Photo-Cross-Linking of CRP to Nonspecific DNA in the Absence of cAMP. DNA Interacts with both the N- and C-Terminal Parts of the Protein

Mehrnaz Katouzian-Safadi,*,‡ Bernadette Blazy,§ Jean-Yves Cremet,‡ Jean-Pierre Le Caer,∥ Jean Rossier,∥ and Michel Charlier‡

Centre de Biophysique Moléculaire, CNRS, 1A avenue de la Recherche Scientifique, 45071 Orléans Cedex 2, France, Laboratoire de Biophysique Interactions Moléculaires, Université Paul Sabatier, 118 route de Narbonne, 31062 Toulouse Cedex, France, and Institut Alfred Fessard, CNRS, 91198 Gif sur Yvette Cedex, France

Received May 4, 1992; Revised Manuscript Received August 27, 1992

ABSTRACT: Adenosine cyclic 3′,5′-phosphate receptor protein (CRP or CAP) is a regulatory protein involved in the transcription of several operons in *Escherichia coli*. cAMP-independent, nonspecific complexes of CRP and DNA were investigated by photochemical cross-linking of the protein to nonspecific DNA, whose thymines are substituted by 5-bromouracil (BrUra). The cross-linked protein was completely digested by trypsin, and the covalently bound peptides were sequenced. We identified two regions of the protein in close contact with DNA: one in the C-terminal part, overlapping the canonical helix–turn–helix motif, and the other one in the N-terminal part, which is usually not considered to belong to the DNA-interacting domain of CRP. This result lead us to propose models for nonspecific interaction, where the DNA is in contact with both the N- and C-terminal parts of the protein.

In *Escherichia coli*, cyclic AMP receptor protein (CRP)[1] is known to regulate either positively or negatively the transcription of several operons. When complexed with cAMP, CRP undergoes a conformational transition (Krakow & Pastan, 1973; Eilen et al., 1978) which stimulates its binding to its specific DNA site located upstream of the promoter.

The structural aspect of the interaction of CRP has been recently reviewed (Steitz, 1990). The protein is a dimer (Anderson et al., 1971) of 2 chemically identical subunits of 209 amino acids (Cossart & Gicquel-Sanzey, 1982). Each subunit is able to bind one molecule of cAMP (Takahashi et al., 1980; McKay & Steitz, 1981). The CRP promoter has a two-domain structure. The larger N-terminal domain (amino acids 1–138) binds cAMP, and is responsible for the subunit–subunit interaction. The smaller C-terminal domain (amino acids 139–209) contains a helix–turn–helix motif involved in DNA binding.

The CRP protein exhibits three modes of interaction with DNA: (i) cAMP-independent, nonspecific binding; (ii) cAMP-dependent, nonspecific binding (Takahashi et al., 1979); and (iii) cAMP-dependent specific binding [see review of Steitz (1990)]. In the case of the nonspecific interaction, the binding is always cooperative. The length of the excluded site determined in these cases was 13–15 base pairs (Takahashi et al., 1979; Saxe & Revzin, 1979).

The crystal structure of the CRP–cAMP complex was solved at 2.9-Å resolution (McKay & Steitz, 1981). Recently, the 3-Å resolution crystal structure of CRP–cAMP complexed with a consensus sequence of 30 base pairs has been published (Schultz et al., 1991). The DNA is bent by about 90°, resulting from two kinks, one on each side of the dyad axis of the complex. Such a bending was previously demonstrated in solution using other techniques [for a review, see Steitz (1991) and papers cited therein]. In this case of specific binding, the interaction occurs through the helix–turn–helix motif (amino acids 168–193) of the protein.

We have previously studied the nonspecific CRP–DNA interaction by photochemical cross-linking (Katouzian-Safadi et al., 1991a). 5-Bromouracil-substituted DNA was used to increase the cross-linking efficiency, and consequently to reduce the direct photolysis of both DNA and proteins. We have shown that the rate of the reaction is not modified upon cAMP binding and that the photo-cross-linking of the CRP involves only one of the two subunits. This photo-cross-linking method has been successfully used in various DNA–protein interaction studies: histones (Weintraub, 1973), RNA polymerase (Simpson, 1980), *lac* repressor (Barbier et al., 1984; Allen et al., 1991), *Eco*RI (Wolfes et al., 1986), and archaebacterial chromosomal protein MC1 (Katouzian-Safadi et al., 1991b).

Our goal in this paper was to identify the region(s) of the protein interacting with nonspecific DNA, by sequence analysis of the peptides remaining bound to DNA after proteolytic cleavage. The involvement of the N-terminal part of the protein lead us to propose models for the nonspecific interaction.

MATERIALS AND METHODS

Preparation of 5-BrUra-Substituted DNA and CRP. BrUra-substituted DNA was purified as described previously (Katouzian-Safadi et al., 1991a) from *E. coli* B/rT⁻ strain grown in M_9 medium supplemented with 5-bromodeoxyuridine and thymidine in a molar ratio of 90:10. The percentage of substitution was determined by measuring the buoyant density of the DNA in a CsCl gradient (Hackett & Hanawalt, 1966) and found to be 75%.

Wild-type CRP was purified from an overproducing strain harboring the plasmid pBS crp2 (Cossart & Gicquel-Sanzey, 1982). The purity of the protein preparation was greater than 99% as estimated from SDS–PAGE.

Concentrations were determined spectrophotometrically, assuming extinction coefficients of 13 000 M^{-1} cm^{-1} at 260

* Address correspondence to this author.
‡ Centre de Biophysique Moléculaire, CNRS.
§ Université Paul Sabatier.
∥ Institut Alfred Fessard, CNRS.

[1] Abbreviations: cAMP, adenosine cyclic 3′,5′-phosphate; CRP, adenosine cyclic 3′,5′-phosphate receptor protein; BrUra, 5-bromouracil; PTH, phenylthiohydantoin.

0006-2960/93/0432-1770$04.00/0

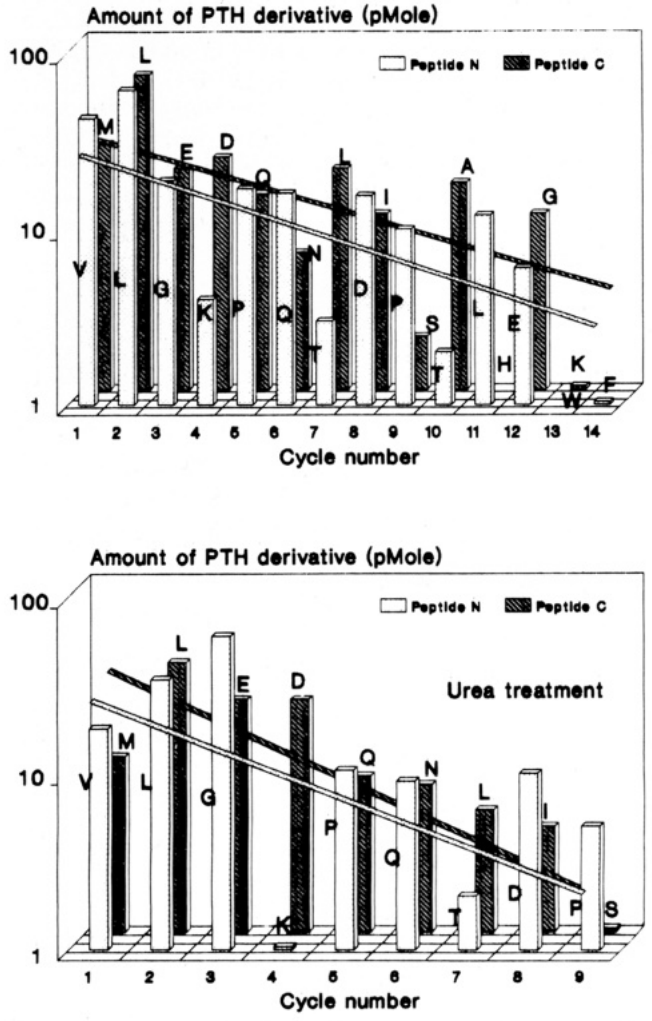

FIGURE 1: Edman degradation of cross-linked peptides generated by tryptic cleavage of protein CRP either in the absence (upper part) or in the presence (lower part) of urea. PTH derivatives of amino acids (indicated by their one-letter code) were identified and quantified by HPLC.

nm per base pair for the DNA and 19 125 M^{-1} cm^{-1} at 275 nm per protomer for the protein.

CRP–DNA Complexes. Complexes between BrUra-substituted DNA and CRP at a final concentration of 80 μM base pairs and 2 μM CRP protein, respectively (ratio of 40 base pairs per protein), were prepared by mixing the two components at high salt concentration (10 mM Tris-HCl/400 mM NaCl, pH 7.5), followed by an extensive dialysis at 4 °C against a low ionic strength buffer (10 mM Tris-HCl/60 mM NaCl, pH 7.5).

Irradiations. Irradiations (2.5-mL solution in Suprasil cuvettes of 1-cm path length) were performed at room temperature, using a 200-W mercury lamp equipped with filters isolating an irradiation band ranging from 300 to 400 nm (Katouzian-Safadi et al., 1991a). The fluence delivered was 2.6 MJ m^{-2} (fluence rate 480 W m^{-2}).

Determination of the Amount of CRP Cross-Linked to DNA. The amount of CRP dimer cross-linked to DNA was determined as previously described (Katouzian-Safadi et al., 1991a).

Tryptic Digestion of the Protein. Samples (2 × 2.5 mL) of either free or DNA-complexed protein (90 μg/mL protein in 10 mM Tris-HCl/60 mM NaCl, pH 7.5) were submitted to trypsin (EC 3.4.21.4, Boehringer) hydrolysis either in the absence or in the presence of 4.8 M urea at 25 °C for 18 h.

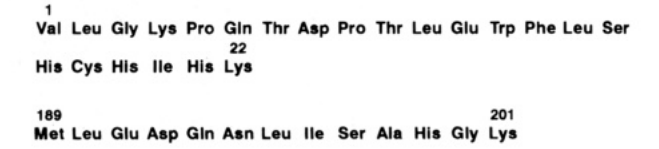

FIGURE 2: Tryptic peptides identified by microsequencing. Peptides N and C (Figure 1) can be identified as the N-terminal part of two tryptic peptides of the CRP.

The final enzyme to substrate ratio was 1/25 (w/w). Digestion of the protein was checked in both conditions by SDS–PAGE [15% acrylamide, acrylamide to bis(acrylamide) ratio of 200/1]. Total hydrolysis of CRP might lead to 23 peptides of various lengths, from 1 to 24 amino acids. Under our experimental conditions, the largest peptides obtained are smaller than 2.6 kDa, as expected from a totally achieved hydrolysis. This procedure was consequently used to digest the irradiated complex containing the cross-linked proteins.

Determination of the Cross-Linked Peptides. After trypsin treatment, the samples containing urea were dialyzed against 10 mM Tris-HCl/60 mM NaCl, pH 7.5. All samples were then adjusted to 0.5 M KCl/0.2 M potassium phosphate, pH 7.4, and chromatographed on Ultrogel ACA 34 (60 × 1.5 cm column) eluted with this high ionic strength buffer. The fractions containing the cross-linked peptides bearing DNA were pooled, concentrated, and successively dialyzed against 200 mM ammonium hydrogen carbonate and ultrapure water. From the DNA absorbance, and taking into account the cross-linking yield, we expect that the amount of peptide remaining is on the order of magnitude of 500–1000 pmol.

These samples were directly submitted to amino acid sequencing, without any digestion of DNA (Katouzian-Safadi et al., 1991b). Automated Edman degradation was carried out on a 470A gas-phase Applied Biosystems sequencer, using the 03RPE1 program, with poly(ethylenimine) (Applied Biosystems) as carrier (Le Caer & Rossier, 1988). Phenylthiohydantoin (PTH) derivatives of amino acids were identified by reversed-phase HPLC with an on-line Applied Biosystems 120 A amino acid analyzer.

RESULTS

Irradiation of nonspecific CRP–BrUra-substituted DNA complexes lead to 70% cross-linked proteins, determined by molecular sieving. We verified that there is no cross-link at high ionic strength, where the protein is not bound (Katouzian-Safadi et al., 1991a).

It has been reported that free CRP is resistant to proteolysis by various proteases, although CRP complexed with DNA can be cleaved by trypsin. In mild conditions, Angulo and Krakow (1986) obtained two peptides of 6 and 9.7 kDa. For our determinations of the cross-linked peptides, a complete hydrolysis of the CRP was necessary. Therefore, we used the procedure tested for the free and complexed protein (see Materials and Methods).

The DNA-cross-linked peptides purified on an ACA 34 gel were then submitted to microsequencing. The PTH derivatives of amino acids identified by sequencing analysis are presented in Figure 1. In both cases (hydrolysis by trypsin with and without urea), we obtained the same mixture of two peptides unambiguously identified on the CRP amino acid sequence (Figure 2). Peptide C (amino acids 189–201) belongs to the C-terminal part of the CRP, and is located near the canonical helix–turn–helix region. Peptide N (amino acids 1–14) belongs to the N-terminal part of the protein. We can remark here that peptide N is not cleaved by trypsin at Lys-4, due to the vicinity of Pro-5.

Both peptides are present in nearly equivalent amounts, and their repetitive yields are quite similar. We remark, however, that the PTH derivative of Lys-4 is lower in the urea-treated sample. As a control, a sample of CRP complexed to nonsubstituted DNA has been irradiated and carried through all steps described previously. In this case, no amino acids were detected by microsequencing.

Another control was to sequence a sample of irradiated CRP. The N-terminal sequence (amino acids 1–25) was obtained. However, the amount of PTH derivative of Trp-13 was 3 times lower than expected (as compared to the adjacent amino acids Glu-12 and Phe-14). This probably results from the photodegradation of tryptophan (Katouzian-Safadi et al., 1991a).

DISCUSSION

Protein–BrUra-substituted DNA photo-cross-linking occurs directly between a BrUra residue and a reactive group of the protein. Its occurrence needs a close contact between DNA and the complexed protein, and a propicious configuration of the two reactive groups. Among all the interaction sites, only those involving a BrUra residue and a cross-linkable amino acid residue located in the major groove are assumed to be selected. In this assumption, only the presence of cross-links can give information concerning the proximity of residues in the complex. The absence of a cross-link cannot lead to pertinent conclusions.

Cross-linking of BrUra-substituted DNA to a protein can probably occur through several pathways. Model systems using amino acids have been studied: cysteine and cystine (Varghese, 1974; Dietz & Koch, 1989); lysine (Saito & Martsuura, 1985); tyrosine, histidine, and tryptophan (Ito et al., 1980; Saito et al., 1981; Saito & Matsuura, 1985; Dietz & Koch, 1987); and even the peptide bond (Dietz et al., 1987).

Our results demonstrated that at least two cross-linking sites exist on the CRP complexed to nonspecific DNA in the absence of cAMP. The first one involves peptide C (amino acids 189–201), located in the C-terminal part of the protein, and belonging partially to the canonical helix–turn–helix motif of the CRP. The second site involves peptide N (amino acids 1–14), located in the N-terminal part of the protein. The photoreactions at the two sites are equally probable, since the amounts of both peptides are similar.

The presence of peptide N in the samples digested without urea was very surprising. In the case of proteins complexed to DNA, our hydrolysis protocol was effective. However, in the case of the cross-linked proteins, where no control was possible, a doubt was subsisting, that a class of cross-linked proteins was resistant to trypsin hydrolysis. To test such an occurrence, digestion in the presence of urea has been done. In these conditions, where the trypsin remains totally active, and the CRP is denatured, we expected a complete proteolysis. Sequencing of the cross-linked peptides after urea treatment shows that the same two peptides are present, in the same proportions, and are sequenced with the same repetitive yield (Figure 1). We concluded that the sequences obtained in the absence of urea were not artifactual.

Our previous results showed that photo-cross-linking of the protein involves only one of the two subunits at one and the same time (Katouzian-Safadi et al., 1991a). In the present case, we show the existence of two different cross-linking points on each subunit, and that both reactions are equally probable. However, we have no information about the simultaneity of two bonds on the same subunit.

Examining the amino acid sequences of the cross-linked peptides, we observed a lack of His-199 in peptide C. This strongly suggests the involvement of histidine-199 in the cross-linking reaction of this peptide. Histidine has been previously identified as a cross-linked amino acid in the case of the *lac* repressor (Barbier et al., 1984; Allen et al., 1991). For this peptide C, the sequencing stops at the tryptic site Lys-201. We remark that the sequencing can bypass the cross-linked histidine.

Sequencing peptide N, we observed a low level of Lys-4 (only for hydrolysis in the presence of urea) and the lack of Trp-13. Although Lys-4 might be responsible for the cross-link, its very low level after urea treatment could be due to the reaction of the ϵ-NH_2 with traces of cyanate present in urea.

As already mentioned, the photodegradation of tryptophan reduces the amount of the PTH derivative by 60–70%. In the case of noninvolvement of tryptophan in the cross-linking reaction, we could expect around 1–2 pmol of PTH derivative, that would be easily detected. The total absence of tryptophan at position 13 suggests its involvement in the cross-linking reaction. With the archaebacterial chromosomal protein MC1, tryptophan has already been found to be cross-linked to BrUra-substituted DNA (Katouzian-Safadi et al., 1991b).

The sequencing stops after Phe-14, beyond Trp-13, but far from the tryptic site Lys-22. The nonsequenced part of the peptide contains three histidine residues, which are also possible sites for photo-cross-linking (Barbier et al., 1984; Allen et al., 1991). The occurrence of several simultaneous cross-links on the same peptide seems unlikely, however, since it would need the presence of clusters of BrUra residues.

In the helix–turn–helix motif, the first helix (E-helix in the CRP) makes mostly nonspecific contacts with the DNA. The second helix (F-helix in the CRP) recognizes the specific sequence through major groove interaction (Harrison & Aggarwal, 1990). Peptide C (amino acids 189–201) is located in the C-terminal domain, and contains a part of the F-helix (amino acids 180–193). Moreover, we show that one of the contact points is His-199. Lee et al. (1991) showed that this His-199 residue is involved in the specific cAMP-dependent interaction with the *lac* promoter. In the model published by Schultz et al. (1991) from crystallographic data, the helix–turn–helix motif interacts with DNA in the major grove, and His-199, six residues below the end of the motif, adjoins the DNA minor grove. Nevertheless, in the absence of cAMP, and with nonspecific DNA, our result shows that His-199 is close to the C5 atom of one bromouracil, located in the major groove. This interacting C-terminal region probably plays a role, although not necessarily the same role, in all types of CRP–DNA interactions.

Our results reveal another contact point with DNA, located in peptide N, i.e., in the N-terminal part of the protein. This fact was a priori unexpected. Nevertheless, the crystal structure of the complex CRP–DNA (Schultz et al., 1991) shows that Lys-26, located in the β-2 sheet, interacts with DNA. Moreover, the large N-terminal peptide of CRP (amino acids 1–116), obtained upon subtilisin treatment, can still interact with poly[d(AT)] (Takahashi et al., 1982). These two results show that the N-terminal domain of the protein plays a role in the cAMP-independent, nonspecific DNA binding process. Our results give direct chemical evidence of this interaction.

We can propose two different models for the cAMP-independent, nonspecific interaction of CRP with long DNA, in which the two peptides N and C are in close contact with

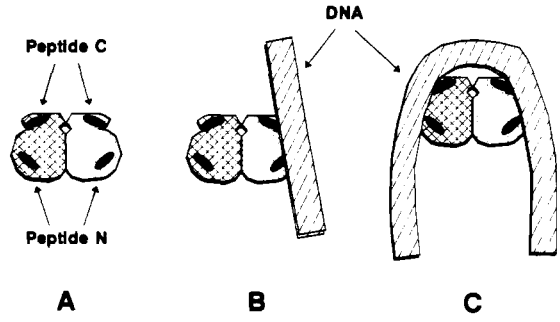

FIGURE 3: Drawing of the cAMP-independent, nonspecific interaction models of the CRP with DNA. (A) The two subunits of the CRP protein, with the marked situation of peptides C and N. (B) Unbent DNA interacts with both peptides of one subunit. (C) Bent DNA interacts with both peptides of both subunits.

the DNA. Considering the X-ray structure of the CRP (McKay & Steitz, 1981), a part of both peptides N and C belongs to a positively charged face of a subunit (His-17, His-19, His-21, Lys-22, Lys-26, Lys-166, His-199, and Lys-201) that could electrostatically interact with about 13–18 base pairs of unbent DNA (Figure 3B). The length of DNA in contact with the protein agrees with the length of the excluded site deduced from circular dichroism or electron microscopy studies (Saxe & Revzin, 1979; Takahashi et al., 1979; Chang et al., 1981). Moreover, this type of model perfectly explains that only one subunit of the CRP was photo-cross-linked to DNA, whatever the cross-linking yield (Katouzian-Safadi et al., 1991).

Another possibility is to consider that the cAMP-independent, nonspecific binding mimics the cAMP-dependent specific binding (Schultz et al., 1991): the DNA is bent in order to contact the two helix–turn–helix motifs of both subunits. As both the protein dimer and the DNA have a dyad axis, it seems tempting to build a model where the complex exhibits such a symmetry. In this case, however, the DNA must be completely bent as a hairpin to closely contact the two peptides C, but also the two peptides N (Figure 3C), and the length of the DNA in contact with the protein would be around 40 base pairs. This model could explain the 4 times packing of DNA complexed with CRP, as observed by electronic microscopy (Chang et al., 1981). In this case, the existence of several interaction points does not implicate the occurrence of several simultaneous cross-linking points, as previously discussed (Katouzian-Safadi et al., 1991).

From our results, it is not possible to choose between the two models. However that may be, both C- and N-terminal parts of the protein must be included in the structural feature of the cAMP-independent, nonspecific binding of CRP with DNA.

ACKNOWLEDGMENT

We thank Françoise Culard for helpful and stimulating discussions.

REFERENCES

Allen, T. D., Wick, K. L., & Matthews, K. S. (1991) *J. Biol. Chem. 266*, 6113–6119.
Anderson, W. B., Schneider, A. B., Emmer, M., Perlman, R. L., & Pastan, I. (1971) *J. Biol. Chem. 246*, 5929–5937.
Angulo, J., & Krakow, J. S. (1986) *J. Biol. Chem. 261*, 11315–11319.
Barbier, B., Charlier, M., & Maurizot, J.-C. (1984) *Biochemistry 23*, 2933–2939.
Chang, J.-J., Dubochet, J., Baudras, A., Blazy, B., & Takahashi, M. (1981) *J. Mol. Biol. 150*, 435–439.
Cossart, P., & Gicquel-Sanzey, B. (1982) *Nucleic Acids Res. 10*, 1363–1378.
Dietz, T. M., & Koch, T. H. (1987) *Photochem. Photobiol. 46*, 971–978.
Dietz, T. M., & Koch, T. H. (1989) *Photochem. Photobiol. 49*, 121–129.
Dietz, T. M., von Trebra, R. J., Swanson, B. J., & Koch, T. H. (1987) *J. Am. Chem. Soc. 109*, 1793–1797.
Eilen, E., Pampeno, C., & Krakow, J. S. (1978) *Biochemistry 17*, 2469–2473.
Hackett, P., & Hanawalt, P. (1966) *Biochim. Biophys. Acta 123*, 356–363.
Harrisson, S. C., & Aggarwal, A. K. (1990) *Annu. Rev. Biochem. 59*, 933–969.
Ito, S., Saito, I., & Matsuura, T. (1980) *J. Am. Chem. Soc. 102*, 7535–7541.
Katouzian-Safadi, M., Blazy, B., & Charlier, M. (1991a) *Photochem. Photobiol. 53*, 611–616.
Katouzian-Safadi, M., Laine, B., Chartier, F., Cremet, J.-Y., Belaiche, D., Sautiere, P., & Charlier, M., (1991b) *Nucleic Acids Res. 19*, 4937–4941.
Krakow, J. S., & Pastan, I. (1973) *Proc. Natl. Acad. Sci. U.S.A. 70*, 2529–2533.
Le Caer, J. P., & Rossier, J. (1988) *Anal. Biochem. 169*, 246–252.
Lee, B. J., Aiba, H., & Kyogoku, Y. (1991) *Biochemistry 30*, 9047–9054.
McKay, D. B., & Steitz, T. A. (1981) *Nature 290*, 744–749.
Saito, I., & Matsuura, T. (1985) *Acc. Chem. Res. 18*, 134–141.
Saito, I., Ito, S., Matsuura, T., & Helene, C. (1981) *Photochem. Photobiol. 33*, 15–19.
Saxe, S. A., & Revzin, A. (1979) *Biochemistry 18*, 255–263.
Schultz, S. C., Shields, G. C., & Steitz, T. A. (1991) *Science 253*, 1001–1007.
Simpson, R. B. (1980) *Nucleic Acids Res. 8*, 759–766.
Steitz, T. A. (1990) *Q. Rev. Biophys. 23*, 205–280.
Takahashi, M., Blazy, B., & Baudras, A. (1969) *Nucleic Acids Res. 7*, 1699–1712.
Takahashi, M., Blazy, B., & Baudras, A. (1980) *Biochemistry 19*, 5124–5130.
Takahashi, M., Gronenborn, A. M., Clore, G. M., Blazy, B., & Baudras, A. (1982) *FEBS Lett. 139*, 37–40.
Varghese, A. J. (1974) *Photochem. Photobiol. 20*, 461–464.
Weintraub, H. (1973) *Cold Spring Harbor Symp. Quant. Biol. 38*, 247–256.
Wolfes, H., Fliess, A., Winkler, F., & Pingoud, A. (1986) *Eur. J. Biochem. 159*, 267–273.