A Characterization of the Lumenal Region of Human Tapasin Reveals the Presence of Two Structural Domains†

Mingnan Chen,‡ Walter F. Stafford,§ Gundo Diedrich,|| Amir Khan,⊥ and Marlene Bouvier*,‡

School of Pharmacy, University of Connecticut, Storrs, Connecticut 06269, Boston Biomedical Research Institute, Watertown, Massachusetts 02472, Section of Immunobiology, Yale University School of Medicine, New Haven, Connecticut 06520, and Department of Molecular and Cellular Biochemistry, Harvard University, Cambridge, Massachusetts 02138

Received August 7, 2002; Revised Manuscript Received October 16, 2002

ABSTRACT: Tapasin is a type I membrane glycoprotein involved with other accessory proteins in the assembly of class I MHC−β_2m−peptide complexes in the endoplasmic reticulum. We have probed the three-dimensional structure of the lumenal region of human tapasin (residues 1−392) tagged with a $(His)_6$ sequence at its C-terminus using biochemical and biophysical techniques. The far-UV circular dichroism spectrum revealed that tapasin possesses well-defined secondary structural elements corresponding predominantly to β-sheets. A thermal denaturation curve recorded at 216 nm showed a midpoint transition centered at ∼45 °C. Sedimentation analysis showed that tapasin is monomeric in solution with a sedimentation coefficient, $S°_{20,w}$, of 2.68 S. This value of $S°_{20,w}$ combined with the value of the molar mass obtained by MALDI mass spectrometry (44.2 kDa) yielded a frictional ratio, f/f_0, of 1.47. Assuming tapasin is a prolate ellipsoid, we calculated an apparent length of 22.5 nm and a diameter of 2.62 nm, consistent with an elongated molecular shape. Controlled proteolysis using various enzymes revealed that a narrow region of tapasin near residue 90 is highly susceptible to digestion, resulting in two fragments that are resistant to further cleavage. The identity of these fragments was determined by amino acid sequencing and MALDI mass spectrometry and revealed a 9 kDa N-terminal fragment and a 34 kDa C-terminal fragment. Collectively, these results suggest that tapasin is comprised of two core domains of different sizes loosely linked by a flexible region.

Tapasin is a 48 kDa type I membrane glycoprotein involved in the conformational maturation of class I major histocompatibility complex (MHC)[1]−β_2m−peptide complexes in the lumen of the endoplasmic reticulum (ER). Along with the chaperone protein calreticulin, the thiol-dependent oxidoreductase ERp57, and the transporter associated with antigen processing (TAP), tapasin is part of a transient multiprotein complex that is responsible for the folding and assembly of class I MHC−β_2m−peptide complexes prior to their transport to the cell surface. The recognition of class I MHC−β_2m−peptide complexes at the surface of cells by CD8+ cytotoxic T-lymphocytes is a process critical to maintaining immunity against invading pathogens.

Studies using cell lines with specific defects have been extremely useful in providing insights into the series of molecular events underlying the conformational maturation of class I MHC−β_2m−peptide complexes (*1*, *2*). From these studies, it has been suggested that the folding, assembly, and stabilization of class I MHC−β_2m heterodimers rely on interdependent interactions between tapasin, calreticulin, and ERp57. It was shown that up to four molecules of the tapasin−class I MHC−β_2m heterotrimer associate with each TAP molecule, with apparent substoichiometric amounts of calreticulin and ERp57 (*3*). The binding of antigenic peptides by class I MHC−β_2m heterodimers causes a conformational change in the class I MHC heavy chain that releases all accessory proteins and completes the formation of class I MHC−β_2m−peptide complexes.

Numerous functional roles have been attributed to tapasin within the class I assembly complex, including mediating interactions between class I MHC−β_2m heterodimers and TAP molecules (*2*), stabilization and retention in the ER of class I MHC−β_2m heterodimers (*3*−*6*), stabilization of TAP molecules in a way that enhances their ability to translocate peptides (*7*, *8*), and editing of antigenic peptides to favor high-affinity peptides (*9*−*13*). To date, although many aspects of the tapasin function remain to be clarified, there is sufficient evidence to suggest that this protein is critically

† This work was supported by NIH Grant AI45070 (M.B.), NSF grant BIR 9513060 (W.F.S.), a Deutsche Forschungsgemeinschaft grant (G.D.), and a Long-Term Fellowship from the Human Frontiers Science Program (A.K.).

* To whom correspondence should be addressed: School of Pharmacy, University of Connecticut, 372 Fairfield Road U-92, Storrs, CT 06269. Phone: (860) 486-4355. Fax: (860) 486-4998. E-mail: bouvier@uconnvm.uconn.edu.

‡ University of Connecticut.

§ Boston Biomedical Research Institute.

|| Yale University School of Medicine. Present address: Targeted Molecules, San Diego, CA 92121.

⊥ Harvard University. Present address: Institut Curie/UMR, Paris 75005, France.

[1] Abbreviations: CD, circular dichroism; DTT, dithiothreitol; ER, endoplasmic reticulum; Ig, immunoglobulin; MHC, major histocompatibility complex; MW, molar mass; TAP, transporter associated with antigen processing.

10.1021/bi020521u CCC: $22.00
Published on Web 11/09/2002

important for normal expression of stable class I MHC−β_2m−peptide complexes at the surface of cells. In fact, tapasin has so far been associated solely with the conformational maturation of class I MHC−β_2m−peptide complexes in the ER. This is in marked contrast to calreticulin and ERp57 which participate in the folding of a large number of newly synthesized proteins in the ER. This apparent specialized role of tapasin is consistent with the mapping of the *Tapasin* gene in the proximity of genes encoding other immunologically relevant proteins such as TAP, LMP proteasome, class I, and class II (*14*−*16*).

Structural analysis has suggested that tapasin is a member of the immunoglobulin (Ig) superfamily (*3*, *15*, *17*, *18*) and consists of a large N-terminal region (residues 1−392) localized in the ER, a single transmembrane domain (residues 393−417), and a short C-terminal cytosolic tail (residues 418−428) with an ER retention motif. Studies aimed at characterizing protein−protein interactions within the class I assembly complex have shown that residues 1−50 of the lumenal region of tapasin are essential to maintaining stabilizing interactions between class I MHC−β_2m heterodimers, calreticulin, and ERp57 (*8*), whereas residues 393−428 of the transmembrane and cytoplasmic regions are required for association with TAP (*7*). Similarly, residues 334−342 of the lumenal region were shown to be critical for association with the class I MHC heavy chain, but are not required for interaction with TAP (*19*). Currently, the complete lack of information about the three-dimensional structure of tapasin limits the development of a more comprehensive understanding of its function.

In this paper, we present a structural characterization of recombinant soluble tapasin (residues 1−392) tagged with a $(His)_6$ sequence at its C-terminus (tapasin). Our results indicate that tapasin is monomeric in solution and adopts an asymmetric, elongated structure. We also provide evidence from circular dichroism (CD) that the protein has a relatively high content of β-sheet structures and a thermal denaturation curve centered at ∼45 °C. Controlled proteolysis shows that tapasin is organized into two major domains linked by a solvent-exposed, flexible region. Results from these studies contribute to our efforts in characterizing the structural properties of individual proteins involved in the class I assembly complex (*20*−*22*) with the goal of reconstituting this macromolecular complex in vitro.

MATERIALS AND METHODS

Construction of an Expression Plasmid for Tapasin. The cDNA encoding the soluble domain (residues 1−392) of human tapasin with a $(His)_6$ sequence at its C-terminus (tapasin) was generated by PCR amplification using the plasmid pCR2.1-tapasin (*3*) as a template and the following two synthetic primers: 5′-ACTGGATCCTATAAATATGAAGTCCCTGTCTCTG-3′ (forward) and 5′-GCGTAGGTCGACTCAATGGTGATGGTGATGATGGTCCTCAAGGGAGGG-3′ (reverse). The amplified cDNA was ligated into the pFastBac-1 vector (Invitrogen) using the *Bam*HI and *Sal*I restriction sites. Generation of recombinant baculovirus-expressing soluble $(His)_6$-tagged tapasin was carried out using the Bac-to-Bac baculovirus expression system (Invitrogen) as recommended by the manufacturer.

Expression and Purification of Tapasin. Sf21 cells were cultured at 27 °C in Sf-900 II serum free medium (Invitrogen). Infection of Sf21 cells with recombinant baculovirus was accomplished in 1 L spinner flasks at a cell density of $1-1.5 \times 10^6$ cells/mL followed by incubation at 27 °C for 72 h. The culture medium was collected by centrifugation (6000*g* and 4 °C for 25 min), and combined supernatants containing soluble tapasin were filtered through a 0.45 μm CA membrane filter unit (Corning) and immediately supplemented with 0.2 mM PMSF. Supernatants were concentrated approximately 6-fold using an S1Y30 spiral-wound membrane cartridge (Amicon).

The concentrated supernatant was dialyzed at 4 °C in 10 mM Tris (pH 8) and 200 mM NaCl with two changes of the dialysis buffer over a period of 48 h. The dialysis buffer was then changed to 10 mM Tris (pH 8), 200 mM NaCl, and 20 mM imidazole, and dialysis was continued for an additional 24 h. The dialyzed protein was directly applied at 4 °C onto a Ni−NTA affinity column (Qiagen) equilibrated with 10 mM Tris (pH 8), 200 mM NaCl, and 20 mM imidazole. The column was washed with 10 mM Tris (pH 8), 500 mM NaCl, and 25 mM imidazole followed by elution of tapasin with 10 mM Tris (pH 8), 500 mM NaCl, and 110 mM imidazole. The eluate containing tapasin was supplemented with 1 mM dithiothreitol (DTT) and concentrated in a stirred cell and Centriprep-10 apparatus (Amicon). The protein was further purified at 4 °C by gel filtration chromatography to remove protein aggregates using a Bio-Prep SE-100/17 (Bio-Rad) FPLC column in 20 mM Tris (pH 7.5) and 150 mM NaCl. The concentrations of tapasin stock solutions in 20 mM Hepes (pH 7.5), 150 mM NaCl, and 1 mM DTT were determined at 280 nm by Edelhoch's method (*23*) using an extinction coefficient of 67 380 M^{-1} cm^{-1} calculated on the basis of the content of Trp, Tyr, and cystine, in recombinant tapasin. The molar mass (MW) of recombinant tapasin was determined to be 44 182 Da by MALDI mass spectrometry in comparison to a value of 42 679 Da calculated from the amino acid sequence of recombinant tapasin. The glycosylated nature of the protein accounts for this discrepancy.

Circular Dichroism. Circular dichroism (CD) experiments were carried out using a Jasco-810 spectropolarimeter equipped with a thermoelectric temperature controller. Concentrations of stock tapasin solutions in 20 mM Hepes (pH 7.5), 150 mM NaCl, 1 mM DTT, and 10% glycerol were determined as described above. The far-UV CD spectrum represents the average of 15 scans recorded at 15 °C between 250 and 190 nm using a 0.2 mm path length cuvette with a protein concentration of 1.25 mg/mL. The thermal denaturation curve was obtained by monitoring the change in ellipticity at 216 nm between 15 and 60 °C using a scan rate of 40 °C/h and a 1 mm path length cuvette with a protein concentration of 0.25 mg/mL. Molar ellipticity values are expressed on a per residue basis.

Analytical Ultracentrifugation. Sedimentation velocity experiments were carried out at 20 °C on a Beckman Instruments Optima analytical ultracentrifuge as described previously (*20*, *21*). A stock solution of purified tapasin at approximately 1 mg/mL was dialyzed at 4 °C in 20 mM Hepes (pH 7.5), 150 mM NaCl, and 1 mM DTT for 16 h prior to analysis and was used to prepare a series of dilutions in the range of 0.1−1.0 mg/mL using the dialysate as a diluent. The value of the partial specific volume ($\nu_{seq} = 0.733$ cm^3/g) was calculated from the amino acid sequence of

recombinant tapasin using the consensus partial volumes of the amino acids (*24*) and taking into consideration the expected electrostriction of −0.0065 cm^3/g caused by the high net positive charge of ∼13/mol of tapasin (*25*, *26*). A ν_{carb} value of 0.625 cm^3/g was used for the partial specific volume of the carbohydrate (*27*). The degree of hydration ($\delta_1 = 0.379$ g of H_2O/g) was also calculated from the amino acid sequence according to the method of Kuntz and Kauzmann (*28*). Values of the sedimentation coefficient, $S^\circ_{20,w}$, were obtained using the program dcdt (*29*). The molar mass of the carbohydrate, M_{carb}, was calculated from the difference between the molar mass of tapasin obtained by MALDI mass spectrometry (44.2 kDa) and that calculated from its amino acid sequence (42.7 kDa). Hydrodynamic modeling was carried out using the equation of Perrin for a hydrodynamically equivalent, hydrated prolate ellipsoid of revolution (*30*). The partial specific volume for tapasin, ν_{tap}, is given by the following equation (*27*, *31*):

$$\nu_{tap} = \{1 - [M_{seq}(1 - \nu_{seq}\rho_o) + M_{carb}(1 - \nu_{carb}\rho_o)]/M_{mass\ spec}\}/\rho_o \quad (1)$$

where M_{seq} is the molar mass of the polypeptide chain, M_{carb} is the molar mass of the carbohydrate bound per mole of polypeptide chain, $M_{mass\ spec}$ is the molar mass of tapasin obtained by MALDI mass spectrometry, and ρ_o is the density of the solvent. The Stokes radius, R_o, of the equivalent hydrated sphere was calculated according to the following equation:

$$R_o = \left[\frac{3M_{mass\ spec}(\nu_{tap} + \delta_1\nu_1^o)}{4N\pi}\right]^{1/3} \quad (2)$$

where δ_1 is the hydration in grams of H_2O per gram, ν_1^o is the partial specific volume of pure water, and N is Avogadro's number. The Stokes radius, R_s, of tapasin was also determined from its elution position on a gel filtration column according to the method of Ackers (*32*).

Proteolysis. Stable proteolytic fragments of tapasin were obtained by incubating the protein (0.13 mg/mL) at 20 °C with chymotrypsin, elastase, and thermolysin at enzyme: substrate ratios (w/w) of 1:200, 1:220, and 1:125, respectively, in 20 mM Hepes (pH 7.5), 150 mM NaCl, and 0.1 mM DTT. Aliquots were taken at different times from the reaction mixtures, and proteolysis was terminated by the addition of the gel loading buffer and boiling samples to 100 °C followed by SDS−PAGE (15%) analysis. Samples corresponding to the 24 h digestion were electroblotted onto a polyvinylidene difluoride membrane (Millipore) and analyzed by N-terminal amino acid sequencing (10 cycles). For chymotrypsin experiments, these membranes were also probed with the R.gp48N antibody (gift of P. Cresswell) which recognizes the 2−20 N-terminal residues of tapasin (*2*) and with the penta-His antibody (Qiagen) which recognizes five or six His residues from the $(His)_6$ sequence at the C-terminus of tapasin. Samples corresponding to the 24 h digestion were purified on an analytical C4 (Vydac) reverse phase HPLC column, and peaks corresponding to stable fragments of tapasin were isolated and analyzed by MALDI mass spectrometry.

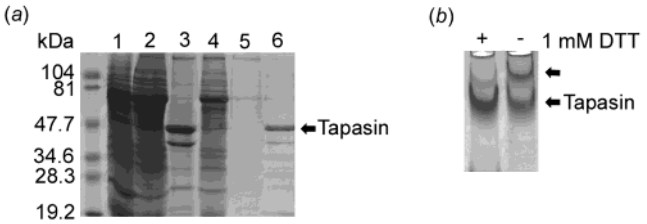

FIGURE 1: (a) SDS−PAGE gel (10%) showing purification of soluble $(His)_6$-tagged tapasin from Sf21 cells: crude supernatant 72 h postinfection (lane 1), flow-through from the Ni−NTA column (lane 2), Ni−NTA matrix showing bound tapasin (lane 3), washes from the Ni−NTA column using a buffer containing 25 mM imidazole (lanes 4 and 5), and tapasin after elution from the Ni−NTA column using a buffer containing 110 mM imidazole (lane 6). Molecular mass markers are indicated at the left of the gel. (b) Native PAGE gel (10%) showing formation of an upper band (indicated with an arrow) corresponding to dimers and oligomers of tapasin in the absence of added DTT in stock solutions. The symbols + and − refer to the presence and absence of 1 mM DTT, respectively.

RESULTS

Expression and Purification of Recombinant Tapasin. The lumenal domain (residues 1−392) of human tapasin tagged with an $(His)_6$ sequence at its C-terminus (tapasin) was optimally secreted from Sf21 cells 72 h postinfection with baculovirus. Purification of tapasin was carried out by Ni−NTA affinity column chromatography (Figure 1a) followed by gel filtration chromatography to remove protein aggregates. The band corresponding to tapasin occasionally appeared as a doublet on SDS−PAGE gels as also reported previously (*19*). The treatment of tapasin with PNGase eliminated this anomalous behavior (data not shown), suggesting that the glycosylated nature of the protein is responsible for the band splitting. Native PAGE analysis (Figure 1b) shows that purified tapasin migrates as a single, compact band, consistent with a well-behaved protein in solution. The addition of DTT in stock solutions of tapasin greatly prevented the formation of a high-mobility band (indicated with an arrow) which was identified as dimers and oligomers of tapasin by nonreducing SDS−PAGE analysis (data not shown).

Circular Dichroism. CD spectroscopy was used to characterize recombinant tapasin by recording a far-UV CD spectrum at 15 °C between 250 and 190 nm and a thermal denaturation curve between 15 and 60 °C (panels a and b of Figures 2, respectively). CD measurements were taken in 20 mM Hepes (pH 7.5), 150 mM NaCl, 1 mM DTT, and 10% glycerol. The addition of glycerol was necessary to prevent precipitation of tapasin upon thermal denaturation and had no apparent effects on the far-UV CD spectrum of the protein.

The far-UV CD spectrum (Figure 2a) of tapasin recorded at 15 °C shows a maximum at 195 nm and a minimum at 215 nm reflecting the presence of well-defined elements of secondary structure consisting predominantly of β-sheets. These results are in agreement with structural analysis predicting that some regions of tapasin are clearly related to the Ig superfamily of proteins (*3*, *15*, *17*, *18*). The thermal denaturation curve of tapasin recorded in the presence of 10% glycerol (Figure 2b) shows a major transition with a midpoint temperature of ∼45 °C. Since the thermal denaturation curve was only moderately reversible, no fit of the

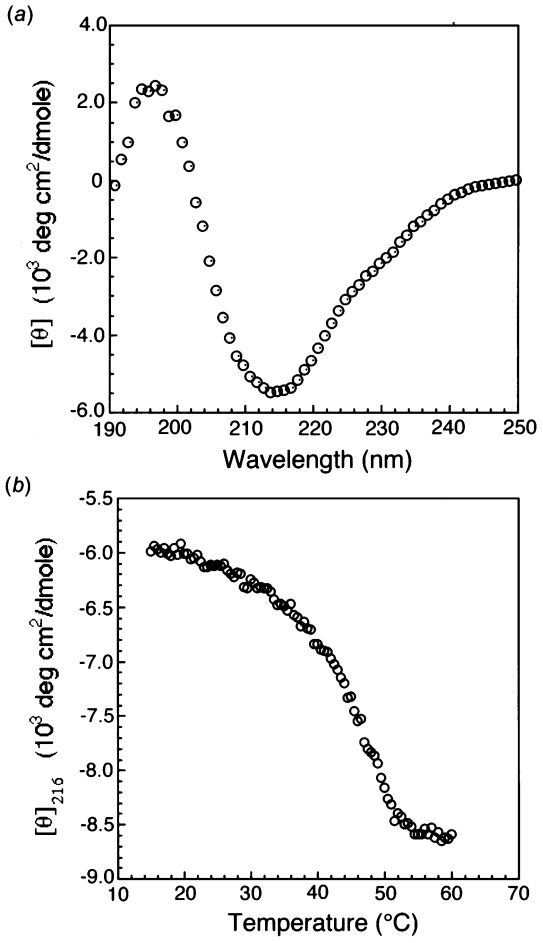

FIGURE 2: (a) CD spectrum (250−190 nm) of tapasin recorded at 15 °C. The tapasin concentration was 1.25 mg/mL in 20 mM Hepes (pH 7.5), 150 mM NaCl, 1 mM DTT, and 10% glycerol. A 0.2 mm path length cuvette was used to record the spectrum. (b) Thermal denaturation curve (15−60 °C) of tapasin obtained by monitoring the change in the CD signal at 216 nm. The tapasin concentration was 0.25 mg/mL in 20 mM Hepes (pH 7.5), 150 mM NaCl, 1 mM DTT, and 10% glycerol. A 1 mm path length cuvette was used to record the denaturation curve.

data points to equations describing equilibrium systems was generated. Collectively, this CD analysis suggests that recombinant tapasin possesses secondary structural elements and conformational properties that are consistent with a folded protein in solution.

Analytical Ultracentrifugation. Sedimentation velocity analyses (Table 1) in 20 mM Hepes (pH 7.5), 150 mM NaCl, and 1 mM DTT were carried out to determine the hydrodynamic properties of tapasin in solution. Results revealed the monodisperse nature of the protein and yielded a sedimentation coefficient, $S°_{20,w}$, of 2.68 S. A value for the Stokes radius, R_s, of 39.4 Å was calculated from the experimentally determined sedimentation coefficient and the molar mass obtained by MALDI mass spectrometry. This is consistent with the R_s value of 36.0 Å calculated on the basis of the elution position of tapasin from the gel filtration column (data not shown). Insights into the size and overall shape of tapasin were obtained from the frictional ratio, f/f_0, which was calculated from the experimentally determined sedimentation coefficient ($S°_{20,w}$ = 2.68 S) and the molar mass obtained by MALDI mass spectrometry (MW = 44.2 kDa). These calculations yielded an f/f_0 value of 1.47. Using values for partial specific volumes and hydration given

Table 1: Hydrodynamic Properties of Tapasin

molar mass (kDa)	
amino acid sequence[a]	42.7
MALDI mass spectrometry[b]	44.2
sedimentation coefficient ($S°_{20,w}$) (S)	2.68
Stokes radius (R_s) (Å)	
sedimentation and MALDI mass spectrometry	39.4
gel filtration	36.0
frictional ratio (f/f_0)	1.47
specific hydration (δ_1) (g of H_2O/g)	0.379
partial specific volume (ν_{tap}) (cm^3/g)[c]	0.729
axial ratio (a/b)	8.6
length 2a (nm)	22.5
diameter 2b (nm)	2.62

[a] Calculated from the amino acid sequence of recombinant tapasin. [b] The glycosylated nature of tapasin accounts for the difference between the value calculated from the amino acid sequence and that obtained by MALDI mass spectrometry. [c] Calculated from eq 1 in Materials and Methods.

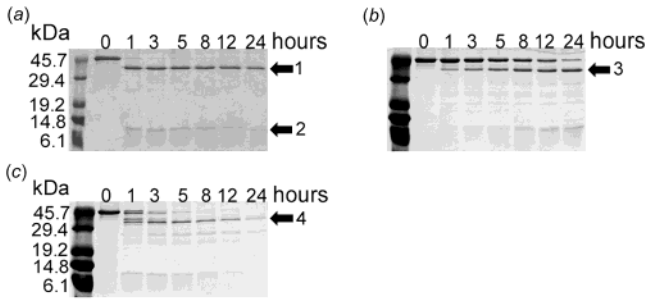

FIGURE 3: SDS−PAGE (15%) analysis of products from (a) chymotrypsin, (b) elastase, and (c) thermolysin digests carried out at 20 °C in 20 mM Hepes (pH 7.5), 150 mM NaCl, and 0.1 mM DTT, using enzyme:substrate ratios (w/w) of 1:200, 1:220, and 1:125, respectively. Aliquots were taken at different times from the reaction mixtures as indicated at the top of the figures. Bands labeled 1−4 were analyzed by N-terminal amino acid sequencing and MALDI mass spectrometry (see Table 2). Molecular mass markers are indicated at the left of panel a and c.

above, tapasin could be represented by an equivalent prolate ellipsoid of revolution with an axial ratio of 8.6, having a length of 22.5 nm and a diameter of 2.62 nm. These hydrodynamic dimensions suggest that monomeric tapasin is likely to be an elongated molecule in solution.

Susceptibility of Tapasin to Proteolysis. The use of proteolytic enzymes as conformational probes is a classical approach to gaining insights into the domain organization of proteins. This is based on the concept that solvent-exposed and flexible regions of proteins are more susceptible to cleavage than more buried or ordered regions. Controlled proteolysis of tapasin with chymotrypsin, elastase, and thermolysin was carried out at 20 °C, and aliquots from the mixtures were taken at different times over a period of 24 h and analyzed by SDS−PAGE (panels a−c of Figure 3, respectively). Results show that although the kinetics of cleavage and the pattern of fragmentation are different among the three enzymes, proteolysis of tapasin leads to two major fragments migrating with apparent MWs of ∼36 (bands 1, 3, and 4) and ∼12 kDa (band 2). These fragments remain stable in solution after 24 h as evidenced on SDS−PAGE gels by the absence of further cleavage, except in the case of thermolysin (Figure 3c), most probably due to the very broad substrate specificity of this enzyme. These proteolysis experiments were repeated in buffers containing EDTA, calcium ions, or zinc ions, and identical results were obtained

Table 2: Identification of Stable Proteolytic Fragments in Tapasin

band[a]	N-terminus cleavage site[b]	MW (kDa)[c]	residues in tapasin[d]
1	Ala86	35.1	86−398
2	Gly1	9.15	1−85
3	Asn93	33.9	93−395
4	Leu89	33.5	89−390

[a] Bands 1−4 in Figure 3. [b] Determined by N-terminal amino acid sequencing (10 cycles). [c] Determined by MALDI mass spectrometry. [d] Amino acid residues are numbered according to the sequence of soluble $(His)_6$-tagged tapasin.

(data not shown), suggesting the apparent absence of binding sites for these ions in tapasin or, if bound to the protein, these metal ions are unlikely to serve a structural role.

Identification of Stable Proteolytic Fragments in Tapasin. The identity of the four proteolytic fragments (bands 1−4 in Figure 3) was unambiguously established by N-terminal amino acid sequencing and MALDI mass spectrometry (Table 2). Characterization of the larger fragments corresponding to bands 1, 3, and 4 reveals C-terminal fragments of 313 residues (residues 86−398), 303 residues (residues 93−395), and 302 residues (residues 89−390), respectively. Similarly, characterization of the smaller chymotryptic fragment corresponding to band 2 reveals that it encompasses the first 85 N-terminal residues of the protein, making its sequence (residues 1−85) entirely continuous with that of the larger chymotryptic fragment (residues 86−398) within the aligned sequence of tapasin. Given the different substrate specificities of the three enzymes, it is particularly noteworthy that the major cleavage sites are restricted to a short region of tapasin that extends from residue ~85 to 93. Collectively, results from proteolysis studies (Figure 3 and Table 2) indicate that the region of tapasin near residue 90 is particularly accessible within the spatial organization of the protein.

The locations of the stable fragments listed in Table 2 have been mapped within the primary sequence of tapasin (Figure 4a). This analysis emphasizes that the structure of tapasin can be divided into two core regions: a small N-terminal fragment comprising the first 85 residues (band 2) of the protein and a larger C-terminal fragment spanning ~300 residues (bands 1, 3, and 4). Samples corresponding to the 24 h digestion by chymotrypsin were further analyzed after electrotransfer onto blotting membranes and probing with the anti-N-terminal R.gp48N (Figure 4b) and anti-C-terminal penta-His (Figure 4c) antibodies. Results show that tapasin (lanes 1) is efficiently recognized by both the R.gp48N and penta-His antibodies, whereas the small (N-terminal) and large (C-terminal) chymotrypsin fragments (lanes 2) are revealed selectively by the R.gp48N and penta-His antibodies, respectively. These immunoblots are consistent with the identity of the two fragments as determined in Table 2 and mapped in Figure 4a (bands 1 and 2).

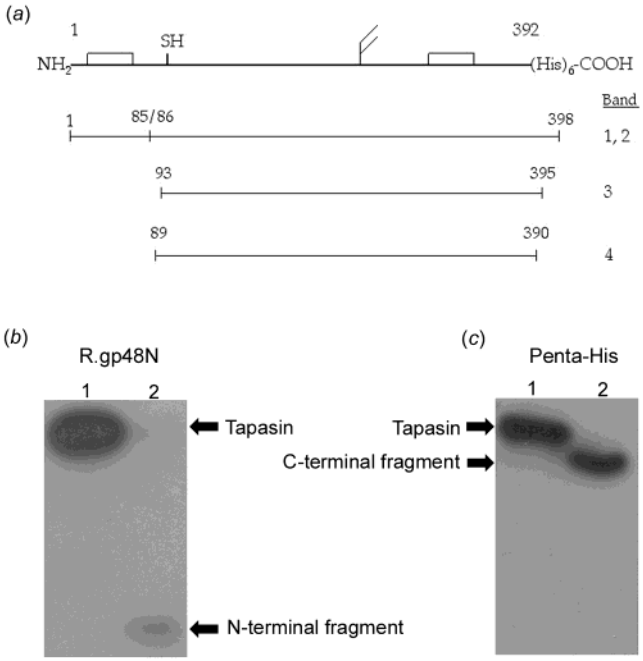

FIGURE 4: (a) Location of the stable fragments (bands 1−4 in Figure 3) in the primary sequence of tapasin. The single free Cys residue (*13*), two disulfide bonds (*13*, *15*), a single oligosaccharide, and the $(His)_6$ sequence are indicated. (b) R.gp48N and (c) penta-His antibodies were used to probe tapasin (lane 1) and the two chymotrypsin fragments (lane 2) after electrotransfer onto blotting membranes.

DISCUSSION

The studies described above have been aimed at characterizing the structural properties of tapasin in an effort to establish relationships with its function. Although secondary and tertiary structural predictions of tapasin have been made (*17*, *18*), no direct experimental analysis has so far been carried out to gain insights into its structure.

Limited proteolysis experiments using enzymes with different specificities have revealed that tapasin is cleaved into two major fragments of 9 and 34 kDa that appear to be resistant to further cleavage, consistent with rather compact structures. The small 9 kDa fragment was mapped within the first 85 N-terminal residues of tapasin, whereas the large 34 kDa fragment comprises the last 300 C-terminal residues of the protein. On the basis of these results, the domain organization of tapasin can be described as an assembly of two core regions of different sizes loosely connected by a linker or loop comprising residues ~85−93. It remains to be investigated whether these major regions are further divided into subdomains which is a strong possibility for the large C-terminal fragment. Sedimentation analysis has revealed that the molecular shape of tapasin in solution is elongated (Table 1). The elongated structure of tapasin suggests that the protein is likely to be flexible in solution which could be an important property for exerting its function within the class I assembly complex.

Studies using deletion mutants of tapasin have shown that the first 50 N-terminal residues are key elements in conferring tapasin with the ability to maintain stable interactions between the peptide-receptive form of class I MHC−β_2m heterodimers, calreticulin, and ERp57 (*8*). On the basis of our results, the 50 N-terminal residues of tapasin are located in the smaller structural domain. It has been shown from in vitro experiments that the peptide binding site (α1/α2 domains) is the most destabilized region of the peptide-receptive form of class I MHC−β_2m heterodimers (*22*). More specifically, a solvent-exposed loop formed by residues 128−136 (α2 domain) is likely to be an element contributing to this destabilization. It has been suggested that this particular loop forms a binding site for one or more of the class I

assembly proteins (*11*, *33*−*37*) and undergoes a conformational change upon peptide binding (*35*, *38*). A functional role for the N-terminal domain of tapasin could therefore be in relation to its ability to stabilize the conformationally immature form of the peptide binding site by interacting with this loop. More recently, disulfide bond isomerization in the α2 domain of the class I MHC heavy chain has been shown to play a role in the assembly of class I MHC−β_2m−peptide complexes, and specific Cys residues in both tapasin and ERp57 have been implicated in this process (*13*). The free Cys95 in tapasin has been identified as a crucial residue, and on the basis of our results, its location within the flexible linker region is consistent with this residue being more readily available to participate in such a dynamic process. Both Cys7 and Cys71, which form a disulfide bond in the N-terminal domain of tapasin (*13*), have also been implicated in the process of disulfide isomerization despite their oxidized state within the protein. In view of these results, an alternative functional role for the N-terminal domain of tapasin could be in relation to its ability to alter the conformational maturity of the peptide binding site through isomerization of the conserved disulfide bond in the α2 domain. This process has been suggested as a plausible mechanism with which tapasin edits the peptide repertoire (*13*).

Although no X-ray crystallographic analysis of tapasin has so far been reported, possibly due to the difficulty in growing crystals of a flexible protein, a hypothetical three-dimensional molecular model of human tapasin has recently been proposed from amino acid sequence comparisons, the location of the gene in the MHC, exon−intron organization, and secondary structure analysis (*17*). In this model, the polypeptide chain of tapasin between residues 105 and 381 is folded into a class I MHC heavy chain-like conformation. Interestingly, no structural homologues could be identified for the first 100 N-terminal residues of the protein. This predicted structural organization of tapasin into two distinct regions is strikingly consistent with the boundaries that delineate the two proteolytic fragments identified in our studies (Table 2 and Figure 4a). The possibility that the first ∼100 N-terminal residues of tapasin form a distinct domain, adopting an as yet unidentified fold, that is connected via a linker or loop region to a larger C-terminal domain predicted to resemble a class I MHC heavy chain would make the structure of this protein unique among nonclassical class I MHC proteins (*39*−*42*). To date, no experimental evidence have been provided to suggest that tapasin binds peptides or associates with β2m, although one study raised the possibility of tapasin−β_2m interactions (*8*). Given these structural predictions, we compared the molecular dimensions of tapasin (Table 1) with those of the class I MHC protein HLA-A*0201 (frictional ratio $f/f_0 = 1.18$, axial ratio $a/b = 4.1$, length $2a = 13.9$ nm, and diameter $2b = 3.4$ nm) as determined by sedimentation analysis under identical conditions. A comparison of these values with those presented in Table 1 shows that the two proteins have different molecular dimensions, consistent with the presence of an additional N-terminal domain in tapasin that contributes to the overall dimensions of the protein. A more meaningful comparison of molecular dimensions between tapasin and class I MHC proteins awaits the cloning and expression of the C-terminal fragment (see below). Additional studies are needed to better understand the immune function and three-dimensional structure of tapasin as well as its evolutionary relationship with class I MHC proteins.

Collectively, these studies have identified distinct structural features of tapasin and have allowed us to discuss the possible functional implications associated with these characteristics. The cloning and bacterial expression of the two proteolytic fragments are currently underway (M. Chen and M. Bouvier, unpublished results), and results should reveal unambiguously whether they correspond to independent folding domains in solution. These isolated fragments will also permit us to investigate whether they represent functionally distinct regions of the protein. Finally, these fragments may also be amenable to X-ray and NMR structural analysis

ACKNOWLEDGMENT

This paper is dedicated to the memory of Professor Don C. Wiley. We thank Kimberly M. Ray for technical assistance with the expression of tapasin, Bill Lane and his staff at the Harvard Microchemistry Facility for MALDI mass spectrometry analysis, and Dr. Peter Cresswell for the gift of the R.gp48N antibody.

REFERENCES

1. Diedrich, G., Bangia, N., Pan, M., and Cresswell, P. (2001) *J. Immunol. 166*, 1703−1709.
2. Sadasivan, B., Lehner, P. J., Ortmann, B., Spies, T., and Cresswell, P. (1996) *Immunity 5*, 103−114.
3. Ortmann, B., Copeman, J., Lehner, P. J., Sadasivan, B., Herberg, J. A., Grandea, A. G., Riddell, S. R., Tampe, R., Spies, T., Trowsdale, J., and Cresswell, P. (1997) *Science 277*, 1306−1309.
4. Grandea, A. G., III, Lehner, P. J., Cresswell, P., and Spies, T. (1997) *Immunogenetics 46*, 477−483.
5. Schoenhals, G. J., Krishna, R. M., Grandea, A. G., III, Spies, T., Peterson, P. A., Yang, Y., and Fruh, K. (1999) *EMBO J. 3*, 743−753.
6. Grandea, A. G., III, Golovina, T. N., Hamilton, S. E., Sriram, V., Spies, T., Brutkiewicz, R. R., Harty, J. T., Eisenlohr, L. C., and Van Kaer, L. (2000) *Immunity 13*, 213−221.
7. Lehner, P. J., Surman, M. J., and Cresswell, P. (1998) *Immunity 8*, 221−231.
8. Bangia, N., Lehner, P. J., Hughes, E. A., Surman, M., and Cresswell, P. (1999) *Eur. J. Immunol. 29*, 1858−1870.
9. Li, S., Sjogren, H.-O., Hellman, U., Pettersson, R. F., and Wang, P. (1997) *Proc. Natl. Acad. Sci. U. S. A. 94*, 8708−8713.
10. Sijts, A. J., and Pamer, E. G. (1997) *J. Exp. Med. 185*, 1403−1411.
11. Lewis, J. W., and Elliott, T. (1998) *Curr. Biol. 8*, 717−720.
12. Williams, A. P., Peh, C. A., Purcell, A. W., McCluskey, J., and Elliott, T. (2002) *Immunity 16*, 509−520.
13. Dick, T. P., Bangia, N., Peaper, D. R., and Cresswell, P. (2002) *Immunity 16*, 87−98.
14. Grandeas, A. G., Comber, P. G., Wenderfer, S. E., Schoenhals, G., Fruh, K., Monaco, J. J., and Spies, T. (1998) *Immunogenetics 48*, 260−265.
15. Herberg, J. A., Sgouros, J., Jones, T., Copeman, J., Humphray, S. J., Sheer, D., Cresswell, P., Beck, S., and Trowsdale, J. (1998) *Eur. J. Immunol. 28*, 459−467.
16. Campbell, R. D., and Trowsdale, J. (1997) *Immunol. Today 14*, 349−352.
17. Mayer, W. E., and Klein, J. (2001) *Immunogenetics 53*, 719−723.
18. Frangoulis, B., Park, I., Guillemot, F., Severac, V., Auffray, C., and Zoorob, R. (1999) *Immunogenetics 49*, 328−337.
19. Turnquist, H. R., Vargas, S. E., Reber, A. J., McIlhaney, M. M., Li, S., Wang, P., Sanderson, S. D., Gubler, B., van Endert, P., and Solheim, J. C. (2001) *J. Immunol. 167*, 4443−4449.
20. Li, Z., Stafford, W. F., and Bouvier, M. (2001) *Biochemistry 40*, 11193−11201.
21. Bouvier, M., and Stafford, W. (2000) *Biochemistry 39*, 14950−14959.

22. Bouvier, M., and Wiley, D. C. (1998) *Nat. Struct. Biol. 5*, 377−383.
23. Edelhoch, H. (1967) *Biochemistry 6*, 1948−1954.
24. Perkins, S. J. (1986) *Eur. J. Biochem. 157*, 169−180.
25. Cohn, E. J., and Edsall, J. T. (1943) in *Proteins, Amino Acids and Peptides*, pp 370−381, Reinhold Publishing Corp., New York.
26. Laue, T. M. (1992) in *Analytical Ultracentrifugation in Biochemistry and Polymer Science* (Harding, S. E., Rowe, A. J., and Horton, J. C., Eds.) pp 63−89, Royal Society of Chemistry, Cambridge, U.K.
27. Shire, S. J. (1992) Technical Note DS-837, Beckman Instruments, Palo Alto, CA.
28. Kuntz, I. D., Jr., and Kauzmann, W. (1974) *Adv. Protein Chem. 28*, 239−345.
29. Stafford, W. F. (1992) *Anal. Biochem. 203*, 295−301.
30. Stafford, W. F., and Schuster, T. M. (1995) in *Hydrodynamic Transport Methods. Introduction to Physical Methods for Protein and Nucleic Acid Research* (Glaser, J. A., and Deutscher, M. P., Eds.) Academic Press Inc., Orlando, FL.
31. Lewis, M. S., and Junghans, R. P. (2000) *Methods Enzymol. 321*, 136−149.
32. Ackers, G. K. (1967) *J. Biol. Chem. 242*, 3237−3238.
33. Peace-Brewer, A. L., Tussey, L. G., Matsui, M., Li, G., Quinn, D. G., and Frelinger, J. A. (1996) *Immunity 4*, 505−514.
34. Lewis, J. W., Neisig, A., Neefjes, J., and Elliott, T. (1996) *Curr. Biol. 6*, 873−883.
35. Yu, Y. Y. L., Turnquist, H. R., Myers, N. B., Balendiran, G. K., Hansen, T. H., and Solheim, J. C. (1999) *J. Immunol. 163*, 4427−4433.
36. Harris, M. R., Lybarger, L., Myers, N. B., Hilbert, C., Solheim, J. C., Hansen, T. H., and Yu, Y. Y. L. (2001) *Int. Immunol. 13*, 1275−1282.
37. Paquet, M.-E., and Williams, D. B. (2002) *Int. Immunol. 14*, 347−358.
38. Elliott, T. (1997) *Immunol. Today 18*, 375−379.
39. Zeng, Z., Castano, A. R., Segelke, B. W., Stura, E. A., Peterson, P. A., and Wilson, I. A. (1997) *Science 277*, 339−345.
40. Li, P., Willie, S. T., Bauer, S., Morris, D. L., Spies, T., and Strong, R. K. (1999) *Immunity 10*, 577−584.
41. Burmeister, W. P., Gastinel, L. N., Simister, N. E., Blum, M. L., and Bjorkman, P. J. (1994) *Nature 372*, 336−343.
42. Sanchez, L. M., Chirino, A. J., and Bjorkman, P. J. (1999) *Science 283*, 1914−1919.

BI020521U