The Primary Structure of Lactobacillus casei Thymidylate Synthetase

II. THE COMPLETE AMINO ACID SEQUENCE OF THE ACTIVE SITE PEPTIDE, CNBr 4*

(Rceived for publication, May 30, 1978, and in revised form, September 6, 1978)

Ronald L. Bellisario, Gladys F. Maley, Don U. Guarino, and Frank Maley
From the Division of Laboratories and Research, New York State Department of Health, Albany, New York 12201

The 102 amino acid residues of CNBr 4, the largest of 5 cyanogen bromide peptides from the Lactobacillus casei thymidylate synthetase, a protein composed of two subunits (2) each with a molecular weight of about 35,000 and probably identical (3, 4). Four of the CNBr peptides were completely sequenced, but the largest of the 5, CNBr 4, with a molecular weight of 11,800 and containing both cysteines present in a subunit (1, 5), posed some unique problems of analysis. These difficulties resulted from the fact that CNBr 4 contained a region of 50 amino acids which could not be cleaved by trypsin or BNPS-skatole, necessitating the preparation of the much smaller and more difficult to isolate staphylococcal protease and chymotryptic peptides. In addition, a region of 11 amino acids near the COOH terminus contained 4 prolines which required particular care in analysis as did the COOH terminus itself, which was finally established with an NBS peptide. Since the latter peptide (NBSGc) overlapped CNBr 4 and 5, it was useful in verifying the location of these peptides in the linear sequence of the enzyme. The isolation and analysis of NBSGc is described in more detail in the following paper (6).

EXPERIMENTAL PROCEDURES

The isolation of CNBr 4 is described in the previous paper as is the methodology for sequencing the peptides employed in this study.

* This investigation was supported in part by United States Public Health Service Grant GM-20371 from the National Institute of General Medical Sciences, United States Public Health Service, Department of Health, Education, and Welfare. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

1 The abbreviations used are: BNPS-skatole, 2-(2-nitrophenylsulfonyl)-3-methyl-3-bromoindolenine; NBS, N-bromosuccinimide; PTH, phenylthiohydantoin; SPITC, 3'-Sulfophenylisothiocyanate; NMAA, N,N'-dimethylaminocellulose; APG, amino-protected controlled pore glass; CPG, controlled pore glass; Quadrol, N,N',N'-tetraakis(2-hydroxypropyl)ethylenediamine; FeUOMP, 5'-fluoro 2'-deoxyuridine 5'-monophosphate; HPLC, high pressure liquid chromatography. For other abbreviations see "Nomenclature" in the miniprint to Ref. 1.

2 Portions of this paper including Figs. 1, 2, and 4 and Tables 1 to 27 are presented in miniprint following the References. Miniprint is easily read with the aid of a standard magnifying glass. Full size photocopies are available from the Journal of Biological Chemistry, 9650 Rockville Pike, Bethesda, Md. 20014. Request Document No. 78M882, cite authors, and include a remittance of $4.65 per set of photocopies.

Additional details on the isolation of peptides from CNBr 4 are described under "Results and Discussion," and also in the miniprint.

RESULTS AND DISCUSSION

CNBr 4 (194-296)

As shown previously (1, 5), this peptide contains both cysteine residues present in each of the two subunits of thymidylate synthetase. Under non-denaturing conditions, only one of the enzyme's four sulphydryl groups reacts with iodoacetate, but when the enzyme is denatured, all 4 cysteine residues are readily converted to their S-[1-14C]carboxymethyl derivatives in the presence of iodo[1-14C]acetate. CNBr 4 isolated from the enzyme so labeled yielded PTH derivatives for 16 continuous cycles of automatic Edman analysis and 4 additional residues to step 24 (Table 1). The sequence pattern and the presence of an S-carboxymethylcysteine residue at the fifth cycle is in agreement with our earlier analysis of the active site peptide (5), referred to as CN2. The remaining sequence of CNBr 4 was determined from a combination of peptides obtained by limited tryptic, chymotryptic, and staphylococcal protease digestions. Isolation of LT Peptides from CNBr 4—Tryptic hydrolysis of citraconylated CNBr 4 yielded 6 peptides, 3 soluble and 3 insoluble. The former peptides were separated by Bio-Gel P-10 chromatography (Fig. 1), with NH2-terminal LT13 (underlined fractions 61 to 70) being further purified by thin layer cellulose chromatography and LT12 by passage through a column of Bio-Gel P-2. The first underlined region, Fractions 50 to 59, represents a peptide of unhydrolyzed LT12 + NH2-terminal LT13. The insoluble peptides were dissolved in 30% acetic acid and were also purified on Bio-Gel P-10 (Fig. 2). The first peak is a large peptide which resulted from an incomplete hydrolysis at Arg-218. The amino acid composition of this peptide closely matches the sum of the amino acid residues (Table 2) from the other two peaks, COOH-terminal LT10, and LT11 (Fig. 2), and would thus appear to encompass residues 194 to 274. The amino acid compositions of the LT peptides from CNBr 4 are given in Table 2. The sequences of these peptides were determined as described below.

COOH-terminal LT10 (194-218)—Automatic sequence analysis of this peptide for 18 steps (Table 3) is basically identical with that obtained for intact CNBr 4 (Table 1), indicating that COOH-terminal LT10 is located at the NH2-terminal end of CNBr 4. Tryptic hydrolysis of COOH-terminal LT10 yielded two peptides (T1 and T2) which were separated on Bio-Gel P-10. T1 was determined to be NH2-terminal mainly from its amino acid composition, which reveals the presence of Cys(Cm) (Table 4). The COOH-terminal end of

3 Where limited tryptic (LT) peptides are used to establish the terminal regions of cyanogen bromide peptides, the NH2-terminal portion is considered synonymous with the COOH terminus of a CNBr peptide, while the COOH-terminal end encompasses the NH2-terminal end of the following CNBr peptide.
LT10, T2 (211-218), was sequenced by a combination of automatic Edman degradation and carboxypeptidase D treatment (Table 3) to complete the entire sequence of COOH-terminal LT10 (Fig. 3).

LT11 (219-274) was sequenced for 43 cycles by automatic Edman analysis. Although most of the residues were quantitated, those that were not could be clearly distinguished on thin layer chromatography (Table 6). The 2nd cysteine residue in CNBr 4 was detected as S-carboxymethylcysteine at cycle number 26 (residue 244) by thin layer chromatography and by radioactivity measurements.

Hydrolysis of LT11 with trypsin yielded four peptides whose amino acid compositions are given in Table 7. T1 and T2 were released by chymotryptic-like cleavage at Phe-228 and Tyr-233, respectively, and as a consequence complicated the peptide separation. However, thin layer chromatography was effective in separating the two peptides. T1 (219-228) was sequenced for seven continuous cycles (Table 8), while T2 (223-233) agreed exactly in amino acid composition with C7, which was previously sequenced (Table 9), and was therefore not studied further (for C peptides see below). Automatic sequence analysis of T3 (234-269) was conducted for 25 continuous cycles (Table 10). However, T3 was contaminated with a peptide (T3a, 229-269) which appears to have resulted from a chymotryptic-like cleavage of Phe-228, and as a consequence yielded a background during the sequencing of T3 (Table 10). This can be confirmed by comparing residues 239-247 of the major peptide with cycles 11 to 19 of the minor peptide. The sequence of T4 (270-274) was determined by automatic Edman analysis (Table 11).

Since the sequence information provided from the above peptides was insufficient to complete the sequence of LT11, the residual sequence was obtained from peptides isolated after treatment of LT11 with staphylococcal protease and from C peptides obtained on digestion of CNBr 4 with chymotrypsin as discussed below.

Sequence Analysis of Staphylococcal Protease Peptides from 13C-SCM LT11—The seven peptides obtained on treatment of LT11 with staphylococcal protease (8) were fractionated on Bio-Gel P-4 (Fig. 4). The individual peptides were further purified by thin layer cellulose chromatography and high voltage electrophoresis at pH 1.9. The amino acid analyses of the SP peptides are given in Table 12. SP1, SP5, and SP7 were isolated in low yields and were not studied further. 13C-SP2, SP3, and SP4 were sequenced by automatic Edman degradation (Tables 13 to 15), while SP6 (271-274) was not sequenced since it clearly matched a portion of T4 (270-274) from LT11 (Table 11).

Isolation and Sequence Analysis of Chymotryptic Peptides from CNBr 4—CNBr 4 (194-296) was hydrolyzed with chymotrypsin (100:1 molar ratio) for 17 h at 37°C as previously described (4) to yield 15 peptides which were separated by a combination of Bio Gel P 4 chromatography, thin layer cellulose chromatography, and high voltage paper electrophoresis at pH 1.9. The amino acid compositions of these peptides are presented in Table 16 and the sequences of 8 peptides are shown in Tables 17 to 25. C1 (194-292) appears from its amino acid composition to contain the FdUMP-binding site, the sequence of which was provided earlier (5) and also in Table 1. The amino acid composition of C5 (Table 16) strongly suggests that this peptide overlaps COOH-terminal LT10 and LT11 from residues 217 to 223. The sequence of C5 would thus be Gin-Ser-Ala-Arg-Ser-Ala-Asp-Ile-Phe. This peptide did not stain with ninhydrin or fluorescine, indicating that the NH2-terminal glutamine residue had cyclized to pyroglutamate. The C peptides were used primarily to confirm the order of placement of the trypptic, limited trypptic, and SP peptides in CNBr 4.

C3 (206-211)—SPIPC treatment was used with 3-aminopropyl-controlled pore glass to enhance the yields of this hexapeptide and although the recovery of lysine was barely detectable because of this treatment (9), the COOH-terminal amino acid could be recovered from the sequenator cup as leucine (Table 17).

C4 (212-216)—This pentapeptide was completely sequenced using just 3-aminopropyl-controlled pore glass beads and the DMAA program. The COOH-terminal residue was recovered from the sequenator cup as tyrosine (Table 18).

C6 (224-228)—This small hydrophobic peptide was sequenced using controlled pore glass beads and the DMAA program, and although proline yields were not high enough to be detected, it was placed at residue 227 based on the amino acid composition of C6 (Table 16). Despite the fact that residue 227 could not be clearly demonstrated, Phe-228 could be recovered from the sequenator cup on the following cycle (Table 19).

C9 (237-251)—Through the use of APG beads and the 0.1 M Quadrol program, C9 was completely sequenced. Because the first 2 residues are normally difficult to detect, manual dansyl Edman degradation was employed to confirm His-238 as the 2nd residue. His-242 was placed by difference from the amino acid composition of C9 (Table 16) and its detection at this residue in LT11 (Table 6). In fact, this peptide is encompassed in LT11 (219-262) and was used primarily to confirm
the presence of the enzyme's second cysteine at residue 244 (Table 20).

C10 (252–255)—This peptide was sequenced by automatic Edman degradation using APG beads and the DMAA program (Table 21), in addition to carboxypeptidase A treatment. Its location within LT11 (Table 6) was thus confirmed.

C11 (256–261)—C11 was sequenced in the same manner as C10, with the COOH-terminal amino acids being established by carboxypeptidase A treatment (Table 22). These residues, which were not convincingly determined in LT11 (Table 6), are now clearly delineated.

C12 (262–272)—This section of CNBr 4 was difficult to obtain as a tryptic, limited tryptic, or even as an overlap peptide, but could be isolated on chymotrypsin digestion of CNBr 4. Sequencing with APG beads and the DMAA program provided excellent yields through residue 269. The last 2 residues (271 and 272) were identified with carboxypeptidase A (Table 23). The information obtained also provided a missing sequence in LT11.

C13 (273–278)—Since C13 overlaps T4 from LT11 (270–274) (Table 11) and NH2-terminal LT13 (275–288) (Table 2), it aided greatly in locating this difficult to analyze sequence of amino acids, which results from the fact that of the 9 residues, 6 are serine, threonine, or proline. The tripeptide, 1-trp-Pro-Arg, which is partially resistant to trypsin (Table 16), is also present in this sequence followed by a Pro-Ala-Pro (Table 24). The close proximity of these 2 prolines to Arg-277 is probably responsible for the observed resistance. With the information obtained from the C peptides, the complete sequence of LT11 was established (Fig. 5).

C14 (292–299)—Residues Phe-292 and Ile-291 were removed with carboxypeptidase A and the residual peptide was placed in the sequenator cup with APG beads and analyzed with the DMAA program. Asp-290 was recovered from the cup and identified on the amino acid analyzer (Table 25). The sequence of C14 agrees exactly with that of NH2-terminal LT13 (278–290) (Table 26) and extends the latter by 2 residues.

C15 (293–296)—As shown in Table 16 this peptide contains 2 aspartates, 1 phenylalanine and 1 homoserine, indicating that it is derived from the COOH terminus of CNBr 4. Although sufficient quantities were not available for analysis, the sequence of C15 was established as Asp-Phe-Asp-Met with both NBS and LT peptides described in the following paper (6).

LT12 (275–277)—Manual subtractive Edman degradation of LT12 (data not presented) revealed an NH2-terminal threonine and established the sequence Thr-Pro-Arg (Table 2) which as indicated below is only partially released by trypsin from the peptide containing residues 270 to 288. This sequence was confirmed by analysis of peptide (LT12 + NH2-terminal LT13) (Table 27) and by its presence in C15 (Table 24).

NH2-terminal LT13 (278–296)—As indicated above, this peptide resulted from the partial cleavage of the Arg-Pro bond at positions 277–278 and differs from the unhydrolyzed purified peptide (275–296) by a Thr-Pro-Arg (LT12). This is shown in both the agreement of the amino acid compositions (Table 2) and sequence analyses of the two peptides (Tables 26 and 27). NH2-terminal LT13 if analyzed directly from the Bio-Gel P-10 column (Fig. 1) without further purification was contaminated with peptide residues 275 to 296, as indicated by the sequence obtained through 14 continuous cycles (Table 27). Since this peptide could only be sequenced through residue 290 (Table 26), the remaining sequence of NH2-terminal LT13 (the COOH terminus of CNBr 4) had to be obtained from an NBS overlap peptide (see Figs. 8 and 9 of Ref. 6). The COOH terminus of CNBr 4 is shown in Fig. 6.

The complete sequence of CNBr 4 is summarized in Fig. 7.

REFERENCES
Thymidylate Synthetase. II

<table>
<thead>
<tr>
<th>FRACTION NUMBER</th>
<th>(-----) RELATIVE RADIOACTIVITY</th>
<th>(-----) RELATIVE FLUORESCENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

<table>
<thead>
<tr>
<th>Table 1: Fraction Activity and Fluorescence</th>
<th>Table 2: Fraction Activity and Fluorescence</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Columns 1-5)</td>
<td>(Columns 1-5)</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

**Table 1:**

<table>
<thead>
<tr>
<th>Fraction Number</th>
<th>Activity</th>
<th>Fluorescence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

**Table 2:**

<table>
<thead>
<tr>
<th>Fraction Number</th>
<th>Activity</th>
<th>Fluorescence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

**Legend:**

1. Activity: Measured in units of 
2. Fluorescence: Measured in arbitrary units

---

*Note: The tables and diagrams are placeholders for the actual data that would be present in the full document.*
The primary structure of Lactobacillus casei thymidylate synthetase. II. The complete amino acid sequence of the active site peptide, CNBr 4.
R L Bellisario, G F Maley, D U Guarino and F Maley


Access the most updated version of this article at http://www.jbc.org/content/254/4/1296

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 0 references, 0 of which can be accessed free at http://www.jbc.org/content/254/4/1296.full.html#ref-list-1