Fry, M., & Loeb, L. A. (1986) *Animal Cell DNA Polymerases*, CRC, Boca Raton, FL.
Karawya, E., Swack, J., Albert, W., Fedorko, J., Minna, J. D., & Wilson, S. H. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 7777–7781.
Kornberg, A. (1969) *Science (Washington, D.C.) 163*, 1410–1418.
Kozak, M. (1984) *Nucleic Acids Res. 12*, 857–872.
Kunkel, T. A., & Loeb, L. A. (1981) *Science (Washington, D.C.) 213*, 765–767.
Larson, K., Sahm, J., Shenkar, R., & Strauss, B. (1985) *Mutat. Res. 150*, 77–84.
Lechner, R. L., Engler, M. J., & Richardson, C. C. (1983) *J. Biol. Chem. 258*, 11174–11184.
Maxam, A. M., & Gilbert, W. (1980) *Methods Enzymol. 65*, 499–560.
Mosbaugh, D. W., & Linn, S. (1983) *J. Biol. Chem. 258*, 108–118.
Planck, S. R., & Wilson, S. H. (1980) *J. Biol. Chem. 255*, 11547–11556.
Sanger, F., Coulson, A. R., Barrell, B. G., Smith, A. J. H., & Roe, B. A. (1980) *J. Mol. Biol. 143*, 161–178.
SenGupta, D. N., Zmudzka, B. Z., Kumar, P., Cobianchi, F., Skowronski, J., & Wilson, S. H. (1986) *Biochem. Biophys. Res. Commun. 136*, 341–347.
Siedlecki, J. A., Nowak, R., Soltyk, A., & Zmudzka, B. (1981) *Acta Biochim. Pol. 28*, 157–173.
Stalker, D. M., Mosbaugh, D. W., & Meyer, R. R. (1976) *Biochemistry 15*, 3114–3121.
Swack, J. A., Karawya, E., Albert, W., Fedorko, J., Minna, J. D., & Wilson, S. H. (1985) *Anal. Biochem. 147*, 10–21.
Tanabe, K., Bohn, E. W., & Wilson, S. H. (1979) *Biochemistry 18*, 3401–3406.
Wang, T. S.-F., & Korn, D. (1982) *Biochemistry 21*, 1597–1608.
Weaver, D. T., & DePamphilis, M. L. (1982) *J. Biol. Chem. 257*, 2075–2086.
Yamaguchi, M., Tanabe, K., Taguchi, Y. N., Nishizawa, M., Takahashi, T., & Matsukage, A. (1980) *J. Biol. Chem. 255*, 9942–9948.
Zmudzka, B. Z., SenGupta, D., Matsukage, A., Cobianchi, F., Kumar, P., & Wilson, S. H. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 5106–5110.

Comparison of the Crystal and Solution Structures of Calmodulin and Troponin C†

Douglas B. Heidorn and Jill Trewhella*
Life Sciences Division and Neutron Scattering Center, Los Alamos National Laboratory, Los Alamos, New Mexico 87545
Received September 11, 1987

ABSTRACT: X-ray solution scattering data from skeletal muscle troponin C and from calmodulin have been measured. Modeling studies based on the crystal structure coordinates for these proteins show discrepancies between the solution data and the crystal structure that indicate that if the size and shape of the globular domains are the same in solution as in the crystal, the distances between them must be smaller by several angstroms. Bringing the globular domains closer together requires structural changes in the interconnecting helix that joins them.

Calmodulin is a small (M_r 16 700), highly conserved protein that is present in all eukaryotic cells and plays a regulatory role in a diverse set of calcium-dependent cellular processes (Cheung, 1980). Troponin C (M_r 18 000) is a calcium-dependent regulator in muscle systems that shows strong sequence homology with calmodulin. These proteins can cross-react in their respectively biological systems [Means and Dedman (1980) and references cited therein]. The crystal structures of calmodulin (Babu et al., 1985) and skeletal troponin C (Herzberg & James, 1985; Sundaralingam et al., 1985a) have been solved, and they are very similar. Both structures show four Ca^{2+} binding sites each with the helix–loop–helix structural motif that has also been found in parvalbumin (Wery et al., 1985; Moews & Kretsinger, 1975) and in the vitamin D dependent calcium regulating protein from bovine intestine (Szebenyi et al., 1987).

The overall tertiary structures of troponin C and calmodulin are unusual in that they each show two globular domains (each having two Ca^{2+} binding sites) connected by a single α-helix, of about eight turns. The interdomain connecting helix is mostly exposed to solvent and forms few contacts with the rest of the molecule. Questions have been raised as to what stabilizes the interconnecting helix in these structures in the crystal forms and what rearrangements might occur in solution (Herzberg & James, 1985; Sundaralingam et al., 1985b). It has also been suggested that structural changes specifically in the interconnecting helix may be important in the calcium regulation (Babu et al., 1985; Herzberg & James, 1985; Sundaralingam et al., 1985a; Seaton et al., 1985).

Changes in the structure of the interconnecting helix can result in large effects on the overall shape of these proteins. Both troponin C and calmodulin are acidic proteins with p*I* values around 4.5, and the crystals used for the crystallographic analyses were grown at low pH (pH 5.0 and 5.5, respectively). In view of the potential flexibility in the interconnecting helix, it is important to examine the effects of a solution environment as well as changes in pH on the overall structure of these proteins. Fluorescence energy transfer measurements (Wang et al., 1986; Wang & Cheung, 1985) on skeletal troponin C suggest that it undergoes a large conformational rearrangement involving a change in the distance between the two calcium

† This project is supported by DOE/OHER Project HA-02-02-03/B04664.

0006-2960/88/0427-0909$01.50/0

binding domains between pH 5.0 and pH 6.7.

Small-angle scattering can be useful in exploring changes in the overall shape of molecules in solution under a variety of conditions. Seaton and co-workers (Seaton et al., 1985) have, in fact, used small-angle X-ray scattering to examine the effects of Ca^{2+} binding on the structure of calmodulin at roughly physiological pH and ionic strength. They observed a small increase in radius of gyration and maximum vector length on Ca^{2+} binding and interpreted this as indicating that the Ca^{2+} binding domains move apart on binding Ca^{2+}. At the time of their analysis the crystallographic coordinates for calmodulin were not available and so a detailed comparison of the solution scattering data with the crystal structure was not possible.

We have undertaken solution scattering experiments that address the relationship between the crystal structures of calmodulin and skeletal troponin C and their structures in solution under physiological pH and ionic strength. We have been specifically concerned with evaluating, and quantitating, the potential flexibility of the structure in the interconnecting helix region. Small-angle X-ray solution scattering data were collected for calmodulin with and without calcium at pH 7.4 (physiological pH) and at pH 5.5 (the pH of the crystallization experiments). Data were also collected on troponin C at pH 7.4 in the presence of Mg^{2+}. Protein aggregation has prevented the analysis of the effects of Ca^{2+} and low pH on the structure of troponin C. The crystallographic α-carbon coordinates were used to model the solution scattering data for both proteins. The model calculations based on the α-carbon coordinates do not, however, give good agreement with the solution scattering data for either protein. In each case, good agreement with the solution scattering data is obtained for models in which the structures of the two globular domains are preserved, but they have been brought into closer proximity by rearrangements in the interconnecting helix region.

Materials and Methods

Sample Preparations. Bovine brain calmodulin was purchased from Calbiochem and used without further purification. Rabbit skeletal troponin C was isolated according to Wilkinson (1985) and provided in pure form by Dr. Herbert Cheung. Samples for solution scattering were prepared in either 50 mM 3-(*N*-morpholino)propanesulfonic acid (MOPS) at pH 7.4 or 50 mM 2-(*N*-morpholino)ethanesulfonic acid (MES) at pH 5.5. All samples contained 100 mM KCl and 0.02% sodium azide. Ca^{2+}-free samples of calmodulin contained 5 mM ethylene glycol bis(β-aminoethyl ether)-*N*,*N*,*N*′,*N*′-tetraacetic acid (EGTA), while the Ca^{2+}-saturated samples contained 50 mM $CaCl_2$. Calmodulin samples were checked for Ca^{2+} binding by monitoring the tyrosine fluorescence [Klee and Vannaman (1982) and references cited therein]. Troponin C samples with Mg^{2+} contained 50 mM $MgCl_2$. Concentrations were determined both by a BIORAD assay and by measurement of the UV absorption. Extinction coefficients were $\epsilon^{1\%} = 0.19$ for calmodulin and $\epsilon^{1\%} = 0.175$ for troponin C (McCubbin et al., 1986). All reagents used were of analytical grade.

X-ray Scattering Measurements. X-ray data were collected on a small-angle X-ray scattering station at Los Alamos. This station uses a line source from a sealed X-ray tube that is driven by an Enraf Nonius generator, typically operated between 1.2 and 1.5 kW; 8.5-keV X-rays from a copper anode pass through a nickel filter and are focused, by glancing angle reflection from a bent glass mirror, at the plane of a Tennelec PSD-100 position-sensitive one-dimensional detector. The samples for measurement were positioned behind guard slits (to remove parasitic scattering) halfway between the mirror and the detector. The samples were centrifuged into a 1-mm glass capillary that was glued into a brass holder. The sample holder was maintained at a constant temperature (23 °C) by circulating water. The sample-to-detector distance was 33 cm. Data were recorded on an IBM-PC installed with a Nucleus PCA-8000 multichannel analyzer board. The response of the position-sensitive detector was checked for uniformity before and after every set of experiments. Backscatter from a lead beam stop was used to monitor the transmitted beam intensity, and this was used to normalize the intensity data. The net scattering from the protein molecules was calculated by subtracting a normalized buffer spectrum measured in the same sample cell. With use of the partial specific volume of the protein, a correction was made for the volume occupied by the protein molecules in the solution. Typical data collection times were 6–18 h, depending on protein concentration.

Remeasurement of the X-ray scattering data for samples showed no change in the $P(r)$ functions, nor in the R_g values. In addition, no changes were observed in the sodium dodecyl sulfate gel electrophoretic bands after X-ray irradiation. X-ray damage was, therefore, assumed to be minimal, consistent with the observations of Seaton et al. (1985) in the case of calmodulin. They also determined that X-ray irradiation did not inhibit the ability of calmodulin to activate bovine brain calmodulin-dependent cyclic nucleotide phosphodiesterase in a manner quantitatively indistinguishable from nonirradiated protein.

Scattering Data Analysis. Data were summed from detector channels at scattering angles of equal magnitude on opposite sides of the direct beam for all analyses. Data that had been corrected for slit smearing (Moore, 1980) were analyzed with the Guinier approximation (Guinier, 1939) as well as an indirect Fourier transform analysis developed by Moore (1980). Guinier analysis of data that had not been corrected for slit smearing generally gave radius of gyration values that were about 1 Å smaller than those obtained from corrected data.

The Guinier approximation says that for scattering from a point source by a homogeneous solution of monodisperse particles, at sufficiently small scattering angles, a plot of ln I versus $(2\theta)^2$ will be linear, with a slope that is related to the radius of gyration R_g by

$$R_g^2 = \frac{3\lambda^2}{4\pi^2} \frac{\Delta \ln I}{\Delta(2\theta)^2}$$

where I is the scattered intensity, 2θ the scattering angle, and λ the wavelength of the scattered radiation (1.54 Å for Cu Kα). R_g values were calculated from the straight-line fit of the scattering data that satisfied the condition $R_g q_{max} \leq 1.0$, where $q = (4\pi \sin \theta)/\lambda$. Statistical errors were propagated from the counting statistics in the fitting procedure, which also gave the net scattered intensity at zero scattering angle, I_0, and its statistical error.

For the $P(r)$ analysis, data were used out to $q^2 = 0.1$ Å^{-2}, and the length distributions, $P(r)$, were calculated direct from the scattering intensity by a Fourier inversion:

$$P(r) = \frac{1}{2\pi^2} \int_0^\infty I(q) qr \sin(qr)\, dq$$

The length distribution is simply the frequency of vectors connecting small-volume elements within the entire volume of the particle weighted according to their X-ray scattering power. Beyond a distance d_{max}, the maximum linear dimension of the particle, the $P(r)$ function vanishes. The radius of gyration can be computed from the second moment of $P(r)$:

$$R_g^2 = \frac{\int_0^{d_{max}} P(r) r^2 \, dr}{2 \int_0^{d_{max}} P(r) \, dr}$$

Length distributions were calculated for a range of d_{max} values. Typically, the computed reduced χ^2 values for the calculated and experimental intensities decreased from values much greater than 1 for small d_{max} to a stable minimum, and the computed R_g reached a plateau as d_{max} was increased. At significantly larger values of d_{max} the $P(r)$ function gave negative values. The final d_{max} was chosen according to the criteria of Glatter (1982), and the corresponding R_g value thus derived was checked for agreement with that determined with the Guinier analysis of the data that had been corrected for slit smearing.

Effects of protein concentration c on the computed R_g values from both the Moore and Guinier analyses were corrected by a least-squares linear extrapolation of plots of R_g^2 versus c to zero protein concentration according to Zaccai and Jacrot (1983). The experimental $P(r)$ functions were also calculated by extrapolating the intensity data to zero protein concentration (Moore, 1980).

Model Calculations. $P(r)$ functions were calculated from models based on the crystallographic α-carbon coordinates by a Monte Carlo integration technique. In this procedure homogeneous spheres representing each amino acid are positioned according to the α-carbon coordinates; the size and X-ray scattering power of each sphere is estimated from the partial specific volume and chemical composition of the amino acid it represents, respectively, with a correction applied for the solvent contribution to the scattering. The model structure is placed in a box, such that it is contained wholly within the box, and the Monte Carlo routine generates random points within the box. If a point is contained within the model, it is saved. Distances between every pair of saved points are weighted according to the product of the electron scattering powers at each point and binned. In the case of overlapping volumes for neighboring amino acids, the scattering power was taken as the average. The resulting histogram is the calculated $P(r)$ function. This calculation accounts for the internal scattering density fluctuations in the molecule.

Model calculations were also done with models that consisted of uniform density ellipsoids constructed for best fit to the molecular boundary determined from the crystal structures. Three or four ellipsoids were used; two of the ellipsoids corresponded to the two globular domains, and the interconnecting helix was modeled either as one ellipse (to represent the crystal structure) or as two ellipses at an angle ("bent" models). The ellipsoids are all prolate ellipsoids of revolution. For the troponin C crystal structure the semimajor and semiminor radii for each of the globular domains and the interconnecting helix were 22 Å and 14 Å, 20 Å and 13 Å, and 30 Å and 7 Å, respectively. For calmodulin the corresponding axes were 20.5 Å and 13 Å, 20.5 Å and 13 Å, and 30 Å and 7 Å. The interconnecting helix in each case penetrated the globular domains at either end in order to approximate a uniformly thick interconnecting region and to give a molecule of the appropriate length. For the uniform density ellipsoid calculations, only the molecular boundary is of importance, since a uniform scattering power is applied to all points within the boundary.

Results

Figure 1 shows Guinier regions of scattering curves for calmodulin and troponin C in each of the solution conditions

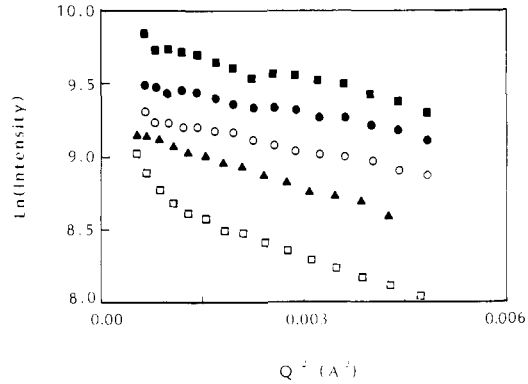

FIGURE 1: Guinier regions for calmodulin at pH 7.4 with EGTA (●) and with Ca^{2+} (■) and at pH 5.5 with EGTA (○) and with Ca^{2+} (□) and for troponin C at pH 7.4 with Mg^{2+} (▲).

Table I: R_g and d_{max} Values Determined for Troponin C and Calmodulin at Infinite Dilution[a]

		R_g (Å)	d_{max} (Å)
troponin C			
Mg^{2+}, pH 7.4	$P(r)$	23.0 ± 0.2	70 ± 3
	Guinier	23.0 ± 0.2	
crystal structure		23.9	72
bent model		22.7	68
calmodulin			
Ca^{2+}, pH 7.4	$P(r)$	21.3 ± 0.2 [21.5 ± 0.2]	63 ± 2 [62]
	Guinier	21.0 ± 0.6	
EGTA, pH 7.4	$P(r)$	19.6 ± 0.12 [20.6 ± 0.2]	59 ± 2 [58]
	Guinier	19.9 ± 0.4	
EGTA, pH 5.5	$P(r)$	19.4 ± 0.1	58 ± 2
	Guinier	19.2 ± 0.4	
crystal structure		22.8	70
bent model		21.2	62

[a] Values in square brackets are from Seaton et al. (1985).

measured. In the case of calmodulin for the pH 7.4 conditions and for the pH 5.5 without Ca^{2+} condition there is no evidence of any upward curvature at low 2θ in the Guinier plots that would result from aggregation in the samples. In addition, plots of I_0/c versus c, where c is the protein concentration for a given set of sample conditions, are linear and, when extrapolated to infinite dilution, give the same value, within errors, for the pH 7.4 samples both with and without Ca^{2+}, as well as the pH 5.5 sample without Ca^{2+}. This is also indicative that the protein is not aggregated in any of these conditions. At pH 5.5 with Ca^{2+}, however, aggregation is evident both in the anomolously high I_0/c values (they are approximately 50% larger than for the other solution conditions) and in the fact that the Guinier region of the scattering data shows more than one linear component.

In the presence of Mg^{2+} at pH 7.4, troponin C behaves as a monomer in solution as evidenced by a linear Guinier region and I_0/c dependence. Attempts were made to collect data from troponin C in the presence of Ca^{2+} as well as at low pH, but these conditions consistently gave problems due to protein aggregation.

The concentration dependence of R_g^2 values determined [using the $P(r)$ analysis] for troponin C and calmodulin under solution conditions for which they are monomeric is shown in Figure 2, and Table I summarizes the R_g values calculated by extrapolation to infinite dilution. R_g values quoted for the Guinier analysis are those determined from data corrected for slit smearing. The $P(r)$ and Guinier analyses give very good agreement for the R_g determinations, even though the Guinier R_g was not used to help select d_{max} in the $P(r)$ analysis.

Shown in Figure 3 are the experimental $P(r)$ data for calmodulin in the presence of Ca^{2+} and for troponin C in the

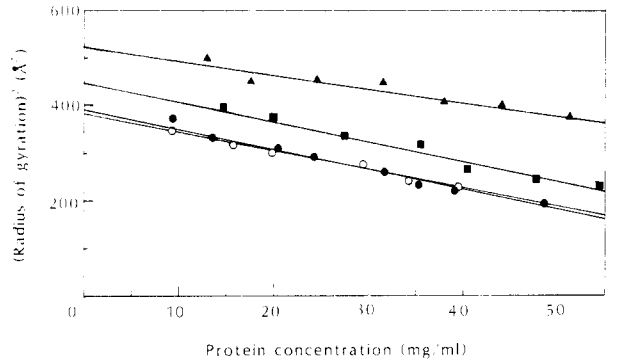

FIGURE 2: Concentration dependence of the R_g values determined with the $P(r)$ analysis for calmodulin, pH 7.4, with Ca^{2+} (■) and without Ca^{2+} (●); for calmodulin, pH 5.5, without Ca^{2+} (○); and for troponin C, pH 7.4, with Mg^{2+} (▲). The errors are represented by the size of the symbols.

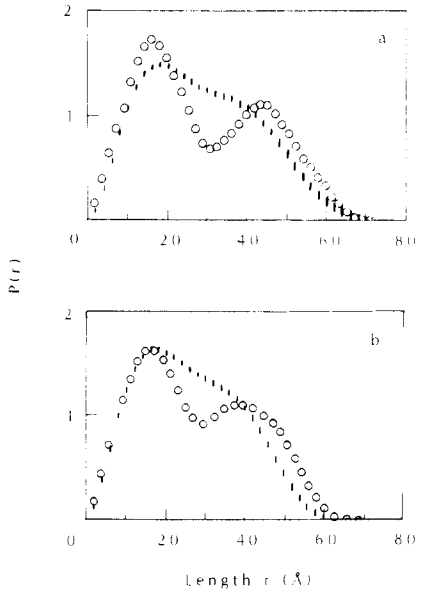

FIGURE 3: $P(r)$ functions calculated from the α-carbon crystal structure coordinates (○) and from the pH 7.4 solution scattering data (|) for (a) troponin C and (b) calmodulin. The size of the vertical bars indicate the propagated errors in the experimental data.

presence of Mg^{2+}, superimposed, in each case, with the calculated $P(r)$ function based on the respective α-carbon crystal coordinates and amino acid sequences. The experimental functions and those calculated from the crystal structures show differences that are well beyond the estimates of the experimental errors which are propagated from statistical errors in the data and from errors due to corrections to the data such as slit desmearing. In addition, the R_g and d_{max} values determined from the crystal structure $P(r)$'s are larger than the values determined experimentally (Table I), though this is more significant for calmodulin than for troponin C.

Figure 4 shows the calmodulin crystallographic α-carbon backbone superimposed with a dot representation (derived from the Monte Carlo calculation) of the uniform ellipsoid model that was used to approximate the crystal structure and its corresponding $P(r)$ function. The uniform ellipsoid $P(r)$ function is very similar to that calculated from the crystal structure coordinates. The $P(r)$ function based on the α-carbon coordinates shows a maximum, a minimum, and another maximum at 16, 29, and 39 Å, with relative intensities of 1.6, 0.9, and 1.1, respectively. The uniform ellipsoid model shows the same features at 16, 30, and 41 Å, with relative intensities of 1.7, 0.9, and 1.0%. Also shown in Figure 4 is the dot representation of a uniform density ellipsoid model that gives

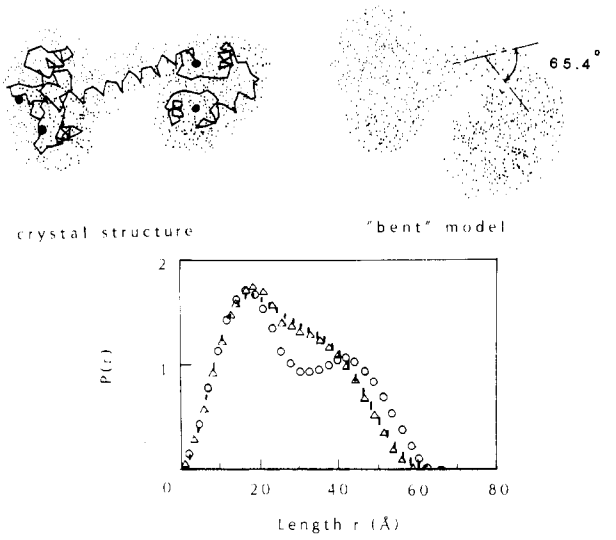

FIGURE 4: $P(r)$ functions calculated for calmodulin with the uniform ellipsoid model that best represents the crystal structure (○), the bent model (△), and the experimental curve (|). Also shown are the dot representations of the models derived from the Monte Carlo procedure. The crystallographic α-carbon backbone of calmodulin is superimposed on the crystal structure model.

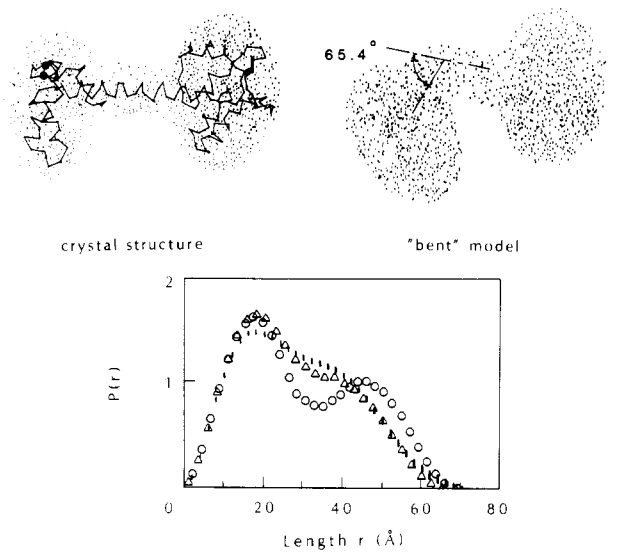

FIGURE 5: $P(r)$ functions calculated for troponin C with the uniform ellipsoid model that best represents the crystal structure (○), the bent model (△), and the experimental curve (|). Also shown are the dot representations of the models derived from the Monte Carlo procedure. The crystallographic α-carbon backbone of troponin C is superimposed on the crystal structure model.

a $P(r)$ function that agrees well with the experimental $P(r)$. This model was obtained by placing a bend in the interconnecting helix and rotating the globular domains with respect to each other. The bent model and the experimental $P(r)$ functions show a maximum at 18 Å, with a shoulder near 38 Å. In addition, the bent model predicts the measured R_g and d_{max} values within the experimental errors (Table I).

Figure 5 shows the results of similar model calculations done for troponin C. Again it was found there is relatively good agreement between the $P(r)$ calculated from the α-carbon coordinates and the uniform ellipsoid model $P(r)$ derived from the crystal structure. Similar to calmodulin, these $P(r)$ functions are characterized by a maximum, a minimum, and a maximum at 17, 31, and 45 Å, with relative intensities of 1.7, 0.7, and 1.1, for the α-carbon $P(r)$ and at 17, 32, and 45 Å, with relative intensities 1.6, 0.8, and 1.0, for the uniform

ellipsoid model. A model that gives good agreement with the experimental $P(r)$ curve was obtained by placing a bend in the interconnecting helix, again similar to calmodulin. This model curve, like the experimental curve, shows a maximum at 18 Å with a shoulder near 40 Å and gives R_g and d_{max} values that agree within error with the experimental values (Table I). One significant difference between this model $P(r)$ and experimental $P(r)$ is that the intensity of the maximum at 18 Å is approximately 10% higher for the model $P(r)$. Adjusting the angle of bend, or the relative rotations of the globular domains, does not improve the agreement in this region of the curve. Increasing the thickness of the interconnecting region does improve the fit, but for this to be real, parts of the globular domains would have to rearrange and contribute to the scattering density in that region. Though the discrepancy between the experimental and model $P(r)$ functions probably does reflect some real structural difference, the limited information content of the scattering data does not justify presenting models that require structural rearrangements within the globular domains without independent supporting data.

Models were also tested for which the radii of the uniform density ellipses were increased by 1–4 Å, maintaining their centers of mass at the locations indicated by the crystal structure. This was done in order to evaluate the potential influence of uniform hydration layers of different thickness on the $P(r)$ function. Such models gave very similar $P(r)$ functions to those of the crystal structures; i.e., they showed two well-resolved maxima, as well as giving d_{max} and R_g values that were significantly larger than the experimental values.

Discussion and Conclusions

The calculated $P(r)$ functions based on the crystal structures for both proteins have two distinct maxima, characteristic of a "dumbell"-shaped object. For such objects there will be a maximum near the value of the radius of each of the end domains and a second, less intense, maximum near the value of the distance between the two end domains. In the experimental curves the position of the first maximum is similar to that calculated for the crystal structure, but the second maximum has moved to shorter vector length, and the valley between the two maxima is filled in showing a higher probability of intermediate length vectors. In addition, there are significantly fewer vectors longer than 40 Å in the experimental curve, which results in smaller d_{max} and R_g values for the solution structure. These observations suggest that while the size and shape of the globular domains may be the same in solution as in the crystal, the distances between them must be smaller. The differences between the experimental and crystal structure $P(r)$ curves are well beyond the errors in the experimental data.

The question arises as to whether the differences between the crystal structure $P(r)$ and the experimental $P(r)$ can be attributed to experimental artifacts that result in a loss of resolution in the $P(r)$ function. This could only occur if there were something in the experimental arrangement that resulted in the scattering from the protein molecules being modulated by some function not related to the protein scattering. This would be the case if one region of the scattering profile was being artificially suppressed due to failure to correct for a nonuniform detector response or possibly insufficient counting statistics such that the low-intensity data at high q were suppressed and the scattering profile was effectively truncated. Failure to collect data over a sufficiently wide q range would have a similar effect. The q range and q resolution required to evaluate the $P(r)$ function adequately for an object of a given size are well defined by Moore (1980). For a scattering particle with a d_{max} of approximately 70 Å, q_{min} and Δq have to be ≤ 0.04 Å^{-1}. For our experiments $q_{min} = 0.01$ Å^{-1} and $\Delta q = 0.01$ Å^{-1}. In addition, since $P(r)$ is represented as Fourier series, data must then be measured to high enough q to determine at least three Fourier coefficients. Our experiments measured data to high enough q to determine at least five Fourier coefficients, more than sufficient to allow for the resolution of features seen in the crystal structure $P(r)$ functions. The Los Alamos X-ray station in its current configuration has a very low, featureless background (the mean background level is 10^{-6} of the main beam intensity). The detector is continually monitored for uniformity and is rewired if there are any nonuniformities greater than 5%; all data are nonetheless corrected for detector nonuniformity. Finally, data aquisition times were carefully chosen so that there were sufficient counts to define the scattering curve adequately over the entire q range required for the $P(r)$ analysis. The differences between the experimental and crystal structure $P(r)$ functions, therefore, cannot be attributed to lack of resolution or to experimental artifacts.

The crystallographic coordinates used for the modeling were those for rat testes calmodulin and turkey skeletal troponin C, whereas the solution scattering data were collected with bovine brain calmodulin and rabbit skeletal troponin C. In the case of skeletal troponin C the sequence homology between rabbit and turkey is very high (90% exact homology), and comparative Ca^{2+} binding and conformational studies using circular dichroism show them to be the same (Collins et al., 1977; McUbbin et al., 1986). In particular, the estimates of α-helix content are the same for both proteins (McUbbin et al., 1986). For calmodulin the sequences for bovine brain and rat testes are the same (Watterson et al., 1980; Dedman et al., 1978).

The crystal structure of calmodulin has all four Ca^{2+} binding sites saturated, while the crystal structure of troponin C only has the two high-affinity sites occupied by Ca^{2+}. Since Mg^{2+} can bind only at the two high-affinity sites of troponin C (Potter & Gergely, 1975) and since Ca^{2+} induces aggregation of troponin C, the Mg^{2+} form was chosen for comparisons with the crystal structure. It might be argued that different cations could have different efects on the overall tertiary structure of troponin C, and the comparison of the Mg^{2+} form to a form with two Ca^{2+} ions bound is not justified. In fact, this may be the reason for the discrepancy observed between our model and experimental $P(r)$ curves for troponin C. The solution scattering data on calmodulin, however, show that the Ca^{2+}-saturated form differs from the Ca^{2+}-free form only very subtly—showing evidence for only a slight extension of the structure on Ca^{2+} binding (Table I). The $P(r)$ functions in essence are very similar between the two forms, in that they show the characteristic maximum at 18 Å with a shoulder near 38–40 Å. In the absence of scattering data on the Ca^{2+} form of troponin C, comparisons with the Mg^{2+} data are reasonable.

The R_g values determined for calmodulin in the presence and absence of Ca^{2+} are systematically smaller than the values determined by Seaton et al. (1985; Table I). This arises from the fact that these workers extrapolated plots of R_g versus protein concentration rather than using R_g^2. If the same treatment is followed for our data, perfect agreement is obtained with their values, within experimental error. Using R_g^2 is a more rigorous treatment for the effects of interparticle interference (Zaccai & Jacrot, 1983). We find the same relative increase in R_g on Ca^{2+} binding as that observed by Seaton and co-workers. In addition, the $P(r)$ functions de-

termined in these independent studies are the same.

No significant change was observed in the $P(r)$ function (not shown) or in R_g and d_{max} at pH 5.5 compared with pH 7.4 for calmodulin in the absence of Ca^{2+}, indicating there is no change in the overall shape of the molecule between the two pH values under these solution conditions. In the presence of Ca^{2+}, however, lowering the pH induced protein aggregation that prohibited analysis of the solution scattering data for the shape parameters. Cheung and co-workers (private communications) measured hydrodynamic data (sedimentation coefficients and frictional ratios) for calmodulin at pH 7.5 and pH 5.2 and concluded that at low pH, in the presence of caldium, dimers formed. The solution scattering data are consistent with this observation.

The R_g value determined for troponin C in the presence of Mg^{2+} is larger than that determined for calmodulin with Ca^{2+} (23.0 ± 0.2 Å compared with 21.3 ± 0.2 Å). The $P(r)$ function determined for troponin C is qualitatively similar to that determined for calmodulin, but it does show a larger d_{max} value (70 Å compared with 63 Å). Both of these observations are consistent with the fact that troponin C is larger than calmodulin, having an additional 14 residues. Uncertainties concerning aggregation effects have, to date, prevented a complete analysis of troponin C data collected at pH 7.4 in the presence of Ca^{2+}, or at pH 5.5 where fluorescent energy transfer measurements indicate a large conformational rearrangement (Wang et al., 1986; Wang & Cheung, 1985). It would be extremely valuable to be able to identify solution conditions for which either calmodulin or troponin C (or both) was stabilized in the crystal conformation in a monodisperse solution, but efforts to identify such conditions by use of buffers related to those used for crystallization have failed due to aggregation problems. This is not surprising since crystallization and aggregation are closely linked phenomena.

The "bent" structures considered in analyzing the solution scattering results are attractive because they require only modifications to the crystal structure in the region that is the most likely region of structural instability. The globular domains in the troponin C and calmodulin crystal structures are similar to globular domains observed in other proteins, and they would be expected to be stable in solution. The interconnecting helix, however, in each case has very few contacts with the rest of the structure, and it is clearly important to understand what stabilizes it in the crystal environment. Sundaralingam and co-workers (Sundaralingam et al., 1985b) have proposed that intrahelical salt linkages might be important in stabilizing the interconnecting helix in the troponin C crystal. Herzberg and James (1985), on the other hand, observed that in the troponin C crystals the N-terminal domain of a symmetry-related molecule packs between the two domains of each troponin C. However, there are few direct intermolecular contacts between the symmetry-related molecule and the interconnecting helix, and so this may not be a factor in its stabilization. Such an interaction does not occur in the calmodulin crystals (Babu et al., 1985). Herzberg and James have also suggested that in troponin C the interconnecting helix may not be stable in solution and could fold around the "helix-breaker" glycine-92 that occurs in this part of the structure. Thus they propose a mechanism by which the N- and C-terminal globular domains of troponin C could interact.

The experimental data can be fitted, at least equally as well as for the bent models above, with more detailed models derived by starting with the crystal structure coordinates and by using specific bond rotations in the interconnecting helix to bring the globular domains closer together and give a new set of α-carbon coordinates. We have deliberately refrained from presenting such models to avoid the implication that the solution scattering data support a specific model but rather that the data suggest the nature of the conformational rearrangement only. Solution scattering data are always isotropically averaged, and it is therefore not usually possible to prove one particular model is correct on the basis of solution scattering data alone. It is possible, however, to test models. What is clear from these studies is that the crystal structures do not predict the solution scattering data on the basis of comparisons of the $P(r)$ functions. The differences in the $P(r)$ functions cannot be attributed to experimental artifacts, nor can including larger uniform hydration layers in the structures improve the agreement between the experimental and crystal $P(r)$ functions. Models that have been developed predict the solution scattering data significantly better than the crystal structures do, but there may be other alternatives. One possibility is that in solution there are a number of conformational substates and that the distances between the globular domains could vary over a range of values and the solution scattering reflects the "average" conformation. Whatever the case, if the structures of the globular domains are preserved with regard to overall size and shape between the crystal and solution states, then for both troponin C and calmodulin they must, on average, be in closer proximity than in the crystal structure in order to be consistent with the solution scattering data. Specifically, the modeling indicates that for troponin C and calmodulin the centers of mass are closer by approximately 5 Å, while the distance of closest approach for the ellisoid surfaces are smaller by approximately 9 and 11.0 Å, respectively.

It is evident from the solution scattering studies that, for this particular structural class of calcium regulating proteins, it will be very important to pursue studies that can directly address the solution structures quantitatively and that this will be particularly important in evaluating models for regulation that involve conformational rearrangements.

ACKNOWLEDGMENTS

We thank Sue E. Rokop for technical assistance in sample preparations. We also thank Herbert C. Cheung for providing us with pure rabbit troponin C. We thank Michael N. G. James for making available the crystal structure coordinates of troponin C, William J. Cook and Charles E. Bugg for the coordinates of calmodulin, David Torney for assistance with the modeling, and Stephen P. Edmondson for computer graphics representations of the crystal structures. We also thank Renu Kowluru for technical advice on calmodulin preparations. Finally, thanks to Gerald Johnson, Philip Seeger, and Murlin Nutter for assistance and technical advice on setting up the small-angle X-ray scattering station, to Peter B. Moore for making available the data analysis software, and Rex Hjelm for reading the manuscript critically.

REFERENCES

Babu, Y. S., Sack, J. S., Greenhough, T. J., Bugg, C. E., Means, A. R., & Cook, W. J. (1985) *Nature (London) 315*, 37.

Cheung, W. Y. (1980) *Science (Washington, D.C.) 207*, 19.

Collins, J. H., Greaser, M. L., Potter, J. D., & Horn, M. J. (1977) *J. Biol. Chem. 252*, 6356.

Dedman, J. R., Jackson, R. L., Schreiber, N. E., & Means, A. R. (1978) *J. Biol. Chem. 253*, 343.

Glatter, O. (1982) in *Small Angle X-ray Scattering* (Glatter O., & Kratky, O., Eds.) pp 119–196, Academic, New York.

Guinier, A. (1939) *Ann. Phys. (Paris) 12*, 161.
Herzberg, O., & James, M. N. G. (1985) *Nature (London) 313*, 653.
Klee, C. B., & Vannaman, T. C. (1982) *Adv. Protein Chem. 35*, 213.
McUbbin, W. D., Oikawa, K., & Kay, C. M. (1986) *FEBS Lett. 195*, 17.
Means, A. R., & Dedman, J. R. (1980) *Nature (London) 258*, 73.
Moews, P. C., & Kretsinger, R. H. (1975) *J. Mol. Biol. 91*, 201.
Moore, P. B. (1980) *J. Appl. Crystallogr. 13*, 168.
Porod, G. (1951) *Kolloid-Z. 124*, 83.
Potter, J. D., & Gergely, J. (1975) *J. Biol. Chem. 250*, 4628.
Seaton, B. A., Head, J. F., Engelman, D. M., & Richards, F. M. (1985) *Biochemistry 24*, 6740.
Sundaralingam, M., Bergstrom, R., Strasburg, G., Rao, S. T., Roychowdhury, P., Greaser, M., & Wang, B. C. (1985a) *Science (Washington, D.C.) 227*, 945.
Sundaralingam, M., Drendel, W., & Greaser, M. (1985b) *Proc. Natl. Acad. Sci. U.S.A. 82*, 944.
Szebenyi, D. M. E., Obendorf, S. K., & Moffat, K. (1987) *Nature (London) 294*, 327.
Wang, C. K., & Cheung, H. C. (1985) *Biochemistry 24*, 3365.
Wang, C.-L. A., Zhan, Q., Tao, T., & Gergely, J. (1986) *Biophys. J. 49*, 48a.
Watterson, D. M., Sharief, F., & Vanaman, T. C. (1980) *J. Biol. Chem. 255*, 962.
Wery, J. P., Dideberg, O., Charlier, P., & Gerday, C. (1985) *FEBS Lett. 182*, 103.
Wilkinson, J. M. (1985) *Biochim. Biophys. Acta 359*, 379.
Zaccai, G., & Jacrot, B. (1983) *Annu. Rev. Biophys. 12*, 139.

^1H and ^{31}P NMR Investigations of Actinomycin D Binding Selectivity with Oligodeoxyribonucleotides Containing Multiple Adjacent d(GC) Sites†

Elwood V. Scott,‡ Robert L. Jones,§ Debra L. Banville,‡ Gerald Zon,∥,⊥ Luigi G. Marzilli,*,‡ and W. David Wilson*,§

Department of Chemistry, Georgia State University, Atlanta, Georgia 30303, Department of Chemistry, Emory University, Atlanta, Georgia 30322, and Molecular Pharmacology Laboratory, Division of Biochemistry and Biophysics, Food and Drug Administration, Bethesda, Maryland 20892

Received March 17, 1987; Revised Manuscript Received September 22, 1987

ABSTRACT: Imino proton and ^{31}P NMR studies were conducted on the binding of actinomycin D (ActD) to self-complementary oligodeoxyribonucleotides with adjacent 5′-GC-3′ sites. ActD showed very high specificity for binding to GC sites regardless of oligomer length and surrounding sequence. For a first class of duplexes with a central GCGC sequence, a mixture of 1:1 complexes was observed due to the two different orientations of the ActD phenoxazone ring system. Analysis of ^1H chemical shifts suggested that the favored 1:1 complex had the benzenoid side of the phenoxazone ring over the G base in the central base pair of the GCGC sequence. This is the first case in which an unsymmetrical intercalator has been shown to bind to DNA in both possible orientations. A unique 2:1 complex, with significantly different ^1H and ^{31}P chemical shifts relative to those of the 1:1 complexes, was formed with these same oligomers, again with the benzenoid side of the ActD molecule over the G base of the central GC base pair. There is considerable anticooperativity to binding of the second ActD in a GCGC sequence. In titrations of oligomers with the GCGC sequence, only the two 1:1 complexes are found up to ratios of one ActD per oligomer. Increasing the ActD concentration, however, resulted in stoichiometric formation of the unique 2:1 adduct. Spectrophotometric binding studies indicated that the apparent binding equilibrium constant for a GC site adjacent to a bound site is reduced by approximately a factor of 20 relative to the ActD binding constant to an isolated GC site. Both upfield and downfield shifts were seen for imino proton signals for base pairs adjacent to ActD binding sites. This suggests that ActD has considerable long-range effects on oligomer conformation. Anticooperativity was also seen in NMR studies with a second class of oligomers containing alternating GC sequences longer than GCGC. It was found that in any three consecutive GC binding sites only two ActD can be bound. Anticooperativity can, thus, define the apparent number of base pairs in the binding site of an intercalator. Binding results with poly[d(G-C)]·poly[d(G-C)] were similar to those obtained with an oligomer containing four consecutive GC sites.

Actinomycin D (ActD)[1] was one of the first antibiotics whose interaction with DNA was investigated in depth [reviewed in Waring (1981)], and its DNA complex has been used as a model for DNA–protein interactions (Sobell, 1973). Müller and Crothers (1968) in a detailed hydrodynamic, kinetic, and thermodynamic study on several ActD analogues proposed that this molecule bound to DNA by intercalation,

† This work was supported by NSF Grant DBM-8603566 to W.D.W. and NIH Grant GM 29222 to L.G.M.
* Author to whom correspondence should be addressed.
‡ Emory University.
§ Georgia State University.
∥ Food and Drug Administration.
⊥ Present address: Applied Biosystems, 850 Lincoln Centre Drive, Foster City, CA 94404.

[1] Abbreviations: ActD, actinomycin D; DNA, deoxyribonucleic acid; PIPES, piperazine-*N*,*N*′-bis(2-ethanesulfonic acid); poly[d(G-C)]$_2$, poly[d(G-C)]·poly[d(G-C)]; EDTA, ethylenediaminetetraacetic acid; 2D, two dimensional; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; TSP, 3-(trimethylsilyl)propionic acid. Interchain base pairs are indicated with a dot, e.g., G·C, while intrachain sequences are indicated without a dot or dash, e.g., GC or GCGC; all sequences are written in the 5′ → 3′ direction.

0006-2960/88/0427-0915$01.50/0