Isolation and Characterization of a Small Catalytic Domain Released from the Adenylate Cyclase from Escherichia coli by Digestion with Trypsin*

(Received for publication, December 1, 1987)

Mitchell M. Holland, Terry K. Leiby, and John A. Gerlt
From the Department of Chemistry and Biochemistry, University of Maryland, College Park, Maryland 20742 and the Center for Advanced Research in Biotechnology, Rockville, Maryland 20850

An expression plasmid containing a hybrid gene encoding a protein having the primary amino acid sequence of the adenylate cyclase from Escherichia coli was constructed. When the gene was induced, the adenylate cyclase could be expressed at high levels in a cya- strain of E. coli. The majority of the enzymatic activity and protein (having a molecular weight of 95,000) induced was insoluble. However, treatment of the insoluble fraction of cell lysates with trypsin resulted in both an increase in and solubilization of the total amount of adenylate cyclase activity. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis of the soluble protein produced by treatment with trypsin revealed a polypeptide having a molecular weight of 30,000. This soluble, catalytically active fragment of adenylate cyclase was purified and subjected to amino-terminal sequence analyses; two amino-terminal sequences were identified beginning at residue 82 and at residue 342 of the intact enzyme. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis analysis of the purified fragment followed by either silver or Coomassie Blue staining revealed the presence of only a single polypeptide having a molecular weight of 30,000; a short oligopeptide associated with the amino terminus at residue 342 could not be detected. Site-directed mutagenesis was used to place a stop codon at residue 341; the truncated enzyme was catalytically active, so the short oligopeptide is not necessary for catalysis. The $K_\text{m}$ for ATP, the $K_\text{a}$ for Mg$^{2+}$, and the $V_{\max}$ determined for the product containing the 30,000-dalton fragment were similar to the values reported for the intact enzyme from E. coli.

Adenylate cyclases (ATP pyrophosphate-lyase (cyclizing), EC 4.6.1.1) catalyze the intramolecular transfer of the adenylyl group of ATP from pyrophosphate to the 3'-hydroxyl group to form cyclic AMP. The activities of the adenylate cyclases from most, if not all, organisms appear to be regulated by various metabolites and stimuli. For example, the enzymes from Escherichia coli and Salmonella typhimurium are subject to catabolite repression (1, 2), and the enzymes from mammalian sources are subject to regulation by a number of hormones (3, 4). In addition, the detailed chemical mechanisms of the reactions catalyzed by enzymes from diverse sources are likely to be analogous, given the identical stereochemical consequences of the reactions catalyzed by the enzymes from Brevibacterium liquefaciens (5, 6), Bordetella pertussis (7), and bovine brain (8). Furthermore, no convenient source of an enzyme that is amenable to mechanistic scrutiny has been identified or developed. Given our interest in studying the mechanism of this metabolically important reaction, we are investigating the feasibility of using recombinant DNA technology to isolate large amounts of enzyme.

Several years ago, this laboratory reported that the enzyme from S. typhimurium contained a catalytic domain (9). The evidence for this conclusion was based upon measurements of catalytic activity associated with extracts of E. coli cells that contained adenylate cyclases suffering deletions of large portions of the carboxyl-terminal ends of the protein. The deletions in the protein were produced by Bal31 digestion of the 3'-end of the cloned gene, and these truncated genes were cloned in pBR322 and transformed into cya- host strain of E. coli. DNA sequence analysis of the smallest truncated but catalytically active enzyme revealed that the protein was composed of 419 amino acids and had a molecular weight of approximately 45,000.

A similar conclusion about the presence of a catalytic domain in the enzyme from E. coli was reached essentially simultaneously by Danchin and his co-workers (10, 11), although the approach used to truncate the carboxyl-terminal end of the protein differed from ours. The gene for the enzyme from E. coli contains a unique HpaI site at codon 556 (the intact gene contains 848 codons). Termination at this site results in a protein having a molecular weight of approximately 64,000 which can complement a cya- phenotype. This truncated gene has been fused to the eighth codon of $\beta$-galactosidase, resulting in production of a hybrid protein having a molecular weight of 175,000 and both adenylate cyclase and $\beta$-galactosidase activities. Exposure of the hybrid protein to trypsin resulted in the rapid loss of adenylate cyclase activity with no significant effect on the $\beta$-galactosidase activity; analysis of this digest by SDS-PAGE revealed the generation of only the $\beta$-galactosidase polypeptide with no tryptic peptides associated with the adenylate cyclase catalytic domain being discernible (12). Furthermore, determination of the kinetic properties of the fusion protein disclosed that the adenylate cyclase activity was not subject to substrate inhibition as had been previously described for the intact enzyme. (Danchin and his co-workers (10, 11) also reported that the gene could be truncated at codon 377 to

*This research was supported by National Institutes of Health Grant GM-34572. 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.

‡ Present address: General Electric Research Laboratories, Schenectady, NY 12301.

† The abbreviations used are: SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis; FPLC, fast protein liquid chromatography.
produce a protein that could complement a cyt− phenotype but did not pursue further characterization of the product of this truncation.

We now report our observations regarding the proteolytic generation of a catalytic domain from the intact adenylate cyclase from E. coli that has been overproduced in an E. coli host that is cyt−. A thermally inducible expression plasmid containing a gene for the intact adenylate cyclase from E. coli under transcriptional control of the bacteriophage λ PPr promoter has been constructed. Induction of the gene leads to the accumulation of large amounts of adenylate cyclase (20–30% of the total cellular protein) as a largely insoluble but at least partially catalytically active protein. Exposure of the insoluble protein to trypsin results in solubilization of the catalytic activity; analysis of the time course of the trypsin reaction by SDS-PAGE reveals the initial generation of a 40,000-dalton polypeptide which is subsequently further proteolyzed to yield a 30,000-dalton polypeptide. Purification of the latter reaction product affords an apparent catalytic domain having two amino termini which are located within the amino-terminal half of the intact enzyme; one can be associated with a polypeptide of molecular weight 30,000 and the second with an oligopeptide (presumed to have a very low molecular weight) which cannot be detected by SDS-PAGE and either silver staining or Coomassie Blue staining. Site-directed mutagenesis was used to place a stop codon at the second amino terminus to ascertain the importance of this oligopeptide, and cells transformed with the plasmid containing the gene for the truncated enzyme produced an insoluble but catalytically active adenylate cyclase. Thus, we conclude that the catalytic domain of the adenylate cyclase is located between codons 82 and 341 of the intact enzyme. Furthermore, the molecular weight 30,000 trypsinic product displays kinetic properties similar to those reported previously for the intact enzyme (13).

EXPERIMENTAL PROCEDURES

Restriction endonucleases and bacteriophage T4 DNA ligase were purchased from New England Biolabs. The Klenow fragment of E. coli DNA polymerase I, bacteriophage T4 DNA ligase, and all restriction endonuclease enzymes from Boehringer Mannheim were purchased from New England Biolabs. Bovine pancreas trypsin and soybean trypsin inhibitor were purchased from Sigma. All enzymes were used according to the suppliers' instructions. Protein concentrations were determined by the method of Lowry et al. (14). Proteins were detected in polyacrylamide gels by silver staining with the reagents and protocol obtained from Bio-Rad. The Mutagenic excision kit was also obtained from Bio-Rad for the mutagenesis studies.

Plasmids and Bacterial Strains—The plasmid pTL203 was previously obtained by BaSal digestion from the 3′-end of gene for the adenylate cyclase from S. typhimurium (9). This plasmid encodes the catalytic domain of this enzyme, which by DNA sequence analysis is composed of the amino-terminal 419 amino acids of this enzyme; a unique EcoRI restriction site is located at the 3′-end of the coding sequence. A unique HindIII restriction site subsequently was placed at the initiation codon of this coding sequence by BaSal digestion from the 5′-end of the gene followed by ligation of a HindIII linker; this plasmid is designated pTL104. The plasmid pDIA100 which contains the gene for the intact adenylate cyclase from E. coli was obtained from Dr. Alan Peterkofsky (National Institutes of Health, Bethesda, MD). The expression vector pCQV2 (15) was obtained from Dr. Cary Queen (National Institutes of Health). The expression vector pCQV2 was linearized at the BamHI cloning site; following digestion of the plasmid with the Klenow fragment of E. coli DNA polymerase I, bacteriophage T4 DNA ligase, and the blunt-ended vector was dephosphorylated with calf intestine alkaline phosphatase. The fragment bearing the gene for the catalytic domain and the vector were ligated, and the reaction mixture was used to transform JG1. This construction inserts the leterapetide Apl-Gln-Leu-Val between the amino-terminal Met and the penultimate Tyr of the intact enzyme; however, codon numbers cited in this paper will refer to the sequence of the intact enzyme. Plasmids isolated from transformants were screened for the desired orientation of the gene by restriction analysis, and one with the proper orientation was designated pMH100 and selected for further manipulations. Induction of the catalytic domain in JG1 transformed with pMH100 can be accomplished by growing the transformed cells to midlog phase (OD_{600 nm} = 1.0) at 30 °C and rapidly raising the temperature of the culture to 42 °C by the addition of hot medium.

Construction of an Expression Plasmid for the Intact Enzyme from E. coli—pMH100 contains a unique BstEII site at codon 51 of the gene for the catalytic domain from S. typhimurium as well as a unique SalI site downstream of the gene. pDIA100 contains the analogous unique BstEII restriction site at codon 51 of the gene for the intact enzyme from E. coli as well as a unique SalI restriction site downstream of the gene. Sequence analyses of the gene for the catalytic domain from the enzyme S. typhimurium and of the gene for the intact enzyme from E. coli reveal that the amino acid sequences from the initiation codon to the unique BstEII restriction sites are identical (although differences in the wobble positions of 12 codons are present). Thus, excision of the BstEII-SalI fragment bearing the bulk of the gene for the catalytic domain of the enzyme from S. typhimurium from pMH100 and replacement with the BstEII-SalI fragment from pDIA100 resulted in the construction of a hybrid gene for the intact enzyme from E. coli in the expression vector pCQV2; this expression plasmid was designated pMH105. Induction of the alkaline phosphatase from E. coli in JG1 transformed with pMH105 was accomplished by growing the transformed cells to midlog phase at 30 °C and rapidly raising the temperature of the culture to 42 °C by the addition of hot medium.

Digestion of Insoluble Intact Enzyme with Trypsin—JG1 transformed with pMH105 and induced for the production of the intact enzyme was harvested after 2 h at 42 °C; the cells were quickly frozen in liquid nitrogen and stored at −70 °C. One gram of frozen cells was thawed in 4 ml of TD buffer (50 mM Tris-HCl, pH 7.8, 5 mM dithiothreitol, 1 mM EDTA, and 20% glycerol) and lysed by passage through a French pressure cell at 18,000 p.s.i. The insoluble fraction was removed by centrifugation at 16,000 × g for 15 min and resuspended in 8 ml of TD buffer. Trypsin (250 μg dissolved in 10 μl of H2O) was added to a 2-ml aliquot of this suspension, and the reaction was allowed to proceed at 20 °C. After 2 h, an additional 250 μg of trypsin was added. After a total reaction time of 4 h, soybean trypsin inhibitor (1 mg dissolved in 20 μl) was added, and the reaction was chilled on ice. The suspension was clarified by centrifugation at 4 °C (15 min at 40,000 × g followed by 1 h at 100,000 × g). The supernatant was either used immediately or stored at −20 °C (where the activity is stable indefinitely).

Digestion with trypsin was also carried out for only 10 min in order to generate early time point products, using the conditions described in the previous paragraph. Upon stopping the reaction after 10 min by the addition of soybean trypsin inhibitor, the soluble fraction was isolated as described in the preceding section and, if necessary, stored at −20 °C.

Expression of Insoluble Intact Enzyme—All steps were carried out at 4 °C. Twenty milliliters of supernatant obtained as described in the previous section (after either 10 min or 4 h digestion with trypsin) was diluted to 100 ml with 80 ml of TD buffer. A solution of 41.8 g of (NH4)2SO4 in 100 ml of TD buffer was added slowly to the diluted supernatant, and the suspension was allowed to stir for 6 h before centrifugation (30 min at 16,000 × g) to collect the precipitated protein. The pellet was dissolved in 20 ml of TD buffer and dialyzed overnight against 2 liters of TD buffer. Approximately 70% of the adenylate cyclase activity was recovered in the resuspended and dialyzed precipitate.

Six milliliters of the dialyzed solution was diluted to 24 ml with 18 ml of TD buffer, and this solution was applied to a Mono Q anion
of the gene) or the unique EcoRI (3'-end of the gene) restriction site followed by ligation of either a HindIII or EcoRI linker and cloning of the HindIII-EcoRI fragment obtained by restriction digestion into intact pBR322. Restriction analysis of transformants allowed selection of deletions suitable for nested DNA sequencing. Commercially available sequencing primers for clockwise sequencing from the EcoRI site and counterclockwise sequencing from the HindIII site in pBR322 were used for DNA sequencing by the chain terminating dideoxy methodology. Using these procedures, the coding sequence for the catalytic domain was determined in its entirety on both strands. After this sequence determination was completed, the complete sequence of the gene for the adenylate cyclase from E. coli was reported by Dankin et al. (18); the sequence differences obtained by comparison of the first 419 codons of the genes from E. coli and S. typhimurium are detailed in Fig. 1. (The figure emphasizes the sequence determined for DNA from E. coli, since the remainder of this paper focuses on the catalytic domain of the enzyme from E. coli.) Clearly, the presumed catalytic domains in the enzymes from the two sources are highly homologous, with only 179 base pair and 15 amino acid changes being observed.

In work that is also described in detail elsewhere (17), we fused the gene for the catalytic domain of the enzyme from S. typhimurium to an initiation codon following the bacteriophage P1 promoter and cII gene ribosome binding site as found in the expression vector pAS1. We expressed the gene for the catalytic domain in a suitable E. coli host whose chromosome contained a lytic deficient λ phage bearing the gene for the cl857 temperature-sensitive mutation of the λ repressor. The catalytically inactive protein that was produced had a molecular mass of approximately 45,000 daltons and was found in the insoluble fraction of the induced cells. However, the protein could be solubilized by exposure to 5 mM guanidinium hydrochloride buffered at pH 8 with Tris-HCl and containing 5 mM β-mercaptoethanol and 1 mM EDTA followed by dialysis against the same buffer without the denaturant. The soluble protein so produced was catalytically active and nearly homogeneous as assessed by SDS-PAGE; however, chromatography of this highly unstable protein on DEAE-Sepharadex revealed considerable heterogeneity with very low recoveries of activity. For these reasons, we have pursued the proteolytic generation of the catalytic domain from intact enzyme as described in this paper.

**RESULTS AND DISCUSSION**

**DNA Sequence Analysis of the Catalytic Domain in the Adenylate Cyclase from S. typhimurium**—The portion of the gene for adenylate cyclase from S. typhimurium contained in the previously described pTL203, the plasmid in which we genetically defined catalytic domain was cloned in pBR322, was subjected to DNA sequence analysis; the details of this sequence analysis are described in detail elsewhere and will be only summarized in this paper (17). The strategy for the sequencing employed progressive deletions using Bal31 from either the unique HindIII upstream sequences at the 5’-end
FIG. 2. Plasmid constructions. The construction of expression plasmids containing the coding sequence for the catalytic domain of the enzyme from *S. typhimurium*, designated pMH100, and the coding sequence for the intact enzyme from *E. coli*, designated pMH103, are summarized. Both plasmids utilize the expression vector pCQV2; the details of the manipulations are described in the text.

with the solubilizing protein being both stable and catalytically active.

pMH100 serves as a convenient vector for cloning and expressing the intact enzyme from *E. coli*. As revealed by the sequence data presented in Fig. 1, the genes for the catalytic domain of the enzyme from *S. typhimurium* and the intact enzyme from *E. coli* both contain unique BstEII restriction sites (GGTTACC) at codon 51; in addition, both genes are followed by unique SalI restriction sites. Thus, restriction digestion of pMH100 and of pDIA100 with BstEII and SalI followed by ligation of the fragment containing the bulk of the gene for the intact enzyme from *E. coli* with the larger fragment derived from pMH100 results in the construction of a plasmid containing a hybrid gene for the intact enzyme from *E. coli*; this plasmid is designated pMH103. Growth of JG1 transformed with pMH103 in LB at 30 °C followed by a rapid increase in the temperature of the medium to 42 °C results in the rapid production of a polypeptide having a molecular weight of approximately 95,000. The protein produced by this induction is largely insoluble but partially catalytically active, in contrast to the activity associated with the previously described catalytic domains produced by heat induction of λ promoters (P<sub>1</sub> and P<sub>R</sub> for *S. typhimurium*). When grown at 30 °C on McConkey plates containing lactose, JG1 transformed with pMH103 gives a pink phenotype.

**Production of a Soluble Catalytic Domain from Insoluble Intact Enzyme by Digestion with Trypsin**—We have not attempted denaturation mediated solubilization of the catalytic domain of the enzyme from *E. coli* for adenylyl cyclase from *E. coli* (top) and *S. typhimurium* (bottom) and of the translated protein sequences are compared; the only protein primary sequence data given for the translated protein from *S. typhimurium* are the differences from the sequence from *E. coli*.
Under "Experimental Procedures," we have subjected the insoluble fraction of induced cells directly to the action of trypsin. This reaction was monitored both by measurement of soluble and total enzymatic activity (Fig. 3) as well as by analysis of the soluble and insoluble proteins by SDS-PAGE (Fig. 4). Both the soluble and total enzyme activity increased significantly (20- and 2-fold, respectively) until these activities were equal; with time the total enzyme activity eventually decreased from the observed maximum value. SDS-PAGE revealed that this increase in soluble and total catalytic activity was accompanied by loss of the insoluble 95,000-dalton intact enzyme and initial production of a polypeptide having a molecular weight of 40,000; upon further digestion by trypsin, this protein disappeared and was replaced by a polypeptide having a molecular weight of 30,000. The digestion was stopped by the addition of a molar excess of soybean trypsin inhibitor. In contrast to the stability of the guanidinium hydrochloride solubilized catalytic domain (having a molecular weight of 45,000), the trypsin solubilized catalytic domain can be stored indefinitely at −20 °C with no detectable loss of catalytic activity.

These observations contrast sharply with those of Crenon et al. (12) on the trypsin digestion of the fusion protein generated by fusion of the first 556 codons of the gene for the enzyme from E. coli to the eighth codon of β-galactosidase. This soluble protein which contains both adenylate cyclase and β-galactosidase activities is rapidly degraded by trypsin, with the product of the reaction being a trypsin resistant polypeptide having the molecular weight of β-galactosidase; this polypeptide retains β-galactosidase activity. The adenylate cyclase is rapidly lost, and analysis of the digestion by SDS-PAGE does not reveal the production of any large fragments of the adenylate cyclase portion of the fusion protein. A definitive reason for the trypsin sensitivity of this protein and our ability to generate a stable and (reasonably) trypsin-resistant catalytic domain from the insoluble intact enzyme is unknown, although we assume that conformational differences between the two proteins must be important.

**Purification and Kinetic Characterization of the Trypsin-Solubilized Catalytic Domain**—As detailed under "Experimental Procedures," the fraction containing the polypeptide having a molecular weight of 30,000 and presumably comprising the catalytic domain of the enzyme from E. coli can be purified to near homogeneity by an initial ammonium sulfate fractionation followed by gel filtration on Superose 12 and ion exchange chromatography on a column of Mono Q resin. Multiple species with adenylate cyclase activity are eluted from the Mono Q column; however, analysis of the fractions containing catalytic activity by SDS-PAGE reveal that all of these fractions contain a polypeptide having a molecular weight of approximately 30,000. The presence of multiple catalytically active species presumably can be explained by multiple sites of trypsin cleavage to produce the carboxyl terminus and/or the presence of an additional site of trypsin cleavage within the isolated protein; documentation of the latter situation will be discussed in the following section. The level of catalytic activity measured in the fractions from the anion exchange column is in agreement with the amount of the 30,000-dalton polypeptide observed by SDS-PAGE. Fractions containing catalytic activity and only the 30,000-dalton polypeptide were selected for further characterization. A photograph of an SDS-PAGE gel is reproduced in Fig. 5 that shows (lane 1) the soluble polypeptides after trypsin digestion, (lane 2) the soluble polypeptides remaining after ammonium sulfate fractionation, (lane 3) the soluble polypeptides remaining after Superose gel filtration, and (lane 4) the pooled fractions from the Mono Q column. In lane 4 the faint protein having a molecular weight of 40,000 is the apparent kinetic precursor to the major species having a molecular weight of 30,000 (see the next section), and the faint protein having a molecular...
molecular weight of 21,500 is the soybean trypsin inhibitor used to stop the digestion with trypsin. The presence of low molecular weight oligopeptides cannot be detected by this analysis (although the presence of such peptides are implicated by the amino-terminal sequence analyses described in the following section). Approximately 5 μg of highly purified catalytic domain can be obtained from 1 g of induced cells.

The dependence of the reaction velocity on both ATP concentration (at a fixed concentration of 20 mM MgCl₂) and MgCl₂ concentration (at a fixed concentration of 1 mM ATP) have been determined for the purified catalytic domain; the results are presented in Figs. 6 and 7, respectively. The ATP dependence of the reaction velocity catalyzed shows substrate inhibition by ATP (Fig. 6), with the data points for low concentrations of ATP yielding a Vmax of 435 nmol/(min·mg) and a Kₘ of 0.25 mM and the data points for high concentrations of ATP yielding a Kc of 2.8 mM. The Mg⁺ dependence of reaction velocity can be described as hyperbolic (Fig. 7), with the Kc calculated from the data being 9.7 mM. The values measured for the Kₘ and Kc for ATP and the Kc for Mg⁺ are in excellent agreement with those reported by Yang and Epstein (13) for the purified intact enzyme (1.0 and 1.5 mM for ATP and 8.0 mM for Mg⁺, respectively). The value of 435 nmol/(min·mg) for Vmax is also in good agreement with the value reported by Yang and Epstein (13) (700 nmol/(min·mg)). In principle, the Vmax we find for the catalytic domain could have been three times that reported for the intact enzyme if removal of the amino- and carboxyl-terminal polypeptides had been without either direct or indirect influence on catalysis; given the significant decrease in molecular weight accompanying the production of the catalytic domain, we regard the approximate 5-fold lower activity of the catalytic domain as insignificant.

The similarity of the kinetic constants determined for the catalytic domain containing the 30,000-dalton polypeptide with those previously reported for the intact enzyme provide persuasive evidence that we have, in fact, produced by tryptic digestion a catalytic domain which has a molecular weight approximately one-third of that of the intact enzyme. Assuming that the small oligopeptide implicated by the amino-terminal sequence analyses described in the next section is no longer than 20 amino acid residues in length (and the data in Fig. 5 strongly suggest that we should have been able to detect an oligopeptide this small) the catalytic domain we have now characterized is the smallest catalyst for a nucleotidyl cyclase reaction yet isolated or generated. The guanylate cyclase from rat liver plasma membrane can be digested with trypsin to yield a catalytic domain having a molecular weight of 68,000 (19) and the adenylate cyclase from ram spermato-
zyme can be digested with chymotrypsin to yield a catalytic domain having a molecular weight of 34,000 (20).

**Placement of the Catalytic Domain within the Primary Sequence of the Intact Enzyme by Amino-terminal Sequence Determination**—A fraction of the catalytic domain obtained by Sepharose 12 gel filtration and Mono Q anion exchange chromatographies has been subjected to automated amino-terminal sequence analysis. Ten cycles of degradation were performed, with the identity of the first residue being uncertain due to contaminating ions in the protein sample. Two amino-terminal sequences were present and could be unequivocally identified within the primary structure of the intact protein deduced from the DNA sequence. The two sequences were identified as Met-Ser-Val-Gln-Asp-Pro-Pro-Lys-Gly and Cys-Val-Gly-Trp-Arg-Arg-Ala-Val-Leu (with the identity of the Cys being based upon the absence of a second amino acid in cycle 2, the known inability of the sequencing procedure to detect Cys, and the exact correspondence of the remaining sequence to the primary structure of the intact protein). The sequence of the first 419 amino acids in the intact enzyme shown in Fig. 1 reveals that the first amino-terminal sequence is at residue 82, with the actual amino terminus of this polypeptide being Gly. The second amino terminus can be located at residue 342, with the actual amino terminus being Ala. These amino-termini are consistent with the use of trypsin to solubilize and produce the catalytic domain since residues 81 and 341 are both Arg.

A molecular mass of 30,000 daltons suggests that the predominant catalytically active polypeptide has its amino terminus at residue 82 and that its carboxyl terminus is located approximately at residue 340 of the intact enzyme. The observation of a second amino terminus beginning at residue 342 suggests that trypsin has also cleaved at Arg-341 to produce an oligopeptide of unknown but presumably low molecular weight since it cannot be detected by SDS-PAGE followed by either silver (Fig. 5) or Coomassie Blue (data not shown) staining. This smaller peptide is presumably noncovalently associated with the 30,000-dalton fragment via hydrophobic or electrostatic interactions, since they are not separated by gel filtration and no difference in electrophoretic mobility can be detected if the SDS-PAGE of the purified trypptic product is performed under nonreducing conditions (data not shown). As noted previously, our catalytically active species containing two polypeptides is somewhat chromatographically heterogeneous (two or three species may exist). We assume that the heterogeneity can be explained by cleavage of trypsin at one of several Arg and Lys residues in the carboxyl-terminal region of the primary sequence of the 30,000-dalton polypeptide.

Using procedures described in detail under “Experimental Procedures,” we have used site-directed mutagenesis to ascertain whether the second oligopeptide beginning at residue 342 is important for catalytic activity. Two successive stop codons and a single base insert were placed after condon 341 to ensure that translation would be terminated at this position. The plasmid pMH40 containing this mutation was transformed into JG1, and the observed pink phenotype on McConkey plates indicates that the adenylate cyclase encoded by the truncated gene is catalytically active. After heat induction of the truncated protein, SDS-PAGE revealed that the amino-terminal half of the intact enzyme that we previously had implicated via deletions at the 3′-end of the gene from S. typhimurium as containing a catalytic domain. To support this conclusion we have also determined the amino-terminal sequence of the early trypsin product containing the 40,000-dalton polypeptide. Two amino-terminal sequences were also found (after the first cycle in which no residue could be unequivocally identified) for this catalytically active species, with the sequences being Asp-Gln-Leu-Val-Tyr-Leu-Tyr-Ile-Glu and Cys-Val-Gly-Trp-Arg-Arg-Ala-Val-Leu. The first of these is the predicted amino terminus of the intact protein encoded by pMH103, and the second is the same sequence previously described as starting at residue 342. Since this species is the kinetic precursor of the species containing the 30,000-dalton polypeptide, we can be assured that the placement of the catalytic domain within the amino-terminal half of the intact enzyme as previously discussed is correct. Thus, the catalytic domain produced by trypsin digestion is cleaved by rapid trypptic cleavage in the region of residue 341 and at an additional site(s) downstream of this residue. Subsequent slower cleavage at residue 81 produces the final species which is relatively resistant to further digestion by trypsin.

In contrast to the ability to purify a soluble catalytic domain from insoluble enzyme encoded by the intact gene, trypsin digests of the insoluble fractions produced by truncating the primary sequence at either codon 342 by site-directed mutagenesis (encoded by pMH40) or at codon 42 (encoded by pMH45) by both site-directed mutagenesis and Bal31 deletions from the 3′-end of the gene have consequently resulted in solubilization of only small amounts of soluble active enzyme having a molecular weight of 30,000 as evidenced by SDS-PAGE and enzyme assays (data not shown). Additionally, attempts to purify the solubilized catalytic domains from those species have been unsuccessful. Although these observations do not detract from our production of a soluble and stable catalytic domain from the insoluble intact enzyme, they emphasize that the success we have achieved with trypytic digestion depends critically upon the identity of the starting primary and tertiary structures that constitute the catalytic domain.

Our characterization of the primary amino acid sequence of the catalytic domain in the adenylate cyclase from E. coli by proteolytic digestion is now directing further manipulation of the gene for the intact enzyme so that large amounts of the catalytic domain may become readily available for detailed structural and mechanistic studies.

**Acknowledgments**—We thank Drs. Antoine Danchin (Institut Pasteur, Paris, France) and Alan Peterkofsky (National Institutes of Health, Bethesda, MD) for pDIA100, the plasmid containing the intact gene for the adenylate cyclase from E. coli, and Dr. P. Shenbagamurthi (The Johns Hopkins University School of Medicine, Baltimore, MD) for assistance with amino-terminal sequence analyses.

**REFERENCES**

Catalytic Domain from Adenylate Cyclase