Review

Function of the transport complex TAP in cellular immune recognition

Rupert Abele, Robert Tampé *

Institut für Physiologische Chemie, Philipps-Universität Marburg, Karl-von-Frisch-Str. 1, 35033 Marburg, Germany

Received 31 August 1999; accepted 1 September 1999

Abstract

The transporter associated with antigen processing (TAP) is essential for peptide loading onto major histocompatibility complex (MHC) class I molecules by translocating peptides into the endoplasmic reticulum. The MHC-encoded ABC transporter works in concert with the proteasome and MHC class I molecules for the antigen presentation on the cell surface for T cell recognition. TAP forms a heterodimer where each subunit consists of a hydrophilic nucleotide binding domain and a hydrophobic transmembrane domain. The transport mechanism is a multistep process composed of an ATP-independent peptide association step which induces a structural reorganization of the transport complex that may trigger the ATP-driven transport of the peptide into the endoplasmic reticulum lumen. By using combinatorial peptide libraries, the substrate selectivity and the recognition principle of TAP have been elucidated. TAP maximizes the degree of substrate diversity in combination with high substrate affinity. This ABC transporter is also unique as it is closely associated with chaperone-like proteins involved in bonding of the substrate onto MHC molecules. Most interestingly, virus-infected and malignant cells have developed strategies to escape immune surveillance by affecting TAP expression or function. © 1999 Elsevier Science B.V. All rights reserved.

Keywords: ABC transporter; Transporter associated with antigen processing; Antigen presentation; Transport mechanism; Virus persistence

Contents

1. Introduction – overview of the MHC class I antigen processing pathway ............... 406
2. Structural organization of the TAP complex .................................................. 408
3. Multistep transport mechanism of TAP .................................................. 412
4. The substrate binding motif of TAP – coevolution of affinity, specificity and diversity .... 412
5. TAP as part of a macromolecular transport and chaperone complex .................. 415
6. Implication in human diseases ........................................................................... 415

References ........................................................................................................... 417

* Corresponding author. Fax: +49-6421-286-4335; E-mail: tampe@mail.uni-marburg.de

0005-2736/99/$ – see front matter © 1999 Elsevier Science B.V. All rights reserved.
PII: S0005-2736(99)00171-6
1. Introduction – overview of the MHC class I antigen processing pathway

Under normal conditions class I molecules of the major histocompatibility complex (MHC) present peptides derived from endogenous proteins on the cell surface of every nucleated cell (Fig. 1). During viral infection or malignant transformation, an additional set of peptides bound to MHC class I molecules is delivered to the cell surface for presentation to cytotoxic T lymphocytes (CTL). The recognition of MHC class I molecules (‘self-component’) loaded with peptides derived from ‘non-self’ proteins by CTL via the T cell receptors eventually leads to the lysis/apoptosis of abnormal cells (for review see [1–4]).

The pathway of antigen presentation on the cell surface of pathogens dwelling in the cytosol comprises proteins from different compartments. The pathogens together with their intrinsic proteins are degraded by the major cytosolic proteolysis machinery, the 20S/26S proteasome (for review see [5–8]). After cleavage, the peptides are translocated by the transporter associated with antigen processing (TAP) into the endoplasmic reticulum (ER) where assembly of MHC class I molecules and peptides occurs. Stable ternary complexes consisting of heavy chain, β2-microglobulin (β2-m) and bound peptide can leave the ER for surface presentation to T cell receptor.

Every human has three to six different MHC class I alleles for the presentation of antigens on the cell surface. This low number of different MHC class I molecules has to bind peptides of every non-self protein for presentation to CTL and protection of the individuals against pathogens. As revealed by X-ray crystallography, the antigenic peptides bind in a groove formed by two α-helices on the rim and eight β-strands on the bottom containing residues of the α1 and α2 domains of the heavy chain of MHC class I molecules [9–11]. The peptides are fixed via the free amino- and carboxy-termini. In addition, anchor residues at position two or three and at the carboxy-terminal residue pointing into the groove are important. In between these anchor residues, the side chains point outside the groove which explains the large pool of peptides that can be presented by one MHC class I allele. Interestingly, the T-cell receptor monitors this variable region [12,13]. The groove is closed on both ends, therefore the size of the peptides is determined to 8–10 residues in length.

The assembly of MHC class I molecules has been well studied (for review see [14–16]). However, less is known about how the peptides are generated and transported into the ER [4,17]. The contribution of the 20S proteasome for the generation of antigenic peptides came from the observation that the two β-type proteasomal subunits, LMP2 and LMP7 (low molecular weight peptides), were identified within the MHC locus [18–20]. The proteasome complex is a multicatalytic macromolecular protease (700 kDa) consisting of 28 subunits arranged in two outer rings of α-type and two inner rings of β-type subunits (for review see [5,7]). LMP2 and LMP7 are upregulated by interferon-γ (INF-γ), which causes a structural rearrangement of the 20S proteasome in which the catalytically active subunits β2, β1 and β5 are replaced by MECL1 (multicatalytic endopeptidase complex-like), LMP2 and LMP7, respectively, and the so-called immunoproteasomes are formed. Contradictory results exist about the functional difference and cleavage specificity of the immunoproteasome compared to the proteasome consisting of constitutively expressed β-type subunits. Due to the substitution of the active subunits, the enzymatic activity of the immunoproteasome cleaving after hydrophobic and basic substrates is increased whereas the peptidyl-glutamyl activity is decreased in comparison to the housekeeping proteasome [21–24]. This alteration may fit with the preference of MHC class I molecules for peptides with hydrophobic and basic carboxy-terminal peptide residues. The peptides produced by the proteasome have a size distribution of approximately 3–30 residues with a maximum of 6–11 residues [21,25–27] which are also in part the size of the antigenic peptides bound in the groove of the binding pocket of MHC class I molecules. In addition to the proteasome, other proteases such as an IFN-γ-inducible leucine aminopeptidase [28] or a giant cytosolic protease system in the cytosol [29,30] may play a role in epitope generation. Vinitsky and coworkers showed that the presentation of antigenic peptides of influenza viral proteins is not influenced by blocking proteasomal activity [26]. In addition, epitopes may also be produced by proteolytic trimming in the ER after TAP-dependent transport [31].

The important role for peptide transport into the
ER lumen became clear from studies of various cell lines with a strongly reduced level of MHC class I molecules on the cell surface [32,33]. Although the expression levels of MHC class I heavy chain and β2-m are normal and exogenously added peptides or peptides introduced in the ER by a signal sequence were efficiently presented, these defective cell lines were unable to present intracellular antigens on the cell surface. The defective phenotypes could be restored by transfection of tap1 and/or tap2 [34,35]. The genes for human TAP1 and TAP2 are located in the MHC II locus of chromosome 6 and comprise 10 kb each. The genes are encoded by 11 exons. Eight exons have the same size and all exon/intron boundaries are identical [36]. All four genes are organized in an array lmp2, tap1, lmp7, tap2, where only lmp2 is encoded on the (−) strand, suggesting a duplication of an ancestral gene for a pro-

Fig. 1. Antigen processing and presentation via MHC class I molecules. Endogenous proteins including viral or tumor-specific proteins are degraded in the ubiquitin-proteasome pathway, and peptides are transported into the ER lumen by TAP. There, several molecules have been implicated in the tightly regulated folding, assembly and loading of MHC class I molecules, including calnexin, calreticulin, tapasin, ERp57, and possibly a peptide trimming activity. Stable MHC-peptide complexes can leave the ER via the Golgi compartment to the cell surface for recognition by cytotoxic T lymphocytes.
teasomal β type and TAP subunit followed by an inversion of lmp2. Transcription of all four genes is induced by IFN-γ, indicating a common regulation and concerted function of these genes in antigen processing.

By using isolated microsomes or semi-permeabilized cells, ATP- and TAP-dependent peptide translocation into the ER was demonstrated [37–39]. In addition, expression of TAP in insect cells and yeast revealed that TAP is functional in the absence of factors of the adaptive immune system [40,41]. In addition to this major TAP-dependent translocation mechanism, alternative, TAP-independent pathways may exist for processing and loading of peptides onto MHC class I molecules [42].

The TAP sequences of human, gorilla (Gorilla gorilla), mouse (Mus musculus/castaneus), rat (Rattus norvegicus), hamster (Mesocricetus auratus), and salmon (Salmo salar) show phylogenetic differences as expected. Human TAP1 possesses 98.8% homology with gorilla TAP1, 69.2% with hamster TAP1 and only 40% with salmon TAP1. The homology between TAP1 and TAP2 in all species is approximately 35% although they have a similar predicted membrane topology. Thus, TAP1 and TAP2 seem to be derived from an ancestral gene by gene duplication that happened before the evolution of the adaptive immune system found in vertebrates.

The tap genes of all species examined are polymorphic, but only the rat alleles [43] as well as the most recently discovered human TAP2iso splice variant [44] possess different substrate specificity. The cim (class I modifier) polymorphism in rat is based on four tap2 alleles grouped into cima and cimb [43]. Cimb comprises the b and u alleles with a specificity for hydrophobic, carboxy-terminal residues of the processing peptide [45,46]. In contrast, the haplotypes of cima (a and l) show a much broader specificity and can transport peptides with basic or hydrophobic residues at the carboxy-terminus. The cima and cimb alleles differ in 25 residues whereas only two of them are located in the nucleotide binding domain of TAP2. The remaining divergent residues are in the amino-terminal half of the subunits. For human or murine TAP approximately 10 polymorphic sites have been found for each subunit. However, these alleles do not show any difference in substrate specificity [47,48].

2. Structural organization of the TAP complex

TAP1 and TAP2 belong to the superfAMILY of ATP-binding cassette (ABC) transporter that comprises a large number of polytopic integral membrane proteins transporting a diverse set of molecules across membranes in an ATP-dependent manner [49,50]. ABC transporters are found in all three domains of life representing in some organisms the largest family of paralogous proteins. All ABC transporters possess two conserved cytoplasmic ATP-binding domains and two hydrophobic domains comprising 5–10 transmembrane stretches, possibly α-helices, lining up the putative translocation pore. Although the structural organization of the ABC transporters can be assumed to be very similar, the hydrophobic transmembrane domains show only very weak sequence homology. TAP1 or TAP2 are so-called half-size transporters composed of a hydrophobic transmembrane domain (TMD) followed by a highly conserved nucleotide binding domain (NBD). There is 60% sequence identity within the NBDs and only 30% between the TMDs of TAP.

Human TAP1 and TAP2 have a length of 748 amino acids (81 kDa) and 686 amino acids (75 kDa), respectively. The molecular weight determined experimentally by SDS-polyacrylamide gel electrophoresis is 71 kDa for TAP1 and 75 kDa for TAP2. Both proteins are located in the ER and cis-Golgi [51] retarded by a so far unknown cryptic ER retention signal. For human TAP2, an allele that is 17 amino acids longer (703 aa) has been identified resulting from a polymorphic single base pair substitution in the stop codon of the shorter TAP2 allele [52]. Recently, a splice variant of hTAP2 with a length of only 653 amino acids was detected [44]. As also found for other ABC transporters synthesized at the ER, both TAP proteins lack an N-terminal signal sequence for the import in the ER, suggesting that an internal signal sequence may exist promoting insertion in the ER membrane. TAP1 and TAP2 are found to be non-glycosylated although human TAP1 has three consensus glycosylation sites, two facing the cytosol and one placed in a short ER loop which are too short for effective glycosylation by the oligosaccharide transferase [40]. A very minor subpopulation of hTAP was found to be N-glycosylated [53] which may reflect misfolded protein.
By immuno-coprecipitation it was shown that TAP1 and TAP2 assemble in the ER membrane to form a heteromeric complex [54,55]. Coexpression of TAP1 and TAP2 in TAP-deficient cell lines demonstrated that both subunits are essential for antigen processing [34,35,56]. Furthermore, no additional factors of the immune system are required for TAP function [40,41]. Gel filtration analysis together with more recent crosslinking experiments suggested that TAP1 and TAP2 form a functional heterodimer within the ER membrane with a stoichiometry of 1:1 [40,57,58].

As mentioned, the homology between the hydrophobic domains of ABC transporters is not very high and so there is a variability in the number of transmembrane domains. Six transmembrane helices are predicted for hemolysin transporter HlyB [59,60]. The subunits MalG and MalF of the maltose transporter are supposed to contain six and eight transmembrane helices, respectively [61,62]. Some eukaryotic ABC transporters show six transmembrane helices of each hydrophobic domain with the amino- and carboxy-terminus in the cytosol. Such a 2×6 transmembrane helix model could also be confirmed for MDR1 [63,64]. As predicted from hydrophobicity analysis, the membrane topology of human TAP1 and TAP2 is supposed to be different (Fig. 2) [65]. Sequence alignments with other ABC transporters exclude the first 175 residues for hTAP1 and 140 residues for hTAP2. This part of TAP (N-domain) displays no sequence homology with any protein. In addition, the N-domain is very hydrophobic comprising four and three predicted transmembrane helices for TAP1 and TAP2, respectively. Although no function of this amino-terminal hydrophobic region is known so far, it can be speculated that it is essential for ER membrane targeting and correct assembly of the complex. The rest of the transmembrane domain shows weak, but significant sequence homology to MDR1 and other ABC transporters of the P-glycoprotein/TAP subfamily. The sequence similarity is increased from TM1 to TM6. Based on the hydrophobicity analysis and sequence alignments with MDR1 for which the membrane topology has been extensively studied, a 2×6 transmembrane helix model of TAP can be derived extended by an additional four and three transmembrane helices predicted for the N-domain [65]. Interestingly, the TAP complex is highly asymmetric and only a very small portion (<10%) is in the ER lumen, whereas large cytosolic loops and both NBDs are in the cytosol.

In a first experimental approach, 10 transmembrane helices were reported for hTAP1 expressed in *Escherichia coli* [66]. The topology was studied using carboxy-terminal truncations of hTAP1 fused to β-lactamase as reporter. This approach has been quite successful in addressing the topology of several membrane proteins of prokaryotes. However, care must be taken if multisubunit ABC transporters from highly specialized eukaryotic intracellular compartments are studied. The topology model of hTAP1 found in this study contradicts the orientation of the loop between TM5 and TM6 that has been identified by peptide photocrosslinking experiments to be involved in peptide binding from the cytosol [67,68]. Differences in membrane topology were also reported for P-glycoprotein expressed in *E. coli* [64] in comparison to the expression in eukaryotic cells [69]. Notably, TAP1 expressed in *E. coli* is not functional as large carboxy-terminal regions and the entire TAP2 protein are missing, both are likely to be essential for correct membrane insertion. It has been further shown that even the insertion of an epitope, a glycosylation targeting sequence or a protease cleavage site into transmembrane loops of ABC transporters resulted in a loss of protein function and possibly incorrect folding into the membrane. Thus, alternative methods have to be applied to address the topology of the TAP complex. Following the elegant approach reported for lactose permease [70] and MDR1 [69,71], it will be interesting to construct cysteine-less mutants of TAP1 and TAP2 that are functional. Subsequently, single cysteines can be introduced in predicted loops and their accessibility can be probed by thiol-specific reagents. Mutation of all 19 cysteines of TAP is very time-consuming, but these cysteine mutations are supposed to be minimally invasive to the structure and function of TAP.

In TAP, the hydrophobic transmembrane domains are linked to the nucleotide binding domains which harbor the highly conserved ATP-binding cassette consisting of the Walker A and B motifs (A/B) for ATP binding and hydrolysis [72]. All ABC transporters have a so-called C-loop (C) that comprises six to eight conserved amino acids which are located sev-
eral residues upstream of the Walker B sequence. The ‘EAA’-like motif (E) found in the last cytosolic loop of the membrane spanning domain of bacterial ABC transporters [73] is less significant for TAP. There is genetic evidence that this ‘EAA’ sequence interacts via the C-loop with the nucleotide binding domain and couples ATP-hydrolysis to the transmembrane spanning domain [74,75].

Peptides are photocrosslinked to TAP1 and TAP2 suggesting that both subunits contribute to peptide binding [76,77]. In more recent studies, the peptide binding site was mapped to regions of human TAP1 and TAP2. The crosslinked subunits were digested by trypsin and/or bromocyan and crosslinked fragments were immunoprecipitated with antibodies against epitopes of putative loops in TMD [67,68]. The proteolytic analysis of the crosslinked TAP subunits revealed a similar binding region for TAP1 and TAP2. The binding region comprises the cytosolic loop between putative TM4 and TM5 and a carboxy-terminal stretch of approximately 15 amino acids following TM6. Due to the topological model (Fig. 2), all these binding regions are exposed to the cytosol. Regarding the transmembrane helices, TM4 to TM6 seem to be involved in peptide binding. For TAP2, the ER-exposed loop between predicted TM1 and

\[\text{Fig. 2. Membrane topology of the human TAP complex. The membrane topology was predicted by sequence alignment with human MDR1 and using TopPredII (M.-G. Claros, Paris, France, based on the algorithm of Gunnar von Heijne). For the hydrophobicity plot the following parameters were used (Kye-Doolittle; full window: 21; core window: 11). From the hydrophobicity plot four and three transmembrane helices can be predicted for the hydrophobic N-terminal domain (N) of TAP1 and TAP2, respectively. The communication between the NBD and TMD leads to tight coupling of peptide binding and ATP hydrolysis probably via contact of the C-loop (C) and the ‘EAA’-like region (E). The Walker A (A) and B (B) regions fix the ATP at the nucleotide binding pocket.}\]
TM2 is also implicated in peptide binding [67]. The identified photorecrosslinked regions also comprise the polymorphic residues 374 and 380 of rat TAP2 which alter the substrate specificity in the rat TAP2a and TAP2a alleles [78,79]. In future, a more detailed analysis of the substrate binding site may also become possible using single cysteine mutants as described above.

To obtain information about which subunit of the TAP complex is involved in substrate specificity, Armandola and coworkers coexpressed TAP subunits from different species and tested their substrate specificity using peptides with different carboxy-terminal residues [80]. By using chimeric rTAP2a or rTAP2u and hTAP2, residues located between residues 1 and 361 of hTAP2 were identified to affect the specificity for peptides with various carboxy-terminal residues. Moreover, two pairs of residues (217/218) and (374/380) of rat TAP and a single point mutation (A374D) of human TAP affect the substrate specificity [78,80]. Taking these data together, transport specificity with regard to the carboxy-terminal residue of the peptide substrate is mainly affected by TAP2, but TAP1 can also contribute. The influence of TAP2 on peptide binding seems to be more complex because TAP2iso resembling a truncated splice variant of hTAP2 showed an altered peptide specificity [44]. Thus, in addition of regions described above, the last 33–51 carboxy-terminal amino acids of TAP2 seem also to contribute to peptide specificity. How this carboxy-terminal region of TAP2 affects the peptide selection remains an open and puzzling question.

Similar to P-glycoprotein [81,82], the region of TM5 and TM6 of both subunits seems to build the substrate binding site (Fig. 2). Less is known about the peptide translocation pathway, but it can be speculated that TM5 and TM6 are also part of a translocation channel. Most strikingly, the putative pore seems to have an extended diameter because peptides with very large side chains such as fluorophores or octapeptides can be efficiently bound and transported by TAP [83–85]. Interestingly, peptides labeled with gold clusters of 1.4 nm in diameter bind specifically to TAP (L. Neumann and R. Tampé, unpublished results). Moreover, some peptides containing hydrophobic side groups have higher affinity for TAP than the unlabelled peptide [83,84].

---

**Fig. 3. Model of substrate binding and translocation by the TAP complex.** Peptide (▲) and ATP bind to TAP from the cytosol in an uncoordinated event. Under physiological conditions, it can be assumed that the NBDs are already loaded with ATP/ADP. Peptide binding to TAP induces a structural reorganization of the TAP-peptide complex triggering ATP hydrolysis in the NBD which subsequently leads to the transport of the peptide across the ER membrane and the release in the ER lumen. However, we cannot distinguish whether ATP hydrolysis at both NBDs is required for substrate translocation (parallel mode) or whether ATP hydrolysis at one NBD drives the translocation leaving the second NBD for completion of the transport cycle (sequential mode).
3. Multistep transport mechanism of TAP

Peptide transport by TAP is a multistep process (Fig. 3). Peptide associates with TAP in an ATP-independent manner following a monophasic 1:1 Langmuir binding model \((A+B\rightleftharpoons AB)\) [83,86]. In direct peptide binding or competition assays as well as photocrosslinking experiments, no indication for a second binding site was found. However, it cannot be entirely ruled out that a second binding site with very similar affinity or alternatively with very low affinity exists. By using peptides labeled with an environmentally sensitive fluorophore, the association pathway to TAP could be kinetically dissected in real time [84]. Peptide binding to TAP is composed of a fast bimolecular association step followed by a slow isomerization of the TAP-peptide complex. This structural reorganization may trigger the ATP hydrolysis and substrate translocation across the membrane. TAP selectivity is primarily determined by the first bimolecular association step. Further evidence for a structural isomerization of the TAP complex was deduced from chemical crosslinking experiments in which the level of crosslinked TAP heterodimers increased in the presence of peptide [57].

Peptide translocation strictly requires the hydrolysis of \(\text{Mg}^2\text{ATP}\) [37-40]. Non-hydrolyzable ATP analogs, such as ATP\(\gamma\)S, AMP-PNP or AMP-PCP, do not promote peptide transport. Peptide transport can be energized by ATP, UTP, CTP and GTP [39,87]. The Michaelis-Menten constant \(K_M\) (\(\text{Mg}^2\text{ATP}\)) for peptide transport is \(0.1-1\) mM [87]. Direct binding of nucleotides was demonstrated by 8-azido-ATP photocrosslinking experiments [53,88,89]. Interestingly, ATP and ADP have similar affinity for TAP, explaining that peptide transport can be inhibited by ADP and other nucleoside diphosphates competing for ATP binding [87]. Both nucleotide binding domains interact with ATP even if separately expressed [88,89]. However, the NBDs of TAP1 and TAP2 are unable to hydrolyze ATP, suggesting that ATP hydrolysis requires a certain arrangement of both NBDs via the transmembrane domains mediated by peptide binding. In addition, both NBDs are essential for TAP function because mutation of one NBD (R659Q of hTAP1) leads to a loss of transport function [90].

By partial purification and reconstitution into proteoliposomes, the ATPase activity of the TAP complex could be analyzed for the first time (S. Gorbulev and R. Tampé, manuscript in preparation). The ATPase activity is substrate-specific and can be blocked by a viral TAP inhibitor (see Section 5). Most strikingly, the ATPase activity is tightly coupled to peptide binding, indicating that substrate binding is a requisite step for subsequent ATP hydrolysis. Thus, peptide binding may cause a structural rearrangement of the NBDs that could function as a molecular switch to activate the ATPase of TAP thereby preventing the waste of ATP without translocation of peptides.

Regarding the final ATP-dependent translocation step, important questions remain open. Do the two NBDs work in sequential or synchronous fashion which implies the question whether one or two ATP are needed for a complete transport cycle? Furthermore, are both NBDs equal in function, or does ATP hydrolysis at one NBD drive the peptide translocation whereas hydrolysis at the other is needed for completion of the transport cycle and reconversion of the initial substrate binding site exposed to the cytosol? It will be interesting to see in the future how closely TAP resembles transport models of MDR or CFTR (for review see [91]).

4. The substrate binding motif of TAP – coevolution of affinity, specificity and diversity

In contrast to our limited understanding of the structure of TAP including the spatial arrangement of the substrate binding pocket, the substrate specificity of TAP has been well studied. The first generation of results came from experiments based on trapping transported peptides in the ER via glycosylation. Comparing the number of glycosylated peptides differing in amino acid composition or length, information about sequence and length preferences of the transported peptides was extracted (for review see [4,92,93]). The most efficient transport was observed for peptides with a length of 8–12 amino acids [94] whereas van Endert et al. [86] showed an optimum of peptides with 8–16 amino acids for peptide binding. However, peptides of six or 40 amino acids in length are also transported with TAP but with lower efficiency as compared to peptides with 8–12
amino acids. In conclusion, TAP preferentially transports peptides similar or slightly larger in length suitable for MHC class I binding. In an alternative approach, TAP selectivity was studied by ATP-independent peptide binding assays [83,86]. The specificity studied with both methods showed the same affinity pattern for different substrates, but with notable differences in absolute values of the affinity for TAP [76,83,95]. The glycosylation method may lead to an underestimation of differences of transported peptides because this assay includes several side reactions such as transport, glycosylation, degradation and peptide export. Each step has its own kinetic parameter and specificity. In contrast, peptide binding to TAP is a simple bimolecular reaction. Thus, the binding assay seems to be more accurate in resolving subtle details of the binding motif than the glycosylation assay.

To resolve the recognition principle and the substrate binding motif of TAP, complex peptide libraries were applied [96]. With this combinatorial method comparing the average affinity of a randomized peptide mixture with one residue in common with a totally randomized peptide mixture, it was possible to determine the influence on the affinity to TAP of each peptide residue independently of a given sequence context. Thus, the effect of each amino acid residue on stabilization of peptide binding to TAP could be determined. The peptide with the highest affinity showed a 45-fold higher affinity than a totally randomized peptide mixture resembling the selectivity of TAP for certain peptides. In a similar approach using an octapeptide library, a murine MHC class I allele showed a 200-fold higher affinity for a high affinity peptide [97]. On the basis of this coarse comparison, TAP seems to be less restrictive than MHC class I molecules. Most interestingly, the effect of amino acids on the binding to TAP is critically dependent on the position in the nonapeptide. As reported earlier, the strongest differences in peptide binding affinity were observed at the carboxy-terminus of the peptides, as seen for some rat and mouse alleles [45,98]. Human TAP showed a preference for peptides with hydrophobic or basic amino acids (Phe, Leu, Arg or Tyr) at the carboxy-terminus which are also preferred peptide anchors for MHC class I binding (Fig. 4). It should be mentioned that none of the disfavored residues Asp, Glu, Asn or Ser at the carboxy-terminus are anchors for peptide binding to MHC class I molecules. Thus it can be speculated that the recognition principle of TAP and MHC class I molecules coevolved to transport preferentially peptides which later on are bound to MHC class I molecules. Moreover, there are hints that the proteasomal subunits LMP2 and LMP7 upregulated by INF-γ upon immune stimulation increase the trypsin and chymotryptic activity of the proteasome while reducing the peptidyl-glutamyl-peptide hydrolyzing activity [99]. Thus more peptides available for MHC class I binding are generated. The amino-terminal residues of peptides binding to TAP and to MHC class I molecules do not agree very well. Human TAP favors Arg at the second peptide position although Arg is only a weak anchor residue for certain MHC class I alleles at this position. In contrast, Pro at position 2, reducing the affinity for TAP dra-
tically, is the preferred amino acid for other alleles at this position. A similar effect for Pro at position 3 was found for murine TAP [100] suggesting that it is an important principle. Supported by a series of indications, it can be speculated that an amino-terminal exopeptidase located in the ER lumen processes the peptides transported by TAP for loading of some epitopes. This model is supported by the fact that TAP transports peptides that are longer than those bound to MHC class I molecules. While MHC class I alleles use positions 2 and 3 as anchors at the amino-terminus, position 1 of the peptide substrate is also important for TAP selectivity. Here, Lys, Asn and Arg are favored, whereas Asp and Glu weaken TAP binding.

To determine the contribution of the peptide backbone and steric constraints to the peptide binding to TAP, substrates with different lengths were modified by D-amino acids in each position and tested for TAP binding. Only D-amino acids in positions 1–3 and the carboxy-terminal position showed a strong decrease in affinity. This result together with the effect of proline in position 2 supports the idea that the peptides are fixed at positions 1, 2 and 3 via contacts mostly to the peptide backbone. The carboxy-terminal of the peptide functions as a second important anchor selecting for peptides with hydrophobic and basic amino acids. Most importantly, the peptides are fixed at the free amino- and the carboxy-terminus via hydrogen bonding [96]. In analogy to MHC class I molecules, it seems likely that the amino- and carboxy-terminal anchor regions do not influence the binding and transport by TAP, indicating that this part of the peptide does not interact with TAP (Fig. 4, lower panel). This binding motif explains that longer peptides can bulge out of the substrate binding pocket and that modifications by very large amino acid side chains are tolerated within this region. Most strikingly, the T cell receptors make contacts mainly between residues 5–8 of MHC class I-associated peptides [12,13]. Thus, the T cell recognition takes place in an area where TAP shows the lowest specificity, therefore enlarging the pool of peptides presented on MHC class I molecules. Thus, the proteasome, TAP and MHC class I molecules might have co-evolved a similar substrate specificity to optimize the antigen processing machinery.

Fig. 5. Pathway of assembly of MHC class I molecules and formation of the macromolecular TAP-tapasin-MHC complex critical for loading of antigenic peptides onto MHC class I molecules. Peptide-loaded MHC molecules can dissociate from the TAP complex. Kinetically unstable MHC-peptide complexes (fast $k_{off}$) are trapped and recycled by association with tapasin and TAP. Only kinetically stable MHC complexes (slow $k_{off}$) leave the ER via the Golgi to the cell surface ('quality control mechanism').
5. TAP as part of a macromolecular transport and chaperone complex

In the assembly and loading of MHC class I molecules a growing number of auxiliary, chaperone-like proteins has been identified (for review see [102,103]). Some of them are in close association with TAP (Fig. 5). Calnexin and calreticulin bind to monoglycosylated core glycans in the ER lumen and increase the efficiency of correct folding of glycoproteins (see review [104]). When the binding of these chaperones is inhibited, for example by glucosidase inhibitors, the folding and subsequent surface expression of MHC class I molecules is drastically decreased [105,106]. Calnexin, a type I membrane protein, binds to nascent MHC class I heavy chain before assembly with β2-m and correct folding occur. The thiol-dependent reductase ERP57 found in complex with MHC class I heavy chain, calnexin and calreticulin also binds in a very early step of the maturation probably supporting the correct formation of the conserved disulfide bridges [107-109]. Then calreticulin, a soluble protein with high sequence homology to calnexin, binds to MHC class I heavy chain perhaps displacing calnexin and supporting the assembly of heavy chain and β2-m. Furthermore, calreticulin seems to retain empty MHC class I molecules in the ER. Most strikingly, MHC class I molecules are found in close association with the TAP complex [110,111]. As reported by Powis, MHC class I molecules interact with TAP1 and TAP2 [112]. Approximately four MHC class I molecules seem to be linked via four tapasins to one TAP complex [58]. Tapasin is a ER-resident type I glycoprotein consisting of two immunoglobulin folds [58,113]. The reason for this unexpected stoichiometry could be the presence of different MHC class I alleles to increase the efficiency of antigen processing and presentation. In addition to its critical role in mediating the TAP-MHC association, tapasin also has an important function in the assembly and peptide loading of MHC class I molecules probably based on a chaperone effect. As shown by Lehner and coworkers, truncated tapasin missing the transmembrane region and the cytosolic tail does not associate with the TAP complex but still rescues MHC class I surface expression [114]. Finally, tapasin increases the level of TAP [114]. It is not unlikely that more, at the moment unknown proteins are involved in the assembly and loading of MHC class I molecules. In addition to the thiol-dependent reductase activity, ERP57 also has a cysteine reductase activity that may possibly process the TAP-derived peptides for optimal binding to MHC class I molecules [107].

6. Implication in human diseases

In the past few years, it has become evident that TAP is involved in several human diseases. Herpes- and adenoviruses are known to interfere with antigen presentation by a downregulation of MHC class I molecules on the cell surface of virus-infected cells (for review see [115]). The immediate early gene product ICP47 of herpes simplex virus type 1 (HSV-1) was identified to inhibit peptide loading onto MHC class I molecules thereby evading detection by cytotoxic T lymphocytes [116]. ICP47 inhibits peptide translocation into the ER [117,118] by blocking the peptide binding site of TAP with high affinity [119,120]. ICP47 is highly species-specific because the affinity for mouse TAP is 100-fold reduced in comparison to human TAP [119,120]. For the efficient inhibition of peptide binding a 32 amino acid fragment comprising amino acid 3-34 of ICP47 is sufficient as seen by binding studies with truncated ICP47 [121]. Which part of the 32mer associates with the peptide binding site is not known so far. However, a photocrosslinker positioned on the amino-terminus of the ICP47 fragment labels both subunits to the same extent whereas 125I-labeled Tyr21 polypeptide crosslinks only TAP1, suggesting that ICP47 is asymmetrically bound to TAP [122]. Upon binding to membranes, ICP47 seems to undergo a conformational change from a loosely folded to an α-helical structure because structural studies showed that the 32mer is loosely structured in aqueous solution. However, by binding to membranes, the polypeptide seems to adopt an α-helical structure [123]. The structure of the active domain of ICP47 was analyzed by NMR. In detergent solution, the active domain of ICP47 consists of two α-helices at residues 3-13 and 23-32 connected by a flexible loop [124].

Human cytomegalovirus encodes at least four different proteins that inhibit the cell surface expression of MHC class I molecules in different ways and at
different stages after infection (for review see [125]). In contrast to ICP47, the ER-resident transmembrane class I glycoprotein US6, expressed late post infection, binds probably to the ER-luminal part of TAP, thereby inhibiting peptide translocation [126–128]. US6 does not affect the association of tapasin, calreticulin and MHC class I molecules to TAP [126]. Moreover, the action of US6 can be overcome by TAP induction with INF-γ. In contrast to ICP47, which competes with peptides for the binding site on TAP, US6 inhibits neither peptide binding nor ATP binding to TAP (Fig. 6). Thus, it seems that US6 blocks the translocation step by interacting with the luminal part of TAP and prohibiting a conformational change required for peptide translocation.

Only little is known about congenital human TAP deficiency. Inherited TAP deficiency in a family caused by a stop codon at position 253 of TAP2 resulting in non-functional TAP has been described [129]. The homozygous TAP2−/− siblings in this family show a 100-fold decreased expression of MHC class I molecules on the cell surface. Cytotoxic
T lymphocytes were present in a reduced amount. Despite this TAP deficiency, these siblings do not show increased susceptibility to viral infection. In further studies with MHC class I molecules expressed in lymphoblastoid cell lines and cytotoxic T cells derived from these individuals it was observed that the antigens seem to be translocated to the ER lumen in a TAP-independent manner [130].

At present, the participation of TAP in autoimmune diseases is a matter of controversy. Whereas some publications report that TAP is linked to autoimmune diseases, others report the opposite. Thus, further studies will be necessary to elucidate the linkage of TAP to autoimmune diseases.

In tumor cell lines and tumor tissues, a loss of antigen presentation by MHC class I molecules is often observed and the downregulation of MHC class I surface expression can have various reasons (for review see [131,132]). In some malignant tumors, the deficient presentation of endogenous antigens on the cell surface is caused by a drastically reduced level of mRNA of TAP1, TAP2, LMP1 and LMP2 that could be restored by IFN-γ stimulation. In these renal cell carcinomas, the transcription of TAP and LMP seems to be downregulated by an unknown mechanism [133]. Another interesting example of a tumor escape mechanism was found in small lung cancer where a TAP1 (R659Q) mutant was identified [90]. Expression of MHC class I molecules on the cell surface could be restored by expression of wild-type TAP1 or adding exogenous peptides. R659 is localized between the C-loop and the Walker B motif, indicating that this mutation affects ATP binding or hydrolysis. The suppression of MHC class I molecules at the cell surface is a tumor escape mechanism from immune recognition for tumor cells. However, it should kept in mind that these mechanisms are more complex as cells lacking MHC class I molecules on the cell surface might become targets for natural killer cells.

References
