THE JOURNAL OF BIOLOGICAL CHEMISTRY

Vol. 260, No. 22, Issue of October 5, pp. 12124–12129, 1985
Printed in U.S.A.

The Bacteriophage P22 Arc and Mnt Repressors

OVERPRODUCTION, PURIFICATION, AND PROPERTIES*

(Received for publication, February 25, 1985)

Andrew K. Vershon‡, Philip Youderian§, Miriam M. Susskind¶, and Robert T. Sauer‡

From the ‡Biology Department, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, the §Department of Biological Sciences, University of Southern California, Los Angeles, California 90089, and the ¶Department of Molecular Genetics and Microbiology, University of Massachusetts Medical School, Worcester, Massachusetts 01605

The *arc* and *mnt* genes of bacteriophage P22 encode small repressor proteins. We have cloned these genes onto plasmids that overproduce Arc and Mnt to greater than 1% of the soluble cellular protein. Both proteins were purified to greater than 95% homogeneity, and N-terminal sequences and amino acid compositions were determined. These data, in combination with previously determined gene sequences, establish the complete protein sequences for Arc (53 residues) and Mnt (82 residues). Both proteins have melting temperatures between 45 and 55 °C and can be renatured to a fully active species. Arc is a dimer in solution and Mnt is a tetramer.

Salmonella phage P22 has two regulatory regions, *imm*C and *imm*I, involved in the establishment and maintenance of lysogeny (1, 2). The *imm*C region is analogous in function to the bacteriophage λ immunity region. The *imm*I region, which has no analog in phage λ, encodes three proteins, antirepressor, Arc, and Mnt. Antirepressor, the product of the *ant* gene, functions to inactivate the P22 *c*2 repressor encoded by the *imm*C region. Arc and Mnt are repressor proteins that regulate gene expression in the *imm*I region. The general organization of genes and regulatory sites in the *imm*I region is shown in Fig. 1. The *arc* and *ant* genes are co-transcribed from the *ant* promoter, P_{ant}. During lytic growth of P22, Arc binds to O_{arc} and represses its own synthesis as well as that of antirepressor. The *mnt* gene is transcribed from its own promoter, P_{mnt}. During lysogeny Mnt binds to O_{mnt} and represses synthesis of both Arc and antirepressor. The DNA sequence of the *imm*I region has been determined and protein sequences have been predicted for Arc, Mnt, and antirepressor (3).

In this paper, we describe the overproduction, purification, and properties of the Arc and Mnt repressor proteins.

MATERIALS AND METHODS

Construction of Plasmids—Plasmids designed to overproduce Arc (pMS274) and Mnt (pMS282) were constructed in several steps using the general procedures described by Youderian *et al.* (4). To construct pMS274 (Fig. 2*a*), P22 DNA was cleaved with *Ava*II, which cuts the DNA 22 base pairs 5′ to the *arc* gene and several thousand base pairs 3′ to the gene. *Ava*II ends were made flush using DNA polymerase I large fragment plus dCTP, dGTP, and dTTP, and phosphorylated *Eco*RI linkers were ligated to the ends. The ligation mixture was then cleaved with *Eco*RI and *Hpa*I and the 295-base pair fragment containing the *arc* gene was gel purified and ligated into *Eco*RI/*Pvu*II-digested pBR322 to produce pMS100. pMS100 was cleaved with *Eco*RI, and the ends were filled in using DNA polymerase I large fragment, dATP and dTTP. The DNA was then cleaved with *Pst*I and the backbone fragment was ligated with a *Pst*I/*Hpa*I fragment containing the λ P_L promoter from plasmid pTP123[1] in the presence of phosphorylated *Xba*I linkers (5′-CTCTAGAG-3′; P-L Biochemicals). Transformants were selected for their ability to confer ampicillin resistance and the structure of pMS274 was verified by restriction mapping. pMS274 was transformed into *Escherichia coli* strain N4830 (5) to construct the Arc producing strain MS2274.

pMS282 was constructed by cleaving P22 *sie*A44 pANT RU454 *arc-am*H1605 DNA with *Taq*I and *Hae*III (the RU454 mutation creates a *Taq*I site 26 base pairs upstream of the AUG initiation codon of the *mnt* gene; 4). The 306-base pair *Taq*I/*Hae*III fragment carrying the *mnt* gene was gel purified and ligated with *Cla*I/*Pvu*II-digested pBR322 DNA. The resulting plasmid, pMS80, was cleaved with *Eco*RI and the ends were filled in with DNA polymerase I large fragment, dATP, and dTTP. The DNA was then digested with *Pst*I. The backbone fragment was purified and ligated with a *Pst*I/*Hpa*I fragment from plasmid pTP123 in the presence of phosphorylated *Xba*I linkers. Following selection for ampicillin resistance, one plasmid, pMS278, was found to contain multiple linkers at the 5′ end of the *mnt* gene. pMS278 DNA was digested with *Xba*I, diluted, and religated. The religation mixture was then digested with *Sst*I to linearize products carrying multiple linkers; pMS282 was obtained from a transformant of this mixture and was shown to contain a single *Xba*I linker. The structure of pMS282 (Fig. 2*b*) was verified by restriction mapping. Plasmid pMS282 was transformed into strain N4830 to construct the Mnt producing strain MS2282.

Buffers—These were composed of the following; lysis buffer: 100 mM Tris-HCl (pH 8.0), 200 mM KCl, 1 mM EDTA, 2 mM $CaCl_2$, 10 mM $MgCl_2$; PCB: 50 mM Tris-HCl (pH 8.0), 0.1 mM EDTA, 5% glycerol, 1.4 mM 2-mercaptoethanol; SB: 50 mM Tris-HCl (pH 7.5), 0.1 mM EDTA, 5% glycerol, 1.4 mM 2-mercaptoethanol; MSB: 50 mM MES[2]-HCl (pH 6.0), 0.1 mM EDTA, 5% glycerol, 1.4 mM 2-mercaptoethanol; GAB: 10 mM Tris-HCl (pH 7.5), 10 mM $MgCl_2$, 0.1 mM EDTA, 200 mM KCl, 1 mM $CaCl_2$, 1 mM dithiothreitol, 100 μg/ml BSA, and 25 μg/ml sonicated calf thymus DNA.

Purification of Arc Protein—Strain MS2274 was grown with aeration at 30 °C in 10 liters of LB broth in a New Brunswick MF-114 fermentor. Ampicillin (0.5 g) was added every hour and growth was monitored by absorbance at 600 nm. At an absorbance of 1.0, induction was initiated by raising the temperature to 42 °C (to inactivate the λ *c*I857 repressor). After 30 min, the temperature was lowered to 37 °C and growth was continued for another 2.5 h.

All purification steps were performed at 4 °C. The cells were harvested by centrifugation in a Beckman J6B centrifuge at 4,000 rpm for 20 min. Unless indicated, all further centrifugations were performed at 10,000 rpm in the GSA rotor of a Sorvall centrifuge. The cell paste was resuspended in 250 ml of lysis buffer and was centrifuged at 6000 rpm for 15 min. The pelleted cells (50 g) were resuspended in 100 ml of lysis buffer. Cells were lysed by sonication

* This work was supported by National Institutes of Health Grants AI-16892 (to R. T. S.), GM32261 (to P. Y.), GM22877, and GM29049 (to M. M. S.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[1] A. R. Poteete, unpublished.

[2] The abbreviations used are: MES, 4-morpholinepropanesulfonic acid; SDS, sodium dodecyl sulfate; PEI, polyethyleneimine; BSA, bovine serum albumin.

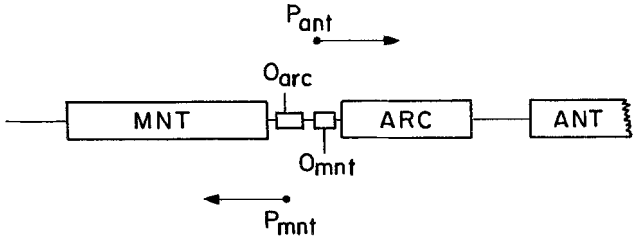

FIG. 1. **The P22 *imm*I region.** The relative positions of the Arc and Mnt operator sites (O_{arc} and O_{mnt}) are indicated with respect to the initiation sites for transcription from the P_{ant} and P_{mnt} promoters. Only a portion of the *ant* gene is shown.

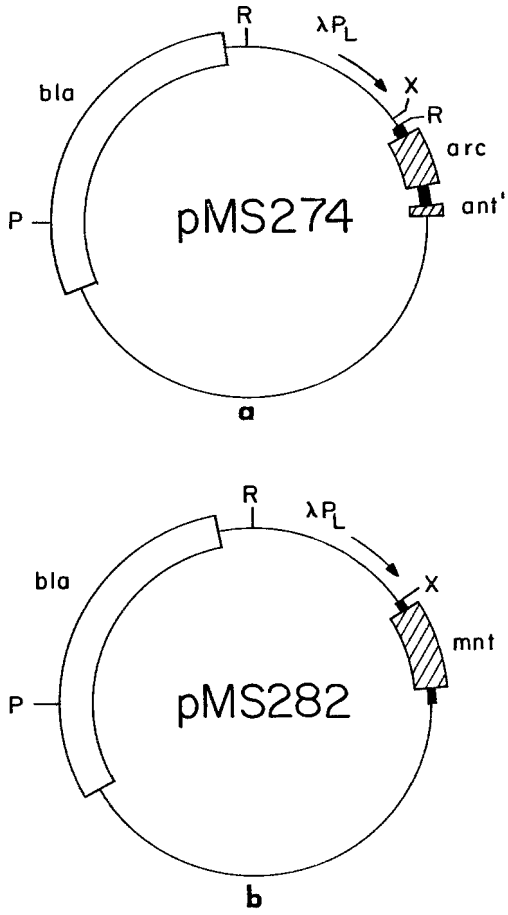

FIG. 2. **Structures of plasmids pMS274 and pMS282.** The *arrow* shows the direction of transcription from the λ P_L promoter. The *hatched* regions indicate sequences coding for the structural genes of Arc, Mnt, and the first 10 amino acids of antirepressor. *bla* is the β-lactamase gene from pBR322. *R*, *Eco*RI; *P*, *Pst*I; *X*, *Xba*I.

until the absorbance at 600 nm of the suspension was reduced 10-fold. This crude lysate was diluted 5-fold with lysis buffer, centrifuged for 1 h, and the supernatant (Pool 1) was saved.

Polyethyleneimine (PEI) was added to the supernatant to a final concentration of 1.0% and the precipitate was removed by centrifugation. Ammonium sulfate (0.45 g/ml supernatant) was added to the PEI supernatant fraction and stirred overnight. The precipitate was collected by centrifugation for 1 h; washed in 100 ml of 0.4 g/ml ammonium sulfate, 50 mM Tris (pH 8.0); and centrifugated for 30 min. The pellet was suspended in 50 ml of 0.2 g/ml ammonium sulfate, 50 mM Tris (pH 8.0) and stirred for 30 min. After 30 min of centrifugation, the supernatant was collected and saved. The pellet was washed twice more in the same manner and the supernatants were pooled with the first supernatant. Solid ammonium sulfate (0.25 g/ml) was added to the combined supernatant pool and stirred for 30 min. After centrifugation for 1 h, the precipitate was dissolved in 20 ml of PCB plus 0.2 M KCl, and was designated as Pool 2.

Pool 2 was chromatographed on a Sephadex G-75 (fine) column (5.0 × 90 cm) at a flow rate of 60 ml/h in PCB plus 0.2 M KCl. Ten-ml fractions were collected. Elution of protein was monitored by A_{280} and by SDS-polyacrylamide electrophoresis on 15% gels using the buffer system of Laemmli (6). Arc protein eluted between 950 and 1150 ml. Fractions containing Arc were pooled and designated Pool 3.

Pool 3 was loaded onto a Whatman P-11 phosphocellulose column (2.5 × 15 cm) at a flow rate of approximately 40 ml/h. After loading, the column was washed with 2 volumes of PCB plus 0.2 M KCl and a gradient was developed from 0.2 to 1.0 M KCl (300 ml each). Ten-ml fractions were collected. Arc protein eluted between 0.5 and 0.7 M KCl. Fractions containing Arc protein were combined and designated Pool 4.

Pool 4 was dialyzed into SB plus 0.05 M KCl and was loaded onto a Bio-Rad Affi-Gel Blue column (1.5 × 20 cm). After loading, the column was washed with 2 volumes of SB plus 0.05 M KCl and a gradient was developed from 0.05 to 1.0 M KCl (200 ml each). Arc eluted as a very broad peak which was pooled to give Pool 5. This pool was dialyzed into SB plus 0.1 M KCl and concentrated in a Amicon ultra filter using a YM-5 filter.

During the purification, portions of each pool were saved and stored at −20 °C. Protein concentrations were determined by the Bio-Rad protein assay, and Arc activity was assayed by the gel binding assay described below. The purified protein (Pool 5) lost no appreciable activity when stored at 4 or −70 °C for a period of 6 months.

Purification of the Mnt Protein—Mnt protein was purified from strain MS2282. Growth, harvest, lysis, PEI, and ammonium sulfate fractionations were performed as described for the Arc purification. The concentrated ammonium sulfate fraction (Pool 2) was chromatographed on a G-75 column in PCB plus 0.2 M KCl. Fractions were assayed for the presence of Mnt protein by SDS-polyacrylamide electrophoresis on 15% gels. Mnt protein was present in two peaks. A small amount (5%) of active Mnt eluted in the void volume (after approximately 500 ml). The majority of the Mnt, however, eluted between 680 and 800 ml. Fractions from this second peak were pooled and designated as Pool 3. Mnt from the void volume peak was not purified further or characterized.

Pool 3 was loaded onto a Whatman P-11 phosphocellulose column, which was washed and eluted as in the Arc purification. Mnt protein elutes in fractions containing between 0.4 and 0.6 M KCl. These fractions were combined and designated Pool 4.

Pool 4 was dialyzed into SB plus 0.05 M KCl to reduce the salt concentration. This material was loaded onto a CM-Sephadex (C-50) column (1.5 × 10 cm) equilibrated in SB plus 0.05 M KCl. After loading, the column was washed with 2 volumes of SB plus 0.05 M KCl buffer and a gradient was developed from 0.05 to 1.0 M KCl (100 ml each). Fractions containing Mnt were combined to give Pool 5.

To remove trace comtaminates, Pool 5 was dialyzed into MSB plus 0.05 M KCl, loaded onto a Bio-Rad P-Cell column (2.5 × 20 cm), and washed with 2 column volumes of the same buffer. A gradient was developed from 0.05 to 1.0 M KCl (200 ml each). Fractions containing Mnt protein were pooled to give Pool 6. Pool 6 was dialyzed into SB plus 0.1 M KCl and stored at −70 °C. Protein concentration and DNA binding activity was assayed as described for the Arc protein.

Protein Characterization—Amino acid analysis and amino-terminal Edman degradation were performed as previously described (7). Measurement of circular dichroism, CD melting curves, and thermolysin digestion experiments were performed as described by Hecht *et al.* (8).

Molecular weights for the native proteins were determined by gel filtration chromatography. Arc was chromatographed on a 1.0 × 50-cm Sephadex G-50 (fine) column in 50 mM Tris-HCl (pH 7.5), 0.2 M NaCl, 0.1 mM EDTA, 5% glycerol. The column was calibrated by determining the elution volumes of BSA, carbonic anhydrase, myoglobin, cytochrome *c*, aprotinin, and vitamin B12. The molecular weight of Mnt was determined on a 1.5 × 50-cm Sephadex G-75 column equilibrated in the same buffer. The column was calibrated with thyroglobulin, BSA, ovalbumin, carbonic anhydrase, myoglobin, cytochrome *c*, aprotinin, and vitamin B12.

Preparation of Labeled Operator DNA—The operators, O_{arc} and O_{mnt}, recognized by the Arc and Mnt proteins reside on a 250-base pair *Eco*RI/*Hin*dIII DNA fragment that can be obtained from plasmid pMS200 (4). ^{32}P-labeled operator fragments were obtained by digesting plasmid pMS200 with *Eco*RI and filling in the ends with [α-^{32}P]dATP using DNA polymerase large fragment. The DNA was then cut with *Hin*dIII and electrophoresed on a 5% acrylamide gel

(cross-linked with 0.1% bisacrylcystamine (Bio-Rad). The appropriate band was excised and the gel was solubilized by vortexing with 0.9 ml of 0.2 M NaCl, 20 mM Tris-HCl (pH 7.5), 1 mM EDTA, and 0.1 ml of 2-mercaptoethanol, followed by incubation at room temperature for at least 1 h. The solubilized mixture was loaded onto a Schleicher & Schuell Elutip which was then washed extensively in 0.2 M NaCl, 20 mM Tris-HCl (pH 7.5), 0.1 mM EDTA and was precipitated with ethanol using 5 μg of sonicated calf thymus DNA as carrier. The DNA pellet was dried and resuspended in 100 μl of 10 mM Tris-HCl (pH 7.5), 0.1 mM EDTA.

DNA Binding Assay—Binding of Arc or Mnt to operator DNA was assayed by a gel electrophoresis method (9, 10). In this assay, free DNA and protein-DNA complexes have different electrophoretic mobilities and can be resolved and quantitated. For the assay used here, the labeled fragment was diluted to approximately 20,000 cpm/ml in GAB buffer. Ten μl of this DNA was mixed with 10 μl of protein sample diluted in GAB and was incubated for at least 15 min at room temperature. Three μl of sample buffer (50% glycerol, 10 mM Tris-HCl (pH 7.5) plus tracking dyes bromphenol blue and xylene cyanol) were added and the entire sample was loaded onto a 5% acrylamide TBE gel (11) and electrophoresed at 250 V until the bromphenol blue dye migrated to the bottom of the gel (approximately 10 cm). The gel was blotted onto filter paper, dried under vacuum for 1 h on a Hoeffer gel dryer, and exposed to Kodak XR-5 film at −70°C overnight. Under the conditions of the binding assay, the apparent equilibrium dissociation constant for Arc-operator binding is 0.7 μM and that for Mnt-operator binding is 0.5 nM. For both proteins, 1 unit of binding activity is defined as the sample dilution at which 50% binding is observed. Activities at various stages of purification were assayed using the same preparation and quantity of labeled DNA.

RESULTS

Purification—After temperature induction of the Arc and Mnt overproducing strains (MS2274 and MS2282), SDS-polyacrylamide gels revealed low molecular weight proteins that were not present in the uninduced lysates. In initial studies, the Arc and Mnt purifications were monitored by DNA binding assays but we found that the purifications could be reliably and more conveniently monitored by SDS gels. Tables I and II show the yields and activities of pools at different stages of purification of the Arc and Mnt proteins. Fig. 3 shows SDS-gel electrophoretic assays of the relative purity of the pools at each step of purification.

The polyethyleneimine precipitation serves chiefly to remove nucleic acids from the lysate which might later interfere with protein binding to the ion exchange columns. Although a significant portion of the Arc protein is lost in the PEI pellet, this loss is tolerable because the PEI step serves to remove contaminants that are difficult to separate from the Arc protein later in the purification. The ammonium sulfate precipitations separate Arc and Mnt from the PEI in solution and from some of the larger protein contaminants. This step also serves to concentrate the protein so that a reasonable volume can be loaded onto the G-75 Sephadex column.

TABLE I
ARC purification

Step	Total protein	Total activity	% activity recovered
	mg	*units*	
1. Crude lysate	6250	80,000	
2. PEI/$(NH_4)_2SO_4$	2150	26,000	32.5
3. G-75 Pool	292	18,000	22.5
4. Phosphocellulose Pool	102	16,500	20.6
5. Affi-Gel Blue	110	12,200	15.2

TABLE II
MNT purification

Step	Total protein	Total activity	% activity recovered
	mg	*units*	
1. Crude lysate	6750	8.0×10^6	
2. PEI/$(NH_4)_2SO_4$	2450	7.0×10^6	87.5
3. G-75 Pool	882	7.0×10^6	87.5
4. Phosphocellulose 1 Pool	163	5.8×10^6	72.5
5. CM Pool	98	6.3×10^6	78.7
6. Phosphocellulose 2 Pool	73	7.0×10^6	87.5

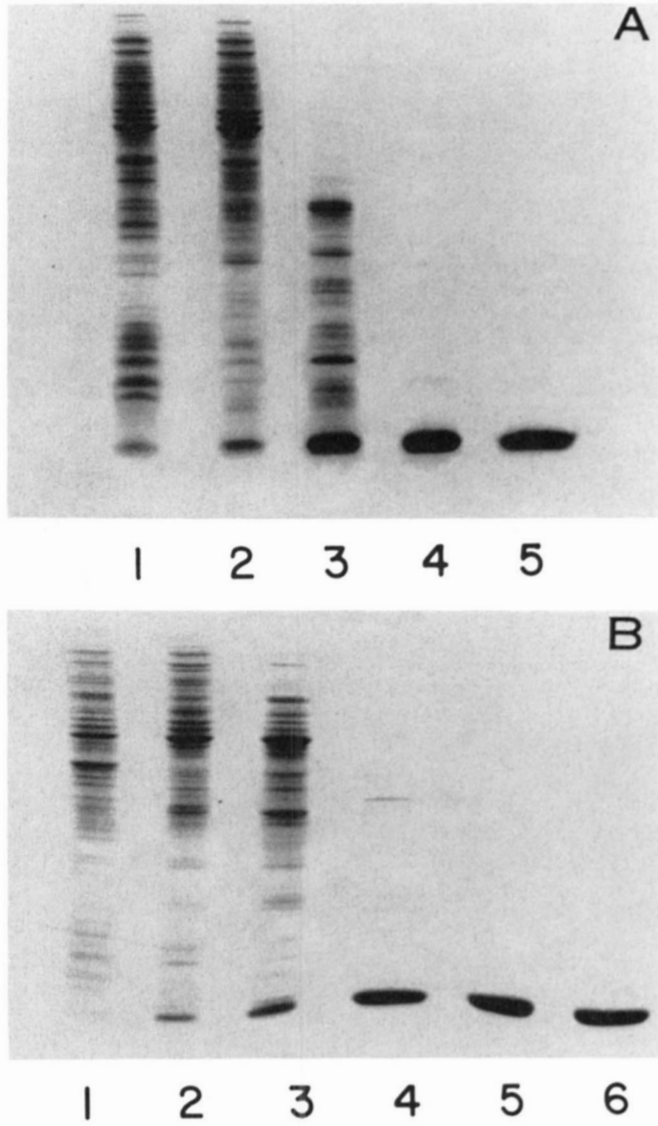

FIG. 3. **SDS-gel electrophoretic assays of the Arc and Mnt protein purification pools.** Approximately 20 μg of total protein from each pool was analyzed. *A*, arc purification; the *number* below each *lane* corresponds to the pool number in Table I; *1*, crude lysate; *2*, PEI/$(NH_4)_2SO_4$; *3*, G-75; *4*, phosphocellulose; *5*, Affi-Gel Blue. *B*, Mnt purification; the *number* below each *lane* corresponds to those in Table II. *1*, crude lysate; *2*, PEI/$(NH_4)_2SO_4$; *3*, G-75; *4*, phosphocellulose 1; *5*, CM-Sephadex; *6*, phosphocellulose 2.

Arc and Mnt are both small proteins and elute after the major protein peak, which is excluded from the G-75 column. As shown in Fig. 3, the G-75 chromatography separates Arc and Mnt from many of the larger contaminating proteins. Arc elutes as a distinct peak of absorbance and is one of the last proteins to elute from the column. Mnt, which is larger than Arc, elutes as a trailing shoulder of the main protein peak. Pools from the G-75 column were loaded directly onto the phosphocellulose columns at the 0.2 M KCl concentration of the G-75 running buffer. At this salt concentration, most proteins do not bind to the phosphocellulose column and are

found in the flow-through fraction. Arc and Mnt bind tightly and elute as major protein peaks near 0.5 M KCl. Conservative pooling of peak fractions at this step removes certain contaminants that are otherwise difficult to remove in later purification steps. At this stage of purification, most absorbance at 280 nm is from Arc or Mnt, but several nonabsorbing contaminants can be detected by SDS gels. Most of these contaminants can be removed from Arc by chromatography on Affi-Gel Blue (Pool 5) and from Mnt by chromatography on CM-Sephadex (Pool 5). After these steps both Arc and Mnt are greater than 95% pure. A second phosphocellulose column can be used in the Mnt purification to remove additional trace contaminants.

N-terminal Sequences and Amino Acid Compositions—Ten cycles of automated Edman degradation were performed on the purified Arc and Mnt proteins. The results are listed in Table III. The amino-terminal sequence of Arc is exactly that predicted from the gene sequence (3). The amino-terminal sequence of Mnt is in agreement with the sequence predicted from the gene sequence allowing for post-translational cleavage of the initiator formyl methionine to produce the mature protein.

The amino acid compositions of Arc and Mnt (Table IV) are also in reasonable agreement with compositions predicted from the gene sequences allowing for the removal of fMet from Mnt. The complete amino acid sequences for Arc and Mnt are shown in Fig. 4.

The molar extinction coefficients at 280 nm for Arc and Mnt were determined by combining absorption data with protein concentrations calculated from amino acid compositions. Arc contains 1 tryptophan residue and 1 tyrosine residue, and has a molar extinction coefficient of 7800. Mnt contains 2 tyrosine and no tryptophan residues and has an extinction coefficient of 4500.

TABLE III

N-terminal sequences

PTH-Arg and PTH-His were not identified in these experiments and are expected at the positions marked XXX. See Fig. 4.

Residue:	1	2	3	4	5	6	7	8	9	10
Arc:	Met	Lys	Gly	Met	Ser	Lys	Met	Pro	Gln	Phe
Mnt:	Ala	XXX	Asp	Asp	Pro	XXX	Phe	Asn	Phe	XXX

TABLE IV

Amino acid compositions

Amino acid	Arc protein		Mnt protein	
	Expected	Found	Expected	Found
Cm-Cys	0	N.D.[a]	0	N.D.
Asp + Asn	4	4.1	11	11.2
Thr	0	0	4	3.9
Ser	4	3.4	5	4.7
Glu + Gln	8	8.3	10	10.6
Pro	2	1.9	4	4.3
Gly	4	3.7	2	2.4
Ala	2	2.0	6	6.1
Val	5	5.3	5	4.8
Met	4	3.3	4	4.0
Ile	2	1.9	1	0.8
Leu	3	3.1	8	8.1
Tyr	1	0.9	2	1.8
Phe	2	2.0	4	3.9
His	0	0	1	1.6
Lys	5	5.7	8	8.0
Arg	6	6.1	7	6.9
Trp	1	N.D.	0	N.D.
Monomer molecular weight		6199		9504

[a] N.D., not determined.

ARC

Met 1 – Lys 2 – Gly 3 – Met 4 – Ser 5 – Lys 6 – Met 7 – Pro 8 – Gln 9 – Phe 10 –
Asn 11 – Leu 12 – Arg 13 – Trp 14 – Pro 15 – Arg 16 – Glu 17 – Val 18 – Leu 19 – Asp 20 –
Leu 21 – Val 22 – Arg 23 – Lys 24 – Val 25 – Ala 26 – Glu 27 – Glu 28 – Asn 29 – Gly 30 –
Arg 31 – Ser 32 – Val 33 – Asn 34 – Ser 35 – Glu 36 – Ile 37 – Tyr 38 – Gln 39 – Arg 40 –
Val 41 – Met 42 – Glu 43 – Ser 44 – Phe 45 – Lys 46 – Lys 47 – Glu 48 – Gly 49 – Arg 50 –
Ile 51 – Gly 52 – Ala 53

MNT

Ala 1 – Arg 2 – Asp 3 – Asp 4 – Pro 5 – His 6 – Phe 7 – Asn 8 – Phe 9 – Arg 10 –
Met 11 – Pro 12 – Met 13 – Glu 14 – Val 15 – Arg 16 – Glu 17 – Lys 18 – Leu 19 – Lys 20 –
Phe 21 – Arg 22 – Ala 23 – Glu 24 – Ala 25 – Asn 26 – Gly 27 – Arg 28 – Ser 29 – Met 30 –
Asn 31 – Ser 32 – Glu 33 – Leu 34 – Leu 35 – Gln 36 – Ile 37 – Val 38 – Gln 39 – Asp 40 –
Ala 41 – Leu 42 – Ser 43 – Lys 44 – Pro 45 – Ser 46 – Pro 47 – Val 48 – Thr 49 – Gly 50 –
Tyr 51 – Arg 52 – Asn 53 – Asp 54 – Ala 55 – Glu 56 – Arg 57 – Leu 58 – Ala 59 – Asp 60 –
Glu 61 – Gln 62 – Ser 63 – Glu 64 – Leu 65 – Val 66 – Lys 67 – Lys 68 – Met 69 – Val 70 –
Phe 71 – Asp 72 – Thr 73 – Leu 74 – Lys 75 – Asp 76 – Leu 77 – Tyr 78 – Lys 79 – Lys 80 –
Thr 81 – Thr 82

FIG. 4. **Sequences of the Arc and Mnt proteins.**

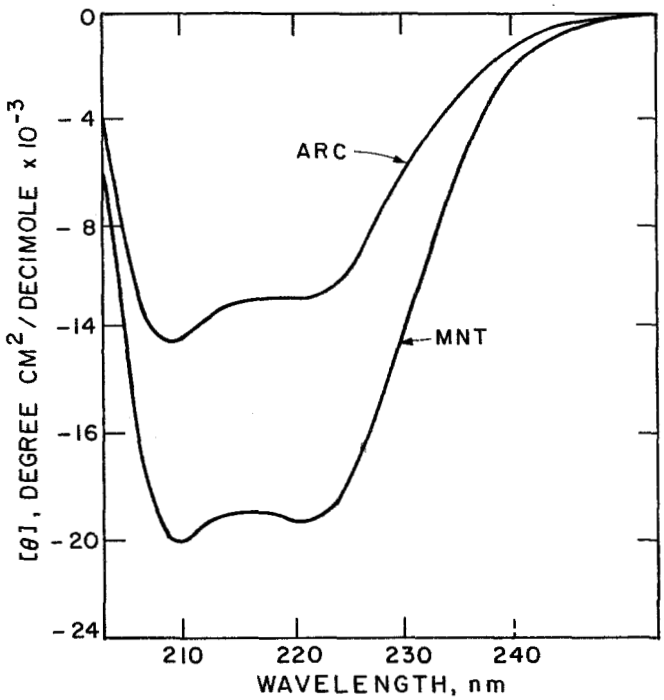

FIG. 5. **Circular dichroism spectra of Arc and Mnt proteins.**

Secondary Structure and Stability—Fig. 5 shows the circular dichroism (CD) spectra of Arc and Mnt. The peaks of negative ellipticity at 208 and 222 nm and the general shape of the curves suggest that each protein contains some α-helix

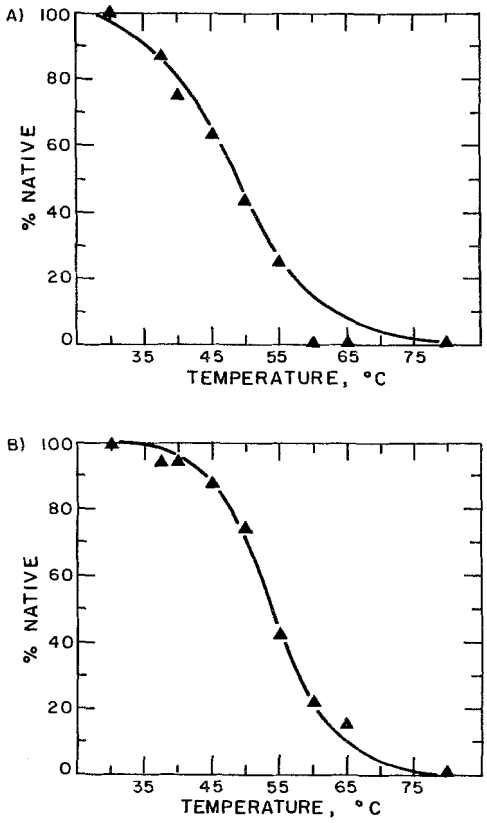

FIG. 6. **Melting curves of the Arc and Mnt proteins.** The per cent of each protein in the native conformation was calculated from the negative ellipticity at 222, at each temperature. *Triangles* represent actual data points. The *solid curves* are theoretical plots calculated from the van't Hoff equation using T_m = 48 °C, ΔH_{app} = 28 kcal/mol for Arc and T_m = 54 °C, ΔH_{app} = 48 kcal/mol for Mnt. *A*, Arc melting curve. *B*, Mnt melting curve.

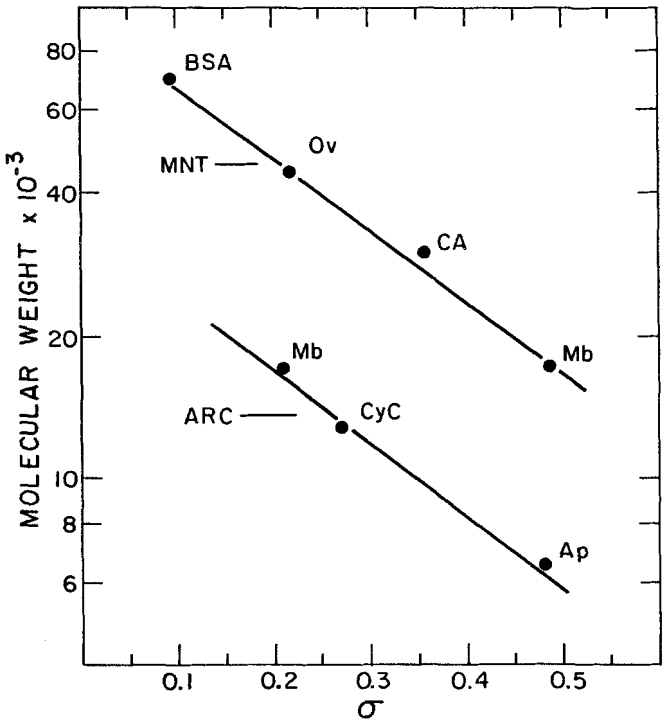

FIG. 7. **Elution positions of Arc and Mnt in gel filtration chromatography.** The *lower curve* represents chromatography on a Sephadex G-50 (fine) column. The *upper curve* represents chromatography on a Sephadex G-75 (fine) column. BSA; 64,000; ovalbumin (*Ov*), 44,000; carbonic anhydrase (*CA*), 29,500; myoglobin (*Mb*), 17,000; cytochrome *c* (*CyC*), 12,500; and aprotinin (*Ap*), 6,500. σ is calculated as $[V_0 - V_e]/[V_0 - V_s]$, where V_0 is the void volume of the column (measured as the elution volume of thyroglobulin on G-75 and BSA on G-50); V_s is the salt volume (measured as the elution volume of vitamin B12 on both columns); and V_e is the elution volume of the included protein.

(12). There is no evidence for the presence of β-sheet in either protein.

The thermal stabilities of the Arc and Mnt proteins were investigated by measuring changes in the CD spectra as a function of temperature. Fig. 6 shows melting curves for both Arc and Mnt. Arc has a melting temperature (T_m) of approximately 48 °C. The curve is rather broad and suggests that a fraction of the Arc molecules are denatured at physiological temperatures. An apparent enthalpy of denaturation of 28 kcal/mol was calculated by fitting the experimental points to the van't Hoff equation. The melting temperature of Mnt (T_m = 54 °C) is slightly higher than that of Arc, and the Mnt denaturation curve is somewhat steeper (ΔH_{app} = 45 kcal/mol). The apparent enthalpy changes for both protein denaturations should be considered approximate values since the experimental data are not sufficiently precise to evaluate possible heat capacity differences between the folded and unfolded proteins.

The denaturation of Arc and Mnt were also monitored by measuring their susceptibility to proteolytic digestion by thermolysin at different temperatures (8). Proteolysis was monitored by SDS-gel electrophoresis. 50% digestion of Mnt occurs between 51 and 55 °C, in good agreement with the T_m obtained from the CD melting curves. Digestion of Arc protein occurred at all temperatures but was slower at the lower temperatures, consistent with the broad melting curve observed by CD. Since Arc and Mnt are small proteins, we assume that each folds into a single domain. We have been unable to generate discrete, stable fragments of either protein by limited digestion with trypsin or thermolysin.

The renaturation of Arc and Mnt after heating to 80 °C was also examined. Both protein renature fully upon cooling, and produce CD spectra identical to the ones obtained before heating. Each protein also recovers over 95% of its original DNA binding activity after heating and cooling. The ability of both proteins to renature after melting can be used as a purification step. When crude lysates are diluted at least 3-fold and heated to 90°C for 10 min, slow cooling results in precipitation of 80–90% of the bulk protein but from 40 to 85% of the Arc or Mnt activity remains soluble. We do not routinely utilize this procedure in our purification since there is considerable variability in the yields, but it can be useful when rapid purification is desirable.

Oligomeric Forms of Native Proteins—The native molecular weight of each protein was established by gel filtration chromatography. Fig. 7 shows the elution positions of Arc and Mnt after chromatography on Sephadex G-50 or G-75 and also shows the elution positions of several proteins of known molecular weight. Variation of the loading concentrations of Arc and Mnt from 0.012 to 1.2 mg/ml caused no significant change in their elution volumes. At high and low loading concentrations, Arc and Mnt eluted as symmetric peaks, with no trailing. Arc chromatographs at an apparent molecular weight of 13 kDa, close to the 12.4 kDa value expected for a dimer. Mnt chromatographs at an apparent molecular weight of 43 kDa, close to the 38 kDa value expected for a tetramer.

DISCUSSION

To overproduce the Arc and Mnt proteins of bacteriophage P22, we fused the structural genes for these proteins to the λ

P_L promoter and cloned them onto multicopy plasmids derived from pBR322. When induced, strains bearing these plasmids produce Arc and Mnt at a level equivalent to about 1–2% by weight of soluble cell protein. The purification protocols that we describe are relatively simple and provide sufficient quantities of the purified proteins for most biochemical and physical studies.

In work reported previously (3), we determined the DNA sequences of the *arc* and *mnt* genes and used amber mutations in these genes to establish the direction of transcription and the proper translational reading frames. The N-terminal sequences and amino acid compositions we report here for Arc and Mnt confirm the amino acid sequences predicted from the DNA sequences and show that the initiator fMet of Mnt is removed in the mature protein.

The Arc and Mnt proteins, like most site-specific DNA binding proteins, are oligomeric. Arc is dimeric and Mnt is tetrameric at concentrations ranging from approximately 2 to 200 μM. However, both proteins bind to their operator sites with equilibrium dissociation constants below 1 μM, and thus we cannot be certain that the Arc dimer and the Mnt tetramer are the active DNA binding species. The amino acid sequences of Arc and Mnt are about 35% homologous (3) and these two proteins are therefore likely to share structural homology. This expectation is supported by the finding that both proteins behave in a similar manner during purification. However, Arc, which contains 53 residues, is significantly shorter than Mnt, which contains 82 residues, and the sequence homology aligns Arc with the N-terminal two-thirds of the Mnt sequence. We speculate that the C-terminal region of Mnt may contain the contacts that stabilize the tetrameric form of the protein.

The CD spectra of Arc and Mnt indicate that both proteins have significant regions of α-helix. The three-dimensional structures of λ Cro, the N-terminal domain of λ repressor, and the *E. coli* catabolite activator protein are known (13–15), and these proteins and many homologous DNA binding proteins appear to use a conserved structure consisting of two linked α-helices for DNA recognition (16, 17). Do Arc and Mnt use similar bihelical units for DNA binding? We cannot answer this question at present but we note that Arc and Mnt do not contain sequences that are homologous to the bihelical sequences of the other DNA binding proteins. Arc and Mnt may use another type of secondary structure for DNA recognition, but structural studies will be needed before this question can be answered.

Acknowledgments—We thank Kathy Hehir and Michael Hecht for their guidance in sequencing and circular dichroism measurements and Anna Jeitler-Nilsson for help with the purifications.

REFERENCES

1. Susskind, M. M., and Botstein, D. (1978) *Microbiol. Rev.* **42,** 385–413
2. Susskind, M. M., and Youderian, P. (1983) in *Lambda II* (Hendrix, R. W., Roberts, J. W., Stahl, F. W., and Weisberg, R. A., eds) pp. 347–363, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
3. Sauer, R. T., Krovatin, W., DeAnda, J., Youderian, P., and Susskind, M. M. (1983) *J. Mol. Biol.* **168,** 699–713
4. Youderian, P., Bouvier, S., and Susskind, M. M. (1982) *Cell* **30,** 843–853
5. Gottesman, M. E., Adhya, S., and Das, A. (1980) *J. Mol. Biol.* **140,** 57–75
6. Laemmli, U. K. (1970) *Nature* **227,** 680–685
7. DeAnda, J., Poteete, A. R., and Sauer, R. T. (1983) *J. Biol. Chem.* **258,** 10536–10542
8. Hecht, M. H., Sturtevant, J. M., and Sauer, R. T. (1984) *Proc. Natl. Acad. Sci. U. S. A.* **81,** 5685–5689
9. Garner, M. M., and Revzin, A. (1981) *Nucleic Acids Res.* **9,** 3047–3060
10. Fried, M., and Crothers, D. M. (1981) *Nucleic Acids Res.* **9,** 6505–6525
11. Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) in *Molecular Cloning: A Laboratory Manual*, pp. 173–177, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
12. Greenfield, N., and Fasman, G. D. (1969) *Biochemistry* **8,** 4108–4115
13. Anderson, W. F., Ohlendorf, D. H., Takeda, Y., and Matthews, B. W. (1981) *Nature* **290,** 754–758
14. Pabo, C. O., and Lewis, M. (1982) *Nature* **298,** 443–447
15. McKay, D. B., and Steitz, T. A. (1981) *Nature* **290,** 744–749
16. Takeda, Y., Ohlendorf, D. H., Anderson, W. F., and Matthews, B. W. (1983) *Science* **221,** 1020–1026
17. Pabo, C. O., and Sauer, R. T. (1984) *Annu. Rev. Biochem.* **53,** 293–321