Conformational Flexibility of the MHC Class I α_1-α_2 Domain in Peptide Bound and Free States: A Molecular Dynamics Simulation Study

Martin Zacharias and Sebastian Springer
International University Bremen, School of Engineering and Science, D-28759 Bremen, Germany

ABSTRACT Major histocompatibility complex class I proteins play a key role in the recognition and presentation of peptide antigens to the host immune system. The structure of various major histocompatibility complex class I proteins has been determined experimentally in complex with several antigenic peptides. However, the structure in the unbound (empty) form is not known. To study the conformational dynamics of the empty major histocompatibility complex class I molecule comparative molecular dynamics simulations have been performed starting from the crystal structure of a peptide bound class I peptide-binding domain in the presence and absence of a peptide ligand. Simulations including the bound peptide stayed close to the experimental start structure at both simulation temperatures (300 and 355 K) during the entire simulation of 26 ns. Several independent simulations in the absence of peptide indicate that the empty domain may not adopt a single defined conformation but is conformationally significantly more heterogeneous in particular within the α-helices that flank the peptide binding cleft. The calculated conformational dynamics along the protein chain correlate well with available spectroscopic data and with the observed site-specific sensitivity of the empty class I protein to proteolytic digestion. During the simulations at 300 K the binding region for the peptide N-terminus stayed close to the conformation in the bound state, whereas the anchor region for the C-terminus showed significantly larger conformational fluctuations. This included a segment at the beginning of the second α-helix in the domain that is likely to be involved in the interaction with the chaperone protein tapasin during the peptide-loading process. The simulation studies further indicate that peptide binding at the C- and N-terminus may follow different mechanisms that involve different degrees of induced conformational changes in the peptide-binding domain. In particular binding of the peptide C-terminus may require conformational stabilization by chaperone proteins during peptide loading.

INTRODUCTION

Major histocompatibility complex (MHC) class I glycoproteins play a key role in the recognition of pathogens by presenting antigenic peptides at the cell surface. MHC class I proteins consist of α_1, α_2, and α_3 domains and a C-terminal membrane anchor region. Before transport to the cell surface (extracellular side) MHC class I proteins associate with β-2-microglobulin (β_2m) and antigenic peptides (8–10 residues). Peptides are bound by the α_1-α_2 domain in a long and narrow cleft located between two α-helices on top of an antiparallel β-sheet (Fig. 1). In humans one can distinguish three classical MHC class I subtypes termed HLA-A, -B, and -C (corresponding subtypes in mice are H-2D, H-2K, and H-2L). Among the subtypes, HLA-A and HLA-B are mainly responsible for presenting antigenic peptides at the cell surface and show also a greater polymorphic variation than type HLA-C. Other structurally related, ''nonclassical class I'' or class 1b MHC proteins bind a more restricted set of peptides or serve other functions (Maenaka and Jones, 1999). Peptide loading and maturation of stable peptide-filled class I proteins occurs in the endoplasmic reticulum (ER) in a multistep process that involves interaction with the chaperone-like proteins calnexin, calreticulin, tapasin, the protein disulfide isomerase Erp57, and the peptide transporter TAP (reviewed in Bouvier, 2003; Ellgaard and Helenius, 2003). The structural details of class I peptide interaction have been investigated extensively using x-ray crystallography of many MHC class I-peptide complexes (Madden, 1995). A major result of these studies is that even though there exists a great allelic variability among class I molecules (several hundred different class I sequence variants are known), the basic features of peptide binding are conserved (Madden, 1995). Peptide binding to the class I binding cleft is mainly mediated through contacts of the peptide N- and C-termini. These are held by networks of hydrogen bonds that are very similar in all class I-peptide structures and make the largest contribution to the binding energy of the peptide (Matsumura et al., 1992; Bouvier and Wiley, 1994; Fahnestock et al., 1994; Madden, 1995). These hydrogen bonds come from the β-sheet and from both lateral helices of the binding groove such that the peptide termini stabilize the structure of the α_1-α_2 peptide binding domain. In addition, two hydrophobic pockets, so-called primary anchor regions, interact with small hydrophobic residues (preferentially Ile, Leu, or Val) at the N- and C-terminus of the bound peptide (positions 2 and the last side chain of the bound peptide). Other contributions to binding come from interactions between class I and the peptide backbone or the variable peptide side chains. These interactions occur in all class I-peptide complexes but differ according to the class I allele and bound peptide (Jones et al., 1998).

Submitted April 20, 2004, and accepted for publication July 2, 2004.

Address reprint requests to Martin Zacharias, International University Bremen, School of Engineering and Science, Campus Ring 6, Bremen 28759, Germany. Tel.: 49-421-2003541; E-mail: m.zacharias@iu-bremen.de.

0006-3495/04/10/2203/12 $2.00

doi: 10.1529/biophysj.104.044743

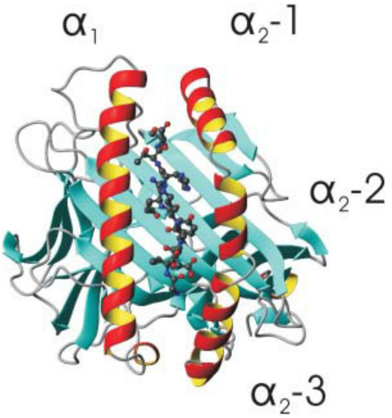

FIGURE 1 Cartoon representation of the MHC class I molecule HLA-A*0201 in complex with an antigenic peptide consisting of 10 residues (Hillig et al., 2001; peptide sequence: GVYDGREHTV; Protein Data Bank entry:1A1O). The α-helices of the MHC protein are in red/yellow (β-strands in *light blue*). The view is into the α_1-α_2 domain peptide binding cleft (peptide in atomic resolution). The β_2m protein and the α_3 domain are located below the α_1-α_2 domain.

The contacts observed in available experimental class I structures allow a structural interpretation of the class I peptide-binding specificity and sequence dependence. However, the mechanistic details of the class I peptide-loading process and the structure of the peptide-free MHC class I molecules (empty form) are less well understood. So far, attempts to crystallize class I receptors in the empty form have been unsuccessful. Various biophysical and biochemical techniques have been used to characterize the peptide-free form (Bouvier and Wiley, 1998; Springer et al., 1998). Circular dichroism spectroscopy indicates that the secondary structure of peptide-free and peptide-bound class I (HLA-B7) are similar at room temperature but differ substantially at higher temperature (70°C). Temperature-dependent tryptophan fluorescence indicates that the empty class I molecules may unfold at lower temperature (~30°C) than the peptide-bound form (~60°C; Springer et al., 1998). Empty MHC class I molecules are also much more susceptible to proteolytic cleavage in particular at sites close to the peptide-binding cleft (Bouvier and Wiley, 1998).

One possible reason for the increased sensitivity of class I molecule to proteolytic cleavage and for the failure to crystallize class I molecules in the empty form might be a significantly enhanced conformational heterogeneity or flexibility in the absence of a peptide. A putative role of the chaperone proteins which participate in forming the class I peptide-loading complex in the ER could be to stabilize specific MHC class I conformations that allow easy access of antigenic peptides to the peptide-binding region. Knowledge of the accessible states of the empty MHC class I molecules and characterization of flexible regions would allow to better understand the conformational transitions necessary to bind an antigenic peptide and may give hints on which parts of the class I MHC protein require stabilization by association with the loading complex.

Molecular dynamics (MD) simulation is a powerful tool to compare the conformational dynamics of MHC class I molecules in ligand-bound and free states at high resolution in space and time. In previous studies the molecular dynamics method has been used to study the interaction between peptides and MHC class I molecules (Rognan et al., 1992, 1994) and the role of water molecules and water structure located at the class I–peptide interface (Meng et al., 1997, 2000; Petrone and Garcia, 2004). The recent study by Petrone and Garcia (2004) emphasizes the role of water molecules for high affinity binding of various peptides to MHC class I molecules. It was found that water molecules make a favorable free energy contribution to peptide binding that increases the affinity of bound peptides without contributing to peptide selectivity. The study also indicated a larger mobility of the HLA-A2 class I molecule in the absence of a peptide compared to simulations in the presence of a peptide from the HIV reverse transcriptase on the 5 ns timescale (Petrone and Garcia, 2004).

In this study several fairly long MD simulations (each >25 ns) on the α_1-α_2 domain of an HLA-A MHC class I protein in the presence and absence of a peptide ligand have been compared. The main focus of the simulation studies is to compare the conformational flexibility of the peptide-binding region in the empty and peptide-bound states and relate it to the mechanism of peptide recognition. The α_1-α_2 domain alone without α_3 and β_2-microglobin has been shown to bind peptides (Rigney et al., 1998). Restriction of our studies to the α_1-α_2 peptide-binding domain allowed us to perform several extensive simulations at two different temperatures reaching a total simulation time (of all simulations) of >150 ns. The simulations reveal significant differences in conformational flexibility between regions of the protein involved in binding the peptide C-terminus (more flexible) compared to those mediating the binding of peptide N-terminal parts. This result indicates that binding of the peptide N- and C-terminus may involve different degrees of induced conformational changes at the class I-binding cleft during peptide binding.

METHODS

Molecular dynamics (MD) simulations were started from a crystal structure of a HLA-A*0201 class I protein in complex with a peptide consisting of 10 amino acids (Hillig et al., 2001; Fig. 1). This system was chosen as a model system since it has been determined to very high resolution of 1.4 Å. The bound peptide has the sequence: GVYDGREHTV and corresponds to a tumor-specific antigenic peptide derived from the melanoma antigen (MAGE-A4 protein; Hillig et al., 2001). The simulations included the first 181 residues corresponding to the α_1-α_2 peptide-binding domain either in the presence (+peptide) or absence of antigenic peptide. All MD simulations were performed with the *sander* module of the Amber6 (Assisted Model Building with Energy Restraints) package (Pearlman et al., 1995) in a periodic box including TIP3 water (Jorgensen et al., 1983) and using the parm99 force field (Cornell et al., 1995; Wang and Kollman, 2001). Initial positions of ten additional sodium and chloride counter ions were placed using the *xleap* module of the Amber package; ~7500 water molecules were added to fill the box. A 9 Å cutoff for the short-range nonbonded interactions was used in combination with the particle mesh Ewald option (Darden et al., 1993) using a grid spacing of ~0.9 Å to account for long-range electrostatic interactions. The conformation of the solvated protein was first relaxed via energy minimization. After minimization the system was gradually heated from 50 to 300 K with positional restraints on the protein atoms over a period of 0.1 ns. During another 0.1 ns simulation time at 300 K the positional restraining force constant was gradually reduced from 50 kcal mol^{-1} $\mathrm{Å}^{-2}$ to zero. The simulations were continued for a total simulation time of 26 ns for both the peptide-bound and free forms. To check the dependence of the results on the initial conditions a separate 26 ns simulation was run starting from the same structure but using different initial velocities. Additional simulations at an elevated temperature of 355 K were initiated starting from the coordinates of the 300 K simulations at 6.2 ns (heating up in 15 K steps within 1 ns). Two such simulations were performed for the empty class I form and one in the case of the peptide-bound form and were continued to reach in each case a final simulation time of 26 ns. All MD simulations were performed at constant pressure of 1 bar with relaxation time of 5 ps. Solute coordinates were stored each 0.4 ps simulation time. Principal component analysis (PCA) of the trajectory (only backbone heavy atoms) was performed as described (Amadei et al., 1993; Wlodek et al., 1997; Sherer et al., 1999; Zacharias, 2000). The positional covariance matrix,

$$C_{ij} = 1/N_k \times \Sigma_k [x_i(k) - \langle x_i \rangle] \times [x_j(k) - \langle x_j \rangle], \tag{1}$$

(x_i atomic positions, summation, k, is over all structures, N_k, in the trajectory) was diagonalized. The resulting eigenvectors with large eigenvalues (principal components, PCs) describe flexible orthogonal degrees of freedom of the protein molecule. The eigenvalues are a measure of the conformational fluctuation or flexibility in the corresponding eigenvector direction.

RESULTS

Molecular dynamics simulations were started from a crystal structure of a HLA-A*0201 class I molecule in complex with a peptide derived from a tumor-specific antigen (Hillig et al., 2001; Fig. 1). The peptide is bound in a cleft on top of a β-sheet and flanked by two α-helices (α_1, residues; 55–85, α_2, residues 138–175). The second α-helix is not completely straight but contains kinks at residues ~150 and 165, respectively, such that one can distinguish three α-helical subsegments (α_2-1, α_2-2, and α_2-3; indicated in Fig. 1). The N- and C-terminal anchor amino acids are valine residues (second and last amino acid of the bound peptide) and are bound at the N-terminus anchor region (between helices α_1 and α_2-3) and C-terminus anchor region (between helices α_1 and α_2-1; see Fig. 1), respectively. In the following, the terms N- and C-terminus peptide binding regions include residues that interact with the peptide termini and residues that contact the N- and C-terminal anchor side chains.

The simulations were performed on the first 181 residues corresponding to the α_1-α_2 peptide binding domain in the absence (simulations on empty α_1-α_2 domain) and presence of the peptide (sequence: GVYDGREHTV) as found in the crystal structure (α_1-α_2 domain-peptide complex simulations). Two independent simulations on the empty α_1-α_2 domain and one in the presence of bound peptide were performed at 300 K (~25°C). Although fairly extensive simulations of up to 26 ns were performed, the simulation time may not be sufficient to sample all relevant states of the empty class I α_1-α_2 domain at a simulation temperature of 300 K (~25°C). To enhance the sampling of available conformational states in particular of the empty α_1-α_2 domain an additional set of simulations at a higher temperature of 355 K (82°C) has been performed (starting from the coordinates of the 300 K simulations at 6.2 ns). Even at this temperature that is slightly above the unfolding temperature of the complex (Springer et al., 1998) a simulation time of 26 ns may not be sufficient to observe a possible unfolding of the protein. However, it might be possible to observe an unset of unfolding in case of the empty α_1-α_2 domain and to characterize the flexible regions of the protein that are mainly affected by the absence or presence of a peptide ligand.

Conformational deviation of simulated structures from experimental start structure

Both simulations of the peptide-bound class I form at 300 K and 355 K showed a stable Cα main chain Rmsd (root-mean square deviation) after ~2–4 ns simulation time of ~1.5–1.8 Å with respect to the start structure (Fig. 2 *A*). Even in case of the elevated temperature simulation (355 K) the peptide-bound form stayed relatively close to the start structure up to the end of the simulation (26 ns, Fig. 2 *A*). No significant conformational changes or dissociation of the bound peptide were observed (not shown). The Rmsd of the empty α_1-α_2 domain reached in the simulations at both temperatures (300 K and 355 K) significantly higher levels compared to the simulation of the peptide-bound form (Fig. 2, *B* and *C*). It reached a stable level of ~2.2 Å after ~4 ns in the case of the two independent 300 K simulation and ~3–3.5 Å in the case of the elevated temperature simulations (Fig. 2 *C*). Interestingly, the Rmsd time course of only the β-sheet backbone was in all simulations (including those of the empty α_1-α_2 domain) significantly smaller than the Rmsd of the complete backbone (compare Fig. 2, *C* and *D*). This result indicates that most of the increased conformational mobility in the absence of the peptide ligand and/or during simulations at elevated temperature is due to the α-helical and loop regions.

Bound water molecules at the peptide binding region

X-ray crystallography of several MHC class I peptide complexes indicates that water molecules at the interface between class I receptor and bound peptides are important for high-affinity binding (Madden, 1995). The importance of water-mediated interactions between class I protein and peptide has also been highlighted in a recent molecular

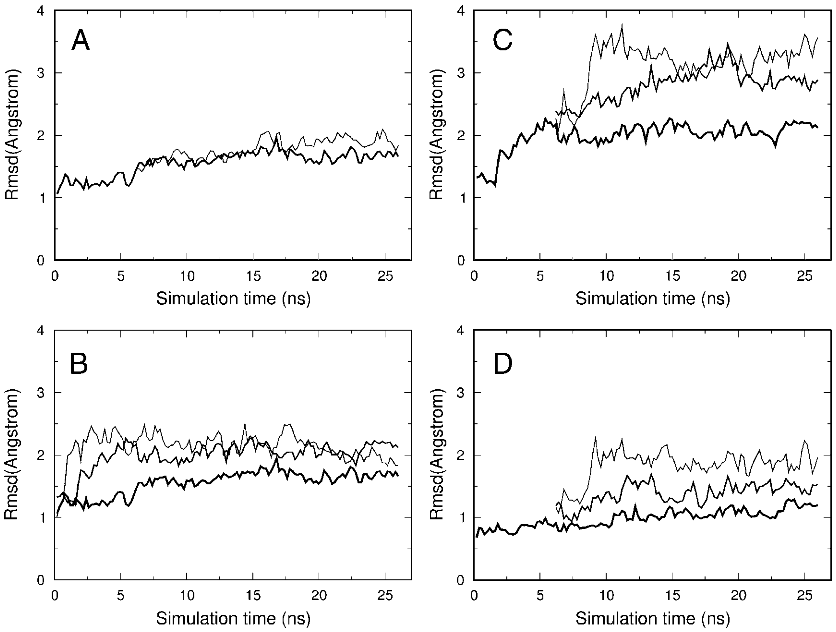

FIGURE 2 (*A*) Rmsd (Cα backbone) time course of the α_1-α_2 domain (with respect to the start structure) during the simulation of the complex at 300 K (*bold line*) and 355 K (*thin line*). (*B*) Rmsd (Cα backbone) time course of the complex simulation at 300 K (*bold line*) and the two 300 K simulations of the empty form. (*C*) Cα backbone-Rmsd of one of the 300 K simulations (*bold*) and the two simulations at 355 K (start at 6.2 ns, *thin lines*) in the absence of peptide. (*D*) Same as in *C* but Rmsd (Cα backbone) of only the β-sheet segments (residues 3–14, 20–28, 31–37, 44–48, 92–104, 109–119, 121–127, 132–135) from the start structure.

dynamics simulation study on another MHC class I peptide complex by Petrone and Garcia (2004). In this study several stably bound water molecules have been identified that stayed at definite positions in the peptide class I interface for >1 ns simulation time. In agreement with the simulation results by Petrone and Garcia (2004) we could also identify several water molecules with very long residence times for definite sites in the class I binding groove. The most stably bound water molecules were found near the bound peptide N-terminus and C-terminus. At the peptide N-terminus water molecules at definite positions were observed in contact with the terminal peptide amino group, with Glu-63, Lys-66, Tyr-59, and Tyr-171 of the class I protein. One buried water molecule which was hydrogen-bonded to the peptide amino group and the ε-carboxyl group of Glu-63 stayed at this position for >5 ns. Several other water molecules that stayed for >1 ns at a definite position were observed near Lys-66, His-70 (floor of the class I peptide-binding groove), and Arg-97 (located in the floor of the peptide-binding groove near to the binding region of the peptide C-terminus). These water molecules form hydrogen bonds between peptide main chain and side chain and main chain groups of the class I protein. At the peptide C-terminus binding region a bound water molecule that stayed at a definite position for >5 ns was found in contact with Thr-73 and Asp-77 and the carboxy-group of the bound peptide. Each of these stable water binding positions can be matched with water molecules identified in the high-resolution x-ray structure of the class I peptide complex (Hillig et al., 2001). The fact that similar positions of bound water molecules with high residence time have also been found by Petrone and Garcia (2004) in an MD simulation of another MHC class I peptide complex indicates that these waters may have a functional role for high-affinity peptide binding. In case of the simulations on the empty class I α_1-α_2 domain the peptide-binding groove is highly solvated including the regions that bind water molecules in complex with the peptide. Even in the empty form several water molecules stayed for >1 ns in the binding groove but did not stay at definite positions that match the stable water-binding sites in case of the peptide-bound class α_1-α_2 domain. The stability of water binding to some sites in case of the peptide-bound class α_1-α_2 domain is also supported by the observation that some of the binding sites observed in case of the simulations at 300 K were also occupied by water molecules in the elevated temperature simulations at 355 K (for residence times >1 ns). In particular, the buried water binding site near the peptide amino terminus was occupied by water molecules throughout the whole simulation with residence life times for individual water molecules of >1 ns. Buried water-binding sites near Arg-97 were also occupied by the same water molecules for residence times >1 ns at 355 K. In contrast, the water-binding sites (not buried) near the binding region for the peptide C-terminus showed more rapid exchange with bulk solvent (no waters were observed with nanosecond residence times at 355 K).

Characterization of conformational fluctuations along the sequence

Similar to the Rmsd time course of the β-sheet regions the Cα-backbone conformational fluctuations in β-sheet segments (indicated as *small rectangular boxes* in Fig. 3) showed also the smallest conformational fluctuations (Rmsf) along the sequence. With the exception of the small β-strand segment consisting of residues 132–135 (which contact the α_2-1 segment) the conformational flexibility of most of the

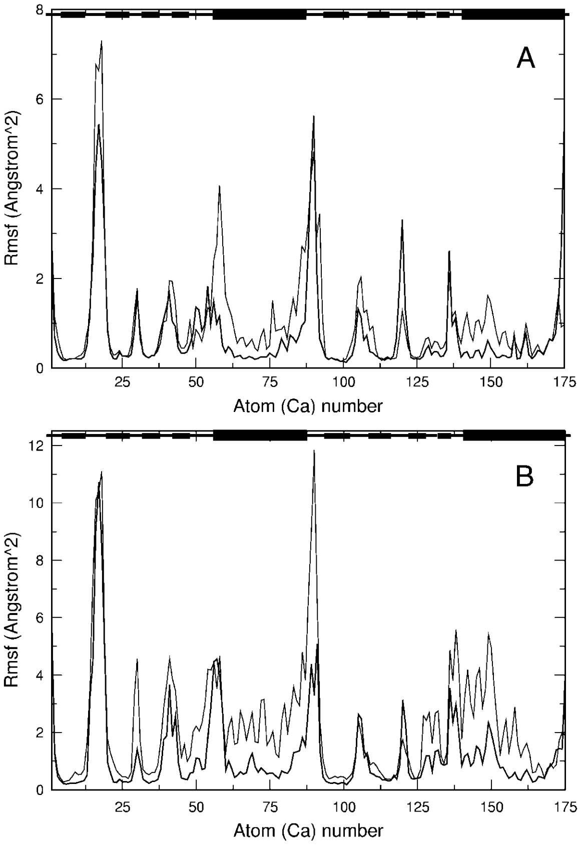

FIGURE 3 Mean square fluctuations of the Cα backbone along the sequence of the α_1-α_2 domain in the absence (*bold line*) and presence (*thin line*) of the peptide ligand for the simulations at 300 K (*A*) and 355 K (*B*), respectively. At the top of each panel the positions of β-strands (*small solid boxes*) and α-helices (*large solid boxes*) along the sequence are indicated. In case of the empty α_1-α_2 domain the average fluctuations from the two independent simulations are shown.

β-sheet regions is not dramatically enhanced in the simulations at elevated temperature (compare *thin* and *bold curves* in Fig. 3, *A* and *B*). In contrast, the fluctuations in the loop regions (connecting β-strand segments and β-strands and α-helices) are strongly enhanced in the simulations at 355 K of both peptide-bound and free-protein simulations compared to the simulations at 300 K. For the simulations at both temperatures the conformational fluctuations of β-sheet regions depended to a much lesser degree on the presence or absence of the peptide ligand compared to the α-helical regions. This is especially true for the α-helical segments that flank the binding cleft (residues 56–86 and 138–175; Fig. 3). The conformational fluctuations of these segments are significantly higher in case of empty receptor MD simulations versus simulations of the complex.

The conformational flexibility observed during the simulation for the α_1-α_2 domain in the presence of peptide showed good qualitative agreement with conformational fluctuations derived from the experimental B-factors of the high-resolution x-ray structure (Fig. 4 *A*). In Fig. 4 *B* the fluctuations of individual amino acid side chains of the bound peptide at the two simulation temperatures have been compared. Residues around the primary anchor regions (close to position 2 and around the last amino acid) showed little fluctuations of <1 Å at both simulation temperatures. The flexibility of residues 4, 6, and 8 (which are partially solvent exposed) was significantly enhanced at the elevated temperature simulations.

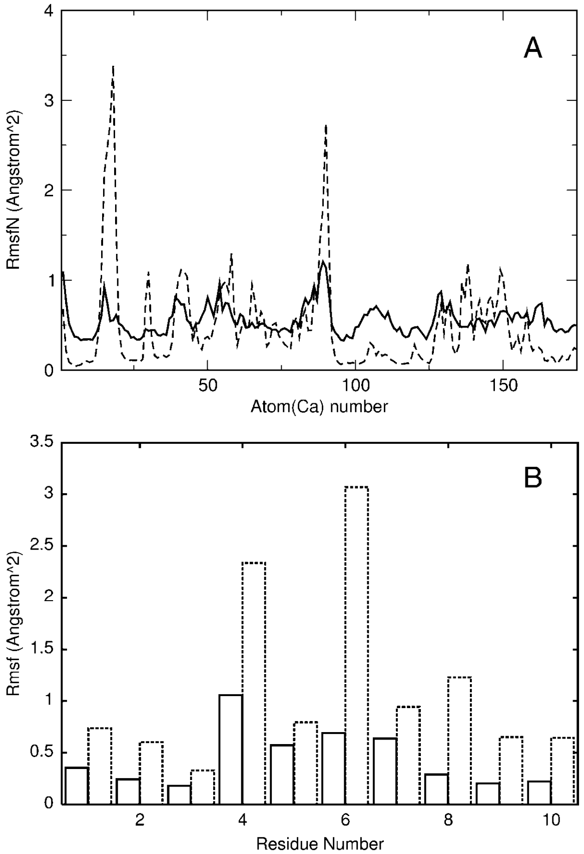

FIGURE 4 (*A*) Comparison of Cα backbone fluctuations derived from crystallographic B-factors (bold line, at a temperature of 100 K, Hillig et al.,2001, Protein Data Bank entry: 1A1O) and corresponding fluctuations observed during the simulations of the peptide bound class I α_1-α_2 domain (300 K simulation, dashed line). RmsfN indicates that the fluctuations from the simulation have been scaled to give the same average fluctuations as those derived from the B-factors. (*B*) Heavy atom conformational fluctuations of the amino acids of the bound peptide during the simulations at 300 K (*continuous line*) and 355 K (*dashed boxes*).

In the complete class I HLA-A molecule the side of the β-sheet opposite to the peptide-binding region is associated with β_2m. The residues that contact β_2m showed only small conformational fluctuations compared to the average fluctuations and residues in the ligand-binding cleft (not shown). This result indicates that the simulations largely preserve the β_2m-binding interface of the α_1-α_2 region at least on the present nanosecond timescale. It further indicates that the results obtained from simulations on α_1-α_2 domain may also be relevant for the complete HLA-A molecule. It also agrees with experimental studies that indicate specific peptide binding to the isolated α_1-α_2 domain (in the absence of β_2m) following a similar mechanism as to the full class I molecule (Rigney et al., 1998).

Mobility of the α-helical segments and residues of the primary peptide binding regions

The simulation results at both 300 K and 355 K indicate that the largest difference in conformational mobility between the peptide-bound and free α_1-α_2 domain can be attributed to the α-helices that form the walls of the peptide-binding cleft. However, the increased mobility of the α-helices in case of the simulations on the empty form is not uniformly distributed along the α-helical segments. Some of the observed transitions in the helical segments are illustrated as snapshots from the trajectories in Fig. 5 *A*. Transitions correspond to shifts and local unfolding of helical segments (e.g., of the last part of α_1 and the α_2-1 segment). Other changes correspond to transient kinks in α-helices, for example of the first helix near the binding region of the peptide C-terminus (position 72-78) and partial closing and wide opening of the binding cleft (snapshots from one simulation at 355 K). Another coupled transition observed during both 300 K and 355 K simulations corresponds to a collective shift of two segments in helix 2 (α_2-1 and α_2-2) leading to a transient narrowing of the peptide binding cleft (Fig. 5 *B*). In the peptide bound class I form the kink formed by these two helix segments points away from the peptide-bind cleft (Fig. 1 and *top panel* of Fig. 5 *A*). During the transient conformational shift this kink points into the binding cleft (Fig. 5 *B*). Interestingly, this conformational transition observed during the simulation of the empty class I α_1- α_2 domain is quite similar to a conformation found experimentally in human CD1a nonpolymorphic MHC class I-like glycoproteins (Wilson and Bjorkman, 1998; Zajonc et al., 2003) and other MHC class I-like proteins that bind nonpeptide ligands (reviewed in Maenaka and Jones, 1999). The CD1a protein is structurally similar to HLA-A or -B subtypes but binds antigenic foreign lipids instead of peptides. Binding of lipids requires a very narrow binding cleft that is realized by a conformational shift of the two helical segments toward the binding cleft. The simulation results indicate that such transitions (among others) are at least transiently also possible in case of HLA-A class I molecules in the unbound state (Fig. 5 *B*).

Comparison of the average structures of all simulations of the free protein indicate an overall average shift of the first α-helical segment of the second helix (α_2-1, residues 135–151) that is part of the peptide C-terminus binding region. The shift moves the α_2-1 segment on average away from the binding cleft (opening of the binding pocket, Fig. 6). No such shift has been observed during simulations of the complex.

The predicted enhanced conformational mobility of the helical 135–151 segment is further supported by the comparison of experimentally determined MHC-like protein structures in complex with peptides and other ligands. Such comparison reveals that most of the structural variation observed in these x-ray structures is found in the 135–151 segment and the next helical segment extending from 152–160 (Elliott, 1997). Regions of enhanced conformational flexibility or heterogeneity are often more accessible to proteolytic cleavage. In this regard the simulation results agree with comparative proteolysis studies of a peptide-bound and empty class I molecule that indicate the presence of a particularly sensitive cleavage site (at position 142) in this helical segment (Bouvier and Wiley, 1998) in case of the empty form.

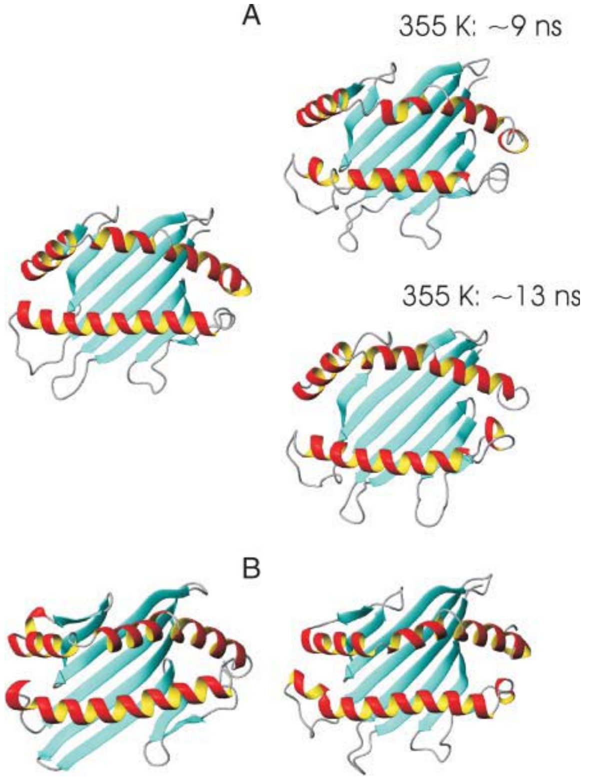

FIGURE 5 (*A*) HLA-A*0210 class I α_1-α_2 domain starting structure (*left panel*) and two snapshots obtained during the simulation at 355 K, respectively. (*B*) Comparison of a snapshot from one 300K MD simulation of the α_1-α_2 domain without peptide at 9.3 ns simulation time (*right panel*) and the α_1-α_2 domain of the CD1a class I-like protein (*left panel*; Protein Data Bank entry: 1ONQ; Zajonc et al., 2003).

During the 300 K and 355 K simulations of the empty form transient interruptions and partial unfolding of the first α-helix around residues 72–79 was observed (close to the peptide C-terminus binding region). An on-average significantly enhanced kinking of the helix in the same region was observed in one of the elevated temperature simulations which is manifested as a pronounced kink in the average structure (Fig. 6) and observed transiently in simulation snapshots (Fig. 5). It is interesting to note that position 81 in this helix close to the predicted region of enhanced helix

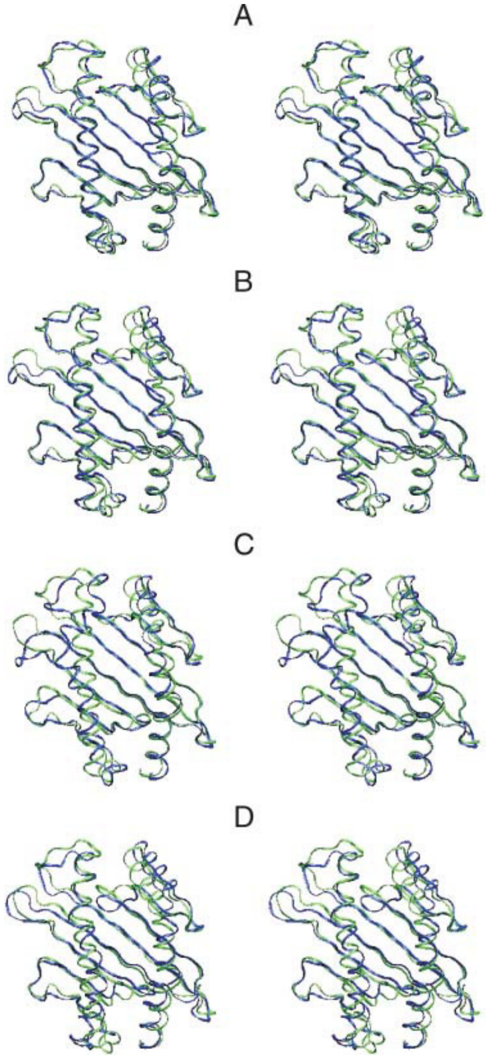

FIGURE 6 Superposition (in stereo) of the HLA-A*0210 α_1-α_2 domain backbone crystal structure (*green*; Hillig et al., 2001) and the average α_1-α_2 domain backbone conformation obtained from the two 26 ns simulations at 300 K (*A* and *B*) and the two simulations at 355 K (*C* and *D*), respectively.

kinking and unfolding was also identified as an enhanced cleavage site for proteolytic digestion of the empty class I molecule (Bouvier and Wiley, 1998).

Mobility of binding regions for peptide termini in the empty α_1-α_2 domain

The two α-helical segments that flank the binding region of the peptide N-terminus (residues 56–70 of α_1 and α_2-3 segment: residues 155–175) show similar center of geometry distance and helix axis angle distributions in the 300 K simulations of the complex and the free protein (Fig. 7, *A* and *B*). In contrast, the two helical segments that flank the binding region of the peptide C-terminus showed not only a broader distance and helix axis angle distribution in case of the MD simulations of the empty protein but also a significant shift of the distribution curve with respect to the one obtained for the simulation of the complex (Fig. 7, *C* and *D*). The shift is such that on average the distance between these two helices is larger than in the bound form (Fig. 7 *C*, see also last paragraph). This result was obtained in both independent 26 ns simulations at 300 K (Fig. 7). In case of the 355 K simulations a significant deviation of the distance and angle distributions of both helical segments that flank the peptide N- and C-terminus binding sites has been observed (not shown). The results suggest that at least at 300 K the C-terminus peptide-binding region adopts a less conserved or less stable structure that differs more significantly from the structure in the peptide-bound form than in case of the N-terminus binding region. It indicates that a larger conformational transition in the empty form is required to form a binding pocket for the peptide C-terminus than for the peptide N-terminus.

This conclusion is further supported by the analysis of the conformational fluctuations of residues that form the binding regions for peptide N-terminus and C-terminus (including the two valine anchor residues). The Rmsd probability distribution of the residues that interact with the peptide N-terminus obtained during both independent simulations at 300 K of the empty form stayed close to the distribution observed during the control simulation of the α_1-α_2 domain-peptide complex and close to the arrangement seen in the x-ray start structure (Fig. 8 *A*). A much larger deviation of the corresponding distributions was observed for the binding region of the peptide C-terminus in case of the empty α_1-α_2 domain simulation compared to the simulation of the α_1-α_2 domain-peptide complex (Fig. 8 *B*). During the simulations at 355 K the arrangement of the anchor residues still stayed close to the start structure only in case of the simulation of the complex but showed significantly larger fluctuations for both binding regions in both simulations of the empty α_1-α_2 domain (Fig. 8, *C* and *D*).

The average conformational shift of the helix formed by residues 138–150 and the larger mobility of the helices that flank the C-terminus binding region suggest, however, also that the C-terminus binding region might be at least transiently more accessible than the protein region that interacts with the peptide N-terminus (in the unbound form of the class I α_1-α_2 domain).

Comparison of correlated motions

The motions observed in case of the empty class I α_1-α_2 domain are not only significantly larger in magnitude than the motions of the protein in complex with the peptide but also overall less correlated as indicated in the 2D plots of Fig. 9. Each spot at position i, j in the plot corresponds to the fluctuation of the distance between Cα atom pairs i and j. The darker a spot the larger the distance fluctuations observed during the simulations. For example, in both plots distance fluctuations of loop regions that connect the

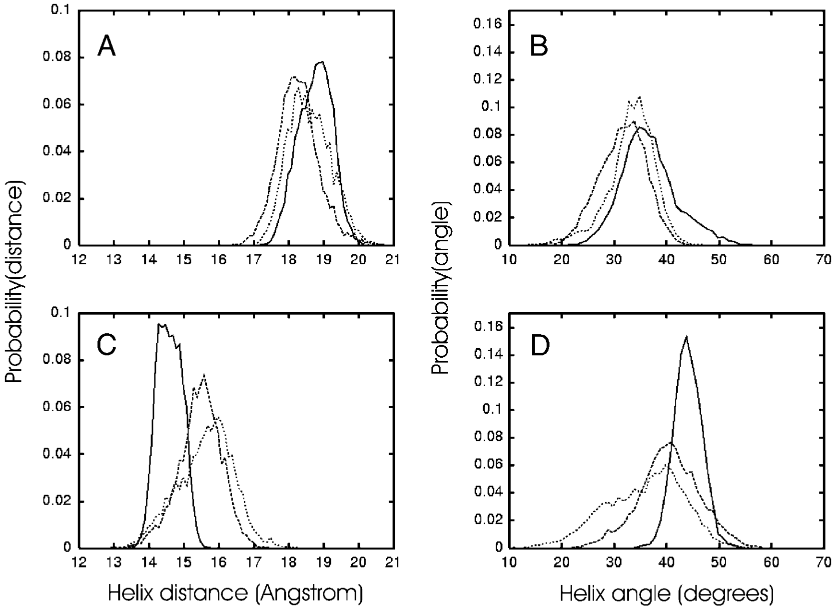

FIGURE 7 Distance (*A* and *C*) and angle (*B* and *D*) distributions observed during simulations at 300 K for the two pairs of helical segments that form the walls of the peptide N-terminus (*A* and *B*) and C-terminus (*C* and *D*) binding regions (*bold*: complex; *dashed and broken lines*: the two simulations of the empty α_1-α_2 domain). The helical segments flanking the N-terminus binding region correspond to residues 57–70 and 161–174 (in case of the peptide C-terminus: 74–85 and 138–151), respectively.

β-strands (e.g., residues around position 16, 29, etc.) show up as dark spots indicating large amplitude and uncorrelated motions of these regions with respect to other loop regions. Relative to these motions most distance fluctuations of the β-sheet regions are smaller (*shaded* or almost *white regions* in both plots). The largest qualitative difference in the character of the observed motion between the two plots is seen for the α -helical regions (indicated as *rectangular frames* in the plot for the empty form, Fig. 9). In particular, the backbone of residues ~140–160 (the first two segments of the α_2 helix) showed large-scale uncorrelated motions with respect to loop regions but also with respect to the α_1 helix (in particular residues ~55–70). The magnitude of this motion is significantly smaller in the complex (see Fig. 3) but also the character changes to a highly correlated type of motion such that the distance between atom pairs in the two α-helical segments undergoes little fluctuation (similar to the level seen for the correlation of β-strand motions). As expected the small β-strands at the outer edges of the β-sheet (residues 3–14 and 132–135) show larger amplitude and also less correlated motion compared to the ''inner-sheet'' β-strands.

Principle component analysis of MD simulations

Principle component analysis (PCA) of the positional covariance matrix (heavy atoms of the backbone) was used to investigate global collective motions of the conformations obtained during the MD simulations (Amadei et al., 1993; Wlodek et al., 1997; Sherer et al., 1999). The first few collective degrees of freedom with largest associated eigenvalues describe large global scale motions observed during the simulations (soft modes or essential dynamics modes, Amadei et al., 1993). The motions associated with the softest mode obtained from the PCA of the two simulations of the empty form at 300 K show qualitative similarity and are illustrated in Fig. 10. The motion corresponds for both simulations to a partial opening/closing of the peptide-binding cleft coupled with a partial local unfolding of helical segments especially those segments that form the binding region for the peptide C-terminus. Similar opening/closing motions of the binding cleft in the absence of a peptide associated with soft principal components of motion were also found by Petrone and Garcia (2004) in 5 ns simulation study of a HLA-A2 class I molecule. However, a partial unfolding of the helix segments near the binding region of the peptide C-terminus associated with the softest PCs was not reported (Petrone and Garcia, 2004). The softest mode with the largest contribution to the overall motion is also associated with a collective motion of the β-sheet. This motion leads to a flattening or bending of the β-sheet and shifts of parts of the helices. Although this type of global motion is also present in the soft modes obtained from the simulation of the peptide bound form its amplitude is much reduced in the presence of the bound peptide. The second and third softest PCs and some higher modes in both simulations are also associated with partial opening/closing motions of the peptide-binding cleft. However, the amplitude of motion associated these collective modes is already by a factor of 2 and more reduced compared to the motion associated with the first mode (not shown).

DISCUSSION

This study aimed at characterizing the conformational flexibility of a peptide bound and empty HLA-A α_1-α_2 domain using molecular dynamics simulations starting from a high-resolution crystal structure of the peptide-bound form (Hillig et al., 2001). Previous shorter simulations studies on

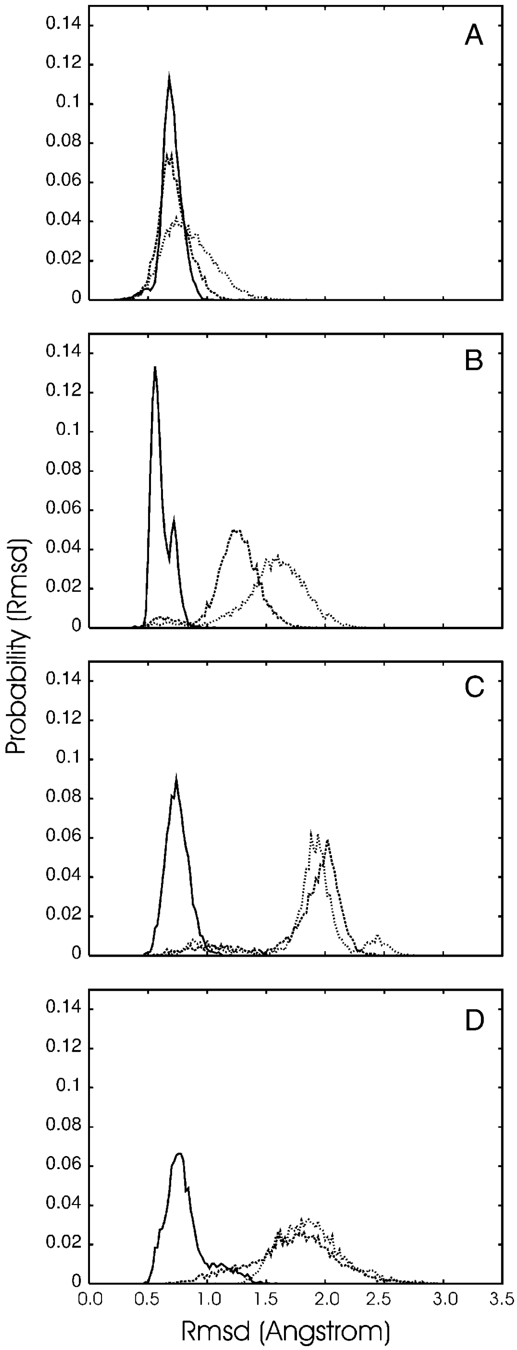

FIGURE 8 Heavy atom Rmsd distribution of residues forming the regions for binding the peptide N-terminus (residues 7, 9, 45, 67, 63, 99; in panel *A* and *C*) and C-terminus (residues 77, 81, 116, 123, 143, 147; in panels *B* and *D*) of the α_1-α_2 domain during the simulations at 300 K (*A* and *B*) and 355 K (*C* and *D*), respectively. The bold curves correspond to the Rmsd distribution of the binding pocket residues for the simulation in the presence of the peptide. The plots obtained from the simulations of the empty α_1-α_2 domain are shown as broken and dashed lines, respectively.

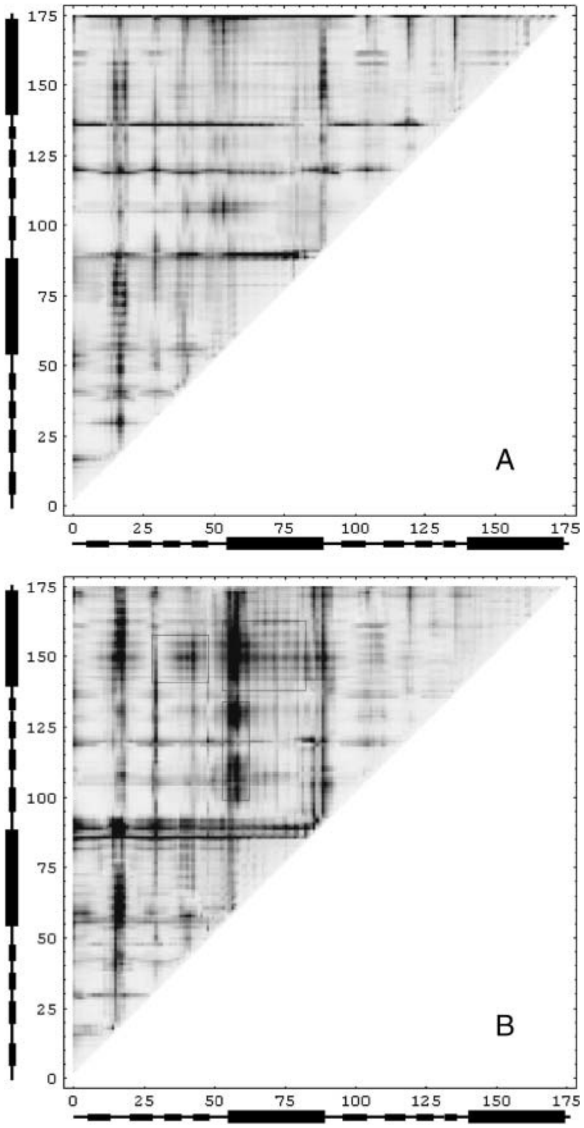

FIGURE 9 Cα-Cα distance correlation plot for the 300 K simulations of the α_1-α_2 domain in complex with peptide (*A*) and of the empty form (*B*). The shaded scale corresponds to the magnitude of the distance fluctuations between a pair of Cα atoms. Regions of most significant difference between plots *A* and *B* are indicated as framed boxes in *B*. The positions of β-strands (*small solid boxes*) and α-helices (*large solid boxes*) along the sequence are indicated at the plot axis.

MHC-class I molecules have mainly focused on studying the interaction between peptides and class I molecule and peptide sequence dependence (Rognan et al., 1992, 1994; Pohlmann et al., 2004) and on the influence of water molecules on peptide binding (Meng et al., 1997, 2000; Petrone and Garcia, 2004). The MD simulation studies by Petrone and Garcia (2004) indicate that water molecules at the peptide interface serve to stabilize and increase the affinity of peptide binding. In this study similar water-binding sites could be identified with water residence times exceeding 1 ns. Several of these water molecules form hydrogen bonds with the peptide backbone and conserved side chains of the class I molecule. This supports the view that bound water molecules at the peptide class I interface are of general importance to stabilize peptide binding. In our

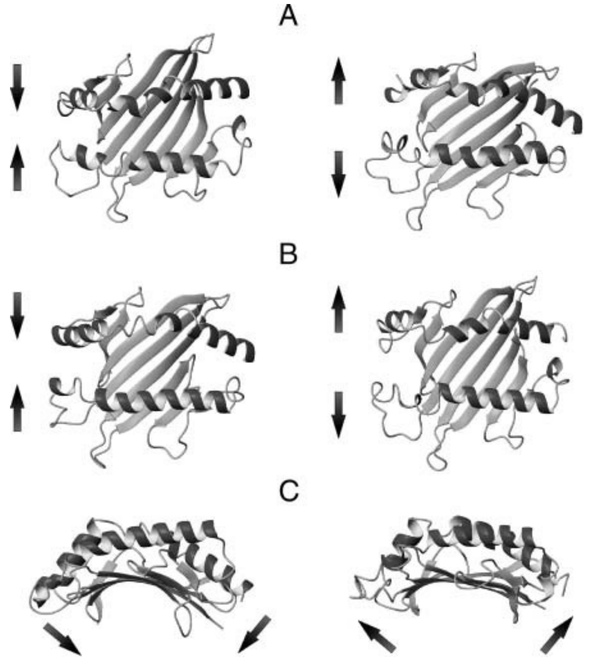

FIGURE 10 (*A*) Deformation of the average α_1-α_2 domain structure in the two possible directions of the softest principal component of motion (*left* and *right panels*, respectively) extracted from the first simulation of the empty domain at 300 K (*A*) and second simulation at 300 K (*B*). (*C*) Same as *B* but the molecule is rotated by 90° with respect to its long axis. The overall direction of motion associated with the softest PCs is indicated by solid arrows.

simulations the largest differences in conformational mobility of peptide-bound and free α_1-α_2 domains were found for the two α-helical segments that form the wall of the peptide-binding cleft. Even at elevated temperature simulations of the empty class I α_1-α_2 domain (355 K) the structural integrity of the β-sheet region was largely preserved whereas the α-helical segments showed large amplitude shift, bending and kinking motions, and partial unfolding. On average the overall content of secondary structural elements was well conserved at least in all simulations at 300K. This was also found in the simulation study of Petrone and Garcia (2004) on an HLA-A2 molecule in the absence of a peptide. The result agrees with conclusions by Bouvier and Wiley (1998) who found similar CD spectra (indicating similar secondary structure) for both empty and peptide-bound HLA class I molecules at low temperature but differences in CD spectra at elevated temperatures. The simulations allow also to explain the experimentally observed increased proteolitic sensitivity especially of the α -helical segments in the class I α_1-α_2 domain in the absence of bound peptides (Bouvier and Wiley, 1998). Even the observed site-specific enhancement of proteolytic sensitivity along the α-helices could be correlated with enhanced conformational heterogeneity at specific regions during the simulations. The α-helices forming the wall of the peptide-binding cleft showed a significantly higher level of uncorrelated motion in case of the empty α_1-α_2 domain than for the complex. This type of motion is more characteristic for partially unfolded than stable-folded proteins and compatible with the idea that the empty class I molecule may adopt an at least partially unfolded of molten globule-type structure (Bouvier and Wiley, 1998; Springer et al., 1998). Binding of a peptide leads to a partial refolding of the α-helical segments and consequently to an enhanced coupling of the motions of the α-helices that form the binding cleft.

The binding region that contacts the peptide N-terminus and the valine anchor residue of the bound peptide was found to adopt a conformation during the simulation of the empty form closer to the conformation in the complex with the peptide than the binding pocket for the peptide C-terminus (for the simulations at 300 K). It was observed for both independent long MD simulations on the empty α_1-α_2 domain. It indicates that the binding region for the peptide N-terminus is more likely to adopt a ''preformed'' pocket (in the empty form) and requires less ''induced fit'' motions for peptide binding than the region that interacts with the peptide C-terminus in the complex. Binding of the peptide C-terminus requires larger conformational transitions of the protein receptor. This simulation result is also supported by the fact that the peptide N-terminal section of the α_2 helix is tied to the β-sheet by the C101-C164 disulfide bond. No such stabilization is present for any helical segment at the region that binds the peptide C-terminal part. It is further supported by structural studies of Khan et al. (2000) on an HLA class I peptide complex with an empty but structurally conserved peptide N-terminal binding site.

The differences in conformational flexibility found for the two major binding regions of the class I α_1-α_2 domain during the simulations at 300 K may allow to draw important conclusions on the mechanism of peptide loading and on protein segment-specific conformational stabilization of the class I molecule by components of the loading complex. Indeed, experimental studies have mapped the class I residues important for the interaction with accessory proteins for peptide loading to regions close to the binding region for the peptide C-terminus (substitution of residues 86, 128–136 abrogates association with the accessory proteins TAP, tapasin, and calreticulin; Lewis et al., 1996; Peace-Brewer et al., 1996; Lewis and Elliott, 1998; Yu et al., 1999; reviewed in Bouvier, 2003) suggesting that this region requires binding of accessory proteins to support peptide binding. The greater apparent flexibility of the binding site at the peptide C-terminus perhaps also explains why the refolding of denatured class I heavy chains in vitro can be supported by C-terminally but not by N-terminally extended peptides (Horig et al., 1999). The increased deformation of the binding region for the C-terminal end of the peptide in the simulations of the empty form can in part be attributed to an increased conformational flexibility and a conformational shift of residues 131–151 in the simulations of the empty class I α_1-α_2 domain. Interestingly, this part of available class I molecule structures (in complex with various peptides) shows the largest conformational variation (Smith et al., 1996; Elliott, 1997). Based on this observation it has been suggested by Elliott (1997) that this region may undergo a conformational shift upon peptide binding and this suggestion is supported by our simulation results. The same region has also been implicated in the recognition by tapasin that is a component of the peptide-loading complex (Ortmann et al., 1997; Yu et al., 1999; Barnden et al., 2000; Grandea and van Kaer, 2001; Harris et al., 2001). A possible role of tapasin binding to this flexible region could be to induce a transition to a geometry close to the structure in the peptide bound form or to stabilize a certain conformation or a subset of conformations that allows effective binding or anchoring of the C-terminus of a peptide ligand.

We thank A. Barthel, N. Riemann, and D. Roccatano for helpful discussions.

REFERENCES

Amadei, A., A. B. M. Linssen, and H. J. C. Berendsen. 1993. Essential dynamics of proteins. *Proteins.* 17:412–425.

Barnden, M. J., A. W. Purcell, J. J. Gorman, and J. McCluskey. 2000. Tapasin-mediated retention and optimization of peptide ligands during the assembly of class I molecules. *J. Immunol.* 165:322–330.

Bouvier, M., and D. C. Wiley. 1994. Importance of peptide amino and carboxyl termini to the stability of MHC class I molecules. *Science.* 265:398–402.

Bouvier, M., and D. C. Wiley. 1998. Structural characterization of a soluble and partial folded class I major histocompatibility chain/b2m heterodimer. *Nat. Struct. Biol.* 5:377–384.

Bouvier, M. 2003. Accessory proteins and the assembly of human class I MHC molecules: A molecular and structural perspective. *Mol. Immunol.* 39:697–706.

Cornell, W. D., P. Cieplak, C. I. Bayley, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman. 1995. A second generation force field for simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.* 117:5179–5197.

Darden, T., D. York, and L. Pedersen. 1993. Particle mesh Ewald: an NlogN method for Ewald sums in large systems. *J. Chem. Phys.* 98:10089–10092.

Ellgaard, L., and A. Helenius. 2003. Quality control in the endoplasmic reticulum. *Nature Reviews-Mol. Cell Biol.* 4:181–191.

Elliott, T. 1997. How does TAP associate with MHC class I molecules? *Immunol. Today.* 18:375–379.

Fahnestock, M. L., J. L. Johnson, R. M. Feldman, T. J. Tsomides, J. Mayer, L. O. Narhi, and P. J. Bjorkman. 1994. Effects of peptide length and composition on binding to an empty class I MHC heterodimer. *Biochemistry.* 33:8149–8158.

Grandea 3rd, A. G., and L. Van Kaer. 2001. Tapasin: an ER chaperone that controls MHC class I assembly with peptide. *Trends Immunol.* 22:194–199.

Harris, M. R., L. Lybarger, N. B. Myers, C. Hilbert, J. C. Solheim, T. H. Hansen, and Y. Y. Yu. 2001. Interactions of HLA-B27 with the peptide loading complex as revealed by heavy chain mutations. *Int. Immunol.* 13:1275–1282.

Hillig, R. C., P. G. Coulie, V. Stroobant, W. Saenger, A. Ziegler, and M. Hulsmeyer. 2001. High-resolution structure of HLA-A*0201 in complex with a tumor-specific antigenic peptide encoded by the MAGE-A4 gene. *J. Mol. Biol.* 310:1167–1176.

Horig, H., A. C. Young, N. J. Papadopoulos, T. P. DiLorenzo, and S. G. Nathenson. 1999. Binding of longer peptides to the H-2Kb heterodimer is restricted to peptides extended at their C-terminus: refinement of the inherent MHC class I peptide-binding criteria. *J. Immunol.* 163:4434–4441.

Jones, E. Y., J. Tormo, S. W. Reid, and D. I. Stuart. 1998. Recognition surfaces of MHC class I. *Immunol. Rev.* 163:121–128.

Jorgensen, W. L., J. Chandrasekhar, J. Madura, R. W. Impey, and M. L. Klein. 1983. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79:926–935.

Khan, A. R., B. M. Baker, P. Ghosh, W. E. Biddison, and D. C. Wiley. 2000. The structure and stability of an HLA-A*0201/octameric tax peptide complex with an empty conserved peptide-N-terminal binding site. *J. Immunol.* 164:6398–6405.

Lewis, J. W., and T. Elliott. 1998. Evidence for successive peptide binding and quality control stages during MHC class I assembly. *Curr. Biol.* 8:717–720.

Lewis, J. W., A. Neisig, J. Neefjes, and T. Elliott. 1996. Point mutations in the α_2 domain of HLA-A.1 define a functionally relevant interaction with TAP. *Curr. Biol.* 6:873–883.

Madden, D. R. 1995. The three-dimensional structure of peptide-MHC-complexes. *Annu. Rev. Immunol.* 13:587–622.

Maenaka, K., and E. Y. Jones. 1999. MHC superfamily structure and the immune system. *Curr. Opin. Struct. Biol.* 9:745–753.

Matsumura, M., D. H. Fremont, P. A. Peterson, and I. A. Wilson. 1992. Emerging principles for the recognition of peptide antigens by MHC class I molecules. *Science.* 257:927–934.

Meng, W. S., H. von Grafenstein, and I. S. Haworth. 1997. A model of water structure inside the HLA-A2 peptide binding groove. *Int. Immunol.* 9:1339–1346.

Meng, W. S., H. von Grafenstein, and I. S. Haworth. 2000. Water dynamics at the binding interface of four different HLA-A2-peptide complexes. *Int. Immunol.* 12:949–957.

Ortmann, B., J. Copeman, P. J. Lehner, B. Sadasivan, J. A. Herberg, A. G. Grandea, S. R. Riddell, R. Tampe, T. Spies, J. Trowsdale, and P. Cresswell. 1997. A critical role for tapasin in the assembly and function of multimeric MHC class I-TAP complexes. *Science.* 277:1306–1309.

Peace-Brewer, A. L., L. G. Tussey, M. Matsui, G. Li, D. G. Quinn, and J. A. Frelinger. 1996. A point mutation in HLA-A*0102 results in failure to bind the TAP complex and to present virus-derived peptides to CTL. *Immunity.* 4:505–516.

Pearlman, D. A., D. A. Case, J. W. Caldwell, W. S. Ross, T. E. Cheatham, S. Debolt, D. Ferguson, G. Seibel, and P. A. Kollman. 1995. *Comput. Phys. Commun.* 91:1–41.

Petrone, P. M., and A. E. Garcia. 2004. MHC-peptide binding is assisted by bound water molecules. *J. Mol. Biol.* 338:419–435.

Pohlmann, T., R. A. Bockmann, H. Grubmüller, B. Uchanska-Ziegler, A. Ziegler, and U. Alexiev. 2004. Differential peptide dynamics is linked to major histocompatibility complex polymorphism. *J. Biol. Chem.* 279:28197–28201.

Rigney, E., M. Kojima, A. Glithero, and T. Elliott. 1998. A soluble major histocompatibility complex class I peptide-binding platform undergoes a conformational change in response to peptide epitopes. *J. Biol. Chem.* 273:14200–14204.

Rognan, D., L. Scapozza, G. Folkers, and A. Daser. 1994. Molecular-dynamics simulation of MHC-peptide complexes as a tool for predicting potential T-cell epitopes. *Biochemistry.* 33:11476–11485.

Rognan, D., N. Zimmermann, G. Junk, and G. Folkers. 1992. Molecular-dynamics study of a complex between the human histocompatibility antigen HLA-A2 and the Imp58–66 nona-peptide from influenza-virus matrix protein. *Eur. J. Biochem.* 208:101–113.

Sherer, E. C., S. A. Harris, R. Soliva, M. Orozco, and C. A. Laughton. 1999. Molecular dynamics studies of DNA A-tract structure and flexibility. *J. Am. Chem. Soc.* 121:5981–5991.

Smith, K. J., S. W. Reid, D. I. Stuart, A. J. McMichael, E. Y. Jones, and J. I. Bell. 1996. An altered position of the alpha 2 helix of MHC class I is revealed by the crystal structure of HLA-B*3501. *Immunity.* 4:203–213.

Springer, S., K. Doring, J. C. A. Skipper, A. R. M. Townsend, and V. Cerundolo. 1998. Fast association rates suggest a conformational change in the MHC class I molecule H-2D^b upon peptide binding. *Biochemistry.* 37:3001–3012.

Wang, J., and P. A. Kollman. 2001. Automatic parameterization of force field by systematic search and genetic algorithms. *J. Comput. Chem.* 22:1219–1228.

Wilson, I. A., and P. J. Bjorkman. 1998. Unusual MHC-like molecules: CD1, Fc receptor, the hemochromatosis gene product, and viral homologs. *Curr. Opin. Immunol.* 10:67–73.

Wlodek, S. T., T. W. Clark, L. R. Scott, and J. A. McCammon. 1997. Molecular dynamics of acetylcholinesterase dimmer complexed with tacrine. *J. Am. Chem. Soc.* 121:5981–5991.

Yu, Y. Y., H. R. Turnquist, N. B. Myers, G. K. Balendiran, T. H. Hansen, and J. C. Solheim. 1999. An extensive region of an MHC class I alpha 2 domain loop influences interaction with the assembly complex. *J. Immunol.* 163:4427–4433.

Zacharias, M. 2000. Comparison of molecular dynamics and harmonic mode calculations on RNA. *Biopolymers.* 54:547–560.

Zajonc, D. M., M. A. Elsliger, L. Teyton, and I. A. Wilson. 2003. Crystal structure of CD1a in complex with a sulfatide self antigen at a resolution of 2.15 Å. *Nature Immunol.* 4:808–815.