Proc. Natl. Acad. Sci. USA
Vol. 82, pp. 7944–7947, December 1985
Biochemistry

Stabilization of the long central helix of troponin C by intrahelical salt bridges between charged amino acid side chains

(α-helix stability/Ca^{2+}-binding proteins/sequence homologies/salt bridges in muscle proteins/protein folding mechanism)

MUTTAIYA SUNDARALINGAM*, WILLIAM DRENDEL*, AND MARION GREASER†

*Department of Biochemistry and †Muscle Biology Laboratory, College of Agricultural and Life Sciences, University of Wisconsin, Madison, WI 53706

Communicated by Robert L. Baldwin, August 5, 1985

ABSTRACT **The unusual dumbbell shape of troponin C is due to the presence of a long α-helix of nine turns that connects the amino- and carboxyl-terminal calcium-binding domains. The center of the long helix appears to be stabilized by several salt bridges. The long helix is also bent about 16° at glycine-92. Calmodulin, which lacks the central glycine, also is predicted to be stabilized by salt bridges in the central helix. The presence of a proline residue in the center of the long helix of ascidian troponin C and the myosin regulatory light chains suggests that a sharper bend may occur in these molecules. The conservation of the bend and salt bridges in the related calcium-binding proteins suggests they may have an important biological function. The structure of troponin C suggests that intrahelix salt bridges between neighboring charged residues may be involved in the stabilization of protein secondary structure. The preponderance of potential salt bridges in other muscle proteins as well may be related to their elongated structures and their participation in the contractile process.**

Recent x-ray diffraction results on troponin C and calmodulin crystals have revealed that these molecules are dumbbell shaped with a long central helix extending between the amino and carboxyl domains (1–4). The nine-turn helix in troponin C has about three turns buried (either wholly or partially) in each of the flanking domains, but this leaves three turns in the center exposed fully to the solvent. It is well known that single α helices are highly unstable in an aqueous environment (5) and, thus, the exposed helix in these calcium-binding proteins was totally unexpected. In fact, there is no precedent for such an arrangement among the large number of proteins whose x-ray structures are known. An examination of the troponin C amino acid sequence in the DE-linker helix, however, shows that the residues have a high potential for helix formation. A proposed structure of troponin C (6) had predicted this region to be a random coil as would have been expected by the corresponding BC-linking region of parvalbumin on which the model was based. It is the unexpected helix in this linker segment that gives rise to the long central helix and in turn the unusual dumbbell shape of these molecules. An examination of the factors that induce and stabilize the central helix—especially the linker segment—is, therefore, essential to understanding the conformation and dynamics of these Ca^{2+}-binding proteins.

Based on our investigation of the x-ray structure of troponin C from chicken skeletal muscle, we suggest that the integrity of the long central helix is derived from numerous intrahelical electrostatic and/or salt-bridge interactions between the basic and acidic amino acid side chains, especially in the exposed DE-linker segment. We support this hypothesis both by reference to our x-ray structure of troponin C and by phylogenetic considerations of the amino acid sequences of various calcium-binding proteins.

It has been shown that the frequency of occurrence of glutamic acid residues in α-helices of proteins markedly exceeds that expected from studies of random copolymers (7). These workers also showed that the probability for a glutamic acid (position *i*) being in an α-helix was significantly greater if an amino acid with a positively charged side chain was three or four residues away ($i \pm 3$ or 4) in the sequence. Similar enhanced probabilities were found with oppositely charged residues at two or three positions away from lysine and at 3 positions away from aspartic acid. Of 25 glutamic acid residues in troponin C, 23 are indeed in helical segments.

Salt Bridges in the Long Helix of Troponin C. The preliminary structure of chicken skeletal muscle troponin C at 3 Å resolution has been reported (1, 2). The electron density map shows clean and well-defined features in the region of the central (DE) helix, permitting an excellent fit of the backbone as well as the proper placement and orientation of most of the side chains. It is not possible at 3 Å resolution to provide a detailed description of the side chain interactions. However, one can readily observe the relative orientations and proximities of the side chains, and this has allowed us to evaluate the likelihood of the putative interactions described above and to tentatively assign them as being either charge–charge or hydrogen bonded or both. In cases where the relevant side chains show an obvious overlapping of electron density, we have designated the interaction as charge–charge/hydrogen bond in Table 1. Where the disposition of the side chains appears to take advantage of the electrostatic attraction, but no overlapping of electron density is seen, the interactions are designated charge–charge. The glutamic acid-85 · · · aspartic acid-89 interaction is designated simply hydrogen bond since no charge–charge interaction is possible.

A total of ten residues within the central helix have the potential for these types of interactions (Table 1 and Fig. 1). There are a striking number of possible salt-bridge combinations: arginine-84 · · · glutamic acid-88; glutamic acid-88 · · · lysine-91; aspartic acid-89 · · · lysine-93; lysine-91 · · · glutamic acid-95; lysine-93 · · · glutamic acid-96. The side chains have been fit into the appropriate x-ray densities using molecular graphics. With the possible exception of lysine-91 · · · glutamic acid-95, the remaining pairs have side chain orientations which are consistent with salt-bridge and/or electrostatic interactions. The electron density maps show that arginine-84 may be disordered and also involved in an interhelical salt bridge with glutamic acid-64 of the C-helix. The structure also shows that the glutamine-85 · · · aspartic acid-89 pair, though not a salt bridge, can be involved in a hydrogen bond. These salt bridges appear to virtually surround the DE-linker helix. In addition there is a potential for a salt bridge between aspartic acid-100 and arginine-103 in the C-terminal third (E-helix) of the long helix where the x-ray results show the carboxyl group of aspartic acid-100 meets the arginine-103 side chain. These interac-

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.

Table 1. Potential salt bridges, electrostatic interactions, and hydrogen bonds between side chains in the x-ray structure of troponin C

Residue, *i*	Arg-84	Gln-85	Glu-88	Asp-89	Lys-91	Lys-93	Glu-95	Glu-96	Asp-100	Arg-103
i − 4			Arg-84 (±/HB)	Gln-85 (HB)		Asp-89 (±)	Lys-91 (±)			
i − 3					Glu-88 (±/HB)			Lys-93 (±)		Asp-100 (±/HB)
i + 3			Lys-91 (±/HB)			Glu-96 (±)			Arg-103 (±/HB)	
i + 4	Glu-88 (±/HB)	Asp-89 (HB)		Lys-93 (±)	Glu-95 (±)					

±, salt bridges and/or electrostatic interactions; HB, hydrogen bonds.

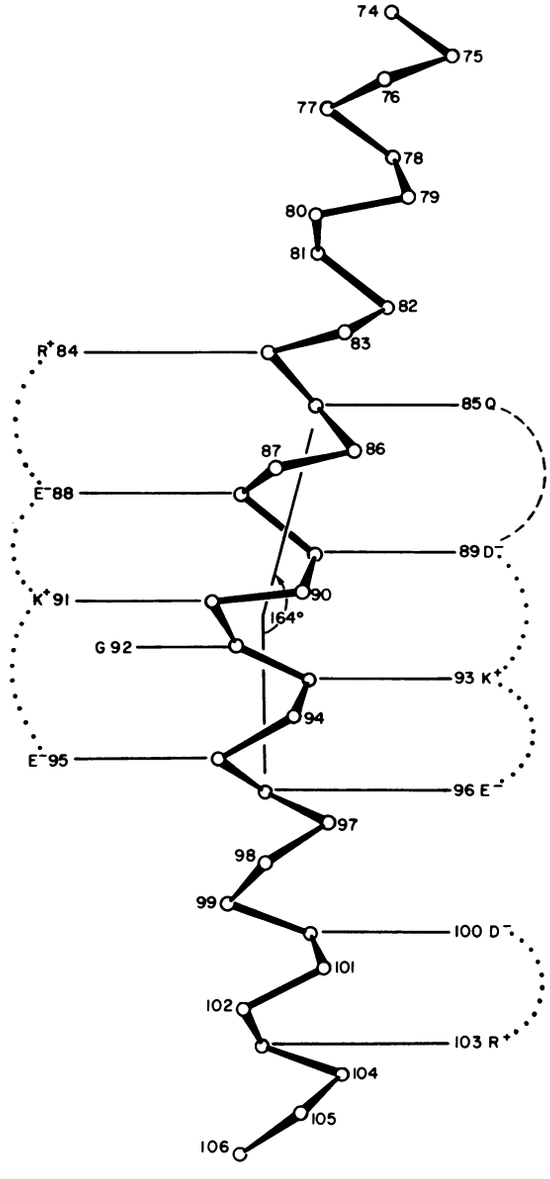

FIG. 1. The nine-turn long central helix of troponin C comprising residues 74–106. This helix is composed of the D-helix at the amino-terminal end, the DE-linker helix in the middle, and the E-helix at the carboxyl-terminal end. Each of these helical segments is about three turns. Notice the possible electrostatic and/or salt-bridge interactions (· · ·) and hydrogen bonds (----) between neighboring side chains ($i \pm 3$ or 4) that are concentrated in the DE-linker, thus, stabilizing the exposed helix. The marked bend in the helix at glycine-92 is indicated.

tions, in addition to stabilizing the helix, will restrain the amino acid side chains to certain preferred conformations.

It may be added that there appears to be some potential for repulsive interactions between pairs of like charges separated by $i \pm 3$ or 4 residues apart—i.e., arginine-84/lysine-87 and glutamic acid-97/aspartic acid-100. But as shown in Fig. 1 and Table 1 the residues arginine-84 and aspartic acid-100 have also the potential to engage in attractive interactions. The attractive interactions between these residues plus the other potential pairs discussed above are expected to outweigh the repulsive interactions. In fact the juxtaposition of such a distribution of charged residues may enhance the formation of potential salt bridges discussed above.

Sequence Comparisons and Potential Salt Bridges in Calcium-Binding Proteins. The amino acid sequences of troponin C, calmodulin, and the myosin regulatory light chains in the region of the long helix are shown in Table 2. The sequences have been aligned with respect to that of chicken troponin C. There is a great degree of homology among the residues that show potential for forming salt bridges. Rabbit and human troponin Cs differ from chicken only at residue 100 where there is a glutamic acid replacing aspartic acid. Frog troponin C has a glutamine at residue 91 instead of a lysine, a replacement that will not allow a salt bridge but the formation of a hydrogen bond with either glutamic acid-88 or -95. Ascidian troponin C has not only more extensive changes in the sequence but also has 4 additional residues in the linker. In spite of these alterations, the potential salt bridges glutamic acid-88 · · · lysine-91 and glutamic acid-100 · · · arginine-103 are retained. An important mutation is the switch of lysine-93 in chicken to glutamic acid in the ascidian protein. This change is accompanied by an insertion of a lysine between residues 96 and 97 to allow the formation of new salt bridges to glutamic acid-93 and/or -100. The switch of chicken serine-94 to arginine in the ascidian further allows an additional salt bridge with glutamic acid-97, whereas the switch in chicken lysine-87 to glutamine in ascidian will not allow a salt bridge with arginine-84, but may permit a hydrogen bond with residue 84. Most of the changes and insertions in the sequence are associated with a high helix potential. In ascidian troponin C the insertions would lengthen the central helix to about 10 turns, and the switch of the glycine-92 to proline will accentuate the break in the helix as discussed below. Cardiac troponin C sequences have only a single change among the potential salt-bridge side chains—an aspartic acid substitution for glutamic acid-88 in chicken skeletal muscle troponin C.

Several calmodulin sequences have also been aligned with troponin C in Table 2. Most of the potential salt-bridge residues are either identical to those in troponin C or conservative replacements. These include arginine-74, lysine-77, aspartic acid-78, glutamic acid-82, glutamic acid-87, and arginine-90. When compared to troponin C, calmodulin has the deletion of three residues and amino acid changes at

Table 2. Amino acid sequences of the central helix region of troponin C, calmodulin, and myosin regulatory light chains from various tissues and species

Troponin C																																					
(Skeletal muscle)	74*						80					85							90						95						100						106
											†	†					†	†		†			†		†	†					†			†			
Chicken (8)	D	F	E	E	F	L	V	M	M	V	R	Q	M	K	–	–	E	D	A	K	–	G	K	S	E	E	–	E	L	A	D	C	F	R	I	F	D
Rabbit (9) & human (10)															–	–					–						–				E						
Frog (11)															–	–				Q	–						–										
Ascidian (12)						C	L			Y				Q	A	Q		E			I	P	E	R			K			S	E	A			L		
(Cardiac)																																					
Bovine (13) & rabbit (14)			D									C			–	–	D		S		–						–			S		L			M		
Calmodulin	64*						70					75							80										85					90			93
											†			†			†		†						†					†	†			†			
Type A (15–19)	D	F	P	E	F	L	T	M	M	A	R	K	M	K	–	–	D	T	D	–	–	–	–	S	E	E	–	E	I	R	E	A	F	R	V	F	D
Type B (20–22)			P				S	L		A		K		R	–	–	D	S	D	–	–	–	–				–		I	R	E	A			V		
Type C (23)			P				S	L		A		K			–	–	D	T	D	–	–	–	–				–			I	E	A		K	V		
Myosin regulatory light chains	78*								85								90							95						100							107
												†						†					†														
Rabbit skeletal (24)	N	F	T	V	F	L	T	–	M	F	G	E	–	K	–	–	L	K	–	–	G	A	D	P	E	D	–	V	I	T	G	A	F	K	V	L	D
Chicken skeletal (25)								–					–		–	–			–	–							–			M							
Chicken cardiac (25)								–					–		–	–			–	–		T				E	–	T		L	N						
Chicken gizzard (25)								–					–		–	–			–	–		T					–			R	N						

The one letter code for amino acid residues is used.

*The numbers correspond to the residue positions of the first species in each protein class.

†Potential salt bridge or hydrogen bond residues.

residues 79 and 80. Calmodulin aspartic acid-80 can now bridge to lysine-77, and the aspartic acid-78 has lost its partner (lysine-93) by deletion. Residue 86 is arginine in two of the three types of calmodulin, which may make a salt bridge with glutamic acid-82. Thus, there appear to be the following possible salt bridges: arginine-74 · · · aspartic acid-78, lysine-77 · · · aspartic acid-80, glutamic acid-82 · · · arginine-86, and glutamic acid-87 · · · arginine-90.

Previous studies have shown sequence homology between troponin C and the myosin regulatory light chains (26). The number of potential salt bridges is more limited in the light chains; only glutamic acid-88 · · · lysine-91 and lysine-91 · · · aspartic acid-94 (Table 2) appear possible. The location of these charged residues is identical among the tissues and species shown.

It is interesting to note the positions of the glycine and proline residues in the various proteins listed in Table 2. As mentioned earlier, the glycine appears to be the pivot point in the bend of the central helix of chicken troponin C. This glycine is among the three residues deleted in the calmodulins. The ascidian troponin C has a proline residue replacing chicken glycine-92. There is also a proline residue at position 95 in the myosin light chain (in all tissues and in the two species shown) immediately adjacent to a potential salt-bridged residue. Proline-95 in the myosin light chains corresponds to serine-94 in troponin C. These observations suggest that this region in troponin C and the light chains may provide a bend point for these molecules. The proline would cause a break (though not necessarily disrupting the helical path of the polypeptide chain) in the long helix and may cause the light chains and ascidian troponin C to adopt a more boomerang shape. The absence of glycine or proline in the corresponding region of calmodulin raises the question of whether the mechanism of action in these proteins involves a major angular displacement of the helix. The salt bridges themselves may be responsible for the bend in the helix since a slight bend can be seen in calmodulin, which lacks glycine and proline in the central region.

The lengths of the central helix of the various calcium-binding proteins in Table 2 appear to be related by integral helical turns. Thus, chicken troponin C has 9 turns, ascidian troponin C has 10 turns (4 amino acid insertions relative to chicken troponin C) and calmodulin (4, ‡) and the myosin regulatory light chains each have 8 turns (3 deletions relative to chicken troponin C). This variation in the length of the long helix maintains similar rotational positions of the amino and carboxyl domains. The length of the helix and the degree of bend may be coupled for the expression of the specific regulatory properties of these molecules.

The helix-loop-helix pattern of troponin C (skeletal and cardiac), calmodulin, and the myosin light chains conform to a four domain structure regardless of whether all the domains are capable of binding divalent metal ions. It is interesting to note that in these structures one finds several different combinations and permutations of metal binding. It may be predicted that cardiac troponin C and myosin light chain 2 will also assume approximate dumbbell shapes like troponin C and calmodulin, although light chain 2 may be more bent in the central helix.

Salt Bridges in Tropomyosin and Troponin T. Intrahelical salt bridges may also be an important structural feature of other muscle proteins. It is interesting to note that tropomyosin, a double stranded fully α-helical protein, has 26 of its 57 glutamic acid side chains in positions with potential for forming intrahelix salt bridges with lysine, arginine, or histidine. An additional 14 potential salt bridges may be formed involving aspartic acid. Hence, there is a potential for as many as 40 intrahelical salt bridges in contrast to only 9 potential interhelical salt-bridge interactions between the two adjacent polypeptide chains (27). Thus, besides the hydrophobic interactions and interchain salt bridges (27) between the two α-helical chains of tropomyosin, intrachain salt bridges may contribute to the protein's stability and properties. It may also be added that inspection of the amino acid sequence of the myosin rod indicates the presence of numerous possible intrahelical salt bridges.

‡Sundaralingam, M. & Rao, S. T. (1985) Book of Abstracts, Fourth International Convention in Biomolecular Stereodynamics, June 4–8, 1985, State Univ. of New York at Albany, pp. 280—281 (abstr.).

The sequence of troponin T (28), the tropomyosin binding component, suggests numerous intrahelical salt bridges, particularly concentrated in the central region, residues 81–145. Thus a long central helix of approximately 18 turns, stabilized by salt bridges, may be present in this protein. Furthermore, circular dichroism measurements on the cyanogen bromide peptide of residues 71–151 indicate a helical content of approximately 80% (29) which seems to be consistent with the proposed salt-bridge stabilization. It is tempting to speculate that the ubiquitous occurrence of salt bridges in the muscle proteins may be related to their unusual elongated structures (virtually totally α-helical or dumbbell shaped) and to the conformational changes involved in the contraction process.

The possibility of interhelical salt bridges has been recognized for some time. In troponin C itself a third of the long helix (helix D) is buried in the N-terminal domain and is stabilized by hydrophobic interactions with the four surrounding helices as well as interhelical salt bridges. The other end of the long helix (helix E) is only partially buried in the C-domain and may derive additional stabilization from the potential intrahelical salt bridge between aspartic acid-100 and arginine-103 as shown in Fig. 1. The exposed middle third of the long helix (DE linker) appears to be mainly stabilized by intrahelical salt bridges. The number of potential intrahelical salt bridges may be related to the degree of exposure of α-helices. The unusual structure of troponin C provides a good case in support of the importance of intrahelical salt bridges between neighboring charged residues in stabilizing protein secondary structure.

This research was supported by Grant AM 34139 from the National Institutes of Health. The continued support of the College of Agricultural and Life Sciences and the Graduate School of the University of Wisconsin-Madison is also acknowledged.

1. Sundaralingam, M., Bergstrom, R., Strasburg, G., Rao, S. T., Roychowdhury, P., Greaser, M. & Wang, B. C. (1985) *Science* **227**, 945–948.
2. Sundaralingam, M., Rao, S. T., Drendel, W. & Greaser, M. L. (1985) *Int. J. Quantum Chem.*, in press.
3. Herzberg, O. & James, M. N. G. (1985) *Nature (London)* **313**, 653–659.
4. Babu, Y. S., Sack, J. S., Greenhough, T. H., Bugg, C. E., Means, A. R. & Cook, W. J. (1985) *Nature (London)* **315**, 37–40.
5. Scheraga, H. A. (1978) *Pure Appl. Chem.* **50**, 315–324.
6. Kretsinger, R. H. & Barry, C. D. (1975) *Biochim. Biophys. Acta* **405**, 40–52.
7. Maxfield, F. R. & Scheraga, H. A. (1975) *Macromolecules* **8**, 491–493.
8. Wilkinson, J. M. (1976) *FEBS Lett.* **70**, 254–256.
9. Collins, J. H., Greaser, M. L., Potter, J. D. & Horn, M. C. (1977) *J. Biol. Chem.* **252**, 6356–6372.
10. Romero-Herrera, A. E., Costillo, O. & Lehmann, H. (1976) *J. Mol. Evol.* **8**, 251–270.
11. Van Eerd, J.-P., Capony, J.-P., Ferraz, C. & Pechere, J.-F. (1978) *Eur. J. Biochem.* **91**, 231–242.
12. Takagi, T. & Konishi, K. (1983) *J. Biochem. (Tokyo)* **94**, 1753–1760.
13. Van Eerd, J.-P. & Takahaski, K. (1976) *Biochemistry* **15**, 1171–1180.
14. Wilkinson, J. M. (1980) *Eur. J. Biochem.* **103**, 179–188.
15. Sasagawa, T., Ericsson, L. H., Walsh, K. A., Schreiber, W. E., Fischer, E. H. & Titani, K. (1982) *Biochemistry* **21**, 2565–2569.
16. Watterson, D. M., Sharief, F. & Vanaman, T. C. (1980) *J. Biol. Chem.* **255**, 962–975.
17. Grand, R. J. A. & Perry, S. V. (1978) *FEBS Lett.* **92**, 137–142.
18. Dedman, J. R., Jackson, R. L., Schreiber, W. E. & Means, A. R. (1978) *J. Biol. Chem.* **253**, 343–346.
19. Putkey, J. A., Ts'ui, K. F., Tanaka, T., Lagace, L., Stein, J. P., Lai, E. C. & Means, A. R. (1978) *J. Biol. Chem.* **258**, 11864–11870.
20. Toda, H., Yazawa, M., Kondo, K., Honma, T., Narita, K. & Yagi, K. (1981) *J. Biochem (Tokyo)* **90**, 1493–1505.
21. Takagi, T., Nemoto, T., Konishi, K., Yazawa, M. & Yagi, K. (1980) *Biochem. Biophys. Res. Commun.* **96**, 377–381.
22. Jamieson, G. A., Bronson, D. D., Schachat, F. H. & Vanaman, T. C. (1980) *Ann. N. Y. Acad. Sci.* **365**, 1–13.
23. Stein, J. P., Munjaal, R. P., Lagace, L., Lai, E. C., O'Malley, B. W. & Means, A. R. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 6485–6489.
24. Matsuda, G., Maita, T., Suzuyama, Y., Setoguchi, M. & Umegane (1977) *J. Biochem. (Tokyo)* **81**, 809–811.
25. Matsuda, G. (1983) *Adv. Biophys.* **16**, 185–218.
26. Collins, J. H. (1976) *Nature (London)* **259**, 699–700.
27. McLachlan, A. D. & Stewart, M. (1975) *J. Mol. Biol.* **98**, 293–304.
28. Pearlstone, J. R., Carpenter, M. R., Johnson, P. & Smillie, L. B. (1976) *Proc. Natl. Acad. Sci. USA* **73**, 1902–1906.
29. Pearlstone, J. R. & Smillie, L. B. (1977) *Can. J. Biochem.* **55**, 1032–1038.