Proc. Nat. Acad. Sci. USA
Vol. 72, No. 4, pp. 1254–1257, April 1975

Molecular Basis of β-Galactosidase α-Complementation

(protein sequencing/protein conformation/deletion mutant)

KEITH E. LANGLEY, MERNA R. VILLAREJO, AUDREE V. FOWLER, PATRICE J. ZAMENHOF, AND IRVING ZABIN*

Department of Biological Chemistry, School of Medicine, and Molecular Biology Institute, University of California, Los Angeles, Calif. 90024

Communicated by Paul D. Boyer, January 13, 1975

ABSTRACT **In previous studies, a cyanogen bromide peptide derived from amino-acid residues 3–92 of β-galactosidase (EC 3.2.1.23; β-D-galactoside galactohydrolase) was shown to have α-donor activity in intracistronic α-complementation. We have now isolated the defective β-galactosidase α-acceptor protein from the deletion mutant strain M15 of *Escherichia coli* and find that it lacks residues 11–41 of β-galactosidase. This is demonstrated by the isolation and sequence determination of a cyanogen bromide peptide from the M15 protein, which is identical to the corresponding peptide from β-galactosidase except for the missing amino acids. We conclude that the α-donor peptide restores the region missing in the M15 protein.**

Intracistronic complementation is the association between differently altered mutant proteins or protein fragments from the same cistron to give biologically functional protein. For β-galactosidase (EC 3.2.1.23; β-D-galactoside galactohydrolase), the enzyme specified by the *Z* gene of the lactose operon in *Escherichia coli*, α-complementation was described by Ullmann, Jacob, and Monod in 1968 (1). Map positions of mutations in complementing strains showed that mutant strains with an intact operator-proximal region would complement a mutant strain such as M15 with a deletion in the same area. This operator-proximal 1/4 or 1/5 of the genetic length of the *Z* gene, which they defined as the α region, specifies the amino-terminal portion of the β-galactosidase polypeptide chain (2).

Morrison and Zipser reported that a soluble peptide fraction with α-donor activity towards extracts of M15 could be obtained by autoclave treatment of extracts of a variety of wild-type and mutant strains. The active material had a molecular weight of about 7400 (3). Similarly, Lin *et al.* in this laboratory found that cyanogen bromide cleavage of β-galactosidase also yields a peptide with α-donor activity (4). The peptide has now been isolated and determined to be the fragment derived from residues 3–92 of β-galactosidase (5).

As an approach towards understanding the molecular nature of the interaction between this peptide (CB2) and defective β-galactosidase from strain M15, we now report the isolation of M15 protein and the identification of the missing portion of the sequence. M15 protein, although it lacks enzyme activity, has a strong substrate binding site (6) and, therefore, could be purified with the aid of affinity chromatography. From the pure protein a cyanogen bromide peptide (Δ_{M15}CB2) was obtained which was identical to CB2 except for a deleted region corresponding to residues 11–41 of β-galactosidase.

Abbreviations: dansyl, 5-dimethylamino-naphthalene-1-sulfonyl; CB and T refer to cyanogen bromide and tryptic peptides, respectively.

* To whom requests for reprints should be sent.

MATERIALS AND METHODS

Materials. [2-^{14}C]Iodoacetic acid (22.6 mCi/mmol) and [2-^{3}H]iodoacetic acid (96 mCi/mmol) were obtained from Amersham/Searle. The sources of all other chemicals, enzymes, and chromatography resins used have been indicated earlier (5–9).

Bacterial Strains. E. coli K12 strain M15 (2320), chromosomally *lac IZM15*, carries a spontaneous deletion mutation located at the extreme operator-proximal end of the *Z* gene (10). It was obtained from J. Beckwith. Strain RV/F'M15, supplied by A. Ullmann, is chromosomally deleted of the *lac* operon and carries the episome F'*lacIZM15*. For enhanced production of M15 protein, the homomerodiploid strain DZ291, *lacIZM15*/F'*lacIZM15*, was constructed from strains M15 (2320) and RV/F'M15, using standard sexduction procedures (11, 12). *E. coli* strain A324-5 (*lacI*/F'*lacI*) was obtained from Eugene P. Kennedy.

Isolation of β-Galactosidase and M15 Protein. β-Galactosidase was isolated from A324-5 as described (13). DZ291 was grown, and an extract containing M15 was prepared and treated with ammonium sulfate in the same way as for β-galactosidase. The ammonium sulfate fraction in 0.1 M sodium phosphate, 70 mM NaCl, 10 mM 2-mercaptoethanol, pH 7.2, was applied to a β-galactosidase substrate analog affinity column (6, 14) (3.5 × 30 cm) equilibrated with the same buffer at room temperature, and the column was washed with the buffer. M15 protein was retained on the column and was eluted with 0.1 M sodium phosphate, 1.12 M NaCl, 10 mM 2-mercaptoethanol, pH 7.2.

Carboxymethylation and Cyanogen Bromide Cleavage. Carboxymethylation of β-galactosidase with [^{3}H]iodoacetic acid, of M15 protein with [^{14}C]iodoacetic acid, and cyanogen bromide cleavage were done as described (15, 16).

α-Complementation Assays. To measure the α-donor peptide CB2, assays were carried out with an excess of M15 extract (17). For the inverse situation, quantitation of M15 protein, the procedure was identical except that an excess (65 μg) of cyanogen bromide-treated β-galactosidase was added to each tube as α-donor. Both assays were linear over the concentration ranges used.

Peptide Purification and Sequencing. Procedures used in peptide purification and sequence determination have been described (5, 7–9).

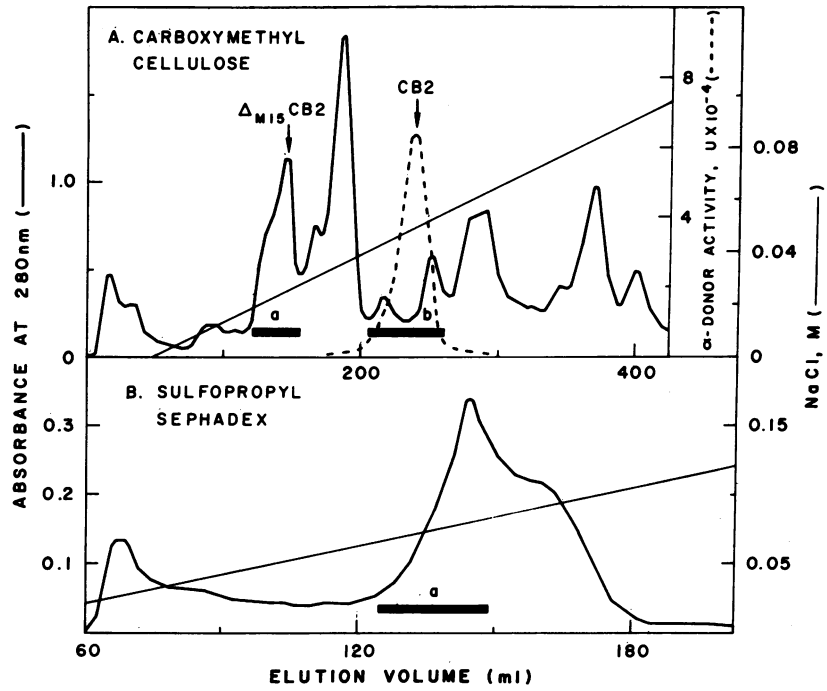

FIG. 1. Purification of Δ_{M15}CB2. (A) Carboxymethyl-cellulose chromatography. The cyanogen bromide digest of 73 mg of M15 [*carboxymethyl*-^{14}C]protein and 2.7 mg of [*carboxymethyl*-^{3}H]β-galactosidase was dissolved in 20 mM ammonium acetate, pH 4.7, 8 M urea, and applied to a 1.5 × 20 cm column equilibrated with the same buffer. The flow rate was 15 ml/hr, and the fraction size 5 ml. Elution was carried out with an NaCl gradient (*narrow solid line*) and monitored by absorbance at 280 nm (*bold solid line*), by α-donor activity (*dashed line*), and by ^{14}C/^{3}H ratio. The *arrows* indicate elution positions of CB2 and Δ_{M15}CB2. (B) Sulfopropyl-Sephadex chromatography. The pool containing Δ_{M15}CB2 from the Sephadex G-50 gel filtration step which followed carboxymethyl-cellulose (see *text*) was applied to a 1.5 × 15 cm column of sulfopropyl-Sephadex (A-25). The sample was dissolved in ammonium formate (10 mM with respect to ammonium ion), pH 2.5, 8 M urea, and the column equilibrated with the same buffer. Flow rate was 40 ml/hr, and a NaCl gradient was used for elution (*narrow solid line*). Fractions (3 ml) were analyzed for absorbance at 280 nm (*bold solid line*) and for ^{14}C/^{3}H.

RESULTS

Purification of M15 protein

By ammonium sulfate treatment and affinity chromatography, M15 protein was purified 35-fold with a yield of 62%. A single band was observed when the protein was analyzed by polyacrylamide gel electrophoresis in the presence of sodium dodecyl sulfate (18). The mobility relative to high-molecular-weight standards indicated a molecular weight of approximately 132,000, slightly less than the monomer molecular weight of 135,000 of β-galactosidase.

Molecular weight estimates by analytical ultracentrifugation were 134,000 in the presence of 6 M guanidine and 265,000 in 0.1 M phosphate, pH 7.2, 70 mM 2-mercaptoethanol. These values indicate that under the latter conditions, M15 protein is a dimer, in contrast to β-galactosidase, which is normally a tetramer (see ref. 19).

The specific activity of the freshly purified protein in the α-acceptor assay was close to values that we observe for pure β-galactosidase. This suggests that under the assay conditions used, β-galactosidase activity can be fully restored to M15 protein by addition of α-donor peptide.

Amino-terminal sequence of M15 protein

To determine the amino-terminal sequence of M15 protein, the method of Bruton and Hartley (20), which had been used successfully for β-galactosidase (9), was applied. The protein was maleylated and hydrolyzed with thermolysin; the digest was passed through a Dowex 50 column in 0.1 M acetic acid. Amino-acid analysis of the eluate showed high levels of threonine and methionine and was equivalent to results obtained with a parallel sample of β-galactosidase. This indicates that the M15 protein, like β-galactosidase, has an amino-terminal sequence of Thr-Met-, and that the region missing from M15 protein is internal.

Isolation of Δ_{M15}CB2

It was anticipated that the deletion in M15 protein would fall within the α-donor peptide CB2, residues 3–92 of β-galactosidase. This meant that cyanogen bromide digests of M15 protein and β-galactosidase would contain identical peptides except for CB2 and the corresponding peptide from M15 protein (Δ_{M15}CB2). CB2 had been isolated by use of its α-donor activity as an assay (5), but cyanogen bromide digests of M15 protein, of course, have no α-donor activity (4). Therefore, a double label technique was used to aid the isolation of Δ_{M15}CB2. We guessed that cysteine at residue 76 of β-galactosidase was not deleted from M15 protein. A small amount of [*carboxymethyl*-^{3}H]β-galactosidase (2.7 mg, 7.9 × 10^{5} cpm/mg) was mixed with 73 mg of M15 [*carboxymethyl*-^{14}C]protein (2.9 × 10^{4} cpm/mg), and a cyanogen bromide digest was prepared. Specific radioactivities of ^{14}C and ^{3}H were virtually identical in the mixture. Consequently, all cyanogen bromide peptides that contain cysteine were labeled equally with ^{14}C and ^{3}H except CB2, which carried only ^{3}H label (and was present in relatively small amount), and the altered peptide Δ_{M15}CB2, which carried only ^{14}C. Δ_{M15}CB2 was thereby detectable during purification by an increased ^{14}C/^{3}H ratio.

TABLE 1. *Amino-acid composition of peptides*

Amino acid	Δ_{M15}CB2	Tryptic peptides from Δ_{M15}CB2 ΔT1–5	T6	T7	T8
Tryptophan	2.75(4)[b]			0.34(1)[c]	1.06(3)[c]
Lysine				0.11	
Histidine					
Arginine	3.82(4)	1.21(1)	1.87(2)	0.82(1)	
Carboxymethylcysteine	0.92(1)				0.69(1)
Aspartic acid	5.72(6)	1.17(1)	1.07(1)	0.98(1)	2.87(3)
Threonine	2.74(3)	0.84(1)	0.98(1)	0.17	0.91(1)
Serine	4.17(5)	0.86(1)	0.87(1)	0.82(1)	1.79(2)
Glutamic acid	8.27(8)	0.35	2.14(2)	1.02(1)	5.21(5)
Proline	6.06(6)		0.97(1)		5.12(5)
Glycine	1.47(1)	0.33	0.26	1.20(1)	0.39
Alanine	5.66(6)	1.88(2)	0.08	0.16	3.85(4)
Valine	4.75(6)	1.57(2)			3.52(4)
Isoleucine	1.12(1)	0.71(1)			
Leucine	5.23(5)	1.06(1)	0.95(1)	1.18(1)	2.08(2)
Tyrosine	0.17				
Phenylalanine	2.13(2)				1.90(2)
Homoserine[a]	0.83(1)				0.85(1)
Amino-terminal residue	Ile	Ile	Thr	Ser	Phe
Total residues	59	10	9	7	33
Yield (%)[d]	45	36	25	72	46
Purification steps[e]		G25, PE	G25, PE	G25	G25, G50
Sequencing procedures[f]		DE, CPA + B	DE, CPA + B		

[a] Determined as homoserine plus homoserine lactone.

[b] Integral number based on analysis of tryptic peptides.

[c] Integral number based on A_{280} and identity to peptide from CB2.

[d] Yields are based on 540 nmol of cyanogen bromide digest, and 220 nmol of tryptic digest.

[e] Purification steps used: G25, Sephadex G-25 in 30% acetic acid; G50, Sephadex G-50 in 30% acetic acid; PE, paper electrophoresis at pH 1.9.

[f] Sequencing procedures used: DE, Edman technique with dansylation at each step; CPA + B, treatment with carboxypeptidases A and B.

Methods used for purification of Δ_{M15}CB2 were identical to those used previously for CB2 (5). The first step was carboxymethyl-cellulose ion exchange chromatography in the presence of 8 M urea (Fig. 1A). The α-donor activity of CB2 eluted in a region with $^{14}C/^{3}H$ ratio *lower* than unity (fraction b). This shows, as expected, that the α-donor peptide is unique to β-galactosidase. Fraction a, on the other hand, had a $^{14}C/^{3}H$ ratio *higher* than unity, and was applied to a Sephadex G-50 column equilibrated with 30% acetic acid. One peak from this column showed a further enhancement in the $^{14}C/^{3}H$ ratio. Two peptides were present in this fraction, and one contained the same amino-terminal residue as CB2 (isoleucine). Sulfopropyl-Sephadex (A25) chromatography in the presence of 8 M urea then gave the elution profile shown in Fig. 1B. Pooled fraction a was judged essentially pure by polyacrylamide gel electrophoresis in urea (21) and by dansyl amino-terminal analysis. The $^{14}C/^{3}H$ ratio was 5.0. The low value for ^{3}H cpm indicates that this peptide is unique to M15 and, therefore, must be the altered peptide.

The yield of Δ_{M15}CB2, estimated from radioactivity and amino-acid analyses, was 45%. Its amino-acid content (Table 1) is similar to CB2, and represents a peptide of 59 residues.

Sequence determination of Δ_{M15}CB2

The eight tryptic peptides indicated in Fig. 2 were obtained previously from CB2 (5). After tryptic digestion of Δ_{M15}CB2, however, dansyl amino-terminal analysis indicated the presence of four amino-terminal residues, Ile, Thr, Ser, and Phe. Tryptic peptides were obtained by procedures comparable to those used with CB2. In contrast to the case with CB2, only four tryptic peptides were present (Table 1). Three of these, with amino-terminal Thr, Ser, and Phe, were identical to peptides T6, T7, and T8 from CB2 with respect to elution position from Sephadex G-25, mobility on thin-layer chromatography, amino-acid composition, and, in the cases of T6 and T7, with respect to electrophoretic mobility on paper at pH 1.9 and amino-acid sequence. The fourth peptide was different from T1 of CB2 in the last few residues; its sequence was found to be Ile-Thr-Asp-Ser-Leu-Ala-Val-Val-Ala-Arg. Note that Ala-Arg replaces the sequence Leu-Gln-Arg at the carboxyl-terminal end of T1. Peptides T2, T3, T4, and T5 of CB2 were absent from the tryptic digest of Δ_{M15}CB2. As diagrammed in Fig. 2, these results demonstrate that residues 11–41 of the α-donor peptide CB2 are missing from the corresponding cyanogen bromide peptide of M15 protein.

DISCUSSION

The α-complementing acceptor protein of the *lac Z* operator-proximal deletion mutant M15 has been purified to homogeneity by ammonium sulfate fractionation and affinity chromatography. Sequence determination of a cyanogen bromide peptide unique to the M15 protein shows that 31 amino acids are missing, beginning with residue 11 of β-galactosidase. Since M15 protein completely lacks β-galactosidase activity, the absent segment must make an essential contribution to the formation of functional enzyme. This contribution may be a general conformational effect rather than a direct role in catalysis, since this portion of the polypeptide chain is not necessary for substrate binding (6). The observation that M15 protein is a dimer, rather than a tetramer, suggests that tetramer formation may be one of the requirements for the restoration of enzyme activity. With the availability in pure form of the two complementing partners, studies are now possible on the role each plays in the restoration of activity.

The primary structures of both the α-donor peptide CB2 from β-galactosidase and the amino-terminal region of the α-acceptor M15 protein are now known. As expected, the M15 deletion falls within CB2; this indicates that CB2 activates M15 protein by supplying the missing region. However, sequences 3–10 and 42–92 are both present in duplicate in the CB2-M15 complex. For each sequence, it seems likely that one of the duplicate sets is not an integral part of the functional complex, as appears to be the case for ω-complementation be-

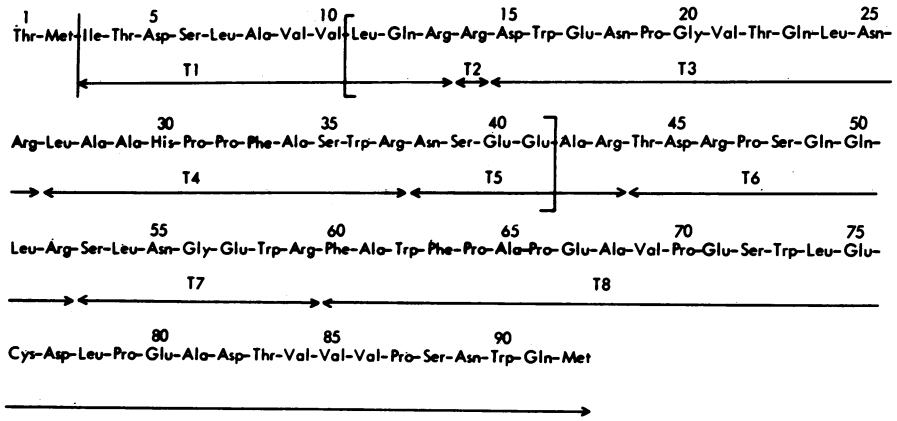

FIG. 2. Sequences of amino-terminal regions of β-galactosidase and M15 protein. CB2 is residues 3–92. Tryptic peptides from CB2 and Δ_{M15}CB2 are indicated. Residues 11–41 (in brackets) of β-galactosidase are missing in the M15 protein.

tween overlapping fragments (22). If so, M15 protein supplies the functionally involved residues, or CB2 supplies them, or M15 protein and CB2 each supply them partially. If CB2 rather than M15 protein supplies residues 3–10, then α-complementation would be similar in concept to complementing systems from ribonuclease (23) and Staphylococcal nuclease (24), in which the amino-terminal region of the functional complex is supplied by an added peptide.

Spontaneous partial deletions in bacterial genes that do not disturb the reading frame of the product polypeptides are not uncommon (25, 26). The partial amino-acid sequence determined for the M15 protein indicates that the M15 deletion is of this type. Clearly, M15 protein could not serve as an α-acceptor if the reading frame had been shifted. *A priori*, such deletions could begin within a codon, provided that the number of base pairs deleted is a multiple of three. However, examination of codon–amino-acid assignments indicates that the deletion in M15 begins at the start of the codon for residue 11. Since a number of highly polar terminator mutants do not recombine with M15 (10), the site of mutation in each such strain must lie at one of the codons corresponding to residues 11–41 of β-galactosidase. The positioning of the M15 deletion can now lead to the placement of these sites. These are examples of genetic fine structure analysis aided by methods of protein chemistry.

Finally, it may be recalled that the original studies of α-complementation raised the possibility that this was a case of inter- rather than intracistronic complementation (1). If so, this would mean that an α-subunit is present in wild-type β-galactosidase. However, it was soon shown that the properties of α-complemented enzyme were different from those of native enzyme. Under dissociating conditions that yield an α-subunit from complemented enzyme, no more than trace amounts could be obtained from the native protein (27). The demonstration that strain M15 produces a single polypeptide chain of molecular weight 132,000 is further support for the conclusion that the *Z* gene is a single cistron. Therefore, α-complementation is intracistronic.

We thank Douglas M. Brown for doing the ultracentrifuge analyses, and Val Schaeffer, Paulette Osborne, and Steve Barta for excellent technical assistance. This investigation was supported by USPHS Grant AI 04181 and by USPHS Molecular Biology Training Grant GM1531.

1. Ullmann, A., Jacob, F. & Monod, J. (1967) *J. Mol. Biol.* **24**, 339–343.
2. Fowler, A. V. & Zabin, I. (1966) *Science* **154**, 1027–1029.
3. Morrison, S. L. & Zipser, D. (1970) *J. Mol. Biol.* **50**, 359–371.
4. Lin, S., Villarejo, M. & Zabin, I. (1970) *Biochem. Biophys. Res. Commun.* **40**, 249–254.
5. Langley, K. E., Fowler, A. V. & Zabin, I. (1975) *J. Biol. Chem.*, in press.
6. Villarejo, M. R. & Zabin, I. (1973) *Nature New Biol.* **242**, 50–52.
7. Fowler, A. V. & Zabin, I. (1970) *J. Biol. Chem.* **245**, 5032–5041.
8. Fowler, A. V. (1972) *J. Biol. Chem.* **247**, 5425–5431.
9. Zabin, I. & Fowler, A. V. (1972) *J. Biol. Chem.* **247**, 5432–5435.
10. Beckwith, J. (1964) *J. Mol. Biol.* **8**, 427–430.
11. Adelberg, E. A. & Burns, S. N. (1960) *J. Bacteriol.* **79**, 321–330.
12. Zamenhof, P. J. & Villarejo, M. (1972) *J. Bacteriol.* **110**, 171–178.
13. Fowler, A. V. (1972) *J. Bacteriol.* **112**, 856–860.
14. Steers, E., Jr., Cuatrecasas, P. & Pollard, H. B. (1971) *J. Biol. Chem.* **246**, 196–200.
15. Crestfield, A. M., Moore, S. & Stein, W. H. (1963) *J. Biol. Chem.* **238**, 622–627.
16. Steers, E., Jr., Craven, G. R. & Anfinsen, C. B. (1965) *J. Biol. Chem.* **240**, 2478–2484.
17. Lin, S. & Zabin, I. (1972) *J. Biol. Chem.* **247**, 2205–2211.
18. Weber, K. & Osborn, M. (1969) *J. Biol. Chem.* **244**, 4406–4412.
19. Zabin, I. & Fowler, A. V. (1970) in *The Lactose Operon*, eds. Zipser, D. & Beckwith, J. R. (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.), pp. 27–47.
20. Bruton, C. J. & Hartley, B. S. (1970) *J. Mol. Biol.* **52**, 165–178.
21. Jovin, T., Chrambach, A. & Naughton, M. A. (1964) *Anal. Biochem.* **9**, 351–369.
22. Goldberg, M. E. (1969) *J. Mol. Biol.* **46**, 441–446.
23. Richards, F. M. & Vithayathil, P. J. (1959) *J. Biol. Chem.* **234**, 1459–1465.
24. Anfinsen, C. B. (1973) *Science* **181**, 223–230.
25. Jacob, F., Ullmann, A. & Monod, J. (1964) *C.R.H. Acad. Sci.* **258**, 3125–3128.
26. Malamy, M. H. (1966) *Cold Spring Harbor Symp. Quant. Biol.* **31**, 189–201.
27. Ullmann, A. & Perrin, D. (1970) in *The Lactose Operon*, eds. Zipser, D. & Beckwith, J. R. (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.), pp. 143–172.