Acta Cryst. (1994). D**50**, 40–49

Structure of Chicken Skeletal Muscle Troponin C at 1.78 Å Resolution

BY K. A. SATYSHUR AND D. PYZALSKA

Crystallography Laboratory, Department of Biochemistry and College of Agriculture and Life Sciences, University of Wisconsin, Madison, WI 53706, *USA*

M. GREASER

Muscle Biology Research Laboratory, College of Agriculture and Life Sciences, University of Wisconsin, Madison, WI 53706, *USA*

AND S. T. RAO AND M. SUNDARALINGAM*

Crystallography Laboratory, Department of Biochemistry and College of Agriculture and Life Sciences, University of Wisconsin, Madison, WI 53706, *USA, and Laboratory of Biological Macromolecular Structure, Department of Chemistry, Biochemistry and Biotechnology Center, The Ohio State University,* 1060 *Camark Road, Columbus, OH* 43210, *USA*

(*Received* 9 *September* 1992; *accepted* 27 *July* 1993)

Abstract

The structure of chicken skeletal muscle troponin C (TnC) has been refined to an *R* value of 0.168, using 14 788 reflections, in the resolution range 8.0–1.78 Å. Our earlier 2 Å resolution structure [Satyshur, Rao, Pyzalska, Drendel, Greaser & Sundaralingam (1988). *J. Biol. Chem.* **263**, 1628–1647] served as the starting model. The refined model includes atoms for all protein residues (1–162), 2 Ca^{2+} ions, 169 water molecules and one sulfate ion. The high-resolution refinement shows more clearly the details of the protein and water structure. The side chains Glu63, Cys101, Arg123, Asp140 and Asp152 adopt two discretely ordered conformations. The long central helix is only slightly curved/bent (7.9°) and all the central helix NH···O=C hydrogen bonds are intact. Seven of the nine carbonyl O atoms of the mid segment of this helix, including the *D/E* linker region, are hydrogen bonded to water molecules which weakens the helix hydrogen bonds. In contrast, in each of the protected upper and lower thirds of the long central helix, only two carbonyl O atoms are hydrogen bonded to water molecules. The hydrogen-bonding patterns displayed by some of the carbonyl O atoms of NT and *A* helices of the N-terminal domain and the *F* and *H* helices of the C-terminal domain, which are on the exposed surface of the protein, are similar. The *B* helix of the calcium-free site I is kinked, with the local helix axes at either end making an angle of 39°, by two inserted water molecules between N—H and O=C groups, breaking the adjacent helix hydrogen bonds. A sulfate ion from the crystallization buffer is also trapped in the *B* helix between the guanidinium group of Arg47 and these two inserted water molecules. The *C* helix of site II is devoid of similar hydration and is probably responsible for the different interhelical angles *A/B* at site I (134°) and *C/D* at site II (149°). Extensive interhelix hydrogen bonds occur between the side chains of the *C* and *D* helices of the 'apo' site II: Gln51–Asp89, Asn52–Asp89, Glu57–Gln85, Glu57–Glu88 and Glu64–Arg84, which apparently are disrupted upon Ca uptake and the resulting rearrangement of the helices expose the side chains, lining the palm of the N-(and C-) terminal domains, for interaction with specific peptide fragment of troponin I (TnI) during muscle contraction. The dominant crystal packing motif involves a head-to-tail interaction between the N-terminal domain *A* helix of one molecule and the palm of the C-terminal domain of the 3_2-related molecule, in a manner similar to that which can be expected for the TnC–TnI complex. Similar interactions may also be responsible for the dimerization of TnC at low pH.

Introduction

Troponin is a trimeric protein consisting of troponin T (TnT), troponin I (TnI) and troponin C (TnC) (Greaser & Gergely, 1971). Troponin C undergoes a Ca^{2+}-activated conformational change which triggers muscle contraction (Leavis & Gergely, 1984). The structure of chicken skeletal muscle TnC (Satyshur *et al.*, 1988) and the closely related turkey TnC (Herzberg & James, 1988) have been determined at 2 Å resolution. In this paper we report

* To whom correspondence should be addressed.

ISSN 0907-4449 ©1994

the refinement of our earlier 2.0 Å resolution X-ray structure of chicken skeletal muscle troponin C at a higher resolution of 1.78 Å, containing nearly 80% more data. The present work confirms our earlier 2 Å structure and additionally reveals the presence of a bound sulfate ion in the *B* helix, multiple conformations for some side chains, and greater details of protein solvation, especially the hydration of the α-helices.

Experimental

Data collection

The crystals of chicken skeletal TnC were grown using the conditions described earlier (Satyshur *et al.*, 1988) and belong to the trigonal space group $P3_221$, $a = b = 66.7$, $c = 60.8$ Å with one molecule in the asymmetric unit. High-resolution intensity data were collected at the multiwire (Xuong–Hamlin type) area-detector facility, University of Virginia, Charlottesville, using a Rigaku RU200 Cu-rotating anode operating at 7.5 kW (50 kV, 150 mA) equipped with a graphite monochromator. A single crystal of dimensions $0.4 \times 0.2 \times 0.2$ mm was mounted with the diagonal of the a^*c^* parallel to the φ axis of the goniostat and cooled to 263 K with a cold air stream. At this temperature the crystal lasted for about 8 d in the X-ray beam, compared to 5–6 d at room temperature with a sealed X-ray tube. Two detectors where placed at 2θ angles of -35.0 and $+34.6°$ at a distance of 427 and 430 mm, respectively, from the crystal and the path filled with helium. This limited the resolution range from 5 to 1.78 Å. Data were collected in two positions of χ, $+20$ and $-15°$, using φ scans from 0 to 360°, in 0.25° intervals and a counting time of 20 s. Nearly 62000 reflections were recorded for each φ rotation on each detector giving a total of 253002 reflections which were scaled together using the area-detector facility programs. The R_{scale} for the equivalent reflections $(\sum\sum|F_i| - \langle|F_i|\rangle|/\sum\sum\langle|F_i|\rangle|)$, as a function of resolution, varied between 0.047 and 0.054, and the data were uniformly complete in the 5.0–1.78 Å resolution range (Table 1). Since the crystal had deteriorated, low-resolution intensity data ($d \geq 5$ Å) could not be collected on the area detector. The data between 8 and 3 Å resolution, collected earlier on an Enraf–Nonius CAD-4 diffractometer (Satyshur *et al.*, 1988), were scaled to the area-detector data, using 1966 common reflections ($F > 3\sigma$, $5 \geq d \geq 3$ Å) and R_{scale} was 0.045. The common reflections between the two data sets were not averaged but the area-detector data were augmented by 555 diffractometer data with $8 \geq d \geq 5$ Å. The composite data set, used in the refinement, contained a total of 14788 reflections with $F > 3\sigma$ and $8 \geq d \geq 1.78$ Å (97.5% complete, Table 1).

Table 1. *Possible and observed number of reflections in TnC*

Resolution range (Å)	Possible reflections	Observed with $F > 3\sigma$	% of total
8.00–5.00	555	476	85.8
5.00–3.45	1465	1435	97.8
3.45–2.91	1418	1413	99.6
2.91–2.60	1439	1436	99.8
2.60–2.39	1403	1396	99.5
2.39–2.23	1400	1382	98.7
2.23–2.11	1415	1393	98.4
2.11–2.01	1392	1373	98.6
2.01–1.93	1413	1388	98.2
1.93–1.86	1377	1338	97.2
1.86–1.80	1363	1309	96.0
1.80–1.78	526	449	85.4
8.00–1.78	15168	14788	97.5

Refinement

The starting model for the refinement was our 2.0 Å structure with 1248 protein atoms for residues 3–162, 2 Ca^{2+} and 68 water molecules. The *PROLSQ* least-squares program (Hendrickson, 1985) was initially used on the San Diego supercomputer, Cray-XMP. Latter rounds of refinement were carried out using *X-PLOR* (Brünger, 1992) on the Ohio State YMP-8/64 supercomputer. All model fitting was carried out on our Evans and Sutherland PS340 and ESV-10 molecular graphics systems, using the program *FRODO* (Jones, 1985). In the initial refinement cycles *PROLSQ* was used with 2 Å data and the first two amino-terminal residues were not included but the two carboxy-terminal residues Val161 and Gln162 were placed in tentative positions. The *R* value dropped from 0.277 to 0.181. Further refinement was carried out using *X-PLOR*. The resolution was gradually increased from 2.0 to 1.78 Å and the protein model was refined using $2w|F_o| - |F_c|$ omit maps, calculated by omitting five residues at a time, where w is the Sim weight (Sim, 1958) for the reflection, which is a function of the ratio of the product $F_o \times F_c$ and the r.m.s. deviation between them. In general, the 1.78 Å resolution electron-density maps showed the molecular features more clearly than in the 2 Å analysis and the apo sites I and II could now be better fitted and the corrected amino-terminal sequence 1-Ala-Ser-Met-Thr-4, deduced from the cDNA (Rainach & Karlsson, 1988) could be built, though the electron density for these regions was still weak. The water molecules in the 2.0 Å structure were checked and five sites with low electron density were removed. Additional solvent sites were located and added to the model, if they had electron densities $> 2\sigma$ in the $F_o - F_c$ difference maps and were within 3.4 Å of any polar atom of the protein, or already established solvent sites, and the hydrogen-bond angles were within 30° of expected values. All solvent sites were assigned unit occupancy, which was not varied during the refinement. Continuing the map fitting

and refinement yielded additional solvent sites, many in the second coordination sphere, which, in general, had lower electron densities and higher thermal factors. At each stage, the electron densities for previously established water molecules in a difference omit map were inspected and those sites which had density below 2σ were discarded from the model. A sulfate ion was identified, which is trapped between two inserted water molecules in the *B* helix and Arg47, also in the *B* helix (Fig. 1). The side chains of residues Glu63, Cys101, Arg123, Asp140 and Asp152 were discretely ordered over two states. Their occupancies were refined and assumed values varying between 0.45 and 0.55. Thus, the occupancies of the discretely ordered side chains were fixed at 0.5 during further refinement. In the final round of refinement the model consisted of all the atoms of the protein

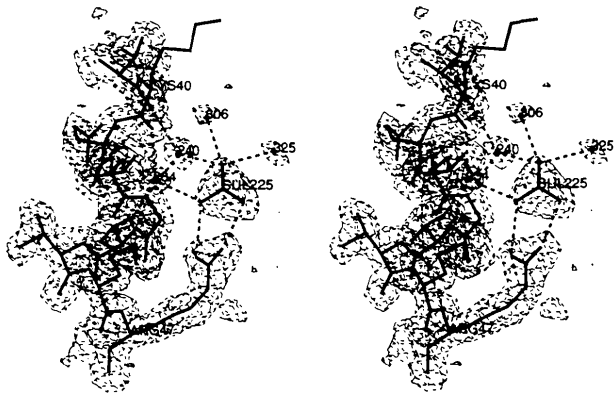

Fig. 1. $F_o - F_c$ electron-density map for the *B*-helix region, the two inserted water molecules (234 and 240) and the sulfate ion (225), which were omitted from the phasing calculations. Contours are at 1.5σ above the mean. The sulfate ion forms hydrogen bonds (dashed line) with NH1 and NH2 of Arg47 and the two inserted water molecules. The two inserted water molecules are 3.44 Å apart.

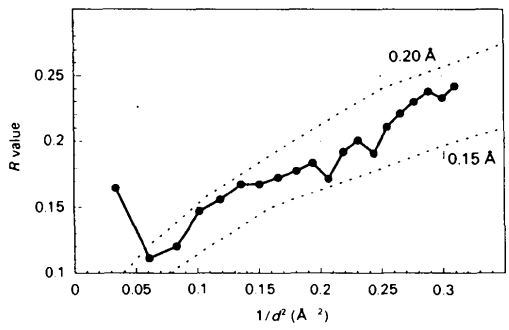

Fig. 2. Luzzati plot for TnC. The *R* values as a function ($1/d^2$) are shown (solid circles, solid lines). Dashed lines represent the *R* values expected for mean positional errors of 0.20 Å (top line) and 0.15 Å (bottom line). The estimated average positional error in the model is 0.18 Å.

Table 2. *Summary of parameters for the present* 1.78 *and the previous* 2.0 Å *studies*

Description	Present 1.78 Å study	Previous 2.0 Å study
Crystal system, space group	Trigonal, $P3_221$	Same
Unit-cell constants (Å)	$a = 66.7$, $c = 60.8$	Same
Data-collection method	Area detector	Diffractometer/ oscillation
Temperature (K)	263	294 (RT)
Limiting resolution (Å)	1.78	2.00
No. of reflections used in the refinement (% completeness)	14788 (97.5)	8100 (75)
Program used for refinement	*X-PLOR/PROLSQ*	*PROLSQ*
Parameter file	param19.pro	—
Final *R* value	0.165	0.172
Model		
Protein residues	1–162 (all)	3–162
Ca^{2+} ions	2	2
SO_4^{2-} ion	1	0
Solvents (water molecules)	169	68
Mean positional error (Å) (Luzzati)	0.18	0.21*
R.m.s. deviations from ideal values		
Bond lengths (Å)	0.012	0.04
Bond angles (°)	2.3	4.6†
Torsion angles (°)	20.6	25.5
Average *B* value for protein atoms (Å²)	25.6	26.9
R.m.s. *B* values/target (Å²)		
Backbone, bonded (1–2)	6.4/5.0	8.2/5.0
Backbone, angle (1–3)	7.8/7.0	8.8/5.0
Side chain, bonded (1–2)	8.2/6.0	11.7/5.0
Side chain, angle (1–3)	9.2/8.5	13.5/5.0

* The previously quoted value of 0.17 Å was an underestimate.

† Deduced from r.m.s. deviation of 0.07 Å in 1–3 distances and assuming average values of 1.5 Å and 110° for bond lengths and bond angles, respectively.

residues 1–162, two Ca^{2+} ions, one sulfate ion and 169 water molecules, with an *R* value to 0.165 for 14788 reflections. Of the 169 water sites, 142 are in the first coordination sphere (at least one polar protein atom within a distance of 3.4 Å), 21 in the second coordination sphere (only other water molecules within a distance of 3.4 Å) and six at distances greater than 3.4 Å from both protein atoms and other solvents. The solvent content of the TnC crystals is 43% by volume and the number of stoichiometric waters in the asymmetric unit is estimated to be ~480. Thus, the ordered water molecules constitute about 35% of the total. The *R* value for all 15168 reflections, in the resolution range 8–1.78 Å, is 0.174. A summary of the present 1.78 Å refinement statistics, compared with our earlier 2.0 Å refinement, is given in Table 2. The Luzzati (1952) plot (Fig. 2) suggests a mean coordinate error of 0.18 Å. The atomic coordinates and the structure amplitudes have been deposited with the Brookhaven Protein Data Bank* (Bernstein *et al.*, 1977).

* Atomic coordinates and structure factors have been deposited with the Protein Data Bank, Brookhaven National Laboratory. Free copies may be obtained through The Technical Editor, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England (Supplementary Publication No. SUP 37096). At the request of the authors, the structure factors will remain privileged until 1 September 1994. A list of deposited data is given at the end of this issue.

Results and discussion

A ribbon diagram of the TnC molecule is shown in Fig. 3. Plots of the average B values of the residues, for only the backbone atoms and all the atoms, are given in Fig. 4(*a*). A C^{α}-atom plot, with the atomic radii proportional to the r.m.s. atomic displacement, is shown in Fig. 4(*b*). The largest atomic motions are at either termini, followed by the loop regions of the two apo sites and the linker region of the central helix, which is exposed to solvent.

Comparison with the 2 Å *model*

The r.m.s. atomic deviations between the present 1.78 model and the earlier 2 Å model are: 0.32 for all atoms, 0.25 for the backbone atoms and 0.19 Å for the C^{α} atoms. The residues with largest deviations are in the loop regions of sites I and II and the linker region of the central helix, all with high thermal parameters and weak electron densities. The mean deviations in the torsion angles between the two models are: φ (7°), ψ (8°) and χ_1 (20°). Significant changes in the backbone torsion angles are at residues Asp30 and Ala31 in site I. Fig. 5 shows the Ramachandran plot, in which the φ, ψ angles for residues 30–41 are shown. Residues 4, 6, 32, 52, 67 and 94 now have χ_1 values differing from the 2 Å model by more than 60°. The Ca-coordination distances for sites III and IV are listed in Table 3 and compared with the values from the 2 Å analysis. The mean values at the two sites are now closer together and also have smaller r.m.s. deviations from the mean value. The average B values for the protein (24.0 Å^2) and backbone atoms (17.9 Å^2) are now

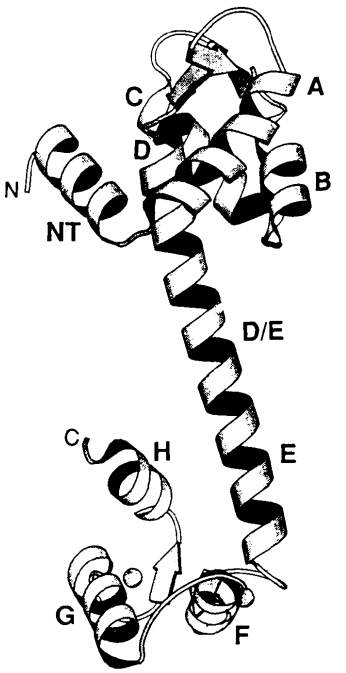

Fig. 3. Ribbon diagram of the troponin C molecule drawn with *MOLSCRIPT* (Kraulis, 1991). The calcium-free apo N-domain is on top and the two-calcium-bound holo C-domain is at the bottom. The two bound metal ions are shown as stippled spheres. There are eight helices: NT (residues 3–12), *A* (16–27), *B* (42–48), *C* (55–65), the single long central helix comprising of the *D*, *D*/*E*-linker helix and the *E* helix (75–105), *F* (115–124), *G* (131–141) and *H* (151–159). The residues in the four calcium-binding loops are: 1 (30–41), 2 (66–77), 3 (106–117) and 4 (142–153). There is a short anti-parallel β-sheet between loops 1 (35–39) and 2 (71–75) in the N-terminal domain and another between loops 3 (113) and 4 (149) in the C-terminal domain.

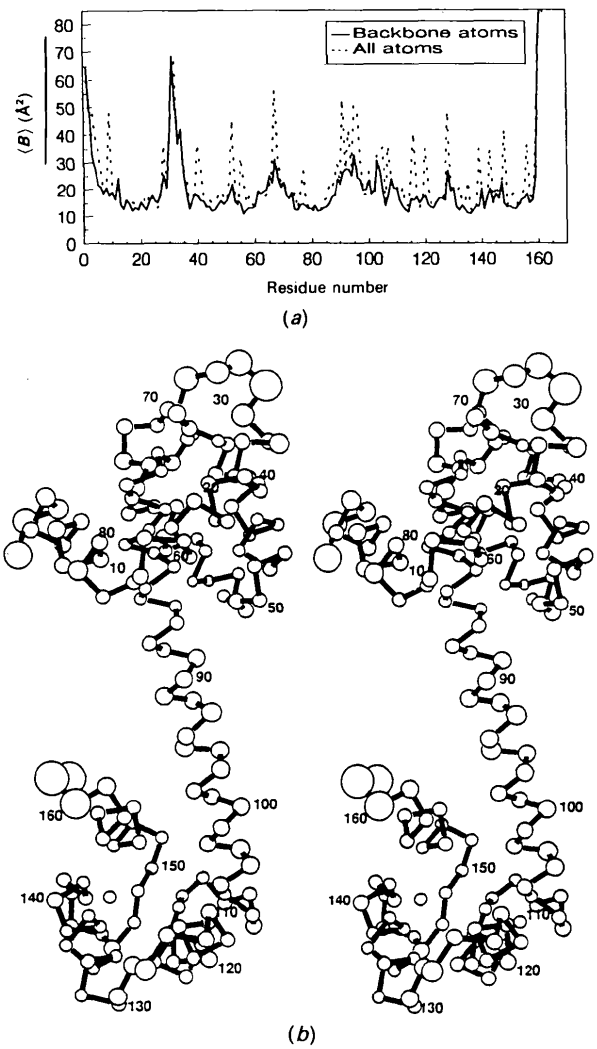

Fig. 4. (*a*) Plot of the $\langle B \rangle$ values for the various residues in TnC molecule. Solid line is for backbone atoms and the dashed line is for all atoms of the residues. (*b*) C^{α} plot of TnC in stereo (view similar to Fig. 3), with atomic radii equal to the r.m.s. atomic displacement deduced from $\langle B \rangle$ for the backbone atoms of each residue. Both termini, as well as the loop regions of the apo sites I and II and the linker region of the central helix, have high mobilities.

lower than the corresponding values (26.9 and 22.2 Å^2) in the 2 Å study, presumably as a result of the lower temperature at which the present intensity data were collected. The 169 water sites have *B* values ranging from 14 to 110 Å^2 and 95 are below the mean value of 50 Å^2. Thus, the present high-resolution model is quite similar to the earlier 2 Å model. However, in the present model, the previously missing N-terminal residues could be modeled, the mean positional error is lower, the r.m.s. deviations in the bond distances and angles from the library values are smaller and more than twice the solvent sites have been found leading to a better characterization of the solvation of the protein. The details of the present model as well as the distortions observed at the helix termini are presented below.

Quality of the electron-density maps and multiple conformations for side chains

The side chains Ala31 and Asp32 of the apo site I and the backbone atoms of the three Gly residues 33, 34 and 35 showed the weakest density. The N-terminal residues (*N*-acetyl-Ala1 and Ser2) as well as the C-terminal residues (Ala61 and Gln162) could be built into the electron density. The long side chains Glu9, Glu67, Glu95, Glu120, Lys91, Lys93 and His128 showed little or no electron density beyond the C$^\gamma$ atoms, while the side chains of Glu63, Cys101, Arg123, Asp140 and Asp152 were discretely ordered over two sites (Table 4). The guanidinium group of Arg123 occupies two positions, one stacked over the hydrogen-bonded pair Glu64···Arg84 and the other over the Glu6 of the N-terminal helix. The

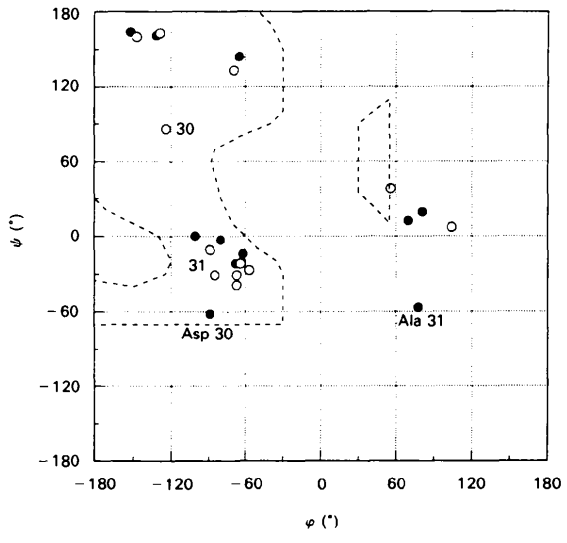

Fig. 5. Ramachandran plot for the residues 30–41 in the 'apo' site I. The φ, ψ angles in the present model are shown as filled circles and in the 2 Å model as open circles. The conformation of residues Asp30 and Ala31 differ in the two models.

Table 3. *Comparison of* Ca–*ligand distances* (Å) *in the* 1.8 *and* 2.0 Å *studies*

	Site 3			Site 4		
Desig-nation	Ligand	Distance (1.8 Å)	Distance (2.0 Å)	Ligand	Distance (1.8 Å)	Distance (2.0 Å)
+*X*	Asp106 OD2	2.22	2.24	Asp142 OD2	2.26	2.09
+*Y*	Asn108 OD1	2.16	2.43	Asp144 OD1	2.24	2.40
+*Z*	Asp110 OD2	2.43	2.45	Asp146 OD2	2.38	2.41
−*Y*	Phe112 O	2.08	2.38	Arg148 O	2.05	2.27
−*Z*	Glu117 OE1	2.53	1.99	Glu153 OE1	2.57	2.68
	Glu117 OE2	2.38	2.49	Glu153 OE2	2.44	2.36
−*X*	*W*176	2.41	2.87	*W*209	2.40	2.48
	Mean	2.32	2.41		2.33	2.37
	R.m.s. deviation	0.15	0.25		0.16	0.17

Table 4. *Conformations of discretely ordered side chains in TnC*

Location	Residue	Discretely ordered atoms*	Torsion angles $\chi_1,\chi_2,\ldots$ (°)
Site II	Glu63	CG, OD1, OD2	−69, 152, 23 150, −101, 103
E helix	Cys101	SG	177 −59
FG link	Arg123	ND, CE, NH1, NH2	−72, −173, 168, 94, −178 −71, −161, −79, 94, −123
G helix	Asp140	CG, OG1, OG2	−62, 180 −170, −54
Site IV	Asp152	CG, OD1, OD2	173, −109 −65, −109

* Refinement of the occupancies of the alternate conformations yielded values between 0.45 and 0.55. Occupancies of 0.5 were assigned to the two conformers and held fixed during further refinement.

stacking interactions between these planar fragments probably involve the π electrons of the guanidinium and carboxylate groups. Both positions of Glu63 are stabilized by hydrogen bonding. The two positions of S$^\gamma$ for Cys101 lie on the surface of the hydrophobic 'palm' of the C domain, near the hydrophobic side chains of the *A* helix from a neighboring molecule (see below). The S$^\gamma$ at either site can form a hydrogen bond with *W*281 (S$^\gamma$···*W* distances 2.50 and 3.05 Å). A similar two-site ordering for the same residue was also found in the related turkey TnC structure (Herzberg & James, 1988) and it is interesting that the major Hg site in the mersalyl derivative is in this hydrophobic cavity and close to both positions of S$^\gamma$. In the chicken TnC structure, this site is occupied by a water molecule (*W*281). The remaining backbone and side chains had clear electron densities.

Distortions at helix termini

All eight helices, including the N-terminal helix, are initiated either by serine (helices NT, *A*, *B*), threonine (helices *C*, *G*) or aspartate (helices *D*, *F*, *H*). The ends of these helices are often frayed (as is generally observed in protein helices) with the backbone carbonyl and amide groups invariably involved in hydrogen bonding to water molecules. Baker & Hubbard (1984) have classified these end distortions in helices into three patterns: the N-terminal (αN)

and two C-terminal (αC1, or αC2). In the αN type, the carbonyl of the first residue (i) forms both 4→1 and 5→1 hydrogen bonds to successive amide groups. In the αC1 type, the α-helical 5→1 hydrogen bond of residue i is followed by a 4→1 hydrogen bond with the ($i+1$) carbonyl. In the αC2 type the 5→1 hydrogen bond of residue i is followed by a 6→1 hydrogen bond involving the ($i+1$) carbonyl. The amino hydrogen that would have formed the 5→1 hydrogen bond with ($i+1$) carbonyl now hydrogen bonds to the carbonyl of residue $i+2$. In TnC, the αN type is found in the *B* helix, the αC1 type in the NT, *A* and *F* helices, and the αC2 type, also in the *B* helix.

Hydration of the backbone carbonyl and amide groups

Most of the ordered solvent molecules hydrate the protein carbonyl and amide groups. Of the 161 carbonyl groups of the backbone, nearly half (80) are hydrogen bonded to water molecules. A survey of proteins (Baker & Hubbard, 1984) shows a similar proportion. Of these carbonyls, 33 are in the helical regions. 29 of these 33 carbonyl groups, which are engaged in helix NH···O=C hydrogen bonds, are also involved in hydrogen bonding to an 'external' water molecule. This extra hydrogen bonding to water distorts the helix geometry and weakens the helix hydrogen bond by lengthening it and making the C=O···N angle less linear, Table 5 (Blundell, Barlow, Borkakoti & Thornton, 1983; Sundaralingam & Sekharudu, 1989). All eight helices of TnC have one or more carbonyl groups bound to a water molcule, particularly on the solvent-exposed side.

In contrast to the hydration of the carbonyl groups, only 20 amide nitrogens (or 12%) are hydrated in TnC (Table 6) and this is lower than about 25% found in other proteins. Two of the amide groups in the *B* helix are involved in hydrogen bonding to the two inserted water molecules.

Long central helix and hydration

The long central helix is composed of the *D* helix, *D/E* linker helix and the *E* helix. The helix is 46.2 Å long and contains about nine turns in 31 residues (Phe75 to Phe105). An ideal α-helix, with 3.6 residues/turn, will correspond to 8.6 turns. The two ends of the helices are involved in a number of interactions. A number of carboxylate–carboxylate interactions are found, arising from the low pH (5.0) of crystallization, in which a proton is shared between the two carboxylate groups. All the carboxylate–carboxylate interactions and salt bridges in chicken TnC, at distances up to 4 Å, are listed in Table 7. Similar interactions have been noted in turkey TnC (Herzberg & James, 1985). The side chains of the *D* helix and *C* helix interact. The first hydrophilic residue on the *D* helix, after emerging from the hydrophobic core of the N-terminal domain, Arg84, has a strong electrostatic interaction with Glu64 of the *C* helix (OE1 64···NH1 84, 2.90 Å). Gln85 interacts with Asp89 of the *D* helix (NE2 85···OD1 89, 3.13 Å) as well as with Glu57 of

Table 5. *Geometry involving backbone carbonyl O atoms in the helical regions hydrogen bonded to water molecules*

	Residue (i)	φ (°)	ψ (°)	Water No.	$O(i)\cdots W$ distance (Å)	$O(i)\cdots N(i+4)$ distance (Å)	C=O···N angle (°)
NT							
	Thr4	−66	−40	331	2.64	2.84	160
	Gln6	−63	−42	271	2.86	2.86	154
	Gln7	−62	−42	181	3.27*	2.93	158
	Ala8	−66	−38	222	2.71	3.03	147
A							
	Ala24	−60	−42	284	3.35*	—	—
B							
	Thr39	−68	−22	240	3.04	(4.78)	—
	Lys40	−63	−21	234	2.67	(5.40)	—
				306	2.98		
	Thr44	−61	−47	280	2.66	2.94	161
				234	2.67		
D/E							
	Glu76	−64	−38	181	2.91	3.30	147
	Val80	−54	−47	270	2.72	3.04	155
	Met81	−68	−40	200	2.93	2.99	153
	Arg84	−56	−48	274	2.81	3.03	156
	Met86	−61	−43	175	2.94	2.94	160
	Lys87	−66	−40	165	3.10	3.16	148
	Glu88	−68	−44	183	2.92	3.13	154
	Asp89	−68	−46	304	2.99	2.88	165
	Ala90	−55	−45	173	2.97	2.93	166
	Gly92	−69	−36	279	2.85	3.17	150
	Glu95	−58	−42	278	3.25*	3.03	155
F							
	Ile115	−67	−33	229	2.99	2.89	152
	Glu120	−69	−44	224	2.83	3.13	150
	Ile121	−64	−44	167	2.69	3.38	106
H							
	Asp152	−60	−47	291	2.86	3.18	154
	Glu153	−66	−31	244	2.74	2.80	152

* $O\cdots W$ distances between 3.2 and 3.4 Å are listed if the hydrogen-bond angle is comparable to those with shorter distances.

Table 6. *Amide groups hydrogen bonded to water molecules*

Water No.	NH of residue	$N\cdots W$ distance (Å)	C^{α}—N···W angle (°)
351	Met3	2.95	115
202	Ser15	2.79	121
221	Glu16	2.81	130
248	Ala31	2.83	111
204	Ser38	2.91	114
240	Gly43	2.85	110
234	Thr44	2.95	113
333	Glu67	2.93	104
334	Glu67	3.16	111
334	Asp68	2.89	118
268	Phe75	3.01	97
268	Glu76	3.28*	132
200	Glu85	3.38*	95
170	Ile115	2.91	121
195	His128	2.81	122
301	Val129	2.96	113
174	Glu131	2.93	123
191	Phe151	2.90	115
290	Phe151	3.27*	105
290	Asp152	3.15	120

* Distances between 3.2 and 3.4 Å are listed only when the corresponding angle lies between 90 and 150°.

Table 7. *Charge–charge (carboxyl–carboxylate and salt-bridge) interactions in chicken TnC*

Carboxyl–carboxylate	d (Å)	Salt bridges	d (Å)
Intramolecular			
Glu57–Glu88	2.76	Arg11–Glu76	2.91
Asp66–Asp68	3.83	Arg11–Glu16	3.95
Glu97–Asp100	3.74	Arg11–Glu76	2.88
Asp106–Asp110	3.09	Lys23–Asp27	2.83
Asp106–Glu117	3.46	Arg84–Glu64	2.90
Asp114–Glu116	3.73	Lys91–Glu95	3.46
Asp114–Glu117	3.77	Arg103–Asp100	3.63
Asp142–Asp146	3.00	Arg123–Glu131	3.09
Asp142–Glu153	3.54	Arg148–Asp146	3.67
Asp150–Glu153	3.64	Lys156–Asp152	3.97
Intermolecular			
Glu16–Glu127	2.76	Arg11–Glu127	2.85
Glu17–Glu63	3.50	Arg123–Glu64	3.47
Glu17–Asp133	3.74		
Glu63–Asp136	3.25		
Glu63–Asp140	2.73		
Glu67–Glu135	3.14		
Glu132–Glu132	3.55		
Glu132–Asp136	2.79		

the *C* helix (NE2 85···OE2 57, 3.27 Å). Glu57 also interacts with Glu88 in a carboxyl–carboxylate interaction (2.76 and 3.22 Å). Asp89 interacts with Gln51 and Asn52 of the *C* helix (OD2 89···NE2 51, 3.03; OD2 89···OD1 52, 2.97 Å). The region between the *C* and *D* helices is highly hydrated. The *E* helix at the C-terminal end is well protected. Arg103 forms an intramolecular salt bridge with Asp100 (3.63 Å) and Asp100 is engaged in a carboxyl–carboxylate interaction with Glu97 (3.73 Å). Glu95 and Glu96 are exposed to solvent and have weak electron densities; Lys91 and Lys93 are not defined beyond the C^{γ} atom. The hydroxyl group of Ser94 forms a hydrogen bond to the carbonyl group of Lys91 (OG 94···O 91, 2.85 Å) thus shielding it from solvent. The side chains Lys91, Lys93, Glu95 and Glu96 are discretely ordered over two sites, but the Lys91–Glu95 pair could still be involved in a weak salt bridge.

Twelve of the 31 carbonyl groups of the long helix are hydrogen bonded to water molecules. The central third of the long helix (Arg84 to Gly92) is fully exposed to solvent and seven out of nine carbonyl groups in this region are hydrated (Fig. 6). The carbonyl groups of the two protected ends of the long helix are not so hydrated and the corresponding C=O···N hydrogen-bond lengths are slightly shorter (3.00 *versus* 3.08 Å) and the hydrogen-bond angles, C=O···N, are somewhat more linear (157.2 *versus* 154.6°). These trends are in agreement with previous studies (Blundell *et al.*, 1983; Sundaralingam & Sekharudu, 1989). The N-terminal region (Phe75 to Val83) of the long helix is well protected by the helices of the N-terminal domain (NT, *A*, *B* and *C*) and only the carbonyl groups of Glu76, Val80 and Met81 are hydrogen bonded to water molecules. At the C terminus of the long helix, the carbonyl groups of Glu95 and Glu96 are accessible to water molecules but only the carbonyl group of the former is hydrated. It has been suggested that these two glutamic residues together with Glu97 form part of the recognition site for TnI (Leavis & Gergely, 1984).

The long helix is amphiphilic at either end and hydrophilic in the middle. The two amphiphilic ends of the long helix have their hydrophobic surfaces out of register (Fig. 7), *i.e.* not on the same side of the long helix. The exposed portion of the central helix is encircled by water molecules, hydrogen bonded to the carbonyl groups. Thus, the long helix is not significantly bent and the radius of curvature is large, 150.1 (2) Å, unlike an amphiphilic helix, where only the polar side is hydrogen bonded to water molecules, bending the helix towards the hydrophobic side (Blundell *et al.*, 1983).

Neighboring bulky side chains can shield the helix carbonyl groups from hydration

Not all helix carbonyl groups are hydrated and bulky side chains at residues i, $i+3$, $i+4$, can adopt conformations that shield the ith carbonyl group from hydrogen bonding to a water molecule. For example, the bulky side chains Lys91 (i) and Glu95 ($i+4$) shield the carbonyl of Lys91 and similarly, the carbonyl of Gln85 is shielded by the hydrophilic side

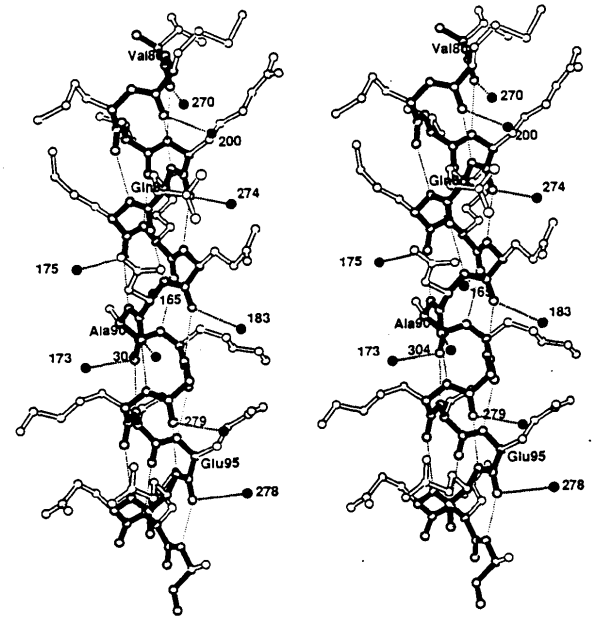

Fig. 6. A stereo diagram of a portion of the long central helix, spanning the residues 80–99, and the ten water molecules directly hydrogen bonded (thick dashed lines) to the helix backbone carbonyl groups without disrupting the Pauling N—H···O=C helix hydrogen bonds (thin dashed lines). The majority of the water molecules surround the exposed three turns of the *D*/*E* linker region of the central helix.

chains of Gln85, Glu88, and Asp89 (i, $i+3$ and $i+4$). The carbonyl of Lys91 (i) is shielded by the formation of an intramolecular hydrogen bond with the hydroxyl of Ser94 ($i+3$).

The absence of calcium ions in the N-terminal sites I and II results in 'collapsed' calcium-binding sites, which are stabilized by several water molecules hydrogen bonding to the metal-ligand sites. In particular, the three carbonyl O atoms at the beginning and end of the two apo metal-binding loops are hydrogen bonded to water molecules.

Water inserted B helix and the sulfate group

As highlighted earlier (Satyshur *et al.*, 1988), a unique feature of the *B* helix of the apo N-domain is the insertion of two water molecules near the amino-terminal end, between Thr39/Gly43 and Lys40/Thr44, disrupting the α-helix hydrogen bonds by increasing the separation between the carbonyl and amide groups to 4.9 and 5.4 Å respectively (Fig. 1 and Table 8). The hydroxyl group of Thr44 is also hydrogen bonded (2.87 Å) to the second inserted water molecule. This results in a kink in the *B* helix of 39° and the changes its course. A sulfate group of crystallization is now found to be locked in place, forming hydrogen bonds to the two inserted water molecules on one side, and to the amino groups of the guanidinium moiety of Arg47 on the other.

Table 8. *Hydrogen-bonding scheme involving the inserted waters and the sulfate group in the B helix*

Residue/atom	Residue/atom	Distance (Å)	Angle (°)	
Thr39 C=O	*W*240	3.04	C=O···Gly43 N	140
*W*240	Gly43 N	2.87	C=O···*W*240	152
Thr39 C=O	Gly43 N	4.78	O···*W*240···Gly43 N	108
Lys40 C=O	*W*234	2.67	C=O···Thr44 N	112
*W*234	Thr44 N	2.95	C=O···*W*234	128
Lys40 C=O	Thr44 N	5.40	O···*W*234···Thr44 N	149
*W*234	Thr44 OG1	2.69		
*W*234	*W*240	3.44*		
Sul225 O1	*W*254	2.57	S—O1···*W*254	106
Sul225 O1	Arg47 NH2	2.82	S—O1···NH2	104
Sul225 O2	*W*254	2.86	S—O2···*W*254	92
Sul225 O2	*W*230	3.06	S—O2···*W*230	137
Sul225 O3	*W*234	2.90	S—O3···*W*234	118
Sul225 O3	Arg47 NH1	2.94	S—O3···NH1	121
Sul225 O3	Arg47 HNH2	2.92	S—O3···NH2	100
Sul225 O4	*W*240	3.09	S—O4···*W*240	116
Sul225 O4	*W*306	2.79	S—O4···*W*306	132
Sul225 O4	*W*325	3.12	S—O4···*W*325	118

* Distance is greater than 3.2 Å, but is listed to show the proximity of the two inserted water molecules.

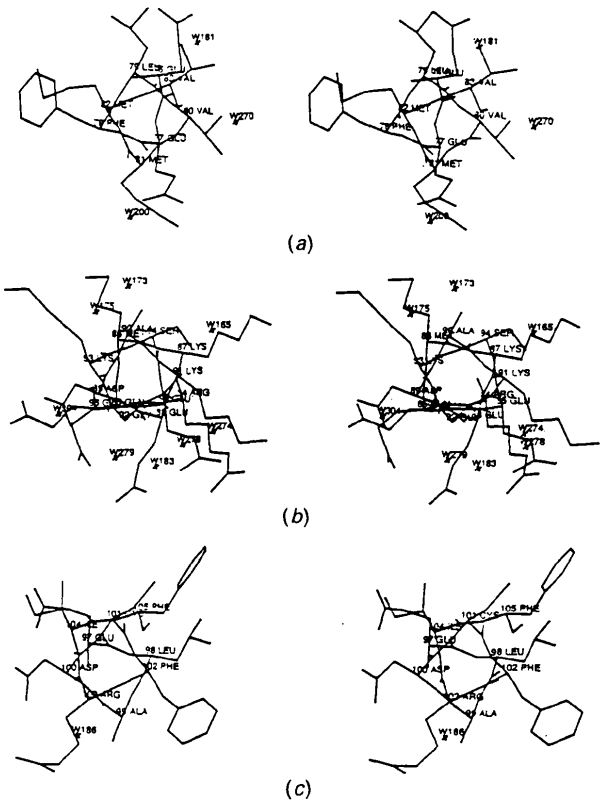

Fig. 7. Stereoviews of the Edmondson wheels of the long central helix. (*a*) The N-terminal residues 76–83 with three waters, (*b*) middle of the helix, residues 84–96, with eight waters and (*c*) C-terminal residues 97–105 with one water. Notice that the hydrophobic residues line opposite faces of the long helix at the two ends (*a*) and (*c*).

Hydration differences in the mini β-sheets of the apo and holo calcium-binding domains

The two *EF*-hand pairs in the amino and carboxyl domain are connected by a short piece of anti-parallel β-sheet through their loops, related by the pseudo-twofold axis. In the apo domain, the β-sheet is engaged in four direct hydrogen bonds (Fig. 8) including the two central hydrogen bonds between Ile37 O–Ile73 N (2.85 Å) and Ile37 N–Ile73 O (3.09 Å), and the two flanking hydrogen bonds Gly35 O–Phe75 N (3.20 Å) and Thr39 N–Gly71 O (3.04 Å). In the holo domain, the β-sheet is less twisted, and only the two central hydrogen bonds between Ile113 O–Ile149 N (2.98 Å) and Ile113 N–Ile149 O (2.87 Å) are present. The other two potential β-sheet hydrogen bonds are bridged by water molecules: Gly111 O–*W*191–Phe151 N (2.81, 2.90 Å) and Ile115 N–*W*170–Gly147 O (2.91, 2.73 Å). This crucial structural change upon calcium binding may be functionally important in determining the nature of the interactions between TnC and the troponin subunit TnI, and the interaction of the complex in turn, with the third components TnT (troponin T) and the actomycin filament (Leavis & Gergely, 1984).

Crystal packing as a model for TnC binding to target peptides

The crystal packing is dominated by head-to-tail interactions between the C-terminal domain of one

molecule and the *A* helix of a neighboring molecule, related by an internal 3_2 axis (Fig. 9). The *A* helix forms several hydrophobic interactions between its side chains and the 'palm' of the C-terminal domain. The interhelix angle between the *A* helix and the long helix is 99.3°. The angles between the *A* helix and other helices of the C-domain are: *A–F* = 104.3, *A–G* = 101.0 and *A–H* = 121.4°. The closest contact between the SD atom of Met28 and the S^{γ} of Cys101 is a hydrogen bond (2.97 Å) (Fig. 10) and the sulfhydryl group in the alternate conformation is at a distance of 4.76 Å. Other interactions include Ala24 with the hydrophobic residues of the *E* and *F* helices, Lys23 to the backbone of the *F* helix, and Glu21 to the *G* and *H* helices. The 'head-to-tail' interaction between the *A* helix of one molecule and the C-terminal cleft of a 3_2-related molecule forms an interlocked chain of TnC molecules spiraling around the 3_2 axis, a left-handed helical 'polymer' with three TnC molecules per turn and a rise of 20.3 Å. Two TnC molecules may also associate in a similar fashion to form a head-to-tail dimer, observed at low pH in solution (Wang, Lebowitz & Cheung, 1989).

Two models for the binding of target peptides to calmodulin have been developed based on this interaction in the TnC crystals (Strynadka & James, 1990; Sekharudu & Sundaralingam, 1993*a*), where the two palms of the N- and C-terminal domain bind to the same target peptide. These two models differ by 100° in the azimuthal angle of the target peptide helix. The latter model is similar to the binding of MLCK peptide to calmodulin (Meador, Means & Quiocho, 1992). A similar mode of interaction, perhaps involving only one palm, may be envisaged between TnC and a peptide segment of TnI in the troponin complex.

Summary

The high-resolution structure of TnC reveals more clearly the details of the geometry and conformation of the main chain and side chains compared to our 2 Å study. The high-resolution refinement has additionally revealed the discrete twofold ordering of a few side chains and more than twice the number of water molecules that solvate the protein. Also, the stereochemistry of the model has been improved and the mean positional error in the atomic coordinates reduced. A unique feature of the structure is the difference in the mini β-sheet structure in the

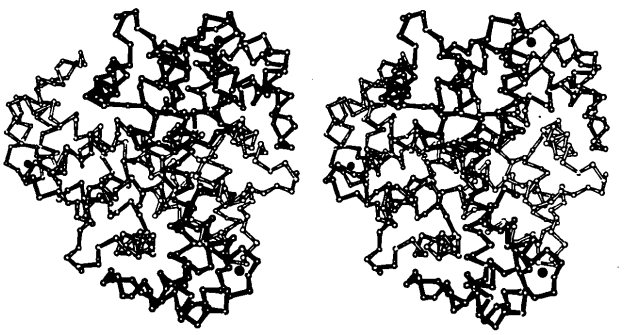

Fig. 9. The dominant theme in the crystal packing is the 'head-to-tail' locking of the TnC molecules around the 3_2 axis that run through the entire crystal. C^{α} traces for three such molecules are shown in stereo, viewed down the 3_2 screw axis. The filled balls represent the Ca^{2+} ions. The *A* helix of the molecule in the middle gets into the palm-shaped cavity of the Ca-bound C domain of the molecule above, while its C domain engages in the same interaction with the *A* helix of the molecule below.

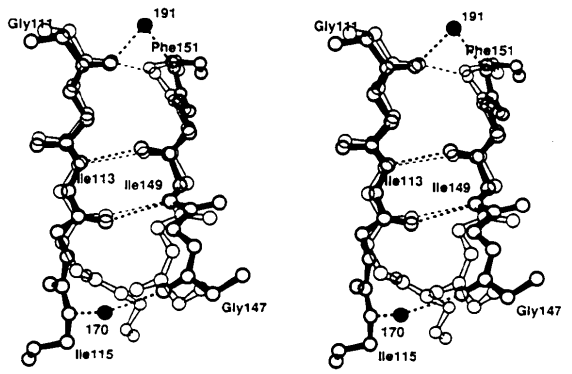

Fig. 8. The two mini anti-parallel β-sheets, dark bonds for the holo C-terminal domain and open bonds for the apo N-terminal domain. The side-chain atoms have been omitted and the residues in the C-terminal domain are numbered. The two sheets were superposed using the backbone atoms of the two central residues (r.m.s. = 0.25 Å). Notice how the water bridges at either end in the holo C-terminal domain change the twist of the sheet.

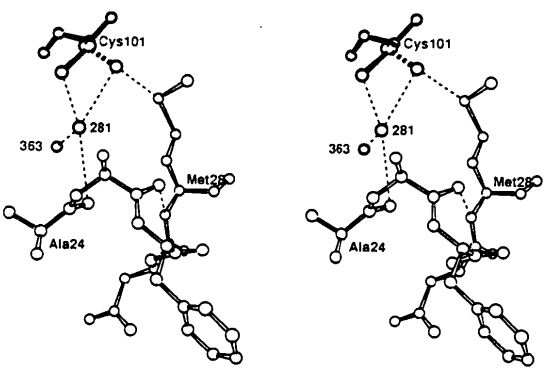

Fig. 10. Stereo diagram of the intermolecular interaction between Cys101, with its side chain discretely ordered over two sites (one shown as solid bond and the other as broken), and residues 24–28 of the *A* helix of the molecule related by a 3_2 axis (open bonds). The water molecule 281 interacts with both S^{γ} sites of Cys101 (2.50, 3.05 Å), the carbonyl O atom of Ala24 (3.35 Å) and water 363 (2.64 Å). One site of S^{γ} (broken bond) forms a hydrogen bond with SD of Met28 (2.97 Å).

calcium-bound C-terminal domain and the calcium-free N-terminal domain, where the water molecules play an important role in changing the twists in the two β-sheets. Upon calcium uptake, one can visualize not only the loops and the adjacent helices in the '*EF*-hands' of the N-terminal domain undergoing conformational rearrangement, ejecting the two inserted water molecules in the *B* helix, but also the β-sheet between the loops reducing its twist and capturing two water molecules to bridge the terminal hydrogen-bonding sites. The interruption of the β-sheet by water molecules at either end provides an insight into probable pathways for the formation and disruption of β-sheets by water molecules. (Sekharudu & Sundaralingam, 1993*b*).

We gratefully thank the National Institute of Health for Arthritis and Musculoskeletal Disease for a grant AR 34139, Professors Robert H. Kretsinger, S. Sobotnik and Dr Ron Chandros for their assistance and use of the area-detector facility at University of Virginia, Charlottesville, San Diego Supercomputer Centre and The Ohio State Supercomputer Center for allocation of time on their Cray computers. Endowment support (to MS) by the Regents of The Ohio State is gratefully acknowledged.

References

BAKER E. N. & HUBBARD, R. E. (1984). *Progr. Biophys. Mol. Biol.* **44**, 97–179.

BERNSTEIN, F. C., KOETZLE, T. F., WILLIAMS, G. J. B., MEYER, E. F. JR, BRICE, M. D., ROGERS, J. R., KENNARD, O., SCHIMANOUCHI, T. & TASUMI, M. J. (1977). *J. Mol. Biol.* **112**, 535–542.

BLUNDELL, T. L., BARLOW, D. J., BORKAKOTI, N. & THORNTON, J. M. (1983). *Nature (London)*, **306**, 281–283.

BRÜNGER, A. (1992). *X-PLOR* 3.0 *Manual*. Yale Univ., New Haven, CT, USA.

GREASER, M. L. & GERGELY, J. (1971). *J. Biol. Chem.* **246**, 4226–4233.

HENDRICKSON, W. A. (1985). *Methods Enzymol.* **115**, 252–270.

HERZBERG, O. & JAMES, M. N. G. (1985). *Nature (London)*, **313**, 653–659.

HERZBERG, O. & JAMES, M. N. G. (1988). *J. Mol. Biol.* **203**, 761–779.

JONES, A. (1985). *Methods Enzymol.* **115**, 157–171.

KRAULIS, P. J. (1991). *J. Appl. Cryst.* **24**, 946–950.

LEAVIS, P. C. & GERGELY, J. (1984). *CRC Crit. Rev. Biochem.* **16**, 235–305.

LUZZATI, V. (1952). *Acta Cryst.* **5**, 802–810.

MEADOR, W. E., MEANS, A. R. & QUIOCHO, F. A. (1992). *Science*, **257**, 1251–1255.

RAINACH, F. C. & KARLSSON, R. (1988). *J. Biol. Chem.* **263**, 2371–2376.

SATYSHUR, K. A., RAO, S. T., PYZALSKA, D., DRENDEL, W., GREASER, M. & SUNDARALINGAM, M. (1988). *J. Biol. Chem.* **263**, 1628–1647.

SEKHARUDU, C. Y. & SUNDARALINGAM, M. (1993*a*). *Protein Sci.* **2**, 620–625.

SEKHARUDU, C. Y. & SUNDARALINGAM, M. (1993*b*). *Hydration of Protein Secondary Structures – Role in Protein Folding, Water Science Reviews* 5, edited by F. FRANKS, pp. 148–162. Cambridge Univ. Press.

SIM, G. A. (1958). *Acta Cryst.* **12**, 813–815.

STRYNADKA, N. J. C. & JAMES, M. N. J. (1990). *Proteins Struct. Funct. Genet.* **7**, 234–248.

SUNDARALINGAM, M. & SEKHARUDU, C. Y. (1989). *Science*, **244**, 1333–1337.

WANG, C. K., LEBOWITZ, J. & CHEUNG, H. C. (1989). *Proteins Struct. Funct. Genet.* **6**, 424–430.