Stephenson, E. W. (1981) *Am. J. Physiol. 240*, C1–C19.
Tada, M., Yamamoto, T., & Tonomura, Y. (1978) *Physiol. Rev. 58*, 1–79.
Tada, M., Yamada, M., Kadoma, M., Inui, M., & Ohmori, F. (1982) *Mol. Cell. Biochem. 46*, 73–95.
Tuana, B. S., & MacLennan, D. H. (1984) *J. Biol. Chem. 259*, 6979–6983.
Volpe, P., Costello, B., Chu, A., & Fleischer, S. (1984) *Arch. Biochem. Biophys. 233*, 174–179.
Winegrad, S. (1982) *Annu. Rev. Physiol. 44*, 451–462.
Yagi, K., Yazawa, M., Kakiuchi, S., Ohshima, M., & Uenishi, K. (1978) *J. Biol. Chem. 253*, 1338–1340.
Yamamoto, N., & Kasai, M. (1982) *J. Biochem. (Tokyo) 92*, 485–496.

Bacteriophage P22 Cro Protein: Sequence, Purification, and Properties†

Anthony R. Poteete*
Department of Molecular Genetics and Microbiology, University of Massachusetts Medical School, Worcester, Massachusetts 01605
Kathleen Hehir and Robert T. Sauer
Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139
Received June 26, 1985

ABSTRACT: The DNA sequence of part of the bacteriophage P22 early regulatory region, including genes *cro* and *c*1, was determined. The protein product of the *cro* gene consists of 61 amino acid residues, and that of *c*1, 92 amino acid residues. Both genes were placed separately in plasmids from which they are expressed from a controllable promoter in vivo. Induced cells bearing the *cro*-expressing plasmid were used as a source for purifying and characterizing the Cro protein. The amino-terminal sequence of this protein was found to be as predicted by the DNA sequence; close agreement was also observed between its predicted and experimentally determined amino acid composition and molar extinction coefficient at 280 nm. In gel filtration experiments, Cro protein at concentrations around 10^{-5} M appears to have a molecular weight of 8600, which is more consistent with monomers (6800) than with dimers (13 600). Cro protein binds specifically to the three repressor binding sites in the P22 right operator; in order of decreasing affinity, these are O_R3, O_R1, and O_R2.

Salmonella phage P22 has a regulatory gene called *cro* that is responsible for turning down synthesis of early phage genes during infection (Winston & Botstein, 1981). Phages that are defective in *cro* function are unable to grow lytically; upon infection of a normal host cell, they overproduce repressor (the product of gene *c*2) and are channeled into lysogeny. The *cro* gene is located between *c*2 and another regulatory gene, *c*1. The *c*1 gene has a role opposite to that of *cro*: it is required for the establishment of lysogeny, though not for lytic growth (Levine, 1957).

Studies of λ, a related phage that grows in *Escherichia coli*, have led to a detailed understanding of the mechanisms that regulate expression of its genes [see Hendrix et al. (1983) for a review]. The basic form of this regulation is the same for both λ and P22 (Susskind & Botstein, 1978). The clusters of early regulatory genes and sites of λ and P22 constitute modules that can be exchanged between the two phages by recombination. In particular, both the *cro* and *c*1 genes of P22 have counterparts in λ, called *cro* and *c*II, respectively. Where the mechanisms of action of P22 regulatory elements have been examined in detail, they are strikingly similar to those of their λ counterparts (Poteete & Ptashne, 1982). It is reasonable to expect that P22 *cro* and *c*1 fall into this pattern. We can predict that P22 Cro protein will be found to exert its regulatory effects by binding to the left and right operators of the phage. In particular, it should bind to the *c*2 repressor binding sites of the right operator with O_R3, the left-most site, having a higher affinity than the other two sites. [The rationale for this regulatory scheme in λ and its generalization to P22 have been discussed by Johnson et al. (1981).]

Although we expect that the Cro proteins of λ and P22 function similarly, the two proteins should exhibit completely distinct specificities. The operator sites of the two phages are quite different in sequence (Maniatis et al., 1975; Poteete et al., 1980); P22 *c*2 repressor does not recognize λ operators, and λ *c*I repressor does not recognize P22 operators (Poteete & Roberts, 1981).

The repressor and Cro proteins of λ and its relatives P22 and 434 constitute an interesting group. Studies of these regulatory proteins have led to insights into the structural basis of sequence-specific DNA binding and the mechanisms of both positive and negative regulation of transcription initiation. In this paper, we present the P22 DNA sequence from the early regulatory region, including the *cro* and *c*1 genes. In addition, we describe the construction of plasmids that express *cro* at high levels and the purification and properties of P22 Cro protein.

Materials and Methods

DNA Sequencing. DNA sequencing was performed by using the chemical method of Maxam & Gilbert (1980) and standard procedures for 3′ and 5′ end-labeling, restriction digestion, and polyacrylamide electrophoresis (Maniatis et al., 1982).

Plasmids. General procedures for plasmid construction and purification were as described previously (Poteete & Roberts, 1981; Poteete, 1982). pBR322 (Bolivar et al., 1977) was the

† This research was supported by Grants AI18234 to A.R.P. and AI16892 to R.T.S. from the National Institutes of Health.

0006-2960/86/0425-0251$01.50/0 © 1986 American Chemical Society

vector; *E. coli* W3110 *lacI*^{q} L8 (Brent & Ptashne, 1981) was the host in procedures described below. Plasmid pTP7 (Poteete et al., 1980) was the source of DNA fragments for sequencing as indicated in Figure 1.

Plasmid pTP153 was constructed from three DNA fragments: (1) a 355-base pair (bp) fragment containing *cro*, produced by cutting pTP7 with *Hin*fI, purifying the 405-bp *cro*-containing fragment, filling in its ends, digesting with *Msp*I, and purifying from the smaller fragment generated by *Msp*I digestion; (2) the origin-containing fragment of pBR322 generated by digestion with *Pst*I and *Cla*I; (3) a *Pst*I- and *Pvu*II-ended fragment from plasmid ptac12 (Amann et al., 1983). The junction between the *Pvu*II and filled-in *Hin*fI ends restores the *Hin*fI site. Plasmid pTP154 was constructed by deletion of sequences between the *Ava*I and *Hin*dIII sites of pTP153, with insertion of *Eco*RI linkers between the filled-in ends.

Plasmid pTP180 was constructed by digesting pTP154 with *Fnu*DII, ligating the digestion products in the presence of a greater than 10-fold molar excess of *Eco*RI linkers, digesting to completion with *Eco*RI, purifying the resulting 480-bp P_{tac} and *cro*-containing fragment, and ligating it into the *Eco*RI site of pBR322.

Plasmid pTP36 was constructed by deletion of sequences between two sites of pTP28 (Berget et al., 1983): The *Bam*HI site in gene 9 and the *Pvu*II site in the pBR322 sequence. The junction of the *Pvu*II and filled-in *Bam*HI sites restores the *Bam*HI site.

Plasmid pTP187 was constructed by ligating the *cro*-containing *Eco*RI fragment from pTP180 into the *Eco*RI site of pTP36. The orientation of the insert is such that *cro* is transcribed from P_{tac} toward the terminator t_{ant} contained in pTP36.

Plasmid pTP270 was constructed from three fragments: the origin-containing fragment of pBR322 generated by digestion with *Pst*I and *Bam*HI, the P_{tac}-containing *Pst*I- and *Pvu*II-ended fragment from ptac12, and a 448-bp fragment containing sequences between the *Pvu*II site in *cro* and the *Sau*3a site 14 bp downstream from the termination codon of *c*1.

Protein Purification. *E. coli* strain W3110 *lacI*^{q} L8/pTP187 was grown with aeration at 37 °C, in 10 L of LB broth to mid-log phase (A_{600} = 1.1); 0.57 g of isopropyl β-D-thiogalactopyranoside (IPTG) was added to induce P22 Cro synthesis, and growth was continued for 100 min at 32 °C. Approximately 35–40 g of cells was harvested by centrifugation and was resuspended in 75 mL of a buffer containing 25 mM tris(hydroxymethyl)aminomethane hydrochloride (Tris-HCl) (pH 8), 50 mM KCl, 0.5 mM ethylenediaminetetraacetic acid (EDTA), and 5% glycerol. Except where indicated, all further steps of purification were performed at 4 °C. The resuspended cells were lysed by five 2-min bursts of sonication. After each burst, 0.3 mL of a 20 mg/mL solution of phenylmethanesulfonyl fluoride in ethanol was added. Following lysis, $MgCl_2$ was added to 10 mM, 1 mg each of DNase I and RNase A was added, and the crude lysate was incubated at 37 °C for 20 min. The lysate was then centrifuged at 42000 rpm for 90 min in a Beckman 50Ti rotor, and the pellet fraction was discarded.

The high-speed supernatant (100 mL) was loaded onto a 1.8-L Sephadex G-75 column (5 × 92 cm) equilibrated in 50 mM Tris-HCl (pH 8), 200 mM KCl, 0.5 mM EDTA, 1.4 mM 2-mercaptoethanol, and 5% glycerol. The column was developed with the same buffer at a flow rate of approximately 100 mL/h, and fractions were screened by sodium dodecyl sulfate (SDS)–polyacrylamide gel electrophoresis. Fractions containing a protein of the size expected for P22 Cro were pooled and were extensively dialyzed, with Spectrapor 3 dialysis tubing (Spectrum Medical Industries), against PC buffer, which contains 20 mM Tris-HCl (pH 8), 0.1 mM EDTA, 1.4 mM 2-mercaptoethanol, and 5% glycerol. The dialyzed material was loaded onto a 50-mL Whatman P11 phosphocellulose column (2.5 × 10 cm) equilibrated in PC buffer; the column was washed with 100 mL of PC buffer, and a linear gradient (250 mL each) from PC buffer to PC buffer plus 300 mM NaCl was run at a flow rate of 30 mL/h. Fractions containing P22 Cro were located by SDS gel electrophoresis and were found at the very beginning of the gradient. These fractions were pooled and were dialyzed as described above into PC buffer. The pooled phosphocellulose fractions were loaded onto a 35-mL Bio-Rad Affigel Blue column (2.5 × 7 cm), and the column was eluted with successive 50-mL volumes of PC buffer containing no salt, 10 mM NaCl, 20 mM NaCl, and 40 mM NaCl. P22 Cro eluted from the Affigel Blue column in the PC buffer plus 20 mM NaCl step.

During purification, protein concentrations were estimated by using the Bio-Rad protein assay. Standard curves for this assay were prepared by using hen egg-white lysozyme. SDS–polyacrylamide gel electrophoresis was performed by using the buffer system of Laemmli (1970). The resolving gels contained 13.05% acrylamide and 0.45% *N*,*N*-methylenebis(acrylamide). Proteins were visualized following staining of the gel with Coomassie Brilliant Blue R-250.

Protein Chemistry. Amino acid compositions were determined by using a Durham D500 analyzer following hydrolysis of samples in vacuo for 24 h at 110 °C in 6 N HCl plus 1% phenol. Automated Edman degradation was performed by using a Beckman 890C sequencer and the 0.1 M Quadrol program described by Brauer et al. (1975). Phenylthiohydantoin amino acid derivatives were identified by gas–liquid and high-performance liquid chromatography as described (Sauer et al., 1981).

Determination of Native Molecular Weight. The native molecular weight of P22 Cro was determined by comparison of its elution volume in gel filtration experiments with the elution volume of bovine serum albumin, myoglobin, cytochrome *c*, aprotinin, and sodium azide. Gel filtration chromatography was performed at room temperature on a Sephadex G-50 fine column (1 × 43 cm) equilibrated in 50 mM Tris-HCl (pH 7.5), 200 mM KCl, 0.1 mM EDTA, and 5% glycerol. Samples wre loaded on the column in a volume of 0.5 mL, and the flow rate was maintained at 6 mL/h by using a peristaltic pump. The absorbance at 280 nm of the column effluent was continuously monitored with an LKB Uvicord S monitor.

DNA Binding Assay. Cro-mediated protection of sites in O_R from digestion by DNase I was measured essentially as described previously for *c*2 repressor; the source of O_R-containing DNA fragments was the plasmid pTP37 (Poteete & Ptashne, 1982). Purified, end-labeled DNA at a concentration of approximately 10 nM was incubated with varying concentrations of Cro protein in 20-μL of a buffer that typically contained 10 mM Tris-HCl, pH 7.4, 50 mM KCl, 10 mM $MgCl_2$, and 1 mM $CaCl_2$, for 5–7 min at room temperature. DNase I (Sigma, EP grade, approximately 2.5 pg) was added in a volume of 2 μL, and the reactions were incubated at 37 °C for 15 min. In other experiments, the pH and the salt concentration were varied in order to examine their effects on the affinity of Cro for sites in O_R. Reactions were stopped by addition of 0.22 mL of 70 μg/mL tRNA in 1.7 M am-

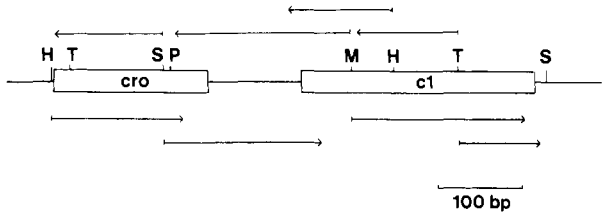

FIGURE 1: DNA sequencing strategy. Arrows show the direction and extent of sequence information obtained from specified fragments. Those above the map represent the strand shown in Figure 2; those below, the other strand. Abbreviations: H, *Hin*fI; T, *Taq*I; S, *Sau*3a; P, *Pvu*II; A, *Alu*I. The location of the *Sau*3a site to the right of *c*1 is from the sequence determined by Kroger and Hobom (personal communication).

monium acetate. DNA was precipitated with ethanol and then redissolved in formamide-containing sample buffer (Maxam & Gilbert, 1980) for electrophoresis.

Results

DNA Sequence of the P22 cro and c1 Genes. The DNA sequence in the region of the P22 *cro* gene was determined by using the sequencing strategy shown in Figure 1. This DNA sequence and the protein sequences encoded by two large open-reading frames are shown in Figure 2. The first open-reading frame, which encodes a protein of 61 amino acids, is in the position expected for the P22 *cro* gene. The second open-reading frame, which encodes a protein of 92 amino acids, is in the position expected for the P22 *c*1 gene.

Construction of Cro and c1-Producing Plasmids. The initial strategy for expressing the P22 *cro* gene at high levels from a plasmid was to fuse the strong inducible promoter P_{tac} to the *cro* gene. A *Hin*fI site located just upstream from *cro* was employed in the plasmid construction. When filled in and fused to a *Pvu*II ended fragment containing the tac promoter from ptac12 (Amann et al., 1983), the *Hin*fI-ended fragment should generate the following sequence at the junction:

CAC*AGGA*AACAG AGTCT*ATG*TACAAG

Such a sequence constitutes a hybrid ribosome binding site of the type that has been used in the construction of plasmids that express λ *cro* at high levels (Roberts et al., 1979). The resulting plasmid, pTP153, contains the tac promoter, the putative hybrid ribosome binding site, the *cro-c1* intercistronic region, and part of *c*1. When fully induced with IPTG, cells bearing this plasmid exhibit a reduced rate of growth as well as immunity to infection by λ*imm*P22 *c*2-5, though not by λ bearing its own operators. Slow growth and immunity are IPTG dependent (data not shown). Preliminary attempts to purify Cro from cells bearing pTP153 were hampered, though, because the production of Cro could not be detected either by SDS gel electrophoresis of whole cell extracts followed by staining with Coomassie Blue or by DNase protection assays of whole cell extracts. That the plasmid produced some Cro protein was verified by specific protection of O_R3 sequences by concentrated fractions obtained following gel filtration of extracts of the induced plasmid-bearing cells (not shown).

In an attempt to increase the level of Cro production, most of the *cro-c*1 intercistronic region was removed. The resulting plasmid, pTP180, is otherwise similar to pTP153; its immunity properties suggest that it makes less Cro (not shown). This

```
  Asn Met
5'ATT CAT AGTTAAGTCATCTTAAATAAACTTGACTAAAGATTCCTTTAGTAGATAATTTAAGTGTTCTTTAATTTC
     c2        OR3                  OR2                  OR1

          Met Tyr Lys Lys Asp Val Ile Asp His Phe Gly Thr Gln Arg Ala Val Ala
GGAGCGAGTCT ATG TAC AAG AAA GAT GTT ATC GAC CAC TTC GGA ACC CAG CGT GCA GTA GCT
          cro

Lys Ala Leu Gly Ile Ser Asp Ala Ala Val Ser Gln Trp Lys Glu Val Ile Pro Glu Lys
AAG GCT TTA GGC ATT AGC GAT GCA GCG GTC TCT CAG TGG AAG GAA GTT ATC CCA GAG AAA

Asp Ala Tyr Arg Leu Glu Ile Val Thr Ala Gly Ala Leu Lys Tyr Gln Glu Asn Ala Tyr
GAC GCA TAC CGA TTA GAG ATC GTT ACA GCT GGC GCC CTG AAG TAC CAA GAA AAC GCT TAT

Arg Gln Ala Ala
CGC CAA GCG GCG TAA GCAAAACGCTCTTTACCAATCTGAACCGCCGACAACGCGGTAAACCTATTTCAAAGCGC
                                          nutR?

                                                Met Glu Leu Thr Ser Thr Arg Lys
ATCAACGAATGCGCACAACTAACTATTAACTACAGGAATGTTCACAT ATG GAA CTC ACA AGC ACT CGC AAG
  tR1?                                          c1

Lys Ala Asn Ala Ile Thr Ser Ser Ile Leu Asn Arg Ile Ala Ile Arg Gly Gln Arg Lys
AAA GCC AAC GCA ATT ACC AGC AGC ATC CTT AAC CGG ATA GCT ATT CGT GGA CAG CGT AAA

Val Ala Asp Ala Leu Gly Ile Asn Glu Ser Gln Ile Ser Arg Trp Lys Gly Asp Phe Ile
GTC GCT GAT GCG TTA GGC ATT AAC GAA TCT CAA ATT TCA CGA TGG AAA GGC GAT TTC ATT

Pro Lys Met Gly Met Leu Leu Ala Val Leu Glu Trp Gly Val Glu Asp Glu Glu Leu Ala
CCG AAG ATG GGG ATG TTA TTG GCG GTT CTG GAG TGG GGT GTC GAG GAT GAG GAG TTG GCA

Glu Leu Ala Lys Lys Val Ala His Leu Leu Thr Lys Glu Lys Pro Gln Asp Cys Gly Asn
GAA CTG GCA AAG AAA GTT GCG CAT CTG CTG ACA AAA GAA AAG CCT CAA GAC TGC GGG AAC

Ser Phe Glu Ala
AGT TTT GAG GCC TGA TGTAGAA 3'
```

FIGURE 2: DNA sequence of the P22 *cro-c*1 region.

Table I: Purification of Cro Protein

	volume (mL)	total protein (mg)
crude lysate	120	6600
high-speed supernatant	100	3000
G-75 pool	160	108
PC pool	135	23
Affigel pool	135	12

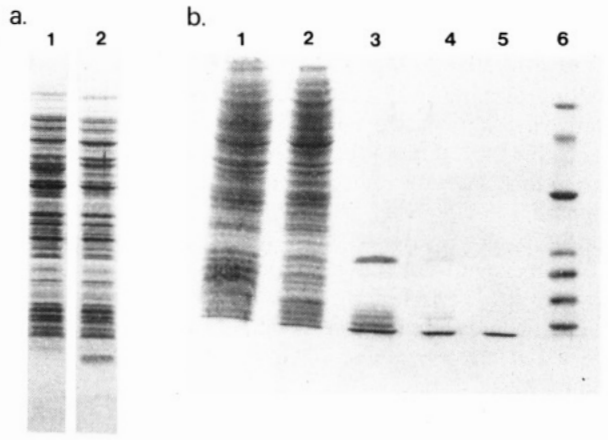

FIGURE 3: Production and purification of P22 Cro protein. (a) Plasmid-bearing cells were induced with IPTG as described under Materials and Methods, concentrated, and lysed by heating in SDS-containing sample buffer; the lysates were subjected to electrophoresis in a gel with a linear gradient of 10–20% acrylamide and stained with Coomassie Blue. Lane 1, pBR322 control; lane 2, pTP187. (b) Samples taken at various stages in the purification were subjected to electrophoresis in a 13.5% acrylamide gel. Lane 1, 1.2 μL of crude lysate; lane 2, 2 μL of high-speed supernatant; lane 3, 15 μL of Sephedex G-75 pool; lane 4, 25 μL of phosphocellulose pool; lane 5, 25 μL of Affigel pool; lane 6, standards: bovine serum albumin (M_r 68 000), ovalbumin (M_r 44 000), carbonic anhydrase (M_r 29 000), soybean trypsin inhibitor (M_r 21 000), myoglobin (M_r 17 000), cytochrome *c* (M_r 12 500), and aprotinin (M_r 6500).

result prompted construction of a third Cro-producing plasmid, pTP187, in which the tac promoter–hybrid ribosome binding site-*cro* fusion, lacking the *cro*-*c*1 intercistronic sequences, was placed upstream from the strong, ρ-independent terminator t_{ant} (Berget et al., 1983). Cells bearing pTP187 produce enough Cro when induced to generate a visible band in Coomassie Blue stained SDS gels of whole cell extracts (see below).

A *c*1-expressing plasmid, pTP270, was constructed by fusing the tac promoter to a fragment of DNA consisting of sequences between the *Pvu*II site in *cro* and the *Sau*3a site 14 bp downstream from the termination codon of *c*1. That this plasmid expresses *c*1 is shown by a complementation test: on *Salmonella typhimurium* bearing both pTP270 and a *lac*I-expressing plasmid, P22 $c1^-$ forms turbid (phenotypically wild-type) plaques, while P22 $c2^-$ forms clear plaques (not shown).

P22 Cro Purification. The purification of P22 Cro protein from IPTG-induced cultures of *E. coli* strain W3110 *lac*I^q/pTP187 is summarized in Table I. Since we could not directly assay P22 Cro operator binding activity in crude lysates, SDS gel electrophoresis was used to monitor the purification procedure. Figure 3 shows SDS gel electrophoretic assays at several steps during the purification. The major purification of P22 Cro was achieved in the Sephadex G-75 gel filtration step, where Cro chromatographed as one of the smallest cellular proteins. Further purification was achieved by cation-exchange chromatography on phosphocellulose and Affigel Blue columns. At pH 6–8, P22 Cro will only bind to these negatively charged columns if it is applied in the absence of salt using buffers of low ionic strength, and it elutes from these columns at ionic strengths corresponding to 10–20 mM NaCl. (P22 Cro flows through a positively charged Sephadex QAE column even when applied at very low ionic strengths; anion-exchange chromatography is thus not apparently useful as a purification step.) The behavior of P22 Cro on cation-exchange columns is somewhat surprising since most DNA binding proteins bind quite tightly to phosphocellulose and require ionic strengths corresponding to 200–500 mM NaCl for their elution.

Table II: N-Terminal Edman Degradation of Purified P22 Cro Protein

step	1	2	3	4	5
residue	Met	Tyr	Lys	Lys	Asp
yield (nmol)	2.5	2.2	3.1	3.4	2.2

Table III: Amino Acid Composition of Purified P22 Cro Protein

amino acid	found[a]	expected
Asx	4.89	5
Thr	1.82	2
Ser	1.94	2
Glx	7.80	8
Pro	0.83	1
Gly	ND[b]	3
Ala	10.63	11
Val	4.23[c]	5
Met	0.88	1
Ile	3.08[c]	4
Leu	3.06	3
Tyr	3.89	4
Phe	1.01	1
His	1.02	1
Lys	6.01	6
Arg	3.25	3
Trp	ND	1
Cys	ND	0
		total 61

[a] Experimental values are composition averages of three 24-h hydrolyses. [b] ND, value not determined. [c] The values for Val and Ile are presumed to be low because of slow hydrolysis of the three Val–Ile or Ile–Val peptide bonds in the sequence.

Amino Acid Sequence and Composition. The DNA sequence of the P22 Cro gene predicts an N-terminal protein sequence Met-Tyr-Lys-Lys-Asp. Automated Edman degradation of 5 nmol of purified P22 Cro protein confirmed this N-terminal sequence as shown in Table II. The amino acid composition of purified P22 Cro is also in good agreement with that predicted from the gene sequence (Table III). Overall, the sequence and composition data suggest that the formyl group of the initiator fMet is removed following translation but that no further N-terminal or C-terminal processing of the protein occurs.

The molar extinction coefficient at 280 nm for P22 Cro (at pH 8) was calculated by using amino acid analysis to determine the molar concentration of protein in a sample of known absorbance. This experimentally determined value is 11 350 $(M \cdot cm)^{-1}$. This value is close to the value of 11 200 $(M \cdot cm)^{-1}$ expected on the basis of the molar extinction coefficients for the four tyrosines and one tryptophan present in the P22 Cro sequence.

Native Molecular Weight. The subunit molecular weight of P22 Cro calculated from its protein sequence is 6800. In gel filtration experiments on Sephadex G-50, P22 Cro chromatographs at a position expected for an approximately spherical protein of molecular weight 8600. These data are shown in Figure 4 for chromatography performed at pH 7.5 and 200 mM KCl. Essentially identical results were obtained when the chromatography was repeated at a KCl concentration of 2 M. These data indicate that P22 Cro is substantially

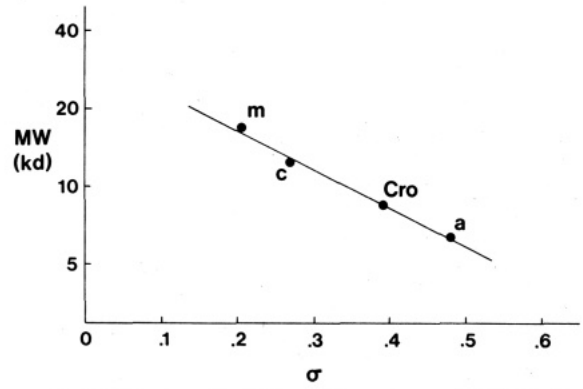

FIGURE 4: Gel filtration of P22 Cro and protein standards in Sephadex G-50. In these experiments, bovine serum albumin and sodium azide were used as void volume (V_0) and salt volume (V_s) markers, respectively. σ values were calculated from protein elution volumes (V_e) by the equation $\sigma = (V_e - V_0)/(V_s - V_0)$. Abbreviations: m, myoglobin; c, cytochrome *c*; a, aprotinin.

monomeric at the concentrations used for chromatography. In the experiment shown in Figure 4, P22 Cro was loaded at an initial concentration of 27 μM and was present at a concentration of 2.8 μM in the peak fraction. Thus, if monomers and dimers of P22 Cro are in equilibrium (see below), the equilibrium dissociation constant for dimerization must be 10 μM or greater.

Operator DNA Binding. Cro protein binds specifically to the repressor binding sites in the P22 right operator O_R, as shown in Figure 5. In this experiment, an O_R-containing DNA fragment labeled with ^{32}P at one end of one strand was incubated with varying amounts of Cro and then subjected to partial digestion with DNase I. The products of the reaction were denatured and subjected to electrophoresis in a urea-containing polyacrylamide gel, which was dried and exposed to film. With increasing concentrations of Cro, first O_R3, then O_R1, and then O_R2 are protected from digestion. The concentrations of Cro required to obtain half-maximal protection are approximately 1.3 μM for O_R3, 2 μM for O_R1, and 5 μM for O_R2. In the course of examining the effects of varying conditions of the DNase protection assay, we have observed half-maximal protection of operator sites at Cro concentrations as low as 10-fold less than these by lowering the pH and Mg^{2+} concentration (not shown).

Discussion

Sequences of the P22 cro and c1 Genes. The open-reading frames shown in Figure 2 are the *cro* and *c*1 genes of P22, as shown by a number of observations. (1) A small fragment of DNA including the *c*1 gene, and no other open-reading frames of significant length, can be induced to complement P22 *c*1$^-$, enabling it to form turbid plaques. (2) The P22 *cro* gene maps between O_R and *c*1 (Winston & Botstein, 1981), as does the open-reading frame designated *cro* in Figure 2. (3) The two genes occupy positions analogous to those of their λ counterparts. (4) The deduced amino acid sequences of the proteins encoded by the genes exhibit limited homology with other site-specific DNA binding proteins. The functional significance of homologies among these proteins (including P22 Cro and c1 proteins) has been discussed previously (Sauer et al., 1982).

The *cro-c*1 intercistronic region of P22 can be expected, on the basis of the analogy with phage λ, to contain a variety of regulatory sites. The sequence shown in Figure 2 contains elements that resemble the nut_R, t_R1, and P_{RE} sites of λ (Rosenberg et al., 1978). While this work was in progress,

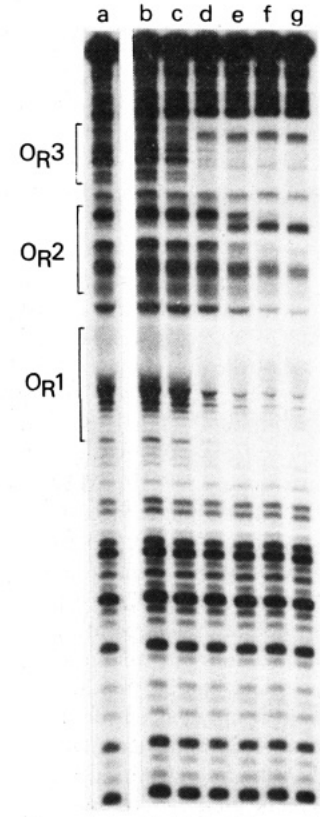

FIGURE 5: Protection of sites in O_R by Cro protein. Bands corresponding to the sites O_R1, O_R2, and O_R3 are identified on the basis of previous experiments (Poteete & Ptashne, 1982). Amounts of Cro protein included in the reactions are (a) 0, (b) 0.6, (c) 1.3, (d) 2.5, (e) 5, (f) 10, and (g) 20 μM.

Backhaus & Petri (1984) published the sequence of a segment of P22 DNA extending about 3000 bp to the right from the *Pvu*II site in *cro*. Their sequence agrees with ours in the area of overlap.

Properties of the Cro Protein. As predicted, Cro binds specifically to the repressor binding sites in O_R, showing the highest affinity for O_R3. Presumably, binding to O_R3 inhibits transcription from P_{RM}, while binding to the other sites inhibits transcription from P_R, as is seen with *c*2 repressor (Poteete & Ptashne, 1982).

The Cro protein of P22 resembles that of λ: it is a small protein, encoded by a gene immediately to the right of O_R, that binds to sites in O_R. It differs, though, in two possibly significant ways (other than in the sequence specificity of its binding). First, P22 Cro appears to exist as a monomer in solution at concentrations up to 10 μM. λ Cro is generally regarded as a stable dimer (Takeda et al., 1977), although Boschelli (1982) has reported λ Cro may undergo a monomer–dimer transition in the micromolar concentration range. Second, λ Cro binds more strongly to its operator sites in vitro than does P22 Cro. For example, half-protection of λ O_R3 is observed at a λ Cro concentration of about 3×10^{-9} M (Johnson et al., 1979) whereas comparable protection of P22 O_R3 by P22 Cro requires concentrations of $(1 - 10) \times 10^{-7}$ M. Since both Cro proteins presumably bind to operator DNA as dimers, λ Cro may exhibit stronger operator binding simply because it forms a more stable dimer. We cannot rule out the possibility that the relatively weak binding that we observe for P22 Cro is due to inactivation during purification. The hypothesized inactivation would have to have occurred at an early stage of the purification, though, as at least 50% of the activity in the G-75 pool (see Materials and Methods) was recovered in the final preparation (not shown). One possible

explanation for the low affinity of P22 Cro, that the protein we characterized is a mutant variant, has been ruled out. We directly determined the DNA sequences of the *cro* genes borne by the Cro-producing plasmid pTP187 and by wild-type P22 phage and found them to be identical (data not shown).

It should be noted that the possibility that P22 Cro has a lower affinity for its binding sites in vivo than λ Cro has for its sites does not imply to a qualitatively different mechanism of gene regulation. P22 Cro would simply have to be synthesized to a higher concentration in order to exert its effects. Alternatively, it may be that P22 Cro can be effective at the same concentration as λ Cro, in spite of its weaker operator binding. This would be the case if much of the λ Cro, but little of the P22 Cro, present in a cell were bound nonspecifically to random DNA sequences. The dissociation constant characterizing the nonspecific binding of λ Cro to DNA has been measured as approximately 5×10^{-5} M (Boschelli, 1982). If we assume that the intracellular concentration of nonspecific Cro binding sites is about 5×10^{-3} M (one binding site per base pair, 3×10^6 bp per cell in an intracellular volume of 10^{-15} L), then we can estimate that 99% of the λ Cro in a cell at any time would be bound to nonoperator sites. We have not measured the nonspecific DNA binding affinity of P22 Cro, but the weakness of its binding to cation exchangers (see above) suggests that it may be considerably lower than that of λ Cro. If both molecules have the same ratio of specific to nonspecific DNA binding affinities, then we estimate that a negligible fraction of the P22 Cro in a cell would be bound to nonoperator sites. This would mean that, for a given intracellular level, P22 Cro would be present at a 100-fold higher effective concentration than λ Cro, compensating for its lower site-specific DNA binding affinity.

ACKNOWLEDGMENTS

We thank Drs. Manfred Kroger and Gerd Hobom, who sent us P22 DNA sequences to the right of those shown in Figure 2 before publication.

REFERENCES

Amann, E., Brosius, J., & Ptashne, M. (1983) *Gene 25*, 167–178.

Backhaus, H., & Petri, J. B. (1984) *Gene 32*, 289–303.

Berget, P. B., Poteete, A. R., & Sauer, R. T. (1983) *J. Mol. Biol. 164*, 561–572.

Bolivar, F., Rodriguez, R. L., Greene, P. J., Betlach, M. C., Heyneker, H. L., Boyer, H. W., Crosa, J. H., & Falkow, S. (1977) *Gene 2*, 95–113.

Boschelli, F. (1982) *J. Mol. Biol. 162*, 267–282.

Brauer, A. W., Margolies, M. N., & Haber, E. (1975) *Biochemistry 14*, 3029–3035.

Brent, R., & Ptashne, M. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 4204–4208.

Hendrix, R. W., Roberts, J. W., Stahl, F. W., & Weisberg, R. A., Eds. (1983) *Lambda II*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Johnson, A. D., Meyer, B. J., & Ptashne, M. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 5061–5065.

Johnson, A. D., Poteete, A. R., Lauer, G., Sauer, R. T., Ackers, G. K., & Ptashne, M. (1981) *Nature (London) 294*, 217–223.

Laemmli, U. K. (1970) *Nature (London) 227*, 680–685.

Levine, M. (1957) *Virology 3*, 22–41.

Maniatis, T., Ptashne, M., Backman, K., Kleid, D., Flashman, S., Jeffrey, A., & Maurer, R. (1975) *Cell (Cambridge, Mass.) 5*, 109–113.

Maniatis, T., Fritsch, E. F., & Sambrook, J. (1982) *Molecular Cloning, a Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Maxam, A., & Gilbert, W. (1980) *Methods Enzymol. 65*, 499–559.

Poteete, A. R. (1982) *Virology 119*, 422–429.

Poteete, A. R., & Roberts, T. M. (1981) *Gene 13*, 153–161.

Poteete, A. R., & Ptashne, M. (1982) *J. Mol. Biol. 157*, 21–48.

Poteete, A. R., Ptashne, M., Ballivet, M., & Eisen, H. (1980) *J. Mol. Biol. 137*, 81–91.

Roberts, T. M., Kacich, R., & Ptashne, M. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 760–764.

Rosenberg, M., Court, D., Shimatake, H., Brady, C., & Wulff, D. L. (1978) *Nature (London) 272*, 414–423.

Sauer, R. T., Pan, J., Hopper, P., Hehir, K., Brown, J., & Poteete, A. R. (1981) *Biochemistry 20*, 3591–3598.

Sauer, R. T., Yocum, R. R., Doolittle, R. F., Lewis, M., & Pabo, C. O. (1982) *Nature (London) 298*, 447–451.

Susskind, M. M., & Botstein, D. (1978) *Microbiol. Rev. 42*, 385–413.

Takeda, Y., Folkmanis, A., & Echols, H. (1977) *J. Biol. Chem. 252*, 6177–6183.

Winston, F., & Botstein, D. (1981) *J. Mol. Biol. 152*, 209–232.