doi:10.1061/S0022-2836(02)00563-6 available online at http://www.idealibrary.com on IDEAL® J. Mol. Biol. (2002) 320, 1109–1117

JMB

A Structure for the Trimeric MHC Class II-associated Invariant Chain Transmembrane Domain

Andreas Kukol[1]*, Jaume Torres[2] and Isaiah T. Arkin[3]

[1]*Department of Biological Sciences, University of Warwick, Coventry CV4 7AL UK*

[2]*School of Biological Sciences Nanyang Technological University, 1 Nanyang Walk Block 5, Level 3, Singapore 637616, Singapore*

[3]*The Alexander Silberman Institute of Life Sciences Department of Biological Chemistry, The Hebrew University, Givat-Ram Jerusalem 91904, Israel*

**Corresponding author*

The major histocompatibility complex (MHC)-associated invariant chain (Ii) contains a single transmembrane domain that forms trimers. Ii is involved in the assembly of the MHC and antigen presentation, and is thus central to the function of the immune system. Here, we show by attenuated total reflectance, Fourier transform infrared (ATR–FTIR) spectroscopy that the transmembrane domain is α-helical and we provide a structural model of the transmembrane domain obtained by a combination of site-specific infrared dichroism and molecular dynamics (MD) simulations. This work resolves the backbone structure of a transmembrane peptide by multiple $^{13}C{=}^{18}O$ labelling at ten different residues. A second purely computational approach, based on MD simulations of Ii transmembrane homologous sequences, yields a similar structure that is consistent with our experimental results. The structure presented forms a left-handed coiled coil with an average helix tilt of $13(\pm 6)°$; the residue Gln47 implicated in trimer formation forms strong interhelical contacts, Thr50 points to the inside of the trimeric coil and forms a network of hydrogen bonds.

© 2002 Elsevier Science Ltd. All rights reserved

Keywords: associated invariant chain; FTIR spectroscopy; molecular dynamics; site-specific dichroism; transmembrane protein

Introduction

The MHC class II-associated invariant chain (Ii) is an accessory protein involved in the pathway of MHC class II maturation and peptide loading.[1] The 279 residue Ii contains a single transmembrane domain and associates after biosynthesis to form a trimer.[2,3] This trimeric complex serves as a scaffold for the binding of three MHC class II αβ heterodimers.[4] The assembled $(\alpha\beta)_3Ii_3$ complex is exported through the endocytic pathway. In the late endocytic compartment, Ii is degraded progressively by proteases, which leaves only the residues 81–104 known as CLIP peptides remaining in the antigen-binding groove. The CLIP peptides are finally released by HLA-DM,[5] which remains associated with MHC until the loading with antigenic peptides occurs.[6]

Trimerisation of Ii is mediated by two different regions of the polypeptide chain: (1) residues 163–183;[7,8] and (2) the transmembrane domain, residues 31–54.[9] The available data can be interpreted such that the C-terminal domain initiates trimerisation and, once a trimer is formed, it remains stable even after cleavage of the C-terminal domain. However, a truncated version of the first 80 residues of Ii has been found sufficient for trimerisation and a patch of hydrophilic residues in the transmembrane domain has been shown to be important for trimerisation.[9]

Taken together, it is highly likely that the transmembrane domain on its own assembles into a native-like structure resembling the transmembrane part of the whole Ii molecule.

This work uses infrared spectroscopy to confirm for the first time the predicted α-helical structure and, more importantly, to obtain a detailed molecular model of the transmembrane domain.

Since structure determination of membrane proteins is still a difficult task for X-ray crystallography and NMR spectroscopy, a new, yet already proven successful approach is undertaken here to obtain a structural model for the Ii transmembrane

Abbreviations used: MHC, major histocompatibility complex; ATR, attenuated total reflectance; FTIR, Fourier transform infrared; MD, molecular dynamics; SSID, site-specific infrared dichroism; HIV-I, human immunodeficiency virus type I; DMPC, dimyristoylphosphocholine; Iitp, Ii transmembrane peptide.

E-mail address of the corresponding author: akukol@bio.warwick.ac.uk

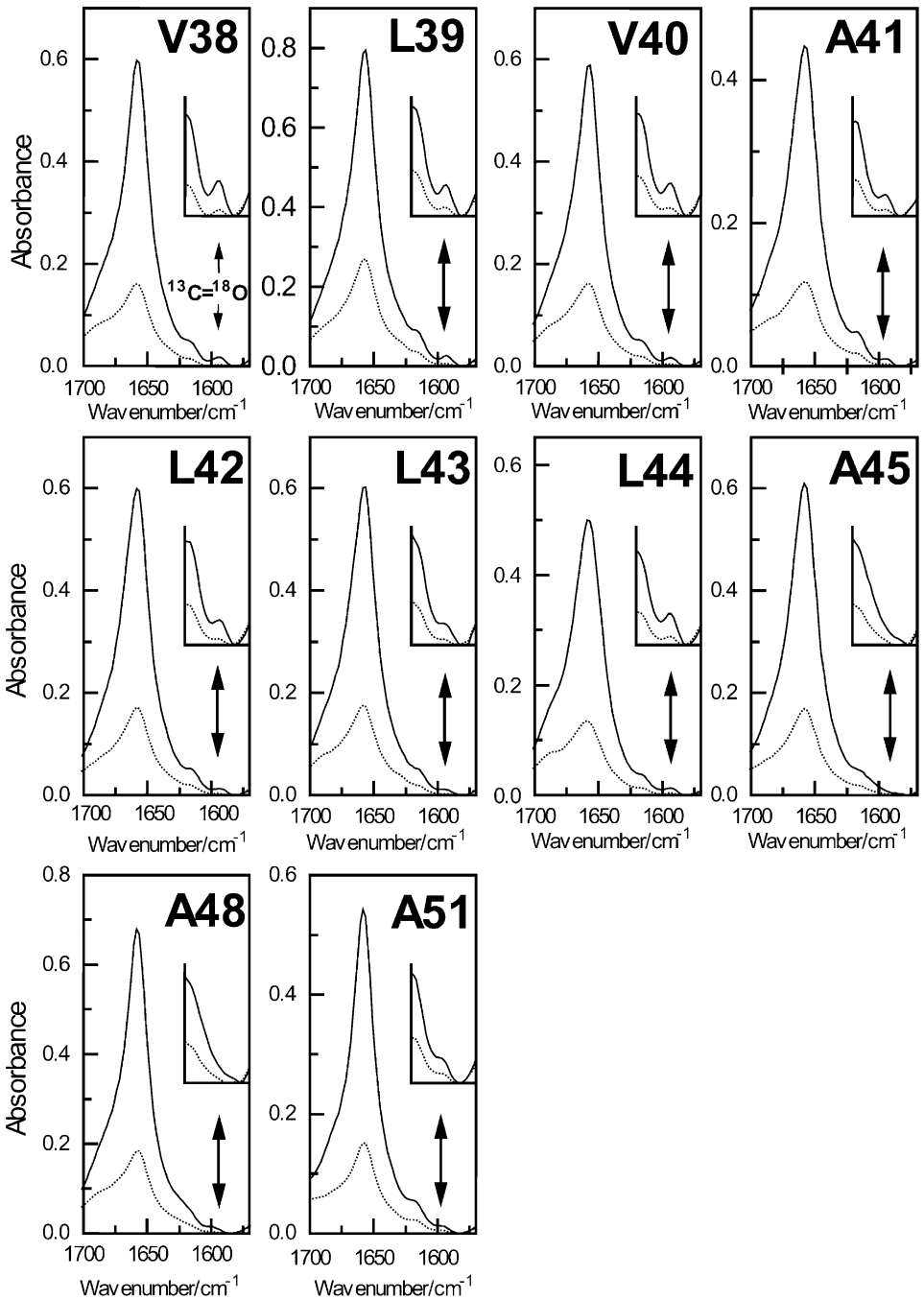

Figure 1. Representative ATR–FTIR-spectra of the lipid reconstituted Ii transmembrane peptide for all ten different labels, at 0° (continuous line) and 90° (broken line) polarisation. The inset shows the region with the $^{13}C{=}^{18}O$ amide I absorption band vertically expanded.

domain trimer. The method of site-specific infrared dichroism (SSID)[10] is applied to an isotopically labelled synthetic transmembrane peptide to yield helix tilt and rotation defining the interaction surface of the helices in the bundle. In the case of the transmembrane domain of glycophorin A[10] and the M2 protein,[11] orientational data have been obtained that are in agreement with NMR data. Introducing this data into molecular dynamics (MD) calculations as energy constraints yields a detailed structural model. This method has been applied to other transmembrane proteins, i.e. the M2 protein of the influenza A virus,[12] the vpu protein of HIV-I virus,[13] the CM2 protein of influenza C virus[14] and phospholamban.[15] Here, an extension of previous applications is described that makes use of $^{13}C{=}^{18}O$ labelled amino acid residues[15,16] at each possible position in the transmembrane sequence, thus yielding a more complete picture of the backbone structure compared to previous examples.

Independently, a different approach has been employed, utilising homologous sequences of the Ii transmembrane domain to select the correct

Table 1. Rotational pitch angle ω and local helix tilt β, determined for all $^{13}C{=}^{18}O$ labelled residues

	RGALYTGVS**VLVALLLA**GQ**A**TT**A**YFLY**QQQ**GR			
Label	ω (deg.)	β (deg.)	*n*	θ (deg.)
V38	290 ± 10	10 ± 2	6	19.5
L39	359 ± 10	10 ± 2	3	12.1
V40	108 ± 12	6 ± 2	4	24.2
A41	216 ± 10	8 ± 2	3	28.3
L42	316 ± 14	14 ± 4	7	13.2
L43	359 ± 12	7 ± 2	3	15.1
L44	83 ± 13	21 ± 5	5	24.4
A45	133 ± 14	25 ± 8	10	39.3
A48	299 ± 20	15 ± 7	12	16.9
A51	239	15		34.1[a]

n is the number of experiments for each label. From these values, the angle between the C=O bond and the *z*-axis, θ, has been calculated. The peptide sequence with labelled residues is shown in boldface. Note that each peptide contains only one label.

[a] Calculated from A48 assuming $\Delta\omega = 300°$.

model.[17] In this approach, MD simulations are carried out for each of the homologous sequences. This leads to several possible structures for each sequence. However, only one structure is common to all simulations, and this is selected as the correct model. Encouragingly, both methods result in a similar structure.

Results and Discussion

Secondary structure and transmembrane orientation

Typical attenuated total reflectance (ATR) infrared spectra of the Ii transmembrane peptide (Iitp) reconstituted in dimyristoylphosphocholine (DMPC) vesicles are shown in Figure 1 at 0° and 90° polarisation. The shape of the amide I absorption band centred at 1658 cm^{-1} is indicative of a highly α-helical protein[18] with the helix oriented perpendicular to the plane of the bilayer membrane. The orientation is confirmed by integration of the amide I absorption bands at 0° and 90° polarisation. Typically the dichroic ratio $R = A_{\parallel}/A_{\perp}$ of the amide I varies between $R = 2.6$ and $R = 4.1$, which is caused by a variation in sample order. Classical analysis of dichroic ratios[19] yields order parameters between $S = 0.36$ and $S = 4.1$, which correspond to helix tilts between 41° and 12°, respectively, if a completely ordered sample is assumed. Note that in SSID the sample order is taken into account and samples with any order above zero can be analysed to yield the absolute helix tilt. The small bands centred at 1618 cm^{-1} and 1594 cm^{-1} correspond to the $^{13}C{=}^{16}O$ and $^{13}C{=}^{18}O$ label in the positions indicated in Figure 1. In all cases, only part of ^{16}O has exchanged to ^{18}O, which does not restrict the analysis as long as the $^{13}C{=}^{18}O$ amide I band can be integrated. In fact, for nine out of ten labelled peptides the $^{13}C{=}^{18}O$

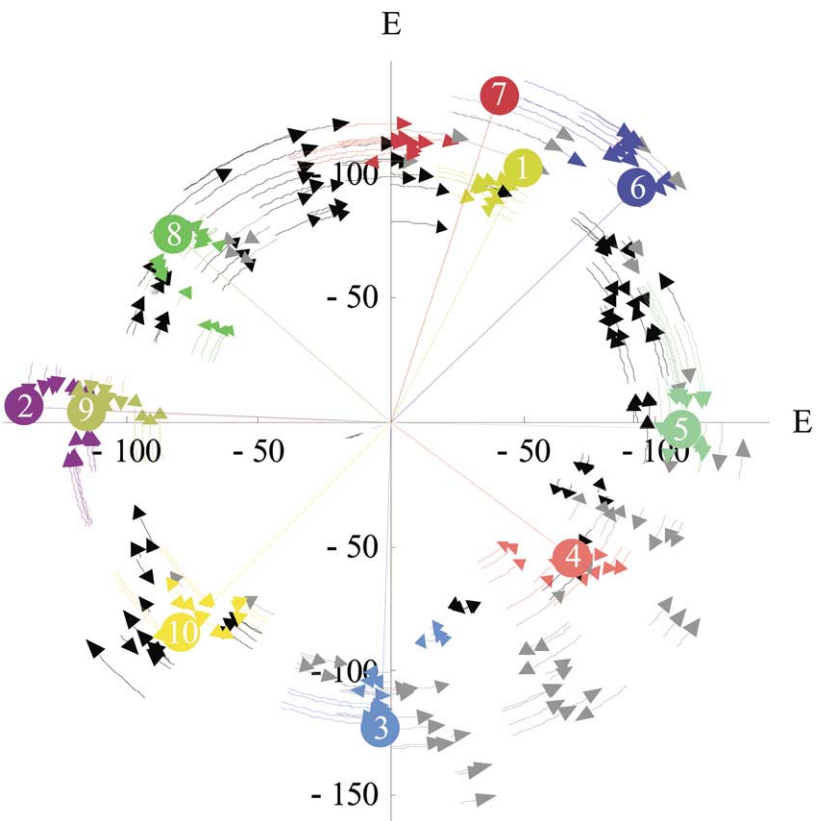

Figure 2. Energy and helix rotation parameter ϕ for the 288 structures of a global molecular dynamics search shown in a polar plot. The arcs represent each individual structure, where the position on the circle indicates ϕ and the distance from the centre indicates the negative energy. Arrows indicate how the structure ϕ changed during simulation. Structures are coloured according to their cluster affiliation, while the cluster averages are shown as circles connected to the origin by azimuthal lines.

Table 2. Rotational pitch angles ω for all structures resulting from the constrained molecular dynamics search

	Structure										
	1	2	3	4	5	6	7	8	9	10	Exp.
A. *Dimer search*											
V38	19	73	324	8.6	35	104	211				290
L39	110	111	78	113	141	285	344				359
V40	207	308	148	210	230	353	34				108
A41	280	2.7	239	298	322	46	78				216
L42	15	75	333	45	85	96	347				316
L43	132	281	98	137	177	297	4.1				359
L44	246	331	190	252	278	349	48				83
A45	318	36.5	305	334	8.0	71	269				133
A48	272	5.6	239	279	310	25	73				299
A51	209	308	182	204	258	325	29				239
Deviation[a]	338	330	260	345	386	359	304				0
B. *Trimer search*											
V38	203	290	30	90	339	9.1	47	92	292	176	290
L39	300	31	151	216	66	102	143	213	60	323	359
V40	52	169	227	294	137	203	233	321	–	26	108
A41	147	251	307	4.6	243	301	330	28	–	88	216
L42	260	331	81	109	342	47	92	97	–	235	316
L43	344	77	183	267	81	136	187	225	–	336	359
L44	83	142	272	336	191	250	282	350	–	47	83
A45	198	291	7.2	44	300	340	16	50	–	117	133
A48	121	246	308	2	222	297	306	349	–	74	299
A51	69	197	240	299	199	214	216	287	–	16	239
Deviation[a]	295	213	384	397	251	335	389	412	–[b]	288	0
C. *Tetramer search*											
V38	18	39	119	214	289	191	27	94	195		290
L39	113	168	292	320	32	322	130	189	339		359
V40	208	250	3.4	50	151	32	228	306	36		108
A41	285	329	54	147	246	94	320	28	91		216
L42	20	93	292	282	333	311	69	110	272		316
L43	128	212	41	5.5	61	355	163	230	13		359
L44	255	301	12	78	157	50	274	343	50		83
A45	329	21	106	216	285	121	8.1	51	113		133
A48	280	332	56	117	228	71	299	13	72		299
A51	220	274	23	75	154	41	248	290	37		239
Deviation[a]	353	441	336	287	226	321	339	420	277		0

[a] The deviation between the simulated ω_s and experimental values ω_e is calculated according to: $D = \sqrt{\sum_{i=1}^{10} (\omega_{s,i} - \omega_{e,i})^2}$.
[b] The rotational pitch angles of structure 9 could not be determined, because of distortions.

amide band could be integrated without the need for resolution enhancement techniques, like Fourier self-deconvolution, which can introduce artefacts into the results. Furthermore, in all cases the band appears at the same position, indicating that all residues chosen as labels are in an α-helical environment.[16] For Ala45 the $^{13}C{=}^{16}O$ amide I band has been integrated after Fourier self-deconvolution.

Quantitative analysis of spectra from adjacent labels according to the theory of SSID, as detailed in Materials and Methods, yields the orientation of each label in terms of rotational pitch angle and local helix tilt shown in Table 1. The average helix tilt is 13(±6)°, with a tendency to higher helix tilts from Leu44 onwards.

Constraint global molecular dynamics search

To find a detailed molecular structure that represents the experimentally determined orientation and corresponds to an energy minimum of an MD simulation, a global search has been carried out,[20] searching all possible interaction surfaces between helices in a trimeric bundle employing the orientational energy refinement in the form of constraints for the angles between the C=O bond and the *z*-axis for all labelled residues. From the search process, described in Materials and Methods, ten energy-minimised structures have been found. The distribution of all structures as a function of energy and rotational search parameter ϕ is shown in Figure 2. The orientations of the labelled sites are compared with the experimental values in Table 2. Note that, although the experimental data have been applied as additional energy refinement terms, energy minima at different orientations have been found, because due to CPU time limitations, MD simulations normally find only local minima and have difficulty overcoming large energy barriers involving the rotation of helices as a whole by angles

more than $\phi = 30°$ to 40°, as can be seen in Figure 2. The experimental constraints are fulfilled in all these structures, because slight distortions from the α-helical geometry can produce a variation of angles between the C═O bond and the *z*-axis. However, comparing simulated structures and experiment with regard to the rotational pitch angle, structure 2 emerges as the one closest to the experimental data as well as possessing the lowest energy. The corresponding right-handed model is structure 9, as can be seen from Figure 2. But, due to distortions from helical geometry, accurate determination of the rotational pitch angles was impossible. In addition, structure 2 has a lower energy than structure 9. Although SSID alone cannot decide between left and right-handed structures, it can be safely concluded from a comparison between simulation and experiment that Iitp forms a left-handed coiled coil trimer.

Molecular dynamics search of functional mutant sequences

In this approach, a similar global MD search has been carried out for a number of homologous sequences, but without experimental constraints. Structures resulting from simulations with different sequences are compared to each other with regard to their backbone rmsd. The structure chosen as the best model in this case is the structure that is common to all simulations performed with different sequences. The reason is, that all sequences form the same native structure, and therefore this structure must be found in all of the simulations. Non-native structures found in the simulations need not persist in all of the sequences, since the sequence variation (also called silent substitutions) are silent only with respect to the native structure.

Figure 3 shows that there is only one structure (left-handed, see open squares), at $\omega_{42} \approx -10°$, appearing in all simulations. In this case, the structures of different simulations differ from each other by less than 1.0 Å backbone rmsd. Two other sets were found at $\omega_{42} \approx 0°$ (right-handed, grey squares) and $\omega_{42} \approx 100°$ (right-handed, grey squares), but the rmsd between the structures is greater than 1.2 Å. Therefore, the left-handed structure at $\omega_{42} \approx 0°$ is chosen as the best model of Iitp, resulting from this approach.

Comparison of the two models

Calculation of the backbone rmsd between the two structures obtained independently from experimental and evolutionary data yields 1.1 Å. Since each structure has been obtained from a different approach, the two structures are assumed to be identical at the level of accuracy that can be achieved with the methods employed here. Although SSID enables us to select the native structure from an outcome of a global search, the closeness of structures from two different approaches provides a further strong argument that the correct structure has been found. For the following discussion, the model obtained from orientational data is chosen. We believe that it is more accurate in detail, because orientational constraints have been applied during the simulation.

Reliability of the model

The viability of the approach using a transmembrane peptide as a model of the transmembrane domain of a 279 residue protein can be questioned. Therefore, in this work, two alternative approaches have been employed. In addition to orientational data obtained from the transmembrane peptide, homologous transmembrane sequences have been used in simulations. The assumption that the structure adopted by the transmembrane domain is similar to that of the transmembrane part of the whole protein is implicitly contained in these variant sequences, because otherwise they would not be found in other organisms. Since both approaches led to the same model, it can be safely concluded that the structure presented here is a reliable model for the transmembrane domain of the MHC-associated invariant chain (Ii). In addition, infrared spectroscopic data show that the transmembrane peptide exists in an oligomeric form and not as a monomer, because the site dichroism R_{site} is always different from R_{helix}. Similar dichroisms would indicate free rotation of the helix, since the dichroism is an average over all orientations in the lipid bilayer.

Structural implications

Figure 4 shows the structure of Iitp, with the residues Tyr33, Gln47, Thr49 and Thr50 highlighted. A triple mutant of Ii at the position Gln47, Thr49 and Thr50 has been shown to fail to assemble with MHC class II because of a failure of Ii trimerisation.[9] From the structure shown in Figure 4, it is clear that Gln47 and Thr50 interact with each other and may stabilise the trimer structure. Thr49 points to the lipid phase and forms, as would be expected, intrahelical hydrogen bonds to main-chain carbonyl oxygen atoms at position i-4 in the sequence. It can be predicted that mutation of Gln47 and Tyr50 only will be sufficient to disrupt trimerisation of Ii. It has been shown recently that polar residues, capable of serving simultaneously as hydrogen bond donor and acceptor (Asn, Asp, Gln, Glu, His), may stabilise transmembrane helix association.[21] The role of Gln47 in our model can be understood in the light of the above study, whereas Thr has been shown not to facilitate helix–helix association. Contrary to that, our model provides an interesting example of the involvement of Thr in helix–helix association.

Other residues involved in stabilising the trimer structure are Tyr33 and Leu43, which are

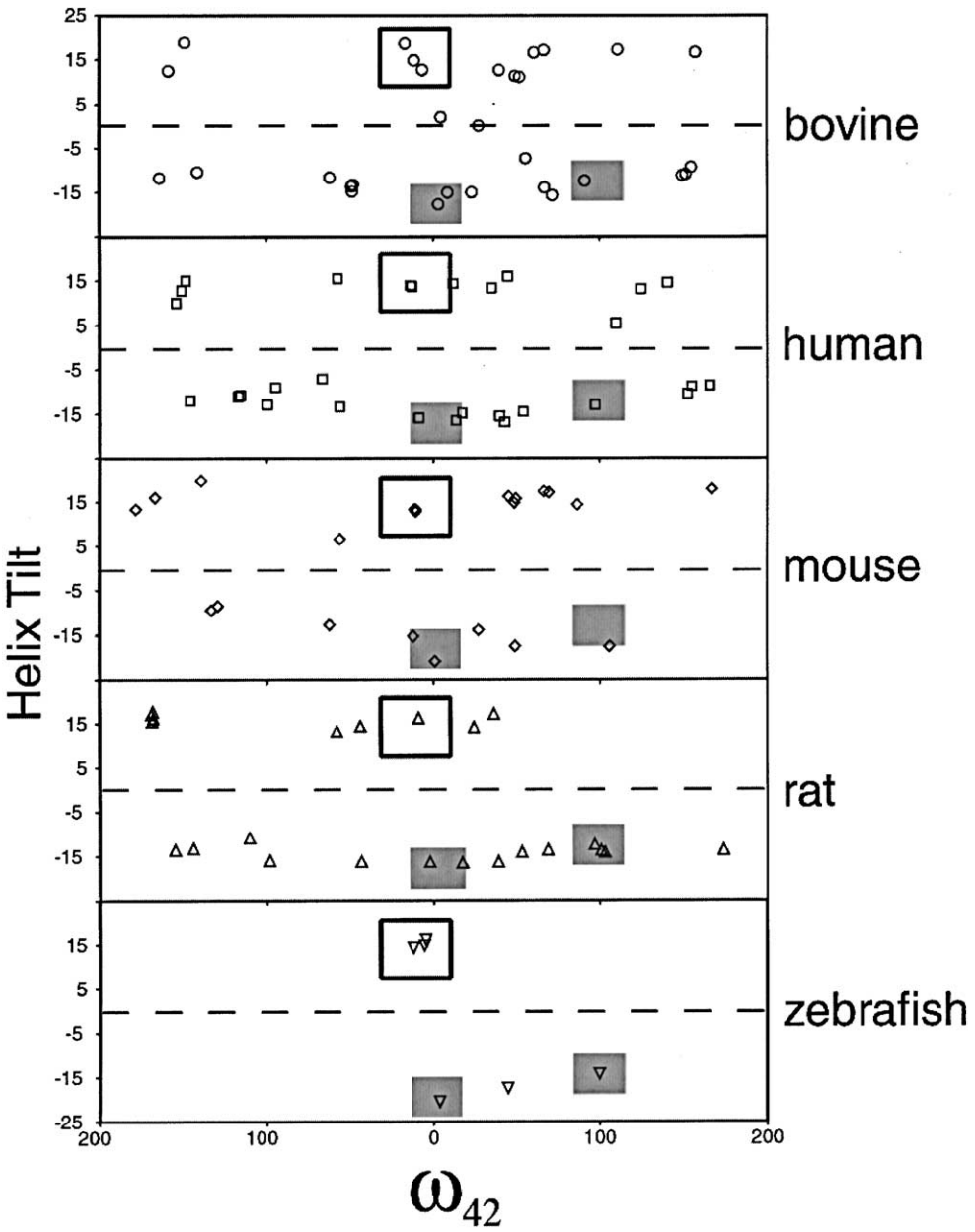

Figure 3. Structures obtained from the simulation of homologous sequences. The obtained cluster averages are shown as a function of helix tilt and rotational pitch angle corresponding to residue 42, ω_{42}. The horizontal broken line in each case separates left-handed bundles (above the line) from right-handed ones (below the line). Open squares indicate the closest model that persists in all simulations. Grey squares indicate models that also persist, but where structures are more disperse (see the text).

highlighted in Figure 4(a). Minor interactions exist between Val36 and Val40 (not shown).

Materials and Methods

Peptide purification

Synthetic peptides corresponding to the predicted transmembrane domain of Ii (mouse) and a few adjacent amino acid residues to reduce the hydrophobic character were made by solid-phase FMOC (*N*-(9-fluorenyl)methoxycarbonyl) chemistry, cleaved from the resin with trifluoroacetic acid, and lyophilised. Ten peptides corresponding to residues 29–60 of Ii with the following sequence RGALYTGVSVLVALLLAGQATTAYFLYQQQGR were synthesised, each containing one 1-$^{13}C{=}^{18}O$ labelled amino acid residue in positions 38, 39, 40, 41, 42, 43, 44, 45 and 48. The $^{13}C{=}^{18}O$ labelled amino acid residues have been prepared from 1-^{13}C labelled amino acid residues (Cambridge Isotopes Laboratories, Andover, MA) by acid-catalysed oxygen exchange in $H_2^{18}O$ as described.[17] The peptides were further purified as described[12] for analogous transmembrane peptides. Briefly, the peptide was dissolved in trifluoroacetic acid and purified by reversed phase chromatography (Jupiter 5C4-300 Å column, Phenomenex, Cheshire, UK). Peptide elution was achieved

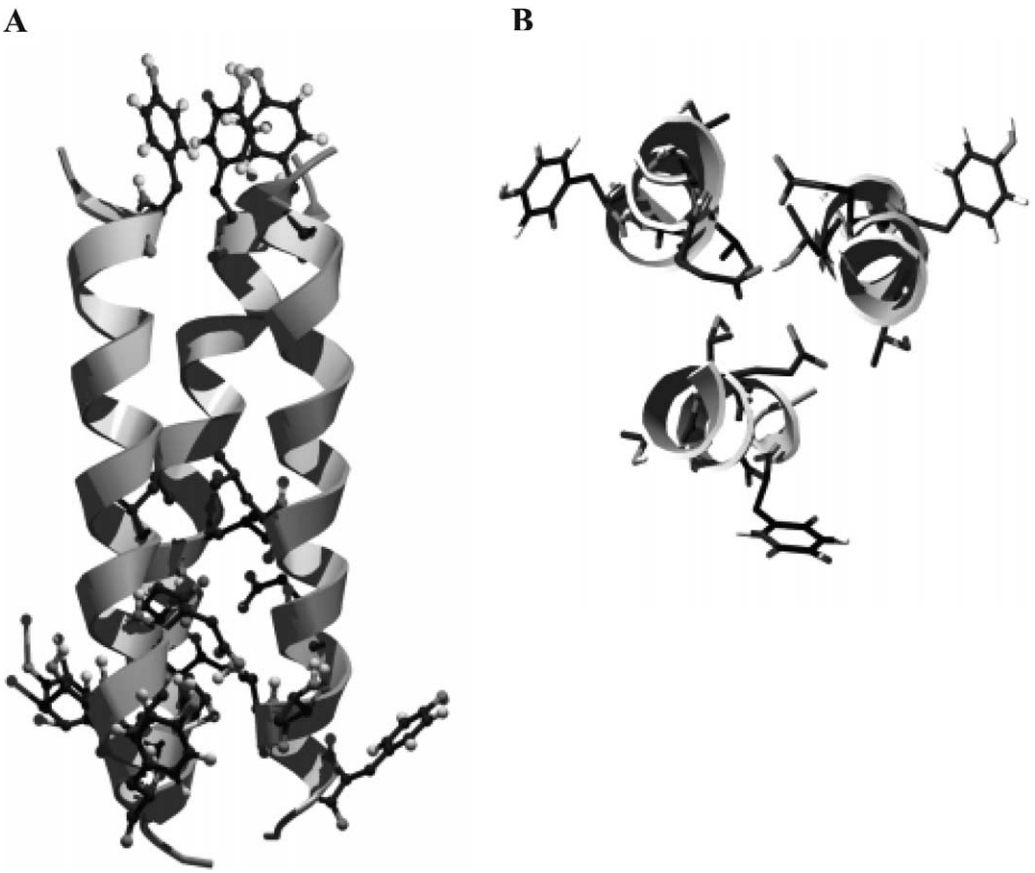

Figure 4. Model of the Ii transmembrane peptide. (a) The whole structure is shown with the residues Tyr33, Gln47, Thr49, Thr50 and Tyr52 in ball-and-stick representation. (b) View of the patch of hydrophilic residues implicated in trimerisation of the helices. Gln47, Thr49, Thr50 and Tyr52 are shown in ball-and-stick representation. The Figure was created with Swiss-PDB viewer.[27]

with a linear gradient to a final solvent composition of 5% (v/v) H_2O, 38% (v/v) acetonitrile and 57% (v/v) 2-propanol. All solvents contained 0.1% (v/v) trifluoroacetic acid. After lyophilisation of the pooled fractions the peptide were dissolved in 1,1,1,3,3,3-hexafluoro-2-propanol, containing dimyristoylphophocholine (Sigma), to a peptide to lipid ration of 1:25 (w/w).

Infrared spectroscopy

Fourier transform infrared (FTIR) spectra were recorded on a Nicolet Magna 560 spectrometer (Nicolet, Madison, WI) equipped with a high-sensitivity liquid nitrogen-cooled MCT/A detector. Attenuated total reflection (ATR)–FTIR spectra were measured with a 25 reflections ATR accessory from Graseby Specac (Kent, UK) and a wire grid polariser (0.25 μm, Graseby Specac). A sample (100 μl) of the peptide/lipid solution in hexafluoroisopropanol was dried onto a germanium trapezoidal internal reflection element (50 mm × 2 mm × 20 mm) under a stream of nitrogen. After the solvent was removed, the lipid/protein film was hydrated for 12 hours with a water-saturated stream of nitrogen. Dichroic ratios similar to those for lipid vesicle preparations were achieved through this procedure (data not shown). After removal of bulk water, an FTIR spectrum was obtained from 1000 averaged interferograms and single-point zero filling Happ–Genzel apodisation. The dichroic ratio, R, was calculated as the ratio between the integrated absorption of parallel and perpendicular polarised light of the absorption bands (between 1670 and 1645 cm^{-1} for the $^{12}C{=}^{16}O$ amide I, and between 1600 and 1583 cm^{-1} for the $^{13}C{=}^{18}O$ amide I mode).

Data analysis

The data were analysed according to the theory of site-specific dichroism presented in detail elsewhere.[10] Briefly, the measured dichroic ratio, the absorption between parallel and perpendicular polarised light $R = A_{\parallel}/A_{\perp}$ of a particular transition dipole moment is a function of its spatial orientation. For the amide I mode (mainly the C=O bond vibration) of an α-helical protein, the geometric relation between the transition dipole moment and the helix is known from fibre diffraction studies.[22] Therefore, by measuring the orientation of the amide I transition dipole moment one can determine the helix tilt angle β and the rotational pitch angle ω of the specific dipole moment about the helix axis. The rotational pitch angle ω is defined arbitrarily as 0° when the transition dipole moment, the helix director, and the z-axis all reside in a single plane (Figure 5). Thus, measuring the site-specific dichroic ratio R_{site} of the $^{13}C{=}^{18}O$ amide I mode from a particular label and the helix dichroic ratio R_{helix} allows calculation of the helix tilt β and the rotational pitch angle ω of a particular label if measurements from two samples with labels at different positions are analysed together. The pair of

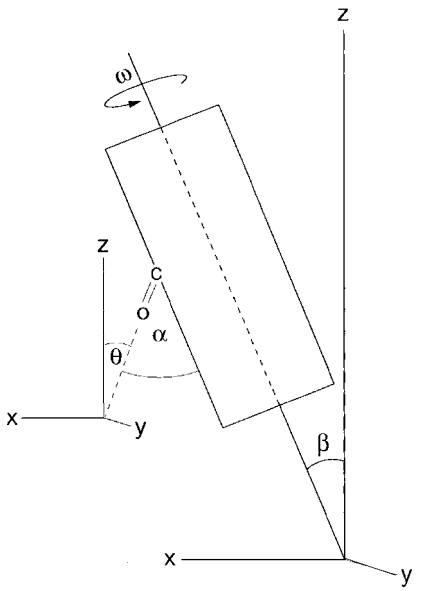

Figure 5. A diagram of a helix shown as rectangular shape, defining the helix tilt angle β and the rotational pitch angle ω in relation to the *z*-axis and to the transition dipole moment of the amide I mode (shown as C═O). The angle $\alpha = 141°$ is fixed in α-helix geometry and known from fibre diffraction studies. The angle θ is the absolute angle between the C═O bond and the *z*-axis; it is used as a constraint in MD simulations.

labels used in the analysis has been chosen such that they are closest in the sequence; i.e. V38/L39, L39/V40, V40/A41, A41/L42, L42/L43, L43/L44, L44/L45, L45/A48, A48/A51. For each pair of labels X_n/Y_{n+1} the calculation yields the rotational pitch angle and local helix tilt for the label in position *n*. The difference of ω between two consecutive residues is assumed to be 100° as in a canonical α-helix and 300° for the pairs L45/A48 and A48/A51.

Global molecular dynamics search

A global search with respect to rotation about the helix axis, assuming trimeric symmetry, was carried out as described.[12,20] In brief, all calculations were performed with the parallel processing version of the Crystallography & NMR System version 0.40[23] and CHI, using the OPLS parameter set with a united atom topology that explicitly represents the polar hydrogen and aromatic side-chain atoms.[24] All calculations were carried out *in vacuo* with the initial co-ordinates of a canonical α-helix (3.6 residues/turn). Symmetric trimers were generated from the sequence ALYTGVSVLVALLLAGQATTAYFL, acetylated at the N terminus and methylaminated at the C terminus, by replicating the helix and rotating it by 360°/3 around the centre of the trimer. An initial crossing angle of 25° for left-handed and −25° for right-handed structures was introduced by rotating the long helix axis with respect to the long bundle axis. The symmetric search was carried out by applying a rotation to all helices simultaneously between $\phi = 0°$ and $\phi = 360°$ in 10° steps. Four trials were carried out for each starting structure, using different random initial atom velocities in each case at both right-handed and left-handed crossing angles yielding $36 \times 4 \times 2$ equalling 288 structures. Each structure was subjected to a simulated annealing and energy minimisation protocol. Clusters of similar structures were defined such that the rmsd of the co-ordinates between all structures within a cluster was not larger than 1 Å; a cluster was formed by a minimum of ten structures. For each cluster, an average structure was calculated, energy minimised and subjected to the same simulated annealing protocol used in the systematic search.

During all simulations, orientation refinement energy terms have been incorporated as described.[12] In this case, for each of the ten amino acid residues, an orientational restraint has been applied by setting the angles between the $^{13}C{=}^{18}O$ bond and the *z*-axis to those obtained from the experiment. This angle is dependent on the rotational pitch angle and the helix tilt. The overall weight constant k_{dichro} for these orientational constraints was chosen to be $k_{dichro} = 1000$ kcal/rad^2 (1 cal = 4.184 J), determined empirically by selecting the minimum value necessary to obtain an identical outcome from MD protocols with different constants.

Molecular dynamics search of mutant sequences

The same global search protocol, as described above (without orientational restraints), was applied to sequences from different organisms. In this case, no orientational refinement energy terms have been applied. The following sequences have been used:Q9XRE4, LKITGLTVLA CLLLAGQALT AYMV (zebra fish); k01144, ALYTGFSILV TLLLAGQATT AYFL (human); jc4796, ALYTGFSVLV ALLLAGQATT AYFL (horse_pig_bovine); p04441, ALYTGVSVLV ALLLAGQATT AYFL (mouse); x13044, VLYTSVSVLV ALLLAGQATT AYFL (rat).

These sequences have been chosen as homologous to but not identical with the human sequence and represent all that could be found in the sequence database OWL.[25] The results from the global searching MD simulations were represented graphically by plotting each cluster representative (candidate models) as a function of two parameters, the helix tilt β, and the rotational orientation ω, as described.[17] The tilt angle of the model, β, was taken as the average of the angles between each helix axis and the *z*-axis, i.e. coincident with the normal to the bilayer and was calculated by CHI.[20,26] The helical axis is a vector with starting and end point above and below the defined residue, these points corresponding to the mean of the co-ordinates of the five α carbon atoms N-terminal and the five α carbon atoms C-terminal to the defined residue.

The rotational orientation ω is defined relative to an arbitrarily chosen, specified residue. The angle ω is defined by the angle between a vector perpendicular to the helix axis, oriented towards the middle of the peptide C═O bond of the residue, and a plane that contains both the helical axis and the normal to the bilayer. This angle is 0° when the residue is located in the direction of the tilt.[10] The structures identified were plotted against the ω angle of residue 42. Precise comparisons between similar clusters obtained from different variants were made by calculating the rmsd between its Cα backbones. In the simulations, the handness of the bundle is indicated by the helix tilt sign, positive or negative, which corresponds to left-handed and right-handed bundles, respectively. Simulations were performed without any restraint.

Acknowledgments

This work was supported by grants from the Welcome Trust and BBSRC to I.T.A.

References

1. Busch, R., Doebele, R., Patil, N. S., Pashine, A. & Mellins, E. D. (2000). Accessory molecules of MHC class II peptide loading. *Curr. Opin. Immunol.* **12**, 99–106.
2. Lamb, C. A. & Cresswell, P. (1992). Assembly and transport properties of invariant chain trimers and HLA-DR-invariant chain complexes. *J. Immunol.* **148**, 3478–3482.
3. Marks, M. S., Blum, J. S. & Cresswell, P. (1990). Invariant chain trimers are sequestered in the rough endoplasmic reticulum in the absence of association with HLA class II antigens. *J. Cell Biol.* **111**, 839–845.
4. Roche, P. A., Marks, M. S. & Cresswell, P. (1991). Formation of a nine-subunit complex by HLA class II glycoproteins and the invariant chain. *Nature*, **354**, 392–394.
5. Wolf, P. R., Tourne, S., Miyazaki, T., Benoist, C., Mathis, D. & Ploegh, H. L. (1998). The phenotype of H-2M-deficient mice is dependent on the MHC class II molecules expressed. *Eur. J. Immunol.* **28**, 2605–2618.
6. Denzin, L. K. & Cresswell, P. (1995). HLA-DM induces clip dissociation from MHC class II alpha-beta dimers and facilitates peptide loading. *Cell*, **82**, 155–165.
7. Bertolino, P., Staschewski, M., Trescol-Biemont, M. C., Freisewinkel, I. M., Schenck, K., Chretien, I. *et al.* (1995). Deletion of a C-terminal sequence of the class II-associated invariant chain abrogates invariant chains oligomer formation and class II antigen presentation. *J. Immunol.* **154**, 5620–5629.
8. Bijlmakers, M. J., Benaroch, P. & Ploegh, H. L. (1994). Mapping functional regions in the lumenal domain of the class II-associated invariant chain. *J. Expt. Med.* **180**, 623–629.
9. Ashman, J. B. & Miller, J. (1999). A role for the transmembrane domain in the trimerization of the MHC class II-associated invariant chain. *J. Immunol.* **163**, 2704–2712.
10. Arkin, I. T., MacKenzie, K. R. & Brünger, A. T. (1997). Site-directed dichroism as a method for obtaining rotational and orientational constraints for orientated polymers. *J. Am. Chem. Soc.* **119**, 8973–8980.
11. Kovacs, F. A. & Cross, T. A. (1997). Transmembrane four-helix bundle of Influenza A M2 protein channel: structural implications from helix tilt and orientation. *Biophys. J.* **73**, 2511–2517.
12. Kukol, A., Adams, P. D., Rice, L. M., Brünger, A. T. & Arkin, I. T. (1999). Experimentally based orientational refinement of membrane protein models: a structure for the influenza A M2 H+ channel. *J. Mol. Biol.* **286**, 951–962.
13. Kukol, A. & Arkin, I. T. (1999). vpu Transmembrane peptide structure obtained by site-specific Fourier transform infrared dichroism and global molecular dynamics searching. *Biophys. J.* **77**, 1594–1601.
14. Kukol, A. & Arkin, I. T. (2000). Structure of the influenza C virus CM2 protein transmembrane domain obtained by site-specific infrared dichroism and global molecular dynamics searching. *J. Biol. Chem.* **275**, 4225–4229.
15. Torres, J., Adams, P. D. & Arkin, I. T. (2000). Use of a new label $^{13}C{=}^{18}O$, in the determination of a structural model of phospholamban in a lipid bilayer. Spatial restraints resolve the ambiguity arising from interpretations of mutagenesis data. *J. Mol. Biol.* **300**, 677–685.
16. Torres, J., Kukol, A. & Arkin, I. T. (2001). Site specific examination of secondary structure and orientation determination in membrane proteins: the peptidic $^{13}C{=}^{18}O$ group as a novel infrared probe. *Biopolymers*, **59**, 396–401.
17. Briggs, J. A., Torres, J. & Arkin, I. T. (2001). A new method to model membrane protein structure based on silent amino acid substitutions. *Proteins: Struct. Funct. Genet.* **44**, 370–375.
18. Byler, D. M. & Susi, H. (1986). Examination of the secondary structure of proteins by deconvolved FTIR spectra. *Biopolymers*, **25**, 469–487.
19. Braiman, M. S. & Rothschild, K. J. (1988). Fourier transform infrared techniques for probing membrane protein structure. *Annu. Rev. Biophys. Biomol. Struct.* **17**, 541–570.
20. Adams, P. D., Arkin, I. T., Engelman, D. M. & Brünger, A. T. (1995). Computational searching and mutagenesis suggest a structure for the pentameric transmembrane domain of phospholamban. *Struct. Biol.* **2**, 154–162.
21. Zhou, F. X., Merianos, H. J., Brünger, A. T. & Engelman, D. M. (2001). Polar residues drive association of polyleucine transmembrane helices. *Proc. Natl Acad. Sci. USA*, **98**, 2250–2255.
22. Tsuboi, M. (1962). Infrared dichroism and molecular conformation of α-form poly-gamma-benzyl-L-glutamate. *J. Polym. Sci.* **59**, 139–153.
23. Brünger, A. T., Adams, P. D., Clore, G., Gros, W., Grosse-Kunstleve, R., Jiang, J. *et al.* (1998). Crystallography and NMR system: a new software suite for macromolecular structure determination. *Acta Crystallog. sect. D*, **54**, 905–921.
24. Jorgensen, W. & Tirado-Rives, J. (1988). The OPLS potential function for proteins, energy minimization for crystals of cyclic peptides and crambin. *J. Am. Chem. Soc.* **110**, 1657–1666.
25. Bleasby, A. J. & Wooton, J. C. (1990). Construction of validated, non-redundant composite protein sequence databases. *Protein Eng.* **3**, 153–191.
26. Adams, P. D., Engelman, D. M. & Brünger, A. T. (1996). Improved prediction for the structure of the dimeric transmembrane domain of glycophorin A obtained through global searching. *Proteins: Struct. Funct. Genet.* **26**, 257–261.
27. Guex, N. & Peitsch, M. C. (1997). SWISS-MODEL and the Swiss-Pdb Viewer: an environment for comparative protein modelling. *Electrophoresis*, **18**, 2714–2723.

Edited by G. von Heijne

(Received 23 November 2001; received in revised form 25 May 2002; accepted 28 May 2002)