[3H]uracil incorporation relative to the rate of tritium release suggested that under nonpermissive conditions DNA synthesis in *am* 37 infected cells was affected more by the nature of the DNA being synthesized than by a decline in the rate of synthesis of nucleotide precursors. Furthermore, this DNA may also be a poor template for RNA polymerase, so that late gene expression would be impaired. Finally, even if there was some expression of late genes, the DNA might prove difficult to package. The pyrophosphorylated DNA is being purified so that its physical–chemical properties may be examined.

The results obtained with *am* 37 confirm that putThy and thymine arise from hmUra by separate routes. The modification of ϕW-14 DNA is thus a complex process, involving at least two steps for the formation of putThy and an unknown number for the formation of thymine. The relationship of this modification system to the replication machinery for ϕW-14 DNA poses some intriguing questions. For example, are the modification reactions base-sequence specific? If ϕW-14 DNA synthesis involves a multienzyme complex, are the modification enzymes found in the complex? Such an association could couple modification to replication.

Acknowledgments

We thank Patty Miller for expert technical assistance and Dr. Bob Miller for advice and criticism.

References

Brandon, C. R. (1974) Ph.D. Thesis, Yeshiva University, New York.
Brown, D. M., Burdon, M. G., & Slatcher, R. P. (1968) *J. Chem. Soc. C*, 1051.
Goodman, DeW. S., & Popjak, G. (1960) *J. Lipid Res. 1*, 287.
Kelln, R. A., & Warren, R. A. J. (1973a) *J. Bacteriol. 113*, 510.
Kelln, R. A., & Warren, R. A. J. (1973b) *J. Virol. 12*, 1427.
Kropinski, A. M. B., & Warren, R. A. J. (1970) *J. Gen. Virol. 6*, 85.
Kropinski, A. M. B., Bose, R. J., & Warren, R. A. J. (1973) *Biochemistry 12*, 151.
Kumamoto, J., & Westheimer, F. (1955) *J. Am. Chem. Soc. 77*, 2515.
Lewis, H. A., Miller, R. C., Jr., Stone, J. C., & Warren, R. A. J. (1975) *J. Virol. 16*, 1375.
Maltman, K. L., Neuhard, J., Lewis, H. A., & Warren, R. A. J. (1980) *J. Virol. 34*, 354.
Marmur, J., Brandon, C., Neubort, S., Ehrlich, M., Mandel, M., & Konvicka, J. (1972) *Nature (London) 239*, 68.
Neuhard, J., Maltman, K. L., & Warren, R. A. J. (1980) *J. Virol. 34*, 347.
Pogolotti, A. L., Jr., & Santi, D. V. (1977) in *Bioorganic Chemistry* (Van Tamelen, E. E., Ed.) pp 277–310, Academic Press, New York.
Randerath, K. (1969) *Anal. Chem. 41*, 991.
Santi, D. V. (1967) *J. Heterocycl. Chem. 4*, 475.
Santi, D. V. (1971) *J. Heterocycl. Chem. 8*, 265.
Stanier, R. Y., Palleroni, N. J., & Doudoroff, M. (1966) *J. Gen. Microbiol. 43*, 159.
Tomich, P. K., Chin, C.-S., Wovcha, M. G., & Greenberg, G. R. (1974) *J. Biol. Chem. 249*, 7613.
Walsh, C. (1979) *Enzymatic Reaction Mechanisms*, pp XV, 978, W. H. Freeman, San Francisco, CA.
Warren, R. A. J. (1979) *J. Bacteriol. 137*, 1053.

Primary Structure of the Phage P22 Repressor and Its Gene *c2*†

Robert T. Sauer,* Jullie Pan, Peggy Hopper, Kathleen Hehir, John Brown, and Anthony R. Poteete‡

ABSTRACT: The amino acid sequence of the Salmonella phage P22 repressor and the DNA sequence of its gene *c2* have been determined. Sequential Edman degradations on intact P22 repressor and repressor peptides generated by proteolytic and chemical cleavages have been overlapped to give approximately 97% of the complete protein sequence. Additionally, the nucleotide sequence of the P22 *c2* repressor gene has been determined by DNA sequencing techniques. The DNA sequence and partial protein sequence are collinear and together define the complete amino acid sequence of P22 repressor. The repressor is a single-chain 216 amino acid polypeptide. Basic residues in the sequence tend to be clustered, and residues 9–20 are highly basic, containing five arginyl and three lysyl residues. The carboxyl-terminal 133 amino acids of the *c2* repressor are homologous to the carboxyl-terminal sequence of the coliphage λ *cI* repressor. The amino-terminal sequences of these two repressors show little similarity.

Upon infection of their hosts, members of the lambdoid family of temperate bacteriophages can either form lysogens or undergo lytic growth. In a lysogen, the phage DNA is integrated into the host chromosome, and all of the phage genes necessary for lytic growth are turned off. This negative regulation of phage transcription is mediated by a repressor protein. In the case of Salmonella phage P22, the repressor protein encoded by the phage *c2* gene is necessary for maintenance of lysogeny (Levine, 1957; Levine & Smith, 1964). Phage with mutant *c2* polypeptides, or phage in which the synthesis of the *c2* repressor is reduced, cannot form lysogens and grow only lytically.

The P22 *c2* repressor controls gene expression by binding to the phage DNA at two operator regions (Bronson & Levine, 1972; Ballivet & Eisen, 1978). The nucleotide sequence of these operator regions has been determined, and the interaction

† From the Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139 (R.T.S., P.H., and K.H.), and the Biological Laboratories, Harvard University, Cambridge, Massachusetts 02138 (J.P., J.B., and A.R.P.). *Received October 11, 1980.* This research was supported in part by a grant from the Athwin Foundation and by National Institutes of Health Grant AI-16892. A.R.P. was a postdoctoral fellow of the Helen Hay Whitney Foundation.

‡ Present address: Department of Molecular Genetics and Microbiology, University of Massachusetts Medical School, Worcester, MA 01605.

0006-2960/81/0420-3591$01.25/0

of *c*2 protein with these sequences has been studied (Poteete et al., 1980). Both operator regions consist of three catenated repressor binding sites. Depending on which site or combination of sites is occupied, P22 repressor can mediate negative control of early phage gene transcription and positive or negative control of its own synthesis (A. R. Poteete and M. Ptashne, unpublished results).

Phages λ and 434 have repressor proteins analogous in function to the *c*2 repressor protein of P22. These repressors are essential for maintenance of lysogeny, act by binding to two operator regions containing catenated repressor binding sites, and control their own synthesis [see Ptashne et al. (1980) and references therein]. Moreover, these repressors are all active as oligomers (Chadwick et al., 1970; Pirrotta et al., 1970; Ballivet & Eisen, 1978), and all can be inactivated by the P22 antirepressor protein (Susskind & Botstein, 1975) and by the *rec*A protease (Roberts et al., 1978; Phizicky & Roberts, 1980; R. Yocum, personal communication). Nevertheless, the operators to which these repressors bind are completely distinct in nucleic acid sequence, and a particular repressor binds only to its operators, not to the operators of a heteroimmune phage. The repressors of P22 and coliphage 21 are an apparent exception to this rule. Each binds to the operators of the other phage as well as its own operators (Ballivet et al., 1977; Ballivet & Eisen, 1978). However, Poteete et al. (1980) have shown that the operators of the two phages are virtually identical in nucleotide sequence, and the repressor proteins are also very similar in amino acid sequence as judged by heteroduplex analysis of their respective genes (Botstein & Herskowitz, 1974) and by comparative tryptic peptide analysis of the two repressor proteins (M. Ballivet, personal communication).

One question of interest is whether the different phage repressor proteins which are related in function are also related in structure. The repressor protein of phage λ has been characterized in some detail. Its amino acid sequence has been determined, and it has been shown to consist of two structural domains (Sauer & Anderegg, 1978; Pabo et al., 1979; Sauer et al., 1979). The amino-terminal domain functions in operator DNA recognition and positive control, while the carboxyl-terminal domain mediates subunit oligomerization. At present, the structures of the 434, 21, and P22 repressors are less well characterized.

In this paper, we present data that determine the complete amino acid sequence of the P22 repressor and the nucleic acid sequence of its gene *c*2. These studies reveal some striking and unexpected homologies between the carboxyl-terminal sequences of the P22 and λ repressors. We discuss the implications of these homologies in terms of the common functions and the evolution of these phage repressors.

Materials and Methods

Repressor Protein. *c*2 repressor was purified from the *Escherichia coli* strain W3110/pTP 69. This strain contains a recombinant plasmid which directs the synthesis of 1–3% of the cellular protein as *c*2 repressor (A. R. Poteete and M. Ptashne, unpublished results). The purification of *c*2 protein is outlined here and will be presented in greater detail elsewhere (R. T. Sauer and A. R. Poteete, unpublished results). A total of 150 g of W3110/pTP 69 cells was lysed by sonication, and *c*2 repressor was purified by differential precipitation with polyethylenimine and ammonium sulfate and by ion-exchange chromatography on CM-Sephadex, hydroxylapatite, and phosphocellulose. Purification steps were monitored by sodium dodecyl sulfate ($NaDodSO_4$) gel electrophoresis. From a single purification, about 200 mg of *c*2 protein was obtained. This material was greater than 95% pure as judged by $NaDodSO_4$ gel electrophoresis.

S-Carboxymethylation. Purified *c*2 protein was dialyzed into 0.2 M NH_4HCO_3 (pH 8); the mixture was frozen and lyophilized. A total of 50 mg of protein was dissolved in 12 mL of 8 M guanidine hydrochloride, and 3 mL of 2.5 M tris(hydroxymethyl)aminomethane hydrochloride (Tris-HCl) (pH 8.1) and 10 mM ethylenediaminetetraacetic acid (EDTA) was added. Iodo[^{3}H]acetic acid (100 μCi; NEN, 0.52 mCi/mg) was added, and the mixture was incubated at 23 °C for 20 min. After this time, 150 μL of unlabeled iodoacetic acid (30 mg/mL in H_2O) was added, and the reaction was allowed to proceed for 30 min before quenching with 10 μL of β-mercaptoethanol. The reaction mixture was diluted with an equal volume of 1 M acetic acid and 1.4 mM β-mercaptoethanol and was dialyzed extensively against this buffer. Total incorporation of counts into protein was 8.3×10^7 cpm (counted in Aquasol). The alkylated repressor protein was stored frozen at −20 °C, and aliquots were lyophilized prior to other treatments.

Maleylation. Alkylated repressor (11 mg) was dissolved in 4 mL of 8 M guanidine hydrochloride and 1 mL of 1 M $NaHCO_3$. The pH was adjusted to 9.0 by addition of 1 N NaOH. Ten 100-μL aliquots of maleic anhydride (100 mg/mL in dioxane) were added to the protein with stirring. After each addition, the pH was adjusted to 9.0 with NaOH. The reaction mixture was then dialyzed extensively against 0.2 M NH_4HCO_3 (pH 8), frozen, and lyophilized.

Proteolytic and Chemical Cleavage of Repressor Protein and Peptides. Cleavage of S-carboxymethylated repressor or of repressor peptides with TPCK-trypsin (Worthington) was performed in 0.2 M NH_4HCO_3 (pH 8.0). The protein concentration was 2–5 mg/mL, and the enzyme/substrate ratio was 1:100 (w/w). Maleylated repressor or repressor peptides were cleaved for 4 h at 37 °C. Unmaleylated repressor was cleaved for 20 h at 37 °C. Following digestion, samples were loaded directly on gel filtration columns for peptide separation.

Cleavage of maleylated, carboxymethylated repressor with staphylococcal protease V8 (Miles Laboratories) was also performed in 0.2 M NH_4HCO_3 (pH 8.0). The protein concentration was 5.5 mg/mL, and the enzyme/substrate ratio was 1:150 (w/w). The digestion was carried out for 48 h at 37 °C.

Cleavage of S-carboxymethylated repressor (6.4 mg) with cyanogen bromide (11 mg) was performed in 1 mL of 70% formic acid. The mixture was sealed under nitrogen and incubated at 20 °C for 20 h in the dark. The reaction was terminated by addition of 20 mL of distilled water, freezing, and lyophilization. Prior to column chromatography, the CNBr digest was redissolved in 200 μL of 90% formic acid and then diluted to 1 mL by dropwise addition of distilled water. Under these conditions, all peptides remained soluble.

Separation of Peptide Mixtures. The overall strategy for peptide isolation was the same for all types of cleavages. The digest was first chromatographed on a 1.5 × 90 cm Sephadex G-75 (superfine) column equilibrated either in 0.2 M NH_4HCO_3 (tryptic, maleylated tryptic, and maleylated staphylococcal protease digests) or in 9% formic acid (CNBr digest). Peptides containing Trp, Tyr, or mal-Lys were located by absorbance at 280 nm and/or 230 nm, and [^{3}H]CM-cysteine containing peptides were located by scintillation counting of a portion of the effluent in Aquasol. Regions of the included volume of the column not containing absorbance or radioactivity were also pooled. Pooled fractions were lyophilized, and homogeneity was assessed by two cycles of manual Edman degradation on 5–10% of the pooled material.

Pools found to be heterogeneous by end-group analysis were further purified by high-performance liquid chromatography. Reverse-phase chromatography was performed on either a 0.39 × 30 cm Waters μBondapak C_{18} or a μBondapak CN column with two Waters Model 6000A pumps, a Waters Model 450 variable wavelength detector, a Waters Model 660 solvent programmer, and a Waters U6K injector. The aqueous phase used in the chromatography was either 0.2 M ammonium acetate (pH 6.5), 0.2 M triethylamine acetate (pH 5.0), or 0.3 M triethylamine acetate (pH 3.0). All reagents were made from water purified by using a Millipore Milli-Q water purification system. Triethylamine was refluxed over phthalic anhydride and distilled twice prior to use. In all cases, the organic phase used for chromatography was 1-propanol and elution was achieved with a linear gradient from 0% to 50% organic phase at a flow rate of 1–1.5 mL/min over 30–60 min. Peptides were detected by absorbance at 230 nm, by radioactivity, and, in some cases, by fluorescence, following reaction of a part of the sample with fluorescamine (Udenfriend et al., 1972). Samples in ammonium buffers were lyophilized thoroughly or dried twice in a vacuum oven at 60 °C prior to fluorescamine treatment. Pooled peptides were frozen and lyophilized directly, and homogeneity was assessed by two cycles of manual Edman degradation. Samples greater than 90% pure by end group or samples containing approximately equal mixtures of only two peptides were used for further sequencing studies.

Amino Acid Sequencing. Manual Edman degradations were performed by a modification of the three-step procedure of Edman (1960) as described by Sauer et al. (1974). Automated degradations were performed by using a Beckman 890C Sequencer and the 0.1 M Quadrol program described by Brauer et al. (1975). Peptides were degraded in the presence of 6 mg of polybrene (Tarr et al., 1978). The polybrene was added to the sequencer cup and subjected to two cycles of automated Edman degradation prior to addition of the peptide. In some cases, 2-anilinothiazolinone amino acid derivatives were converted to Pth amino acid derivatives by using the Sequemat P6 autoconverter. In these cases, methanol–HCl was used for the conversion reaction and aspartic acid, glutamic acid, and N^{ϵ}-maleyllysine were identified as their methyl ester derivatives. Otherwise, conversion was carried out for 10 min at 80 °C by using 1 N HCl (Ilse & Edman, 1963) and Pth derivatives except those of Arg and His were extracted with ethyl acetate.

Pth amino acids were identified by gas–liquid chromatography by using a 10% SP-400 column and a Hewlett-Packard 5830A gas chromatograph and high-performance LC employing a Waters System including two Model 6000A pumps, a Model 710A injector, a Model 440 absorbance detector, and a Model 740 system controller. The dehydro derivatives of Pth-Thr and Pth-SCMC were detected by absorbance at 313 nm. Other Pth amino acid derivatives were detected by absorbance at 254 nm. Reverse-phase chromatography was performed by using a Waters RCM-C_{18} column and a two-buffer system. Buffer A was 20 mM sodium acetate (pH 3.0) and 10% propanol, and buffer B was 100% propanol. The elution program was as follows: 3.9-min isocratic elution (90% A, 10% B) at 1.5 mL/min, followed by a 4-min linear gradient (90% A, 10% B to 68% A, 32% B) at 1.8 mL/min, followed by a 4-min isocratic elution (68% A, 32% B) at 1.8 mL/min, followed by reequilibration to starting conditions. The turnaround time for the program was 20 min. In our hands, this system gives adequate separation of all the Pth amino acids with the exception of Pth-Arg and Pth-His which chromatograph poorly. These amino acids (which remain in the aqueous phase following conversion) were analyzed separately by using a Waters C_{18} mBondapak column and isocratic elution at 1.5 mL/min with 20 mM NaOAc (pH 5) and 18% methanol.

Amino Acid Analysis. Peptide or protein samples were hydrolyzed in vacuo at 105 °C in 6 N HCl and 1% phenol. Hydrolysis was for either 24, 48, or 72 h. Amino acids were analyzed on a Durham D500 analyzer equipped with PDP 8M software.

DNA Sequencing. Two plasmids, pPR4 (A. R. Poteete and T. M. Roberts, unpublished results) and pJB1, were used in the DNA sequencing studies described here. Both plasmids carry the intact P22 *c2* gene. pJB1 was constructed by cloning the *Hin*dIII–*Pvu*II fragment of P22 [extending from 0.699 to 0.755 on the P22 map (Chisholm et al., 1980)] into a plasmid vector derived from pBR322. pBR322 was digested with *Bam*H1, and the sticky ends were filled in with DNA polymerase large fragment (New England Biolabs) and all four deoxynucleotide triphosphates. This linear molecule was then digested with *Hin*dIII. The P22 *Hin*dIII–*Pvu*II fragment was then ligated with the *Hin*dIII and filled-in *Bam* ended fragments of pBR322, and *E. coli* strain 294 was transformed with the ligation mix. Transformants were selected for ampicillin resistance and for immunity to superinfection by phage λimm^{P22}. Plasmid DNA was prepared from single transformants and was shown by restriction analysis to contain both parental restriction fragments in the expected orientation.

DNA sequencing was carried out by using the chemical methods of Maxam & Gilbert (1977, 1980). End-labeled restriction fragments were obtained by digesting intact plasmid DNA and then labeling the resulting mixture of fragments either with [γ-^{32}P]ATP and T_4 polynucleotide kinase or with DNA polymerase large fragment and [γ-^{32}P]dXTPs. Singly end-labeled fragments were generated by digesting with a second restriction endonuclease, and these fragments were purified by polyacrylamide gel electrophoresis prior to sequencing.

Results

Amino-Terminal Sequence. S-Carboxymethylated *c2* repressor was subjected to 44 cycles of automated Edman degradation. The results of this degradation are presented schematically in Table I and establish the amino-terminal sequence of the P22 repressor.

Peptide Sequences. A number of peptides of *c2* repressor were purified by a combination of gel filtration and reverse-phase, high-performance LC following digestion of the protein with either trypsin, staphylococcal protease, or cyanogen bromide (see Materials and Methods). The results of automated Edman degradation on these purified peptides are represented in Tables I and II. The amino acid compositions of many of these peptides are given in Table III. These data define two continuous stretches of amino acid sequence. The first includes amino acids 1–75. This sequence is generated by direct overlap of the amino-terminal sequence of P22 repressor with the sequences determined for peptides CN4 and MT7. The position of these peptides in the sequence of the repressor is illustrated schematically in Figure 1. A second stretch of continuous sequence encompassing 133 amino acids results from overlapping of sequences determined for peptides MT9, CN5, MT10, CN7, MT11, SP21, MT12, CN9, T25, and T26 (see Figure 1). The overlap of peptides MT11 and SP21 is only three amino acids. However, the alignment of these two peptides is completely supported by the DNA sequence data (see below). The contiguity of peptides T25 and T26 is based on the amino acid compositions of peptides MT12

Table I: Edman Degradations[a]

intact *c*2	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 Met-Asn-Thr-Gln-Leu-Met-Gly-Glu-Arg-Ile-Arg-Ala-Arg-Arg-Lys-Lys-Leu-Lys-Ile-Arg-Glu-Ala-Ala-Leu-Gly-Lys-Met-Val- 350 187 116 193 200 215 158 153 91 125 112 141 97 108 135 147 118 122 107 72 55 79 139 100 34 74 88 69	
	29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 Gly-Val-Ser-Asn-Val-Ala-Ile-Ser-Gln-Trp-Glu-Arg-Ser-Glu-Thr-Glu. . . 76 75 39 48 52 52 55 23 41 16 34 22 34 21 8 29	
CN4	28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 Val-Gly-Val-Ser-Asn-Val-Ala-Ile-Ser-Gln-Trp-Glu-Arg-Ser-Glu-Thr-Glu-Pro-Asn-Gly-Glu-Asn-Leu-Leu-Ala-Leu-Ser-Lys-Ala-Leu- 48 18 44 23 29 39 42 34 29 21 9 23 7 21 21 9 17 10 12 10 14 13 16 21 13 14 12 19 17 13	
	58 59 60 61 62 63 64 65 66 67 Gln-Cys-Ser-Pro-Asp-Tyr-Leu-Leu-Lys-Gly. . . 19 13 16 2 7 8 12 17 6 5	
MT7	41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 Ser-Glu-Thr-Glu-Pro-Asn-Gly-Glu-Asn-Leu-Leu-Ala-Leu-Ser-Lys-Ala-Leu-Gln-Cys-Ser-Pro-Asp-Tyr-Leu-Leu-Lys-Gly-Asp-Leu- 43 90 21 82 55 63 59 65 44 55 66 51 48 32 30 39 38 34 19 17 25 28 22 29 38 23 16 13 16	
	70 71 72 73 74 75 Ser-Gln-Thr-Asn-Val-Ala. . . 8 10 7 9 11 13	
MT10	108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 Ala-Ile-Glu-Asn-Trp-His-Asp-Thr-Thr-Val-Asp-Cys-Ser-Glu-Asp-Ser-Phe-Trp-Leu-Asp-Val-Gln-Gly-Asp-Ser-Met-Thr-Ala-Pro-Ala-Gly- 282 284 159 128 95 82 63 57 65 213 139 155 181 145 75 181 94 31 82 62 80 60 78 43 75 64 21 51 24 71 23	
	139 140 141 142 143 144 145 146 147 148 149 Leu-Ser-Ile-Pro-Glu-Gly-Met-Ile-Ile-Leu-Val. . . 42 37 44 14 27 11 39 40 68 35 34	
MT11	157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 Asn-Gly-Lys-Leu-Val-Val-Ala-Lys-Leu-Glu-Gly-Glu-Asn-Glu-Ala-Thr-Phe-Lys-Lys-Leu-Val-Met-Asp-Ala-Gly-Arg 91 90 21 88 98 102 105 35 98 50 26 40 42 54 69 28 40 13 17 40 40 47 29 43 20 16	

[a] Numbers under residues are yields of the Pth-amino acid derivatives × 10^{-10} mol. Numbers above residues represent residue positions in the final sequence (Figure 2).

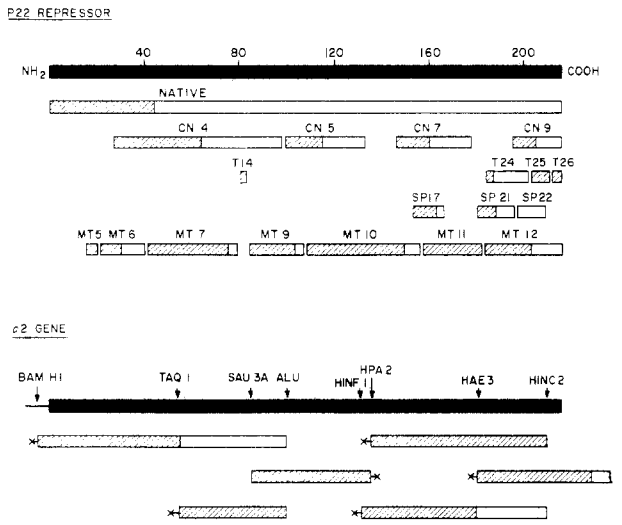

FIGURE 1: Schematic representation of positions of peptide and restriction fragments used for sequence analysis. The hatched areas represent regions of the fragments for which reliable sequence information was obtained. Only selected restriction endonuclease cleavage sites in the *c*2 gene are shown. The *Bam*H1 site shown is present only in the recombinant plasmid pPR4. All DNA fragments were 3′ labeled with the exception of the *Hae*III fragment which was 5′ labeled. The labeled termini of restriction fragments are indicated by ×-.

Table II: Edman Degradations[a]

MT5	15 16 17 18 19 20 Lys-Lys-Leu-Lys-Ile-Arg 260 204 281 200 146 63	
MT6	21 22 23 24 25 26 27 28 29 30 Gln-Ala-Ala-Leu-Gly-Lys-Met-Val-Gly-Val. . . 47 42 36 49 39 17 23 33 34 37	
T14	80 81 82 83 His-Glu-Pro-Arg 163 208 111 85	
MT9	84 85 86 87 88 89 90 91 92 93 94 95 96 97 Gly-Ser-Tyr-Pro-Leu-Ile-Ser-Trp-Val-Ser-Ala-Gly-Gln-Trp- 94 46 63 28 40 46 20 13 30 14 29 18 12 4	
	98 99 100 101 102 103 Met-Glu-Ala-Val-Glu-Pro. . . 16 20 13 8 5 3	
CN5	99 100 101 102 103 104 105 106 Glu-Ala-Val-Glu-Pro-Tyr-His-Lys- 38 63 59 27 23 40 18 31	
	107 108 109 110 111 112 Arg-Ala-Ile-Glu-Asn-Trp. . . 9 33 25 13 18 5	
CN7	146 147 148 149 150 151 152 153 154 Ile-Ile-Leu-Val-Asp-Pro-Glu-Val-Glu- 41 47 37 35 26 21 16 24 16	
	155 156 157 158 159 160 Pro-Arg-Asn-Gly-Lys-Leu. . . 10 6 2 3 2 2	
SP17	153 154 155 156 157 158 159 160 161 162 163 Val-Glu-Pro-Arg-Asn-Gly-Lys-Leu-Val-Val-Ala. . . 82 44 34 18 8 10 6 8 12 16 8	
SP21	180 181 182 183 184 185 186 187 188 Ala-Gly-Arg-Lys-Phe-Leu-Lys-Pro-Leu. . . 215 165 60 108 83 70 42 28 46	
MT12	183 184 185 186 187 188 189 190 191 Lys-Phe-Leu-Lys-Pro-Leu-Asn-Pro-Gln- 65 220 160 110 84 170 58 84 42	
	192 193 194 195 196 197 Tyr-Pro-Met-Ile-Glu-Ile. . . 72 83 145 125 130 170	
CN9	195 196 197 198 199 200 201 202 203 204 205 Ile-Glu-Ile-Asn-Gly-Asn-Cys-Lys-Ile-Ile-Gly. . . 79 43 56 8 11 12 9 5 5 6 4	
T25	203 204 205 206 207 208 209 210 211 Ile-Ile-Gly-Val-Val-Val-Asp-Ala-Lys 273 289 197 240 282 310 171 112 76	
T26	212 213 214 215 216 Leu-Ala-Asn-Leu-Pro 185 179 120 91 30	

[a] Numbers under residues are yields of the Pth-amino acid derivatives × 10^{-10} mol. Numbers above residues represent residue positions in the final sequence (Figure 2).

and CN9 (Table III) in conjunction with the sequences of MT12, CN9, T25, and T26. This stretch of 133 amino acids occupies residues 84–216 in the repressor sequence (Figure 3). Peptide T26 was generated by cleavage of P22 repressor with trypsin. However, this peptide contains no arginine or lysine (Tables II and III), and thus it must be the carboxyl-terminal peptide of P22 repressor. The carboxyl-terminal residue of T26, proline, is thus the carboxyl-terminal residue of the repressor protein.

DNA Sequence of c2. Six singly end-labeled restriction endonuclease fragments bearing *c*2 sequences were sequenced by the chemical methods of Maxam & Gilbert (1980). The positions of these restriction fragments in the *c*2 gene are represented in Figure 2, as are the extents of reliable nucleic acid sequence information derived from each fragment. Direct overlapping of these nucleic acid sequences defines a sequence of 648 bases beginning with an ATG codon and followed by a single ochre termination codon. The nucleic acid sequence of *c*2 is shown in Figure 2. The sequence determined for amino acids 1–75 of P22 repressor is collinear with the nucleic acid sequence from base 1 to base 225, and the amino acid sequence from residues 84 to 216 is collinear with the nucleotide sequence from base 250 to base 648. The nucleic acid sequence

Table III: Peptide Amino Acid Compositions[a]

	MT5	MT6	MT7	MT9	MT10	MT11	MT12	T24
CM-Cys	(0)	(0)	0.99 (1)	(0)	0.60 (1)	(0)	0.70 (1)	0.87 (1)
Asx	(0)	1.06 (1)	5.14 (5)	(0)	7.07 (7)	2.74 (3)	5.09 (5)	2.90 (3)
Thr	(0)	(0)	1.83 (2)	(0)	2.86 (3)	1.14 (1)	(0)	(0)
Ser	(0)	1.69 (2)	4.20 (5)	2.88 (3)	3.77 (4)	(0)	(0)	(0)
Glx	(0)	2.88 (3)	4.90 (5)	3.32 (3)	6.19 (6)	3.00 (3)	2.04 (2)	2.21 (2)
Pro	(0)	(0)	1.64 (2)	1.84 (2)	3.79 (4)	(0)	3.80 (4)	2.85 (3)
Gly	(0)	2.03 (2)	2.06 (2)	2.02 (2)	3.10 (3)	3.11 (3)	2.03 (2)	1.11 (1)
Ala	(0)	2.93 (3)	3.09 (3)	2.02 (2)	3.05 (3)	2.99 (3)	1.97 (2)	(0)
Val	(0)	2.89 (3)	0.83 (1)	2.07 (2)	4.01 (4)	2.76 (3)	2.56 (3)	(0)
Met	(0)	0.83 (1)	(0)	0.79 (1)	1.91 (2)	0.73 (1)	0.86 (1)	0.93 (1)
Ile	0.93 (1)	1.01 (1)	(0)	1.09 (1)	3.24 (4)	(0)	3.44 (4)	1.72 (2)
Leu	1.05 (1)	1.13 (1)	7.20 (7)	1.22 (1)	3.19 (3)	3.14 (3)	4.20 (4)	2.15 (2)
Tyr	(0)	(0)	1.94 (2)	0.86 (1)	(0)	(0)	0.89 (1)	0.88 (1)
Phe	(0)	(0)	(0)	(0)	0.95 (1)	0.92 (1)	0.87 (1)	0.83 (1)
His	(0)	(0)	0.81 (1)	0.88 (1)	0.99 (1)	(0)	(0)	(0)
Lys	2.86 (3)	0.96 (1)	2.10 (2)	0.82 (1)	(0)	4.06 (4)	3.76 (4)	1.91 (2)
Arg	0.95 (1)	0.94 (1)	0.93 (1)	0.95 (1)	1.02 (1)	0.98 (1)	(0)	(0)
Trp	(0)	(1)	(0)	(2)	(2)	(0)	(0)	(0)
sequence position	15–20	21–40	41–79	84–107	108–156	157–182	183–216	184–202

	T25	T26	CN4	CN9	SP17	SP21	SP22
CM-Cys	(0)	(0)	0.74 (1)	0.69 (1)	(0)	(0)	0.82 (1)
Asx	1.05 (1)	0.92 (1)	6.08 (6)	4.25 (4)	1.20 (1)	1.08 (1)	3.00 (3)
Thr	(0)	(0)	1.83 (2)	(0)	(0)	(0)	(0)
Ser	(0)	(0)	8.87 (10)	0.23 (0)	0.20 (0)	(0)	(0)
Glx	(0)	(0)	8.95 (9)	1.34 (1)	2.01 (2)	2.01 (2)	0.24 (0)
Pro	(0)	0.96 (1)	3.85 (4)	1.28 (1)	0.78 (1)	2.91 (3)	(0)
Gly	0.97 (1)	(0)	5.10 (5)	2.06 (2)	1.20 (1)	1.09 (1)	2.01 (2)
Ala	0.95 (1)	1.07 (1)	5.07 (5)	2.05 (2)	1.00 (1)	1.01 (1)	0.19 (0)
Val	2.61 (3)	(0)	4.90 (5)	2.36 (3)	2.56 (3)	(0)	2.90 (3)
Met	(0)	(0)	(1)	(0)	(0)	0.72 (1)	(0)
Ile	1.73 (2)	(0)	1.80 (2)	2.87 (4)	0.17 (0)	0.95 (1)	2.36 (3)
Leu	(0)	2.17 (2)	8.52 (8)	1.97 (2)	1.87 (2)	2.05 (2)	0.19 (0)
Tyr	(0)	(0)	2.62 (3)	(0)	(0)	0.78 (1)	(0)
Phe	(0)	(0)	(0)	(0)	(0)	0.91 (1)	(0)
His	(0)	(0)	2.03 (2)	(0)	(0)	(0)	(0)
Lys	0.99 (1)	(0)	1.82 (2)	1.79 (2)	1.69 (2)	1.86 (2)	0.98 (1)
Arg	(0)	(0)	2.98 (3)	(0)	0.97 (1)	0.92 (1)	(0)
Trp	(0)	(0)	(2)	(0)			
sequence position	203–211	212–216	28–98	195–216	153–166	180–196	197–209

[a] The numbers in the columns are the observed compositional values expressed in residues per mole of peptide. Values less than 0.15 residue per mol are not shown. Numbers in parentheses are the expected compositional values calculated from the primary sequence. All experimental data was derived from amino acid analysis of 24-h acid hydrolysates except for peptide CN9 which is an analysis following a 72-h acid hydrolysis.

additionally predicts the amino acid sequence of residues 76–83 of P22 repressor as being Tyr-His-Ser-Arg-His-Glu-Pro-Arg. The amino acid sequence of peptide T14 (Table II) matches the sequence predicted for amino acids 80–83 of P22 repressor. Moreover, the amino acid compositions of peptides MT7 and CN4 (Table III) are consistent with the amino acid sequence predicted for residues 76–83 from the nucleic acid sequence data. We conclude that the amino acid and nucleic acid sequences are completely collinear.

Complete Sequence and Confirming Data. The data presented here define a continuous sequence of 216 amino acids for the P22 *c*2 repressor. This sequence and the nucleic acid sequence of the *c*2 gene are shown in Figure 2. The composite amino acid composition of P22 repressor (Table IV) is in good agreement with the composition predicted from the protein sequence, although there are minor differences in the values for Glx, Met, and Lys. Overall, we feel that the protein and nucleic acid sequences are highly reliable because of the complementary nature of the protein and DNA sequencing data.

At physiological pH, the P22 repressor protein appears to be close to electrically neutral. There are 27 aspartyl and glutamyl residues and 26 arginyl and lysyl residues. The four histidyl residues, depending upon their ionization state, can thus swing the balance from a slightly acidic to a slightly basic molecule. The basic residues in the sequence, however, tend to be clustered. Residues 9–20, 87–93, 105–107, and 174–186 account for 16% of the P22 repressor sequence and contain 66% of the basic residues.

Three *S*-(carboxymethyl)cysteine residues were located in the P22 amino acid sequence studies. These have all been assigned as cysteine residues in the final sequence since they were alkylated in near-stoichiometric yield (see Table IV) in the absence of prior reduction (see Materials and Methods).

The molecular weight calculated from the sequence for P22 repressor is 24 041. However, on $NaDodSO_4$–polyacrylamide gel electrophoresis, the P22 *c*2 repressor has been reported to migrate at an apparent molecular weight of about 31 000 (Ballivet & Eisen, 1978). In our hands, P22 repressor migrates at an apparent molecular weight of about 29 000 and runs consistently as if it were larger than the λ *c*I repressor, a protein of 236 amino acids. There is thus a considerable discrepancy between the apparent monomer molecular weight calculated from $NaDodSO_4$ gel electrophoresis and the actual

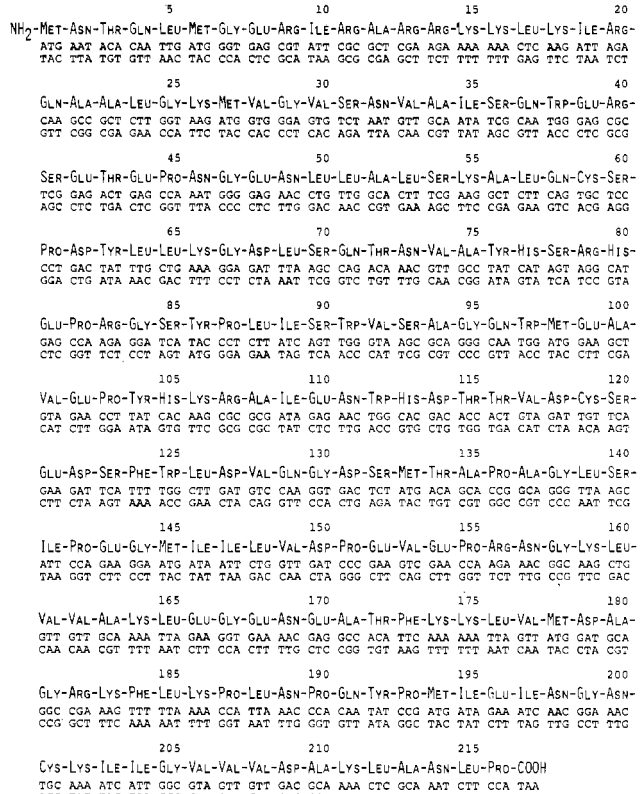

FIGURE 2: Protein and gene sequences of the phage P22 *c*2 repressor. The numbers refer to residue positions in the protein sequence.

Table IV: P22 Repressor Amino Acid Composition[a]

	24 h	48 h	72 h	PAO	composite[b]	sequence
Cys	ND	ND	ND	3.03[d]	3.03	3
Asx	22.30	22.30	22.40	23.92	22.33	22
Thr	6.58	6.09	6.05	6.66	6.80	7
Ser	12.69	11.29	10.30	13.50	13.87	14
Glx	25.56	25.54	25.46	26.21	25.52	25
Pro	12.65	12.59	12.54	14.05	12.58	13
Gly	15.17	15.20	15.39	16.15	15.25	15
Ala	16.96	16.70	17.01	16.72	16.89	17
Val	15.45	16.06	16.37	15.48	15.96	16
Met	6.12	6.92	6.88	7.15[c]	6.63	8
Ile	10.72	11.30	11.47	10.09	11.47	12
Leu	21.31	21.10	21.30	20.84	21.23	21
Tyr	5.22	5.24	5.07	1.08	5.17	5
Phe	2.80	2.81	2.72	2.54	2.77	3
His	4.11	4.09	4.07	4.02	4.09	4
Lys	14.23	14.01	14.13	13.57	14.12	15
Arg	10.90	10.92	10.88	10.38	10.89	11
Trp	ND	ND	ND	ND	ND	5

[a] Numbers in columns are compositional values expressed in residues per mole of P22 repressor monomer. ND, not determined. PAO, 24-hydrolysis of performic acid oxidized repressor. [b] Composite values determined by averaging 24-, 48-, and 72-h values except for Thr and Ser values, which were extrapolated to zero time, Ile, which was taken from the 72-h analysis, and Cys, which was taken from the cysteic acid value of the 24-h hydrolysis of performic acid oxidized repressor. [c] As methionine sulfone. [d] As cysteic acid.

molecular weight of the P22 repressor. We presume that this discrepancy reflects the fact that the *c*2 repressor binds a nonaverage amount of $NaDodSO_4$ and thus migrates anomalously in electrophoresis.

Discussion

A comparison of the P22 repressor sequence and the *c*2 gene DNA sequence suggests that maturation of the repressor protein requires only removal of the formyl group from the initiator formylmethionine. No protein processing appears to occur at either terminus of the repressor since the protein starts with a methionine and terminates with a proline whose codon precedes an ochre termination codon in the gene sequence.

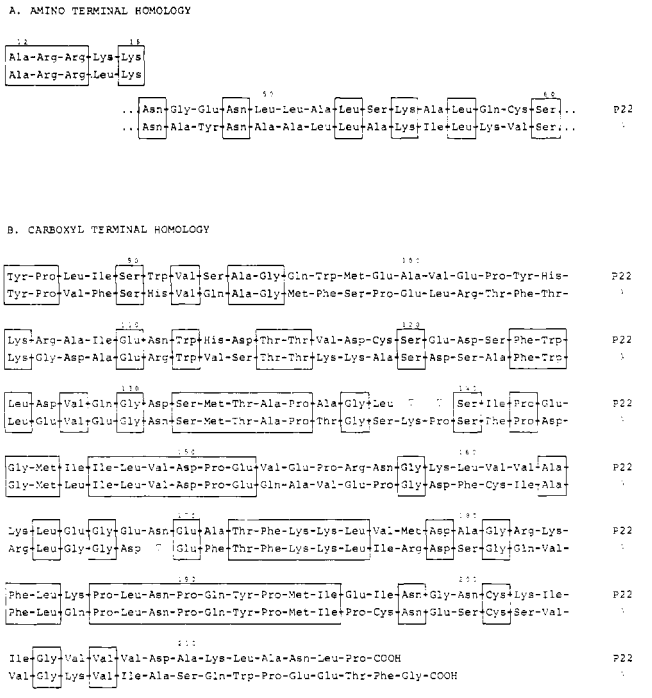

FIGURE 3: Comparison of the amino acid sequences of the P22 *c*2 repressor and the phage λ *c*I repressor. Numbers refer to residue positions in the P22 repressor only.

In lysogens, *c*2 repressor is synthesized from a maintenance promoter located immediately adjacent to the *c*2 gene. Messenger RNA initiated in vitro at this promoter has been sequenced (A. R. Poteete, unpublished results). This mRNA initiates four bases from the AUG coding for Met[1] in our protein and DNA sequence. Thus, P22 *c*2 maintenance mRNA, like λ *c*I maintenance mRNA, does not contain an extensive leader sequence (Ptashne et al., 1976) nor is there evidence for a strong ribosome binding site in either mRNA. In the case of λ repressor, it has been argued that this lack of a strong mRNA ribosome binding site results in lower but more uniform synthesis of repressor protein in lysogens and thus in more stable steady-state repressor concentrations [R. Schlief, unpublished results, cited in Ptashne et al. (1976)].

As discussed in the introduction, the *c*2 and *c*I repressor proteins of phages P22 and λ are highly analogous in function and in their mechanism of action. Yet, each specifically binds to its own operators, which although similar in overall structure are completely distinct in DNA sequence. In the λ *c*I repressor, the amino acids which mediate specific operator recognition are contained within the amino-terminal 92 residues of the protein, and the first 30 residues of this region are highly basic (Sauer et al., 1979; Sauer & Anderegg, 1978). The amino-terminal region of the P22 *c*2 repressor is also involved in operator DNA binding (J. DeAnda, A. R. Poteete, and R. T. Sauer, unpublished results), and contains regions rich in basic amino acids (Figure 2). The most striking of these sequences includes residues 9–20 and contains five arginine and three lysine residues. Residues 12–16 of this region are partially homologous to residues 15–19 of the λ *c*I repressor. Residues 46–60 of P22 repressor also show partial homology to residues 58–72 of λ repressor. The alignment of these sequences is shown in Figure 3A. Apart, however, from these

two regions of partial homology, the amino-terminal sequences of the P22 and repressors show little overall similarity.

Both λ and P22 repressors mediate positive control of their own synthesis. In λ repressor, this positive control function resides in the amino-terminal domain of the protein (Sauer et al., 1979), and it is likely that the P22 repressor is similar in this respect. One model for positive control is that the repressor makes a protein–protein contact with the RNA polymerase molecule (Ptashne et al., 1980). If this model is correct, then it is plausible that both λ and P22 repressors make similar contacts with polymerase. In this case, some elements of the small amino-terminal homologies shown in Figure 3A might serve as a polymerase recognition sequence.

In contrast to the lack of overall homology between amino-terminal sequences, there is striking homology between the carboxyl-terminal regions of the P22 and λ repressors. These two sequences have been aligned in Figure 3B where two small deletion–insertion events have been included to maximize the homology. As shown in Figure 3, about 45% of the carboxyl-terminal residues of the two repressors are homologous. This homology can be increased slightly if additional insertion–deletion events are allowed or if acidic, basic, hydrophobic, or aromatic amino acid pairs are counted as homologous. The amino acid sequence homology between the P22 and λ repressors is greater than the homology between their respective genes; even in regions of amino acid homology, the nucleotide sequences diverge. For example, residues 132–136, 147–152, 172–176, and 187–195 of the P22 repressor are completely homologous to λ repressor sequences. However, a comparison of the nucleotide sequences coding for these homologous residues reveals 16 silent base changes between the 25 codons in each gene. This finding is consistent with previous studies which show that divergence of nucleic acid coding sequences is often greater than divergence of the protein sequences themselves (Nichols & Yanofsky, 1979).

The carboxyl-terminal homology between the P22 and λ repressors suggests that these proteins share a common ancestral origin and further suggests that these common sequences play an important role in some function or functions of these repressor proteins. The exact nature of these conserved functions is presently unclear. We know that the carboxyl-terminal domains of the λ and P22 repressors mediate subunit oligomerization, but these domains may also serve other functions such as recognition of antirepressors, the *rec*A protease, or other macromolecules. Clearly, further studies are required before specific residues or regions of the carboxyl-terminal repressor sequences can be assigned particular functions.

The overall lack of amino-terminal homology between the P22 and λ repressors is somewhat surprising given their carboxyl-terminal homologies. One possibility is that these sequences have not arisen by divergence from a common sequence but rather by recombination events which serve to join different DNA binding proteins or domains to common or similar oligomerization domains. If the amino-terminal sequences of these heteroimmune repressors have arisen by divergence, then the rate of this divergence is certainly greater than that of the carboxyl-terminal sequences of these proteins. One can imagine that there might be positive selective pressures for the establishment of distinct heteroimmunity and that such heteroimmunity requires rather drastic changes in the protein structure and not simply the change of a few residues which make critical contacts with the operator DNA. Our results do not imply that there are not common structural features shared by the amino-terminal domains of the P22 and λ repressors but merely that such common structures, if present, are not reflected in the primary sequences of the two proteins. Alternatively, there may simply be fewer constraints on the actual tertiary structure of the amino-terminal domains of these proteins so that many mutational events are not selected against. By this model, we would assume that the three-dimensional folding of the two heteroimmune DNA binding domains could actually be quite different.

The carboxyl-terminal homology between the P22 and λ repressors indicates that certain conserved sequences play an important role in repressor function. The λ and 434 repressors are also likely to share common carboxyl-terminal sequences since the genes of these repressors form a small heteroduplex in a region encoding carboxyl-terminal sequences of both proteins (Westmoreland et al., 1969). We anticipate that many of the lambdoid phage repressors will share amino acid sequence homology and that a study of these homologies will provide important insights into structure–function relationships of this family of proteins.

Acknowledgments

We thank Linda Capuno for help with the amino acid analysis and Patty Rich for help in preparing this manuscript. We also thank Mark Ptashne in whose laboratory some of the experiments described here were performed.

References

Ballivet, M., & Eisen, H. (1978) *Eur. J. Biochem. 82*, 175.
Ballivet, M., Reichardt, L. F., & Eisen, H. (1979) *Eur. J. Biochem. 73*, 601.
Botstein, D., & Herskowitz, I. (1974) *Nature (London) 251*, 584.
Brauer, A. W., Margolies, M. N., & Haber, E. (1975) *Biochemistry 14*, 3029.
Bronson, M. J., & Levine, M. (1972) *Virology 47*, 644.
Chadwick, P., Pirrotta, V., Steinberg, R., Hopkins, N., & Ptashne, M. (1970) *Cold Spring Harbor Symp. Quant. Biol. 35*, 283.
Chisholm, R. L., Deans, R. J., Jackson, E. N., Jackson, D. A., & Rutila, J. E. (1980) *Virology 102*, 172.
Edman, P. (1960) *Ann. N.Y. Acad. Sci. 88*, 602.
Ilse, D., & Edman, P. (1963) *Aust. J. Chem. 16*, 411.
Levine, M. (1957) *Virology 3*, 22.
Levine, M., & Smith, H. O. (1964) *Science (Washington, D.C.) 146*, 1581.
Maxam, A. M., & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 560.
Maxam, A. M., & Gilbert, W. (1980) *Methods Enzymol. 65*, 499.
Nichols, B. P., & Yanofsky, C. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 5244.
Pabo, C. O., Sauer, R. T., Sturtevant, J., & Ptashne, M. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 1608.
Phizicky, E. M., & Roberts, J. (1980) *J. Mol. Biol. 139*, 319.
Pirrotta, V., Chadwick, P., & Ptashne, M. (1970) *Nature (London) 227*, 41.
Poteete, A. R., Ptashne, M., Ballivet, M., & Eisen, H. (1980) *J. Mol. Biol. 137*, 81.
Ptashne, M., Backman, K., Humayun, M. Z., Jeffery, A., Maurer, R., Meyer, B., & Sauer, R. T. (1976) *Science (Washington, D.C.) 194*, 156.
Ptashne, M., Jeffrey, A., Johnson, A. D., Maurer, R., Meyer, B. J., Pabo, C. O., Roberts, T. M., & Sauer, R. T. (1980) *Cell (Cambridge, Mass.) 19*, 1.
Roberts, J., Roberts, C., & Craig, N. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 4714.

Sauer, R. T., & Anderegg, R. (1978) *Biochemistry 17*, 1092.
Sauer, R. T., Niall, H. D., Hogan, M. L., Keutmann, H. T., O'Riordan, J. L. H., & Potts, J. T., Jr. (1974) *Biochemistry 13*, 1994.
Sauer, R. T., Pabo, C. O., Meyer, B., Ptashne, M., & Backman, K. C. (1979) *Nature (London) 279*, 396.
Susskind, M. M., & Botstein, D. (1975) *J. Mol. Biol. 98*, 413.
Tarr, G. E., Beecher, J. F., Bell, M., & McKean, D. J. (1978) *Anal. Biochem. 84*, 622.
Udenfriend, S., Stein, S., Bohlen, P., Dairman, W., Leimgruber, W., & Weigele, M. (1972) *Science (Washington, D.C.) 178*, 871.
Westmoreland, B. C., Szybalski, W., & Ris, H. (1969) *Science (Washington, D.C.) 163*, 1343.

Single-Strand Deoxyribonucleic Acid Binding Protein from Rat Liver Changes the Helical Structure of Deoxyribonucleic Acid†

Michel Duguet,* Catherine Bonne, and Anne-Marie de Recondo

ABSTRACT: Incubation of rat liver single-strand DNA binding protein S25 with covalently closed relaxed SV40 DNA in the presence of rat liver topoisomerase I induced a decrease in the linking number L_K of DNA, so that it appeared negatively supertwisted after removal of the protein. ΔL_K was found to be a linear function of protein to DNA ratio and reached a plateau corresponding to about 15 superhelical turns. The same result was obtained when S25 was incubated with form I or form Ir before addition of topoisomerase I or when SV40 was replaced by PM2 DNA. The observed reduction in the linking number of DNA when it is closed in the presence of rat liver protein S25 can be explained either by supercoiling of DNA induced by S25 or by detorsion or unwinding of DNA.

A number of single-strand DNA binding proteins (SSB)[1] isolated from eukaryotes and virus-infected eukaryotes as well as from prokaryotes or phage-infected prokaryotes have been studied today or are under investigation [for review, see Champoux (1978); Coleman & Oakley, 1980)]. For some of these proteins, their in vivo role in replication or in genetic recombination has been firmly established (Alberts, 1970; Sigal et al., 1972; Cavalieri et al., 1976; Meyer et al., 1979). For others it was only suggested (Banks & Spanos, 1975; Herrick et al., 1976). Some are melting proteins, i.e., are able to lower the melting point of DNA (Sigal et al., 1972; Kelly et al., 1976); some, in contrast, stabilize the DNA (Fowlkes et al., 1979); others have no reported effect on the melting point (Novak & Baril, 1978). Some bind cooperatively; others do not. The way by which each of these proteins binds to DNA and "shapes" it thus appears original. For instance, T_4 gene 32 protein binds cooperatively to single-stranded DNA and puts the strand in an extended and rigid conformation (Alberts, 1970), while gene 5 protein of fd binds by wrapping of single-stranded DNA around protein dimers in rodlike structures (Pratt et al., 1974). Little is known on the way by which eukaryotic proteins bind to DNA. Calf thymus single-strand binding proteins (or HDP) were reported to bind noncooperatively to ss-DNA and to bring it in an extended conformation (Herrick et al., 1976), while *Ustilago maydis* binding protein binds cooperatively to ss-DNA (Banks & Spanos, 1975).[2]

In previous works, we have isolated and studied some properties of a single-strand binding protein from rat liver. The protein isolated from regenerating liver, called HD25, was able both to lower the melting point of polyd(A-T) (Duguet & de Recondo, 1978) and to stimulate rat liver DNA polymerases (Duguet et al., 1977). These results supported the idea that it could be involved in DNA replication. In contrast, the protein from normal rat liver called S25, although it is a single-strand binding protein, was not found to be a melting protein and, in some conditions, inhibits homologous DNA polymerases (Bonne et al., 1979; de Recondo et al., 1980). At this time, the only difference found between these proteins is the ability of HD25 to oligomerize, a property that is not shared by S25. Otherwise, HD25 and S25 proteins are undistinguishable, even by the pattern of proteolytic degradation (Bonne et al., 1979; Bonne et al., 1980).

In search for a possible function of S25, we decided to look more carefully at the way it binds to DNA. Using competition experiments with ^{32}P-labeled M13 single-stranded DNA and various relaxed double-stranded DNAs, we have estimated that, at low ionic strength, S25 binds about 1000-fold better on single-stranded DNA (Bonne et al., 1980). We also know from previous studies that both HD25 and S25 bind to SV40 DNA; this was followed both by sedimentation analysis and by nitrocellulose filter assay. The use of a covalently closed circular DNA like SV40 or PM2 provides a very sensitive probe to test any change in the DNA helix parameters that is immediately reflected in the shape of the molecule and consequently in its electrophoretic mobility (Germond et al., 1975).

† From the Unité d'Enzymologie, Institut de Recherches Scientifiques sur le Cancer, 94800 Villejuif, France. *Received September 18, 1980.* This work was supported by grants from the Délégation à la Recherche Scientifique et Technique (1 650 1524 and A 650 1631).

[1] Abbreviations used: SSB, single-strand DNA binding protein; $NaDodSO_4$, sodium dodecyl sulfate; EtBr, ethidium bromide; HDP, helix-destabilizing protein; DTT, dithiothreitol; ss-DNA, single-stranded DNA; Tris, tris(hydroxymethyl)aminomethane; EDTA, ethylenediaminetetraacetic acid.

[2] Both are reported to be melting proteins. In contrast, the 72K protein from adenovirus, although it binds specifically to ss-DNA, is not a melting protein (Fowlkes et al., 1979).

0006-2960/81/0420-3598$01.25/0