Overexpression of native human $\beta_2$-microglobulin in Escherichia coli and its purification

(Circular dichroism; HLA heavy chain; periplasmic space; preprotein; recombinant DNA; secretion; signal peptidase; signal sequence)

Kenneth C. Parker and Don C. Wiley

Department of Biochemistry and Molecular Biology, Howard Hughes Medical Institute, Harvard University, Cambridge, MA 02138 (U.S.A.)

Received by M. Bagdasarian: 8 February 1989
Revised: 27 April 1989
Accepted: 1 May 1989

SUMMARY

$\beta_2$-Microglobulin ($\beta_2$M), the small subunit of human leukocyte antigen (HLA) class-I proteins, has been synthesized in *Escherichia coli* and purified in mg amounts. A $\beta_2m$ cDNA clone was fused in-frame behind DNA encoding the signal sequence for the outer membrane protein, OmpA. Three different constructions were made, whose products differed by the insertion of either an extra Ala residue, the hexapeptide AEFLEA [single-letter amino acid (aa) code], or no aa between the OmpA signal sequence and $\beta_2$M-coding sequence. All three protein products were correctly processed by bacterial signal peptidase, as determined by N-terminal sequencing, and all three were secreted as soluble proteins into the periplasmic space. However, the signal sequence of the preprotein with the inserted hexapeptide, AEFLEA, was cleaved to a much greater degree than the other two preproteins. When there was no insertion, the mature protein was identical to human $\beta_2$M, as analyzed by sodium dodecyl sulfate–polyacrylamide-gel electrophoresis, circular dichroism, and native isoelectric focusing. This 'bacterial $\beta_2$M', radiolabeled with Bolton–Hunter reagent, was able to exchange into papain-solubilized HLA-B7, as determined by Sephadex G-75 chromatography and immune precipitation, indicating that bacterial $\beta_2$M could complex with the heavy chain of HLA-B7.

INTRODUCTION

$\beta_2$M is a monomeric 99-aa serum protein, originally isolated from the urine of patients with kidney disease (Berggård and Bearn, 1968). It is also the small subunit of HLA class-I histocompatibility antigens (see Nathenson et al., 1986, for a recent review). $\beta_2$M is necessary for the cell surface expression of FPLC, fast protein liquid chromatography (Pharmacia); HLA, human leukocyte antigen; IPTG, isopropyl-$\beta$-D-thiogalactopyranoside; LB, Luria–Bertani (broth); mAb, monoclonal antibody; oligo, oligodeoxynucleotide; Omp, outer membrane protein; PAGE, polyacrylamide-gel electrophoresis; PolIk, Klonev (large) fragment of *E. coli* DNA polymerase I; SDS, sodium dodecyl sulfate; UV, ultraviolet; [], denotes plasmid-carrier state.
class-I heavy chains, though it is noncovalently attached and is freely exchangeable with the \( \beta_2 \)M in blood in vivo (Kimura et al., 1983). The tertiary structure of the bovine milk homologue has been determined by x-ray crystallography (Becker and Reeke, 1985), and that of human \( \beta_2 \)M by the crystal structure study of human HLA (Bjorkman et al., 1987).

The function of both HLA class-I and class-II proteins is to bind one of a set of internally- or externally-generated peptides, derived from normal cell proteins, viral proteins or tumor antigens (Townsend et al., 1986; Buus et al., 1987). The complex of peptide plus class-I or class-II protein is recognized by the T-cell receptor, resulting in diverse aspects of the cellular immune response. To date, it has been possible to demonstrate peptide binding to purified class-II proteins, but not to purified class-I proteins*. Recently, the three-dimensional structure of the class-I protein, HLA-A2, was determined by x-ray crystallography (Bjorkman et al., 1987). The structure revealed a cleft on the surface of the protein, presumably suited for binding peptide, and some unknown substance already bound in the cleft. Because of the difficulty of isolating large amounts of pure class-I proteins from lymphocyte tissue culture cells, and because of the dilemma of peptide binding, it is important to develop an alternative source of starting material for these studies. In this report, we demonstrate the successful synthesis and large-scale purification, from bacteria, of human \( \beta_2 \)M, the small subunit of the HLA class-I proteins.

The expression/secretion vector, pIN-III-ompA2, was obtained from M. Inouye (Ghahreyn et al., 1984). Oligos were synthesized and purified by B. Lane at the Harvard Protein Chemistry facility and were used without further purification. M13mp18/19 was purchased from New England Biolabs. Human \( \beta_2 \)M was a kind gift of Dr. Arnold Sanderson. LB was from Gibco BRL. Papain-solubilized HLA-B7 was prepared by Anastasia Haykova as described (Parham et al., 1977). Bolton–Hunter reagent was obtained from New England Nuclear.

(b) Construction of plasmids

The 441-bp \( \text{HaeIII}\)-polI fragment encoding mature \( \beta_2 \)M and 4 aa of the signal sequence was inserted into pIN-III-ompA2 that was cut with \( \text{EcoRI} \) and filled with PolI to generate p715. The 542-bp \( \text{XbaI}\)-HindIII fragment was inserted into M13mp19, and site-directed mutagenesis was carried out with oligo ACC-GTA-GCT-CAG-GCC-ATC-CAG-CGT-ACC, or oligo ACC-GTA-GCT-CAG-GCC-GCT-ATC-CAG-CGT-ACC, to generate the p714 or p741 sequences, respectively, according to Norrander et al. (1983). The \( \text{XbaI}\)-HindIII fragment was then reinserted into pIN-III-ompA2 to generate p714 or p741 (see Fig. 1). All DNA manipulations were according to standard protocols (Maniatis et al., 1982).

(c) Expression and purification of bacterial \( \beta_2 \)M

In a typical experiment, several liters of LB, supplemented with 50 \( \mu \)g Ap/ml, were inoculated with an overnight culture of JM103 carrying either p714, p715, or p741, and grown at 30°C. In the case of p714 and p741, induction by IPTG resulted primarily in overexpression of preprotein, and had no effect on the level of mature protein; consequently, the cultures were simply harvested just before lag phase. In contrast, JM103[p715] was induced with 2 mM IPTG at an \( A_{650} \) of 0.8, and harvested just before lag phase. The cells could be stored for several days at 0°C without affecting the yield of intact b-\( \beta_2 \)M.

To release the contents of periplasmic space, the cells were subjected to osmotic shock (Nossal and Heppel, 1966). Briefly, the cells were pelleted, washed in 30 mM Tris·HCl pH 7.0, and resus-

---

* For recent progress in this area, see Chen and Parham (1989); Bouillot et al. (1989).
pended in 10 ml of 30 mM Tris · HCl pH 7.0, per liter of medium. The cell suspension was then diluted into an equal volume of 40% sucrose/30 mM Tris · HCl pH 7.0/1 mM EDTA. After 30 min at room temperature, the cells were pelleted and resuspended in a minimum volume of supernatant. This solution was rapidly diluted into 100 vols. of ice-cold 0.5 mM MgCl₂. The cells were again pelleted and the supernatant was passed through a 0.45-μm filter and saturated to 90% by the addition of 63.6 g of dry solid (NH₄)₂SO₄/100 ml of solution. After incubation overnight at 4°C, the precipitate was pelleted at 5000 rev./min for 20 min in a Beckman JA-14 rotor, and dissolved in 2 ml of water. This solution was dialyzed against several changes of 5 mM Tris · HCl pH 9.0 and applied to a Pharmacia Mono Q anion exchange column running in 10 mM triethanolamine · HCl pH 7.0. The b-β₂M was eluted with a gradient of 0 to 120 mM NaCl in the running buffer. JM103[p714]- or [p741]-derived b-β₂M eluted at 20 mM NaCl, whereas JM103[p715]-derived b-β₂M eluted at 30 mM NaCl. Fractions containing b-β₂M were concentrated over Centricon-10 ultrafilters (Amicon), and purified further on a Pharmacia Superose 12 column running in 10 mM Na phosphate pH 7.0/140 mM NaCl.

(d) Analytical techniques

Five cycles of aa sequencing were performed by Bill Lane on the peak fraction from the Mono Q column, and in all three cases the predicted sequence was obtained. Native isoelectric focusing was generously performed by Michael Silver. Immune precipitation was carried out using the mAb W6/32 (Brodsky et al., 1979), using Pansorbin (Calbiochem) as the precipitating agent.

RESULTS AND DISCUSSION

(a) Generation and analysis of p715

To produce human β₂M in E. coli, a cDNA clone corresponding to the coding sequence of the mature protein was inserted into the expression-secretion vector, pIN-III-ompA2 (Ghrayeb et al., 1984). This vector was chosen because it contains DNA encod-

ing the OmpA signal peptide, followed by a unique EcoRI site, a unique HindIII site, and a unique BamHI site, available in all three reading frames. A cDNA clone that is inserted into one of these sites, in the correct reading frame, encodes a fusion protein that should be secreted to the periplasmic space. It is desirable to synthesize a protein like β₂M in the periplasmic space, because normal disulfide bond formation can take place, and purification is simplified. Since β₂M has a single disulfide bond, the correct disulfide bond is the only possible one. A convenient HaeIII site, located four codons 5' to DNA encoding the N-terminal Ile of mature β₂M, was chosen, that could be directly fused to the EcoRI site of the pIN-III-ompA2 created by Ghrayeb et al. (1984). The resulting plasmid (p715) should code for a fusion protein that has the OmpA signal peptide followed by a 6-aa linker region in-frame with mature β₂M (Fig. 1). When this construct was first made, it was immediately noted that the colonies containing this plasmid appeared much larger and more mucoid than those containing vector alone, especially in strains JM103 and MM294. SDS-PAGE analysis revealed that a new protein at 12 kDa was being made, which corresponds to the expected Mr of β₂M. Induction with IPTG led to a several-fold increase in

![Diagram](image)

Fig. 1. Plasmids designed to produce β₂M. (A) A diagram of the vector pIN-III-ompA2. lpp and lpo stand for the lipoprotein promoter and the lac promoter-operator, respectively. ompA is the gene encoding E. coli OmpA; lacI refers to the gene for lac repressor, and amp stands for the Ap-resistance gene. (B) Amino acid sequences of β₂M constructions. The diagram uses one-letter aa sequence code for the signal peptide and the N terminus of the mature protein from p714, p715 and p741. The N terminus of mature human β₂M begins IQRT... (see arrows).
the intensity of this band (Fig. 3). The bulk of the 12-kDa protein was released into the osmotic-shock supernatant as expected, and was by far the major component (Fig. 2, lane 1). Anion exchange chromatography was used to obtain a b-β2M preparation that appeared substantially pure by SDS-PAGE (Fig. 2). N-terminal sequencing confirmed the sequence, AEFLEAIIQR, of which the first 6 aa are derived from the signal sequence and splice junction, and the last 3 aa correspond to the N-terminal sequence of mature β2M.

(b) Generation and analysis of p714 and p741

Since the 6-aa N-terminal extension might have some effect on biological properties, site-directed mutagenesis was used either to delete precisely the DNA coding for the extension (p714), or to delete all but one alanine of the extension (p741). Colonies containing p714 and p741 were not as unusual as those containing p715. Analysis of the product of these plasmids by SDS-PAGE indicated that a smaller amount of protein, running at the expected M_r of β2M, was being synthesized. Instead, a new band running at a M_r slightly larger than lysozyme appeared (Fig. 3). This new band became especially prominent after induction by IPTG. Analysis of the periplasmic fraction following osmotic shock demonstrated that this new band remained in the pellet, unlike the material running at the expected M_r for β2M. Western-blot analysis using the mAb, BBM.1 (Brodsky et al., 1979), demonstrated that both new bands had β2M determinants (not shown). The obvious implication of this is that the higher M_r band represents unclaved precursor protein. Mature b-β2M released into the periplasmic space of JM103[p714] and [p741], was purified by a scheme similar to that developed for the product of p715. Both the p741- and p714-derived b-β2M variants eluted at a lower salt concentration, upon anion exchange chromatography, as expected for proteins that had two fewer negatively-charged residues.

Fig. 2. 0.1% SDS–12% PAGE (Laemmli, 1970) stained with Coomassie brilliant blue R-250, of Mono Q separation of the periplasmic fraction from JM103[p715] prior to (NH₄)₂SO₄ precipitation. Lanes: 1, whole periplasmic fraction; 2, human β2M; 3, BioRad low M_r standards (lysozyme at 14.4 kDa); 4–20, aliquots from successive fractions from the Mono Q column. Fraction 11 contained the largest amount of b-β2M, and was used for further sequence analysis without further purification. The arrowheads mark the position of b-β2M.
N-terminal sequencing showed that b-β2M derived from p714 and p741 had the sequences listed in Fig. 1. It can be estimated by Coomassie blue staining, that b-β2M is being synthesized at the level of about 10 mg per liter of medium for p715, and perhaps fivefold less for p714 and p741.

(c) Exchange of p714-derived b-β2M into HLA-B7

To determine if b-β2M was correctly folded, an aliquot was labeled with the Bolton–Hunter 125I-reagent. Sephadex G-75 chromatography revealed that 30% and 20% of the p714 and p715 products, respectively, were able to exchange into HLA-B7 (Fig. 4). Exchange was separately confirmed for the p714 product by immune precipitation with the mAb W6/32, which recognizes the HLA heavy chain when complexed with β2M (Brodsky et al., 1979; data not shown).

(d) Comparison of p714-derived b-β2M with human β2M

To compare p714-derived b-β2M with β2M isolated from urine, circular dichroism spectra of p714-derived b-β2M were taken in both the far UV and in the aromatic region (Fig. 5). Comparison of these spectra to the previously published spectra of human β2M (Karlsson, 1974; Isenman et al., 1975; Johnson et al., 1979) indicates that the p714-derived b-β2M spectra are as similar to those previously published as to one another. All of these spectra are easily distinguishable from spectra of reduced and alkylated β2M, and of guanidine-denatured β2M (Isenman et al., 1975). As a separate test, HLA-B7, p714-derived b-β2M, and p715-derived b-β2M were run together on an analytical native isoelectric focusing gel. As expected, p714-derived b-β2M runs identically to the β2M from HLA-B7, whereas the p715-

Fig. 4. Sephadex G-75 chromatography of Bolton–Hunter 125I-labeled b-β2M incubated with HLA-B7. Open circles, p714-derived b-β2M alone. Solid circles, p714-derived b-β2M with HLA-B7. Squares, p715-derived b-β2M with HLA-B7. Five μg of b-β2M were dissolved in 2.5 μl of 0.1 M Na2 borate pH 8.4, and 0.5 μCi of Bolton–Hunter 125I-reagent, dissolved in the same buffer, was added and incubated at 0°C for 20 min. Following desalting on a Sephadex G-50 column, the labeled b-β2M was purified on a Sephadex G-75 column. To measure exchange of b-β2M into HLA-B7, 150,000 cpm of 125I-labeled b-β2M (2 x 10^-12 mol) were incubated with 1.2 μg of HLA-B7 (2.6 x 10^-11 mol) for 16 h at 37°C, and then loaded onto a Sephadex G-75 column in 10 mM Tris·HCl pH 8.0/mM Na2SO4/140 mM NaCl/0.1 mg/ml BSA (Parker and Strominger, 1983). 15-drop fractions were collected and counted.
one of the major factors that determine the efficiency of cleavage by signal peptidase. Summers (1988) found that triosephosphate isomerase from chicken muscle could not be secreted or processed when fused to the β-lactamase signal peptide, unless the Arg at position 3 of the mature protein was replaced. Addition of a sequence containing one or two Glu residues immediately after the signal peptide and in front of the Arg, greatly improved transport across the membrane. This suggested that the extreme pK’ of Arg may in some cases interfere with secretion (Summers, 1988), and that this interference can be overcome by ion-pairing. Our data are entirely consistent with these observations. In p714-derived b-β2M, which is poorly processed, there is an Arg at position 3. In p715-derived b-β2M, which is processed and secreted extremely well, there are Glu residues at positions 2 and 5 in front of the Arg at position 9. Presumably, in cases where an Arg at the N terminus does not prevent transport, such as alkaline phosphatase (Kikuchi et al., 1981), other features of the N terminus dominate to allow proper secretion and processing.

(f) Factors that affect expression-secretion

The pIN-III-ompA vector was developed as a general purpose expression-secretion system that is especially advantageous for the synthesis of proteins with disulfide bonds (see Duffaud et al., 1987, for review). Two general problems encountered in the high-level synthesis of proteins in E. coli are protein aggregation and/or incomplete processing by bacterial signal peptidase. Altering the growth conditions has been shown to influence favorably the yield of correctly processed material; e.g., by means of addition of protein synthesis inhibitors (Lee and Beckwith, 1986), by lowering the temperature (Schein and Noteborn, 1988), or by incomplete induction (Ghreayeb et al., 1984). Because induction with 2 mM IPTG leads almost entirely to an increase in the amount of precursor in the case of p714 and p741 (Fig. 3), we relied on the lower level of natural lac inducers in LB (Duffaud et al., 1987) and reduced the growth temperature to 30°C to synthesize the b-β2M variants at an intermediate level. Under those conditions no sign of DNA rearrangements due to selection against expression have been observed.

derived b-β2M has a much more acidic pI, due to the two extra Glu residues in the 6-aa N-terminal extension (not shown).

(e) Effect of amino acid sequence on cleavage by signal peptidase

The aa sequence around the junction of the signal peptide with the N terminus of the mature protein is
(g) Conclusions

Human β₂M and two N-terminal variants have been synthesized in E. coli and can be purified in native form in mg amounts. The protein that is processed by signal peptidase is soluble in aqueous solution, has the native disulfide bond intact, associates with the class-I histocompatibility antigen in exchange reactions, and has the same circular dichroism spectra as β₂M isolated from human urine. p714-derived b-β₂M (Fig. 1b) has the same N-terminus as mature human β₂M and cannot be distinguished from it by any of the techniques used. On the other hand, p715-derived b-β₂M (Fig. 1b) is experimentally distinguishable, because of two additional negative charges in the 6-aa N-terminal extension. The N-terminal extension of p715-derived b-β₂M, however, appears to have little effect on its ability to associate with the heavy chain of the class-I antigen, HLA-B7.

With the expression-secretion system reported here, it is now possible to investigate by site-directed mutagenesis which parts of β₂M are important for correct folding and solubility in the periplasmic space, and which parts are important for association with the HLA chain in vitro. β₂M mutants expressed using this system could also be used to study amyloid fibril formation in vitro (Connors et al., 1985). Work is in progress to extend this same expression system to the HLA heavy chain.

ACKNOWLEDGEMENTS

We thank Kyusung Park and Gerald Fasman for carrying out the circular dichroism analysis. We thank Dan Gewirth and Ted Jardetzky for commenting on the manuscript. K.C.P. was a fellow of the Leukemia Society of America for part of the time during which this work was carried out.

REFERENCES


