(4) Matsuda, K. *Polym. Prepr. (Am. Chem. Soc., Div. Polym. Chem.)* **1979**, *20*(1), 122.
(5) Blackwell, J.; Gutierrez, G. A.; Chivers, R. A. *Macromolecules* **1984**, *17*, 1219.
(6) Blackwell, J.; Gutierrez, G. A.; Chivers, R. A.; Ruland, W. *J. Polym. Sci., Polym. Phys. Ed.* **1984**, *22*, 1343.
(7) Blackwell, J.; Biswas, A.; Bonart, R. *Macromolecules* **1985**, *18*, 2126.
(8) Mitchell, G. R.; Windle, A. H. *Colloid Polym. Sci.* **1985**, *263*, 230.
(9) Hummel, J. P.; Flory, P. J. *Macromolecules* **1980**, *13*, 479.
(10) Takahashu, N.; Yoon, D. Y.; Parrish, W. *Macromolecules* **1984**, *17*, 2583.
(11) Northolt, M. G. *Eur. Polym. J.* **1974**, *10*, 799.
(12) Vijayalakshmi, B. K.; Srinivasan, R. *J. Cryst. Mol. Struct.* **1973**, *3*, 147.
(13) Gardner, K. H.; Blackwell, J. *Acta Crystallogr., Sect. B: Struct. Crystallogr. Cryst. Chem.* **1980**, *B36*, 1972.
(14) Biswas, A.; Blackwell, J., *Macromolecules*, in press.
(15) Dobb, M. G.; Johnson, D. J.; Saville, B. P. *J. Polym. Sci., Polym. Symp.* **1977**, *No. 58*, 237.
(16) Matheson, R. R., personal communication.
(17) Yoon, D. Y.; Bruckner, S.; Volksen, W.; Scott, J. C.; Griffin, A. C. *Faraday Discuss. Chem. Soc.* **1985**, *79*, 41.
(18) Flory, P. J.; Matheson, R. R., Jr. *J. Phys. Chem.* **1984**, *88*, 6606.
(19) Hendricks, S.; Teller, E. *J. Chem. Phys.* **1942**, *10*, 147.
(20) Chivers, R. A.; Blackwell, J.; Gutierrez, G. A. *Polymer* **1984**, *25*, 435.

A Statistical Test of the Proposition That Intrahelix Salt Bridges Constitute a Significant Stabilizing Feature of the Tropomyosin Coiled-Coil Structure

Alfred Holtzer* and Marilyn Emerson Holtzer

Department of Chemistry, Washington University, St. Louis, Missouri 63130.
Received May 3, 1986

ABSTRACT: A statistical analysis is made of the correlation of the charge state of residue i in the amino acid sequence of α-tropomyosin with that of residues $i \pm 3$ and $i \pm 4$. Since the chains in this protein are essentially completely α-helical (3.6 residues/turn), the side chains of such pairs of residues are brought into juxtaposition by the secondary structure. It has therefore been proposed (Sundaralingam, M., et al. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 7944–7947) that such salt linkages are a major principle in stabilizing the structure in highly α-helical, two-chain, coiled-coil proteins. The statistical analysis, however, shows that the distribution of charged residues in α-tropomyosin with respect to these interactions is essentially random. The statistical property χ^2 is used to measure deviations from expectations based on the hypothesis of randomness. It is found that the value of χ^2 obtained for both the $i{\cdot}(i + 4)$ and $i{\cdot}(i + 3)$ interactions in tropomyosin is no larger than it would be for about 50% of the chains in a huge collection in which each chain has the same amino acid composition and length as tropomyosin, but in which the sequences are random. Thus, there is no reason to believe that the $i \pm 3$ and $i \pm 4$ positions near a given charged group at position i in tropomyosin are biased in favor of opposite charges. This analysis therefore does not support the idea that such intrahelical salt bridges are a major factor in stabilizing the coiled-coil structure. In contrast, it is also shown that a similar statistical test indicates a highly significant deviation from randomness for the interhelical salt bridges that have long been supposed to be a significant feature of the structure. An examination of other existing evidence also does not encourage assigning a significant role to intrahelical salt bridges in coiled-coil proteins.

The proteins tropomyosin and paramyosin and the large tail of the protein myosin have a common molecular structure, called the two-chain coiled coil.[1] In each case, the molecule comprises two poly(amino acid) chains that are essentially completely α-helical. The two constituent helices are arranged in parallel and in register with a slight supertwist. Because of its simplicity, the structure represents an opportunity for development of ideas concerning protein structure in general. In spite of its simplicity, however, the structure is a sharp challenge to our present understanding.

To date, very few general principles have become apparent that serve to define a given poly(amino acid) chain as one in which the two-stranded, coiled coil could be safely predicted to be the preferred conformation. Only four such principles, none quantitatively predictive, seem to have gained general acceptance. First, there must be an absence of α-helix-disruptive residues. That no coiled-coil poly(amino acid) can contain proline residues seems axiomatic. Not quite so understandably, but probably for similar reasons, tryptophan is absent as well and residues of low intrinsic helix-forming potential are few.[2] Second, it was noticed in even the earliest determinations of amino acid composition that such proteins have a greater fraction of charged residues than do most globular proteins.[3] This seemed understandable, since a molecule of extended form has a greater surface/volume ratio and therefore requires a greater fraction of hydrophilic residues in order to destabilize more compact conformations.

Third, hydrophobic residues must appear in a definite pattern in the sequence.[4-7] The sequence, as first reported for tropomyosin, is demonstrably based on a pseudorepeating heptet of residues. Position within each heptet is designated by letters a–g.[7] Positions a and d are ordinarily hydrophobic. This results in an extremely amphipathic helix, providing a hydrophobic streak to which another such helix can adhere. The a and d side chains of each helix readily articulate by "knobs-into-holes" packing.[8] It is generally agreed that this interhelix, hydrophobic interaction is a major factor in holding the chains together and probably also in maintaining the helical state in each.

Fourth, the sequence shows that residues in heptet positions e and g tend to be oppositely charged; in tropomyosin, those in e tend to be negatively charged, and those in g positively charged.[5-7] In the double-helical molecule this leads to salt links, predominantly between residue e in the nth heptet of one helix and residue g' of

0024-9297/87/2220-0671$01.50/0

the $(n - 1)$th heptet of the other (primed) helix.[6,7] These salt bridges also seem to be a characteristic feature of coiled coils. Indeed, since the packing of hydrophobes at the helix–helix interface is, to first approximation at any rate, equally feasible with the helices parallel or antiparallel,[8] these interhelix salt bridges appear to be the only means so far proposed by which parallel assembly can be assured.[2]

Recently, a new principle has been proposed for the maintenance of the coiled-coil structure. In a recent paper, Sundaralingam et al. suggest that a large number of intrahelical salt links may be a significant general feature of the coiled-coil structure.[9] The idea is based on the well-known "helical-wheel" effect. Since an α-helix has 3.6 residues/turn, succeeding turns bring the side chains of residues $i \pm 4$ and $i \pm 3$ into proximity to that of residue i. It is maintained that the amino acid sequence of coiled coils provides a large number of attractive intrahelix charge–charge interactions, which may be a general, stabilizing influence on the structure. In the following, we present a simple statistical test of this idea.

Methods

The possible importance of intrahelix, $i\cdot(i \pm 4)$ and $i\cdot(i \pm 3)$ side-chain interactions was appreciated even before any sequence information was available.[10,11] Publication of the first sequence of a coiled coil (tropomyosin[4]) made possible the compilation of a catalog of these pairwise interactions. At first, we obtained counts by hand of, for example, the number of glutamate–lysine, $i\cdot(i \pm 4)$ pairs. Such census-taking is, however, greatly expedited by entering the information into a computer, using spreadsheet or database software. We employed LOTUS 123. The basic template was a catalog of the complete sequence.[5] Each row in the catalog gives the single-letter code for the amino acid residue (e.g., K for lysine), the number of the residue in the sequence (1–284 in the case of rabbit α-tropomyosin), the number of the heptet in which it is located (1–41), the intraheptet position-designating letter (a–g), the electrostatic charge at near-neutral pH (−1, 0, or +1), and the hydrophobicity.

The spreadsheet/database software allows a great many properties of the sequence to be extracted at the touch of a few keys. For example, one can arrange that a copy of the sequence, shifted by three residues from the original, be added to the original. Then one can query the data in various ways, e.g., have the program select those pairs whose charge products are negative, thus obtaining a list of attractive $i\cdot(i \pm 3)$ interactions, etc. In this way, our primitive catalog was checked and extended.

In the present instance, in addition to the usual −1 charge for aspartate and glutamate and +1 for arginine and lysine, values that apply anywhere in the vicinity of neutral pH, we used a charge of −1 for cysteine and +1 for histidine. Since both of the latter two groups titrate near neutrality, their charge state depends on precisely what value of the near-neutral pH is of interest. However, since α-tropomyosin has only one cysteine and only two histidines,[5] the decision to use the charged form in both cases has no material bearing on the conclusions.

Results and Discussion

The basis for the proposal that intrahelical salt bridges are a highly significant feature of the coiled-coil structure in tropomyosin is the large number of examples in the sequence in which a negatively charged group i has a positively charged neighbor at $i \pm 4$ or $i \pm 3$.[9] At first sight, the numbers indeed seem impressive. One can state the case in various ways. For example, there are a total of 37 instances in the sequence of potential, $i\cdot(i \pm 4)$, attractive, intrahelix interactions.[12] That represents a large number of potential salt bridges without even considering $i\cdot(i \pm 3)$ interactions. However, there are also 30 instances of repulsive interactions in the same category. The net result, 7 attractive interactions, is not nearly so impressive. Of course, neighbors of like charge will simply avoid one another and reduce the statistical weight of their contribution. However, when one also reflects that the two helices, both of net negative charge, then have to be packed together, the case for a large net electrostatic stabilization seems less compelling.

Table I
Contingency Table for $i \bullet (i + 4)$ Intrahelix Charge Interactions: Randomization over Entire Molecule

i position	$i + 4$ position: −	0	+	i totals
−	23 (22.78)	38 (41.62)	20 (15.46)	81 (79.86)
0	40 (41.62)	76 (76.04)	28 (28.26)	144 (145.92)
+	17 (15.46)	31 (28.26)	7 (10.50)	55 (54.22)
$i + 4$ totals	80 (79.86)	145 (145.92)	55 (54.22)	280 (280)

Result: $\chi^2 = 3.30$; $P = 0.51$

In another way of putting the case, Sundaralingam et al. point out that, for example, a large number of glutamate residues in tropomyosin have positively charged neighbors at $i \pm 4$ or $i \pm 3$.[9] This is undeniable. There are 56 glutamates in the tropomyosin sequence and they provide 23 potential ion pairs with positively charged neighbors at $i \pm 4$.[12] This is at first sight a large number, and, of course, one would have to add to it potential glutamate–positive pairs at $i \pm 3$ and, similarly, aspartate–positive pairs. The number cited, however, becomes less impressive upon statistical examination. There are 55 positively charged residues in the molecule. There is, therefore, a probability of 55/284 that a positively charged residue will appear at any given site in a random arrangement. Since we examined both $i + 4$ and $i - 4$ sites for the ith glutamate, and there are 56 glutamates, the expected number of potential attractive interactions in a random sequence is $56 \times 2 \times (55/284) \cong 22$.[13] Thus, the expected number in a random sequence is, in this instance, almost identical with that actually found.

Since this simple statistical calculation does not support the idea that any principle is at work, we performed a χ^2 test[14] on the charge distribution with respect to these helical-wheel interactions. Each residue has a charge of −1, 0, or +1. Successive residues in the sequence were examined with respect to their own charge and that of the residue at $i + 4$. The results are shown in Table I. In this contingency table, the charge state of residue i appears at the left, and that of residue $i + 4$ at the top. The table thus records, for example, that there are 17 positive residues at i that have a negative residue at $i + 4$. Row and column totals are also shown. The grand total is 280 (not 284, the total number of residues) because residues beyond 280 have no neighbor at $i + 4$.

The values expected on the assumption of a random distribution of residues over the whole chain appear in parentheses in Table I. For example, since there are a total of 81 negative residues and 55 positive ones, the probability of finding a positive residue at position $i + 4$ vis a vis a negative residue at i is (81/284)(55/284); since we examine 280 positions, we expect therefore to find (280)(81/284)(55/284) = 15.46 instances of negative–positive pairs in a random sequence. The other expectation values are obtained similarly.

The standard test for such contingency tables is the χ^2 test,[14] in which the quantity $\chi^2 = \sum[(x_{ij} - E_{ij})^2/E_{ij}]$ is

Table II
Contingency Table for $i \cdot (i + 3)$ Intrahelix Charge Interactions: Randomization over Entire Molecule

i position	$i + 3$ position			i totals
	−	0	+	
−	22 (22.86)	42 (41.77)	17 (15.52)	81 (80.14)
0	38 (41.77)	76 (76.31)	31 (28.36)	145 (146.44)
+	20 (15.52)	28 (28.36)	7 (10.54)	55 (54.42)
$i + 3$ totals	80 (80.15)	146 (146.44)	55 (54.42)	281 (281.00)

Result: $\chi^2 = 3.25$; $P = 0.52$

computed. Summation is over all entries in the contingency table. x_{ij} is the table entry in cell i,j; E_{ij} is the corresponding expectation value. For Table I, we find $\chi^2 = 3.30$. Perfect agreement with random expectations would give $\chi^2 = 0$. Therefore, small values indicate agreement with randomness, and large ones disagreement. It remains to demonstrate whether 3.30 is, in this case, large or small. This is normally accomplished by consulting statistical tables that give values for integrals of the distribution of χ^2 values. The form of the distribution is known and it depends upon the number of degrees of freedom of the system, which is given, for an $r \times s$ table, by $f = (r - 1)(s - 1)$. In our case, $f = 4$. In that case the form of the distribution is elementary and is directly integrable. The integral from χ^2 to infinity, i.e., the probability of a value of χ^2 greater than some given one, is given by $P = [1 + (\chi^2/2)]e^{-\chi^2/2}$.

We find for $f = 4$ that the integral from $\chi^2 = 3.30$ to infinity is $P \cong 0.51$. One statistical interpretation of this can be stated as follows. If one generated an enormous number of 284-residue chains, each with precisely the amino acid composition of tropomyosin but of random sequence, 51% of them would have a χ^2 equal to or greater than tropomyosin's. Thus, $\chi^2 = 3.30$ cannot be considered large enough to be judged to be a result of other than chance.

A similar table for the $i \cdot (i + 3)$ intrahelix interactions is shown in Table II. Here we find $\chi^2 = 3.25$ and $P \cong 0.52$. Again, for these interactions, more than half the chains with randomly arranged residues have charge distributions that differ from the random expectation values by as much as or more than tropomyosin does. Once again, we cannot conclude that anything but chance is operating in placing oppositely charged residues at positions favoring their interaction in the α-chains of tropomyosin. Moreover, the β-tropomyosin chain differs from the α at 39 sites. Of those, only 9 involve charged residues, and of those, 6 of the changes conserve charge. The resultant charge changes at only 3 sites are immaterial, so our conclusion clearly extends to β-tropomyosin chains as well. A different conclusion apparently applies in globular proteins.[15]

In the development of contingency Tables I and II, expectation values were calculated on the assumption of random distribution of residues over the whole molecule. It could be argued that this is unrealistic, since no one seriously doubts the predominantly hydrophobic nature of the residues in the a and d positions. We therefore constructed two corresponding tables in which expectation values were determined from the hypothesis that charged residues are randomly distributed, but only among sites actually bearing charges in the α-tropomyosin chain. The results differ negligibly from the previous. We found for $i \cdot (i + 4)$ interactions $\chi^2 = 3.02$ and $P = 0.56$ and for $i \cdot (i + 3)$ interactions $\chi^2 = 3.13$ and $P = 0.54$. The conclusion is the same.

While emphasizing intrahelix salt linkages, Sundaralingam et al. also downplay the role of interhelix salt linkages. They contrast the large number of potential intrahelix bridges (see above) with "only 9 potential interhelical salt-bridge interactions between the two adjacent polypeptide chains". This conclusion apparently rests, in part, upon a misunderstanding of the fourth principle enunciated above. Our own count indeed verifies that only 9 examples exist in the structure of favorable $ne \cdot (ng)'$ interactions. First, it must be pointed out, however, that there are also the $(ne)' \cdot ng$ interactions, so the actual number of such stabilizing contacts per molecule is 18. Second, and crucially, while those interactions may contribute, the canonical, stabilizing, interhelix, salt bridges are not those, but the $ne \cdot [(n - 1)g]'$ and $(ne)' \cdot (n - 1)g$ interactions, which not only exist in much larger number but in a distribution that is certainly statistically distinguishable from chance.[6,7,16,17] This nonrandom property of the distribution of $ne \cdot [(n - 1)g]'$ residue pairs has, in fact, already been subjected to one statistical test.[7]

Table III
Contingency Table for $ne \cdot [(n - 1)g]'$ Interhelix Charge Interactions: Separate Randomization over ne and $(n - 1)g$ Sites

$(n - 1)g'$ position	ne position			$(n - 1)g'$ totals
	−	0	+	
−	2 (6.49)	8 (3.67)	1 (0.85)	11 (11.0)
0	6 (5.31)	2 (3.00)	1 (0.69)	9 (9.0)
+	15 (11.2)	3 (6.33)	1 (1.46)	19 (19.0)
ne totals	23 (23.0)	13 (13.0)	3 (3.0)	39

Result: $\chi^2 = 12$; $P = 0.017$

Table IV
Contingency Table for $ne \cdot [(n - 1)g]'$ Interhelix Charge Interactions: Combined Randomization over ne and $(n - 1)g$ Sites

$(n - 1)g'$ position	ne position			$(n - 1)g'$ totals
	−	0	+	
−	2 (7.41)	8 (4.80)	1 (4.80)	11 (17.01)
0	6 (4.80)	2 (3.10)	1 (3.10)	9 (11.00)
+	15 (4.80)	3 (3.10)	1 (3.10)	19 (11.00)
ne totals	23 (17.01)	13 (11.00)	3 (11.00)	39 (39.01)

Result: $\chi^2 = 34.4$; $P = 6.2 \times 10^{-7}$

To verify the extant results on interhelix salt interactions, we show the contingency Table III, in which frequencies of ne residue charge types (top) and that for their $(n - 1)g'$ neighbors (left side) are given. This table was determined on the same basis as in ref 7. That is, expectation values were obtained with the assumption that the existing ne residues are distributed at random among the 39 ne sites and that the existing $(n - 1)g'$ residues are distributed at random among the 39 $(n - 1)g'$ sites. Note again that all numbers in the table would double if the corresponding $(ne)' \cdot (n - 1)g$ pairs were included. Using the same test as before, we find $\chi^2 \cong 12$ and $P = 0.017$. Thus, one can imagine a huge number of coiled-coil molecules in which each is made of two identical chains, but, from molecule to molecule, the ne residues are randomly arranged among the 39 e sites and the $(n - 1)g'$ residues randomly arranged among the 39 $(n - 1)g'$ sites. In such an assemblage, less than 2% of the molecules would have a distribution of charges in the ne and $(n - 1)g'$ positions, leading to a χ^2 value as large as or larger than in tropomyosin. Plainly, the $ne \cdot [(n - 1)g]'$ salt bridges are there as the fittest survivors of evolutionary processes, not by chance.

In order to probe the interhelical cross-bridge question further, we constructed contingency Table IV on a somewhat different basis. Here, the expectation values were

calculated on the assumption that the *ne* and $(n - 1)g$ side chains on each chain were stripped off and restored at random to the combined set of stripped sites. Then, an identical mate is imagined to provide the $ne\cdot[(n - 1)g]'$ interactions. For this table, we find $\chi^2 = 34.36$ and $P = 6.2 \times 10^{-7}$. Thus, the probability is virtually zero that the existing pattern of $ne\cdot[(n - 1)g]'$ residues occurred as a result of random distribution of all the *ne* and $(n - 1)g'$ residues among all those sites.

We also constructed a contingency table for $ne\cdot[(n - 1)g]'$ charge interactions on the hypothesis that the positions of all the charges in the molecule are randomized over the whole molecule instead of only the charges at *e* and *g'* being randomized over their own *e* and *g'* sites. The results are spectacular. We find $\chi^2 \cong 88$ and $P = 3.5 \times 10^{-18}$. There seems to be no question that $ne\cdot[(n - 1)g]'$ salt links are a strongly preferred interaction in the structure.

It is, of course, not possible to prove by statistical arguments alone that intrahelix salt linkages, or any other interactions, are or are not present or significant. Definitive proof of the existence of salt links can only be provided by determination at high resolution of the spatial coordinates of the charged groups. Even then, proof of a significant overall resultant stabilization would only follow from free energy determinations using a definitive, well-verified theory or experiments that allow the electrostatic contribution to be dissected from the total. None of these conditions holds at present for tropomyosin or is likely to in the near future. Thus, one is forced to adopt less certain criteria, such as statistical tests, in attempting to discover which effects may be broadly applicable. It is impossible to rule out, on this basis, effects that may occur at a few local, but important, sites. It is also impossible to say that randomly occurring interactions cannot be used, opportunely, as stabilizers. Nevertheless, it would seem to be more productive to suspect as broadly significant those few interactions that pass *some* test (as $ne\cdot[(n - 1)g]'$ interactions do) rather than those that pass none.

Finally, it might be added that there are additional arguments that cast doubt on electrostatic interactions in general and intrachain salt links in particular as a major principle of stabilization in coiled coils. First, there is the enormous increment in stability that accompanies reduction of the pH from 7 to 2.[10,11,18,19] At pH 2, all salt links are, perforce, destroyed and the chains acquire a net positive charge. Ironically enough, the earliest conjecture concerning this result, made before the tropomyosin sequence was known, was that side-chain carboxyl–carboxyl hydrogen bonds involving $i\cdot(i \pm 4)$ and/or $i\cdot(i \pm 3)$ pairs might be involved.[10,11] This conjecture remains neither proven nor disproven, but also remains strictly ad hoc. More recently, an explanation has been put forward based on the experimental finding that short-range helix-forming tendencies of aspartic and glutamic residues considerably exceed those for aspartate and glutamate, respectively.[18,19] The latter, while it has not been definitively verified, is at least not completely ad hoc. If intrahelical salt linkages are a major source of stabilization, the enormous stability of coiled coils in acid becomes mysterious.

Second, a major energetic contribution from widespread salt links would necessarily alter the p*K*'s of the protein's acidic groups (which titrate in a region where the structure is stable) on a large scale. In fact, they titrate normally.[20] This result does not vitiate the canonical view of the role of interhelix $ne\cdot[(n - 1)g]'$ salt links. It is not necessary to assume that the latter make up a sizeable fraction of the total stabilization. Indeed, their contribution is thought to be a small fraction even of only the interhelical part of the overall stabilization free energy.[19] Their role is supposed to be one of fine-tuning, i.e., to dictate parallel, rather than antiparallel, assembly.[2]

Third, the classic test for electrostatic interactions is their sensitivity to ionic strength. Because tropomyosin only becomes truly molecularly dispersed above an ionic strength of ~0.5 M, effects of this variable have not been well studied. However, it is clear that the structural stability not only persists up to at least 1 M NaCl, but is almost independent of ionic strength above 0.5 M.[21] Moreover, such studies as do exist indicate a decrease in structural stability below 0.5 M,[22] precisely opposite to expectations for a structure whose stability results in major part to salt links.

Fourth, poly(amino acid) chains have been synthesized that mimic the canonical heptet structure of tropomyosin.[16,23] These chains principally consist of iterations of the heptet sequence LEALEGK. Such chains satisfy all four of the canons laid down above for coiled coils. Heptet positions *a* and *d* are invariably leucines. Heptet positions *e* are all occupied by negatively charged glutamates. Heptet positions *g* are occupied by positively charged lysines. Chain lengths are short compared with tropomyosin, but lengths up to 43 residues have been achieved. Physical studies indicate that they form two-chain coiled coils. Thermal curves indicate, moreover, that the structure is very stable in spite of the relatively short length, probably because the high degree of regularity ensures efficient packing and maximal free energy loss due to the strong hydrophobicity of the leucines in the contact region. It is easy to see that such a chain has no favorable $i\cdot(i \pm 4)$ or $i\cdot(i \pm 3)$ electrostatic interactions whatever. The lysines experience no charge interactions at those intervals at all. Glutamates in the *b* heptet positions experience only repulsive helical-wheel electrostatic interactions (at $i + 3$ and $i - 4$). Glutamates in the *e* heptet positions also experience only repulsive intrahelix helical-wheel interactions (at $i + 3$ and $i - 3$). It seems unlikely that a stability principle broadly applicable and significant in magnitude in the native proteins could be so thoroughly violated in these peptide models and yet lead to the same structure.

Fifth, and last, it seems unwise to attempt a partial description of electrostatic interactions. Even if regularities can be discerned in the individual helices, their energetic effects will change magnitude, or even sign, when the helices are brought together. Experience with globular proteins suggests the necessity of a global approach in calculating electrostatic interactions.[24] Unfortunately, even the limited success achieved in cases of near-spherical symmetry seems not achievable for molecules that are highly extended in shape.[25] Moreover, it may be necessary to include not only charges on ionizable groups but partial charges as well. The backbone carbonyl and amino groups in an α-helix as well as the terminal amino and carboxylate would conspire to give an α-helix a net dipole moment even at zero net charge. Assembly of the tropomyosin helices in parallel thus entails an unfavorable dipole–dipole contribution. The canonical view is that the $ne\cdot[(n - 1)g]'$ salt links overcome this repulsion with enough to spare to dictate overwhelming parallelism.

The demonstrably nonrandom charge pattern at the *ne* and $(n - 1)g'$ heptet positions clearly supports this view but cannot prove it. It is tempting to look for other ways of negating this unfavorable backbone dipole–dipole interaction. In that connection, it might be noticed in Table I that there is an excess of $i\cdot(i + 4)$ negative–positive pairs (20) over corresponding positive–negative pairs (17), potentially providing a dipole moment tending to cancel that

of the backbone. However, inspection of Table II shows precisely the opposite trend for i·(i + 3) interactions.

Even a basically linear protein structure such as tropomyosin's is rife with possible combinations. One can see examples in it of just about anything one wishes. Under the circumstances, it is probably most practical to focus upon those that satisfy some appropriate criterion other than the fact of their existence. It has not been demonstrated that intrahelical salt linkages pass any such test.

Acknowledgment. This work was supported in part by Grant No. GM-20064 from the Division of General Medical Sciences, U.S. Public Health Service, and in part under a grant from Muscular Dystrophy Association. A.H. acknowledges an informative discussion of contingency tables with a colleague, Prof. Edward Spitznagel, Jr.

References and Notes

(1) Fraser, R. D. B.; Mac Rae, T. P. *Conformation in Fibrous Proteins*; Academic: New York, 1973; pp 419–468.
(2) Parry, D. A. D. *J. Mol. Biol.* **1975**, *98*, 519–535.
(3) Szent-Györgyi, A. G. In *The Structure and Function of Muscle*; G. H. Bourne, G. H., Ed.; Academic: New York and London, 1960; Chapter I.
(4) Hodges, R. S.; Sodek, J.; Smillie, L. B.; Jurasek, L. *Cold Spring Harbor Symp. Quant. Biol.* **1972**, *37*, 299–310.
(5) Mak, A. S.; Smillie, L. B.; Stewart, G. R. *J. Biol. Chem.* **1980**, *255*, 3647–3655.
(6) Stone, D.; Sodek, J.; Johnson, P.; Smillie, L. B. *Proc. FEBS Meeting* **1974**, *31*, 125–136.
(7) McLachlan, A. D.; Stewart, M. *J. Mol. Biol.* **1975**, *98*, 293–304.
(8) Crick, F. H. C. *Acta Crystallogr.* **1953**, *6*, 689–697.
(9) Sundaralingam, M.; Drendel, W.; Greaser, M. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 7944–7947.
(10) Noelken, M. Ph.D. Thesis, Washington University, 1962.
(11) Noelken, M.; Holtzer, A. In *Biochemistry of Muscle Contraction*; Gergely, J., Ed.; Little, Brown and Co.: Boston, 1964; pp 374–378.
(12) We include *all* potential ion pairs. For example, glutamate-33, which has a lysine at 33 + 4 = 37 *and* another lysine at 33 − 4 = 29, contributes two pairs of potential i·($i \pm 4$) attractive pairs to the total number. The total obtained in the following paragraph for i·($i \pm 4$) attractive pairs that involve glutamate was obtained on the same basis.
(13) The factor of 2 is a bit too large. It only holds for an interior glutamate residue, i.e., one that *has* a neighbor at both $i - 4$ and $i + 4$. Thus, only residue positions 5–280 would qualify. A glutamate in positions 1–4 would only have one such neighbor (at $i + 4$) and one in positions 281–284 would also have only one such neighbor (at $i - 4$). Thus, the average number of such neighbors per residue is [(2)(276) + 8]/284 = 1.972. Use of this factor instead of 2 gives an expectation value of 21.4 instead of 22, an immaterial change.
(14) Any standard work on statistics covers this topic. We used: Bulmer, M. G. *Principles of Statistics*; Dover: New York, 1979; p 154f. For numerical values, one can employ: *Handbook of Tables for Probability and Statistics*, 2nd ed.; Beyer, W. H., Ed.; Chemical Rubber Co.: Cleveland, 1968. Similar tables can be found in many standard compilations. Alternatively, one can recognize that the required integral is of standard form for contingency tables of four degrees of freedom and integrate directly.
(15) Maxfield, F. R.; Scheraga, H. A. *Macromolecules* **1975**, *8*, 491–493.
(16) Lau, S. Y. M.; Taneja, A. K.; Hodges, R. S. *J. Biol. Chem.* **1984**, *259*, 13 253–13 261.
(17) Sanders, C.; Smillie, L. B. *J. Biol. Chem.* **1985**, *260*, 7264–7275.
(18) Holtzer, M. E.; Holtzer, A.; Skolnick, J. *Macromolecules* **1983**, *16*, 462–465.
(19) Skolnick, J.; Holtzer, A. *Macromolecules* **1985**, *18*, 1549–1559.
(20) Iida, S.; Imai, N. *J. Phys. Chem.* **1969**, *73*, 75–80.
(21) Isom, L. L.; Holtzer, M. E.; Holtzer, A. *Macromolecules* **1984**, *17*, 2445–2447.
(22) Betteridge, D.; Lehrer, S. *J. Mol. Biol.* **1983**, *167*, 481–496.
(23) Hodges, R. S.; Saund, A. K.; Chong, P. C. S.; St. Pierre, S. A.; Reid, R. E. *J. Biol. Chem.* **1981**, *256*, 1214–1224.
(24) Matthew, J. B.; Gurd, F. R. N.; Garcia-Moreno, B.; Flanagan, M. A.; March, K. L.; Shire, S. J. *CRC Crit. Rev. Biochem.* **1985**, *18*, 91–197.
(25) Skolnick, J.; Fixman, M. *Macromolecules* **1978**, *11*, 867–871.

Influences of the Initiation and Termination Reactions on the Molecular Weight Distribution and Compositional Heterogeneity of Functional Copolymers: An Application of Monte Carlo Simulation

M. Neil Galbraith, Graeme Moad,* David H. Solomon, and Thomas H. Spurling*

CSIRO, Division of Applied Organic Chemistry, Melbourne, Victoria, Australia 3001.
Received September 15, 1986

ABSTRACT: Monte Carlo simulations have been used to predict the molecular weight, molecular weight distribution, and composition of copolymers. The method is particularly suited to the simulation of copolymerizations carried out in the presence of chain-transfer agents (e.g., thiols). Calculations have been performed to show that selectivity shown in the initiation and termination reactions can have a dramatic influence on the composition and molecular weight distribution of cooligomers and copolymers.

Introduction

The aim of this work has been to examine the factors that determine the composition and molecular weight distribution of multicomponent copolymers, in particular, to discover what influences selectivity shown in the initiation and termination steps might have on the distribution of monomer units within the copolymer chain. This is of particular relevance in the synthesis of functional copolymers and cooligomers, which find widespread use in the coatings and adhesives industry.[1,2]

That such considerations should be of significance, particularly where low molecular weight materials are concerned, can be readily appreciated given the knowledge that initiator and transfer agent derived radicals can show a high degree of selectivity for reaction with a given monomer.[3] Thus the initiating end of the polymer or oligomer

0024-9297/87/2220-0675$01.50/0