Kan, L.-S., Barrett, J. C., Miller, P. S., & Ts'o, P. O. P. (1973) *Biopolymers 12*, 2225–2240.
Kan, L.-S., Cheng, D. M., Jayaraman, K., Leutzinger, E. E., Miller, P. S., & Ts'o, P. O. P. (1982) *Biochemistry 21*, 6723–6732.
Kan, L.-S., Chandrasegaran, S., Pulford, S. M., & Miller, P. S. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 4263–4265.
Kan, L.-S., Cheng, D. M., Chandrasegaran, S., Pramanik, P., & Miller, P. S. (1987) *J. Biomol. Struct. Dyn. 4*, 785–796.
Kearns, D. R. (1984) *CRC Crit. Rev. Biochem. 15*, 237–290.
Kearns, D. R., Patel, D. J., & Shulman, R. G. (1971) *Nature (London) 229*, 338–339.
Miller, P. S., Braiterman, L. T., & Ts'o, P. O. P. (1977) *Biochemistry 16*, 1988–1996.
Miller, P. S., Chandrasegaran, S., Dow, D. L., Pulford, S. M., & Kan, L.-S. (1982) *Biochemistry 21*, 5468–5474.
Patel, D. J., Pardi, A., & Itakura, K. (1982) *Science (Washington, D.C.) 216*, 581–590.
Pless, R. C., & Ts'o, P. O. P. (1977) *Biochemistry 16*, 1239–1250.
Redfield, A. G., Kunz, S. D., & Ralph, E. K. (1975) *J. Magn. Reson. 19*, 114–117.
Singer, B., & Grunberger, D. (1983) in *Molecular Biology of Mutagens and Carcinogens*, Plenum Press, New York.
Summers, M. F., Powell, C., Egan W., Byrd, R. A., Wilson, W. D., & Zon, G. (1986) *Nucleic Acids Res. 14*, 7421–7436.

Assignment of Phosphorus-31 and Nonexchangeable Proton Resonances in a Symmetrical 14 Base Pair *lac* Pseudooperator DNA Fragment†

Stephen A. Schroeder, Josepha M. Fu, Claude R. Jones, and David G. Gorenstein*
Department of Chemistry, Purdue University, West Lafayette, Indiana 47907
Received October 10, 1986; Revised Manuscript Received January 14, 1987

ABSTRACT: The ^{31}P chemical shifts of all 13 phosphates and the chemical shifts of nearly all of the nonexchangeable protons of a symmetrical 14 base pair *lac* pseudooperator DNA fragment have been assigned by regiospecific labeling with oxygen-17 and two-dimensional NMR techniques. At 22 °C, 8 of the 13 phosphorus resonances can distinctly be resolved while the remaining 5 resonances occur in two separate overlapping regions. The ^{31}P chemical shifts of this particular 14 base pair oligonucleotide do not follow the general observation that the more internal the phosphate is located within the oligonucleotide sequence the more upfield the ^{31}P resonance occurs, as shown from other ^{31}P assignment studies. Failure of this general rule is believed to be a result of helical distortions that occur along the oligonucleotide double helix, on the basis of the analysis of Callidine [Callidine, C. R. (1982) *J. Mol. Biol. 161*, 343–352]. Notable exceptions to the phosphate position relationship are 5′-Py-Pu-3′ dinucleotide sequences, which resonate at a lower field strength than expected in agreement with similar results as reported by Ott and Eckstein [Ott, J., & Eckstein, F. (1985) *Biochemistry 24*, 253]. A reasonable correlation exists between ^{31}P chemical shift values of the 14-mer and the helical twist sum function of Calladine. The most unusual ^{31}P resonance occurs most upfield in the ^{31}P spectrum, which has been assigned to the second phosphate position (5′-GpT-3′) from the 5′ end. This unusual chemical shift may be the result of the predicted large helical twist angle that occurs at this position in the 14-mer sequence. Further, it is believed that the large helical twist represents a unique structural feature responsible for optimum binding contact between *lac* repressor protein and this 14-mer *lac* pseudooperator segment. Assignments of proton resonances were made from two-dimensional ^{1}H–^{1}H nuclear Overhauser effect (NOESY) connectivities in a sequential manner applicable to right-handed B-DNA, in conjunction with two-dimensional homonuclear and heteronuclear *J*-correlated spectroscopies (^{1}H–^{1}H COSY and ^{31}P–^{1}H HETCOR). Most nonexchangeable base proton and deoxyribose proton (except for some unresolved H4′, H5′, and H5″ protons) resonances were assigned.

Since the development of the classical model for gene regulation by Jacob and Monod (1961), the regulation of gene expression in the lactose (*lac*) operon has served as the archetypal example of a negatively controlled operon in prokaryotes (Gilbert et al., 1975). The binding of the operator DNA prevents transcription of the corresponding mRNA, thus preventing expression. It has been found that the *lac* repressor binds to this specific, operator section of DNA, consisting of approximately 20 base pairs, about 10 million times more tightly than to the rest of the DNA and diffuses to the operator region about 1000 times more rapidly than that by simple diffusion (Berg et al., 1982). Although this interaction has been extensively studied [see Caruthers (1980) and Takeda et al. (1983) for additional references], it is not well understood at a molecular level.

The *lac* operator possesses an approximate twofold axis of symmetry, allowing two subunits of tetrameric *lac* repressor to simultaneously bind to the operator (Piltz et al., 1980). Proton NMR studies in conjunction with related DNA-binding protein X-ray models (Adler et al., 1972; Wade-Jardetzky et al., 1979; Buck et al., 1983; Takeda et al., 1983; Arndt et al., 1981; Zuiderweg et al., 1985a,b) have suggested that various residues in a recognition α-helix of one subunit of the repressor interact in the vicinity of base pairs AT-4, CG-5, and AT-6

†Supported by the National Institutes of Health (GM36281). Support of the Purdue Biochemical NMR facility by NIH (RR01077) is acknowledged.

0006-2960/87/0426-3812$01.50/0

Chart I: Numbering Scheme of Natural Operator and Symmetric Pseudooperator

```
Natural operator   -5        1        5        10       15       20

5'T-G-T-G-T-G-G-A-A-T-T-G-T-G-A-G-C-G-G-A-T-A-A-C-A-A-T-T-T-C-A-C-A
  A-C-A-C-A-C-C-T-T-A-A-C-A-C-T-C-G-C-C-T-A-T-T-G-T-T-A-A-A-G-T-G-T

Symmetric operator -5        1        5        10       15       20       25

5'T-G-T-G-T-G-G-A-A-T╫T-G-T-G-A-G-C---G-C-T-C-A-C-A╫A-T-T-C-C-A-C-A
  A-C-A-C-A-C-C-T-T-A╫A-C-A-C-T-C-G---C-G-A-G-T-G-T╫T-A-A-G-G-T-G-T
```

in a nonintercalative mode (see Chart I). A major portion of the specific and nonspecific binding of *lac* repressor and other DNA binding proteins to DNA is believed to involve phosphates along the DNA backbone (Takeda et al., 1983).

Our initial study, as described in this paper, focuses on a symmetrical 14 base pair analogue of the natural *lac* operator (Simons et al., 1984). The symmetric pseudooperator shown in Chart I differs from the natural operator sequence by a deletion of one base pair at its center and by base pair changes at positions 13 and 15. These sequence alterations resulted in a tenfold increase in affinity for the *lac* repressor (Simons et al., 1984).

Recent analysis of the crystal structures of oligonucleotides has revealed potentially significant sequence-dependent perturbations in the B-DNA double-helix geometry (Calladine, 1982; Dickerson, 1983). As described in this paper, ^{31}P chemical shifts appear to be sensitive to these sequence-dependent helix distortions, apparently as a result of local variations in the phosphate ester conformations in the sugar–phosphate backbone, and thus could serve as a useful probe of protein–DNA recognition.

Nuclear magnetic resonance spectroscopy is now able to provide detailed three-dimensional structures and dynamics of oligonucleotide duplexes and nucleic acid complexes (Kearns, 1984; James, 1984). Unfortunately, of the six torsional angles that largely define the backbone structure, only the four involving the deoxyribose ring have been shown to be directly amenable to analysis by NMR techniques. In fact for modest-sized oligonucleotide duplexes, determination of these torsional angles through measurement of three-bond coupling constants is very difficult. Using the time development of cross-peak intensities in 2D-NOESY NMR spectra, a number of workers (Hare et al., 1983; Feigon et al., 1983a,b; Frechet et al., 1983; Pardi et al., 1983; Wemmer et al., 1983; Broido et al., 1984; Scheek et al., 1984; Hilbers et al., 1985) have been able to qualitatively or semiquantitatively determine whether two protons are within 3–5 Å of each other (Kumar et al., 1979). From these NOESY-derived distances and distance geometry (Havel et al., 1983) calculations, models for these duplexes may be determined. While base pairing and stacking interactions can often be adequately demonstrated, detailed conformational information, especially with regard to the deoxyribose phosphate backbone, is often lacking.

We have proposed that ^{31}P NMR spectroscopy is capable of providing information on the most important remaining two torsional angles involving the phosphate ester bonds [see also Patel (1976, 1979)]. Our studies [see Gorenstein (1978, 1981, 1983, 1984) and Gorenstein et al. (1976a–c, 1981, 1982a,b)] indicated that a phosphate diester monoanion in a gauche,-gauche (*g,g*) conformation should have a ^{31}P chemical shift 1.5–2.5 ppm upfield from an ester in a non-*g,g* conformation.

In this paper, ^{31}P NMR and two-dimensional ^{1}H NMR methods, COSY and NOESY in particular, have been used to examine the structure of the symmetrical 14 base pair *lac* pseudooperator segment. This duplex represents the first structure of this size for which both the ^{31}P NMR signals and most of the nonexchangeable ^{1}H signals have been assigned.

Experimental Procedures

Synthesis and Sample Preparation. The self-complementary 14-mer duplex d(TGTGAGCGCTCACA)$_2$ (**1**) and dinucleotide monophosphates were synthesized by a manual modification of the phosphite triester method on a solid support (Shah et al., 1984; Gorenstein et al., 1984; Caruthers, 1980). We have introduced the ^{17}O label into the phosphoryl group by replacing the I_2/H_2O in the oxidation step of the phosphite by $I_2/H_2{}^{17}O$ (Monsanto $H_2{}^{17}O$, 51.5% ^{17}O, 35.6% ^{18}O, 12.9% ^{16}O) as previously described [Shah et al., 1984; Gorenstein et al., 1984; see also Petersheim et al. (1984) and Ott and Eckstein (1985)]. After cleavage from the support and deprotection, the resulting oligonucleotides were purified by reversed-phase HPLC as previously described (Gorenstein et al., 1984) followed by dialysis. The sample used in the COSY experiment was prepared by dissolving approximately 800 OD units of 14-mer in 400 μL of D_2O buffer containing 150 mM KCl, 0.1 mM NaN_3, and 55 mM potassium phosphate at pH* 7.5 (uncorrected pH meter reading). The 14-mer was then passed through a cation-exchange column (Dowex 50-X8, K^+ form), dialyzed, lyophilized, and redissolved in 400 μL of D_2O buffer prior to the NOESY experiment. The sample used in the ^{31}P–^{1}H HETCOR experiment contained approximately 640 OD units of 14-mer dissolved in 0.5 mL of D_2O containing 150 mM KCl and 1 mM NaN_3 at pH* 7.4.

NMR. The COSY and NOESY spectra were recorded on a Nicolet NT-470 (470-MHz ^{1}H) spectrometer at ambient temperature. The COSY spectrum was acquired with 2048 points in the t_2 dimension and 512 t_1 increments. The data were apodized with an unshifted sine bell before Fourier transformation in both dimensions. The spectrum is presented in the absolute value mode in Figure 1. The phase-sensitive TPPI (Marion & Wuthrich, 1983) NOESY spectrum was acquired with 2048 points in the t_2 dimension and 512 t_1 increments. Mixing times of 100–400 ms were used. The 400-ms mixing time NOESY data were treated with a Gaussian multiplication in the t_1 and t_2 dimensions, and the resulting spectrum is presented in Figure 2.

The ^{31}P–^{1}H HETCOR experiment was performed on a Varian XL-200A (200-MHz ^{1}H) spectrometer at ambient temperature. The data set was acquired with 256 points in the ^{31}P dimension and 45 t_1 increments and then zero filled to 512 × 256. The data set was multiplied by a combination of an increasing exponential and a Gaussian function before Fourier transformation in both dimensions. The resulting spectrum is presented in Figure 3. ^{1}H NMR resonances are referenced to 3-(trimethylsilyl)propanesulfonic acid sodium salt (DSS), and ^{31}P NMR resonances are referenced to an external sample of trimethyl phosphate (TMP).

Results

^{31}P Resonance Assignments of 14-mer. The assignment of the resonances in the 14-mer ^{31}P spectrum is based upon an oxygen isotopic labeling method (Petersheim et al., 1984; Shah

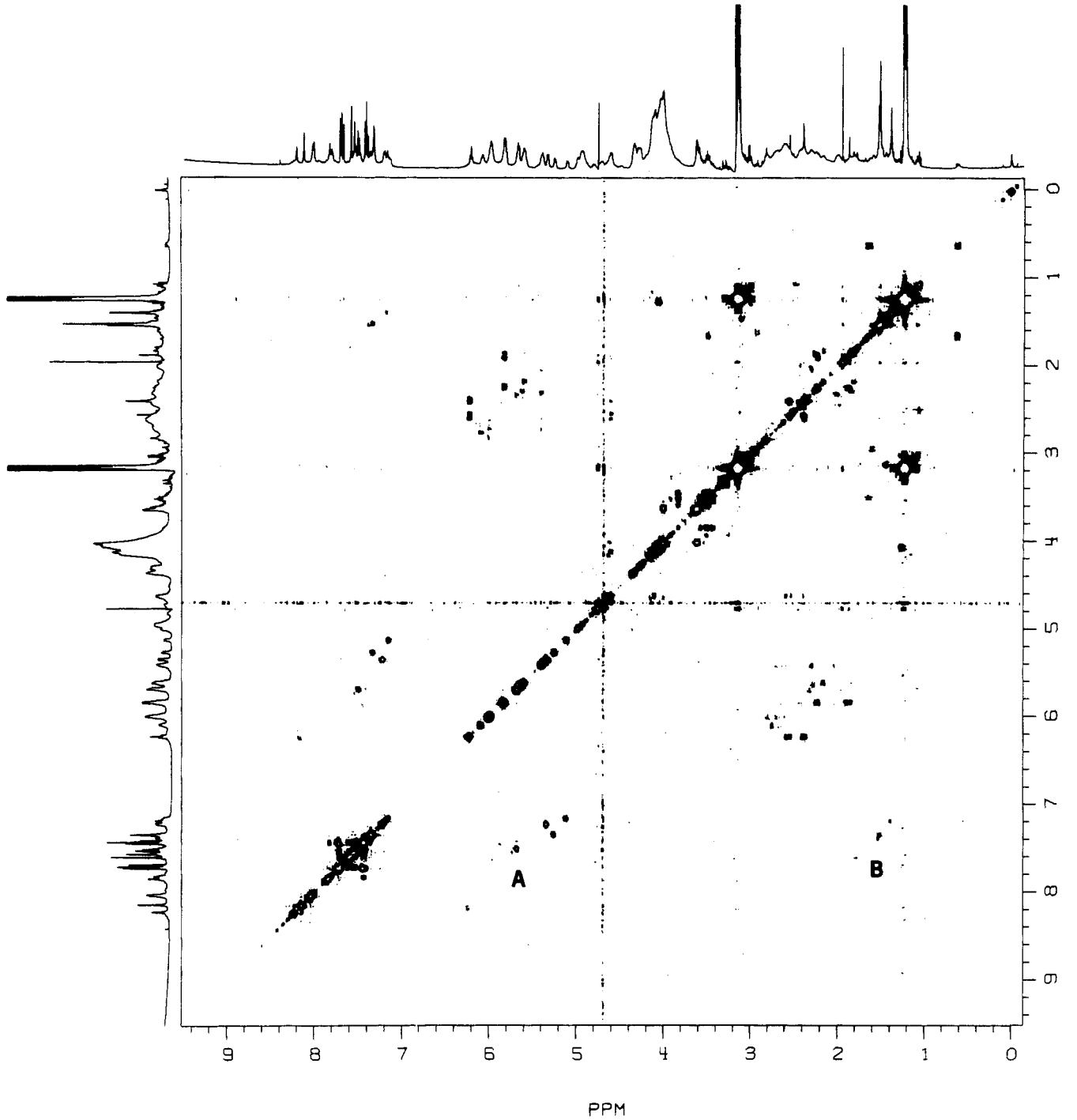

FIGURE 1: Absolute value ^1H–^1H COSY NMR spectrum of duplex 14-mer, d(TGTGAGCGCTCACA)$_2$, at 470 MHz.

et al., 1984; Ott & Eckstein, 1985). The essence of the method is to replace each of the phosphoryl oxygens with an ^{17}O label in a regiospecific manner. The interpretation of the ^{31}P spectra is based on the fact that since ^{17}O is a quadrupolar nucleus and directly bonded to the phosphorus atom, the corresponding phosphorus resonances in the ^{31}P spectrum are broadened to an extent that they disappear from the spectrum. However, one observes only a reduction in resonance intensity of about one-half, since the available ^{17}O-enriched water contains approximately 50% actual $H_2{}^{17}O$.

Figure 4 shows representative ^{31}P spectra of unlabeled and singly ^{17}O-labeled 14-mers. Spectrum A is a spectrum of a 14-mer without any ^{17}O label that can be used as a comparison between labeled spectra. As can be seen in spectra B–D, a decrease in intensity of a single resonance is observed. Eight resonances can clearly be distinguished, each integrating for one phosphorus resonance. Two regions exist where peak overlapping occurs. The most intense region in the middle corresponds to three unresolvable resonances, while the region just upfield from this position has two overlapping peaks. It is interesting to note that the resonance of the labeled phosphate is at times observed as two reduced-intensity resolved peaks associated with ^{16}O (unlabeled) and ^{18}O-labeled phosphorus resonances. This can most easily be seen in Figure 4D where the larger ^{18}O phosphorus resonance is shifted slightly upfield relative to the remaining ^{16}O phosphorus resonance [see Gorenstein et al. (1984)].

In Table I the ^{31}P chemical shift assignments for the 13 phosphates are listed. In addition, ^{31}P chemical shift values of dinucleoside phosphates as monomeric units are listed ad-

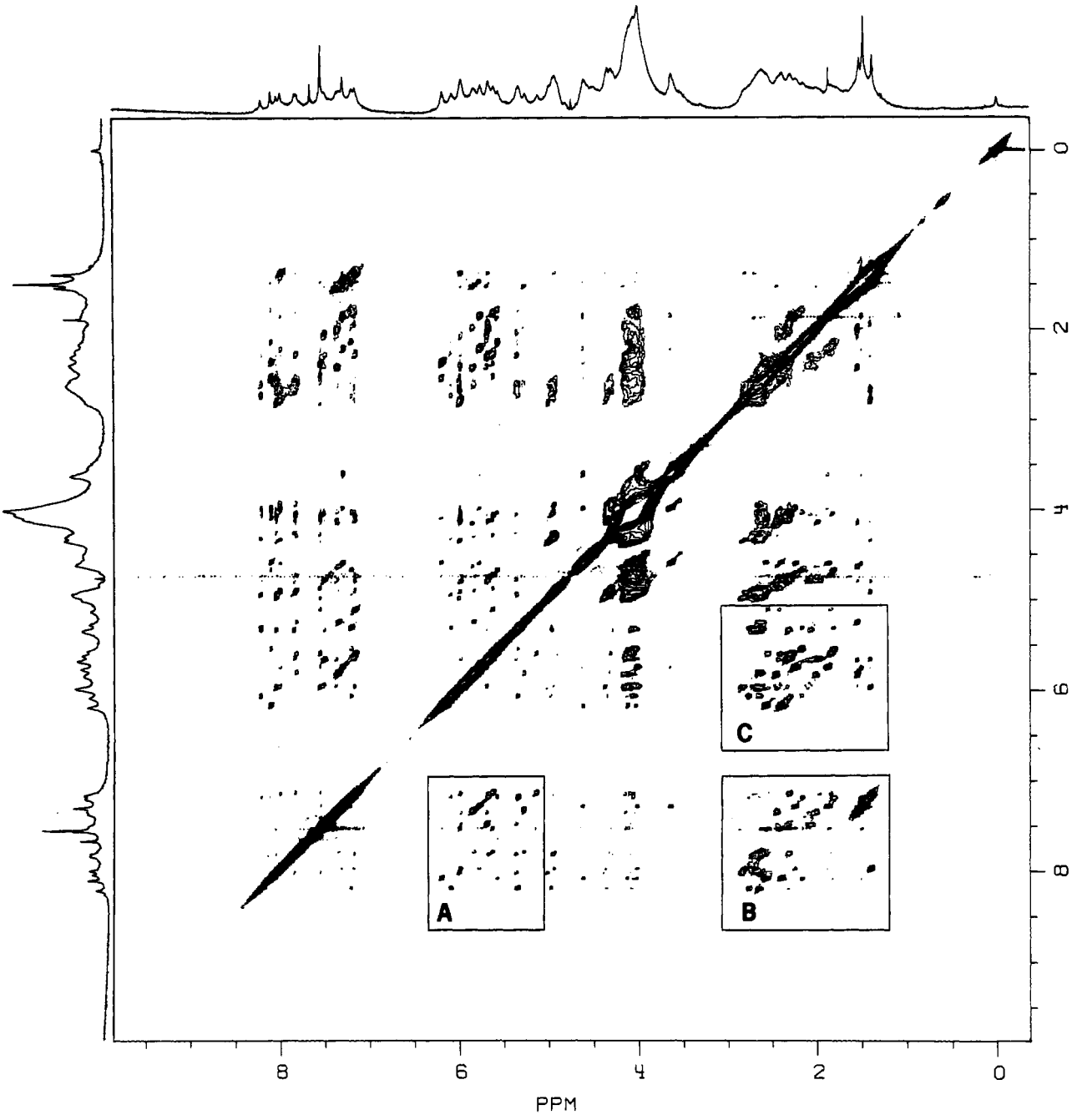

FIGURE 2: Pure absorption phase ^1H–^1H NOESY NMR spectrum of duplex 14-mer, at 470 MHz. Regions labeled A–C are expanded in Figures 6–8.

jacent to the corresponding dinucleoside phosphate in the 14-mer sequence. Since chemical shift values were determined under identical conditions, the values for the dinucleoside phosphates can be used as a comparison between identical dinucleoside sequences in the 14-mer without influence of position or neighboring base effects. The temperature dependence of the ^{31}P chemical shifts is shown in Figure 5. As expected (Gorenstein, 1981, 1984), all resonances shift downfield with increasing temperature.

Assignment of the proton signals of the 14 base pair pseudo *lac* operator duplex was accomplished through analysis of the two-dimensional COSY and NOESY NMR spectra following a sequential assignment methodology (Feigon et al., 1983a,b; Hare et al., 1983; Broido et al., 1984; Scheek et al., 1984). In the COSY spectrum (Figure 1), scalar couplings between protons are manifested as off-diagonal cross-peaks. Each of the four cytosines, in one of the symmetrical halves of the spectrum, gives rise to a cross-peak representing the H5–H6 coupling (region A). Similarly, the long range H5–CH_3 couplings in the three thymines give rise to the cross-peaks in region B. No cross-peaks from the purine bases were present since adenine and guanine do not possess groups of coupled nonexchangeable protons. The various couplings among the deoxyribose protons, i.e., H1′–H2′,2″, H2′,2″–H3′, H3′–H4′, and H4′–5′,5″, may be traced through their COSY connectivities.

While the COSY spectrum can be used to assign the protons on a particular base or sugar, it does not provide any information on the relative position of the base or sugar in the 14-mer sequence. This information was obtained, however, through analysis of the cross-relaxation networks delineated in the NOESY spectrum (Figure 2). Because the NOESY experiment (Kumar et al., 1979) utilizes through-space connectivities rather than through-bond connectivities, correlations between base and sugar protons on neighboring residues can be seen. This NOE information, taken together with the type

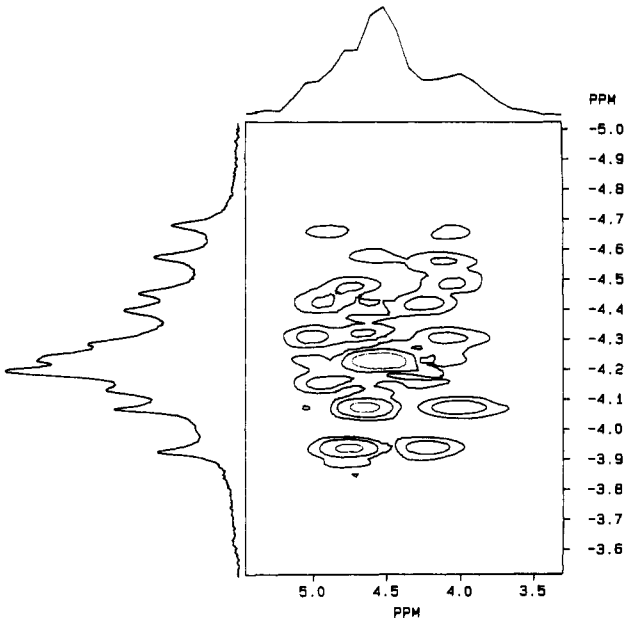

FIGURE 3: Two-dimensional ^{31}P–^{1}H HETCOR NMR spectrum of duplex 14-mer at 200 MHz (^{1}H).

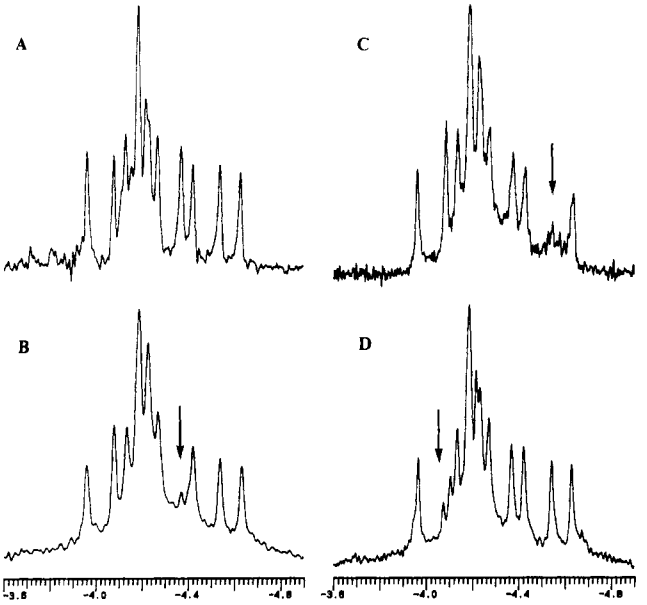

FIGURE 4: ^{31}P NMR spectra of three ^{17}O-labeled 14-mer samples and unlabeled 14-mer: (A) unlabeled 14-mer; (B) labeled at position 5; (C) labeled at position 9; (D) labeled at position 13. Samples contained 4–6 mg of purified 14-mer in 0.4 mL of D_2O containing 25 mM Hepes, 10 mM EDTA, 75 mM KCl, and 0.1 mM NaN_3, pH* 8.0. NMR parameters were as follows: sweep width 172 Hz; acquisition time 2.98 s; block size 1K zero filled to 16K; pulse width 7.0 μs; spectra resolution enhanced with a combination of positive exponential and Gaussian apodization functions; number of acquisitions between 2000 and 3000.

of base assignments from the COSY spectrum and compared to the known sequence of the 14-mer, permitted assignment of nearly all the protons in the *lac* pseudooperator segment.

In B-DNA, each pyrimidine H6 or purine H8 base proton is spatially situated so as to give rise to an NOE correlation with the H1′ sugar proton of the same nucleotide as well as with the H1′ sugar proton of the adjacent nucleotide on the 5′ side (region A). The base at the 5′ end of the chain can be identified by its lack of an NOE cross-peak to a 5′ neighboring sugar H1′ proton, and the sugar H1′ proton at the 3′ end can be identified by its lack of an NOE cross-peak to a 3′ neighboring base proton. Using these as starting points,

Table I: ^{31}P Chemical Shifts of Dinucleoside Phosphates in d(TGTGAGCGCTCACA)$_2$ and as Monomeric Units

dinucleoside phosphate (ppm)	chemical shift in 14-mer (ppm)[a,b]	chemical shift in monomeric unit (ppm)[a]	chemical shift in polymer unit (ppm)[c]
d(TpG)	−4.23 (1)	−4.15	
	−4.19 (3)		
d(CpA)	−3.97 (11)	−4.02	
	−4.09 (13)		
d(GpC)	−4.19 (6)	−3.91	−4.26 (a)
	−4.19 (8)		
d(ApG)	−4.37 (5)	−4.22	
d(GpT)	−4.63 (2)	−4.11	−4.36 (b)
d(GpA)	−4.14 (4)	−4.16	
d(CpG)	−4.27 (7)	−4.03	−4.26 (a)
d(CpT)	−4.54 (9)	−3.99	
d(TpC)	−4.42 (10)	−3.87	
d(ApC)	−4.23 (12)	−3.94	−4.36 (b)

[a] In 25 mM Hepes, 10 mM EDTA, 75 mM KCl, and 0.1 mM NaN_3, pH* 8.0 at 23 °C. [b] Numbers in parentheses indicate phosphorus position (from 5′ end). [c] Chemical shift values for polymeric units were taken from Ott and Eckstein (1985) and were obtained from the following polynucleotides: (a) poly[d(G-C)]; (b) poly[d(A-C),d-(G-T)].

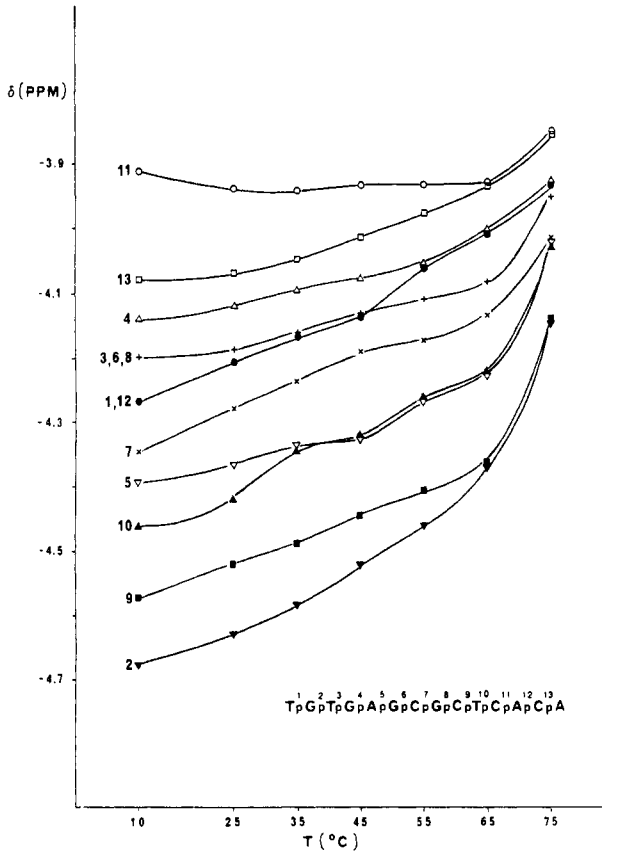

FIGURE 5: Temperature dependence of ^{31}P chemical shifts of d(TGTGAGCGCTCACA)$_2$. Spectra were recorded on a Nicolet NT-470 operating at 191 MHz. The sample was dissolved in D_2O buffer (see legend, Figure 4) and contained approximately 100 A_{260} units.

it was possible to step through the entire helix via the sugar–base–sugar connectivities, as diagramed in Figure 6. The assignments of the base and sugar protons are listed in Table II. This assignment procedure was straightforward except for the sugar H1′ protons of G6 and C7, which possess identical chemical shifts.

The thus-assigned sugar H1′ protons are, in turn, correlated to the H2′ and H2″ protons of the same nucleotide (Figure 7), allowing these to be assigned. For several nucleotides,

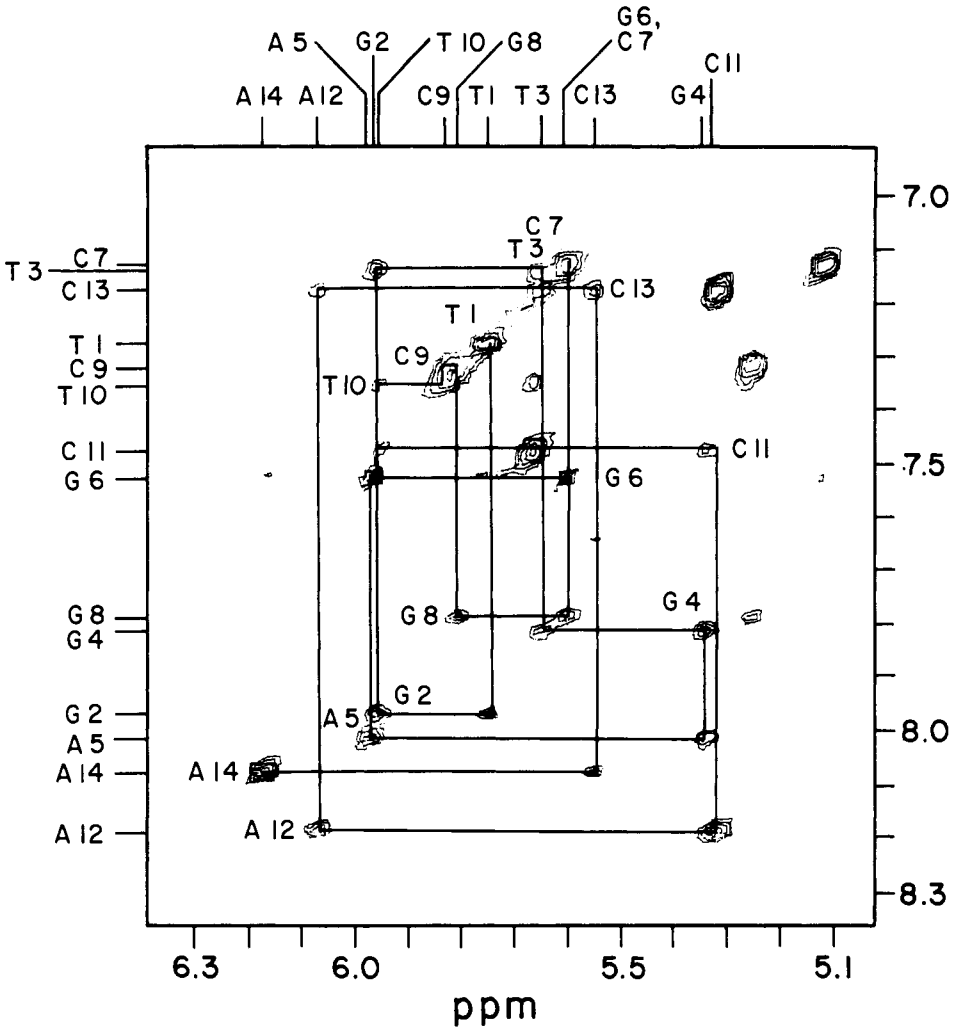

FIGURE 6: Expansion of region A of the NOESY spectrum shown in Figure 2. The sequential assignment of the base and deoxyribose H1′ protons is diagramed.

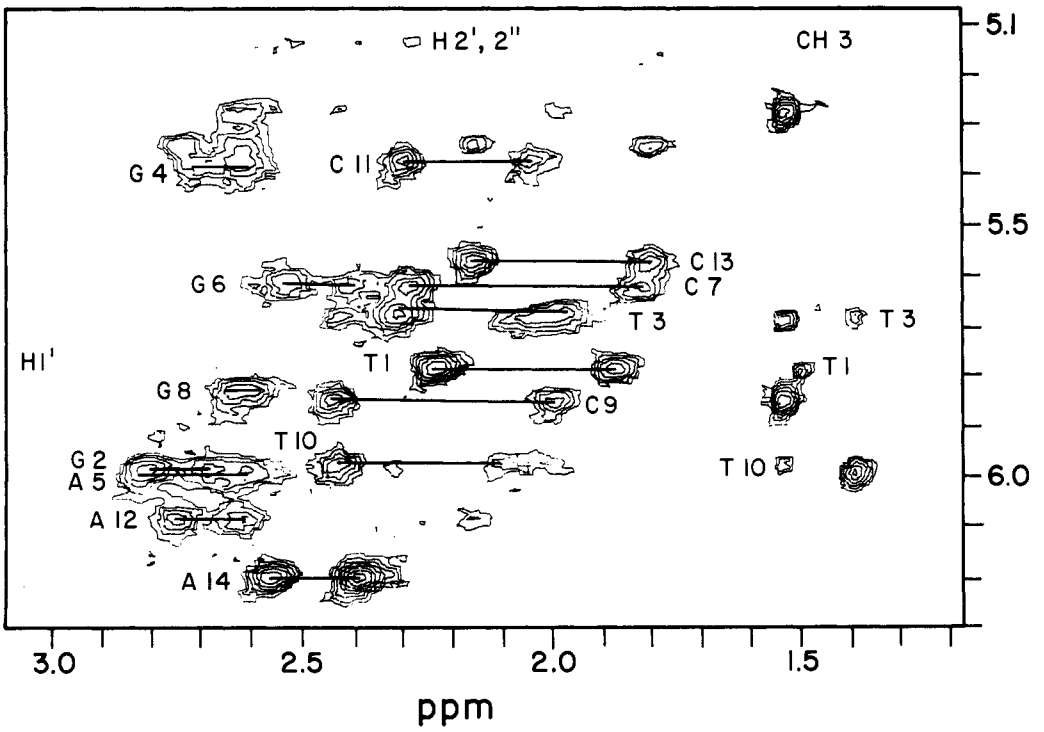

FIGURE 7: Expansion of region C of the NOESY spectrum shown in Figure 2. NOE correlations between the deoxyribose H1′ protons and the deoxyribose H2′,2″ protons, as well as thymine methyl protons, are represented by the cross-peaks in this region. The horizontal line segments connect the H2′ and H2″ protons of the same residue.

Table II: Chemical Shifts (in ppm) of Assigned ^{1}H Resonances of d(TGTGAGCGCTCACA)$_2$

residue	H8 or H6	CH_5 or TCH_3	H1′	H2′	H2″	H3′	H4′	H5′, H5″[a]
T1	7.33	1.50	5.76	1.88	2.24	4.63	4.00	3.63
G2	8.01		5.97	2.69	2.81	4.95	4.35	4.00
T3	7.19	1.40	5.66	1.98	2.30	4.82	4.15	4.02[b]
G4	7.86		5.36	2.62	2.72	4.96	4.28	4.07
A5	8.06		5.98	2.62	2.84	5.01	4.37	4.07
G6	7.57		5.61	2.41	2.54	4.91	4.30	4.15[b]
C7	7.20	5.12	5.61	1.83	2.29	4.77	4.30	NA
G8	7.83		5.81	2.61	2.66	4.93	4.31	4.23
C9	7.36	5.26	5.84	2.00	2.45	4.66	4.17	4.05, 4.16
T10	7.40	1.55	5.97	2.11	2.44	4.82	4.04	4.13
C11	7.52	5.67	5.35	2.06	2.31	4.82	4.08	4.03, 4.07
A12	8.23		6.07	2.63	2.76	4.98	4.34	4.06, 4.17[b]
C13	7.23	5.33	5.56	1.82	2.17	4.70	4.06	NA
A14	8.12		6.18	2.38	2.57	4.61	4.11	4.02

[a] NA, not assigned. [b] Assignment obtained from HETCOR spectrum. Error in measurement approximately ±0.1 ppm.

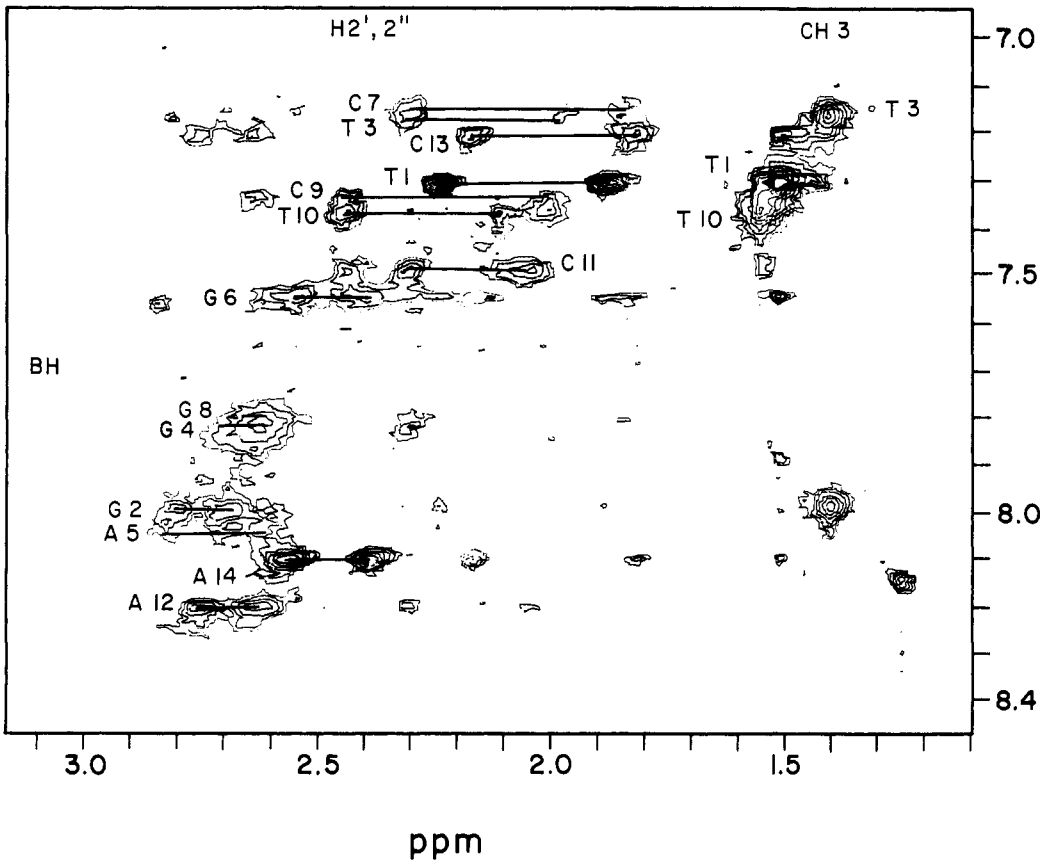

FIGURE 8: Expansion of region B of the NOESY spectrum shown in Figure 2. NOE correlations between the base protons and the deoxyribose H2′ and H2″ protons are indicated by cross-peaks in this region. The horizontal line segments connect the H2′ and H2″ protons of the same residue. Correlations between base protons and thymine methyl protons are also seen.

nearly degenerate H1′ proton chemical shifts did not permit unambiguous determination of the H2′ and H2″ assignments. Fortunately, redundancies exist in the spectrum that facilitated these assignments. Each previously assigned base proton possesses an NOE correlation to the H2′ and H2″ protons of the sugar in the same nucleotide as well as to the H2′ and H2″ protons of the sugar of the 5′ adjacent nucleotide (Figure 8).

Discrimination between the H2′ and H2″ protons was achieved by analysis of cross-peak intensities. In B-DNA, the sugar H1′ proton is nearer in space to the H2″ proton than the H2′ proton, thereby leading to a more intense NOESY cross-peak for the former interaction (Feigon et al., 1983a,b; Scheek et al., 1984). In addition, the base proton to sugar H2″ distance is significantly larger than the base proton to sugar H2′ distance, resulting in a more intense cross-peak in the latter case.

Once the assignments of the sugar H1′ and H2′,2″ and of the base protons were known, the H3′ protons were identified through their NOE correlations to these protons. The H4′ protons were assigned through their connectivities with H1′. Some of the H5′ and H5″ protons were assigned as a group, but complete assignment was not possible because of inadequate resolution in the NOESY and COSY spectra. However, additional assignments of the H5′ and H5″ protons and verification of the H3′ assignments were achieved through analysis of the two-dimensional ^{31}P–^{1}H heteronuclear shift correlation (HETCOR) NMR spectrum (Figure 3). Three- and four-bond scalar couplings between the ^{31}P nuclei in the phosphate backbone and the H3′, H4′, H5′, and H5″ deoxyribose protons are manifested as cross-peaks in this spectrum. From the ^{31}P resonances that were identified by ^{17}O labeling, we were able to assign several additional H5′ and H5″ protons to particular

residues via the HETCOR correlations (see Table II).

Discussion

Using the solid-phase phosphite triester method (Caruthers, 1980), we have synthesized the symmetrical tetradecamer **1**. We have introduced ^{17}O labels into the phosphoryl groups by replacing the I_2/H_2O in the oxidation step of the phosphite by $I_2/H_2{}^{17}O$ (Gorenstein et al., 1984a,b; Ott & Eckstein, 1985). By synthesizing the corresponding mono[^{17}O]-phosphoryl-labeled oligonucleotide (each phosphate is separately substituted along the chain), we can identify the ^{31}P signal of that phosphate diester (Petersheim et al., 1984; Shah et al., 1984; Ott & Eckstein, 1985). The quadrupolar ^{17}O nucleus (generally 40–50% enriched) broadens the ^{31}P signal of the directly attached phosphorus so that only the high-resolution signal of the remaining 50–60% nonquadrupolar broadened phosphorus is observed (Tsai, 1984; Petersheim et al., 1984; Gorenstein et al., 1984; Shah et al., 1984). In this way each synthesized oligonucleotide with a different mono-substituted [^{17}O]phosphoryl group allows identification of all phosphate ^{31}P signals (Figure 4).

Since the phosphate residue to which each ^{31}P signal is associated was identified by ^{17}O labeling, assignment of the ^1H signals coupled to each phosphorus atom was made possible. The two-dimensional ^{31}P–^1H chemical shift correlated (HETCOR) spectrum of **1** provides identification of the 3′- and 5′-deoxyribose protons coupled to the identified ^{31}P signals of the phosphates (H3′, H4′, H5′, and H5″ via ^{31}P–^1H heteronuclear COSY). Recent advances in NOE techniques, particularly two-dimensional NOESY (Feigon et al., 1983; Hare et al., 1983; Broido et al., 1984; Scheek et al., 1984), have shown that NOEs between the sugar and stacked base pairs can allow the identification of the base proton and most of the sugar proton signals. The relative intensities of the two-dimensional NOESY cross-peaks are consistent with the expected right-handed B-DNA geometry of the duplex 14-mer (Clore & Gronenborn, 1986).

^{31}P Shifts and the Calladine Rules. Nucleotides at the ends of short oligonucleotides are, in general, more conformationally flexible than those in the middle. Presumably, at room temperature some equilibrium melting at the ends allows the phosphate ester to exist in a variety of conformations. In a right-handed B-DNA duplex geometry, the P–O ester conformation (ω,ω') is g^-,g^-. In the completely denatured structure at high temperatures ($>T_m$), the oligonucleotide strands exist largely as unstacked, random coils with the phosphate ester in an increased proportion of nongauche [i.e., gauche,trans (g,t) etc.] conformations. The nongauche conformations have been predicted (Gorenstein et al., 1976a–c, 1984; Gorenstein, 1984) and are observed to be downfield of the gauche P–O ester conformation. Greater conformational freedom of the "frayed" ends at temperatures below T_m implies that the phosphate esters at the ends can exist in a number of non-g^-,g^- conformations. The ^{31}P chemical shift of these end phosphates should therefore in the absence of other effects be downfield of the middle nucleoside phosphates.

In addition, as shown in Table I, the ^{31}P chemical shift of a dinucleoside monophosphate occurs at lower field than the corresponding sequence in the 14-mer. This is observed in every case except for the phosphate at position 4. Under the experimental conditions, the dinucleoside monophosphate exists largely as flexible single-strand monomers (although stacking in a number of the dinucleoside monophosphates may be nonnegligible; Gorenstein, 1981). The phosphate is little constrained by any base stacking and pairing requirements in the double helix to the g^-,g^- conformation. Because the g^-,g^- conformation is lower in energy than nongauche conformations, the dinucleoside monophosphates, however, are still predominantly in the g^-,g^- conformation—the proportion, though, is less than in the longer duplex oligonucleotide (Gorenstein, 1981, 1984).

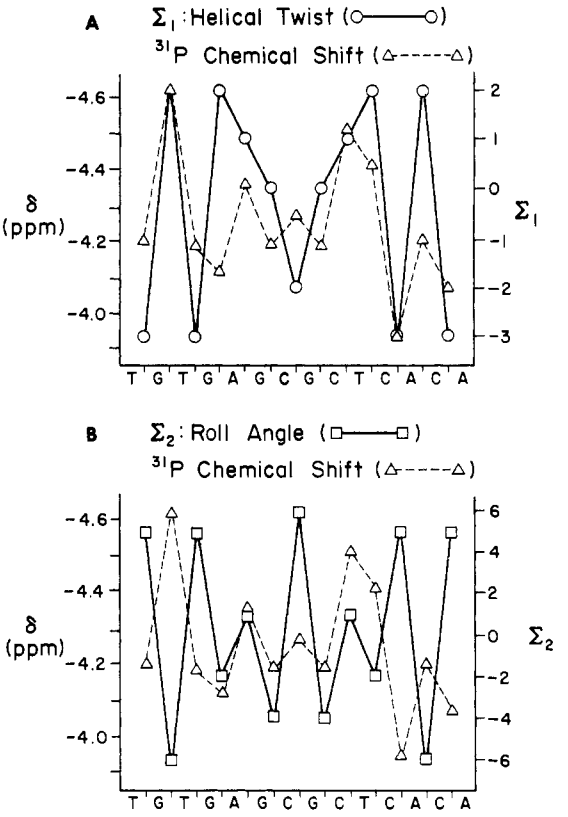

FIGURE 9: Comparison of ^{31}P chemical shifts with helix twist and roll angle sum functions as calculated from the Dickerson/Calladine sum function rules: (A) helix twist; (B) roll angle.

However, as noted by Ott and Eckstein (1985) and shown in Figures 4 and 9, numerous important exceptions to this picture are observed. A recent very exciting development is the suggestion made by Ott and Eckstein that ^{31}P chemical shifts appear to correlate with the Calladine/Dickerson rules (Calladine, 1982; Dickerson, 1983) for local helix distortion. Ott and Eckstein have noted, as discussed above, that "normal" B-DNA g^-,g^- phosphate esters are generally upfield of non-g,g phosphates. However, in purine–pyrimidine or pyrimidine–purine sequences, propeller twisting of the base pairs results in purine–purine steric clashes on opposite strands of the double helix (Calladine, 1982). In order to relieve the steric purine–purine clash, changes in the local helix twist angles and propeller twist angles (as well as several other geometric parameters) are observed. Local unwinding and changes in the torsional angles along the deoxyribose phosphate backbone thus result and may be responsible for the observed variation in the ^{31}P chemical shifts. Helix twist angles can vary by more than ±10° from the "average" B-DNA twist angle of 36° as a result of the purine–purine clash (Dickerson, 1983). These are *substantial* variations, and as shown in Figure 9, ^{31}P shifts are apparently quite sensitive to these variations. The ^{31}P NMR spectrum could therefore be viewed as a possible "fingerprint" of the operator DNA geometry.

In particular it appears that a number of chemical shift anomalies occur in the ^{31}P spectrum of the 14-mer. By far the most unusual chemical shift corresponds to the second phosphate position from the 5′ end (GpT), where it occurs most upfield in the ^{31}P spectrum. With respect to the generalized positional relationship, a much greater downfield shift would

be expected. In addition, this particular phosphorus is also observed to resonate most upfield in a similar assignment study (unpublished) where the 5′ end sequence is identical for the first four bases, that is, 5′-Py-Pu-Py-Pu-3′. This dinucleotide position occurs as a 5′-Pu-Py-3′ sequence; however, no other purine–pyrimidine sequence in the 14-mer has an unusual upfield chemical shift. We suggest that a unique structural feature that exists at or around this position in the sequence could be responsible for the higher than expected chemical shift. From the ^{31}P temperature study (Figure 5), it does not appear to be the result of any unusual thermodynamics of this phosphate. An interesting aspect related to this position in the 14-mer is that studies have indicated that repressor protein binds in the major groove near this phosphate (Takeda et al., 1983). Thus, one might suspect that a unique conformation at this site allows for optimal recognition of the operator DNA sequence by the repressor protein.

The middle phosphorus position, position 7, resonates just upfield from the main cluster region in the ^{31}P spectrum. If based solely on the phosphate position relationship, this appears to be another anomaly. However, this position in the sequence is a 5′-Py-Pu-3′ sequence and results in a more downfield ^{31}P chemical shift than expected, in agreement with results obtained by Ott and Eckstein (1985). Just as interesting, the two flanking phosphorus positions, 6 and 8, having identical base sequences (GpC), both resonate downfield at identical chemical shifts from position 7. It is interesting to note that the phosphates are identical in both sequence and position and have the same chemical shift value.

Other notable features of the ^{31}P spectrum involve the other four 5′-Py-Pu-3′ dinucleotide sequences, positions 1, 3, 11, and 13. Since the chemical shift pattern seen in the 14-mer ^{31}P spectrum does not necessarily follow the position relationship, it is difficult to ascertain the degree to which these phosphorus positions deviate from the generalized relationship. For instance, positions 1 and 3, both being 5′-TpG-3′ sequences, resonate in the middle of the ^{31}P spectrum. Thus, it is uncertain whether these are actual chemical shift anomalies based solely on being 5′-Py-Pu-3′ sequences or from other unknown factors. In contrast, both positions 11 and 13, being 5′-CpA-3′ sequences, resonate most downfield. It is interesting to note that position 11, being the most downfield resonance, occurs more downfield than position 13. In this case one would expect a reversal of chemical shift for these positions if one considers position as the only influencing factor.

Analysis of the four 5′-Py-Pu-3′ sequences further suggests an unusual structural feature within the region 5′-TpGpTpGp-3′/3′-ApCpApCp-5′. If pyridmidine–purine sequence and position are considered solely, than one would certainly expect position 1 and probably position 3 to resonate even more downfield than they actually occur, since these positions are both 3′-Py-Pu-5′ sequences and are located at the end of the 14-mer sequence.

With the discovery of sequence-dependent structural distortions that occur along the oligonucleotide chain (Calladine, 1982; Dickerson, 1983), is it likely that any correlation or relationship exists between ^{31}P chemical shifts and these local helical distortions? Ott and Eckstein (1985) have noted, on the basis of the phosphorus assignments of two dodecamers, that a correlation exists between the roll angle sum function of Calladine (1982) and the ^{31}P chemical shift. No significant correlation is noted between ^{31}P shifts and the other three Dickerson/Calladine sum functions for these dodecamers.

However, from the ^{31}P analysis of our 14-mer, there does appear to be a modest correlation with the helical twist sum function, as shown in Figure 9A. Most interesting, the best correlation points are associated with positions 2 and 10, resonances that occur most upfield and most downfield, respectively. Figure 9B is a plot of the sum function roll angle vs. chemical shift. Though the correlation is poorer than the helical twist, some correlation exists toward the middle region of the 14-mer.

Significantly, Klug and co-workers (Lomonosoff et al., 1981) have shown that the sites of DNase I catalyzed hydrolysis of DNA correlate with these Calladine/Dickerson rules. As pointed out by Dickerson (1983), because DNase I is sensitive to local geometry changes, it would be surprising if DNA binding proteins were not equally if not more sensitive to these geometric variations.

ACKNOWLEDGMENTS

We thank Dr. Robert Santini, Director of NMR Instrumentation, Purdue University, for his assistance in this work.

Registry No. 1, 107939-74-2.

REFERENCES

Adler, K., Beyreuther, K., Fanning, E., Geisler, N., Gronenborn, B., Klemm, A., Muller-Hill, B., Pfahl, M., & Schmitz, A. (1972) *Nature (London) 72*, 322–327.

Arndt, K. T., Boschelli, F., Lu, P., & Miller, J. H. (1981) *Biochemistry 20*, 6109.

Berg, O. G., Winter, R. B., & Von Hippel, P. H. (1982) *Trends Biochem. Sci. (Pers. Ed.) 7*, 52.

Broido, M. S., Zon, G., & James, T. L. (1984) *Biochem. Biophys. Res. Commun. 119*, 663–670.

Buck, F., Hahn, K.-D., Zemann, W., Ruterjans, H., Sadler, J. R., Beyreuther, K., Kaptein, R., Scheek, R., & Hull, W. (1983) *Eur. J. Biochem. 132*, 321–327.

Calladine, C. R. (1982) *J. Mol. Biol. 161*, 343–352.

Caruthers, M. H. (1980) *Acc. Chem. Res. 13*, 155–160.

Clore, G. M., & Gronenborn, A. M. (1986) *Biomolecular Stereodynamics IV, Proceedings of the Fourth Conversation in the Discipline Biomolecular Stereodynamics*, pp 139–155, Adenine, New York.

Dickerson, R. E. (1983) *J. Mol. Biol. 166*, 419–441.

Feigon, J., Leupin, W., Denny, W. A., & Kearns, D. R. (1983a) *Biochemistry 22*, 5930–5942.

Feigon, J., Leupin, W., Denny, W. A., & Kearns, D. R. (1983b) *Biochemistry 22*, 5943–5951.

Frechet, D., Cheng, D. M., Kan, L.-S., & Ts'o, P. O. P. (1983) *Biochemistry 22*, 5194–5200.

Gilbert, W., Gralla, J., Major, J., & Maxam, A. (1975) in *Protein-Ligand Interactions* (Sund, H., & Beaver, G., Eds.) pp 270–288, de Gruyter, Berlin.

Gorenstein, D. G. (1978) *P-31 NMR of Nucleic Acids, Jerusalem Symposium, NMR in Molecular Biology* (Pullman, B., Ed.) pp 1–15, Reidel, Dordrecht, The Netherlands.

Gorenstein, D. G. (1981) *Annu. Rev. Biophys. Bioeng. 10*, 355.

Gorenstein, D. G. (1983a) *Bull. Magn. Reson. 5*, 161.

Gorenstein, D. G. (1983b) *Prog. Nucl. Magn. Reson. Spectrosc. 16*, 1–98.

Gorenstein, D. G., Ed. (1984) *^{31}P NMR: Principles and Applications*, Academic, New York.

Gorenstein, D. G., Kar, D., & Momii, R. K. (1976a) *Biochem. Biophys. Res. Commun. 73*, 105.

Gorenstein, D. G., Kar, D., Luxon, B. A., & Momii, R. K. (1976b) *J. Am. Chem. Soc. 98*, 1668.

Gorenstein, D. G., Findlay, J. B., Momii, R. K., Luxon, B. A., & Kar, D. (1976c) *Biochemistry 15*, 3796.

Gorenstein, D. G., Goldfield, E. M., Chen, R., Kovar, K., & Luxon, B. A. (1981) *Biochemistry 20*, 2141.

Gorenstein, D. G., Luxon, B. A., Goldfield, E., & Vegeis, D. (1982a) *Biochemistry 21*, 580.
Gorenstein, D. G., Luxon, B. A., Goldfield, E. M., & Vegeais, D. (1982b) *Biochemistry 21*, 580.
Gorenstein, D. G., Lai, K., & Shah, D. O. (1984) *Biochemistry 23*, 6717.
Hare, D. R., Wemmer, D. E., Chou, S. H., Drobny, G., & Reid, B. (1983) *J. Mol. Biol. 171*, 319.
Havel, T. F., Kuntz, I. D., & Crippen, G. M. (1983) *Bull. Math. Biol. 45*, 665–720.
Hilbers, C. W., Haasnoot, C. A. G., de Bruin, S. H., Joordens, J. J. M., van der Marel, G. A., & van Boom, J. H. (1985) *Biochimie 67*, 685–695.
Jacob, F., & Monod, J. (1961) *J. Mol. Biol. 3*, 318–356.
James, T. L. (1984) in *Phosphorus-31 NMR: Principles and Applications* (Gorenstein, D., Ed.) pp 349–400, Academic, Orlando, FL.
Kearns, D. R. (1984) *Crit. Rev. Biochem. 15*, 237–290.
Kumar, A., Wagner, G., Ernst, R. R., & Wuthrich, K. (1979) *J. Am. Chem. Soc. 103*, 3654–3658.
Lomonosoff, G. P., Butler, P. J. G., & Klug, A. (1981) *J. Mol. Biol. 149*, 745–760.
Marion, D., & Wuthrich, K. (1983) *Biochem. Biophys. Res. Commum. 113*, 967–974.
Ott, J., & Eckstein, F. (1985) *Biochemistry 24*, 253.
Pardi, A., Walker, R., Rapoport, H., Wider, G., & Wuthrich, K. (1983) *J. Am. Chem. Soc. 105*, 1652.
Patel, D. J. (1976) *Biopolymers 15*, 533.
Patel, D. J. (1979) *Acc. Chem. Res. 12*, 118.
Petersheim, M., Mehdi, S., & Gerlt, J. A. (1984) *J. Am. Chem. Soc. 106*, 439.
Piltz, I., Goral, K., Kratky, O., Bary, R. P., Wade-Jardetzky, N., & Jardetzky, O. (1980) *Biochemistry 19*, 4087–4090.
Scheek, R. M., Boelens, R., Russo, N., Van Boom, J. H., & Kaptein, R. (1984) *Biochemistry 23*, 1371–1376.
Shah, D. O., Lai, K., & Gorenstein, D. G. (1984) *J. Am. Chem. Soc. 106*, 4302.
Simons, A., Tils, D., von Wilcken-Bergmann, B., & Muller-Hill, B. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 1624.
Takeda, Y., Ohlendorf, D. H., Anderson, W. F., & Matthews, B. W. (1983) *Science (Washington, D.C.) 221*, 1020.
Tsai, M.-D. (1984) in *^{31}P NMR: Principles and Applications* (Gorenstein, D. G., Ed.) Chapter 6, Academic, New York.
Wade-Jardetzky, N., Bray, R. P., Conover, W. W., Jardetzky, O., Geisler, N., & Weber, K. (1979) *J. Mol. Biol. 128*, 259–264.
Wemmer, D. E., Chou, S.-H., Hare, D. R., & Reid, B. R. (1983) *Biochemistry 22*, 5943–5951.
Zuiderweg, E. R. P., Boelens, R., & Kaptein, R. (1985a) *Biopolymers 24*, 601.
Zuiderweg, E. R. P., Scheek, R. M., Boelens, R., van Gunsteren, W. F., & Kaptein, R. (1985b) *Biochemie 67*, 707.

Proton NMR and Electrophoretic Studies of the Covalent Complex Formed by Cross-Linking Yeast Cytochrome *c* Peroxidase and Horse Cytochrome *c* with a Water-Soluble Carbodiimide†

Susan J. Moench,‡ James D. Satterlee,*,§,‖ and James E. Erman§
Department of Chemistry, University of New Mexico, Albuquerque, New Mexico 87131, and Department of Chemistry, Northern Illinois University, DeKalb, Illinois 60115

Received August 18, 1986; Revised Manuscript Received December 18, 1986

ABSTRACT: The 1:1 covalently cross-linked complex between horse cytochrome *c* and yeast cytochrome *c* peroxidase (ccp) has been formed by a slight modification of the method of Waldmeyer and Bosshard [Waldmeyer, B., & Bosshard, H. R. (1985) *J. Biol. Chem. 260*, 5184–5190]. This earlier study has been extended to show that efficient cross-linking of the two proteins can occur in a variety of buffers over a broad ionic strength range. The substitution of ferrocytochrome *c* for ferricytochrome *c* in the cross-linking studies resulted in an increased yield of 1:1 complex (~10–20%) under the conditions studied. An improved method for purifying the covalent complex in relatively large quantities is presented here as are the results of electrophoresis and proton NMR studies of the complex. Both electrophoresis and NMR studies indicate modification of some surface acidic amino acids in the covalent complex by the carbodiimide. The proton hyperfine-shifted resonances of cytochrome *c* are broadened in the covalent complex relative to free cytochrome *c*, and the resonances corresponding to the cytochrome *c* heme 3-CH_3 and 8-CH_3 groups are shifted closer together in the complex. Integration of NMR resonances confirms a 1:1 complex as the primary cross-linking reaction product. However, we also demonstrate that the covalent complex can be further coupled to ccp and to cytochrome *c* to form higher molecular weight aggregates.

Cytochrome *c* peroxidase (ferrocytochrome *c*:hydrogen peroxide oxidoreductase, EC 1.11.1.5; ccp)[1] from yeast catalyzes the hydrogen peroxide oxidation of ferrous cytochrome *c* (Yonetani, 1965; Conroy & Erman, 1978; Yonetani & Ray, 1966; Nicholls & Mochan, 1966; Kang et al., 1977; Kang & Erman, 1982) by a mechanism that is believed to involve specific complex formation between the two proteins (Mochan,

† Work supported in part by grants from the National Science Foundation (PCM-DMB-8403353 and CHE-8201374, J.D.S.) and the Alfred P. Sloan Foundation (J.D.S.).
* Author to whom correspondence should be addressed.
‡ University of New Mexico.
§ Northern Illinois University.
‖ Fellow of the Alfred P. Sloan Foundation.

[1] Abbreviations: ccp, cytochrome *c* peroxidase; EDC, 1-ethyl-3-[3-(dimethylamino)propyl]carbodiimide hydrochloride; SDS, sodium dodecyl sulfate; PAGE, polyacrylamide gel electrophoresis; Tris, tris(hydroxymethyl)aminomethane.

0006-2960/87/0426-3821$01.50/0