphs of *Results*—i.e., to the sequence (*A* to *G*)—*L*—*H*—*K*—*J*—*I*—*M*. The order of formation of regions *A* to *G* cannot be decided by this rule, as discussed above. Such a rigorous application of the rule of increasing $\mathrm{Min}|i-j|$ may lead to inconsistencies with known chemical features of the molecule, as illustrated above for RNase S. In the case of proteins that can be subdivided into domains, it is not evident that this strict rule should be applied to the entire protein (see below).

(*ii*) Alternative rules take into account the existence of several domains, which may differ in stability. As an extreme case, it can be assumed that, under the influence of short- and medium-range interactions, several domains along the polypeptide chain can fold nearly simultaneously and independently of each other. After these domains fold, they associate under the influence of long-range interactions. The assumption of Tanaka and Scheraga still holds, but it is applied separately to each domain. The application of this rule to the contact map of Fig. 1 would suggest up to three independent folding domains for RNase A: residues 69–124 (folding to form contact regions *E* + *F* + *K*), residues 39–67 (folding to form contact regions *C* + *D* + *H*), and residues 1–38 (folding to form contact regions *A* + *B* + part of *L*). The three domains could form in any order. Subsequent association of the domains would establish the remaining contact regions. Alternatively, residues 1–38 and 40–124 could form two independent domains (the latter containing contact regions *C* + *D* + *E* + *F* + *H* + *J* + *K*).

(*iii*) The presence of a folded domain might be a prerequisite for the formation of certain short- or medium-range contacts. For example, it might happen that residues 21–49 can take up a stable conformation only if they can fold against a core formed by residues 50–124—i.e., that contact regions *B*, *G*, and *I* form simultaneously and subsequent to the formation of all contact regions to their right in Fig. 1.

Of course, the energy of stabilization of contact regions and domains depends on the nature and strength of noncovalent interactions between residues in contact, and it cannot be determined from the contact map alone. The presence of some favorable interactions may be more important than the total number of contacts (23).

Combined Use of the Contact Map and of Experimental Information on Folding. It is not possible to decide between these and other alternatives on the basis of the contact map alone. Even limited experimental information, however, may be sufficient to exclude some of the alternatives and to reduce the number of possible pathways. For example, knowledge about the folding of RNase S excludes the possibility that residues 1–38 form an independent folding domain [as exemplified in case (*ii*)] because this segment does not occur in the S-protein. The α-helix constituting contact region *A*, however, can form independently both in the S-peptide (prior to its association with the S-protein) and in RNase A. The pathway described in Table 2 is the result of a combination of structural information (from the contact map) and experimental data on folding.

In a reverse application the contact map may be used to fill gaps in the sequence of folding steps derived from experimental studies. If the experimental data are not sufficient to establish all steps of folding, additional constraints can be derived from the contact map. For example, experiments did not indicate the stage at which contact regions *H* and *I* are formed (Table 2). Inspection of Fig. 1 shows that contact region *H* is likely to form after the folding of regions *C* and *D*, but it must precede the formation of contact region *J* because both $\mathrm{Min}|i-j|$ and $\mathrm{Max}|i-j|$ for region *H* lie between the corresponding limits of *C* and *D* on the one hand and those of *J* on the other hand (Table 1). It is seen in a similar manner that contact region *I* forms after regions *H* and *G*. This information puts region *H* in stage IV or III and region *I* in stage III or II of the pathway shown in Table 2.

The four antigenic sites in native RNase occur in three distinct portions of the sequence, corresponding to three of the domains described above (19). The site occurring in segment 1–10 is contained in region *A*, the two sites occurring in segments 40–61 and 63–75 are contained in the domain formed by regions *C*, *D*, and *H*, and the site occurring in segment

87–104 is contained in the domain formed by regions *E*, *F*, and *K*. Cyanogen bromide treatment (19) cleaves the molecule after residues 13, 30, and 79. Three of the resulting fragments contain the antigenic sites mentioned. They also correspond to three separable domains (slightly different in size from those described above) on the contact map. This makes it likely that there is some intrinsic stabilization of native-like conformations in these fragments, resulting in antigenic activity. The fourth fragment, segment 13–29 (in which no antigenic site was found, even though most of it is exposed on the surface), does not correspond to a well-defined domain in the contact map. It is likely that it lacks conformational preference by itself in solution.

CONCLUSIONS

We have shown that the contact map of RNase can be used to predict a well-defined theoretical pathway for the folding of the molecule. This pathway is consistent with the experimental information available on intermediates in the thermal unfolding of RNase and in immunochemical studies of the oxidation of the reduced protein by glutathione and with the pathway (19, 21, 22) derived from this information.

Possible pathways and intermediates in folding can be predicted from the contact map. It can be used, therefore, to suggest experimental tests for intermediates or conformationally stable proteolytic fragments. It may resolve ambiguities in the interpretation of some experimental data when features of the contact map indicate that certain regions have to fold before others.

Experimental data may be used to decide between alternative pathways derived from contact maps. Conversely, contact maps are useful to indicate intermediates between experimentally observed partially folded states. Unfolding experiments and the analysis of the contact map supplement each other in the elucidation of the pathway of folding.

Note Added in Proof. A contact map for RNase A was constructed from the crystallographic coordinates that were recently submitted by R. Wlodawer to the Brookhaven Protein Data Bank and recently became available to us. The map is nearly identical to that shown in Fig. 1. The only major change is the appearance of some new contacts in contact regions *L* and *I*, between residues 16–23 and residues of the S-protein moiety, as would be expected from the presence of the peptide bond between residues 20 and 21 in RNase A. Elsewhere, there are only a few local changes of individual contacts that do not influence the appearance or extent of the contact regions shown in Fig. 1. All conclusions reached in this paper apply, therefore, equally well to RNases A and S.

We are indebted to S. Rumsey for help wth the computations. This work was supported by research grants from the National Institute on Aging (AG-00322) and the National Institute of General Medical Sciences (GM-14312) and from the National Science Foundation (PCM75-08691).

1. Tanaka, S. & Scheraga, H. A. (1975) *Proc. Natl. Acad. Sci. USA* **72,** 3802–3806.
2. Tanaka, S. & Scheraga, H. A. (1977) *Macromolecules* **10,** 291–304.
3. Levitt, M. & Warshel, A. (1975) *Nature (London)* **253,** 694–698.
4. Gō, N. & Taketomi, H. (1978) *Proc. Natl. Acad. Sci. USA* **75,** 559–563.
5. Yčas, M., Goel, N. S. & Jacobsen, J. W. (1978) *J. Theor. Biol.* **72,** 443–457.
6. Phillips, D. C. (1970) *British Biochemistry, Past and Present,* ed. Goodwin, T. W. (Academic, London), pp. 11–28.
7. Nishikawa, K., Ooi, T., Isogai, Y. & Saito, N. (1972) *J. Phys. Soc. Jpn.* **32,** 1331–1337.
8. Kuntz, I. D. (1975) *J. Am. Chem. Soc.* **97,** 4362–4366.
9. Kuntz, I. D., Crippen, G. M., Kollman, P. A. & Kimelman, D. (1976) *J. Mol. Biol.* **106,** 983–994.
10. Rossmann, M. G. & Liljas, A. (1974) *J. Mol. Biol.* **85,** 177–181.
11. Scheraga, H. A. (1979) in *Protein Folding,* Symposium at Regensburg, Germany, Sept. 1979, ed. Jaenicke R. (Elsevier, Amsterdam), in press.
12. Anfinsen, C. B. & Scheraga, H. A. (1975) *Adv. Protein Chem.* **29,** 205–300.
13. Némethy, G. & Scheraga, H. A. (1977) *Quart. Rev. Biophys.* **10,** 239–352.
14. Wyckoff, H. W., Tsernoglou, D., Hanson, A. W., Knox, J. R., Lee, B. & Richards, F. W. (1970) *J. Biol. Chem.* **245,** 305–328.
15. Richards, F. M. & Wyckoff, H. W. (1973) in *Atlas of Molecular Structures in Biology,* eds. Phillips, D. C. & Richards, F. M. (Clarendon, Oxford), Vol. 1.
16. Carlisle, C. H., Palmer, R. A., Mazumdar, S. K., Gorinsky, B. A. & Yeates, D. G. R. (1974) *J. Mol. Biol.* **85,** 1–18.
17. Rackovsky, S. & Scheraga, H. A. (1978) *Macromolecules* **11,** 1–8.
18. Kato, I. & Anfinsen, C. B. (1969) *J. Biol. Chem.* **244,** 1004–1007.
19. Chavez, L. G., Jr. & Scheraga, H. A. (1979) *Biochemistry,* in press.
20. Matheson, R. R., Jr. & Scheraga, H. A. (1978) *Macromolecules* **11,** 819–829.
21. Burgess, A. W. & Scheraga, H. A. (1975) *J. Theor. Biol.* **53,** 403–420.
22. Matheson, R. R., Jr. & Scheraga, H. A. (1979) *Biochemistry* **18,** 2437–2445.
23. Krigbaum, W. R. & Komoriya, A. (1979) *Biochim. Biophys. Acta* **576,** 229–246.