

# Complete Amino Acid Sequence of Bovine $\beta_2$ -Microglobulin\*

Merton L. Groves and Rae Greenberg

From the Eastern Regional Research Center, Agricultural Research Service, United States Department of Agriculture, Philadelphia, Pennsylvania 19118

$\beta_2$ -Microglobulin has been isolated and crystallized from bovine colostrum and represents the only crystalline form of this protein reported. The complete sequence of bovine  $\beta_2$ -microglobulin was determined using only one proteolytic enzyme, *Staphylococcus aureus* V8 protease. Automated degradation of the intact molecule and two large peptides produced by enzymic digestion provided unequivocal placement of all residues. Bovine  $\beta_2$ -microglobulin, molecular weight 11,630, contains 98 residues as compared with 99 for the human, rabbit, guinea pig, and murine proteins. The valine residue at position 49 in all the aforementioned species is deleted in the bovine variant. This crystalline protein is also uniquely characterized by three di-prolyl sequences in the first third of the molecule.  $\beta_2$ -Microglobulin is a highly conserved protein, and in a comparison of absolute amino acid sequence differences among species, bovine  $\beta_2$ -microglobulin ranges from a low of 24 substitutions as compared with the rabbit homologue and 26 for human to 32 for mouse and guinea pig.

$\beta_2$ -Microglobulin is a low molecular weight protein found both free in various body fluids and bound to cell surface proteins.  $\beta_2$ -Microglobulin has been shown to be structurally related to the constant domains of the immunoglobulins and to be noncovalently associated with the major and minor histocompatibility antigens (for a review, see Peterson *et al.* (1)). There is also an accumulation of data which indicates that elevated concentrations of  $\beta_2$ -m<sup>1</sup> in body fluids correlate with certain diseases (2-4). The precise role of  $\beta_2$ -m in the immune system is unclear, although some evidence suggests that it may stabilize the tertiary structure of histocompatibility antigens (5) or that the  $\beta_2$ -m-histocompatibility antigen complex is necessary for subsequent processing and intracellular transport of the antigen after synthesis (6). Another theory assigns to  $\beta_2$ -m the role of interacting with a complement-like killing structure on the T-cell to initiate cell destruction in a manner analogous to the interaction between the Fc portion of IgG and complement (7).

In 1963, we described the isolation and characterization of a new crystalline protein, lactollin, from bovine milk (8). Subsequently, lactollin was shown to be  $\beta_2$ -microglobulin, the only crystalline homologue reported (9). This molecule dem-

onstrates some unusual physical chemical properties. At very low concentrations, it is present as a monomer with a molecular weight of 11,800. At higher concentrations, it undergoes a concentration-dependent monomer to tetramer reversible self-association (10). Becker *et al.* (11) have presented preliminary results on the crystallographic properties of the protein. Recently, x-ray studies on the structure of papain-solubilized histocompatibility antigen were also initiated (12). This will make it possible to compare the three-dimensional structure of  $\beta_2$ -m and soluble histocompatibility antigen and envision better the structure of the complex. Elucidation of the amino acid sequences of these two proteins may supply insight into the nature of their binding sites and perhaps contribute to a final functional assignment.

The present report describes an improved method for the isolation of bovine  $\beta_2$ -m from colostrum and the determination of the complete amino acid sequence of this protein. The primary structures of  $\beta_2$ -microglobulins of human (13), rabbit (14), guinea pig (15), and mouse (16) have been described. A comparison of the bovine sequence with the other species confirms the high degree of structural homology among  $\beta_2$ -microglobulins.

## EXPERIMENTAL PROCEDURES<sup>2</sup>

### RESULTS

Colostrum casein is a good source of bovine  $\beta_2$ -microglobulin. Its isolation was accomplished in four steps: chromatography on DEAE-cellulose, chromatography on CM-cellulose, gel-filtration, and crystallization. The crystalline material shows a single band on polyacrylamide gel electrophoresis, pH 4.3, 8 M urea, and by sodium dodecyl sulfate polyacrylamide gel electrophoresis. About 100 mg of  $\beta_2$ -microglobulin can be expected from 100 g of colostrum casein.

Sequencing of bovine  $\beta_2$ -microglobulin was greatly simplified when two relatively large overlapping peptides (V8-(4)-2 and V8-(8)-2, Table I) were isolated as described under "Experimental Procedures" by limited cleavage with *Staphylococcus aureus* V8 protease. At pH 8 and pH 4 in ammonium bicarbonate and ammonium acetate buffers, respectively, cleavage by this enzyme has been reported to be specific for glutamyl bonds (22). This limited selective hydrolysis resulted when CM- $\beta_2$ -m was digested in ammonium acetate at pH 4. A chromatographic pattern of the digest on Sephadex G-50 (Fig. 5S) revealed two peaks, the first corresponded to the

\* 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.

<sup>1</sup> The abbreviations used are:  $\beta_2$ -m,  $\beta_2$ -microglobulin; CM, S-carboxymethyl; V8 peptides, peptides from *S. aureus* V8 protease digestion at pH 4 or pH 8; HVE<sub>1</sub> high voltage paper electrophoresis. TPCCK, L-1-tosylamido-2-phenylethyl chloromethyl ketone; GLC, gas-liquid chromatography; PTH, phenylthiohydantoin.

<sup>2</sup> Portions of this paper (including "Experimental Procedures," Figs. 1S-7S, Tables IS-IIIIS, and Footnote 3) are presented in miniprint at the end of this paper. 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 20814. Request Document No. 81M-2121, cite authors, and include a check or money order for \$6.40 per set of photocopies. Full size photocopies are also included in the microfilm edition of the Journal that is available from Waverly Press.

## Sequence of Bovine $\beta_2$ -Microglobulin

TABLE I  
Amino acid composition of *S. aureus* V8 protease peptides

	V8-(4)-1 <sup>a</sup> V8-(8)-1	V8-(4)-2	$\beta_2$ -m <sup>(9)</sup>	V8-(8)-2	V8-(8)-3	V8-(8)-4
S-Carboxymethyl-cysteine	0.7 (1)	.6 (1)	2	0.8 (1)		
Aspartic acid	3.0 (3)	7.7 (8)	11	3.9 (4)		2.3 (2)
Threonine	0.2	1.9 (2)	2	1.8 (2)		
Serine	0.9 (1)	6.7 (7)	8	1.5 (2)	1.0 (1)	
Glutamic acid	5.1 (5)	7.2 (7)	12	3.1 (3)	1.2 (1)	1.3 (1)
Proline	7.0 (7)	1.9 (2)	9	2.0 (2)		1.0 (1)
Glycine	1.8 (2)	1.3 (1)	3	0.3		
Alanine	0.1	1.0 (1)	1	0.2		
Valine	1.9 (2)	2.7 (3)	5	2.6 (3)		0.9 (1)
Isoleucine	2.8 (3)	2.6 (3)	6	0.9 (1)	1.0 (1)	0.8 (1)
Leucine	1.3 (1)	7.0 (7)	8	2.2 (2)		1.0 (1)
Tyrosine	3.6 (4)	2.0 (2)	6	1.0 (1)		
Phenylalanine	1.0 (1)	3.0 (3)	4	1.0 (1)		
Lysine	2.4 (2)	6.7 (7)	9	3.3 (3)	1.7 (2)	1.1 (1)
Histidine	1.7 (2)	2.0 (2)	4	0.9 (1)		
Arginine	2.0 (2)	2.9 (3)	5	2.9 (3)		2.0 (2)
Tryptophan		+ <sup>b</sup>	2	+ <sup>b</sup>		+ <sup>b</sup>
Residues	1-36	37-98	1-98	69-98	45-49	88-98

<sup>a</sup> The integer values are the same for both peptides although the actual data is for V8-(4)-1.

<sup>b</sup> + = Qualitative determination.

larger peptide, V8-(4)-2, and the second represented a smaller peptide, V8-(4)-1. Polyacrylamide gel electrophoretic patterns of these peptides showed that the two chromatographic peaks each represent a major peptide band (Fig. 1A, gels 2 and 3). S-Carboxymethyl  $\beta_2$ -m is also shown for comparison in Fig. 1A, gel 1. The sum of the amino acid compositions of V8-(4)-1 and V8-(4)-2 (Table I) accounted for the published composition of  $\beta_2$ -microglobulin (9). The preliminary NH<sub>2</sub>-terminal sequence of  $\beta_2$ -m previously demonstrated that 2 glycyl residues were present in the first 32 amino acids while threonine and alanine were absent (9). Since peptide V8-(4)-1 contains two glycines, it was obvious from the amino acid composition and NH<sub>2</sub>-terminal sequence that this peptide comprised residues 1 to 36. Therefore, the larger peptide V8-(4)-2 which contains threonine and alanine must represent the remainder of the molecule, residues 37 to 98.

When CM- $\beta_2$ -m was digested by *S. aureus* V8 protease and at pH 8 for a longer period of time, the elution pattern of the digest on Sephadex G-50 was more complex (Fig. 6S). Polyacrylamide gel electrophoretic bands of pooled fractions (a) indicated that it contained among other peptides, a strong band corresponding in mobility to peptide V8-(4)-1 and another band of equal intensity with slower mobility, suggesting a more basic peptide. Gel electrophoresis of pooled fraction (b) (Fig. 6S) showed a major band corresponding to the basic peptide while the band corresponding in mobility to peptide V8(4)-1 was significantly reduced. This suggested that the basic peptide was somewhat smaller than the peptide similar in mobility to V8-(4)-1. The two major peptides in pool a, Fig. 6S, were resolved on DEAE-Sephadex A-25 into two peaks (Fig. 7S) in which the basic peptide V8-(8)-2 eluted first followed by peptide V8-(8)-1. The amino acid composition of peptide V8-(8)-1, together with its relative mobility, showed that it was identical with peptide V8-(4)-1, residues 1 to 36. The composition of the basic peptide V8-(8)-2, Table I, indicated an absence of glycine and alanine and the presence of 2 threonyl residues. From the known sequence data of other  $\beta_2$ -microglobulin species and with calculations resulting in two threonyl groups and the presence of an S-carboxymethyl-cysteine, peptide V8-(8)-2 was placed at the carboxyl-terminal portion of  $\beta_2$ -m, residues 69 to 98. Fig. 1B, gel 1, shows the polyacrylamide gel electrophoretic pattern of basic peptide, V8-(8)-2 and Table I lists its composition. Results of the automated sequencing of CM- $\beta_2$ -m, V8-(4)-2, and V8-(8)-2 are given in Tables IS to IIIS. Two other peptides from the S.

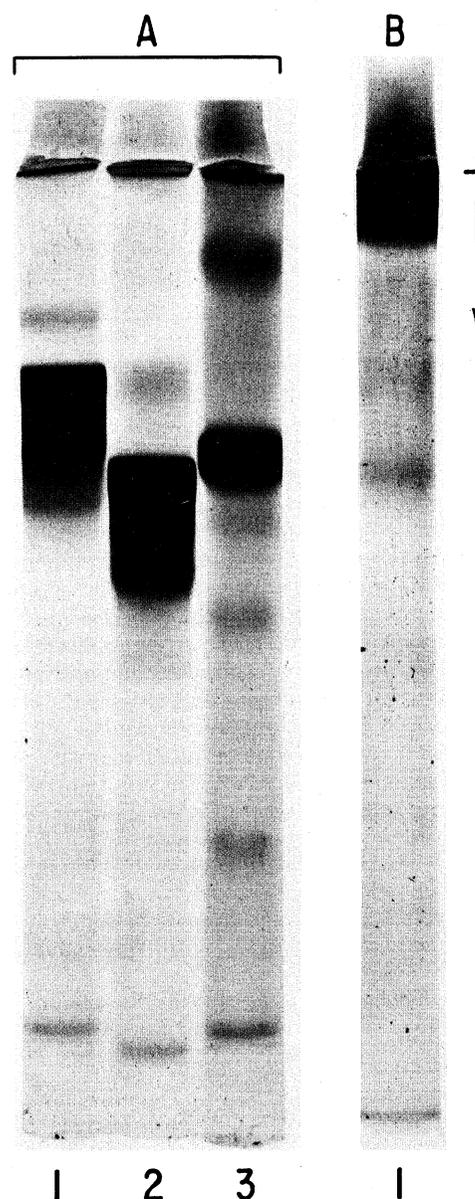


FIG. 1. Polyacrylamide gel electrophoretic patterns, pH 8.9, 4 M urea. A, gel 1, CM- $\beta_2$ -m; gel 2, peptide V8-(4)-2; gel 3, peptide V8-(4)-1. B, gel 1, V8-(8)-2.

Sequence of Bovine  $\beta_2$ -Microglobulin

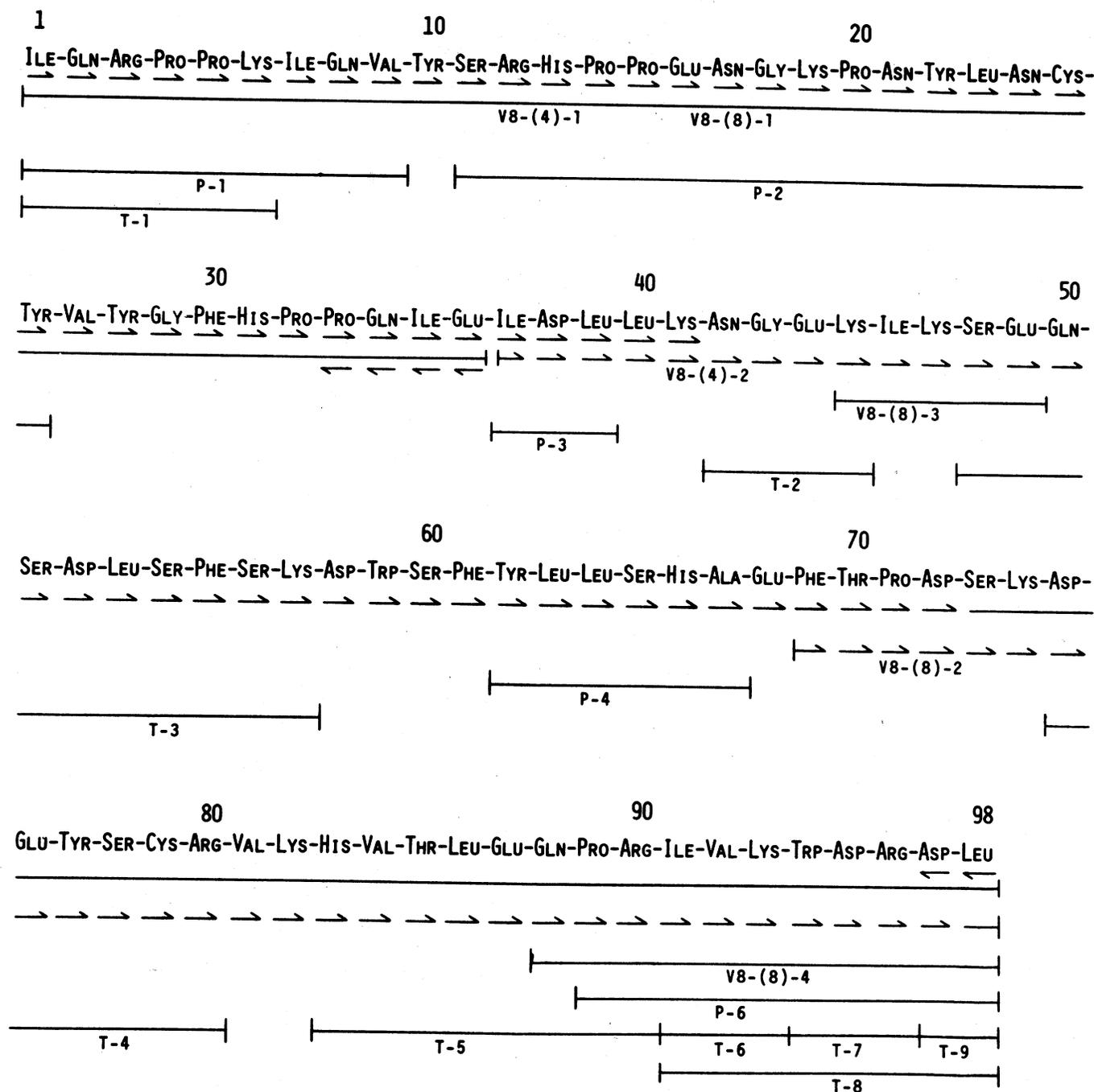


FIG. 2. Strategy for determination of the amino acid sequence of bovine  $\beta_2$ -m. The staphylococcal V8 protease (V8), trypsin (T), and pepsin (P) peptides are indicated. The forward arrows indicate those peptides sequenced; the reverse arrows indicate the use of carboxypeptidases. Solid lines indicate residues placed by peptide composition only.

*aureus* protease, pH 8, digest of  $\beta_2$ -m were purified by HVE, V8-(8)-3, and V8-(8)-4.

Peptides obtained by peptic and tryptic digestion of CM- $\beta_2$ -m were not necessary for determination of the primary structure but a brief description of their isolation was included since the compositions of these purified peptides are added proof of the accuracy of the sequence data. The composition of the peptides isolated after digestion of CM- $\beta_2$ -m with pepsin and trypsin is shown in Table II.

DISCUSSION

The complete amino acid sequence of bovine  $\beta_2$ -microglobulin and the strategy employed are shown in Fig. 2. This protein with 98 amino acid residues has a molecular weight of 11,630. To elucidate the overlaps, two relatively large peptides

(V8-(4)-2 and V8-(8)-2) obtained by limited cleavage of  $\beta_2$ -m with *S. aureus* V8 protease were subjected to automatic sequence analysis. These data together with the NH<sub>2</sub>-terminal 41 residue sequence of intact  $\beta_2$ -m provided enough information to establish clearly the complete sequence. One peptide (residues 69-98, Fig. 2) was sequenced to the penultimate amino acid, leaving only the terminal leucine to be assigned by carboxypeptidase. The amino acid totals determined by the sequence agreed with the published amino acid composition except that there are 9 serines instead of the 8 previously reported (9).

Limited digestion of  $\beta_2$ -m by *S. aureus* protease occurred at pH 4 in ammonium acetate and only the glutamyl-isoleucyl bond (residues 36-37) was efficiently cleaved. At pH 8 in ammonium bicarbonate with increased enzyme/protein ratio,

## Sequence of Bovine $\beta_2$ -Microglobulin

TABLE II  
Amino acid composition of peptic (P) and tryptic (T) peptides

	P-1	P-2	P-3	P-4	P-5	P-6	T-1	T-2	T-3	T-4	T-5	T-6	T-7	T-8	T-9
S-Carboxy-methyl-cysteine		0.6 3.0	1.0		0.6 4.2 1.8	2.0		1.0	1.0	0.7 1.1			1.0	1.8	1.0
Aspartic acid		0.9		0.8	3.0				3.1 <sup>a</sup>	0.8					
Threonine	2.0	1.2			4.4		1.1	1.2	2.1	1.1	2.2				
Serine	2.0	3.1			1.8	1.0	2.0				1.2				
Glutamic acid		1.1						1.0							
Proline				1.0	1.2										
Glycine	0.8				2.8	0.8					0.9	0.8		1.1	
Alanine	1.8		0.8		1.0	0.8	1.0					1.0		0.8	
Valine		1.1	1.0	2.0	3.9	1.2			1.0		1.0			1.0	1.1
Isoleucine		1.9		1.0	1.8					1.0					
Leucine					1.0				0.9						
Tyrosine	1.0	1.0			3.2	1.0	1.1	1.0	1.0			1.2		1.1	
Phenylalanine		0.9		1.0	1.5						0.9				
Lysine	1.0	1.0			2.6	1.9	1.0			1.0	1.0		1.0	1.0	
Histidine					1 <sup>b</sup>	1 <sup>b</sup>							1 <sup>b</sup>	1 <sup>b</sup>	
Arginine	1-9	11-26	37-39	62-67	62-98	89-98	1-6	42-45	48-57	75-80	83-90	91-93	94-96	91-98	97-98
Tryptophan Residues															

<sup>a</sup> Should be 4 serine, probably hydrolytic destruction.

<sup>b</sup> Assumed, not determined.

### COMPARISON OF $\beta_2$ -M SEQUENCE DATA

	10	20	30	40	50
BOVINE	I Q R P P K I Q V Y S R H P P E N G K P N Y L N C Y V Y G F H P P Q I E I D L L K N G E K I K S <span style="border: 1px solid black; display: inline-block; width: 1em; height: 1em; vertical-align: middle;"></span> E Q				
GUINEA PIG	V L H A — R V ————— A ————— Q — F I ————— S ————— V E ————— K — D N V — M				
RABBIT	V — A — N V ————— A ————— F ————— S ————— D — E ————— V — E N V —				
HUMAN	———— T ————— A ————— S — F ————— S ————— S D — V ————— D — R — E K V — H				
MOUSE	———— K T — Q ————— I ————— T Q ————— H ————— Q M ————— K — P K V — M				50
	60	70	80	90	98
BOVINE	S D L S F S K D W S F Y L L S H A E F T P D S K D E Y S C R V K H V T L E Q P R I V K W D R D L				
GUINEA PIG	———— T ————— V — A ————— N D S ————— S — I — S E — K ————— P N K				
RABBIT	———— N ————— V — T ————— N N — N ————— K E — M T ————— Y				
HUMAN	———— Y Y T ————— T E ————— A ————— N ————— S — K ————— M				
MOUSE	———— M ————— I — A — T ————— T E T — T — A ————— A S M A E — K T — Y ————— M				99

FIG. 3. Sequences of  $\beta_2$ -microglobulins aligned for homologies. Numbering at top refers to the bovine sequence. Deletion in the alignment is indicated by  $\square$ . The standard one-letter symbols for amino acids are used (23). A solid line indicates identity with the bovine sequence. Sequences other than bovine  $\beta_2$ -m were determined by other investigators: guinea pig (15); rabbit (14); human (13); mouse, allelic form from strain C57BL/6 (16).

and with a longer digestion time, the  $\text{NH}_2$ -terminal peptide (residues 1-36) and carboxyl-terminal peptide (residues 69-98) were produced in relatively large yields. In this instance, besides the cleavage at residues 36 to 37, the glutamyl-phenylalanyl bond (68-69) was also cleaved. With still longer digestion times, hydrolysis at glutamyl-lysine and glutamyl-glutamine were observed. Under the conditions of digestion, not only was *S. aureus* V8 protease specific for glutamyl bonds, but limited to cleavage of residues lying in favorable environments. The application of this enzyme alone to elucidate a complete sequence demonstrates its uniqueness and suggests that it should be considered in any protocol for protein structure investigations.

$\beta_2$ -Microglobulin has proved to be a highly conserved molecule among divergent species. Fig. 3 shows a comparison of the sequence of bovine  $\beta_2$ -microglobulin (residues 1 to 98)

with the published sequences (residues 1 to 99) of guinea pig (15), rabbit (14), mouse (16), and human (13)  $\beta_2$ -microglobulins. The bovine molecule is characterized by a deletion at position 49 where valine is present in the other species. The serine at 67 for human  $\beta_2$ -m (13) has been deleted because recent studies<sup>4</sup> suggest that the amino acid originally assigned to that position is not present. A total of 48 residues show no substitution among  $\beta_2$ -microglobulins. Bovine  $\beta_2$ -m shows a repeat sequence at residues 13-16 and 31-34 with the exception of a glutamine for glutamic acid at residue 34. Another repeat sequence is found at residues 56-58 and 73-75. Probably by coincidence, in each set, the repeat sequence occurs after 14 intervening residues. At residue 47, the bovine molecule exhibits a striking charge reversal, and at position 48, a

<sup>4</sup> J. A. Ziffer and B. A. Cunningham, personal communication.

**Complete amino acid sequence of bovine  $\beta_2$ -microglobulin.**

*Merton L. Groves and Rae Greenberg*

**Page 2622, Table II:**

Due to a printer's error, the table was set with the amino acid headings out of alignment with the data presented. The correct table is reprinted here:

TABLE II  
*Amino acid composition of peptic (P) and tryptic (T) peptides*

	P-1	P-2	P-3	P-4	P-5	P-6	T-1	T-2	T-3	T-4	T-5	T-6	T-7	T-8	T-9
S-Carboxymethylcysteine		0.6			0.6					0.7					
Aspartic acid		3.0	1.0		4.2	2.0		1.0	1.0	1.1			1.0	1.8	1.0
Threonine					1.8						0.9				
Serine		0.9		0.8	3.0				3.1 <sup>a</sup>	0.8					
Glutamic acid	2.0	1.2			4.4		1.1	1.2	2.1	1.1	2.2				
Proline	2.0	3.1			1.8	1.0	2.0				1.2				
Glycine		1.1						1.0							
Alanine				1.0	1.2										
Valine	0.8				2.8	0.8					0.9	0.8		1.1	
Isoleucine	1.8		0.8		1.0	0.8	1.0					1.0		0.8	
Leucine		1.1	1.0	2.0	3.9	1.2			1.0		1.0			1.0	1.1
Tyrosine		1.9		1.0	1.8					1.0					
Phenylalanine					1.0				0.9						
Lysine	1.0	1.0			3.2	1.0	1.1	1.0	1.0			1.2		1.1	
Histidine		0.9		1.0	1.5						0.9				
Arginine	1.0	1.0			2.6	1.9	1.0			1.0	1.0		1.0	1.0	
Tryptophan					1 <sup>b</sup>	1 <sup>b</sup>							1 <sup>b</sup>	1 <sup>b</sup>	
Residues	1-9	11-26	37-39	62-67	62-98	89-98	1-6	42-45	48-57	75-80	83-90	91-93	94-96	91-98	97-98

<sup>a</sup> Should be 4 serine, probably hydrolytic destruction.

<sup>b</sup> Assumed, not determined.

## Sequence of Bovine $\beta_2$ -Microglobulin

	<u>COW</u>		<u>G.P.</u>		<u>RABBIT</u>		<u>HUMAN</u>		<u>MOUSE</u>	
COW	-									
GUINEA PIG	32	-								
RABBIT	24	25	-							
HUMAN	26	32	27	-						
MOUSE	32	39	34	32	-					

FIG. 4. Amino acid difference matrix for  $\beta_2$ -microglobulins.

charge change (neutral *versus* basic) when compared to most of the other species. One cystine and two tryptophan residues are invariant in all species examined. The constant cystine would be expected since the —S—S— bond in the native protein provides structural stability. A matrix showing the absolute amino acid differences between  $\beta_2$ -m species is presented in Fig. 4. The values vary from a low of 24 substitutions between rabbit and cow  $\beta_2$ -m to a high of 39 for the guinea pig and mouse comparison. Bovine  $\beta_2$ -m contains no methionine while human, guinea pig, and rabbit have 1 and mouse has 4 residues. The bovine protein is also uniquely characterized by three di-prolyl sequences in the first third of the molecule. Two areas of  $\beta_2$ -m are highly conserved, residues 8–19 and 51–64. Perhaps they are involved in forming complexes with the histocompatibility antigens.

*Acknowledgments*—We thank H. J. Dower for outstanding contributions in the amino acid analyses and for highly efficient record keeping. Thanks also to A. I. Rivenburgh, A. J. Menna, and D. K. Koch for precision illustrations.

### REFERENCES

1. Peterson, R. A., Rask, L., and Ostberg, L. (1977) *Adv. Cancer Res.* **24**, 115–163
2. Plesner, T. (1980) *Allergy* **35**, 627–637
3. Mavligit, G. M., Stuckey, S. E., Cabanillas, F. F., Keating, M. J., Tourtellotte, W. W., Schold, S. C., and Freireich, E. J. (1980) *N. Engl. J. Med.* **303**, 718–722
4. Viberti, G. C., Keen, H., and Mackintosh, D. (1981) *Br. Med. J.* **282**, 95–98
5. Lancet, D., Parham, P., and Strominger, J. L. (1979) *Proc. Natl. Acad. Sci. U. S. A.* **76**, 3844–3848
6. Owen, M. J., Kissonerghis, A.-M., and Lodish, H. F. (1980) *J. Biol. Chem.* **255**, 9678–9684
7. Goding, J. W., and Walker, I. D. (1980) *Proc. Natl. Acad. Sci. U. S. A.* **77**, 7395–7399
8. Groves, M. L., Basch, J. J., and Gordon, W. G. (1963) *Biochemistry* **2**, 814–817
9. Groves, M. L., and Greenberg, R. (1977) *Biochem. Biophys. Res. Commun.* **77**, 320–327
10. Kumosinski, T. F., Brown, E. M., and Groves, M. L. (1981) *J. Biol. Chem.* **256**, 10949–10953
11. Becker, J. W., Ziffer, J. A., Edelman, G. M., and Cunningham, B. A. (1977) *Proc. Natl. Acad. Sci. U. S. A.* **74**, 3345–3349
12. Krangel, M. S., Orr, H. T., and Strominger, J. L. (1980) *Scand. J. Immunol.* **11**, 561–571
13. Cunningham, B. A., Wang, J. L., Berggård, I., and Peterson, P. A. (1973) *Biochemistry* **12**, 4811–4822
14. Gates, F. T., III, Coligan, J. E., and Kindt, T. J. (1979) *Biochemistry* **18**, 2267–2272
15. Wolfe, P. B., and Cebra, J. J. (1980) *Mol. Immunol.* **17**, 1493–1505
16. Gates, F. T., III, Coligan, J. E., and Kindt, T. J. (1981) *Proc. Natl. Acad. Sci. U. S. A.* **78**, 554–558
17. Davis, B. J. (1964) *Ann. N. Y. Acad. Sci.* **121**, 404–427
18. Schechter, Y., Patchornik, A., and Burstein, Y. (1973) *Biochemistry* **12**, 3407–3413
19. Pisano, J. J., Bronzert, T. J., and Brewer, H. B., Jr. (1972) *Anal. Biochem.* **45**, 43–59
20. Jeppsson, J. O., and Sjöquist, J. (1967) *Anal. Biochem.* **18**, 264–269
21. Smithies, O., Gibson, D., Fanning, E. M., Goodflesh, R. M., Gilman, J. G., and Ballantyne, D. L. (1971) *Biochemistry* **10**, 4912–4921
22. Drapeau, G. R. (1977) *Methods Enzymol.* **47**, 189–191
23. IUPAC-IUB Commission on Biochemical Nomenclature (1968) *J. Biol. Chem.* **243**, 3557–3559

# Sequence of Bovine $\beta_2$ -Microglobulin

## SUPPLEMENTARY MATERIAL TO

### COMPLETE AMINO ACID SEQUENCE OF BOVINE $\beta_2$ -MICROGLOBULIN

Merton L. Groves and Rae Greenberg

#### EXPERIMENTAL PROCEDURES

##### Isolation of $\beta_2$ -Microglobulin

Casein was precipitated from skim colostrum after dilution with an equal volume of water by the addition of 1 N HCl to a pH of 4.5. If casein did not precipitate, more dilution with water or a fine adjustment of pH were recovered by centrifugation and then lyophilized.  $\beta_2$ -microglobulin was isolated from this casein in four steps at 3°. 1. Colostrum casein (24 g) was dissolved in .005 M sodium phosphate, pH 8.3, dialyzed, and applied to a column of fibrous DEAE-cellulose previously equilibrated with the same buffer. Protein recovery from this column totaled 6.4 g, Fig. 15. Of this amount 5 g were obtained as precipitate when earlier fractions were acidified with 1 N acetic acid to pH 5. The remaining 1.4 g fraction, enriched in  $\beta_2$ -m, was the starting material for the next step. The precipitate, 5 g, contained no  $\beta_2$ -m as shown by disc gel electrophoretic patterns, pH 4.3, 8 M urea, Fig. 15-A, gels 1 to 4. Lactoferrin, among other proteins begins to elute before  $\beta_2$ -microglobulin while  $\beta_2$ -m continued to elute after lactoferrin was off the column. 11. The fractions enriched in  $\beta_2$ -m (Fig. 15-B) from 4 DEAE-cellulose fractionations (about 5.6 g) were combined, dissolved in .05 M potassium phosphate, pH 5.5, dialyzed, and applied to a CM-cellulose column. With this buffer, 4 g protein (low in  $\beta_2$ -m) were recovered before the gradient was started. In the elution profile of the gradient, Fig. 25, lactoferrin was concentrated in the second peak. Polyacrylamide gel electrophoretic patterns, pH 4.3, 8 M urea, of combined fractions, Fig. 35, showed that  $\beta_2$ -m was concentrated in fractions 7 and 8, fractions 1 to 4 represented material eluted before the gradient. 111. Fractions 7 and 8, Fig. 25 and 35, were each applied to a Biogel P-60 column in a .025 M sodium acetate buffer, pH 5.5. Fig. 45 illustrates the elution pattern for fraction 8. Bovine  $\beta_2$ -m undergoes a concentration dependent reversible aggregation (10) so that the elution position for  $\beta_2$ -m will vary depending on its relative concentration. 1V.  $\beta_2$ -Microglobulin from the gel filtration column was dissolved by the addition of acetic acid to pH 5.0. Exposure time for  $\beta_2$ -m at pH 5.0 or less and at 25° was minimized to avoid denaturation (10). The pH of the solution was slowly increased by the addition of .1 N NaOH; the solution became turbid, then showed strong birefringence and shortly after that  $\beta_2$ -m crystals were evident. After a few days at 3°,  $\beta_2$ -m crystals were harvested by centrifugation. About 1 mg of crystalline  $\beta_2$ -microglobulin was obtained from 1 g of colostrum casein.

##### Reduction and Alkylation

Crystalline  $\beta_2$ -m was reduced and S-carboxymethylated by a modification of the method of Schechter *et al.* (18) in which 30 mg  $\beta_2$ -m was used per batch. The reduced and S-carboxymethylated protein (CM  $\beta_2$ -m) was dialyzed in the dark at 3° and then recovered by lyophilization. Amino acid analysis confirmed the presence of two S-carboxymethyl cysteine residues.

##### Polyacrylamide Gel Electrophoresis

Polyacrylamide disc gel electrophoresis was performed by the method described by Davis (17) with 7.5% gels at either pH 8.9 or 4.3. The alkaline gels contained 4 M urea while pH 4.3 gels were 8 M in urea.

##### Amino Acid Analysis

Protein and peptide samples were hydrolyzed for 24 h with 5.7 N HCl containing phenol (.05%) in sealed evacuated tubes. Analyses were performed on a Beckman<sup>3</sup> 119 CL amino acid analyzer and results reported as molar ratios.

##### High Voltage Paper Electrophoresis (HVE)

Both analytical and preparative HVE were carried out on Whatman 3 MM paper for 1.5 h at 40 V/cm in a horizontal water-cooled apparatus with pyridine-acetate buffers pH 6.4, 4.4, or 3.7.

##### Digestion by *S. Aureus* V8 Protease at pH 4

CM  $\beta_2$ -m at a concentration of .1% was dissolved in .1 M ammonium acetate, 2 mM EDTA, pH 4 and *S. aureus* V8 protease (Miles) solution in the same buffer was added at an enzyme to protein ratio of 1/50. The solution was incubated at 37° for 1 h and lyophilized. The peptides produced by the this cleavage were separated on Sephadex G 50 superfine and combined fractions were designated V8-(4)-2 and V8-(4)-1 as shown in Fig. 55.

##### Digestion by *S. Aureus* V8 Protease at pH 8

To a .4% solution of  $\beta_2$ -m in .5%  $\text{NH}_4\text{HCO}_3$ , 2mM EDTA, V8 protease was added to give an enzyme to protein ratio of 1/100. After 2 h at 37° a small amount of precipitate was removed and the supernatant lyophilized. After removal of  $\text{NH}_4\text{HCO}_3$  by repeated lyophilization, the peptides recovered from this cleavage were also separated on Sephadex G 50 superfine, Fig. 65. The peptides in pool (a) were dissolved in .1 M  $\text{NH}_4\text{HCO}_3$ , pH 8.5 and applied to a DEAE-Sephadex column. Elution was effected with a gradient of 75 ml of the .1 M buffer and 75 ml 1.0 M  $\text{NH}_4\text{HCO}_3$ , pH 8.5. The combined fractions were designated V8-(8)-2 and V8-(8)-1, Fig. 75. Digestion by *S. aureus* V8 protease at pH 8 of another sample of CM  $\beta_2$ -m was carried out as described but the digestion time was increased to 4 h in order to obtain smaller peptides. Separation of the peptides on a Sephadex G 50 superfine column as described yielded two peaks and the peptides in the second peak, enriched in smaller molecules were finally purified by high voltage paper electrophoresis and designated peptides V8-(8)-4 and V8-(8)-5.

##### Digestion by Pepsin

The digestion was carried out on a solution of .8% CM  $\beta_2$ -m in 5% formic acid, pH 2, with pepsin (Worthington) at an enzyme to protein ratio of 1/100 for 6 h at 37°. The resulting lyophilized peptides were separated on a Sephadex G 25 fine column equilibrated with .1 N acetic acid. Three major peaks determined by absorption at 280 nm were eluted. The peptides in the second and third peaks were each further fractionated by HVE. Peak 2 yielded pepsin peptide P-6 and peak 3 contained pepsin peptides P-4 and P-3. The peptides in the first major peak from Sephadex were fractionated on a DEAE-Sephadex A 25 column with the  $\text{NH}_4\text{HCO}_3$  gradient described earlier. An early peak detectable only at 220 nm and a major peak at 280 nm yielded P-1 and P-2, respectively, after HVE. Another digestion was carried out at pH 2.2, .5% protein with an enzyme to protein ratio of 1/100 for 30 min, 25°. The lyophilized material was fractionated as described for the previous pepsin experiment with one major peak, peptide P-5, isolated from the Sephadex ion-exchanger.

##### Digestion by Trypsin

CM  $\beta_2$ -m (1.7% protein) at pH 8.2, was treated with TPCK-trypsin (Worthington) at an enzyme to protein ratio of 1/100 for 1 h at 37°, then another aliquot of trypsin was added and digestion continued another h. The recovered peptides were separated on a Sephadex G front, several lesser fractions containing smaller peptides were obtained and three of these pools were each fractionated on an ion exchange column of AG 50W-X4, 200-400 mesh (.6 X 17 cm) at 37° with a gradient of .2 M pyridine acetate, pH 3.1 and 2 M pyridine acetate, pH 4.6. Some fractions were separated by preparative HVE. Purity of peptides was conformed by HVE and when possible by polyacrylamide gel electrophoresis and with the exception of T-8, these peptides were designated T-1 to T-9. Peptide T-8 was isolated after digestion of a .3% solution of CM  $\beta_2$ -m in .1 M  $\text{NH}_4\text{HCO}_3$ , 4 M urea, pH 8, with trypsin at an enzyme to protein ratio of 1/100 37° for 15 min. The solution containing the peptides was adjusted to pH 3.3 and applied directly to a Sephadex G 25 column equilibrated with .1 N acetic acid and the second major peak eluted was designated peptide T-8.

##### Carboxypeptidase Digestion

Carboxypeptidases A and B (Worthington) were used at a ratio of 1/50 at pH 8.2 in unbuffered systems. Carboxypeptidase Y (Worthington) digestion was carried out in .1 M pyridine-acetate at pH 6.5 also at a ratio of 1/50. Since both native and CM  $\beta_2$ -m were quite insoluble between pH 6 and 8, limited carboxyl-terminal sequence information was obtained on the whole molecule.

#### Amino Acid Sequence

Sequencing was carried out on a Beckman 890 C sequencer using the double cleavage 1 M Quadrol protein program of 042772 and when appropriate a .25 M Quadrol single cleavage peptide/protein program using polybrene and a simultaneous benzene and ethyl-acetate wash. Identification of PTH amino acids was accomplished by GLC and thin-layer chromatography (19, 20) and/or hydrolysis with HI back to the parent amino acid (21). All protein and peptide samples were sequenced at least twice and more often three times. With sufficient material available for sequencing (250 to 500 nm per run) and yields ranging from 94 to 96%, there was no difficulty obtaining meaningful information up to 40 cycles. Backgrounds on the GLC quantitation did not exceed 20%. Results of the major sequencing operation are shown in Tables 15-1115.

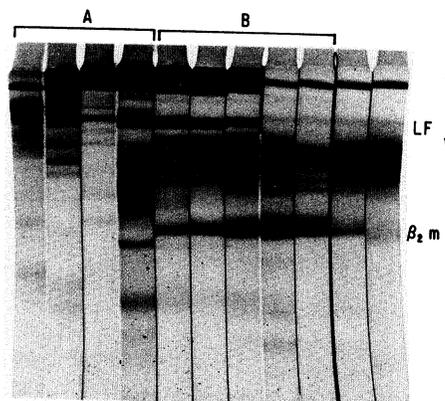


Fig. 15. Polyacrylamide gel electrophoresis of combined fractions of colostrum casein, eluted from a column of fibrous DEAE-cellulose (4 X 50 cm) with .005 M sodium phosphate, pH 8.2. Elution rate was 120 ml/h and 20 ml fractions were collected. See text for gels under A and B. LF refers to lactoferrin.

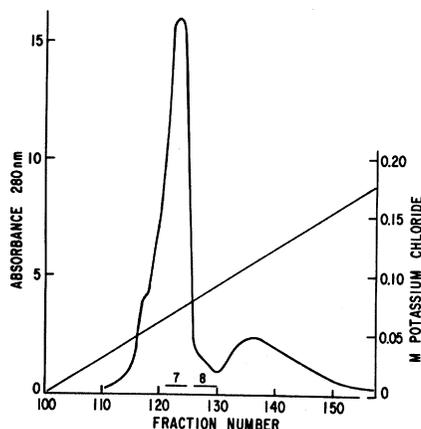


Fig. 25. Fractionation of proteins from DEAE-cellulose enriched  $\beta_2$ -m on CM-cellulose (2 X 36 cm). After protein elution with the starting buffer a gradient of 250 ml .05 M potassium phosphate, pH 5.5 and 250 ml .1 M potassium chloride, 2 M potassium chloride, pH 7.7 was used. The flow rate was 24 ml/h with a volume of 8 ml/tube. The gradient indicated is that calculated for potassium chloride. The fractions were pooled as indicated.

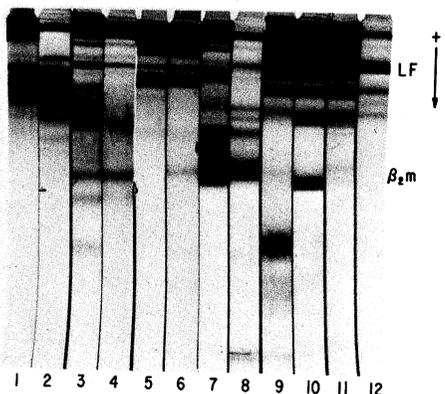


Fig. 35. Polyacrylamide gel electrophoretic patterns, pH 4.5, 8M urea of combined fractions eluted from CM-cellulose, Fig. 25. Gels 1-4 represent fractions eluted before the gradient was started while 7 and 8 represent fractions enriched in  $\beta_2$ -m.

<sup>3</sup>Reference to brand or firm name does not constitute endorsement by the U.S. Department of Agriculture over others of a similar nature not mentioned.

## Sequence of Bovine $\beta_2$ -Microglobulin

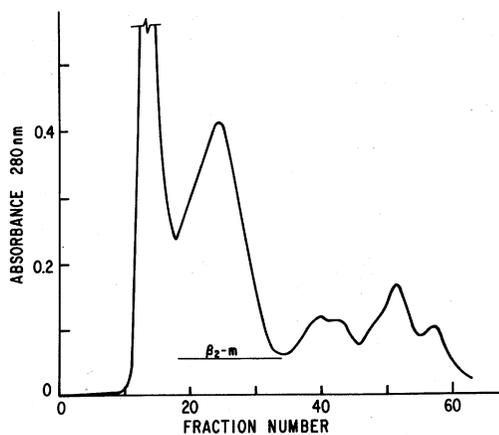


Fig. 45. Fractionation of proteins associated with  $\beta_2$ -m on a column (2 X 68 cm) of Biogel P. 60, 100-200 mesh in .025 M sodium acetate, pH 5.5. Flow rate was 15 ml/h and 5 ml fractions were collected. The pool containing  $\beta_2$ -m was collected as indicated.

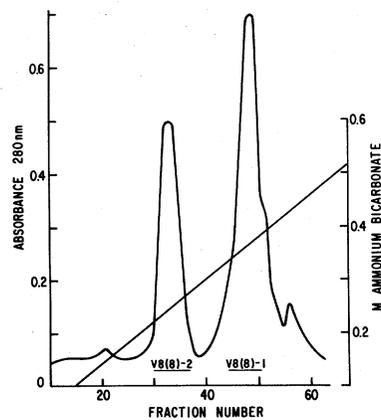


Fig. 75. Fractionation of peptides produced by *S. aureus* V8 digestion at pH 8 of CM  $\beta_2$ -m, Fig. 6S, pool (a) on a DEAE-Sephadex A 25 column (.8 X 25 cm). A gradient of .1 M  $\text{NH}_4\text{HCO}_3$ , pH 8.5 and 1.0 M  $\text{NH}_4\text{HCO}_3$ , pH 8.5 was run over a period of 16 h, 25°. Flow rate was 8 ml/h and 1.4 ml fractions were collected. Fractions were combined as indicated.

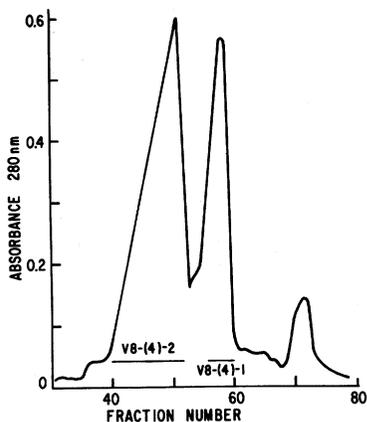


Fig. 55. Fractionation of peptides produced by *S. aureus* V8 cleavage at pH 4 of CM  $\beta_2$ -m on a column (9 X 163 cm) of Sephadex G 50 superfine in 1 M HCOOH. Flow rate was 8 ml/h and 1.6 ml fractions were collected. Pooled fractions are as indicated.

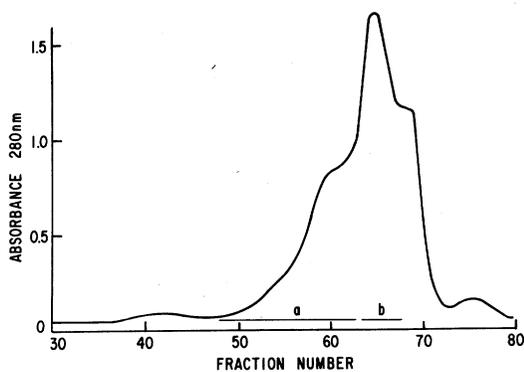


Fig. 65. Fractionation of peptides produced by *S. aureus* V8 cleavage at pH 8 of CM  $\beta_2$ -m on a column (.9 X 163 cm) of Sephadex G 50 superfine in .1 M HCOOH. Flow rate was 8 ml/h and 1.3 ml fractions were collected. The pools were collected as indicated.

TABLE IS  
Automated Edman Degradation of CM- $\beta_2$ -m

Cycle	Amino acid	Yield (nmol)	Method of identification <sup>a</sup>
1	Ile	250	GC, TLC
2	Gln		GC, TLC, AAA
3	Arg		Spot test
4	Pro	228	GC, TLC
5	Pro		GC, TLC
6	Lys		GC, TLC
7	Ile		GC, TLC
8	Gln		TLC, AAA
9	Val		GC, TLC
10	Tyr	178	GC, TLC
11	Ser		AAA
12	Arg		Spot test
13	His		Spot test
14	Pro	140	GC, TLC
15	Pro	116	GC, TLC
16	Glu		GC, TLC
17	Asn		TLC, AAA
18	Gly		GC, TLC
19	Lys		GC, TLC
20	Pro	99	GC, TLC
21	Asn		TLC, AAA
22	Tyr	109	GC, TLC
23	Leu	90	GC, TLC
24	Asn		TLC, AAA
25	Cys		GC, TLC, AAA
26	Tyr	92	GC, TLC
27	Val	77	GC, TLC
28	Tyr	85	GC
29	Gly		GC, TLC
30	Phe		GC, TLC
31	His		Spot test
32	Pro		GC, TLC
33	Pro		GC, TLC
34	Gln		TLC
35	Ile		GC, TLC
36	Glu		TLC
37	Ile	40	GC, TLC
38	?		
39	Leu		TLC
40	Leu		TLC
41	Lys		TLC

Yield 95.7%, Tyr 10 + 26.

<sup>a</sup> GC is gas chromatography, TLC is thin layer chromatography, AAA is amino acid analysis after HI hydrolysis. Spot tests are diazotized sulfanilic acid test for histidine and phenanthrene quinone test for arginine.

# Sequence of Bovine $\beta_2$ -Microglobulin

TABLE IIS  
Automated Edman Degradation of V8-(4)-2

Cycle	Amino acid	Yield (nmol)	Method of identification
1	Ile	480	GC, TLC
2	Asp		GC, TLC
3	Leu	437	GC, TLC
4	Leu		GC, TLC
5	Lys		GC, TLC
6	Asn		GC, TLC
7	Gly		GC, TLC
8	Glu		GC, TLC
9	Lys		TLC
10	Ile	365	GC, TLC
11	Lys		GC, TLC, AAA
12	Ser		AAA
13	Glu		GC, TLC
14	Gln		GC, TLC
15	Ser		AAA
16	Asp		GC, TLC
17	Leu	250	GC, TLC
18	Ser		AAA
19	Phe	211	GC, TLC
20	Ser		AAA
21	Lys		GC, TLC
22	Asp		GC, TLC
23	Trp		GC, TLC
24	Ser		AAA
25	Phe	128	GC, TLC
26	Tyr		GC, TLC
27	Leu	108	GC, TLC
28	Leu		GC, TLC
29	Ser		AAA
30	His		Spot test
31	Ala		GC, TLC
32	Glu		GC, TLC
33	Phe	65	GC, TLC
34	Thr		GC, TLC
35	Pro		GC, TLC
36	Asp		GC, TLC

Yield 94.3%, Leu 3 + 27.

See footnote - Table IS.

TABLE IIIS  
Automated Edman Degradation of V8-(8)-2

Cycle	Amino acid	Yield (nmol)	Method of identification
1	Phe	225	GC, TLC
2	Thr		AAA
3	Pro		GC, TLC
4	Asp		GC, TLC
5	Ser		AAA
6	Lys		GC, TLC
7	Asp		GC, TLC
8	Glu		GC, TLC
9	Tyr		GC, TLC
10	Ser		AAA
11	Cys		GC, TLC
12	Arg		Spot test
13	Val	72	GC, TLC
14	Lys		GC, TLC
15	His		Spot test
16	Val	56	GC, TLC
17	Thr		GC, TLC, AAA
18	Leu	55	GC, TLC
19	Glu		GC, TLC
20	Gln		GC, TLC
21	Pro		Spot test
22	Arg		GC, TLC
23	Ile	35	GC, TLC
24	Val	37	GC, TLC
25	Lys		TLC
26	Trp		GC, TLC
27	Asp		GC, TLC
28	Arg		Spot test
29	Asp		GC, TLC

Yield 94.1%, Val 13 + 24.

See footnote - Table IS.