Determination of the Solution Structure of a Synthetic Two-Site Calcium-Binding Homodimeric Protein Domain by NMR Spectroscopy‡

Gary S. Shaw, Robert S. Hodges, and Brian D. Sykes*

Department of Biochemistry and MRC Group in Protein Structure and Function, University of Alberta, Edmonton, Alberta T6G 2H7, Canada

Received May 14, 1992; Revised Manuscript Received July 14, 1992

ABSTRACT: The solution structure of a 34-residue synthetic calcium-binding peptide from site III of chicken troponin-C has been determined by ^{1}H NMR spectroscopy. In solution and in the presence of calcium this peptide forms a symmetric two-site homodimeric calcium-binding domain (Shaw et al., 1990). The solution structure of this dimer was determined from the measurement of 470 NOEs from a 75-ms NOESY data set. For the dimer structure determination, the constraint list included 868 distance restraints, 44 ϕ angles, and 24 χ_1 and 2 χ_2 angles. Seven structures were calculated by restrained molecular dynamics using a procedure in which intramonomer distances were used first and then all distances, intra- and intermonomer, were input during further dynamics. The structures exhibited a fold very similar to the C-terminal domain of troponin-C comprised of a pair of helix–loop–helix calcium-binding sites. The rms deviation of these structures for backbone atoms between residues 97–122 and 97′–122′ for the dimer was 0.82 Å. The dimer structure was also calculated to be more symmetric than sites III and IV in troponin-C.

Calcium-binding proteins make up a large class of regulatory proteins which usually contain between one and four metal-ion binding sites. These proteins control important processes such as muscle contraction, enzyme activation, neurotransmitter release, and cell growth through a chain of events comprising binding of calcium to the protein, induction of a significant conformational change in the protein, and the triggering of the biological response. In many cases, one of the key features of calcium-binding proteins is the helix–loop–helix structural motif (Kretsinger & Nockolds, 1973), a contiguous stretch of about 30 amino acids which forms the calcium-binding site. Further, these calcium-binding sites are frequently paired so that two helix–loop–helix calcium-binding sites form a two-site domain. X-ray crystallographic studies have revealed that the calcium-binding proteins troponin-C (Herzberg & James, 1988; Satyshur et al., 1988) and calmodulin (Babu et al., 1988) have two such domains while calbindin D_{9k} (Szebenyi et al., 1986) and parvalbumin (Kretsinger & Nockolds, 1973; Swain et al., 1989; Declercq et al., 1988) have a single two-site domain. In each case, these calcium-binding domains have the added feature that a pseudo-2-fold symmetry exists between the pair of helix–loop–helix calcium-binding sites.

For several years now, synthetic peptides and proteolytic fragments have been used to study the calcium-binding properties of an isolated helix–loop–helix or the smaller 12-residue calcium-binding loop. In this manner, Hodges and co-workers have investigated the effects of amino acid sequence (Reid et al., 1981), metal-ion size and ionic charge (Gariepy et al., 1983, 1985), and ligating and nonligating residue composition (Marsden et al., 1988; Shaw et al., 1991a) of the loop toward calcium affinity for peptides comprising calcium-binding site III from the muscle contractile protein troponin-C (TnC). In all of these cases, the peptides were shown, by either NMR techniques or circular dichroism spectroscopy, to undergo a significant conformational change upon metal-ion binding. Further, the solution structure of a metal-ion complexed 13-residue peptide, corresponding to the calcium-binding loop of site III from TnC, has been shown to be monomeric and have a conformation similar to that found in the X-ray structure of this protein (Marsden et al., 1989). However, a different situation arises for the longer helix–loop–helix peptides, where it has recently been observed, using two-dimensional NMR spectroscopy, that a 34-residue peptide from site III of troponin-C (SCIII) not only undergoes a significant conformational change upon calcium-binding but also self-associates to form a homodimeric two-site domain (Shaw et al., 1990). This proposal has been confirmed by calcium titration experiments (Shaw et al., 1991b) and guanidine denaturation studies (Monera et al., 1992; Shaw et al., unpublished results) which clearly show that the calcium-saturated form of the peptide is a dimer. Further, two-dimensional NMR spectroscopy and sedimentation equilibrium experiments have been used to show that a 39-residue proteolytic fragment of rabbit skeletal troponin-C which includes calcium-binding site IV forms a symmetric dimer in the calcium-saturated form (Kay et al., 1991). We have also shown that it is possible to preferentially form a site III/site IV heterodimer from two 34-residue peptides representing sites III and IV from TnC using one- and two-dimensional NMR techniques (Shaw et al., 1992).

The examination of a symmetric, dimeric protein such as SCIII by NMR spectroscopy and the determination of its three-dimensional solution structure can present several difficulties. Among the most problematic is the unambiguous determination of intra- vs intermolecular contacts at the dimer interface. Both of these will have a critical impact on the final structures. Further complications can arise for proteins which are mostly α-helical in structure as shown in recent investigations of the *Escherichia coli* trp repressor (Arrowsmith et al., 1990), ColE1 rop protein (Eberle et al., 1990, 1991), arc repressor (Berg et al., 1989; Zagorski et al., 1989), and the transcription factors LFB1 (Pastore et al., 1991) and GCN4 (Oas et al., 1990) were it is difficult to determine how the dimers are assembled. In the present work, we have made use of distance restraints obtained from two-dimensional

‡ PDB coordinates for the seven final structures of SCIII and the average structure will be submitted to the Brookhaven Protein Data Bank.

* Author to whom correspondence should be addressed.

0006-2960/92/0431-9572$03.00/0

NOESY spectra and a model derived from the X-ray coordinates of turkey TnC to resolve ambiguous resonances at the interface of the SCIII homodimer and have applied restrained molecular dynamics to determine its three-dimensional solution structure.

EXPERIMENTAL PROCEDURES

Sample Preparation. The synthetic 34-residue peptide comprising site III (residues 93–126) of chicken skeletal troponin-C (SCIII) was synthesized using an Applied Biosystems Peptide Synthesizer (Model 430A). Purification was accomplished by reversed-phase HPLC using a C8 semi-preparative column, and the peptide composition was verified by amino acid analysis and mass spectrometry. Details of the synthesis and purification have been previously published (Shaw et al., 1991a,b). For NMR experiments, purified SCIII peptide was dissolved in 600 μL of D_2O, 50 mM KCl, and 4.5 mM $CaCl_2$ at pH 7.35 to give a final peptide concentration of about 1.5 mM. Following the acquisition of NMR data in D_2O, the peptide solution was lyophilized, and the resulting material redissolved in 80% H_2O/20% D_2O.

NMR Spectroscopy. All NMR spectra were obtained on a Varian VXR-500 spectrometer in the phase-sensitive mode. Sequence-specific assignments were carried out at 15 °C; however, spectra were also recorded at 30 °C to resolve overlapping amide resonances. In all experiments, water suppression was accomplished by presaturation. DQF-COSY (Piantini et al., 1982), RELAYED-COSY, and NOESY experiments were acquired in both 80% H_2O/20% D_2O and D_2O at each temperature. Typical data sets were comprised of 64 or 96 transients for each of 256 or 512 t_1 increments. Spectral widths of 7000 Hz in 80% H_2O/20% D_2O and 5200 Hz in D_2O were used. DQF-COSY spectra were processed using a sine-bell weighting function yielding final digital resolutions of 3.4 Hz/point (H_2O) or 2.5 Hz/point (D_2O) in F_1 and F_2. RELAYED-COSY spectra were collected using a 30-ms mixing time to optimize magnetization transfer from NH to βCH_3 for Ala and α to γCH_3 for Thr and Ile residues (Bax & Drobny, 1985). NOESY spectra were collected by the hypercomplex method (States et al., 1982) for mixing times (τ_m) of 75, 150, and 300 ms in H_2O and 150 ms in D_2O solution. In each case a random delay of ±15 ms (±20 ms for τ_m = 300 ms) was employed to suppress zero-quantum coherence. Spectra were processed using either a $\pi/8$-shifted sine bell or a cosine-squared function (Kline et al., 1988) and baseline corrected in F_2. This latter method proved particularly useful for observing weak cross peaks and was used for quantitative analysis of NOESY spectra.

Data Analysis. Spin–spin coupling constant information was obtained from DQF-COSY spectra in H_2O ($^3J_{NH\alpha}$) and D_2O ($^3J_{\alpha\beta}$). In each case, data were processed using a sine-bell weighting function and zero-filled in F_2 to final digital resolutions of 0.85 Hz/point (H_2O) and 0.63 Hz/point (D_2O). Traces of each resolved cross peak were taken along F_2 and the two antiphase components fitted using an in-house iterative curve fitting program written by Robert Boyko at the University of Alberta. Typically, this approach fitted about 10–20 points and yielded coupling constants which were 1–2 Hz smaller than those obtained from direct measurement of cross peak maxima and minima, known to result in a larger coupling constant than actual (Neuhaus et al., 1985).

Stereospecific assignments for several prochiral methylene centers were obtained using the method of Wagner et al. (1987), making use of $^3J_{\alpha\beta}$ coupling constants derived from curve fitting the appropriate cross peaks from DQF-COSY spectra and intraresidue NOEs taken from NOESY spectra with short mixing times (τ_m = 75 ms).

Structure Calculations. In order to calculate ^1H–^1H distances from NOESY spectra, NOE cross peaks were assigned from 300-ms NOESY spectra and the NOE intensities then measured by taking the volume integral of each cross peak from 75-ms NOESY spectra. In each case, the cross peak integral was corrected using an average volume integral for the baseline in that particular region. Volume integrals were corrected for the number of protons involved in the NOE (Williamson & Neuhaus, 1987; Yip, 1990) and translated into distances using the average 2.55-Å distance for β-δ proton pairs in phenylalanine residues (Billeter et al., 1982). Distances were then categorized as strong (1.8–2.7 Å), medium (1.8–3.3 Å), or weak (1.8–5.0 Å) for restrained molecular dynamics calculations (see below). The appropriate distance corrections were made to methyl protons, aromatic ring protons, and nonstereospecifically assigned methylene protons to allow centroid averaging (Wüthrich et al., 1983).

Angular restraints for ϕ backbone torsion angles were derived from $^3J_{NH\alpha}$ coupling constants. Specifically, 18 residues in SCIII which had $^3J_{HN\alpha} \leq 6.5$ Hz were restrained in the range $-85 \leq \phi \leq -40$. Four residues (D110, Y112, I113, and D114) with $^3J_{HN\alpha} > 8.0$ Hz were restrained in the range $-160 \leq \phi \leq -80$. Side-chain torsion angle restraints (χ_1, χ_2) were determined from an analysis of $^3J_{\alpha\beta}$ coupling constants in DQF-COSY spectra and/or intraresidue NOE data. A total of 12 χ_1 and 1 χ_2 restraints (24 χ_1 and 2 χ_2 restraints in the dimer) were included, comprising six in the range $-120 \leq \chi_1 \leq -20$, three in the $120 \leq \chi_1 \leq 240$ range, three in the $20 \leq \chi_1 \leq 120$ range, and one in the range $120 \leq \chi_2 \leq 240$.

Three-dimensional structure calculations were carried out using the energy minimization and restrained molecular dynamics algorithms of the GROMOS program (van Gunsteren et al., 1985; de Vlieg et al., 1986) employing a pseudoharmonic distance restraint potential and square-well dihedral angle restraining potential (Baleja et al., 1990). The starting coordinates for the SCIII homodimer were obtained from a model derived from the X-ray coordinates of calcium-binding sites III and IV of turkey skeletal troponin-C (Herzberg & James, 1988). In the model the coordinates for residues 93–126 were retained, except those for residues 100, 101, and 112 which are different in SCIII compared to turkey skeletal troponin-C. These residues were modified using the program MUTATE, written by Randy Read at the University of Alberta. Residues 129–162 were then changed to those in the SCIII sequence maintaining the coordinates for N, Cα, CO, O, and Cβ. The coordinates for residues 127 and 128 were deleted, and coordinates for an N-terminal acetyl and C-terminal amide were added to each 34-residue fragment (Shaw et al., 1990). This starting model was then subjected to 100 steps of steepest decents energy minimization followed by 200 steps of conjugate gradients energy minimization to remove severe nonbonding interactions and allow molecular dynamics using the SHAKE algorithm. Preliminary investigations into the dimer structure and the distinction of some intra- vs interchain NOEs were based largely on this starting model. The general protocol for molecular dynamics simulations involved heating the model coordinates rapidly to 600 K and then subjecting them to 50 ps of unrestrained molecular dynamics where coordinates were collected every 5 ps. Each of these 10 structures was then energy minimized and used as input structures for restrained molecular dynamics simulations incorporating distances and angles from the NMR

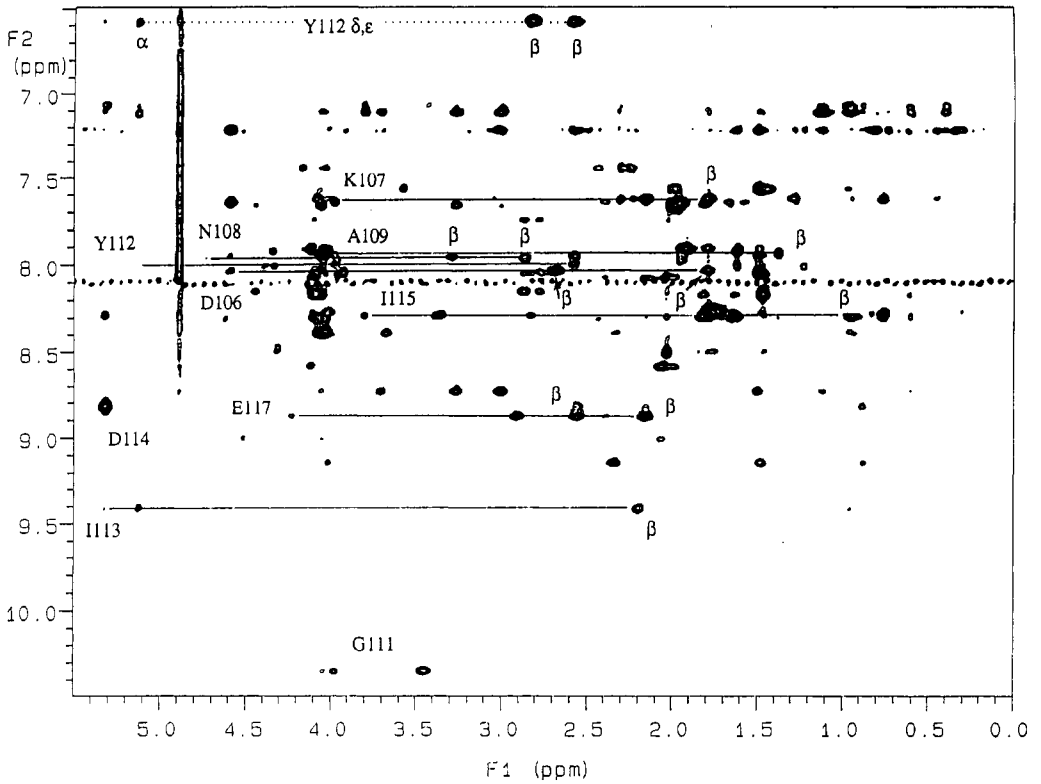

FIGURE 1: Contour plot for the NH–aliphatic region of the 75-ms ^{1}H NOESY spectrum of 1.5 mM SCIII acquired at 500 MHz in 80% H_2O/20% D_2O, 4.5 mM $CaCl_2$, 50 mM KCl, and 30 mM imidazole-d_4, pH 7.35, at 15 °C. The spectrum shows several of the intraresidue correlations for residues in the calcium-binding loop. The intense cross peak labeled D114 is comprised of D114NH, αCH and I113αCH, D114NH contributions. Near the top of the figure are cross peaks between the aromatic protons of Y112 and its αCH and βCH_2 protons. The "t_1 noise" displayed near F_2 = 8.05 ppm results from the 4,5-protons of the imidazole ring in the buffer which has partially exchanged during the experiment. All the other conditions are described under Experimental Procedures.

data. Each run was initiated at 300 K and consisted of increasing the distance restraining force constant for the *intramonomer NOEs only* from 100 to 20 000 kJ mol^{-1} nm^{-2} by doubling its value approximately every 1 ps during the first 16 ps. The dihedral angle restraint force constant was treated in a similar manner starting with an initial value of 25 kJ mol^{-1} rad^{-2} and a maximum value of 1000 kJ mol^{-1} and rad^{-2}. The temperature of the system was then increased to 600 K, and dynamics continued for a further 5 ps at the maximum force constants before allowing the system to cool back to 300 °C (over 4 ps). This sequence was then repeated incorporating *all intra- and intermonomer distance restraints*, and the system was subjected to 200 steps steepest decents and 200 steps conjugate gradients energy minimization.

RESULTS AND DISCUSSION

Sequence-Specific Assignment. The assignment of calcium-bound SCIII was carried out through the identification of the various spin systems via DQF-COSY and RELAYED-COSY spectra at 15 and 30 °C. Although SCIII is only 34 residues in length, these two temperatures were required to resolve resonances, especially in the amide and αCH regions where resonance overlap was a problem. In particular, 22 of the 34 residues in the peptide including the nine AMX spin systems (three Asp, two Asn, two Phe, one Tyr, one Ser), four isoleucine, three leucine, four alanine, and two of three glycine residues could be identified through this approach. In addition, the N-terminal residue, Lys93, was identified from the strong NOE between the N-terminal acetyl methyl group and the Lys93 amide proton. The sequential connectivity of SCIII was established from 150- and 300-ms NOESY spectra. In Figure 1, a contour plot for the NH/aliphatic region of the 75-ms NOESY spectrum is shown to illustrate several of the

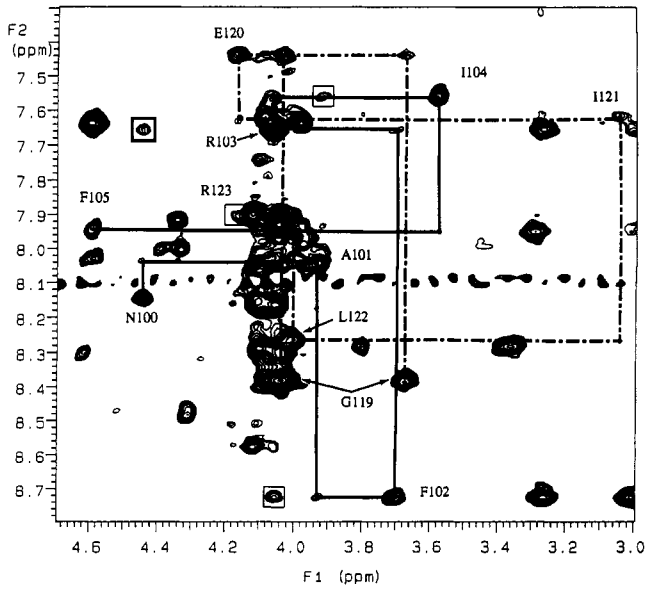

FIGURE 2: NH–αCH region of the contour plot from the 75-ms ^{1}H NOESY spectrum of 1.5 mM SCIII acquired at 500 MHz in 80% H_2O/20% D_2O, 4.5 mM $CaCl_2$, 50 mM KCl, and 30 mM imidazole-d_4, pH 7.35, at 15 °C. The spectrum shows connectivities for N100 to F105 in the E-helix and E120 to R123 in the F-helix of SCIII and the accompanying $d_{\alpha N}(i,i+3)$ cross peaks (boxed). The "t_1 noise" displayed near F_2 = 8.10 ppm results from the 4,5-protons of the imidazole ring in the buffer which has partially exchanged during the experiment.

residue types in the calcium-binding loops of SCIII. An expansion of the NH/αCH region is shown in Figure 2, where the sequential connectivities via $d_{\alpha N}$ and $d_{\alpha N}(i,i+3)$ cross peaks are noted for residues Asn100–Phe105 and Gly119–Arg123.

Table I: ^{1}H Chemical Shifts for SCIII (pH 7.35, 15 °C)[a]

residue	chemical shift (ppm)			
	NH	αCH	βCH	other
K93	8.49	4.31	1.85, 1.76	γCH 1.46; δCH 1.65; εCH 2.98
S94	8.47	4.51	4.17, 3.95	
E95	9.00	4.05	2.07	γCH 2.34
E96	8.58	4.12	2.07, 1.99	γCH 2.29
E97	8.07	4.11	2.16, 2.05	γCH 2.40, 2.27
L98	8.29	4.11	1.81, 1.64	γCH 1.61; δCH$_3$ 0.97, 0.93
A99	8.16	4.06	1.47	
N100	8.15	4.45	2.86,[2] 2.79[3]	
A101	8.05	3.94	1.50	
F102	8.73	3.71	3.26,[3] 3.00[2]	γCH 7.10; εCH 7.09; ζCH 6.82
R103	7.65	4.06	1.99	γCH 1.83, 1.66; δCH 3.27; εNH 7.30
I104	7.56	3.58	1.50	γCH 1.44, 0.98; γCH$_3$ 0.34; δCH$_3$ 0.63
F105	7.95	4.60	3.02,[3] 2.57[2]	δCH, εCH, ζCH 7.21
D106	8.03	4.60	2.68,[3] 1.79[2]	
K107	7.63	3.98	1.99	γCH 1.67, 1.58; δCH 1.82; εCH 3.10
N108	7.96	4.72	3.29,[2] 2.86[3]	
A109	7.93	4.05	1.38	
D110	8.30	4.62	3.07,[2] 2.43[3]	
G111	10.34	3.98, 3.45		
Y112	8.00	5.14	2.87,[3] 2.57[2]	δCH, εCH 6.57
I113	9.41	5.31	2.20	γCH 1.00, 0.96; γCH$_3$ 0.88; δCH$_3$ 0.40
D114	8.81	5.32	3.36,[3] 2.83[2]	
I115	8.28	3.81	0.93	γCH 1.14, 0.90; γCH$_3$ 0.77; δCH$_3$ 0.61
E116	7.62	4.09	2.23, 2.16	γCH 2.40, 2.29
E117	8.87	4.23	2.55,[2] 2.15[3]	γCH 2.91,[3] 2.41[2]
L118	9.14	4.02	2.34, 1.48	γCH 1.63; δCH$_3$ 1.11, 0.80
G119	8.38	4.04, 3.68		
E120	7.44	4.17	2.27	γCH 2.49, 2.44
I121	7.62	3.05	1.79	γCH 1.30, –0.62; γCH$_3$ 0.31; δCH$_3$ 0.45
L122	8.26	4.01	1.78	γCH 1.73; δCH$_3$ 0.84, 0.74
R123	7.90	4.12	1.92	γCH 1.80, 1.57; δ 3.21, 3.24
A124	7.92	4.35	1.62	
T125	8.00	4.39	4.33	γCH$_3$ 1.24
G126	8.07	4.01		

[a] Chemical shifts are referenced to the methyl resonances of DSS at 0.0 ppm. Stereospecific assignments are indicated as either β2 or γ2 (2) and β3 or γ3 (3) by the convention of Wagner et al. (1987).

The complete list of assignments for SCIII at 15 °C is given in Table I.

Stereospecific assignments for eight of the nine AMX spin systems were completed on the basis of an examination of the αβ cross peak coupling patterns, measurement of $^3J_{\alpha\beta}$ coupling constants, and analysis of NH, β, and αβ NOE intensities. In general, these resonances were well resolved and gave very distinct cross peak patterns (Figure 3). In particular, the αβ2 and αβ3 cross peaks from Asp106, Asn108, Asp110, and Asp114 have significant differences in chemical shift between β2 and β3 protons, ranging from 0.43 (Asn108) to 0.89 ppm (Asp106). In addition, the chemical shift difference between γ protons for Glu117 is about 0.50 ppm. These observations are significant since these residues correspond to the five ligating residues in the calcium-binding loop of SCIII which coordinate the metal ion through their side-chain carboxylate or amide carbonyl groups. In contrast, the sixth ligating residue, Tyr112, which coordinates the calcium ion via its backbone carbonyl, and two other AMX spin systems identified in other regions of the peptide have smaller chemical shift differences between β2 and β3 protons (Tyr112, 0.30 ppm; Phe102, 0.26 ppm; Asn100, 0.07 ppm). This trend has been observed for AMX spin systems in the calcium-binding loops of other calcium-binding proteins (Kay et al, 1991; Kordel et al., 1989; Drakenberg et al., 1989; Padilla et al., 1988; Ikura et al., 1991) and likely reflects the ordered nature of the side chains of the ligating residues when coordinated to the calcium ion. It is anticipated that this large distinction in chemical shift between β2 and β3 protons will be sufficient evidence that the calcium ion is coordinated to the side chains of these residues.

Secondary Structure. Following the complete assignment of calcium-saturated SCIII, elements of secondary structure were derived from NOE connectivities and $^3J_{HN\alpha}$ coupling constants. As shown in Figure 4, two significant stretches of $d_{\alpha N}(i,i+3)$ and $d_{\alpha\beta}(i,i+3)$ connectivities between residues 96–106 and 113–123 were found. Further, strong $d_{NN}(i,i+1)$ connectivities and $^3J_{NH\alpha}$ coupling constants ≤6.5 Hz in these regions suggest that both of these regions are helical in SCIII.

The sequence between the two helical regions in SCIII shows a significantly different NOE connectivity pattern. For example, no $d_{\alpha N}(i,i+3)$ or $d_{\alpha\beta}(i,i+3)$ NOE cross peaks were observed for residues 106–113, but rather several $d_{NN}(i,i+1)$ cross peaks were found. In addition, several large $^3J_{NH\alpha}$ coupling constants were determined for residues Asp110, Tyr112, Ile113, and Asp114, suggesting this region of the sequence is in an extended conformation. This is confirmed by the αCH chemical shift of Tyr112 (5.14 ppm), Ile113 (5.31 ppm), and Asp114 (5.32 ppm) and the amide resonance for Ile113 (9.41 ppm), which are all representative of residues in a β-sheet structure (Wishart et al., 1991). Further, the distinctive chemical shifts for the β2 and β3 protons, described earlier, are consistent with the side chains of Asp106, Asn108, Asp110, and Asp114 in this region of SCIII being coordinated to the calcium ion. These results suggest qualitatively that in the presence of calcium SCIII adopts a helix–loop–helix conformation, consistent with that observed for other calcium-binding proteins.

Homodimer Formation. Upon completion of the sequential assignment of SCIII and the establishment of secondary structural elements, long-range NOEs ($i-j \geq$ five residues) were identified in order to establish the tertiary fold of SCIII.

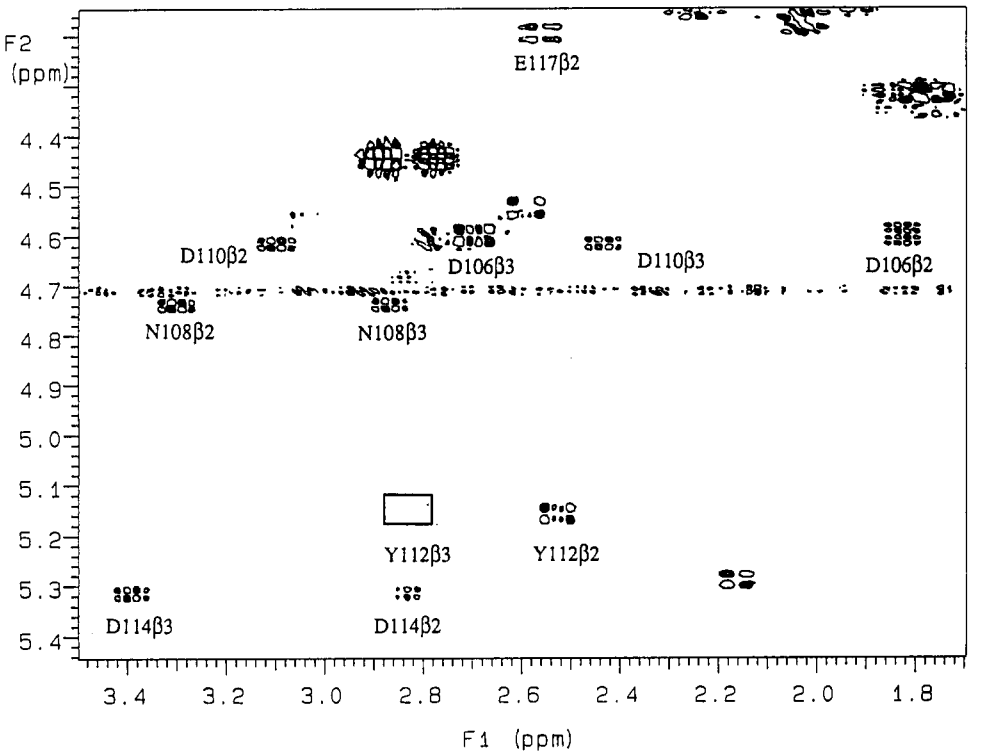

FIGURE 3: Region of the DQF-COSY spectrum of SCIII at 30 °C recorded in D_2O showing the $\alpha\beta$ cross peak coupling patterns for the six calcium coordinating residues: D106, N108, D110, Y112, D114, and E117. The β proton in each pair of cross peaks has been labeled either $\beta 2$ or $\beta 3$ on the basis of its stereospecific assignment according to the nomenclature of Wagner et al. (1987). Note the large differences in chemical shift between the $\beta 2$ and $\beta 3$ protons for residues Asp106, Asn108, Asp110, and Asp114, which coordinate the metal ion through their side chains, and the smaller difference for the $\beta 2$ and $\beta 3$ protons of Y112, which coordinates the calcium via its backbone carbonyl. The box shown for Y112$\beta 3$ indicates an extremely weak cross peak which is only visible at lower contour levels.

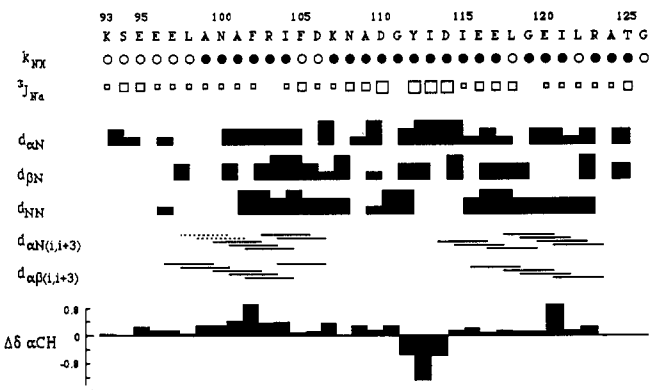

FIGURE 4: Summary of the sequential NOE connectivities, backbone amide exchange rates, and backbone coupling constants ($^3J_{HN\alpha}$) for SCIII. The amino acid sequence and residue numbers are shown at the top of the figure. The $d_{\alpha N}$, d_{NN}, and $d_{\beta N}$ are classified by the NOE volume (strong, medium, or weak) measured from the 75-ms NOESY spectrum at 15 °C. The $d_{\alpha N}(i,i+3)$ and $d_{\alpha\beta}(i,i+3)$ connectivities are shown only if the cross peak was observed at this mixing time. The two dashed lines indicate that the cross peak may be present but overlaps with other cross peaks. Backbone coupling constants are shown as boxes and reflect the relative size of the coupling where the small boxes are for $^3J_{HN\alpha} < 6.5$ Hz and the larger boxes are for $^3J_{HN\alpha} > 8$ Hz. Amide exchange is based on the change in chemical shift for that particular NH resonance between 15 and 30 °C. Those residues exhibiting a shift of <6 ppb/°C over this range were classified as slowly exchanging and are shown as filled circles. Differences in chemical shift between the observed and random coil values ($\delta_{rc} - \delta_{obs}$) for the α proton (Wishart et al., 1991) for each residue are shown on the second to last line. A positive value denotes that the observed chemical shift was upfield of the random coil shift.

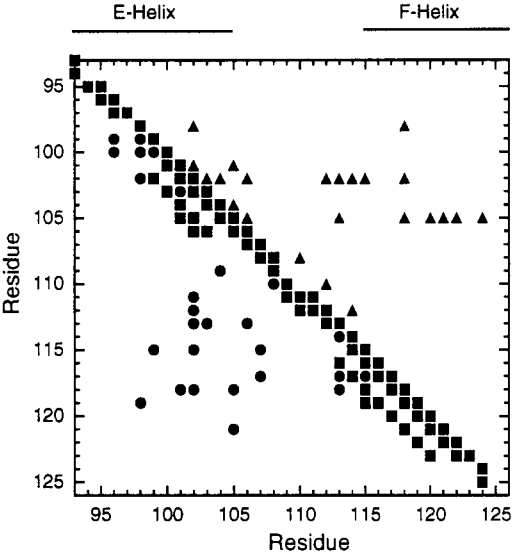

FIGURE 5: Diagonal plot of the NOE distance constraints used for SCIII based on the peptide sequence. Indicated are the observed NOEs (τ_m 75 ms) between residue pairs of backbone protons (■), backbone–side chain residue pairs (▲), and side chain–side chain residue pairs (●).

As shown in Figure 5, a significant number of these NOEs were found. An important observation from Figure 5 is that nearly all of these long-range NOEs are between the two helices on opposite sides of the calcium-binding loop. Specifically, several contacts are found between the side chains of Phe102, Phe105, and Leu98 in the first helix to the side chains of other hydrophobic residues such as Ile115, Leu118, and Leu122 in the second helix. Since the helix–loop–helix structural motif had already been identified in SCIII, it seemed reasonable to compare these long-range NOE assignments with contacts observed in the X-ray structures of TnC (Herzberg & James, 1988; Satyshur et al., 1988). This approach revealed that in the X-ray structures of TnC the only side chain interactions (≤4.3 Å) which occur in these helices of site III are between Ile104 and Ile121 and between

Phe105 and both Leu118 and Ile121 [Figure 18 of Satyshur et al. (1988) and Table IV of Monera et al. (1992)]. These observations are in marked contrast to contacts inferred by the long range NOEs in SCIII, shown in Figure 5.

The apparent discrepancies between the experimental NOE data and the contacts expected for a single helix–loop–helix structural motif were resolved from a series of observations. First, it was shown in the X-ray structures of TnC that residues Leu98 and Phe102 in the E-helix of site III had numerous contacts with hydrophobic residues Phe151, Phe154, Leu155, and Met158 in the H-helix of site IV [see Figure 18 of Satyshur et al. (1988)]. Further, sequence alignment of SCIII with that of site IV has revealed that these hydrophobic residues in site IV align perfectly with the hydrophobes Ile115, Leu118, and Leu122 and residue Gly119 in SCIII (Shaw et al., 1990). In addition, detailed calcium titration studies (Shaw et al., 1991) and sedimentation equilibrium and circular dichroism studies (Monera et al., 1992; G. S. Shaw, unpublished results) have shown that the calcium-bound form of SCIII is a dimer. This is in agreement with preliminary X-ray diffraction studies of a 34-residue site III peptide from rabbit TnC, where it has been suggested that the peptide may exist as a dimer in the crystalline form (Delbaere et al., 1989). On the basis of these observations, a model was proposed whereby two helix–loop–helix SCIII peptides associate in a head-to-tail fashion to form a symmetric dimer (Shaw et al., 1990). The fold of this dimer would be similar to that observed in the X-ray structures of troponin-C for sites III and IV of the C-terminal domain (Herzberg & James, 1988; Satyshur et al., 1988) wherein the backbone atoms exhibit a pseudosymmetric relationship. Using this approach, many of the NOEs observed between hydrophobes of the E-helix and residues of the F-helix in SCIII could now be resolved as interactions at the dimer interface (Shaw et al., 1990). However, some NOEs proved to be ambiguous and could arise from either intrachain or interchain contacts or both. These were resolved during molecular dynamics simulations (see below).

An essential requirement of the SCIII dimer model is a short three-residue β-sheet which must be present involving residues Tyr112, Ile 113, and Asp114 of each peptide. This would correspond to the β-sheet formed between residues 112–114 of site III and residues 148–150 in site IV as observed in the X-ray structures of TnC (Herzberg & James, 1988; Satyshur et al., 1988). To support this idea, NOEs were found for Tyr112 αCH to Asp114 αCH and for the Tyr112 aromatic protons to Asp114 βCH. As described earlier, the chemical shift and coupling constant data for these residues are also consistent with a β-sheet conformation.

Tertiary Structure Determination. While the SCIII dimer model could account for many of the observed NOEs and observations from calcium titration experiments, it also had several weaknesses. Since the model was generated from the X-ray coordinates of turkey TnC, it was not inherently symmetric. In fact, a comparison of the two SCIII monomers (residues Glu97–Leu122) at this stage revealed a backbone atomic rms deviation of 0.90 (Table II). Also, while many calculated distances from NOE integrals fell within the range expected for the model, there were several discrepancies, including 15 violations >1.0 Å and 48 violations >0.5 Å. These observations showed that while the dimer model was a good approximation, it was far from a final solution structure.

The solution structure of the SCIII homodimer was determined from a set of 470 from a total of 495 measured NOEs ($\alpha\alpha'$, $\beta\beta'$, and $\gamma\gamma'$ NOEs were not used). These resulted in a total of 798 intramonomer and 70 intermonomer upper bounds distance restraints for the dimer. A summary of the number of distance restraints per residue and the long-range NOEs is shown in Figure 6. Several molecular dynamics simulations including NOE distance constraints and angular constraints were calculated. However, to ensure that a sufficient conformational space was being sampled, the initial model from the X-ray coordinates was subjected to 50 ps of unrestrained molecular dynamics, during which coordinates

Table II: Structural Statistics for SCIII

(A) Residual Violations

range	av no. of constraint violations
$0.1 \le d < 0.2$	47.5
$0.2 \le d < 0.3$	9.2
max distance violation	0.41 Å
av distance violation	0.015 Å

(B) SCIII Structures[a]

	atomic rms deviations (Å)	
	backbone	all
EM vs EMa[b]	0.82 ± 0.32	
EMa (97–122) vs EMa (97′–122′)	0.30	0.99
EMa vs EMa (helices)[c]	0.43 ± 0.08	0.83 ± 0.41
EMa vs EMa (loops)[d]	0.50 ± 0.20	0.80 ± 0.37

(C) Comparison with Turkey TnC Crystal Structure

	atomic rms deviations (Å)	
	backbone	all
EM vs TnC	1.64	2.33
EM (97–122) vs TnC (97–122)	1.33	2.04
EM (97–122) vs TnC (133–158)	1.00	1.85
TnC (97–122) vs TnC (133–158)	0.90	1.61
EM (97–122) vs EM (97′–122′)	0.30	0.99
TnC (106–117) vs TnC (142–153)	0.90	1.61
EM (106–117) vs EM (106′–117′)	0.17	0.86

[a] In all superpositions residues Glu97–Leu122 and Glu97′–Leu122′ were used. [b] EM refers to the seven restrained energy minimized structures; EMa is the energy minimized average SCIII dimer structure. [c] Helices refer to residues Ala101–Phe105, Ile115–Leu122, Ala101′–Phe105′, Ile115′–Leu122′. [d] Loops refer to residues Asp106–Glu117 and Asp106′–Glu117′.

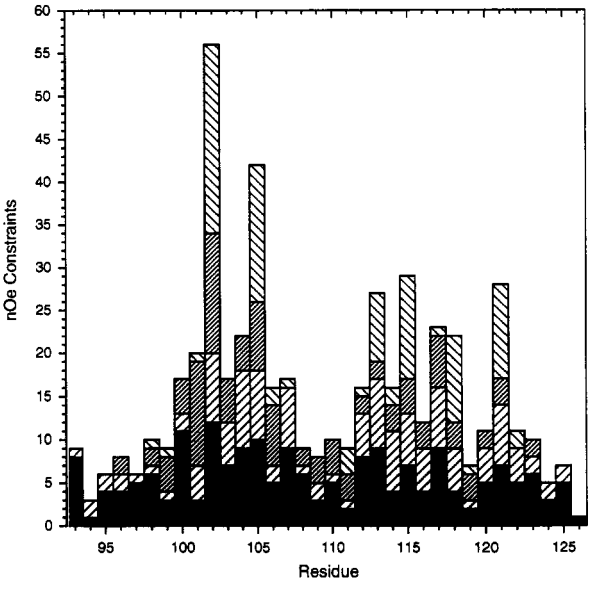

FIGURE 6: Histogram of the number of NOE distance constraints used per residue for SCIII. In this figure all NOEs are indicated twice to reflect the two residues involved in the NOE. The four types of bars indicate intraresidue (solid), sequential (broad cross-hatching, lower left to upper right), medium-range, $2 \le i \le 4$ (narrow cross-hatching, lower left to upper right) and long-range contacts, $i > 5$ (broad cross-hatching upper left to lower right).

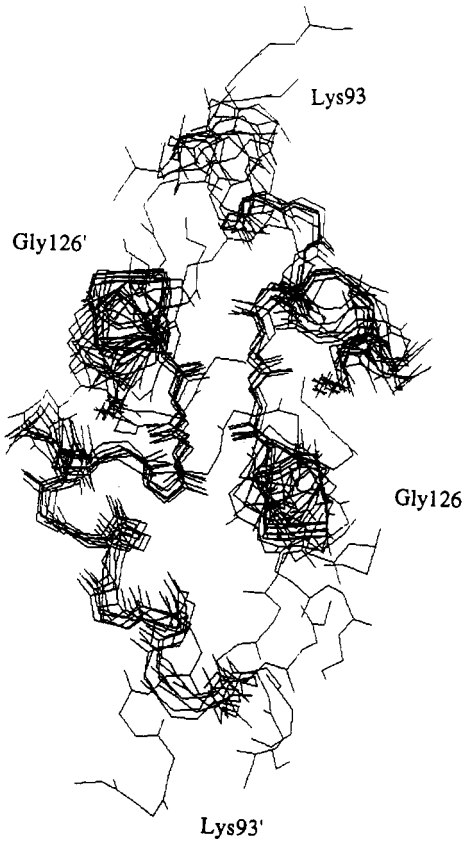

FIGURE 7: Superposition of the seven rMD structures for the SCIII homodimer, residues 93–126. The structures were superimposed for a minimum rmsd with the average structure for the backbone atoms (N, Cα, C, and O) of residues Glu97–Leu122 and Glu97′–Leu122′.

were collected every 5 ps. This approach yielded 10 starting structures which possessed little to no α-helix, no distinctive calcium-binding loops, and exhibited no visible symmetry.

A set of seven restrained molecular dynamics structures were calculated from the 10 starting structures. The average structure from these was then examined to account for ambiguous NOE assignments at the dimer interface which included F102 δCH to I113 β, I113 γCH_3, I113 NH, Y112 β and β′, Y112 αCH, F102 εCH to I113 NH, I113 βCH, I113 γCH_3, Y112 αCH and L118 δCH_3 and Y112 αCH to I113 NH, Y112 β′CH to I113 NH, and Y112 δCH to I113 αCH and D114 βCH cross peaks. Once these NOEs could be definitively assigned as arising from intra- or intermonomer, they were included in the calculations. The resulting seven rMD structures are illustrated in Figure 7, which clearly shows that the SCIII homodimer is comprised of two helix–loop–helix monomers. From this superposition of residues Glu97–Leu122 (and Glu97′–Leu122′), a well-defined backbone structure is present, having a rms deviation of 0.82 ± 0.32 Å about the mean for this region (Table II). Within this portion of the dimer, it is apparent that the two α-helices within each SCIII monomer are the most defined regions of the molecule. This is shown graphically in Figure 8, where the rms deviation for the backbone atoms is lowest between residues Ala101–Asp106 and Ile115–Leu118. From these figures it is also clear that the first and last four residues of each polypeptide chain are not ordered. From an analysis of the NOEs (Figures 5 and 6), this is likely a result of a lack of medium- and long-range NOEs, probably because of the mobility of these regions in solution. Similar observations have been made in calbindin D_{9k} (Akke et al., 1992) and in a proteolytic fragment comprising site IV from TnC, which forms a dimer in solution (Kay et al., 1991).

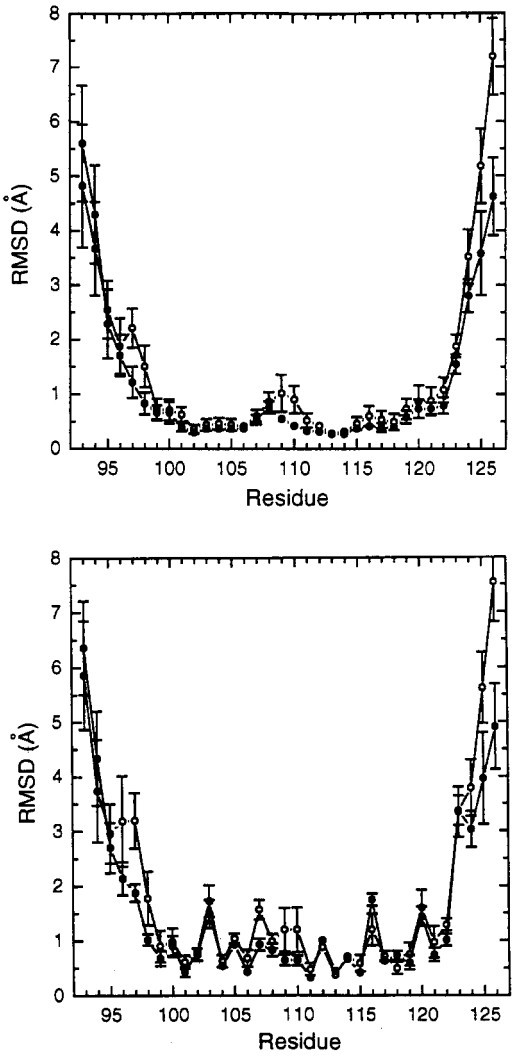

FIGURE 8: Atomic rms deviations per residue for the seven rMD structures of the SCIII homodimer. In panel A (top), the pairwise rms deviation for the backbone atoms (N, Cα, C, and O) about the average structure are shown, and in panel B (bottom) all heavy atoms are included. In each case the seven structures were superimposed with the average structure for the backbone atoms (N, Cα, C, and O) of residues Glu97–Leu122 and Glu97′–Leu122′, and the rms deviation was calculated for the particular residue. For both panels A and B two plots are shown denoting to two helix–loop–helix monomers within the homodimer.

It is interesting that the calcium-binding loops for the SCIII dimer are slightly less well-defined in this assessment. In particular, residues K107–Gly111 immediately preceding the β-sheet in each calcium-binding loop show the greatest deviation. This has also been observed in calbindin D_{9k} for residues Glu17–Asn21 in site I and Lys55–Glu60 in site II and in the site IV homodimer for residues Lys140–Gly144. In both of these cases, these residues also precede the β-sheet in the calcium-binding loop. All of these observations are consistent with the X-ray structures of TnC (Herzberg & James, 1988; Satyshur et al., 1988), which show that in general the helices immediately preceding and following calcium-binding loops III and IV have slightly better temperature factors than the residues at the beginning of each calcium-binding loop.

While the calcium-binding loops for the SCIII homodimer are slightly less well defined than the adjacent helices when residues Glu97–Leu122 are superimposed, a superposition of the individual calcium-binding loops results in an excellent fit (backbone rms of 0.50 ± 0.20). This region of the peptide also had the bulk of the angular constraints and stereospecific

assignments, and as a result the overall rms deviation is also very good (0.80 ± 0.37 for all atoms). In particular, the calcium-coordinating side chains for Asp106, Asn108, Asp110, Asp114, and Glu117 are very well defined. Meanwhile, the side chains of several of the other noncoordinating residues are more poorly defined. This observation is consistent with each of the coordinating residues having nondegenerate β-methylene protons which have been stereospecifically assigned as well as the incorporation of χ' torsional angles (also χ_2 for Glu117).

Important observations in the SCIII homodimer structure can also be made at the interface of the two peptides. Since no covalent bonds exist between the two SCIII peptides, it is clear that interactions at this interface promote and stabilize the association of the peptides. Consistent with this a large van der Waals energy between monomers of –345.7 ± 42.5 kJ/mol is calculated. In this regard only the hydrophobic residues Phe102, Ile113, Ile115 and Leu118 have a significant number of contacts at the dimer interface (Figure 6). This corresponds well with the rms deviation shown in Figure 8 which indicates that the side chains of these residues are among the most well defined. Furthermore, this figure also shows that some of the poorest defined side chains arise from residues which are located on the exposed sides of the helices, namely, Arg103, Lys107, Glu116, and Glu120.

In each structure there are no NOE violations greater than 0.5 Å, and in only one structure is there a violation greater than 0.3 Å. The average violations in the seven structures are similar and range between 0.014 and 0.018 Å. In terms of energy, the average total potential energy for the structures is –3686.9 ± 215.8 kJ/mol. In the calculations, no force potential was used to maintain symmetry as was done with the site IV dimer (Kay et al., 1991), but rather the NOEs were used to determine if the rMD procedure would force the dimer to be symmetric. This seems to have been quite successful as the average hydrophobic components of each monomer in the SCIII homodimer are nearly equal (–897.5 ± 25.3 vs –905.4 ± 37.0 kJ/mol). It is probable that given enough simulations of longer time these averages would become closer.

Comparison with the TnC Crystal Structures. The average SCIII structure was superimposed with the C-terminal domain from the crystal structure of turkey skeletal TnC by minimizing the rms deviation between backbone atoms for residues Glu97–Leu122 and Glu97′–Leu122′ in SCIII with residues Glu97–Leu122 and Asp133–Met158 in TnC (Figure 9). This comparison illustrates that each helix–loop–helix in SCIII is quite similar to those in sites III and IV in TnC although the overall rms deviation for this region is not particularly good (1.64 Å). As a result there are several distinct similarities for the two molecules and some obvious differences. For example, the interhelix crossing angle for helices E and F in SCIII appears to be quite close to that for TnC (109°; Strynadka & James, 1989). A similar observation was noted for the dimeric structure of fourth calcium-binding site, TH2 (Kay et al., 1991). In this regard it is also clear that the helices in SCIII are regular and very similar to TnC, at least between residues Glu97 and Leu122. This observation is consistent with the helices having the bulk of the NOE information and the best rms deviation for SCIII so that they are not affected greatly by the averaging of the seven structures. Beyond this region of SCIII, the comparison with TnC breaks down quickly, and the two structures are not very similar. This arises for two reasons: the lack of NOEs for SCIII in these regions results in poorly defined structures (Figures 7 and 8) and the lack of similarity in these regions between sites III and IV in TnC.

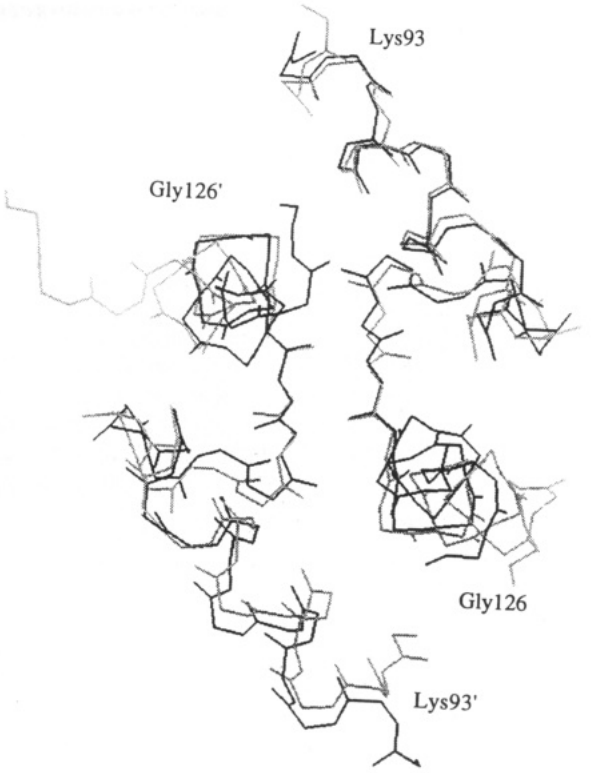

FIGURE 9: Superposition of the average rMD structure for the SCIII homodimer after energy minimization (dark line) with the C-terminal domain of TnC (residues 127 and 128 deleted). The structures were superimposed for a minimum rmsd for the backbone atoms (N, Cα, C, and O) of residues Glu97–Leu122 and Glu97′–Leu122′ in SCIII with those of residues Glu97–Leu122 and Asp133–Met158.

The similarity of the two calcium-binding sites in SCIII is apparent from a superposition of site III with site III′ in the SCIII dimer. This reveals a backbone rms deviation of 0.30 Å compared to 0.90 Å for the corresponding regions in sites III and IV in TnC. A superposition of the backbone atoms of site III in SCIII with those for site III in TnC shows an rms deviation of 1.33 Å. Together these comparisons show that small alterations in the overall structure of SCIII give rise to its better symmetry compared with sites III and IV in TnC and that neither site in SCIII has simply adopted an identical conformation of site III in TnC.

Perhaps the most interesting feature in comparing the SCIII dimer with TnC is the difference in calcium affinities between the two and the correlation of this with structure. In SCIII, it has been shown that calcium binds with two distinctive dissociation constants (K_d of 3 μM and about 1 mM) while in TnC calcium binding is significantly tighter (K_d of 20 nM). It has also been shown that the calcium binding to isolated site IV of TnC is quite weak, and values near 1 mM have recently been measured (Kay et al., 1991; G. S. Shaw, unpublished results). These observations serve to show that site III clearly has a stronger "natural" affinity for calcium and in TnC most certainly binds calcium before site IV. In TnC this seems to result in an enhanced apparent affinity for site IV or positive cooperativity. In SCIII, it appears that the opposite is true and that initial calcium binding perturbs calcium binding to the second site. There are significant differences between SCIII and TnC. First, the SCIII dimer is not a contiguous sequence as is TnC. In this vein it is also interesting that the dimerization constant for TH2 appears to be significantly higher than for SCIII. Both of these dimers

have hydrophobic interactions at their helix interfaces which occur twice (due to the nature of the dimers) and are mixed. In SCIII, the interactions are mainly between Phe102 and Ile115 and Leu118, and in TH2 the interactions are between Ile134, Met138, and Phe154. In TnC these hydrophobic interactions are not mixed and are comprised of two clusters: aromatic including Phe102, Phe151, and Phe154 and aliphatic I115, Leu118, I134, and Met138. It is probable that the nature of these interactions is very important in the stability of the dimers and the C-terminal of TnC and that this is reflected in the large discrepancies in calcium dissociation constants between sites III and IV of TnC and the SCIII and TH2 dimers. Indeed, we have found that it is possible to preferentially form a heterodimer comprised of a 34-residue site III peptide and a 34-residue site IV peptide, and while the dissociation constant for site III is identical to the first calcium binding to SCIII (3 μM), calcium binding to site IV is greatly increased (2 μM).

ACKNOWLEDGMENT

We thank Robert Boyko for his invaluable computing assistance and Jim Baleja and Frank Sönnichsen for help with molecular dynamics simulations. We also thank Sue Smith for typing the manuscript.

SUPPLEMENTARY MATERIAL AVAILABLE

Table of all NOEs measured and corresponding distances and dihedral restraints for SCIII (16 pages). Ordering information is given on any current masthead page.

REFERENCES

Akke, M., Drakenberg, T., & Chazin, W. J. (1992) *Biochemistry 31*, 1011–1020.

Arrowsmith, C. H., Pachter, R., Altman, R. B., Iyer, S. B., & Jardetzky, O. (1990) *Biochemistry 29*, 6332–6341.

Babu, Y. S., Bugg, C. E., & Cook, W. J. (1988) *J. Mol. Biol. 203*, 191.

Baleja, J. D., Pon, R. T., & Sykes, B. D. (1990) *Biochemistry 29*, 4828–4839.

Bax, A., & Drobny, G. (1985) *J. Magn. Reson. 61*, 306–320.

Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, A., & Haak, J. R. (1984) *J. Chem. Phys. 81*, 3684–3690.

Billeter, M., Braun, W., & Wüthrich, K. (1982) *J. Mol. Biol. 155*, 321–346.

Breg, J. N., Boelens, R., George, A. V. E., & Kaptein, R. (1989) *Biochemistry 28*, 9826–9833.

Declercq, J. P., Tinant, B., Parello, J., Etienne, G., & Huber, R. (1988) *J. Mol. Biol. 212*, 349.

Delbaere, L. T. J., Vandonselaar, M., & Reid, R. E. (1989) *J. Biol. Chem. 264*, 10261–10263.

de Vlieg, J., Boelens, R., Scheek, R. M., Kaptein, R., & van Gunsteren, W. F. (1986) *Isr. J. Chem. 27*, 181–188.

Drabikowski, W., Brzeska, H., & Venyaminov, S. Y. (1982) *J. Biol. Chem. 257*, 11584–11590.

Drakenberg, T., Hofmann, T., & Chazin, W. J. (1989) *Biochemistry 28*, 5946–5954.

Eberle, W., Klaus, W., Cesareni, G., Sander, C., & Rösch, P. (1990) *Biochemistry 29*, 7402–7407.

Eberle, W., Pastore, A., Sander, C., & Rösch, P. (1991) *J. Biomol. NMR 1*, 71–82.

Gariépy, J., Sykes, B. D., & Hodges, R. S. (1983) *Biochemistry 22*, 1765–1772.

Gariépy, J., Kay, L. E., Kuntz, I. D., Sykes, B. D., & Hodges, R. S. (1985) *Biochemistry 24*, 544–550.

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

Ikura, M., Spera, S., Barbato, G., Kay, L. E., Krinks, M., & Bax, A. (1991) *Biochemistry 30*, 9216–9228.

Kay, L. E., Forman-Kay, J. D., McCubbin, W. D., & Kay, C. M. (1991) *Biochemistry 30*, 4323–4333.

Kline, A. D., Braun, W., & Wüthrich, K. (1988) *J. Mol. Biol. 204*, 675–724.

Kördel, J., Forsén, S., & Chazin, W. J. (1989) *Biochemistry 28*, 7065–7074.

Kretsinger, R. H., & Nockolds, C. E. (1973) *J. Biol. Chem. 248*, 3313–3326.

Leavis, P. C., Rosenfeld, S. S., Gergely, J., Grabarek, Z., & Drabikowski, W. (1978) *J. Biol. Chem. 253*, 5452–5459.

Marsden, B. J., Hodges, R. S. & Sykes, B. D. (1988) *Biochemistry 27*, 4198–4206.

Marsden, B. J., Hodges, R. S. & Sykes, B. D. (1988) *Biochemistry 28*, 8839–8847.

Monera, O. D., Shaw, G. S., Zhu, B.-Y., Sykes, B. D., Kay, C. M., & Hodges, R. S. (1992) *Protein Sci. 1*, 945–955.

Neuhaus, D., Wagner, G., Vasak, M., Kagi, J. H. R., & Wüthrich, K. (1985) *Eur. J. Biochem. 151*, 257–273.

Oas, T. G., McIntosh, L. P., O'Shea, E. K., Dahlquist, F. W., & Kim, P. S. (1990) *Biochemistry 28*, 2891–2894.

Padilla, A., Cavé, A., & Parello, J. (1988) *J. Mol. Biol. 204*, 995–1017.

Pastore, A., De Francesco, R., Barbato, G., Morelli, M. A. C., Motto, A., & Cortese, R. (1991) *Biochemistry 30*, 148–153.

Piantini, U., Sørensen, O. W., & Ernst, R. R. (1982) *J. Am. Chem. Soc. 104*, 6800–6801.

Reid, R. E., Gariépy, J., Saund, A. K., & Hodges, R. S. (1981) *J. Biol. Chem. 256*, 2742–2751.

Ryckaert, J. P., Ciccotti, G., & Berendsen, H. J. C. (1977) *J. Comput. Phys. 23*, 327–337.

Satyshur, K. A., Rao, S. T., Pyzalska, D., Drendel, W., Greaser, M., & Sundaralingam, M. (1988) *J. Biol. Chem. 263*, 1628–1647.

Shaw, G. S., Hodges, R. S., & Sykes, B. D. (1990) *Science 249*, 280–283.

Shaw, G. S., Hodges, R. S., & Sykes, B. D. (1991a) *Biochemistry 30*, 8339–8347.

Shaw, G. S., Golden, L. F., Hodges, R. S., & Sykes, B. D. (1991b) *J. Am. Chem. Soc. 113*, 5557–5563.

Shaw, G. S., Findlay, W. A., Semchuk, P. D., Hodges, R. S. & Sykes, B. D. (1992) *J. Am. Chem. Soc. 114*, 6258–6259.

States, D. J., Haberkorn, R. A., & Ruben, D. J. (1982) *J. Magn. Reson. 48*, 286–292.

Strynadka, N. C. J., & James, M. N. G. (1989) *Annu. Rev. Biochem. 58*, 951–998.

Swain, A. L., Kretsinger, R. H., & Amma, E. L. (1989) *J. Biol. Chem. 264*, 16620.

Szebenyi, D. M. E., & Moffat, K. J. (1986) *J. Biol. Chem. 261*, 8761.

Tsuda, S., Hasegawa, Y., Yoshida, M., Yagi, K., & Hikichi, K. (1988) *Biochemistry 27*, 4120–4126.

van Gunsteren, W. F., Boelen, R., Kaptein, R., Scheek, R. M., & Zuiderweg, E. R. P. (1985) in *Molecular Dynamics and Protein Structure* (Hermans, J., Ed.) pp 92–99, Polycrystal Book Service, Western Springs, IL.

Wagner, G., Braun, W., Havel, T. F., Schaumann, T., Go, N., & Wüthrich, K. (1987) *J. Mol. Biol. 196*, 611–641.

Weiss, M. A. (1990) *J. Magn. Reson. 86*, 626–632.

Williamson, M. P., & Neuhaus, D. (1987) *J. Magn. Reson. 72*, 369–375.

Wishart, D. S., Sykes, B. D., & Richards, F. M. (1991) *J. Mol. Biol. 222*, 311–333.

Wüthrich, K., Billeter, M., & Braun, W. (1983) *J. Mol. Biol. 169*, 949–961.

Yip, P. F. (1990) *J. Magn. Reson. 90*, 382–383.

Zagorski, M. G., Bowie, J. U., Vershon, A. K., Sauer, R. T., & Patel, D. J. (1989) *Biochemistry 28*, 9813–9825.