Nucleic Acids Res. 14, 1107–1126.
Randall, S. K., Eritja, R., Kaplan, B., Petruska, J., & Goodman, M. F. (1987) *J. Biol. Chem. 262*, 6864–6870.
Sagher, D., & Strauss, B. (1983) *Biochemistry 22*, 4518–4526.
Scheek, R. M., Boelens, R., Russo, N., van Boom, J. H., & Kaptein, R. (1984) *Biochemistry 23*, 1371–1376.
States, D. J., Haberkorn, R. A., & Ruben, D. J. (1982) *J. Magn. Reson. 48*, 286–292.
Takeshita, M., Chang, C. N., Johnson, F., Will, S., & Grollman, A. P. (1987) *J. Biol. Chem. 262*, 10171–10179.
Weiss, B., & Grossman, L. (1987) *Adv. Enzymol. Relat. Areas Mol. Biol. 60*, 1–34.
Weiss, M. A., Patel, D. J., Sauer, R. T., & Karplus, M. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 130–134.
Zacharias, W., Larson, J. E., Klysik, J., Stirdivant, S. M., & Wells, R. D. (1982) *J. Biol. Chem. 257*, 2775–2782.

Crystal and Solution Structures of the B-DNA Dodecamer d(CGCAAATTTGCG) Probed by Raman Spectroscopy: Heterogeneity in the Crystal Structure Does Not Persist in the Solution Structure†

J. M. Benevides,‡ A. H.-J. Wang,§ G. A. van der Marel,‖ J. H. van Boom,‖ and G. J. Thomas, Jr.*,‡

Division of Cell Biology and Biophysics, School of Basic Life Sciences, University of Missouri—Kansas City, Kansas City, Missouri 64110, Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, and Gorlaeus Laboratory, Leiden State University, 2300RA Leiden, The Netherlands

Received June 8, 1987; Revised Manuscript Received September 22, 1987

ABSTRACT: The self-complementary dodecamer d(CGCAAATTTGCG) crystallizes as a double helix of the B form and manifests a Raman spectrum with features not observed in Raman spectra of either DNA solutions or wet DNA fibers. A number of Raman bands are assigned to specific nucleoside *sugar* and *phosphodiester* conformations associated with this model B-DNA crystal structure. The Raman bands proposed as markers of the crystalline B-DNA structure are compared and contrasted with previously proposed markers of Z-DNA and A-DNA crystals. The results indicate that the three canonical forms of DNA can be readily distinguished by Raman spectroscopy. *However, unlike Z-DNA and A-DNA, which retain their characteristic Raman fingerprints in aqueous solution, the B-DNA Raman spectrum is not completely conserved between crystal and solution states.* The Raman spectra reveal greater heterogeneity of nucleoside conformations (sugar puckers) in the DNA molecules of the crystal structure than in those of the solution structure. The results are consistent with conversion of one-third of the dG residues from the C2′-endo/anti conformation in the solution structure to another conformation, deduced to be C1′-exo/anti, in the crystal. The dodecamer crystal also exhibits unusually broad Raman bands at 790 and 820 cm^{-1}, associated with the geometry of the phosphodiester backbone and indicating a wider range of (α, ζ) backbone torsion angles in the crystal than in the solution structure. The results suggest that backbone torsion angles in the CGC and GCG sequences, which flank the central AAATTT sequence, are significantly different for crystal and solution structures, the former containing the greater diversity. The Raman data also suggest no significant change in the geometry of the central AAATTT domain between crystal and solution structures.

Methods of single-crystal X-ray diffraction analysis have provided high-resolution structures of a number of DNA oligomers in left-handed Z and right-handed A conformations (Wang et al., 1979, 1982; Wang & Rich, 1985; Shakked et al., 1983; McCall et al., 1985). Typically, the nucleotide repeat of A-DNA and the dinucleotide repeat of Z-DNA exhibit characteristic backbone conformations and, with the possible exception of terminal residues, a regularly repeating pattern of nucleoside sugar puckers. Many conformational properties of the Z and A families of DNA structure are also not particularly sensitive to the base sequence. On the other hand, the X-ray-determined secondary structure of the dodecamer sequence d(CGCGAATTCGCG), which has been recognized as that of B-DNA, manifests much less regularity in helix parameters and nucleoside conformations (Wing et al., 1980). Refinement of this structure by Dickerson and Drew (1981) has revealed a broader distribution of sugar puckers and base orientations than ordinarily found in oligomeric Z- and A-DNA crystals. It is not known whether such "microheterogeneity" of structure is a fundamental property of the B form of DNA, an attribute of the particular dodecamer sequence investigated, or a consequence of solvent–helix and helix–helix interactions in the crystal. In order to address questions of this sort, we are examining the crystal and solution structures of several sequences selected as candidates to form right-handed double helices of the B form. This paper reports the use of laser Raman spectroscopy to investigate the conformation of one such sequence, d(CGCAAATTTGCG), in both its crystal and its aqueous solution states.

The target sequence, d(CGCAAATTTGCG), is similar to that investigated by Dickerson and co-workers (Wing et al., 1980). Its composition differs only in 2 of the 12 nucleotide positions, with A replacing G at the fourth residue and T replacing C at the ninth residue from the 5′ end. The double

† Part XXXII in the series "Raman Spectral Studies of Nucleic Acids". This work was supported by grants from the National Institutes of Health (G.J.T. and A.H.-J.W.) and by a grant from the Burroughs Wellcome Co. (A.H.-J.W.).
* Author to whom correspondence should be addressed.
‡ University of Missouri—Kansas City.
§ Massachusetts Institute of Technology.
‖ Leiden State University.

0006-2960/88/0427-0931$01.50/0

helix of the present dodecamer contains 50% GC pairs and 50% AT pairs, as well as an internal sequence (AAATTT) considered to have the potential for forming "curved DNA" (Trifonov, 1984).

In previous studies, we have employed laser Raman spectroscopy in conjunction with results of X-ray diffraction analysis to identify bands in the vibrational Raman spectrum that are distinct indicators of Z-DNA and A-DNA structures (Thomas & Benevides, 1985; Benevides et al., 1984, 1986). A composite tabulation of these Raman markers (Thomas et al., 1986) has shown that many are distinct from the corresponding vibrational bands found in spectra of aqueous solutions of double-stranded calf thymus and viral DNAs and of complementary polydeoxynucleotides and oligodeoxynucleotides (Prescott et al., 1984, 1986; Thomas et al., 1987). Insofar as the latter structures are categorized generally as B-DNA,[1] it has been possible to infer that all three of the canonical structures of DNA yield distinctive Raman markers (Thomas & Wang, 1988). The present study examines for the first time the Raman spectrum of a DNA model for which the X-ray crystal structure indicates the B-DNA conformation. The results provide direct evidence that the Raman conformation markers of the B-form structure of DNA are clearly different from those of Z-DNA and A-DNA and that the B structure is not invariant to effects of crystallization.

Experimental Procedures

Materials. The oligonucleotide d(CGCAAATTTGCG) was synthesized by an improved phosphotriester technique in which 1-hydroxybenzotriazole was used as an activating reagent (van der Marel et al., 1981). After the deblocking of the fully protected molecule, the fragment was purified by Sephadex G-50 column chromatography and converted into the ammonium salt. The purity of the dodecamer was greater than 95%, as judged by high-performance liquid chromatography (HPLC) analysis. The crystallization of the dodecamer followed procedures described in detail elsewhere (Wang et al., 1982).

Crystal Structure. The crystal structure of the d(CGCAAATTTGCG) double helix has been refined by Hendrickson–Konnert refinement procedures at 2.5-Å resolution to a current R factor of 20.5%, and the refinement is continuing to include solvent molecules. The double-helical oligomer adopts an overall B-DNA conformation not unlike that of the dodecamer d(CGCGAATTCGCG) (Wing et al., 1980). While a detailed comparison between the two dodecamers must await the final refinement of the d(CGCAAATTTGCG) structure, some conclusions can be drawn from the current results. For example, the base pair stacking patterns along the double helix are quite similar to each other. However, when we examined the backbone torsion angles of the molecule of present interest, it was apparent that a large degree of variability occurs in them, particularly in the conformational angles α, ζ, and χ, which are associated with 5′-O–P, 3′-O–P, and glycosyl torsions, respectively.

Sample Handling for Raman Spectroscopy. Crystals of varying sizes up to 0.2 × 0.5 × 1.0 mm, authenticated by X-ray diffraction as referenced in the preceding section, were suspended in mother liquor in a sealed glass capillary. The mother liquor contained 2-methyl-2,4-pentanediol (40%) + 25 mM sodium cacodylate (pH 7) + 2 mM spermine + 10 mM $MgCl_2$. Spectra were excited with the 514.5-nm line of an argon laser (Coherent Innova 70-2) with 200–300 mW of radiant power at the sample. The Raman scattering at 90° was collected and analyzed on a Spex Ramalog spectrometer under the control of a North Star Horizon-II microcomputer. Spectral data were collected at intervals of 1.0 cm^{-1} with an integration time of 1.5 s and with a spectral slit width of either 8.0 or 4.0 cm^{-1} as indicated in the figure legends. All data collected from the crystals were corrected for Raman scattering of the mother liquor as described previously (Thomas & Benevides, 1985; Benevides et al., 1986).

Similar procedures were employed for collection of Raman spectra of solutions, obtained by dissolving the sodium salt of the dodecamer in 0.1 M NaCl to a DNA concentration of 25 mg/mL (2.5 wt %). Further details of instrumentation, sample handling, data collection, signal averaging, and Fourier deconvolution techniques are described in previous publications (Thomas & Benevides, 1985; Benevides et al., 1984, 1986; Thomas et al., 1986; Prescott et al., 1984; Thomas & Agard, 1984).

Results and Discussion

Comparison of the B-DNA Dodecamer in Crystal and Solution States. Figure 1a shows the Raman spectrum of the d(CGCAAATTTGCG) crystal. The frequencies of many of the Raman bands of this B-DNA model structure are labeled in cm^{-1} units at the peak positions. Principal assignments to base (A, T, G, or C) or backbone groups (OPO, PO_2^-, etc.), derived from earlier studies of mono-, oligo-, and polynucleotides [Thomas et al. (1986) and references cited therein], are listed in Table I.

Figure 1b shows the Raman spectrum of the d(CGCAAATTTGCG) solution, expected to represent also a B-DNA structure. Remarkably, the solution spectrum is significantly different from the crystal spectrum, and these differences are evident in the digital difference spectrum shown in Figure 1c. Here, the crystal is the minuend and the solution is the subtrahend, with the 1093-cm^{-1} band of the PO_2^- group employed for normalization of intensities. Many of the prominent difference bands occur in the regions 600–900 and 1200–1400 cm^{-1}. These regions are sensitive fingerprints of nucleoside conformation (Thomas et al., 1986) and are redisplayed on an expanded scale in Figure 2.

Comparison with Other Model Structures Shows That d(CGCAAATTTGCG) Assumes a B-DNA Conformation in Aqueous Solution. In previous work we have discussed the Raman spectra of aqueous solutions of poly(dG-dC)·poly(dG-dC) and poly(dA-dT)·poly(dA-dT) in relation to the Raman spectrum of calf thymus DNA (Thomas & Benevides, 1985; Benevides et al., 1986; Thomas et al., 1986). Like calf thymus DNA, both polydeoxynucleotides are believed to adopt structures of the B form in solutions of low ionic strength, and both exhibit Raman markers characteristic of B-DNA, as judged from comparison of their solution Raman spectra with spectra of wet DNA fibers that yield B-type fiber diffraction patterns (Prescott et al., 1984; Erfurth et al., 1972). In fact, it has been shown that the Raman spectrum of aqueous calf thymus DNA can be well represented as a simple sum of spectra of aqueous poly(dG-dC)·poly(dG-dC) and poly(dA-dT)·poly(dA-dT) (Thomas et al., 1986).

In view of the above, it is of interest to compare the spectrum of the d(CGCAAATTTGCG) solution (Figure 1b) with the

[1] We use the term "B-DNA" in its conventional context, viz., to designate the most stable conformation of DNA. Structural characteristics usually associated with the B form are 10.5-fold helix pitch, *anti*-glycosyl torsion, C2′-endo family sugar pucker (including C3′-exo, C1′-exo, and, occasionally, O4′-endo), low base tilt (~5–10°), and phosphodiester torsion angles (α, ζ) in the range (−50 ± 10°, −120 ± 30°). Further discussion is given in the work of Arnott et al. (1976) and Saenger (1984).

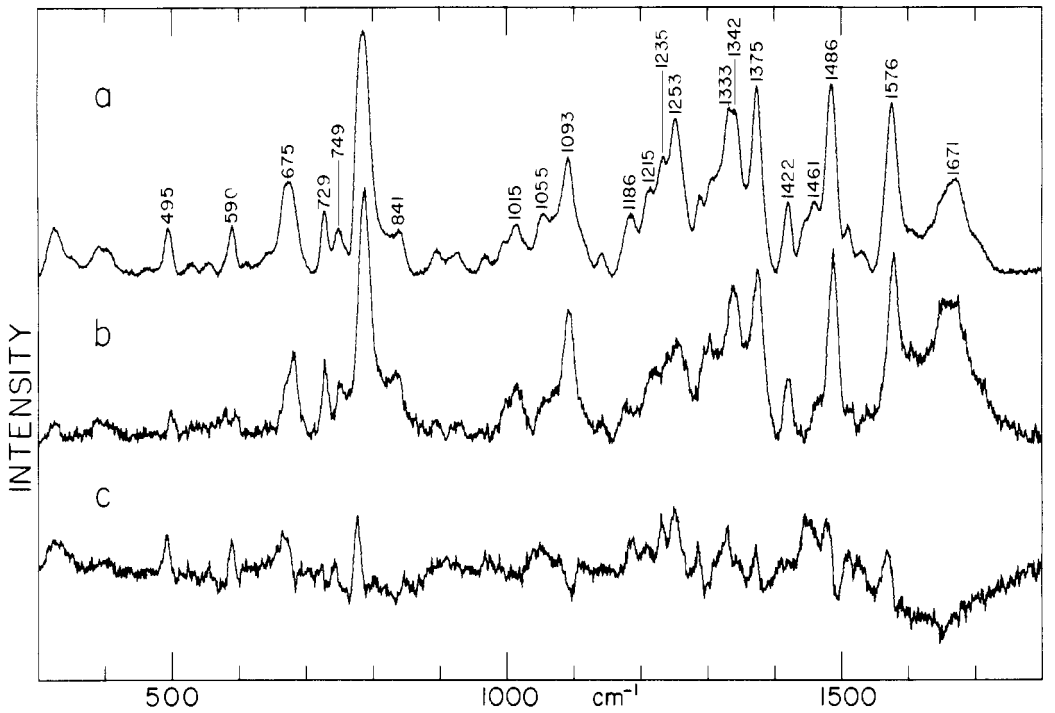

FIGURE 1: Raman spectra in the region 300–1800 cm^{-1} of the dodecamer d(CGCAAATTTGCG) in crystal (a) and solution (b) forms and their difference spectrum (c). Experimental conditions are described in the text. Methods for correction of contributions from mother liquor and solvent have been described (Benevides et al., 1985). The band near 1093 cm^{-1} has been used as the basis for normalization of intensities in computing the difference spectrum. Band assignments and relative intensities are given in Table I.

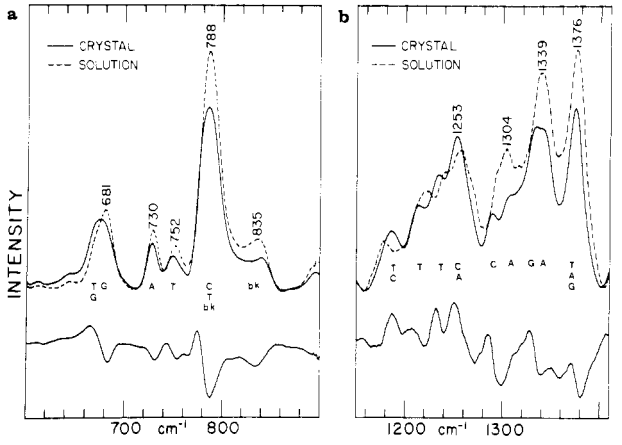

FIGURE 2: (a) Comparison of the Raman spectra of the dodecamer crystal (—) and solution (– –) in the region 600–900 cm^{-1} plotted on an expanded scale and the corresponding difference spectrum normalized to compensate positive and negative difference bands due to dG residues in the 660–690-cm^{-1} interval. (See text.) (b) Similar presentation of data for the region 1200–1400 cm^{-1}. Abbreviations indicate bands assigned to the nucleosides (A, C, G, or T) or backbone (bk).

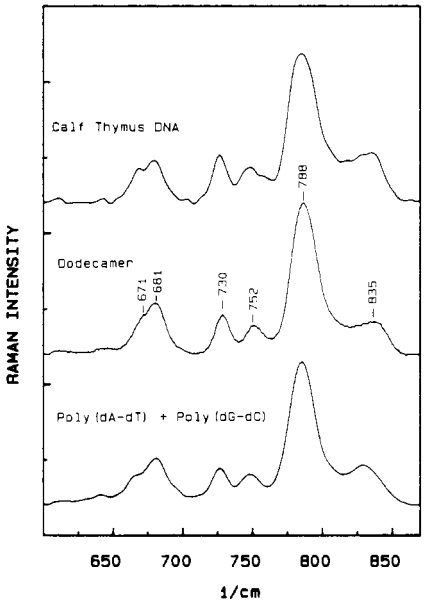

FIGURE 3: Comparison of the Raman spectrum (600–900 cm^{-1}) of the dodecamer solution (middle) with the spectrum of a calf thymus DNA solution (top) and with a composite of spectra of poly(dG-dC)·poly(dG-dC) and poly(dA-dT)·poly(dA-dT) solutions (bottom). The virtual identity of all three Raman profiles is apparent.

spectra of aqueous calf thymus DNA and with the sum of spectra of the aqueous deoxynucleotide copolymers. Such a comparison is shown in Figure 3, for the 600–900-cm^{-1} region. Figure 3 shows a close correspondence between Raman bands of aqueous d(CGCAAATTTGCG) and those of aqueous calf thymus DNA. Further, Figure 3 shows that the spectrum of the aqueous dodecamer can be well represented as a simple sum of spectra of the polynucleotides poly(dG-dC)·poly(dG-dC) and poly(dA-dT)·poly(dA-dT). Although the data shown in Figure 3 span only the 600–900-cm^{-1} interval, similar correlations exist throughout the 300–1800-cm^{-1} region. These results indicate that in aqueous solution the secondary structure of the d(CGCAAATTTGCG) duplex is similar to that of calf thymus DNA and to a composite of secondary structures of the deoxynucleotide copolymers. *Thus, d-(CGCAAATTTGCG) assumes a B-DNA secondary structure in aqueous solution.*

From previous studies of mononucleotide and oligonucleotide model compounds it has been determined that the nucleoside conformation common to aqueous solutions of calf thymus DNA and deoxynucleotide copolymers contains a furanose pucker of the *C2'-endo* family, with *anti*-glycosyl bond orientation and phosphodiester dihedral angles (α, ζ) in the range ($-50 \pm 10°$, $-120 \pm 30°$). These features are considered characteristic of B-DNA (Thomas et al., 1986; Benevides et al., 1984; Arnott et al., 1976). Figure 3 shows that the same conformational properties apply also to the miniature double

Table I: Raman Frequencies and Assignments of the B-DNA Dodecamer d(CGCAAATTTGCG)

crystal[a]	solution[a]	assignment[b]
325 (1.8 B)	325 (0.6)	A
344 (0.8)		d
391 (1.2)	389 (0.9)	d
402 (1.1)	402 (0.8)	d
463 (0.3)		d
495 (1.9)	500 (1.2)	G, T
529 (0.4)		A
556 (0.5)		d
	580 (0.9)	G, C
590 (2.0)	596 (1.0)	G, C
612 (0.4)		A, C, T
645 (1.0)		A, C
	671 (2.3)	T
675 (3.8 B)		G, T
	681 (3.2)	G
729 (2.6)	730 (2.6)	A
749 (1.9)	752 (2.2)	T
786 (10.0)	788 (10.0)	C, T, bk (OPO)
841 (1.8)	835 (2.8)	bk (OPO)
896 (1.1)	894 (0.8)	d
926 (0.9)	926 (0.6)	d
969 (0.8)		d
996 (1.4 S)	999 (1.6 S)	d
1015 (2.1)	1015 (2.3)	d
1055 (2.5)	1058 (1.6)	d (CO)
1093 (4.7)	1094 (5.2)	bk (PO_2^-)
	1109 (1.6)	?
1143 (0.8)	1143 (1.0)	d (CC), A, T
1186 (2.5)	1179 (1.3)	T, C
1215 (3.6)	1220 (2.7)	T
1235 (4.9)	1242 (3.3)	T
1253 (6.4)	1256 (3.7)	C, A
1290 (3.1)		C
1307 (3.9 S)	1304 (4.3)	A
1333 (6.9)		G
1342 (6.7)	1339 (6.1)	A, G
1375 (7.6)	1376 (6.9)	T, A, G
1422 (2.9)	1421 (2.4)	d (CH_2)
1448 (2.2 S)		d (CH_2)
1461 (3.0)	1468 (1.5)	d (CH_2)
1486 (7.7)	1489 (7.8)	A, G
1511 (1.9)	1513 (1.2)	A
1531 (1.0)	1540 (0.9)	C[c]
1576 (7.1)	1579 (7.6)	G, A
1606 (1.8)	1607 (3.7 S)	C[d]
1671 (3.9 B)		T (C=O)
1696 (1.6 S)	1699 (2.6 S)	G (6C=O)[c]

[a] Frequencies are in cm^{-1} units. Values in parentheses are relative intensities on a 0–10 scale. S = shoulder; B = broad band. [b] From Prescott et al. (1984) and references cited therein. A = adenine, C = cytosine, G = guanine, T = thymine, d = deoxyribose, and bk = backbone. Symbols in parentheses indicate specific group vibrations where known. [c] This work. [d] Benevides & Thomas, 1983.

helix of d(CGCAAATTTGCG) in aqueous solution. The pertinent Raman markers in the 600–900-cm^{-1} interval and their conformational significance are as follows: shoulder at 671 and band at 752 cm^{-1} (C2′-endo/anti dT), band at 681 cm^{-1} (C2′-endo/anti dG), band at 730 cm^{-1} (C2′-endo/anti dA), band at 788 cm^{-1} (C2′-endo/anti dC, except as noted below), and broad band centered at 835 cm^{-1} (OPO geometry of B-form backbone) (Thomas et al., 1986). In addition, the B-form backbone contributes a band near 790 cm^{-1} which cannot be resolved from the overlapping dC band except by Fourier deconvolution methods. Deconvolution shows that the dC component of the band is actually centered close to 780 cm^{-1} (Benevides et al., 1986). The presence of two overlapping band components accounts for the high intensity and unusual breadth of the composite 788-cm^{-1} band (Thomas et al., 1986).

In the C2′-endo family we include C1′-exo and C3′-exo puckers. The O4′-endo pucker is sometimes also included in this family, although in terms of the pseudorotation angle P this conformation is intermediate between C2′-endo and C3′-endo puckering (Saenger, 1984). A recent study has suggested that the Raman spectrum might be used to distinguish the various members of the C2′-endo family of furanose puckers from one another in mononucleoside and mononucleotide crystals (Nishimura et al., 1986). However, it has not yet been established that the Raman spectrum can be used to definitively distinguish these closely related geometric conformers from one another in oligomeric DNA, even though as a group they can be easily distinguished from C3′-endo/anti nucleosides, as well as from C3′-endo/syn nucleosides, in both low and high molecular weight DNAs (Benevides et al., 1984, 1986). Further discussion of this subject has been given elsewhere (Thomas et al., 1986; Nishimura & Tsuboi, 1986).

Crystal Structure of d(CGCAAATTTGCG) Deviates from the Aqueous B Conformation. The Raman results discussed in the two preceding sections imply that the structure of the d(CGCAAATTTGCG) crystal contains significant differences from the B-form structures of the d(CGCAAATTTGCG) solution and of DNA solutions. The nature of the spectral differences between dodecamer crystal and dodecamer solution are shown for two informative regions of the vibrational spectrum in Figure 2. On the other hand, X-ray diffraction of the d(CGCAAATTTGCG) crystal suggests predominantly a B-DNA structure (Coll et al., 1987). This apparent conflict is resolved by considering the probable origins of the bands in the difference Raman spectrum of Figure 2 and the limits in resolution of both the Raman and X-ray structures.

Figure 2a shows that the principal OPO marker band of the d(CGCAAATTTGCG) crystal occurs at 841 cm^{-1}, which is within the interval expected for a structure containing the B-form backbone (Thomas et al., 1986). The band intensity, half-width, and peak position are all consistent with the B-DNA backbone, insofar as the local geometry at the phosphodiester groups is concerned. Further, the absence of a strong band near 807 ± 3 cm^{-1} conclusively rules out the OPO geometry of right-handed A-form DNA [cf. Figure 1 in Benevides et al. (1986)]. Likewise, a left-handed Z-DNA conformation can be categorically rejected on the basis of the absence of Raman markers for C3′-endo/syn dG conformers [Table I in Benevides et al. (1984)]. In summary, the Raman spectrum shows that the crystalline dodecamer contains primarily the OPO geometry associated with the right-handed B-DNA backbone. Yet, the fact that the 841-cm^{-1} B-DNA marker in the crystal is significantly displaced from the 835-cm^{-1} B-DNA marker in the d(CGCAAATTTGCG) solution spectrum is evidence that these two B backbones are not identical.

The difference spectrum of Figure 2a also indicates a much greater half-width for the composite 786-cm^{-1} band in the crystal than for its solution counterpart. Since the contribution from deoxycytidine (ca. 780 cm^{-1}, see above) to this composite band is not highly structure sensitive (Benevides & Thomas, 1983), the simplest explanation for the observed dissimilarity is that the OPO contribution to the band (ca. 790 cm^{-1}) is much broader in the crystal than in the solution spectrum. Increased breadth of a Raman band usually results from an increase in the distribution of molecular structures that generate the band, in this case greater structural "heterogeneity" of the population of OPO groups in the crystalline dodecamer than in the solution dodecamer. The precise distribution of OPO geometries in the d(CGCAAATTTGCG) crystal cannot be deduced unambiguously from the available data because discrete peaks are unresolved in the 790–830-cm^{-1} interval.

Nonetheless, the crystal's larger bandwidth at 786 cm^{-1} and its apparently higher OPO frequency at 841 cm^{-1} (versus 835 cm^{-1} in the solution spectrum) both indicate a different B backbone geometry vis-ã-vis the solution structure.

Another important dissimilarity between crystal and solution spectra of d(CGCAAATTTGCG) is evident in the shape of the complex band near 665–685 cm^{-1}. In aqueous B-DNA, including the calf thymus DNA solution shown in Figure 3, we expect and find a principal peak at 681 ± 2 cm^{-1} due to C2′-endo/anti dG and a weaker shoulder at 668 ± 2 cm^{-1} due to C2′-endo/anti dT (Benevides et al., 1986; Thomas & Benevides, 1985). The same is observed for the d-(CGCAAATTTGCG) solution (Figure 3). Such a band profile is generally regarded as an unambiguous fingerprint of B-DNA nucleoside conformations, i.e., C2′-endo/anti dG and dT conformers (Thomas, 1986; Nishimura & Tsuboi, 1986). *However, in the d(CGCAAATTTGCG) crystal the band in question is much broader, and its peak appears at a significantly lower frequency (675 cm^{-1}) than that observed for the solution.*

The difference spectrum of Figure 2a clearly shows that the change in band shape consists of a shift of Raman intensity from 681 cm^{-1} in the solution to about 670 cm^{-1} in the crystal. Model compound studies have shown that a band shift of this kind can occur from either of two types of discrete conformation change in the deoxyguanosine residue: (i) when the sugar pucker of dG is altered from C2′-endo to C1′-exo (Nishimura & Tsuboi, 1987) or (ii) when the glycosyl torsion is rotated from the mid-anti range toward the syn range (Benevides et al., 1986; Nishimura et al., 1986). Thus, by either interpretation, the data indicate a more heterogeneous distribution of dG nucleoside conformers in the crystal structure than in the solution structure. The amplitude of the difference band suggests that about one-third of the dG residues exhibit the altered conformation in the crystal. We may distinguish between which of the two alternatives noted above, i.e., change in sugar pucker or change in glycosyl torsion, is the more likely by considering the following. There is no evidence for syn or very high *anti*-glycosyl torsions of dG residues from the present level of refinement of the X-ray data (Coll et al., 1987). Neither are syn dG conformers indicated by conformation-sensitive Raman bands in other regions of the spectrum. We propose, therefore, that one-third of the dG residues of the dodecamer crystal exhibit sugar pucker outside the normal C2′-endo range that prevails for dG residues of B-DNA in solution. We emphasize that the Raman data indicate the altered dG conformation in the crystal to be in the C1′-exo range, which is generally considered within the C2′-endo family. Supporting evidence for this interpretation is provided by the appearance of a Raman doublet at 1333 and 1341 cm^{-1}, characteristic of a mixture of C2′-endo and C1′-exo dG in the crystal, in lieu of the sharp singlet at 1339 cm^{-1} for the solution spectrum (Figure 2b).

The preceding discussion has focused on conformation-sensitive Raman bands of the sugar–phosphate backbone and dG nucleosides in the region 600–900 cm^{-1}, because spectra–structure correlations for these are the best understood (Thomas, 1986; Thomas & Wang, 1988). However, it is clear from Figures 1c and 2b that a large number of other Raman bands are also sensitive to the crystal/solution structure change of the B-DNA dodecamer. It is improbable that the nucleoside conformational changes are restricted exclusively to deoxyguanosine residues. Unfortunately, a basis does not yet exist for reliable interpretation of the rich pattern of difference bands in the 1200–1400-cm^{-1} region (Figure 2b). In addition,

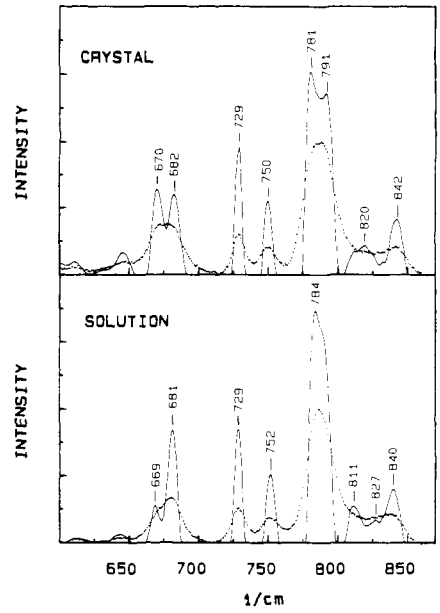

FIGURE 4: (Top) High-resolution spectrum (4-cm^{-1} slit width) of the dodecamer crystal (···) compared with its Fourier deconvolution (—). (Bottom) Spectrum of the dodecamer solution (···) compared with its Fourier deconvolution (—). A Gauss–Lorentz deblurring function of 20-cm^{-1} half-width was employed in the constrained, iterative Fourier deconvolution procedure (Thomas & Agard, 1984).

the extensive overlap of multiple bands in the 1200–1400-cm^{-1} region severely complicates the potential usefulness of the data, even if spectra–structure correlations for Raman bands of individual nucleosides could be established. The latter limitation may be partly overcome by the use of curve decomposition methods, such as Fourier deconvolution as next illustrated, to enhance the resolution of overlapping bands.

Enhanced Resolution of d(CGCAAATTTGCG) Raman Spectra. Figure 4 (top) shows an improved spectrum of the dodecamer crystal (600–900-cm^{-1} interval), achieved by narrowing the spectral band-pass to 4 cm^{-1} and signal averaging the data collected. With this approach, it is possible to just begin resolving the doublet near 665–685 cm^{-1} into two discrete components. Application of Fourier deconvolution to this partially resolved spectrum achieves further band separation, with peaks of nearly equal intensity discriminated at 670 and 682 cm^{-1}. By contrast, Fourier deconvolution of the solution spectrum separates peaks of greatly different intensities (Figure 4, bottom). Since in both cases the lower frequency member of the doublet contains a contribution from C2′-endo/anti dT residues, it is the change in *relative* intensity that accurately reflects the shift of the dG conformation marker. The deconvolution results of Figure 4 confirm the conclusion reached earlier by difference spectroscopy that approximately 33% of the Raman intensity, or one-third of the dG residues per dodecamer, is shifted from the C2′-endo/anti dG marker position in the solution spectrum (681 cm^{-1}) to the C1′-exo/anti position in the crystal spectrum (670 cm^{-1}). The same conclusion is reached by application of a nonlinear least-squares curve-fitting method (Thomas & Wang, 1988) to find the best fit of the experimental data of Figure 4 to Lorentzian peaks centered at the positions of the partially resolved components.

Estimation of the Water Content of the Dodecamer Crystal. Prescott et al. (1984) have shown that the relative Raman intensities of bands due to nucleotide CH stretching (2800–3100 cm^{-1}) and water OH stretching vibrations (3000–3700 cm^{-1}) provide a reliable basis for estimating the weight percent compositions of DNA and water in solutions

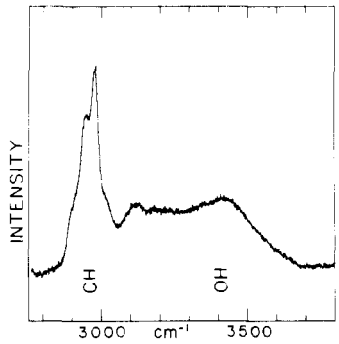

FIGURE 5: Raman spectrum of the dodecamer crystal in the region 2750–3800 cm^{-1}, containing bands due to DNA residues (base and sugar CH stretching vibrations) and water of crystallization (OH stretching vibrations). The data are corrected for Raman scattering of the mother liquor as described (Benevides et al., 1985). These results should be compared with corresponding data obtained from DNA fibers at controlled relative humidity (Prescott et al., 1984).

or fibers (Prescott et al., 1984). Here we use the Raman spectrum to approximate the water content in the crystal of the B-DNA dodecamer. The appropriate Raman data on the dodecamer crystal are shown in Figure 5, and the necessary reference data are given elsewhere [Figure 1 of Prescott et al. (1984)].

The ratio of the peak heights of the DNA and water bands in Figure 5 is 2.0. The reference value of the Raman peak ratio for DNA at 75% relative humidity (RH) is 2.7. Accordingly, we find the water content of the crystal to be approximately 2.0/2.7 = 0.74 ± 0.12 times that of DNA at 75% RH. Falk et al. (1962) have accurately determined by BET gravimetric analysis that DNA contains 36% by weight H_2O at 75% RH, which corresponds to seven water molecules per nucleotide. Thus, we calculate that the crystal contains approximately 26 ± 5% by weight H_2O, or roughly five water molecules per nucleotide. The present estimate is subject to a rather large uncertainty, mainly because of arbitrariness in establishing base lines for peak measurements. However, a similar value for the weight percent H_2O in the crystal is reached from the X-ray-determined unit cell parameters (Coll et al., 1987).

Conclusions

The principal conclusion of the present work is that the differences between Raman spectra of crystal and solution forms of d(CGCAAATTTGCG) reflect differences in structural details of the dodecamer. In crystal and solution, the Raman bands sensitive to *backbone* phosphodiester geometry, though not identical, fall into the ranges expected of B-DNA structures (Arnott et al., 1976), from which we have concluded that in both cases the dodecamer forms a B-DNA secondary structure; yet, the structures are not identical. Similarly, Raman bands sensitive to dG nucleoside pucker are significantly altered in the crystal spectrum from their positions in the solution spectrum, from which we have concluded that in the crystal the dodecamer double helix contains as many as two of its six dG residues in C1′-exo-related conformations not present in solution. The Raman spectra, though sensitive to dT pucker, give no indication of other than C2′-endo/anti dT in either solution or crystal. For dC and dA residues, C2′-endo/anti conformations are also indicated in the solution structure, but no definitive conclusions can be drawn about possible differences in these puckers in the crystal structure. However, the large number of unassigned bands in the difference spectrum suggests that dC and/or dA residues of the crystal structure, like dG, may differ from the C2′-endo/anti puckers observed in the solution structure. Thus, we view the dodecamer as a relatively uniform B-DNA double helix (all C2′-endo) in solution but as a more irregular (heterogeneous) B-DNA double helix in the crystal.

Table II: Selected Raman Line Frequencies Diagnostic of A-, B-, and Z-DNA Structures[a]

residue	A-DNA	B-DNA	Z-DNA
G	664 ± 2[b]	682 ± 2	625 ± 3
	1318 ± 2	1333 ± 3[c]	1316 ± 2
A			624 ± 3
	1335 ± 2	1339 ± 2	1310 ± 5
C	780 ± 2	782 ± 2	784 ± 2[d]
	1252 ± 2	1255 ± 5	1265 ± 2
T	745 ± 2	748 ± 2	
	777 ± 2	790 ± 3	
	1239 ± 2	1208 ± 2	
backbone	705 ± 2	790 ± 5	745 ± 3
	807 ± 3[e]	834 ± 7[f]	
	1099 ± 1	1092 ± 1	1095 ± 2
	1418 ± 2	1422 ± 2	1425 ± 2

[a] Frequencies in cm^{-1} units are determined from Raman spectra of DNA and RNA crystals and fibers of known structure. [This work and references cited in Thomas (1986) and Thomas et al. (1986).] [b] This line is usually observed at 668 ± 1 cm^{-1} in structures containing rG. [c] A weak companion line near 1316 cm^{-1} is also observed in B-DNA structures. [d] The virtual invariance of this frequency to structure change makes it of limited usefulness. [e] This line occurs at 813 ± 3 cm^{-1} in A-RNA structures. A very weak line also occurs at ca. 810 cm^{-1} in Z-DNA structures. [f] The line is centered close to 830 or 840 cm^{-1}, for DNA containing GC or AT, respectively.

Table II contains a partial list of B-DNA Raman markers for comparison with Raman bands recognized previously as indicators of Z-DNA and A-DNA. The Raman marker of the B-form phosphodiester backbone of d-(CGCAAATTTGCG), at 841 cm^{-1} in the crystal or at 835 cm^{-1} in the solution, and markers of the C2′-endo/anti dG and dT nucleosides, at 681 and 668 cm^{-1}, respectively, in both crystal and solution, are likely to be the most useful for secondary structure analyses. [Table II excludes many of the conformation-sensitive bands of A-, B-, and Z-DNA outside the 600–900-cm^{-1} interval. A more comprehensive tabulation will appear in a future paper (Wang, Benevides, and Thomas, unpublished results).]

Previous studies of oligonucleotide and polynucleotide solutions have shown that the B-DNA backbone marker (OPO symmetric stretching vibration) usually occurs in the interval 825–845 cm^{-1}. However, the precise position of the band within this interval depends upon the base sequence. The pattern that has emerged from the numerous examples studied [reviewed in Thomas (1986)] is the following: In structures containing only AT base pairs, the marker is located at 841 ± 2 cm^{-1}. In structures containing only GC base pairs, the marker is located at 829 ± 2 cm^{-1}. In structures containing both AT and GC pairs in comparable numbers, the marker occurs at ca. 835 cm^{-1}. In the last category are found calf thymus DNA, P22 DNA, and other DNAs of mixed-base composition. Although two discrete components to the 835-cm^{-1} band cannot ordinarily be resolved by experimental means, Fourier deconvolution can be employed to successfully resolve two overlapping components assignable to the AT and GC domains (Thomas et al., 1986). These results indicate at least two distinguishable OPO geometries for B-DNA, one specific to AT sequences and the other specific to GC sequences. By contrast, the corresponding A-DNA (807 ± 3 cm^{-1}) and Z-DNA (745 ± 3 cm^{-1}) markers are invariant to the base composition. The precise configurational features of the OPO group which account for two regimes of the B-DNA marker band are not known, but evidently reside in the values of backbone (α, ζ) torsion angles.

Table III: Average Torsion Angles and Deviations for AT and GC Residues in Two B-DNA Dodecamer Crystals

	deg			
	d(CGCAAATTTGCG)[b]		d(CGCGAATTCGCG)[c]	
torsion angle[a]	GC	AT	GC	AT
α	−59.1 ± 60.2	−39.2 ± 23.9	−65.6 ± 7.5	−57.9 ± 8.1
β	164.9 ± 20.1	−167.5 ± 27.2	168.9 ± 13.8	175.8 ± 14.4
γ	58.8 ± 60.2	19.6 ± 29.9	53.4 ± 8.4	56.4 ± 6.3
δ	129.1 ± 30.3	144.0 ± 26.3	124.4 ± 23.5	119.6 ± 14.7
ϵ	−177.2 ± 28.6	−185.9 ± 14.1	−160.3 ± 28.4	−184.0 ± 3.2
ζ	−100.9 ± 50.0	−90.1 ± 22.1	−117.9 ± 40.7	−92.0 ± 5.1
χ	−111.8 ± 22.4	−104.2 ± 13.1	−118.8 ± 14.1	−120.8 ± 9.1

[a] Dihedral torsion angles are defined as follows (Saenger, 1984):

$$O3'-P\overset{\alpha}{-}O5'\overset{\beta}{-}C5'\overset{\gamma}{-}C4'\overset{\beta}{-}C3'\overset{\epsilon}{-}O3'\overset{\zeta}{-}P-O5'$$

[b] Coll et al., 1987. [c] Dickerson & Drew, 1981.

The dodecamer examined here contains equal percentages of AT and GC pairs and is therefore expected to exhibit, like other DNA of equal base composition, a broad Raman band near 835 cm^{-1} as the OPO marker of B-form backbone geometry. In the dodecamer crystal we find instead that the OPO marker is centered at 841 cm^{-1}, which is the position expected for DNA containing only AT pairs. This suggests that the OPO geometry specific to the central AAATTT sequence dominates the crystal structure or, stated otherwise, that the flanking CGC and GCG sequences contain such a broad distribution of OPO geometries that there is insufficient cumulative intensity at 829 cm^{-1} to produce a discrete Raman peak characteristic of the GC domains. This explanation is consistent with the preliminary X-ray results for the OPO torsion angles (α, ζ), summarized in Table III (Coll et al., 1987). The X-ray data indicate a much broader range of torsion angles for the flanking GC domains than for the central AAATTT domain. In the solution structure, on the other hand, the OPO marker occurs at 835 cm^{-1}, which is the value expected for DNA containing comparable numbers of AT and GC pairs. Apparently, the solution structure contains greater consistency in the OPO torsion angles of the flanking GC segments, resulting in an accumulation of Raman intensity ca. 829 cm^{-1}, which in turn combines with the AT contribution ca. 841 cm^{-1} to generate a resultant band that is peaked at the mean value of 835 cm^{-1}. The interpretation given here is well supported by the results of Fourier deconvolution of the Raman band envelope in the 800–850-cm^{-1} interval, as shown in Figure 4. These results further support the general notion of greater heterogeneity in the dodecamer crystal structure than in the solution structure.

The Raman spectra reported here do not contain features that can yet be interpreted as markers of curved DNA, even though the dodecamer contains a central sequence (AAATTT) considered to have a higher than average probability of inducing DNA curvature (Trifonov, 1984). On the other hand, the strong influence of the central six AT pairs on the position of the Raman backbone marker in the crystal (841 cm^{-1} as noted above) indicates that in this flexible double helix the topological properties are dictated by the same portion of the sequence which also is potentially curved. Certainly, the present results are not inconsistent with the possibility of a different degree of curvature in the crystal and solution structures. Further study of model DNA sequences known to exhibit different degrees of curvature should reveal whether discrete Raman bands are sensitive to this conformational property.

The manipulation of crystal and solution spectra to generate a difference spectrum (Figures 1 and 2) is somewhat arbitrary. Thus, the validity of the interpretation developed here rests upon the accuracy of our assumption that the sum of all difference bands should be a minimum. This assumption is equivalent to assuming that both the solution and crystal samples contain only one molecular structure. There is little doubt that this holds true for the solution. However, in the case of the crystal, ambiguity exists as to whether the observed structural heterogeneity actually arises from within a given dodecamer (as we have assumed) or between nonequivalent dodecamers in different unit cells of the lattice. Neither Raman nor X-ray methods can definitively resolve this point. The Raman spectrum, on the other hand, can resolve the question of whether the observed heterogeneity is static or dynamic, i.e., a consequence of heterogeneity in molecular conformation or a consequence of rapid interconversion between conformers. We may reject the latter alternative on the grounds that a dynamic basis for the observed effects would lead to broadening of all Raman bands in the spectrum, not just the few actually noted (Figure 2a).

Finally, we mention that the present results allow us to predict some features of the Raman spectrum of the Dickerson dodecamer crystal d(CGCGAATTCGCG). We expect that the Dickerson structure would contain its OPO marker frequency below 841 cm^{-1} by virtue of its lesser AT content and the lesser structural influence anticipated from its central AATT domain upon backbone OPO geometry. In solution, we would expect the OPO marker band to be more nearly the same as in the crystal structure. In fact, the significant shift of the OPO marker observed for our dodecamer would not be expected to occur for the Dickerson dodecamer. The complex band near 665–685 cm^{-1}, which is anomalously broadened in the present crystal, should not exhibit the same breadth in the Dickerson crystal, because the sugar puckers and glycosyl torsions span a greater range in the present case (Table III) than in the Dickerson model (Dickerson & Drew, 1981). Again, this would render more similar the spectra of crystal and solution forms of the Dickerson dodecamer. In effect, we expect that it would be difficult to detect significant structural differences by comparing the crystal and solution Raman spectra of the Dickerson dodecamer.

References

Arnott, S., Smith, P. J. C., & Chandrasekaran, R. (1976) in *CRC Handbook of Biochemistry and Molecular Biology*, 3rd ed. (Fasman, G. D., Ed.) Vol. 2, pp 100–110, Chemical Rubber Co., Cleveland, OH.

Benevides, J. M., & Thomas, G. J., Jr. (1983) *Nucleic Acids Res. 11*, 5747–5761.

Benevides, J. M., Wang, A. H.-J., van der Marel, G. A., van Boom, J. H., Rich, A., & Thomas, G. J., Jr. (1984) *Nucleic Acids Res. 12*, 5913–5925.

Benevides, J. M., Wang, A. H.-J., Rich, A., Kyogoku, Y., van der Marel, G. A., van Boom, J. H., & Thomas, G. J., Jr. (1986) *Biochemistry 25*, 41-50.
Coll, M., Frederick, C. A., Wang, A. H.-J., & Rich, A. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 8385-8389.
Dickerson, R. E., & Drew, H. R. (1981) *J. Mol. Biol. 149*, 761-786.
Erfurth, S. C., Kiser, E. J., & Peticolas, W. L. (1972) *Proc. Natl. Acad. Sci. U.S.A. 69*, 938-941.
Falk, M., Hartman, K. A., & Lord, R. C. (1962) *J. Am. Chem. Soc. 84*, 3843-3846.
McCall, M., Brown, T., & Kennard, O. (1985) *J. Mol. Biol. 183*, 385-396.
Nishimura, Y., & Tsuboi, M. (1986) in *Spectroscopy of Biological Systems* (Clark, R. J. H., & Hester, R. E., Eds.) Vol. 13, pp 177-232, Wiley, New York.
Nishimura, Y., Tsuboi, M., Sato, T., & Aoki, K. (1986) *J. Mol. Struct. 146*, 123-153.
Prescott, B., Steinmetz, W., & Thomas, G. J., Jr. (1984) *Biopolymers 23*, 235-256.
Prescott, B., Benevides, J. M., Weiss, M. A., & Thomas, G. J., Jr. (1986) *Spectrochim. Acta, Part A 42A*, 223-226.
Saenger, W. (1984) in *Principles of Nucleic Acid Structure* (Cantor, C. R., Ed.) pp 51-104, Springer-Verlag, New York.
Shakked, Z., Rabinovich, D., Kennard, O., Cruse, W. B. T., Salisbury, S. A., & Viswamitra, M. A. (1983) *J. Mol. Biol. 166*, 183-201.
Thomas, G. J., Jr. (1986) in *Spectroscopy of Biological Systems* (Clark, R. J. H., & Hester, R. E., Eds.) Vol. 13, pp 233-309, Wiley, New York.
Thomas, G. J., Jr., & Agard, D. A. (1984) *Biophys. J. 46*, 763-768.
Thomas, G. J., Jr., & Benevides, J. M. (1985) *Biopolymers 24*, 1101-1105.
Thomas, G. J., Jr., & Wang, A. H.-J. (1988) in *Nucleic Acids and Molecular Biology* (Eckstein, F., & Lilley, D. M. J., Eds.) Vol. 2, Springer-Verlag, West Berlin.
Thomas, G. J., Jr., Benevides, J. M., & Prescott, B. (1986) *Biomol. Stereodyn. 4*, 227-253.
Trifonov, E. N. (1984) *CRC Crit. Rev. Biochem. 19*, 89-106.
Van der Marel, G. A., van Boeckel, C. A. A., Wille, G., & van Boom, J. H. (1981) *Tetrahedron Lett.*, 3887.
Wang, A. H.-J., & Rich, A. (1985) in *Biological Macromolecules and Assemblies* (Jurnak, F. A., & MacPherson, A., Eds.) Vol. 2, pp 127-170, Wiley, New York.
Wang, A. H.-J., Quigley, G. J., Kolpak, F. J., Crawford, J. L., Van Boom, J. H., van der Marel, G., & Rich, A. (1979) *Nature (London) 282*, 680-686.
Wang, A. H.-J., Fujii, S., van Boom, J. H., & Rich, A. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 3968-3972.
Wing, R., Drew, H., Takano, T., Broka, C., Tanaka, S., Itakura, K., & Dickerson, R. E. (1980) *Nature (London) 287*, 755-758.

A Novel Actin Label: A Fluorescent Probe at Glutamine-41 and Its Consequences†

Reiji Takashi

Cardiovascular Research Institute, University of California, San Francisco, San Francisco, California 94143

Received July 28, 1987; Revised Manuscript Received October 15, 1987

ABSTRACT: By peptide isolation and analysis, it has been shown that the dansyl fluorophore of dansylcadaverine [*N*-(5-aminopentyl)-5-(dimethylamino)naphthalene-1-sulfonamide] transfers to Gln-41 of actin from rabbit skeletal muscle when the reaction is catalyzed by guinea pig liver transglutaminase. As a function of time, the degree of labeling asymptotically approaches 1 mol of dansyl/1 mol of actin. About 80-85% of the attached dansyl fluorophore was found at Gln-41. Such labeled G-actin polymerizes to the same extent as control actin, but the polymerization rate is greater and the critical concentration is less than for control actin. Complete polymerization is accompanied by a 1.5-2.0-fold increase in the emission intensity of the attached fluorophore. Labeled F-actin thus obtained activates myosin subfragment 1 (S-1) Mg^{2+}-ATPase activity with the same K_{app}, and to the same V_{max}, as control actin; moreover, when such labeled F-actin is cross-linked to S-1 by 1-ethyl-3-[3-(dimethylamino)propyl]carbodiimide, the resulting superactivation of Mg^{2+}-ATPase is the same as that attained with control actin. The attributes of this label thus make it an ideal reporter of events in the N-terminal 10-kilodalton region of actin, and a new topological point for proximity mapping.

Actin is one of the most important proteins of eukaryotic cells; it is essential in cytoskeletal phenomena and in muscle cells; it is, along with myosin, a partner in the contractile event. Domain structure has been ascribed to actin in two senses: Crystallographic studies (Kabsch et al., 1985) show that actin monomer consists of two somewhat unequal lobes; on the other hand, proteases of different specificity all cut initially between two quite unequal fragments of actin (Jacobson & Rosenbusch, 1976; Mornet & Ue, 1984; Konno, 1987). For example, Konno (1988) found chymotrypsin to cleave between Met-44 and Val-45, thus generating isolatable, N-terminal "10-kilo-

† This research was supported by Grants HL-16683 (National Heart, Lung, and Blood Institute) and NSF INT-8414375.

0006-2960/88/0427-0938$01.50/0