Two Regions of the Bifunctional Protein Aspartokinase I-Homoserine Dehydrogenase I Are Connected by a Short Hinge*

(Received for publication, June 22, 1981)

Lise Sibiliò, Gisèle Le Bras, Gérard Le Bras, and Georges N. Cohen

From the Unité de Biochimie Cellulaire, Département de Biochimie et Génétique Moleculaire, Institut Pasteur, 75724 Paris, Cedex 15, France

Four proteases differing in their specificity, i.e. subtilisin, trypsin, α-chymotrypsin and V8 staphylococcal protease, cleave the bifunctional protein Escherichia coli aspartokinase I-homoserine dehydrogenase I (composed of 820 residues) producing an active homoserine dehydrogenase fragment. This cleavage occurs within a short segment of the polypeptide chain extending from residue 293 to residue 300.

Escherichia coli K12 aspartokinase (EC 2.7.2.4) I-homoserine dehydrogenase I (EC 1.1.1.13) is a bifunctional homotetramer involved in the biosynthesis of L-threonine, which regulates both its synthesis and activities (1). Each subunit is composed of 820 amino acid residues (2) and carries, in discrete regions, the aspartokinase I and homoserine dehydrogenase I activities (3, 4).

A tetrameric fragment (M. = 4 x 48,000) called the aspartokinase I fragment, synthesized by an ochre mutant carries the aspartokinase I activity only and has the same NH₂ terminus as the native enzyme (4). On the other hand, limited proteolysis of the native enzyme with subtilisin leads to a homodimeric fragment (homoserine dehydrogenase I fragment) of 523 residues/subunit (Ser297-Val820) carrying only the homoserine dehydrogenase I activity, while the aspartokinase I activity is lost (4, 5). The homogeneous subtilisin fragment has the same COOH-terminal sequence as aspartokinase I-homoserine dehydrogenase I (4).

Limited proteolysis by trypsin, α-chymotrypsin, and papain also leads to homoserine dehydrogenase fragments of approximately the same size, also endowed with homoserine dehydrogenase I activity (4). We have verified that this is also true of limited proteolysis by Staphylococcus aureus V8 protease.

The existence of the two active fragments aspartokinase I and homoserine dehydrogenase I, which have been shown to be able to fold independently (6), is taken as evidence of an ancestral fusion between two genes coding for two separate proteins (4).

Now that the total sequence of aspartokinase I-homoserine dehydrogenase I is known (2), we have verified that the different targets of limited proteolysis lie indeed in a very short segment of the bifunctional protein. It suggests that this segment is an exposed region of polypeptide chain, susceptible to proteolysis, which might be a "hinge" linking two regions of aspartokinase I-homoserine dehydrogenase I.

MATERIALS AND METHODS

Preparation of the Homoserine Dehydrogenase Fragments—The homoserine dehydrogenase fragments were prepared by proteolysis at 27 °C of the native aspartokinase I-homoserine dehydrogenase I either with α-chymotrypsin (1%, w/w), trypsin (5%, w/w), or S. aureus V8 protease (4%, w/w) essentially as previously described for subtilisin (4).

Proteolysis by α-chymotrypsin, as for subtilisin, was stopped by addition of phenylmethanesulfonyl fluoride (2 x 10⁻³ M) (4); trypsin by soybean trypsin inhibitor (5 times w/w excess). The action of V8 protease was arrested by cooling at +4 °C.

Sequence Determination—The NH₂-terminal amino acids of the different fragments were determined by the dansyl¹ technique (10).

The NH₂-terminal sequences were determined by automated Edman degradation in a Beckman 890 C sequencer (N,N-dimethylbenzamido single cleavage program); after denaturation by 6 M guanidine hydrochloride containing 0.2 M Tris-HCl buffer, pH 8.3, 1 mg/ml of EDTA, and dithiothreitol (2 molar excess over total sulfhydryl), the samples (3-5 mg) were extensively dialyzed at 4 °C against distilled water, lyophilized, and applied to the spinning cup in pure formic acid to allow complete solubilization. The phenylthiohydantoin amino acids were identified by thin layer chromatography (11) and often by high performance liquid chromatography (12).

RESULTS

Proteolysis by α-Chymotrypsin—The main dansyl NH₂-terminal was isoleucine. Lower quantities of alanine were found, but the only interpretable sequence was that starting with isoleucine 294 (Table 1).

Proteolysis by Trypsin—The main dansyl NH₂-terminal was Asx. Lower quantities of valine, glycine, and threonine were also found, but no sequences starting with these amino acids were interpretable from the results. Table II indicates the principal residues obtained at each cycle for the tryp tic fragment; traces of other amino acids were also present at each cycle and are probably due to secondary cleavage points. The only interpretable sequence starts at Asp 299.

Proteolysis by S. aureus V8 Protease—We have first verified that limited proteolysis of native aspartokinase I-homoserine dehydrogenase I by S. aureus V8 protease yields an active homoserine dehydrogenase fragment of the same molecular weight as obtained by the other proteases previously studied (result not shown). The main dansyl NH₂-terminals were Gly, Asx, Leu, and Ala. No sequences starting from Gly and Leu were interpretable from the results. The sequence starting with Asp 301 was readily interpretable. The sequence starting with Ala 296 (although weaker than that starting at Asp 301) is extremely clear; such an abnormal cleavage is probably due to a contam-

¹ The abbreviation used is: dansyl, 5-dimethylaminonapthalene-1-sulfonyl.
TABLE I

Automatic NH2-terminal sequence of the chymotryptic homoserine dehydrogenase I fragment

| Cycle | Residues | Identification method |
|-------|----------|-----------------------|
| 1     | Ile, Ala (w)* | TLC HPLC |
| 2     | Gly      | TLC HPLC |
| 3     | Ala      | TLC HPLC |
| 4     | Ser      | TLC HPLC |
| 5     | Arg      | TLC HPLC |
| 6     | Asp      | TLC HPLC |
| 7     | Glu      | TLC HPLC |
| 8     | Asp      | TLC HPLC |
| 9     | Glu      | TLC HPLC |
| 10    | Leu      | TLC HPLC |
| 11    | Pro      | TLC HPLC |
| 12    | Val      | TLC HPLC |
| 13    | Lys      | TLC HPLC |
| 14    | Gly (w)  | TLC HPLC |
| 15    | Ser (w)  | TLC HPLC |
| 16    | Asp (Asn, Gly, Ala, Thr, Val (w))* | TLC HPLC |
| 17    | Asn (w)  | TLC HPLC |
| 18    | Leu (w)  | TLC HPLC |
| 19    | Asn (w)  | TLC HPLC |

* Abbreviations are as in Table I.

TABLE II

Automatic NH2-terminal sequence of the tryptic homoserine dehydrogenase I fragment

| Cycle | Residues | Identification method |
|-------|----------|-----------------------|
| 1     | Asp, (Asn, Gly, Ala, Thr, Val (w))* | TLC HPLC |
| 2     | Glu, Ile (w) | TLC HPLC |
| 3     | Asp      | TLC HPLC |
| 4     | Glu      | TLC HPLC |
| 5     | Leu      | TLC HPLC |
| 6     | Val      | TLC HPLC |
| 7     | Ile      | TLC HPLC |
| 8     | Leu      | TLC HPLC |
| 9     | Pro      | TLC HPLC |
| 10    | Asn, Val | TLC HPLC |
| 11    | Leu, Asn | TLC HPLC |
| 12    | Asn, Glu | TLC HPLC |
| 13    | Asn, Ile | TLC HPLC |
| 14    | Met, Ser (w) | TLC HPLC |
| 15    | Ala, Asn | TLC HPLC |
| 16    | Met, Leu | TLC HPLC |

* Abbreviations are as in Table I.

The determination of the NH2-terminal sequences of the various homoserine dehydrogenase I fragments and the knowledge of the total sequence of aspartokinase I-homoserine dehydrogenase I enable us to determine the cleavage points by the different proteases used (Fig. 1). It is striking to note that they all lie in a very short segment of the native polypeptide chain, although there exist many other potential sites in the vicinity and although the various proteases vary greatly in their specificity. This result suggests that this short segment is especially sensitive to proteolysis and corresponds to the start of a globular region resistant to further proteolysis.

The cleavage sites observed are in agreement with the known specificities of the proteases used; however, using the staphylococcal protease, in addition to the canonical cleavage observed at position 300, a minor cleavage is observed between Gly 295 and Ala 296. A contaminant endowed with such specificity has already been described in preparations of this protease (7). Additional minor cleavage sites are probably present but will not be further discussed for the sake of clarity, since the components they yield are exceedingly minor. It should be recalled that during the limited proteolysis, aspartokinase I activity is destroyed concomitantly with the digestion of the segment comprising approximately the first 293 amino acids (2, 4). On the other hand, in the presence of threonine, the physiological allosteric modulator of the two activities carried by the bifunctional protein, proteolysis does not occur at a detectable rate (4). This protection is not due to an intrinsic stabilization of the NH2-terminal part of the molecule, but to the stabilization of the tetrameric state, the real substrate of proteolysis being the dimer (4). In addition, by changing the conditions of limited proteolysis, we have been able to detect the transient existence of a second large fragment corresponding to the NH2-terminal segment of aspartokinase I-homoserine dehydrogenase I, suggesting that the first nicks leading to the production of homoserine dehydrogenase I fragments actually take place in the region defined above. A very plausible conclusion is the existence of a flexible hinge segment between two compact globules. This hinge is exposed to the action of the proteases, at least in the absence of threonine, i.e. under conditions which stabilize the relaxed form of the bifunctional allosteric protein (8) and favor the dissociation of the tetramer into two dimers (4).

The site of the ancestral genetic fusion has been suggested to occur at a certain distance from the hinge region, in the vicinity of Met 249, on the basis of a Shine and Dalgarno (9) sequence before the corresponding ATG in the DNA sequence, followed by codons which could be converted into termination codons by a single base change (2). Limited
A Hinge in Aspartokinase I—Homoserine Dehydrogenase I

proteolysis is thus not sufficient to define precisely the site of gene fusion.

Acknowledgment—Our thanks go to André de Wolf who ran the sequenator.

REFERENCES

1. Cohen, G. N., and Dautry-Varsat, A. (1980) in Multifunctional Proteins (Bisswanger, H., and Schmincke-Ott, E., eds) pp. 49-121, J. Wiley and Sons, New York
2. Katinka, M., Cossart, P., Sibilli, L., Saint-Girons, I., Chalvignac, M. A. Le Bras, G., Cohen, G. N., and Yaniv, M. (1980) Proc. Natl. Acad. Sci. U. S. A. 77, 5730-5733
3. Falcoz-Kelly, F., Janin, J., Saari, J. C., Veron, M., Truffa-Bachi, P., and Cohen, G. N. (1972) Eur. J. Biochem. 28, 507-519
4. Veron, M., Falcoz-Kelly, F., and Cohen, G. N. (1972) Eur. J. Biochem. 28, 520-527
5. Bridey, P. A., Sibilli, L., Chalvignac, M. A., Cossart, P., Le Bras, G., de Wolf, A., and Cohen, G. N. (1978) J. Biol. Chem. 253, 8867-8871
6. Dautry-Varsat, A., and Garel, J. R. (1981) Biochemistry 20, 1395-1401
7. Sahel, S. A. (1976) Biochimie 58, 793-804
8. Janin, J., and Cohen, G. N. (1969) Eur. J. Biochem. 11, 520-529
9. Shine, J., and Dalgarno, L. (1975) Nature 254, 34-38
10. Zanetta, J. P., Vincendon, G., Mandel, P., and Gombos, G. (1970) J. Chromatogr 51, 441-458
11. Bridgen, J., Graffe, A. P., Karger, B. L., and Waterfield, M. D. (1975) in Instrumentation in Amino Acid Analysis pp. 111-145, (Perham, R. N., ed) Academic Press, New York
12. Frank, G., and Strubert, W. (1973) Chromatographia 6, 522-524
13. Yamada, S., and Itano, H. A. (1966) Biochim. Biophys. Acta 130, 538-540