# Cystic Fibrosis Transmembrane Conductance Regulator: Nucleotide Binding to a Synthetic Peptide

PHILIP J. THOMAS, P. SHENBAGAMURTHI, XAVIER YSERN, PETER L. PEDERSEN*

**Multiple mutations in the gene responsible for cystic fibrosis are located within a region predicted to encode a nucleotide-binding fold in the amino terminal half of the cystic fibrosis transmembrane conductance regulator protein. A 67–amino acid peptide (P-67) that corresponds to the central region of this putative nucleotide binding site was chemically synthesized and purified. This peptide bound adenine nucleotides. The apparent dissociation constants ($K_d$'s) for the trinitrophenyl (TNP) adenine nucleotides, TNP-adenosine triphosphate, TNP-adenosine diphosphate, and TNP-adenosine monophosphate, were 300 nanomolar, 200 nanomolar, and greater than 1 micromolar, respectively. The $K_d$ for adenosine triphosphate was 300 micromolar. Circular dichroism spectroscopy was used to show that P-67 assumes a predominantly β sheet structure in solution, a finding that is consistent with secondary structure predictions. On the basis of this information, the phenylalanine at position 508, which is deleted in approximately 70 percent of individuals with cystic fibrosis, was localized to a β strand within the nucleotide binding peptide. Deletion of this residue is predicted to induce a significant structural change in the β strand and altered nucleotide binding.**

CYSTIC FIBROSIS (CF) IS AN INHERITED disorder that results in impaired epithelial ion transport. The disease is serious in both scope and outcome: more than 1 in 2000 Caucasians are affected, with death usually occurring before age 30 from damage to airway epithelia (*1*). The gene responsible for CF has been identified, cloned, and sequenced (*2*). The predicted gene product is the cystic fibrosis transmembrane conductance regulator (CFTR). This putative 1480–amino acid protein contains five regions with amino acid sequences that are similar to previously described functional domains: two nucleotide-binding domains, two transmembrane domains, and a regulatory domain (*2*). At least one of the nucleotide-binding domains may catalyze the hydrolysis of adenosine triphosphate (ATP), which may be coupled to the transport of a solute across the plasma membrane through the transmembrane domains (*2, 3*). The solute for this putative ATP-driven transport activity is unknown, but it has been suggested that it may be $Cl^-$ (*1*) or a hydrophobic regulatory molecule (*3*).

One of the predicted nucleotide-binding domains of CFTR is located in the $NH_2$-terminal half of the protein and the other is in the COOH-terminal half. Both CFTR nucleotide-binding domains include the A and B consensus sequences that are found in many other nucleotide-binding proteins (*4*): these include transport adenosine triphosphatases (ATPases), such as the multidrug resistance P-glycoprotein (MDR) (*5*), which is responsible for the development of drug resistance in mammalian cells; and the α and β subunits of the mitochondrial F-type ATP synthase (*6*), which is responsible for formation of most of the ATP required by most eukaryotic cells (Fig. 1A). In cases where the tertiary structure of nucleotide binding sites are known, the A and B consensus sequences, along with the intervening sequences, form an alternating β-α-β-α-β structure with a parallel β sheet at its core (*7*). In adenylate kinase, the A consensus sequence forms a glycine flexible loop between one of the β strands and an α helix, whereas the B consensus sequence lies at the end of a hydrophobic β strand (*4, 8*). In several nucleotide-binding proteins, the glycine flexible loop interacts with the β phosphate of ATP, whereas the core of the parallel β sheet interacts with the adenosine moiety of the nucleotide (*9, 10*).

$Phe^{508}$, which is deleted in 70% of individuals with CF (*2*), is located between the

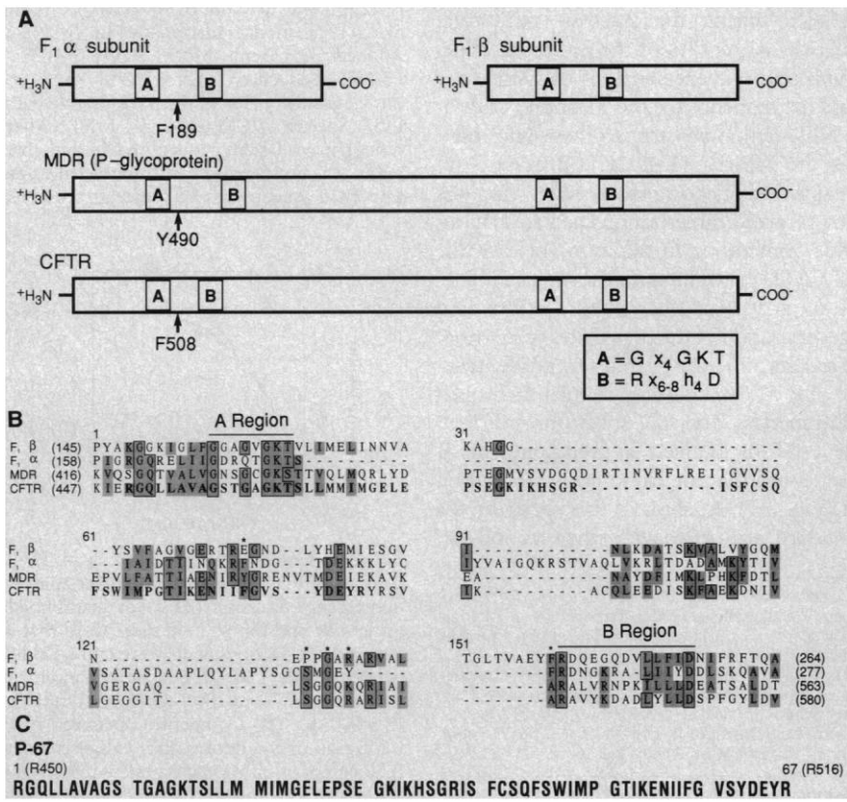


**Fig. 1.** Primary structure of P-67 and similarity to other nucleotide-binding proteins. (**A**) Schematic representation of sequence similarities in nucleotide-binding regions of CFTR (*2*), MDR (*5*), and the α and β subunits of the mitochondrial F, ATP synthase (*6*). The A and B consensus sequences are shown in the box, where x represents any amino acid and h represent hydrophobic residues. Arrows indicate the residues that correspond to $Phe^{508}$ in each protein when the sequences are aligned. (**B**) Sequence alignment of the four proteins showing five regions of similarity. Sequences were aligned with the Align program (Scientific and Educational Software) and gaps (-) were removed in regions of low similarity. The sequence that corresponds to the synthetic peptide P-67 is indicated by bold type. Boxed shading indicates residues that are identical and shading alone represents residues that are conservative replacements in at least three of the four sample proteins. Conservative amino acid substitution groups were as follows: A, F, I, L, M, V, and W (hydrophobic); N and Q [uncharged polar (hydrogen bond acceptors)]; S, T, and Y [uncharged polar (hydrogen bond donors)]; D and E (acidic); K, H, and R (basic); and F, Y, and W (aromatic). Mutations in the nucleotide-binding region of CFTR (*2, 11*) are indicated by an asterisk above the sequences. Abbreviations for the amino acids are A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr. (**C**) Primary sequence of the synthetic peptide P-67.

A and B consensus sequences. $Phe^{508}$ is the 43rd residue on the COOH-terminal side of the glycine flexible loop, whereas four other CF-causing mutations (*11*) occur immediately on the $NH_2$-terminal side of the B consensus sequence (Fig. 1B). In addition to the A and B consensus sequences, three other regions are similar in CFTR, MDR1, and the α and β subunits of mitochondrial ATP synthase (Fig. 1B). These regions of sequence similarity suggest a common function. Furthermore, the clustering of the CF-causing mutations in CFTR indicates that the proposed nucleotide-binding domain may perform a crucial function that is altered in the disease state.

In order to test directly whether the putative nucleotide-binding sequences of CFTR bind adenine nucleotides, we chemically synthesized and purified a 67–amino acid peptide, P-67 (*12*). This peptide extends from $Arg^{450}$ to $Arg^{516}$ (Fig. 1C) and, thus, includes the A consensus sequence and the region of sequence similarity that surrounds $Phe^{508}$. The synthetic peptide was purified on a reversed-phase high-performance liquid chromatography (HPLC) column and tested for nucleotide binding activity. The purified peptide fraction was >95% pure (Fig. 2A). We further characterized P-67 by amino acid analysis and peptide sequencing (*12*). Amino acid analysis both confirmed that the peptide fraction was >93% P-67 as designed (Fig. 2B) and showed that for all of the amino acids—except tryptophan and cysteine, which are degraded during the analysis—the actual value was within 7% of the predicted value. Amino acid sequence analysis confirmed the predicted sequence for the 33 amino acids at the $NH_2$-terminus where synthesis uncertainty is the greatest (Fig. 2C). Because P-67 contains a single cysteine residue, care was taken to avoid dimerization mediated by disulfide formation; in addition, neither the HPLC chromatogram nor the high solubility of P-67 at various pH values (*13*) indicated oligomerization of the peptide was occurring.

To determine the overall secondary structure of P-67, we collected circular dichroism (CD) spectra. The CD spectrum indicated that P-67 maintained a predominantly β sheet conformation in solution at neutral pH (Fig. 2D). Analysis of this spectrum for secondary structure with the use of the PROSEC program (*14*) revealed that P-67 consists of nearly 80% β sheet and <10% α helix. This is not unprecedented: the CD spectrum of the chemically synthesized nucleotide-binding domain of yeast hexokinase, in accordance with the x-ray structure, indicates a β sheet conformation (*15*).

A peptide that included the A consensus sequence of the β subunit of mitochondrial ATP synthase was shown to interact with nucleotide analogs by Garboczi and coworkers (*12*). It therefore seemed plausible that P-67 may also bind adenine nucleotides. The ability of P-67 to bind adenine nucleotides was tested by monitoring the fluorescence enhancement observed on interaction of the peptide with trinitrophenyl (TNP) adenine nucleotides. The fluorescence enhancement, which results from an increase in the quantum yield when ligand binds to the peptide, is proportional to the occupancy of the nucleotide binding sites (*16*). P-67 bound TNP-ATP with an apparent dissociation constant ($K_d$) (*16*) of 300 nM (Fig. 3A). No binding was observed when P-67 was denatured with 4 M urea (Fig. 3A), indicating that peptide structure is necessary for binding. Also, in control

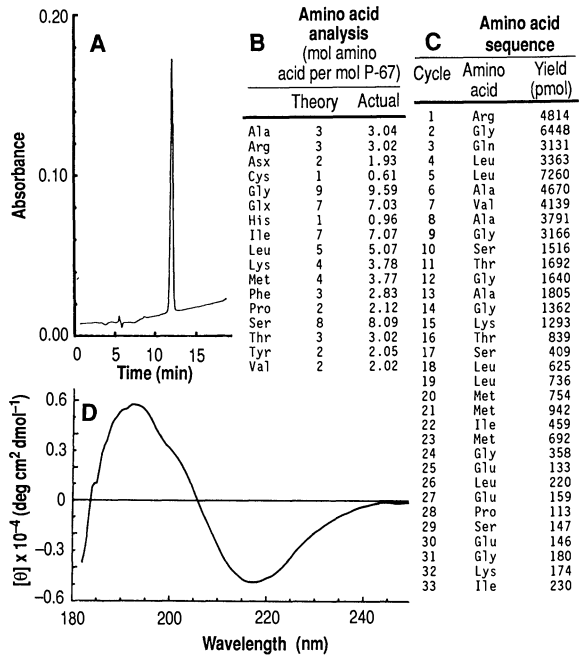


**Fig. 2.** Purity and structure of P-67. (**A**) HPLC chromatogram of purified P-67. Reversed-phase chromatography of purified P-67 (10 μg) was performed on a Waters μBondapack $C_{18}$ column. P-67 was eluted with 0.03% trifluoroacetic acid in a gradient of 30 to 80% acetonitrile in 25 min. The flow rate was 2 ml/min and absorbance at 220 nm was monitored. (**B**) Amino acid analysis of purified P-67. The composition of the synthetic peptide was determined and compared to the predicted product. The absolute amounts of each residue in 100 μg of P-67 were determined by pre-column phenylisothiocyanate derivatization after exhaustive acid hydrolysis (*25*). Tryptophan was degraded and cysteine concentrations diminished during the hydrolysis. (**C**) The amino acid sequence of purified P-67 (7 nmol) was determined by automated Edman degradation on an Applied Biosystems Model 470A Protein Sequencer (*12*). Values shown are the yields of the phenylthiohydantoin amino acid derivatives. (**D**) CD spectrum of purified P-67. Spectra of 120 μM P-67 in 50 mM tris-HCl (pH 7.0) were collected on an Aviv spectropolarimeter in a 0.01-cm pathlength demountable cell. Because P-67 contained a single cysteine residue, the buffer was deoxygenated with nitrogen to avoid dimerization resulting from disulfide formation. The smoothed average of seven separate spectra is shown.

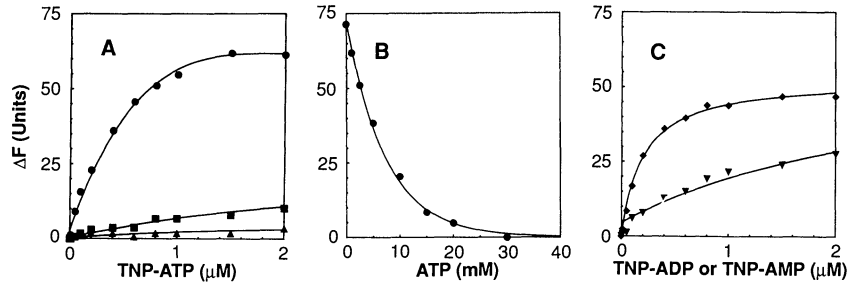


**Fig. 3.** Adenine nucleotide binding by P-67. Fluorescence emission at 540 nm was measured on an Aminco SPF-125 spectrofluorometer at an excitation wavelength of 410 nm. The enhancement of the fluorescence emission (ΔF) was determined by subtracting the fluorescence emission of the TNP-nucleotide and the peptide alone from that of the TNP-nucleotide in the presence of the peptide. All fluorescence measurements were corrected for the inner filter effect and for dilution. (**A**) Interaction of peptides with TNP-ATP. The enhancement of TNP-ATP fluorescence was measured in 50 mM tris-HCl (pH 7.4) with P-67 (5 μM) in the absence (●) or presence of 4 M urea (■) and with insulin (5 μM) (▲). (**B**) Competition between ATP and TNP-ATP for the nucleotide binding site of P-67. Inhibition of the fluorescence enhancement of TNP-ATP (5 μM) with P-67 (5 μM) resulting from ATP competition was determined in 50 mM tris-HCl (pH 7.4). (**C**) Interaction of TNP-ADP (◆) and TNP-AMP(▼) with 5 μM P-67 in 50 mM tris-HCl (pH 7.4).

P. J. Thomas and P. L. Pedersen, Laboratory for Molecular and Cellular Bioenergetics, Department of Biological Chemistry, The Johns Hopkins University School of Medicine, Baltimore, MD 21205.
P. Shenbagamurthi, Protein-Peptide Facility, Department of Biological Chemistry, The Johns Hopkins University School of Medicine, Baltimore, MD 21205.
X. Ysern, Department of Biophysics, The Johns Hopkins University School of Medicine, Baltimore, MD 21205.

*To whom correspondence should be addressed.

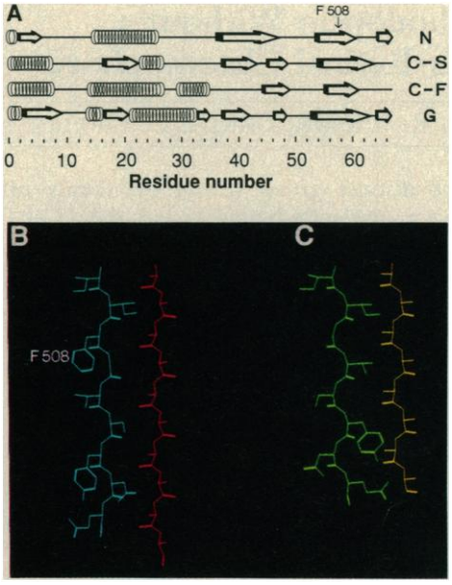


**Fig. 4.** Predicted higher order structure of P-67. (**A**) Predicted secondary structure of P-67. Secondary structure was predicted with the algorithms of Novotny and Auffray (N), Chandrasegaran and Smith (C-S), Chou and Fasman (C-F), and Garnier *et al.* (G) (*23*). The position of predicted α helical segments is indicated by circular coils, β strands are indicated by arrows, and turns and random coils are indicated by a straight line. (**B**) The peptide sequence I I F$^{508}$ G V S Y D E is shown (left) in the predicted β strand conformation in parallel with another β strand (polyalanine was used for illustration) (right). (**C**) Effect of deletion of Phe$^{508}$ on the proposed parallel β sheet structure. The deletion caused a loss of aromatic residues from one face of the β sheet and reorientation of conserved acidic residues at either end of the β strand to the opposite face of the β sheet.

studies, TNP-ATP failed to bind to insulin, which is of approximately the same size as P-67 but does not have a nucleotide-binding domain (Fig. 3A). The natural ligand, ATP, competed with TNP-ATP for binding to P-67 (Fig. 3B). The apparent $K_d$ for ATP, when corrected for competition with TNP-ATP (*17*), was 300 μM. P-67 also bound TNP-adenosine diphosphate (TNP-ADP) and TNP-adenosine monophosphate (TNP-AMP) (Fig. 3C). The affinity for TNP-ADP ($K_d$ = 200 nM) was of the same order as that for TNP-ATP, whereas the affinity for TNP-AMP ($K_d$ > 1 μM) was several fold lower. The equivalent affinities of P-67 for TNP-ATP and TNP-ADP and the lower affinity for TNP-AMP suggest that the specific interactions occur mainly between the peptide and the ADP moiety of TNP-ATP. The contribution of TNP to the binding was reflected in the difference between the affinity for TNP-ATP and that for ATP (Fig. 3, A and B).

We could not detect ATP hydrolysis catalyzed by P-67 with the use of a spectrophotometric assay that couples the production of ADP to the oxidation of reduced nicotinamide adenine dinucleotide (NADH) via the pyruvate kinase and lactate dehydrogenase reactions (*18*). It is possible that most of the residues necessary for binding are present in the peptide, whereas essential catalytic residues are missing. A recently identified mutant CFTR was shown to be CF-causing because it lacks the COOH-terminus of the protein (*19*). In addition, MDR from human myelogenous leukemia is an ATPase (*20*); and both the $NH_2$- and COOH-terminal portions of MDR1, each of which contains a nucleotide-binding domain, are required for ATP-dependent drug transport (*21*). Moreover, although the β subunit of the mitochondrial ATP synthase complex contains the catalytic site, both the α and β subunits are required for hydrolytic activity (*22*). Thus, two nucleotide binding sites, whether on one polypeptide chain, as in the case of CFTR and MDR, or on two polypeptide chains, as in the ATP synthase, may be required for catalysis to occur. The $NH_2$-terminal nucleotide binding site of CFTR in several regions is more similar to the noncatalytic α subunit of the ATP synthase complex than the catalytic β subunit (Fig. 1B).

In an attempt to relate the general structural information present in the CD data to the function of the peptide, we used four separate algorithms (*23*) to predict the secondary structure of P-67. In accordance with the CD spectrum (Fig. 2D), all of the algorithms predicted that P-67 contains significant regions of β strand structure (Fig. 4A). Although the four different predictions varied in the $NH_2$-terminal region of the peptide, they coincided in the COOH-terminal region of the peptide around Phe$^{508}$, where a β strand conformation was predicted (Fig. 4B). This is in contrast to a recent hypothetical scheme, based on the structure of adenylate kinase (*8*), which placed the Phe$^{508}$ in an α helix within a loop that lies outside the compact core of the protein (*24*).

Deletion mutations are most destructive in regions of secondary structure: most spectacularly in regions of β sheet because residues are reoriented from one face of the sheet to the other (*7*). On deletion of Phe$^{508}$, two large aromatic residues would be removed from one face of the β sheet (Fig. 4C). Furthermore, this deletion may cause reorientation of other conserved residues: Glu$^{504}$, or Asp$^{513}$ and Glu$^{514}$. The consequence of such a structural change in the CFTR protein may help explain the apparent loss of function caused by the deletion mutation.

## REFERENCES AND NOTES


1. T. F. Boat, M. J. Welsh, A. L. Beaudet, in *The Metabolic Basis of Inherited Disease*, C. L. Scriver, A. L. Beaudet, W. S. Sly, D. Valle, Eds. (McGraw-Hill, New York, ed. 6, 1989), pp. 2649–2680.
2. J. M. Rommens *et al.*, *Science* **245**, 1059 (1989); J. R. Riordan *et al.*, *ibid.*, p. 1066.
3. D. Ringe and G. A. Pestko, *Nature* **346**, 312 (1990).
4. J. E. Walker, M. Saraste, M. J. Runswick, N. J. Gay, *EMBO J.* **1**, 945 (1982).
5. P. Gros, J. Croop, D. E. Housman, *Cell* **47**, 371 (1986); C. Chen *et al.*, *ibid.*, p. 381.
6. J. E. Walker *et al.*, *J. Mol. Biol.* **184**, 677 (1985); T. M. Duncan, D. Parsonage, A. E. Senior, *FEBS Lett.* **208**, 1 (1986); D. N. Garboczi, A. H. Fox, S. L. Gerring, P. L. Pedersen, *Biochemistry* **27**, 553 (1988); J. H. Lee, D. N. Garboczi, P. J. Thomas, P. L. Pedersen, *J. Biol. Chem.* **265**, 4664 (1990).
7. M. G. Rossmann, A. Lijas, C. I. Brändén, L. J. Banaszak, in *The Enzymes*, P. D. Boyer, Ed. (Academic Press, New York, ed. 3, 1975), vol. 11, pp. 61–102.
8. G. E. Schulz *et al.*, *Nature* **250**, 120 (1974); E. F. Pai, W. Sachsenheimer, R. H. Schirmer, G. E. Schulz, *J. Mol. Biol.* **114**, 37 (1977); D. Dreusicke, P. A. Karplus, G. E. Schulz, *ibid.*, p. 359.
9. A. G. Redfield and M. Z. Papastavros, *Biochemistry* **29**, 3509 (1990); U. Enger, A. G. Tomasselli, G. E. Schulz, *J. Mol. Biol.* **185**, 649 (1987).
10. D. C. Fry, S. A. Kuby, A. S. Mildvan, *Biochemistry* **24**, 4680 (1985); *Proc. Natl. Acad. Sci. U.S.A.* **83**, 907 (1986); *Biochemistry* **26**, 1645 (1987).
11. G. R. Cutting *et al.*, *Nature* **346**, 366 (1990).
12. The peptide was synthesized on an Applied Biosystems Model 430A Peptide Synthesizer by the solid-phase method [R. B. Merrifield, *J. Am. Chem. Soc.* **85**, 2149 (1968)]. The details of the synthesis, analysis, and sequencing are as described [D. N. Garboczi, P. Shenbagamurthi, W. Kirk, J. Hullihen, P. L. Pedersen, *J. Biol. Chem.* **263**, 812 (1988)].
13. P. J. Thomas, P. Shenbagamurthi, P. L. Pedersen, unpublished observations.
14. C. T. Chang, C.-S. C. Wu, J. T. Yang, *Anal. Biochem.* **91**, 13 (1978).
15. K. A. Arora, P. Shenbagamurthi, M. Fanciulli, P. L. Pedersen, *J. Biol. Chem.* **265**, 5324 (1990).
16. L. D. Ward, *Methods Enzymol.* **117**, 400 (1985); R. A. Stinson and J. J. Holbrook, *Biochem. J.* **131**, 719 (1973).
17. G. P. Mullen, P. Shenbagamurthi, A. S. Mildvan, *J. Biol. Chem.* **264**, 19637 (1989).
18. M. E. Pullman, H. S. Penefsky, A. Datta, E. Racker, *ibid.* **235**, 3322 (1960).
19. M. B. White *et al.*, *Nature* **344**, 665 (1990).
20. H. Hamada and T. Tsuruo, *Cancer Res.* **48**, 4926 (1988).
21. S. J. Currier, K. Ueda, M. C. Willingham, I. Pastan, M. M. Gottesman, *J. Biol. Chem.* **264**, 14376 (1989).
22. M. K. Al-Shawi, D. Parsonage, A. E. Senior, *ibid.* **265**, 5595 (1990); Y. Kagawa, S. Ohta, Y. Otawara-Hamamoto, *FEBS Lett.* **249**, 67 (1989); K. Miwa and M. Yoshida, *Proc. Natl. Acad. Sci. U.S.A.* **86**, 6484 (1989).
23. J. Novotny and C. Auffray, *Nucleic Acids Res.* **12**, 243 (1984); S. Chandrasegaran and H. O. Smith, *Second New England Biolabs Workshop on Biological DNA Modifications*, 2 to 7 September 1990 (New England Biolabs, Berlin, 1990), p. 106 (abstr.); P. Y. Chou and G. D. Fasman, *Adv. Enzymol.* **47**, 45 (1978); J. Garnier, D. J. Osguthorpe, B. Robson, *J. Mol. Biol.* **120**, 97 (1978).
24. S. C. Hyde *et al.*, *Nature* **346**, 362 (1990).
25. B. A. Bidlingmeyer, S. A. Cohen, T. L. Tarvin, *J. Chromatogr.* **336**, 93 (1984).


26. We thank L. M. Amzel and W. B. Guggino for reading the manuscript before submission; J. Hullihen for technical assistance; J. Sondek for assistance with the CD spectroscopy; S. Chandrasegaran and H. O. Smith for allowing us to use the pairwise secondary structure prediction program; and M. Baumann and C. Filburn for helpful discussions. Supported by NIH grant CA 10951 to P.L.P.

28 September 1990; accepted 15 November 1990