The TyrR Protein of Escherichia coli, Analysis by Limited Proteolysis of Domain Structure and Ligand-Mediated Conformational Changes*

(Received for publication, August 17, 1992)

Jisong Cui and Ronald L. Somerville
Department of Biochemistry, Purdue University, West Lafayette, Indiana 47907-1153

The TyrR protein of Escherichia coli K12 is a homodimer containing 513 amino acids/subunit. This protein is important in the transcriptional regulation of several genes whose protein products catalyze steps in aromatic amino acid biosynthesis or transport. Methods were developed for efficiently purifying the TyrR protein to apparent homogeneity. We analyzed the pattern of cleavage of the TyrR protein by trypsin, either in the absence of ligands or in the presence of saturating levels of L-tyrosine, ATP, or poly(dI-dC). At low (1:200 ratio by weight) trypsin levels, in the absence of ligands, two major digestion products accumulated. These were polypeptides of 22 and 31 kDa, shown to contain amino acid residues 1-190 and 191-467, respectively. The pattern of trypsin cleavage was unaffected by tyrosine. In the presence of ATP, an intermediate species of 53 kDa, probably containing amino acid residues 1-467, was observed. The kinetics of appearance of the 53-kDa species were consistent with a role for ATP in accelerating the hydrolysis of the R*467-F*468 peptide bond. The 53-kDa polypeptide underwent further tryptic hydrolysis to yield fragments of 22 and 31 kDa. When both tyrosine and ATP were present, the rate of formation of the 22- and 31-kDa fragments was more rapid than in the absence of these ligands. It appears that when both ligands are bound, the rates of hydrolysis of peptide bonds R*190-Q*193 and R*240 and 291-297 were found by computer analysis in the TyrR protein (residues 468-513) contains the operator recognition elements, probably arranged in the form of a helix-turn-helix motif. This polypeptide segment was not detected as a discrete tryptic hydrolysis product.

The tyrosine repressor (TyrR) protein of Escherichia coli, a homodimer of 513 amino acids/subunit (Fig. 1) (Cornish et al., 1986), regulates the rates of transcription of nine genes, eight of which are involved in the synthesis or transport of aromatic amino acids (reviewed by Pittard and Davidson, 1991). The tyrR gene is autogenously regulated (Camakaris and Pittard, 1982). The TyrR protein represses aroF, aroL, tyrB, aroP, aroG, aroH, and tyrR (reviewed by Pittard and Davidson, 1991)) and activates the mtr gene (Heatwole and Somerville, 1991; Sarsero et al., 1991). The TyrR protein either represses or activates the expression of the tyrP gene, depending on whether phenylalanine (activation response) or tyrosine (repression response) is present (Kasian et al., 1986; Whipp and Pittard, 1977). The TyrR protein exerts its regulatory functions by binding to specific 22-base pair target sequences (designated TyrR boxes) that are either within or immediately upstream of the regulated promoters. There are two classes of TyrR box, designated “strong” and “weak” (reviewed by Pittard and Davidson, 1991). The binding of the TyrR protein to strong boxes in vitro is ligand-independent, but ATP and either tyrosine or phenylalanine must be present in order for the TyrR protein to bind to weak boxes (Argyropoulos, 1989). Generally, strong and weak boxes are present as adjacent pairs at the regulated promoters. The mechanisms of repression and activation mediated by the TyrR protein have not been established.

In solution, the TyrR protein binds ATP with a KD of 9 μM. There is no significant interaction between the TyrR protein and tyrosine/phenylalanine in vitro in the absence of ATP. In the presence of ATP, the TyrR protein binds tyrosine (KD = 63 μM) or phenylalanine (KD = 280 μM). Phenylalanine interacts with tyrosine for binding to the TyrR protein (Argyropoulos, 1989). It was recently established that some or all of the 9 amino acids at the N-terminal end of the TyrR protein are absolutely required for its activation function (Cui and Somerville, 1993a). Two presumptive ATP-binding sites (residues 234-240 and 291-297) were found by computer analysis in the TyrR protein (reviewed by Pittard and Davidson, 1991) (Fig. 1). The central region (residues 207-425) has a high degree of amino acid sequence similarity to a family of activators specific for the σ43 form of RNA polymerase (Fig. 1) (Stock et al., 1989). However, each of the genes regulated by the TyrR protein is transcribed by the σ70 form of RNA polymerase. The central domains of the σ43 activators contain sites where the binding and subsequent hydrolysis of ATP must take place in order for promoter activation to occur (Ninfa and Magasanik, 1986; Popham et al., 1989). In the case of the TyrR protein, it is not known whether the presumptive ATP-binding sites were correctly identified, whether hydrolysis of ATP occurs, or whether the binding of ATP to the TyrR protein plays a role in the activation function of this protein. The C-terminal region of the TyrR protein contains a polypeptide segment that serves as a DNA-binding element, pre-

5040
sumably in the form of a helix-turn-helix motif (reviewed by Pittard and Davidson, 1991). Deletion of the C-terminal 23 amino acids resulted in the loss of both activation and repression functions (Cutler and Sorokin, 1993b).

The chemical properties of the TyrR protein in solution have not been investigated. Here, we report a method for the large scale preparation of homogeneous TyrR protein. Limited trypsin digestion was used to elucidate its domain structure. The TyrR protein was found to contain an N-terminal domain of about 190 amino acids and a second domain of 277 amino acids. The C-terminal region of the TyrR protein appears to be loosely folded and does not exhibit those attributes of a domain that are operationally defined in limited proteolysis studies.

Limited trypsin digestion was also employed in detecting conformational changes within the TyrR protein in response to the binding of ligands. Our studies suggest that 1) the TyrR protein undergoes specific conformational changes upon the binding of ligands and 2) the second domain of the TyrR protein is protected by ATP from cleavage by trypsin.

**EXPERIMENTAL PROCEDURES**

**Materials**

Trypsin and soybean trypsin inhibitor were purchased from Sigma. Polyclonal antibodies were purchased from Boehringer Mannheim. Phosphocellulose P11 was obtained from Whatman. Hydroxyapatite (BioGel HTP) was obtained from BioRad. Chromatographic columns and accessories were purchased from Pharmacia LKB Biotechnology Inc.

**Buffers**

Buffer A was 10 mM K2HPO4/KH2PO4 (pH 6.6) with 100 mM NaCl, 1 mM EDTA, 0.01% NaN3, 7 mM 8-mercaptoethanol, and 1 mM phenylmethylsulfonyl fluoride. Buffer B was 50 mM K2HPO4/KH2PO4 (pH 7.5) with 1 mM EDTA, 0.01% NaN3, 7 mM β-mercaptoethanol, and 1 mM phenylmethylsulfonyl fluoride.

2× loading buffer (for SDS-PAGE) contains 125 mM Tris-chloride (pH 6.8), 4% SDS, 0.4% glycerol, 1.2 mM β-mercaptoethanol, and 0.1% bromphenol blue.

**Large Scale Preparation of ACE6**

E. coli strain ED3799 (Rosenberg et al., 1987) was grown overnight at 37 °C in 5 ml of TM medium (Gottesman and Varonolsky, 1968). 1 ml of cells was added to 10 μl of phage ACE6 (108/m) (Rosenberg et al., 1987). After 20 min at 37 °C, the phage-cell mixture was centrifuged at 10000 x g for 10 min at 4 °C. The supernatant, consisting of a high titer phage lysate, was stored at 0-5 °C for subsequent use. This procedure reproducibly yielded lysate at a titer of about 0.3.

**Purification of the TyrR Protein**

The procedure for the purification of the TyrR protein reported here contains some improvements over the previously described method (Argyropoulos, 1989). To overexpress the TyrR protein, strain BL21 (Rosenberg et al., 1987) carrying the tyrR+ construct pJCl00 (Somerville et al., 1991) was first grown overnight at 37 °C in 10 ml of LB medium (Lennox, 1955) supplemented with ampicillin (100 μg/ml) and maltose (0.2%). The overnight culture was then inoculated into 1 liter of the same medium contained in a 4-liter flask. The culture was grown at 37 °C on a rotary shaker to an OD600 of about 0.3. Glucose was then added to a final concentration of 4 mg/ml. The culture was then grown at 37 °C for another hour, and MgSO4 was added to a final concentration of 10 mM. The cells were then infected with ACE6 at a multiplicity of 5 phage/cell. After 3 h at 37 °C, the cells were harvested by centrifugation at 4 °C, 8000 rpm (Sorval RC2B, GSA rotor) for 20 min. The TyrR protein in these cells comprised 20-30% of the total protein. Protein purification, described below, was performed at 4 °C.

**Step 1**—The cells harvested from a 1-liter culture (3 g of paste) were resuspended in 30 ml of Buffer A. Cells were broken in a French press (three passages at 20 000 psi), then spun at 15,000 rpm (Beckman, J21C) for 30 min. The pellet was discarded, and the supernatant material was treated with streptomycin sulfate (10% solution in Buffer A), which was added slowly with stirring to a final concentration of 1%. The resulting suspension was centrifuged at 15,000 rpm (Beckman, J21C) for 30 min, and then the supernatant was dialyzed against Buffer A, the sample was loaded onto an 80-ml phosphocellulose P11 column. The column was developed by an NaCl gradient, rising from 0.1 to 1 M in 400 ml. The TyrR protein emerged between 375-450 mM NaCl. At this stage, the TyrR protein was 90% pure. A second identical phosphocellulose P11 column step further improved the purity of this protein.

**Step 2**—Solid ammonium sulfate was added slowly with stirring to the NaCl eluate to 80% saturation. This treatment precipitated the TyrR protein and eliminated some of the unwanted proteins. After dialysis against Buffer A, the sample was loaded onto an 80-ml hydroxyapatite column, whereas virtually all of the proteins that had coeluted with the TyrR protein from the phosphocellulose column were retained. At this point, the TyrR protein was over 95% pure, as determined by SDS-PAGE (Figs. 2-4; zero time lanes). From 1 liter of cells (about 3 g), 40 mg of homogeneous TyrR protein could be reproducibly purified using the above protocol.

**Trypsin Digestion**

Preparations of Trypsin and Trypsin Inhibitor—Trypsin was dissolved in 0.1 mM HCl at a concentration of 1 mg/ml. The stock trypsin solution (100-μl aliquots) was stored at −20 °C. Prior to use, the solution was thawed in an ice bath. Each thawed trypsin stock was used only once. Trypsin inhibitor was dissolved in 20 mM Tris-chloride buffer (pH 7.5) at a concentration of 1 mg/ml and stored at −20 °C.

**Trypsin Digestion**—All digestion experiments were carried out at room temperature (25 °C). Incubation mixtures contained (final volume, 200 μl) 20 mM Tris-chloride (pH 7.5), 1 mM EDTA, 0.5 mM β-mercaptoethanol, and 100 μg TyrR protein. Trypsin (either 0.5 or 0.9 μg) was then added to initiate digestion. Tyrosine (500 μM), ATP (200 μM), and poly(dI-dC) (20 μg) were added as appropriate. When these ligands were used, the TyrR protein was preincubated with ligand(s) for 10 min at room temperature prior to initiating digestion. A zero time sample was taken before the addition of trypsin. Samples (20 μl) were removed at the time points indicated on the individual figures. Samples removed during the course of digestion were immediately placed into 500-μl Eppendorf tubes in dry ice/ethanol that contained trypsin inhibitor (final concentration, 25 μg/ml). Prior to electrophoresis, 2× loading buffer (20 μl) was added to each sample tube and then samples from all time points were heated at 100 °C for 4 min. The resulting polypeptides were separated electrophoretically on 10% SDS-polyacrylamide gel (see below).

**Characterization of Tryptic Fragments**

**Protein Sequencing**—To prepare polypeptides for protein sequencing, a polyvinylidine difluoride-type support, ProBlott (Applied Biosystems) was used with minor modifications, according to the protocols recommended by the supplier. Briefer, SDS-PAGE was conducted according to our routine method (see below) except for the following. 1) A 2-h period of pre-electrophoresis was employed at 3
mM constant current and glutathione (50 \mu M) was present in the cathode buffer. 2) During electrophoresis, fresh cathode buffer with 0.1 mM sodium thioglycolate was used. After electrophoresis, the polypeptide bands on the gel were transferred electrophoretically to ProBlott membrane in 10 mM CAPS buffer with 10% methanol. The membrane was stained with Coomassie Blue R-200 (0.1%), acetic acid (1%), and methanol (40%). Stained bands were excised and subjected to Edman degradation analysis. Six reaction cycles were performed.

**Purification of the 31-kDa Tryptic Fragment**—Tryptsin digestion of the TyrR protein was carried out as described above. Low levels of trypsin (1:200) were used. Both ATP (200 \mu M) and tyrosine (500 \mu M) were added to protect the 31-kDa fragment from further digestion. After 16 min of incubation, the reaction was stopped with trypsin inhibitor as described above. Tryptic fragments were separated on a 10% native polyacrylamide gel. The native gel and running conditions were identical with those employed in SDS-PAGE, except that no SDS was added to either the gel or the electrode buffers. The gel was stained with Coomassie Blue R-200 (0.1%) for 10 min and destained briefly with distilled H2O until the 31-kDa fragment became visible. The fragment was cut out of the gel and eluted using an S&S ELUTRAP apparatus (Schleicher and Schuell), following the supplier’s instructions. The eluted 31-kDa polypeptide was dialyzed overnight at 4 °C against Buffer B. Frozen samples (−20 °C) were used for molecular mass determination (see below).

**Analytical Methods**

Protein concentrations were determined by the method of Bradford (1976) using bovine serum albumin as standard. Reagents were purchased from Bio-Rad.

SDS-PAGE was carried out in a Mini Protean II cell (Bio-Rad). A tricine buffer system (Schägger and Von Jagow, 1987) was used at a gel concentration of 10%. 2 x loading buffer was used to prepare the samples.

Polypeptide molecular weights were determined by electrospray mass spectrometry by Dr. K. M. Swiderek at the Beckman Research Institute of the City of Hope Medical Center, Duarte, California.

**RESULTS**

**Limited Tryptic Digestion of the TyrR Protein**

**Effects of Ligands**—Limited proteolysis has been extensively used to investigate the structures of multidomain proteins (Rossman and Argos, 1981). The availability in quantity of highly purified TyrR protein (see "Experimental Procedures") made it feasible to utilize trypsin digestion to study the domain structure of the TyrR protein. There are over 50 possible sites of tryptic cleavage in the TyrR protein (Cornish et al., 1986; Fig. 1). Selective hydrolysis of a subset of these peptide bonds would constitute evidence for their exposure to solvent. This feature of interdomain polypeptide segments is commonly observed (Rossmann and Argos, 1981). Two levels of trypsin were employed. Low levels (trypsin:TyrR protein, 1:200 by weight) were used to investigate ligand-mediated conformational changes. Higher levels of trypsin (1:11 by weight) were used to observe the protective effects of ATP. Each set of digestion studies was carried out in parallel using a stock solution of trypsin from the same vial. Identical conditions with respect to volume, buffer components, and protein concentration were employed in each set of digestions. The only variable factor was the nature of the TyrR ligands.

Initially, a series of four digestions (Fig. 2) were carried out to test whether ligands affected the susceptibility of the TyrR protein to trypsin. In the absence of any cofactors, the 58-kDa TyrR protein (Fig. 2A, lane 0) was gradually converted into two major trypsin-resistant fragments of 22 and 31 kDa (Fig. 2A). The half-life of the TyrR protein was about 16 min. Complete conversion of the TyrR protein to stable subfragments of 22 and 31 kDa required about 2 h (data not shown).

In the presence of tyrosine (500 \mu M) (Fig. 2B), the pattern of digestion resembled that which was observed in the absence of ligands (Fig. 2A). As before, two main fragments of 22 and 31 kDa were generated. Under the above two conditions, a minor species of 36 kDa was also observed early in the course of trypsin digestion. The nature of this polypeptide will be explained later. In the presence of ATP (200 \mu M) (Fig. 2C), the TyrR protein was gradually degraded into the same two species as before. However, a new species, (53 kDa) was observed in the early stage of this digestion. The half-life of the TyrR protein in the presence of ATP was 4–8 min, slightly less than the previous two cases. In the presence of both ATP (200 \mu M) and tyrosine (500 \mu M) (Fig. 2D), the 22- and 31-kDa subfragments were still the two main products. A 53-kDa polypeptide, identical with that observed in the presence of ATP, was also detected, as was a species of 36 kDa at digestion times of 2 and 4 min (Fig. 2D). In the presence of both ligands, the conversion of the 58-kDa TyrR protein to fragments of 22 and 31-kDa was much more rapid (t½, 2–4 min) than in their absence (Fig. 2A).

An identical result was obtained when phenylalanine (2 mM) was substituted for tyrosine. Leucine (2 mM), provided in place of tyrosine, had no effect on the course of trypsin digestion (data not shown). The initial study led to the following conclusions. 1) Two major core fragments of the TyrR protein, of 22 and 31 kDa, were resistant to cleavage by low levels of trypsin. 2) Tyrosine alone altered neither the kinetics nor the pattern of tryptic digestion of the TyrR protein. These

---

**Fig. 1.** Amino acid sequence of the TyrR protein (Cornish et al., 1986). The features of particular relevance to the present study (see text) are a 22-kDa N-terminal domain (residues 1-189) operationally defined by the trypsin-sensitive peptide bond (A) connecting Arg200 and Gln191, a 31-kDa second domain having strong homology to the \( \sigma^4 \) transcriptional activators (residues 191-467) operationally defined by the trypsin-sensitive peptide bonds A and C connecting Arg190 and Gln196, and Arg287 and Phe198, the presumptive ATP-binding sites (bracketed residues 234-240 and 291-297), and the operator recognition element (bracketed residues 492-502) (Pittard and Davidson, 1991). The peptide bond connecting Lys286 and Leu289 is a minor point of trypsin cleavage (see text and Fig. 6).

1 MELVEFCDELRGLTRELLDDLVLRLGIDLRGIEIDPIGRYLYWFALEFESFSSLMHAEIR
1 61 IAGVTDXRTVPWMSPERERBLASALSAAEALPSEPVLYSVMKSVMDMAPASQGFLGQKLDRI
2 121 RIHTAAGLIRGFFPLWLESEEPQASHHREPHVINGQQLMLEITPVYQLDEODERVLGTY
2 A
3 181 MLRLSTRMQDLQLQVAYADYSASFSQIVASYWPWMHEVQQAQQLWMLAAPLTLTGTGTO
3 GL
4 241 DLFATACEQASPRAGPFYLMWGCASEPVAESLEFSGHAPLCKGGFEQANGSVYLDRI
4 B
5 301 QEMSPLANGKLLRLNLDQTFPRQVEHCCVHYVRDICAVTAQKRNLEVQKGFMREDLYRL
5 C
6 361 NVTLLNLPLNRPQDIMPTELTFYRFADEQGVPYPRRCCALDLQVLTTRYAMYGNVQLK
7 421 NATYRALTQLGIDYERLPDLTLPPDTDAATVAVGEDAMEGLDIEFTERSVLTQLTRNY
8 481 PSTRKLAKRLYSGSTIAMLKREFTLSQGKNEE

---

**Table 2.** Summary of experimental conditions employed in limited tryptic digestion of TyrR, of (58 kDa) molecular mass, in the absence of any ligand (A), in the presence of ATP (B), and in the presence of ATP and tyrosine (C). The half-lives of digestion, expressed as mean ± SE, are shown in min.

| Condition | Half-life (min) |
|-----------|----------------|
| A (control) | 16 ± 1.2 |
| B (ATP) | 8 ± 0.4 |
| C (ATP + tyrosine) | 4 ± 0.2 |

---

**Table 3.** Summary of peptides identified by electrospray mass spectrometry in the TyrR protein following limited trypsin digestion.

| Peptide | Molecular mass (Da) | Identification |
|---------|---------------------|----------------|
| A | 53 | Arg200-Gln191 |
| B | 22 | Lys286-Leu289 |
were visualized by staining with Coomassie Blue. The arrows refer to the molecular weight (in thousands) of the tryptic fragments. The times (min) when each sample was removed are shown along the top of each panel. Lanes marked 0 show the TyrR protein samples before trypsin was added. Low levels of trypsin (ratio, 1:200) were used in each of the runs. The precise digestion conditions are described under “Experimental Procedures.” A, no ligands; B, tyrosine (500 μM); C, ATP (200 μM); D, ATP (200 μM) plus tyrosine (500 μM).

FIG. 3. Protection of the 31-kDa tryptic fragment of the TyrR protein by ATP. Samples, collected as described under “Experimental Procedures,” were subjected to 10% SDS-PAGE. The resulting fragments were visualized by staining with Coomassie Blue. Arrows refer to the molecular weights (thousands) of the tryptic fragments. Numbers along the top of each gel indicate the time (min) when samples were taken. Lanes marked 0 were loaded with TyrR protein samples taken before trypsin was added. High levels of trypsin (ratio, 1:11) were employed in this study. Digestion conditions are described under “Experimental Procedures.” A, no ligands; B, with ATP (200 μM).

data are consistent with a previous report that tyrosine cannot bind to the TyrR protein in the absence of ATP (Argyropoulos, 1989). 3) A tryptic fragment of 53 kDa was liberated from the TyrR protein when digestion was carried out in the presence of ATP. 4) The rate of conversion of the TyrR protein to fragments of 22 and 31 kDa was accelerated when both ATP and tyrosine were present.

ATP Protection—In a previous study (Argyropoulos, 1989), it was shown that ATP binds to the TyrR protein. By computer analysis, two presumptive binding sites (residues 234–240 and 291–297) were found in the TyrR protein (Pittard and Davidson, 1991; Fig. 1). If ATP binds to the central region of the TyrR protein, this event might affect the cleavage of nearby trypsin-sensitive bonds. To investigate this possibility, two parallel digestions were carried out at a concentration of trypsin 20 times higher than before (trypsin:TyrR = 1:11).

In the absence of ATP, the TyrR protein was rapidly converted to two species (22 and 31 kDa) (Fig. 3A). The 22-kDa fragment was relatively resistant to further cleavage by trypsin. However, the 31-kDa fragment was completely degraded after 8 min. The half-life of the 31-kDa species was 2–4 min. A new species, of 23 kDa, arose from the 31-kDa subfragment of the TyrR protein. The 29-kDa fragment was itself highly susceptible to further tryptic hydrolysis. This experiment suggested that the 31-kDa tryptic fragment of the TyrR protein, in the absence of ATP, was susceptible to further attack by trypsin when this protease was present at the appropriate concentration. In the presence of ATP (200 μM), the 31-kDa subfragment was resistant to hydrolysis by high concentrations of trypsin (Fig. 3B). The half-life of the 31-kDa fragment was greater than 1 h. Even after incubation for 2 h, substantial amounts of undigested 31-kDa fragment remained (Fig. 3B). The stability of the 22-kDa fragment in the presence of trypsin was unaffected by ATP. The susceptibility to trypsin hydrolysis of the 31-kDa fragment was unaffected by tyrosine (500 μM) (data not shown); when ATP (200 μM) and tyrosine (500 μM) were both present, the course of hydrolysis by trypsin was indistinguishable from that observed in the presence of ATP (data not shown). These data indicate that the 31-kDa tryptic fragment of the TyrR protein, although susceptible to cleavage at high concentrations of trypsin in the absence of ligands, was protected by ATP from trypsin attack. The 22-kDa tryptic fragment was very resistant to digestion by trypsin, even when the ratio of trypsin to the TyrR protein was as high as 1:11.

Effect of Poly(dI-dC)—The behavior of the TyrR protein
during the early steps of purification (see "Experimental Procedures") as well as specific, segments of DNA. To investigate whether the TyrR protein alters its conformation upon binding to DNA, two parallel digestions at low levels of trypsin were carried out in the presence or absence of poly(dl-dC). This synthetic DNA-like polymer was chosen because the amounts of DNA that would have been required for complete binding of the TyrR protein precluded the use of material containing authentic TyrR boxes. In the absence of poly(dl-dC), the TyrR protein was rapidly converted into a species of 45 kDa (Fig. 4A). A minor species of 36 kDa was observed transiently in the early stage of the digestion (Fig. 4A). In the presence of poly(dl-dC), the TyrR protein was rapidly converted into a species of 53 kDa (Fig. 4B), previously observed (Fig. 2, C and D) in the presence of ATP. Trypsin digestion in the presence of poly(dl-dC) led to the accumulation of a previously undetected minor species of 45 kDa (Fig. 4B). The final trypsin-resistant core fragments in this digestion were also polypeptides of 22 and 31 kDa (Fig. 4B). These data suggested that the binding of poly(dl-dC) led to a conformational change in the TyrR protein. In particular, the yield of the 53 kDa tryptic fragment was greatly increased, as compared with the pattern of digestion without poly(dl-dC) (Fig. 4A). Moreover, an intermediate species of 45 kDa was observed only in the presence of poly(dl-dC).

Chemical Identity of the Tryptic Fragments of the TyrR Protein

To characterize the tryptic digestion products of the TyrR protein, each fragment was electrophoretically purified and subjected to sequential Edman degradation analysis. The identity and amounts of the first 6 amino acids at the N terminus of each fragment are presented in Table I. The N-terminal amino acid sequence of the 22-kDa core tryptic fragment (Table I) proved that this species originated from the N-terminal end of the TyrR protein. There are three possible sites of trypsin cleavage (R186, R187, and R190) that could be the C terminus of this fragment. The apparent molecular weight of 22 kDa closely matches the estimated molecular weight of a polypeptide containing amino acid residues 1-187 or 1-190 of the TyrR protein. This observation makes it likely that the 22-kDa fragment terminates at R187 or R190. The N-terminal amino acid sequence of the other main core tryptic fragment (31 kDa) was QLQNVP (Table I), which is that of residues 191-196 of the TyrR protein. This result suggests that R190-Q191 is the preferred site of trypsin cleavage in the formation of the 22- and 31-kDa fragments of the TyrR protein. The apparent molecular weight of the 31-kDa fragment was appreciably lower than the calculated molecular weight (35,974) of a polypeptide containing amino acid residues 191-513, making it unlikely that the 31-kDa fragment included the C terminus of the TyrR protein. The minor trypsin digestion product of 36 kDa (Fig. 2) had the same N-terminal sequence as the 31-kDa fragment (Table I). This polypeptide, in all likelihood, contains amino acid residues 191-513. Therefore, the 31-kDa core fragment must be derived from segment 191-513. To identify the C-terminal end of the 31-kDa fragment, its precise molecular weight was determined by electrospray mass spectrometry. The 31-kDa preparation contained two distinguishable fragments (Fig. 5). The molecular weight of the major form was 30,617, which is very close to the predicted molecular weight (30,597) of segment 191-467 of the TyrR protein. The molecular weight of the minor form was 31,044, which matches the estimated molecular weight (31,029) of amino acid residues 191-470. These data suggest that the TyrR protein consists of two principal domains, namely an N-terminal domain (approximately residues 1-190), and a second domain (residues 191-467).

To detect small polypeptides that might reveal the existence of a domain structure within the C-terminal region, a careful series of experiments deliberately aimed at detecting such a species were conducted. Modifications to the standard protocol included the use of trichloroacetic acid as a protein precipitant and varying the nature of the SDS-PAGE system (data not shown). In no case were trypsin digestion products of 5 kDa or smaller, which might indicate an ability of the C terminus of the TyrR protein to exist as a stable trypsin-resistant species, observed.

The 53-kDa fragment that was detected in tryptic digests conducted in the presence of ATP or poly(di-dC) had the same N terminus as the full-length TyrR protein (Table I). The near identity in molecular weight between the 53-kDa fragment and the combined molecular weight of the 22 and 31-kDa species supports the notion that the 53-kDa fragment includes residues 1-467. Owing to the difficulty of purifying it, the precise molecular weight of the 53-kDa fragment could not be determined by electrospray mass spectrometry.

DISCUSSION

Ligand Effects—The effects of ligands on the pattern of tryptic digestion of the TyrR protein are summarized in Fig. 6. In the absence of any ligands, the major cleavage sites are A and C; B is a minor cleavage site. At low levels of trypsin, the 36-kDa tryptic peptide was not observed in the presence of ATP (Fig. 2C). However, a 53-kDa fragment was readily detectable. Evidently, the binding of ATP enhances the cleavage rate of the R187-P188 peptide bond without affecting the digestion rate of the R190-Q191 peptide bond.

Fig. 4. Effect of poly(di-dC) on the tryptic digestion pattern of the TyrR protein. Samples were analyzed as described in the legends to Figs. 2 and 3. Arrows refer to the molecular weights (thousands) of the tryptic fragments. Numbers along the top of each panel indicate the time (min) when samples were taken. Lanes marked 0 were loaded with TyrR protein samples untreated with trypsin. Low levels of trypsin (1:200) were employed in each case. Digestion conditions are described under "Experimental Procedures." A, no ligands; B, with poly(di-dC) (20 μg/200 μl).
When both ATP and tyrosine were present, the conversion of the TyrR protein to fragments of 22 and 31 kDa was enhanced. Under this condition, fragments of 53 and 36 kDa also became detectable (Fig. 2D). These data suggest that the binding of ATP and tyrosine accelerates the rate of cleavage of both the R₁⁹₀₋Q₁⁹₁ and R₄⁶⁷₋F₄₆₈ peptide bonds. In all likelihood, the binding of ATP to the TyrR protein leads to greater solvent exposure of the R₄⁶⁷₋F₄₆₈ peptide bond; after this has happened, tyrosine binding leads to a conformational change that makes the R₁⁹₀₋Q₁⁹₁ peptide bond more accessible.

Another effect of ATP was the protection of the 31-kDa second domain of the TyrR protein from attack by high levels of trypsin (Fig. 3B). This result is consistent with the predicted locations of two presumptive ATP-binding sites that exist within this region of the TyrR protein. When the TyrR protein was exposed to high levels of trypsin in the absence of ATP, a 23-kDa fragment was generated from the 31-kDa fragment (Fig. 3A). This 23-kDa species had the same N terminus as the 31-kDa fragment (GLQNVA, Table I). From its estimated molecular weight, the 23-kDa fragment probably contains amino acid residues 191–398 (Fig. 6). This segment could well be a subdomain of the second domain.

When trypsin digestion of the TyrR protein was carried out in the presence of poly(d1-dC), there were two major consequences. First, there was substantial accumulation of a 53-kDa fragment and a comparable reduction in the levels of the 22- and 31-kDa fragments (Fig. 4, A and B). The second consequence of digestion in the presence of poly(d1-dC) was the appearance of a new polypeptide of 45 kDa and an absence of the 36-kDa fragment (Fig. 4, A and B). The 45-kDa fragment had the same N terminus as the full-length TyrR protein (Table I). These observations suggest that poly(d1-dC) greatly accelerates the digestion rate of the R₁⁹₀₋Q₁⁹₁ peptide bond and greatly diminishes the digestion rate of the R₁⁹₀₋Q₁⁹₁ peptide bond. The production of the 53-kDa fragment, attributable to hydrolysis at site C, occurs prior to digestion at site A (Fig. 6). The 45-kDa fragment arises from a hydrolytic event at minor site B prior to cleavage at site A (Fig. 6). Evidently, upon binding to poly(d1-dC), conformational change occurs within the TyrR protein such that site C becomes the most solvent exposed site, and site A becomes much less exposed, in comparison to unliganded TyrR protein.

To a first approximation, the poly(d1-dC) effect (Fig. 4) can be considered to mimic the effect of the natural TyrR boxes. In a previous spectral study, the immunity repressor of phage λ was shown to undergo identical conformational changes either upon interaction with specific operator DNA or with DNA-like polymers (Saha et al., 1992). A similar finding was made in a study of the cyclic AMP-binding protein (Angulo and Krakow, 1986). In another study, poly(d1-dC) was shown to stimulate the ATPase activity of NtrC, a σ⁴₅-specific transcriptional activator, in a manner similar to that of authentic target DNA for this protein (Austin and Dixon, 1992). It is therefore not unreasonable to suppose that the binding of the TyrR protein to poly(d1-dC) is associated with a conformational change that is similar to that which occurs normally during transcriptional regulation.

Both ATP and poly(d1-dC) led to a greater solvent exposure of the R₄⁶⁷₋F₄₆₈ peptide bond. The hydrolysis of the R₁⁹₀₋Q₁⁹₁ peptide bond increased in the presence of ATP plus tyrosine and was diminished in the presence of poly(d1-dC). These data raise the possibility that the accessibility of the R₁⁹₀₋Q₁⁹₁ peptide bond to trypsin reflects conformational events that occur when this protein exercises its transcriptional regulatory function.

**Domain Structures—**Our results suggest that the TyrR protein is organized as two distinct domains, encompassing about 91% of the amino acid residues. The two substructures consist of an N-terminal domain (approximately residues 1–190) and a second domain (residues 191–467). The remaining 9%, at the C terminus, containing the operator recognition elements, was never observed as a discrete tryptic peptide, despite numerous attempts to detect a species of the predicted size.

---

**Table I**

| Edman degradation cycle amino acid | PTH-derivative | Amount (pmol) | PTH-derivative | Amount (pmol) |
|-----------------------------------|----------------|--------------|----------------|--------------|
| 53 kDa                            | M              | 18           | M              | 2            |
| 46 kDa                            | R              | 17           | R              | 2            |
| 36 kDa                            | L              | 13           | L              | 2            |
| 31 kDa                            | E              | 15           | E              | 2            |
| 23 kDa                            | V              | 22           | V              | 1            |
| 22 kDa                            | F              | 29           | F              | 2            |

**Fig. 5.** Mass spectrometric analysis of the 31-kDa tryptic fragment of the TyrR protein. For details on the preparation and analysis of this species, see text.

For a description of the isolation of the peptides whose analysis is described, see "Experimental Procedures."
The second domain of the TyrR protein shows substantial identity with the so-called "central" domains of the activators specific for the $d^{4}$ form of RNA polymerase (Fig. 1) despite the fact that the regulated promoters utilize the $d^{70}$ form. These segments of the $d^{4}$ activators bind ATP. In one such activator, NtrC, ATP hydrolysis accompanies the conversion of the RNA polymerase-promoter complex from the "closed" to the "open" form, thereby enabling transcription to proceed (Popham et al., 1989). Two ATP-binding sites were predicted to lie within the second domain of the TyrR protein. Our observation that ATP could prevent a trypsin cleavage event within the second domain of the TyrR protein confirms and extends previous in vitro studies wherein it was demonstrated that ATP binds to the TyrR protein (Argyropolous, 1989).

Despite the accumulated evidence that the second domain of the TyrR protein binds ATP, the in vivo physiological role of this event is unknown.

Although the first 190 amino acids of the TyrR protein are organized in the form of a structurally discrete domain, the function of this segment, or how it interacts with the remainder of the protein, has not been established. The role of the N-terminal domains of the $d^{4}$ activators is to participate in regulatory interactions with other factor(s). For example, the N-terminal domain of NtrC contains a residue (Asp$^4$) that is phosphorylated in a reaction catalyzed by NtrB. NtrC must be phosphorylated in order to carry out its activation function (Ninfa and Magasanik, 1986). The N-terminal domain of the TyrR protein contains amino acid residues that are essential for the activation function (Cui and Somerville, 1993a, 1993b). By analogy with the NtrC case, it is reasonable to suggest that the role of the N-terminal domain of the TyrR protein is to originate transcriptional regulatory signals that are transmitted via the second domain to an operator recognition region.

The $d^{4}$-specific activators are thought to contain glutamine-rich connectors between their N-terminal and central domains (Wootton and Drummond, 1989). These connectors are considered to play a role in transmitting regulatory signals between domains. The insertion of numerous extra amino acids into these connectors does not lead to a loss of function (Wootton and Drummond, 1989). From an inspection of the amino acid sequence of the TyrR protein, it is likely that a similar connector (residues 186-200) is situated between the two major domains of the TyrR protein. Such linkers are also present in certain $d^{70}$-specific activator proteins that are members of two-component bacterial regulatory systems (Wootton and Drummond, 1989).

The concept that large proteins may contain multiple, distinct, independently folding structural regions was enunciated about 20 years ago (Wetlaufer, 1973). One serviceable definition for a protein domain is "a structurally independent compact globular region consisting of a continuous stretch of polypeptide chain, from 100–200 amino acids long" (Coggins and Hardie, 1986). Structural domains of proteins are frequently resistant to digestion by low levels of certain proteases. Limited proteolysis has not previously been employed to study either the TyrR protein (a $d^{70}$-specific transcriptional regulatory factor) or any of the homologous prokaryotic activator proteins, all of which happen to be specific for the $d^{4}$ form of RNA polymerase. It has been suggested that this family of proteins is organized in the form of three structural domains, namely an N-terminal domain, a central domain, and a C-terminal domain (Drummond et al., 1986; Nixon et al., 1986). We were able to detect only two domain structures in the TyrR protein. The C-terminal region, rich in trypsin-sensitive peptide bonds, appears to be loosely folded. Our evidence does not exclude the possibility that amino acid residues 468-513 are organized as a distinct substructure. However, by the criterion of limited proteolysis, the C terminus of the TyrR protein fails to qualify as a domain. The same situation may well apply to NtrC, DctD, NiFA, and other members of this family of proteins. The present study is the first chemical demonstration of an authentic domain structure in any member of the family of prokaryotic transcriptional activators homologous to TyrR. Our results provide a chemical basis for localizing domain boundaries. This information may be useful in the future functional analysis of proteins that have structural homologies to the TyrR protein.

Acknowledgments—For advice and technical assistance, we thank Weiping Yang and Ttie-Leou Shieh. For assistance with computational tasks, we thank Yhor Skrypka. For careful mass spectrometric analyses, we thank Kristine Swiderek of the City of Hope National Cancer Institute. We are grateful to Virginia Heatwole, Guo-Ping Zhao, and Gaochao Zhou for advice and comments during the course of this work and the preparation of the manuscript and to Janell Rex for helpful discussions.

REFERENCES

Angulo, J., and Krakow, J. S. (1986) J. Biol. Chem. 261, 11315–11319

Argyropolous, V. P. (1989) Studies on the Regulatory Protein TyrR from...
Domain Structure of the TyrR Protein

Escherichia coli K12, Ph.D. thesis, University of Melbourne, Melbourne, Australia
Austin, S., and Dixon, R. (1992) *EMBO J.* **11**, 2219-2228
Bradford, M. (1976) *Anal. Biochem.* **72**, 248-254
Camakaris, H., and Pittard, J. (1982) *J. Bacteriol.* **150**, 70-75
Coggins, J. R., and Hardie, D. G. (1986) in *Multidomain Proteins: Structure and Evolution* (Hardie, D. G., and Coggins, J. R., eds) pp. 1-12, Elsevier Science Publishers B. V., Amsterdam
Cornish, E. C., Argyropoulos, V. P., Pittard, J., and Davidson, B. (1986) *J. Biol. Chem.* **261**, 403-410
Cui, J., and Somerville, R. L. (1989a) *J. Bacteriol.* **175**, 303-306
Cui, J., and Somerville, R. L. (1993b) *J. Bacteriol.*, in press
Drummond, M. H., Whitty, P., and Wootton, J. (1986) *EMBO J.* **5**, 441-447
Gottesman, M. E., and Yarmolinsky, M. B. (1968) *J. Mol. Biol.* **31**, 487-505
Heatwole, V. M., and Somerville, R. L. (1991) *J. Bacteriol.* **173**, 3601-3604
Kasian, P. A., Davidson, B. E., and Pittard, J. (1986) *J. Bacteriol.* **167**, 556-561
Lennox, E. S. (1955) *Virology* **1**, 190-206
Ninfa, A. J., and Magasanik, B. (1986) *Proc. Natl. Acad. Sci. U. S. A.* **83**, 5909-5913
Nixon, B. T., Ronson, C. W., and Ausubel, F. M. (1986) *Proc. Natl. Acad. Sci. U. S. A.* **83**, 7550-7554
Pittard, A. J., and Davidson, B. E. (1991) *Mol. Microbiol.* **5**, 1585-1592
Popham, D. L., Szeto, D., Keener, J., and Kustu, S. (1989) *Science* **243**, 629-635
Rosenberg, A. H., Lade, B. N., Choi, D.-S., Lin, S.-W., Dunn, J. J., and Studier, F. W. (1987) *Gene (Amst.*) **66**, 125-135
Rossmann, M. G., and Argos, P. (1981) *Annu. Rev. Biochem.* **50**, 497-532
Saha, B., Banik, U., Bandopadhyay, S., Mandal, N. C., Bhattacharyya, B., and Roy, S. (1992) *J. Biol. Chem.* **267**, 5862-5867
Sarsro, J. P., Wooley, P. J., and Pittard, A. J. (1991) *J. Bacteriol.* **173**, 4133-4143
Schägger, H., and Von Jagow, G. (1987) *Anal. Biochem.* **166**, 368-379
Somerville, R. L., Shieh, T. N., Hageswood, B., and Cui, J. (1991) *Biochem. Biophys. Res. Commun.* **181**, 1066-1069
Stock, J. B., Ninfa, A. J., and Stock, A. M. (1988) *Microbiol. Rev.* **63**, 450-490
Wetlaufer, D. B. (1973) *Proc. Natl. Acad. Sci. U. S. A.* **70**, 697-701
Williams, B. C., and Blaftner, F. R. (1979) *J. Virol.* **29**, 555-575
Whipp, M. J., and Pittard, A. J. (1977) *J. Bacteriol.* **132**, 453-461
Wootton, J. C., and Drummond, M. H. (1989) *Protein Eng.* **2**, 535-543