Biochem. J. (1980) **191**, 867–868
Printed in Great Britain

Sequence of residues 400–403 of bovine serum albumin

Roberta G. REED,*‡ Frank W. PUTNAM† and Theodore PETERS, Jr.*
**The Mary Imogene Bassett Hospital, Cooperstown, NY, U.S.S., and †Department of Biology, Indiana University, Bloomington, IN, U.S.A.*

(*Received 2 September 1980/Accepted 12 September 1980*)

A large tryptic peptide of bovine serum albumin (residues 377–582) was subjected to 32 cycles of Edman degradation to determine the sequence of the last remaining unknown segment of this protein. Residues 400–403 were identified as Gly-Phe-Gln-Asn. Amide assignments were also made at positions 388 (glutamine), 389 (asparagine), 391 (aspartic acid) and 392 (glutamine).

Since the first published version of the sequence of bovine serum albumin appeared in 1975, a gap of three (Brown, 1975) or four (Brown, 1977) residues extending from position 400 has existed. By analogy to human serum albumin these residues were predicted to be Lys-Phe-Gln-Asn (Meloun *et al.*, 1975; Behrens *et al.*, 1975), but their identity was not confirmed.

The widespread use of bovine albumin as a protein and molecular-weight standard, as well as the extensive work on physical and chemical properties of this important protein, make it especially desirable to have the complete sequence known.

In order to complete the sequence of bovine serum albumin we undertook the sequential analysis of the first 32 residues of a large tryptic peptide extending from residue 377 to 582 (the *C*-terminus).

Experimental

The peptide corresponding to residues 377–582 was isolated as previously described (Peters & Feldhoff, 1975) from a limited tryptic digest of crystalline bovine serum albumin bound to palmitoyl–aminoethylamino-agarose. The peptide was identified as extending from residue 377 to the *C*-terminus of Brown's (1975) sequence on the basis of *N*-terminal sequence (His-(Leu/Val)-Asp), *C*-terminal sequence (Thr-Gln-Thr-Ala-Leu-Ala), size, composition, and results of further digestion with pepsin and cleavage with CNBr (Peters & Feldhoff, 1975). The electrophoretic pattern of this fragment on SDS/polyacrylamide gel, with and without β-mercaptoethanol, was a single band corresponding to a molecular weight of 23000, with no evidence of impurities or 'nicks' in the peptide chain.

Sequence analysis of a 116 nmol sample was performed on a Beckman sequencer [model 890B (updated)] by using a 0.1 M-Quadrol program (no. 011576) supplied by Beckman. The amino acid phenylthiohydantoin derivatives were identified by h.p.l.c. on a Hewlett–Packard model-1084A liquid chromatograph equipped with an octadecylsilane column; elution was with a gradient of 0.2 M-sodium acetate buffer, pH 4.8, and acetonitrile.

Abbreviation used: SDS, sodium dodecyl sulphate; h.p.l.c., high-pressure liquid chromatography.

‡ To whom reprint requests should be addressed.

Results and discussion

Results of the sequence determination are given in Table 1. The findings for 32 cycles confirmed the published sequence, and indicated that the missing residues 400–403 are Gly-Phe-Gln-Asn. The amino acid composition of residues 377–582 reported by Peters & Feldhoff (1975) is in good agreement with this assignment of the residues 400–403, including the presence of glycine rather than lysine at residue-400, as in human serum albumin.

Four amide residues have also been assigned by the sequence analysis of this fragment. They are: 388 (glutamine); 389 (asparagine); 391 (aspartic acid) and 392 (glutamine). This assignment is consistent with the expectation of three amides and one acid based on the amide NH_3 from amino acid analysis of residues 377–582 (Peters & Feldhoff, 1975).

The amino acid composition of bovine albumin according to the sequence of Brown (1977) and including the residues reported herein becomes: Lys_{59}, His_{17}, Arg_{23}, Asp_{39}, Asn_{12}, Asx_3, Thr_{34}, Ser_{28},

0306-3275/80/120867-02$01.50/1

Table 1. *Identification of the first 32 residues of tryptic peptide 377–582 of bovine serum albumin*

Cycle	Position	Expected residue*	Experimental residue: Identity	Experimental residue: Amount (nmol)†
1	377	His	His	89
2	378	Leu	Leu	57
3	379	Val	Val	77
4	380	Asp	Asp	112
5	381	Glu	Glu	85
6	382	Pro	Pro	28
7	383	Gln	Gln	46
8	384	Asn	Asn	30
9	385	Leu	Leu	51
10	386	Ile	Ile	41
11	387	Lys	Lys	61
12	388	Glx	Gln	21
13	389	Asx	Asn	14
14	390	Cys	‡	—
15	391	Asx	Asp	31
16	392	Glx	Gln	26
17	393	Phe	Phe	30
18	394	Glu	Glu	32
19	395	Lys	Lys	36
20	396	Leu	Leu	16
21	397	Gly	Gly	9.2
22	398	Glu	Glu	21
23	399	Tyr	Tyr	19
24	400	Xxx‖	Gly	10
25	401	Xxx	Phe	26
26	402	Xxx	Gln	15
27	403	Xxx	Asn	8.6
28	404	Ala	Ala	18
29	405	Leu	Leu	17
30	406	Ile	Ile	12
31	407	Val	Val	19
32	408	Arg	Arg	§

* Based on the published sequence (Brown, 1977).

† 116 nmol sequenced.

‡ No amino acid was detected at this position. Since disulphide bonds were not reduced and blocked before sequencing, a cysteine residue would not be isolated at this position, but would remain attached to cysteine-436 through a disulphide bond.

§ Qualitative identification only.

‖ Xxx, unknown amino acid.

Glu_{59}, Gln_{19}, Glx_1, Pro_{28}, Gly_{16}, Ala_{46}, Cys_{35}, Val_{36}, Met_4, Ile_{14}, Leu_{61}, Tyr_{19}, Phe_{27}, Trp_2.

The calculated molecular weight is 66267, assuming amide groups on two of the four aspartic acid and glutamic acid residues that remain unassigned.

This work was supported by U.S. National Institutes of Health grants nos. HL 02751 and AM 19221 and the Stephen C. Clark Research Fund of The Mary Imogene Bassett Hospital. We are indebted to Mrs. S. Dorwin, J. Madison and P. Davidson for technical assistance.

References

Behrens, P. Q., Spiekerman, A. M. & Brown, J. R. (1975) *Fed. Proc. Fed. Am. Soc. Exp. Biol.* **34**, 591

Brown, J. R. (1975) *Fed. Proc. Fed. Am. Soc. Exp. Biol.* **34**, 591

Brown, J. R. (1977) *FEBS Meet. 11th* **50**, 1–20

Meloun, B., Moravek, L. & Kostka, V. (1975) *FEBS Lett.* **58**, 134–137

Peters, T., Jr. & Feldhoff, R. C. (1975) *Biochemistry* **14**, 3384–3391