Effect of Variation of the Strength of the Aromatic Interactions of Tryptophan on the Cooperative Structural Refolding Behavior of a Peptide from HIV 1

Simon Schweizer and Jennifer Reed
Mechanisms of Biomolecular Interactions, Department A060, German Cancer Research Center, Heidelberg, Germany

ABSTRACT A 15-residue sequence (LPCRIKQFINMWQEV) forming the principal CD4-binding domain of gp120 from HIV 1 displays unusual, highly cooperative refolding from β-hairpin to 3_{10} helix when the polarity of the surrounding medium drops below a critical point, the so-called conformational switch. The tryptophan at position 12 has been shown to be essential for the cooperativity of the refolding process, and several lines of evidence from earlier work had suggested that it was the aromatic quadrupole that was responsible for this. To define more precisely what physico-chemical properties of tryptophan brought about the unique behavior of this peptide, nonproteogenic aromatic amino acids have been selected based on desired alterations in quadrupole moment, electrostatic potential surface, and binding energy to ions. These were built into the peptide in the place of tryptophan and their effect on switch behavior examined. It could be shown that a minimal strength of the quadrupole moment is necessary but not sufficient to enforce cooperativity of refolding, with other properties of tryptophan playing a role in the optimum interaction of this residue with other side chains of the peptide.

INTRODUCTION

Certain naturally occurring sequences found in a number of proteins act as polarity-dependent conformational switches (1,2). That is, they refold rapidly and cooperatively from one form of secondary structure to another when the polarity of the medium changes beyond a threshold value. The index case for such switches is the 15-residue peptide LPCRIKQFINMWQEV that comprises the principal CD4 binding site of the HIV 1 envelope protein, gp120, and changes in a water/trifluoroethanol gradient from a 60% β-conformation in purely aqueous solution to 55% helix under apolar (100% trifluoroethanol, TFE) conditions. The entire refolding process occurs within a narrow polarity range corresponding to 65 $\pm$ 5% TFE (2). This peptide has been extensively studied and much of the physico-chemical basis for its switch behavior has been identified (2–5). Chief in importance are two areas: the four N-terminal amino acids, LPCR, that constitute an initiation site for the helical fold and without which helical structure is not formed even at high TFE concentrations, and the tryptophan residue, located eight residues C-terminal to it, that appears to stabilize the initial β-fold against lowered polarity so that the full 60% is maintained until the critical point is reached. Charged-to-alanine experiments indicated that the tryptophan is interacting with more than one—indeed, the majority—of the other side chains to accomplish this (3).

Several lines of evidence support the conjecture that the ability of the tryptophan residue to stabilize the polar β-structure might be related to its aromatic properties, in particular the quadrupole strength. While substitution of this residue by valine leads to a complete loss of cooperativity in the folding process (and loss of infectivity in the virus), substitution by phenylalanine, which has a weak quadrupole, causes only a partial loss (1). Further, modification of the tryptophan residue by saturating the C2–C3 double bond so the aromatic character of the indole is eliminated (Hückel rule) results in a loss of folding cooperativity as well (3). In the latter case, however, interpretation is complicated by the fact that the modified tryptophan will no longer be planar in character, introducing steric as well as charge distribution changes. It must also be born in mind that in addition to the quadrupole, the π-electron system of an aromatic ring possesses dipole characteristics and an effective electrostatic potential surface, all of which would be affected by the modifications described above. It seemed desirable to examine what effect systematic changes of the strength of these properties of tryptophan would have on the cooperativity of the switch peptide.

Using computational methods to assess the effect of various proposed modifications of the tryptophan side chain on its physical properties, a series was selected that in theory encompassed a wide spectrum of potential changes. These modified tryptophans were incorporated in the switch peptide LPCRIKQFINMWQEV in the place of the wild-type residue and the consequences to the cooperative folding behavior monitored using circular dichroism (CD) spectroscopy and polarity titration.

METHODS

Calculation methods

The molecules were ab initio calculated by the TURBOMOLE program package, using the def-SV(P) and TZVPP basis sets and the B-P and B3-LYP

Submitted January 25, 2008, and accepted for publication May 27, 2008.

Address reprint requests to Jennifer Reed, Tel.: +49-6221-423256; E-mail: j.reed@dkfz-heidelberg.de.

Editor: Steven D. Schwartz.

0006-3495/08/10/3381/10 $2.00

doi: 10.1529/biophysj.107.126375

functionals (6–12). Every calculation was done first with def-SV(P)/B-P, then repeated with TZVPP/B-P and TZVPP/B3-LYP providing calculation results based on different basis sets and functionals. If they do not significantly differ, this indicates that the results are independent from the basis set and functional used to generate them and are thus reliable. In Table 1, the results of the TZVPP/B3-LYP calculations are shown. The molecules were taken as PDB files, converted to XYZ files using Open Babel (13), and oriented for a proper interpretation of the results. First, an appropriate origin was chosen. It is chemical convention to use the center of mass as origin, but here it was decided to self-select the origin and place it at a point of high symmetry within the aromatic ring. (For instance, it is desirable to have the indole and fluoroindole at the same origin. Following chemical convention would imply having a different origin since the center of mass is shifted by the fluorine atom's being heavier than the hydrogen atom.) The molecule was then rotated through the Eulerian angles for the aromatic ring being in the *x-y* plane and symmetry axes or functional groups aligned to the *x* or *y* axis. The particular orientation of the calculated molecules is depicted in Fig. 1. To check whether it was really sufficient to use indole alone and not the complete side chain of tryptophan, calculations on 3-ethylindole were carried out for comparison, with the CH_2 group saturated by adding a methyl group.

Measurements

Members of the modified aromatics in Table 1 were chosen for inclusion in the LAV peptide on the basis of significant alterations in the three calculated properties and their commercial availability. The 5-fluoro-L-tryptophan and 5-methoxy-L-tryptophan were purchased from Biosynth (Staad, Switzerland); β-(1-naphthyl)-L-alanine from Sigma-Aldrich (St. Louis, MO), and pentafluorophenyl-L-alanine from Chemgo Organica (Basel, Switzerland). CD spectra of the modified peptides were taken from 190 to 240 nm (far UV light) using a JASCO Spectropolarimeter J-710 (Hachioji, Japan). Peptide solutions at a concentration of 100 μg/mL were measured in 1-mm quartz cuvettes at room temperature. In the case of the 5-methoxytryptophan-LAV-peptide measured at pH 9.0, due to high solvent absorbance at low wavelengths a concentration of 10 μg/mL was measured in a 1-cm cuvette, thus providing a comparable number of molecules along the optical path. Each spectrum presented is the result of fourfold signal averaging with a similarly averaged solvent baseline subtracted. The spectra were then converted to mean residue ellipticity and subjected to fast Fourier transform to remove high-frequency noise. Polarity titrations were carried out using water/TFE mixtures, with spectra being taken at 20% steps from 0–100% TFE in most cases. Every titration series was measured twice. The spectra were fitted using a newly written program, PepFit-The Next Generation, which essentially follows the same principles as the old PepFit program (14) but has a much more convenient user interface and uses 1-nm subdivisions (personal communication, S. Schweizer, 2007).

RESULTS

In silico calculations

The results of the TURBOMOLE calculations are presented in Table 1. The values displayed (in atomic units) are the dipole moment and its absolute value, and the quadrupole moment, both as tensor and diagonalized (showing the values along the principal axes). As the aromatic rings of the molecules lie in the *x-y* plane the Q_{zz} is mostly unaffected by diagonalization. Exceptions are molecules with an angled functional group such as hydroxyl-, methoxy-, and amino groups. In these the quadrupole is somewhat altered in its alignment. For the sake of completeness, the anisotropy of the quadrupole moment is also given although it is of limited significance for the problem at hand. Boldface in Table 1 indicates those molecules used in this or previous studies.

As values for the quadrupole moment are dependent on the choice of origin for the coordinate system, comparisons must be limited to calculations within a given system. Thus, it is not possible, for example, to directly compare the value for indole given in Table 1 with others in the literature. Values for different moieties calculated within a single constant coordinate system, on the other hand, can be used to gain an idea of relative strengths. Thus one can conclude that fluoroindole has a weaker quadrupole moment than indole, indole a higher than benzene, and cyclohexane almost no quadrupole moment, as theory predicts. For comparative electrostatic surfaces and ion binding energies of aromatic molecules, the calculations of Mecozzi et al. (15) and Ma and Dougherty (16) were used.

Circular dichroism (CD) spectra

The program PepFit, used to calculate the secondary structure content of all the LAV variants for this study, is specifically designed for the deconvolution of peptide as opposed to globular protein. Circular dichroism (CD) spectra used a weighted average of synthetic peptide homopolymer CD spectra for the principal curve types seen in Fig. 2. Initial fitting is carried out with these alone; the fit is later refined using spectra representing the various types of reverse turns (for a complete description see (14)). Although the higher energy transitions of aromatic residues can extend below 240 nm and aromatic side chains are underrepresented in the synthetic peptides used to generate the master curves of Fig. 7, this has not in the past proven an obstacle to obtaining accurate fits, probably because the dichroism arising from aromatic residues is induced rather than intrinsic and thus much weaker than that associated with the peptide bond, and such low wavelength absorption bands constitute in any case a minor component of aromatic absorption. This has held true even for the peptides in this study using modifications of tryptophan that might have been expected to introduce unusual absorption bands. Fig. 3 presents an example of a typical fit using the PepFit program, in this case that of the 5-fluoro-tryptophan LAV peptide in 90% TFE.

The 5-fluorotryptophan derivative, as seen in Table 1, has in theory a weaker quadrupole than the wild-type residue (-3.9 vs. -6.2). In addition, the concentration of negative charge at the electrostatic surface above and below the plane of the ring should be significantly weaker, given the effect of addition of a single fluorine atom to benzene (15). The calculated dipole moment, however, is somewhat stronger than in unmodified tryptophan. Its effect on the folding behavior of the LAV peptide, when incorporated in the place of wild-type Trp, can be seen in Fig. 4 *b*. First, the initial amount of β-structure present in fully aqueous solution is lower than for the LAV peptide, 36% as opposed to 60%. Next, the cooperativity of the $\beta \rightarrow$ helix transition is almost completely eliminated, the conversion being linear with rising TFE

TABLE 1 List of molecules subjected to TURBOMOLE calculations

Molecule/related amino acid	Dipole moment $\vec{p}$	Abs. value $\lvert\vec{p}\rvert$	Quadrupole moment Q	Main axes of Q	Anisotropy of Q
Indole **Tryptophan**	$\begin{pmatrix}0.597\\0.610\\0.000\end{pmatrix}$	0.854	$\begin{pmatrix}2.41 & 2.07 & 0.00\\2.07 & 3.77 & 0.00\\0.00 & 0.00 & -6.18\end{pmatrix}$	0.91 5.27 −6.18	10.0
5-Fluoroindole **5-Fluorotryptophan**	$\begin{pmatrix}1.193\\0.939\\0.000\end{pmatrix}$	1.518	$\begin{pmatrix}-1.43 & -1.02 & 0.00\\-1.02 & 5.35 & 0.00\\0.00 & 0.00 & -3.92\end{pmatrix}$	−1.58 5.50 −3.92	8.5
7-Fluoroindole 7-Fluorotryptophan	$\begin{pmatrix}0.665\\-0.002\\0.000\end{pmatrix}$	0.665	$\begin{pmatrix}4.67 & 3.41 & 0.00\\3.41 & 0.12 & 0.00\\0.00 & 0.00 & -4.79\end{pmatrix}$	6.49 −1.71 −4.79	10.1
5-Chloroindole 5-Chlorotryptophan	$\begin{pmatrix}1.312\\0.969\\0.000\end{pmatrix}$	1.631	$\begin{pmatrix}-1.58 & -1.01 & 0.00\\-1.01 & 5.50 & 0.00\\0.00 & 0.00 & -3.92\end{pmatrix}$	−1.72 5.64 3.92	8.7
6-Chloroindole 6-Chlorotryptophan	$\begin{pmatrix}1.263\\0.236\\0.000\end{pmatrix}$	1.285	$\begin{pmatrix}-1.29 & 5.23 & 0.00\\5.23 & 5.28 & 0.00\\0.00 & 0.00 & -3.99\end{pmatrix}$	−4.18 8.17 −3.99	12.3
7-Chloroindole 7-Chlorotryptophan	$\begin{pmatrix}0.633\\-0.084\\0.000\end{pmatrix}$	0.638	$\begin{pmatrix}4.67 & 3.37 & 0.00\\3.37 & 0.36 & 0.00\\0.00 & 0.00 & -5.02\end{pmatrix}$	6.51 −1.49 −5.02	10.2
5-Bromoindole 5-Bromotryptophan	$\begin{pmatrix}1.365\\0.988\\0.000\end{pmatrix}$	1.685	$\begin{pmatrix}-1.64 & -1.02 & 0.00\\-1.02 & 5.59 & 0.00\\0.00 & 0.00 & -3.96\end{pmatrix}$	−1.78 5.73 −3.96	8.8
7-Bromoindole 7-Bromotryptophan	$\begin{pmatrix}0.621\\-0.122\\0.000\end{pmatrix}$	0.632	$\begin{pmatrix}4.70 & 3.32 & 0.00\\3.32 & 0.48 & 0.00\\0.00 & 0.00 & -5.18\end{pmatrix}$	6.52 −1.34 −5.18	10.3
5-Iodoindole 5-Iodotryptophan	$\begin{pmatrix}1.361\\0.982\\0.000\end{pmatrix}$	1.678	$\begin{pmatrix}-1.18 & -0.63 & 0.00\\-0.63 & 5.48 & 0.00\\0.00 & 0.00 & -4.30\end{pmatrix}$	−1.24 5.54 −4.30	8.7
7-Iodoindole 7-Iodotryptophan	$\begin{pmatrix}0.612\\-0.102\\0.000\end{pmatrix}$	0.621	$\begin{pmatrix}4.38 & 3.18 & 0.00\\3.18 & 1.22 & 0.00\\0.00 & 0.00 & -5.59\end{pmatrix}$	6.34 −0.75 −5.59	10.4
5-Methylindole 5-Methyltryptophan	$\begin{pmatrix}0.489\\0.554\\0.000\end{pmatrix}$	0.739	$\begin{pmatrix}2.57 & 2.36 & 0.00\\2.36 & 3.48 & 0.00\\0.00 & 0.00 & -6.05\end{pmatrix}$	0.62 5.43 −6.05	10.0
5-Methoxyindole **5-Methoxytryptophan**	$\begin{pmatrix}0.721\\0.712\\0.435\end{pmatrix}$	1.103	$\begin{pmatrix}1.36 & 1.37 & -4.13\\1.37 & 3.67 & -1.76\\-4.13 & -1.76 & -5.03\end{pmatrix}$	1.57 5.58 −7.15	11.3
5-Hydroxyindole 5-Hydroxytryptophan	$\begin{pmatrix}0.850\\0.243\\0.135\end{pmatrix}$	0.895	$\begin{pmatrix}-1.80 & 4.29 & -0.85\\4.29 & 7.27 & -1.64\\-0.85 & -1.64 & -5.46\end{pmatrix}$	−3.50 9.20 −5.71	13.9
Aminobenzene Aminophenylalanine	$\begin{pmatrix}0.525\\0.000\\0.376\end{pmatrix}$	0.646	$\begin{pmatrix}3.85 & 0.00 & 2.76\\0.00 & 1.60 & 0.02\\2.76 & 0.02 & -5.45\end{pmatrix}$	4.60 1.60 −6.20	9.7
5-Aminoindole 5-Aminotryptophan	$\begin{pmatrix}0.036\\0.471\\-0.349\end{pmatrix}$	0.488	$\begin{pmatrix}5.57 & 3.84 & 3.98\\3.84 & 2.63 & 2.44\\3.98 & 2.44 & -8.20\end{pmatrix}$	9.44 −0.60 −9.43	16.3

(*Continued*)

TABLE 1 (*Continued*)

Molecule/related amino acid	Dipole moment $\vec{p}$	Abs. value $\lvert\vec{p}\rvert$	Quadrupole moment Q	Main axes of Q	Anisotropy of Q
Dearomatized indole **Dearomatized tryptophan**	$\begin{pmatrix}0.601\\0.098\\0.282\end{pmatrix}$	0.672	$\begin{pmatrix}1.60 & 0.22 & 0.55\\0.22 & 2.70 & 0.63\\0.55 & 0.63 & -4.30\end{pmatrix}$	1.58 2.82 −4.41	6.7
Benzothiophene β-(3-Benzothienyl)-alanine	$\begin{pmatrix}-0.003\\-0.234\\0.000\end{pmatrix}$	0.234	$\begin{pmatrix}3.40 & -1.52 & -0.02\\-1.52 & 1.95 & 0.01\\-0.02 & 0.01 & -5.35\end{pmatrix}$	4.36 0.99 −5.35	8.5
Benzothiazole Benzothiazol-2-ylalanine	$\begin{pmatrix}-0.199\\-0.489\\0.000\end{pmatrix}$	0.528	$\begin{pmatrix}4.95 & -1.13 & 0.00\\-1.13 & -0.82 & 0.00\\0.00 & 0.00 & -4.14\end{pmatrix}$	5.16 −1.03 −4.14	8.2
Benzindole Benzotryptophan	$\begin{pmatrix}0.859\\0.562\\0.000\end{pmatrix}$	1.026	$\begin{pmatrix}4.79 & 3.52 & 0.00\\3.52 & 4.00 & 0.00\\0.00 & 0.00 & -8.79\end{pmatrix}$	7.94 0.85 −8.79	14.5
3-Ethylindole Tryptophan	$\begin{pmatrix}0.631\\0.523\\-0.022\end{pmatrix}$	0.820	$\begin{pmatrix}2.36 & 1.91 & -0.12\\1.91 & 3.68 & 0.07\\-0.12 & 0.07 & -6.04\end{pmatrix}$	5.03 1.01 −6.04	9.7
Benzene **Phenylalanine**	$\begin{pmatrix}0\\0\\0\end{pmatrix}$	0	$\begin{pmatrix}1.91 & 0.00 & 0.01\\0.00 & 1.91 & 0.00\\0.01 & 0.00 & -3.81\end{pmatrix}$	1.91 1.91 −3.81	5.7
Phenol tyrosine	$\begin{pmatrix}-0.007\\0.508\\-0.046\end{pmatrix}$	0.510	$\begin{pmatrix}1.31 & -3.56 & 0.90\\-3.56 & 2.51 & -0.98\\0.90 & -0.98 & -3.81\end{pmatrix}$	−1.71 5.71 −4.00	8.8
Cyclohexane	$\begin{pmatrix}0.000\\0.000\\0.000\end{pmatrix}$	0.000	$\begin{pmatrix}-0.16 & 0.00 & 0.01\\0.00 & -0.16 & -0.01\\0.01 & -0.01 & 0.33\end{pmatrix}$	−0.16 −0.16 0.33	0.5
Pentafluorobenzene **Pentafluorophenylalanine**	$\begin{pmatrix}-0.544\\0.000\\-0.001\end{pmatrix}$	0.544	$\begin{pmatrix}1.17 & 0.00 & 0.00\\0.00 & -4.10 & 0.00\\0.00 & 0.00 & 2.93\end{pmatrix}$	1.17 −4.10 2.93	6.3
Hexafluorobenzene	$\begin{pmatrix}0\\0\\0\end{pmatrix}$	0	$\begin{pmatrix}-2.10 & 0.00 & -0.01\\0.00 & -2.07 & 0.00\\-0.01 & 0.00 & 4.17\end{pmatrix}$	−2.10 −2.07 4.17	6.3
Heptafluoroindole	$\begin{pmatrix}0.039\\-0.203\\0.065\end{pmatrix}$	0.217	$\begin{pmatrix}-2.02 & -0.95 & 0.01\\-0.95 & -3.05 & -0.08\\0.01 & -0.08 & 5.07\end{pmatrix}$	−1.45 −3.62 5.07	7.8
Naphthalene **1-Naphthylalanine**	$\begin{pmatrix}0\\0\\0\end{pmatrix}$	0	$\begin{pmatrix}3.02 & 0.00 & 0.00\\0.00 & 2.87 & 0.00\\0.00 & 0.00 & -5.90\end{pmatrix}$	3.02 2.87 −5.90	8.8
Pyridine β-(3-pyridyl)-alanine	$\begin{pmatrix}-0.434\\0.767\\0.000\end{pmatrix}$	0.881	$\begin{pmatrix}2.50 & 2.15 & 0.01\\2.15 & -0.09 & 0.00\\0.01 & 0.00 & -2.41\end{pmatrix}$	3.72 −1.31 −2.41	5.7
Biphenyl 4-phenyl-phenylalanine	$\begin{pmatrix}0\\0\\0\end{pmatrix}$	0	$\begin{pmatrix}3.54 & 0.00 & 0.01\\0.00 & 3.69 & 0.01\\0.01 & 0.01 & -7.23\end{pmatrix}$	3.54 3.69 −7.23	10.8
Nitrobenzene Nitrophenylalanine	$\begin{pmatrix}-1.853\\0.001\\-0.002\end{pmatrix}$	1.853	$\begin{pmatrix}-7.60 & 0.00 & -0.01\\0.00 & 5.64 & 0.00\\-0.01 & 0.00 & 1.96\end{pmatrix}$	−7.60 5.64 1.96	11.8
Azulene β-Azulylalanine	$\begin{pmatrix}-0.405\\0.000\\-0.002\end{pmatrix}$	0.405	$\begin{pmatrix}3.20 & 0.00 & 0.04\\0.00 & 3.08 & 0.03\\0.04 & 0.03 & -6.28\end{pmatrix}$	3.20 3.08 −6.28	9.4

Some molecules, such as hexafluorobenzene or cyclohexane, are included for didactic purposes. If a molecule has no dipole moment due to symmetry, a 0 is used instead of 0.000. The orientation of the molecules is displayed in Fig. 1. Boldface indicates those molecules used in this or previous studies.

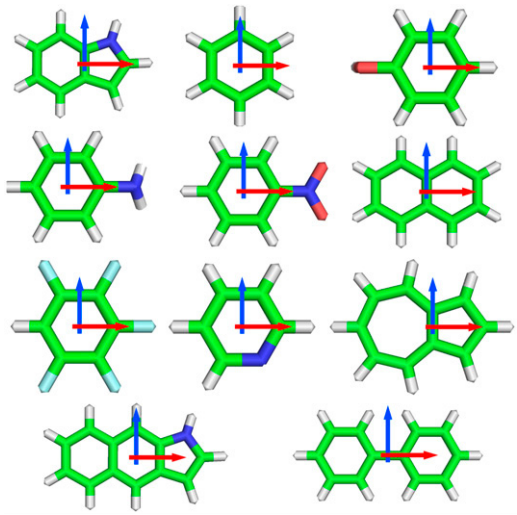

FIGURE 1 Diagram showing the origin and orientation of the calculated molecules: indole (and analogs), benzene (and analogs), phenol, aminobenzene, nitrobenzene, naphthalene, pentafluorophenylalanine, pyridine, azulene, benzindole, and biphenyl. The *x* axis is colored in red and the *y* axis in blue. The program PyMOL (21) was used to generate the molecular images.

concentration except for a steeper area between 10% and 30% TFE. In this respect, its effect is quite similar to the Trp → Phe substitution discussed earlier. There, too, was a near linear $\beta \rightarrow$ helix conversion as a function of TFE concentration except for a steeper area between 20% and 40% TFE. This shift of the more nearly cooperative region to higher TFE concentrations plus the fact that the Trp → Phe mutant has nearly the wild-type level of β-structure in aqueous solution would seem to indicate that Phe in this position is better at maintaining cooperativity than the 5-fluoro derivative.

CD TFE titration spectra and secondary structure content of the naphthylalanine LAV peptide are shown in Fig. 5, *a* and *b*. The quadrupole moment of naphthalene, the aromatic molecule in the side chain of β-(1-naphthyl)-alanine, at −5.9, is not too different from that of wild-type tryptophan. In comparison to the indole ring, however, the distribution of negative charge at the surface is weaker and more diffuse in naphthalene and there is no effective dipole moment, which symmetry reasons do not allow. Incorporated into the LAV peptide, the effect on conformational behavior is twofold. Despite the moderately strong quadrupole at the tryptophan site, the peptide again has a lower content of β-structure at 0% TFE and this structure is lost in a linear fashion with no obvious steeper declivity. The helix content, on the other hand, rises earlier (lower TFE concentration) and more sharply than in the wild-type LAV peptide.

The methoxy group in the 5-methoxytryptophan derivative should strengthen the π-electron system by pushing electrons inside, and therefore increase the quadrupole moment. TURBOMOLE modeling in fact indicated that this derivative should theoretically possess a stronger quadrupole moment than the wild-type side chain, −7.2 rather than −6.2. The dipole moment is also significantly stronger. However, its

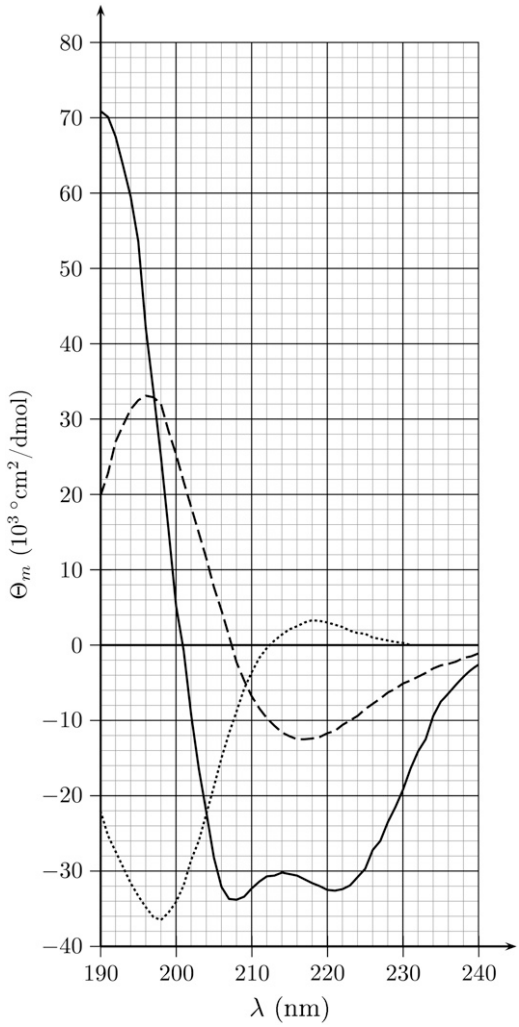

FIGURE 2 CD master curves for α-helix (*solid line*), β sheet (*dashed line*), and random coil (*dotted line*) the data points are in 1-nm subdivisions (personal communication, J. Reed, 2007; however, the same data points every 2.5 nm are also published in Reed and Reed (14)).

effect on the folding behavior of the LAV peptide, shown in Fig. 6 *b*, confounds expectations. It was thought the stronger dipole/quadrupole would increase the strength of the interactions that serve to stabilize the β-structure, maintaining the cooperative nature of the polarity-driven refolding and possibly shifting it to lower polarity, i.e., ≥60% TFE. Instead, while there is still a fair level of cooperativity in the $\beta \rightarrow$ helix folding transition, it occurs over a broader range than for wild-type LAV and at lower TFE concentrations (20–40%). In addition, the amount of β-structure initially present, while higher than for the other two derivatives, is still well below 60%. One problem with the introduction of the methoxy group is that it provides a potential hydrogen-bonding partner not present in the native peptide; this might work to disrupt some interactions stabilizing the β-fold. Accordingly, the experiment was repeated, this time at pH 9.0; by approaching the pK_a of potential side-chain amide donors it was hoped to weaken any nonnative interactions. In fact, the TFE

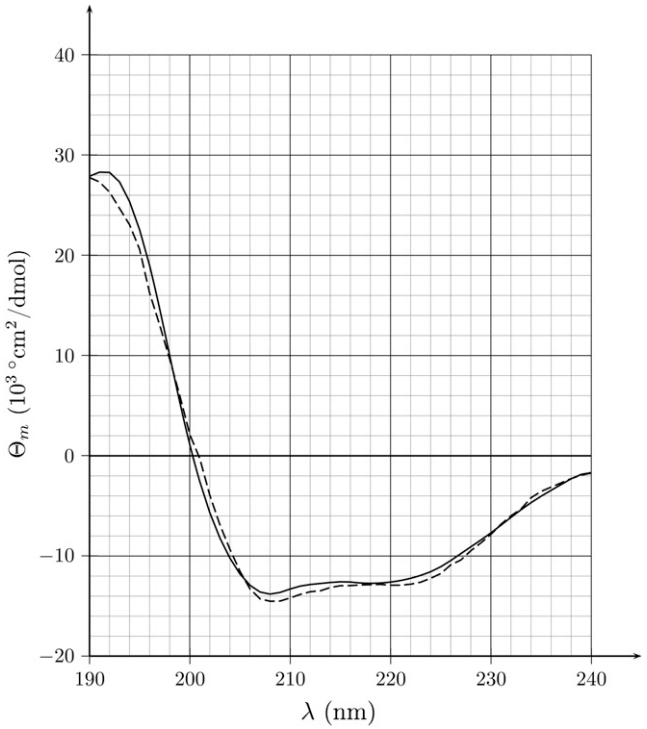

FIGURE 3 Example of measured curve and fit curve. (*Solid line*) A single CD measurement of the 5-fluoro-tryptophan LAV peptide in 90% TFE (*dashed line*) fit curve containing 35% α-helix, 4% β sheet, 14% random coil, and remainders of other master curves.

titration at pH 9.0 (Fig. 7 *b*) displays a higher initial level of β-structure, a more pronounced maintenance of β and absence of helix at low TFE concentrations, and a shift of the critical transition point to ~45% TFE.

In the wild-type LAV peptide, the positively charged side chains of the two basic residues (Arg and Lys) are thought to interact with the negative areas of the quadrupole above and below the ring plane in a classic amide-aromatic interaction. (Basic residues at these positions are a strongly conserved feature of natural HIV strains.) The effect of introducing a strongly electronegative atom like fluorine at all possible sites on the benzene molecule is to invert the quadrupole; hexafluorobenzene has positive charge above and below the plane of the ring and negative charge within the plane (17). The effect of an inverted quadrupole was investigated in a peptide where the two basic residues were replaced by two acidic residues to maintain the proposed interaction (18). Interpretation is complicated by the fact that a totally F-substituted tryptophan was not available, and pentafluorophenylalanine was used instead. The folding behavior of the peptide, if the principal β-stabilizing interactions remained intact, would thus be expected more nearly to resemble the W → F mutant than wild-type LAV. The results are presented in Fig. 8

What is immediately clear is that complimentary inversion of the charge on both the quadrupole and two of the charged side chains with which it is likely to interact does not reproduce the conditions necessary for cooperative refolding.

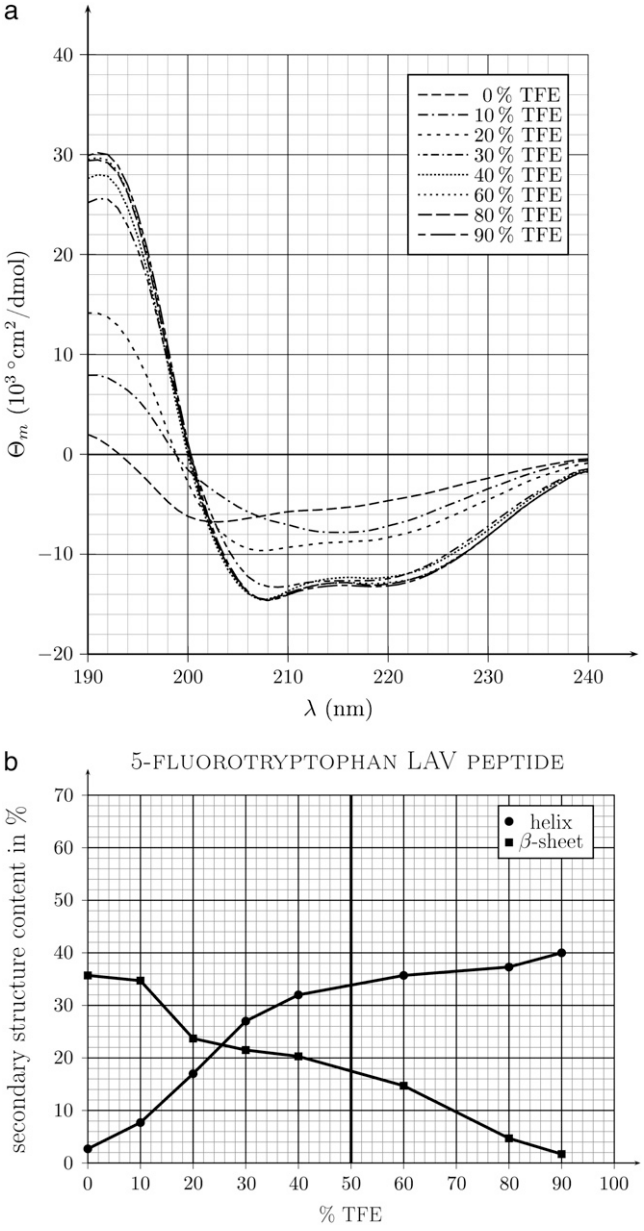

FIGURE 4 The 5-fluorotryptophan LAV peptide.

There is no β-structure at all present in aqueous solution while there is a small amount of helix. The amount of helix increases up to 30% TFE; that of β increases between 30% and 80% TFE. There is no obvious interconversion. Whatever interactions with the tryptophan residue are responsible for stabilizing the β-structure, inverting the aromatic quadrupole destroys them completely and substituting oppositely charged side chains does not compensate for the effects of the inversion.

DISCUSSION

Any interpretation of the effect of modified aromatics will be complicated by the fact that the alterations introduced may not

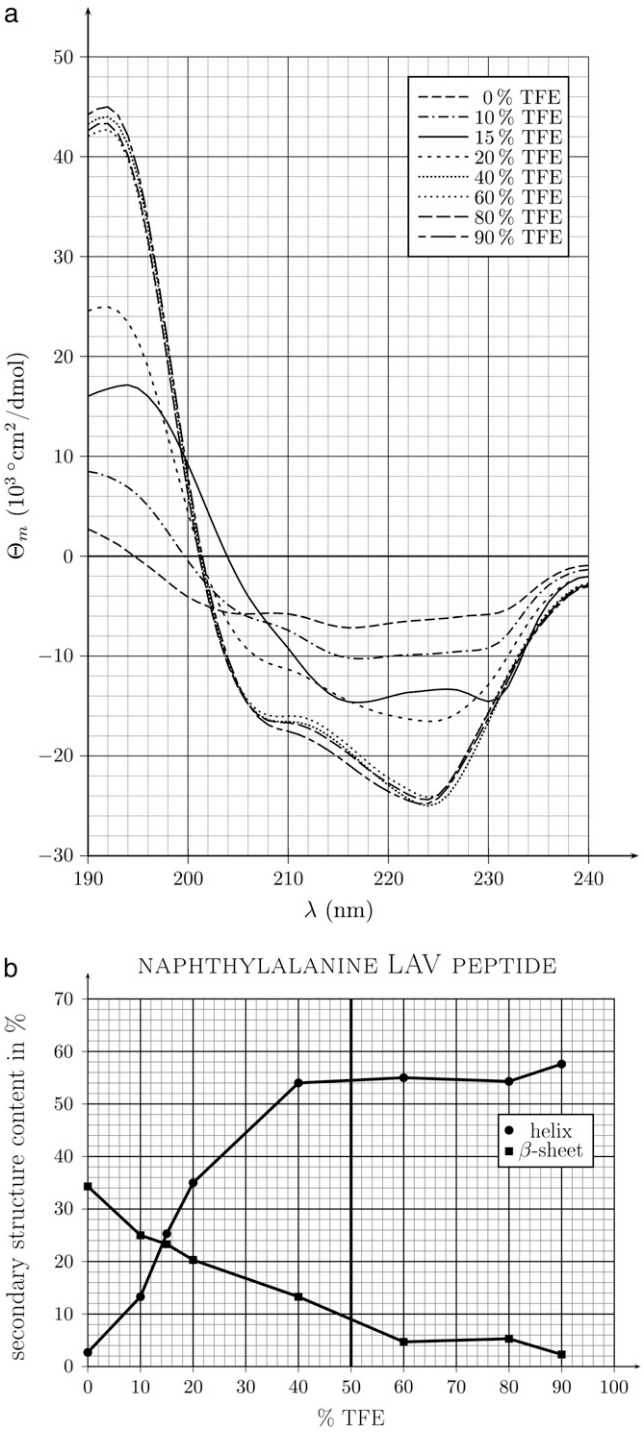

FIGURE 5 The 1-naphthylalanine LAV peptide.

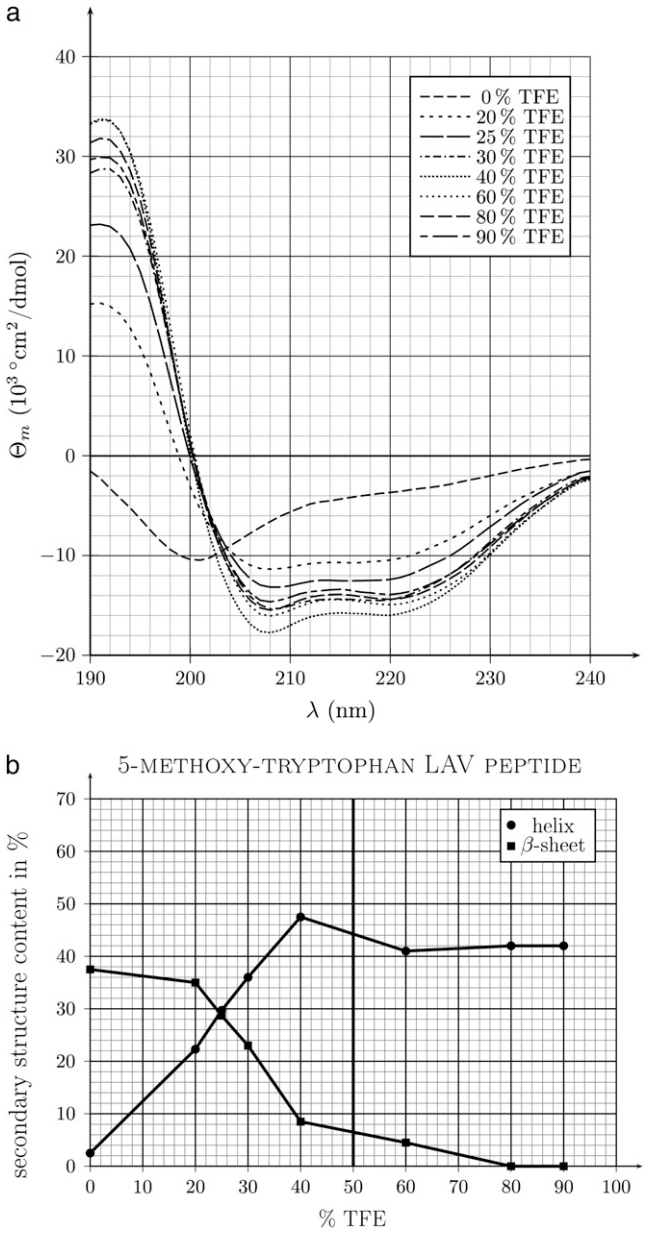

FIGURE 6 The 5-methoxytryptophan LAV peptide.

only affect the properties under examination—here quadrupole strength, dipole moment, etc.—but also alter the relative hydrophobicity of the side chain and/or its secondary structure preferences. With the exception of the pentafluorophenylalanine, available data suggests that the effects on hydrophobicity will be minor (19,20). With regard to secondary structure propensities, no data is available to indicate how strongly this might be affected. As wild-type tryptophan, however, has no strong preferences, being weakly inclined to both helix and β-sheet and lying close to the median for turn formation, the modifications we have introduced will in no case eliminate a strong tendency to one or another conformation. With this in mind, it is possible to draw certain conclusions.

The series of experiments described here have served to establish that a major, if not the only, factor governing the cooperativity of the polarity-driven $\beta \rightarrow$ helix folding transition in the native LAV peptide is stabilization of the original β-fold against decreasing polarity. All alterations that lowered the amount of β-structure in aqueous solution also showed decreased cooperativity and there is a roughly pro-

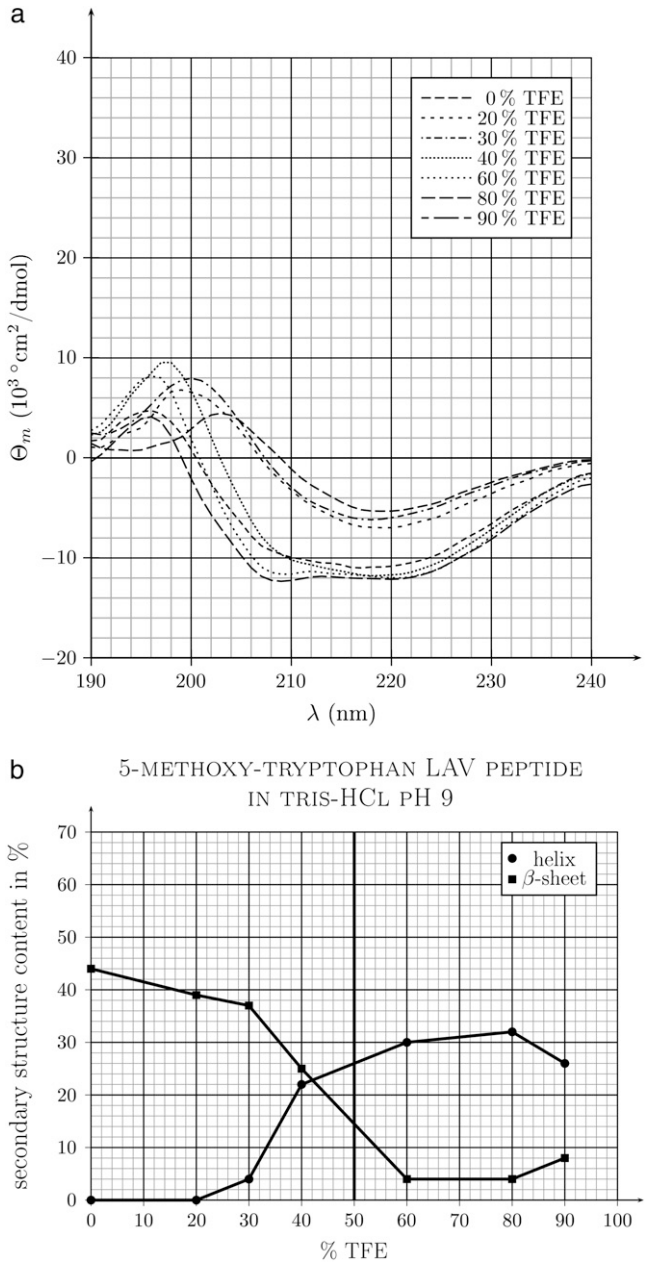

FIGURE 7 The 5-methoxytryptophan LAV peptide in Tris HCl, pH 9.0.

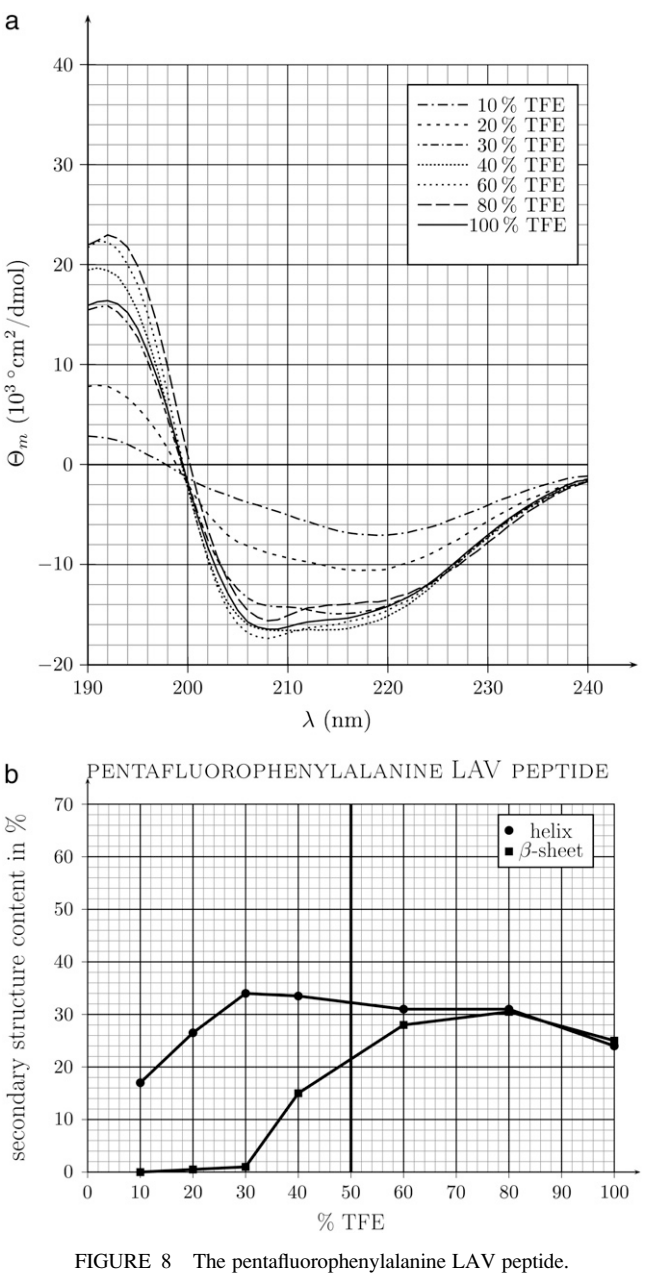

FIGURE 8 The pentafluorophenylalanine LAV peptide.

portional relationship between the amount of β present at 0% TFE and cooperative refolding. The results suggest that although the LAV peptide does have a tendency to adopt β-structure in water, there is a degree of flexibility so that the stabilizing effect of an aromatic side chain is required to raise the conformational average to 60%. The forces that rigidify the initial conformation also appear to maintain the β-fold in the face of decreasing polarity. The effects of the several aromatic modifications, however, show that the role of tryptophan is more complex than originally conceived. There is no simple, consistent association between global β-stabilization in water and any single physico-chemical property of the tryptophan residue.

The results of the experiment with 5-fluorotryptophan incorporated into the LAV peptide did seem to bear out the initial contention that it is the quadrupole strength of the aromatic ring system that is decisive in stabilizing β-structure and ensuring cooperativity. Not only is there a moderately linear correspondence between these two qualities on the one hand and the quadrupole strength on the other, both following a Trp $>$ Phe $>$ 5-F-Trp progression, but the dipole moment seems relatively unimportant, as the fact that the dipole moment of 5-F-Trp (1.5) is stronger than the wild-type residue (0.9) while Phe even lacks an effective dipole, is not apparently able to compensate for the weaker quadrupole.

When one looks at the results with the naphthylalanine LAV peptide, though, this relatively simple picture is con-

tradicted. Although the quadrupole strength of the naphthyl derivative (−5.9) is quite close to that of Trp (−6.2), the folding behavior of the modified peptide is even less like that of native LAV than the 5-F-Trp-LAV (less β-structure under polar conditions and no cooperativity). Thus although a higher dipole strength cannot compensate for a weak quadrupole, complete absence of the dipole moment, even when the quadrupole moment should be adequate, also seems to disrupt whatever forces act to stabilize β-structure as the polarity drops. (It should be noted that the helix content tends to rise at lower TFE concentrations in these cases, an indication that the default position of LAV peptide in the absence of firm β-stabilization is to form helix once the environment is not fully aqueous.) The remarkable sensitivity of folding behavior in this peptide to any modification of the tryptophan residue shows that it possesses an ideal combination of several characteristics intrinsic to the process.

Yet another facet of the uniqueness of the balance of forces present in the tryptophan residue is revealed when, in an attempt to strengthen the indole quadrupole, a methoxy group is introduced at the ring carbon C5. Although the calculated quadrupole and dipole strength of this derivative are both stronger than those of the native side chain, the cooperativity of the folding transition in this peptide is weaker and accompanied, as usual, by less β at 0% TFE. Rather than preserving the β-structure against higher TFE concentrations, the effect of the substitution is to weaken it. One possibility was that nonnative hydrogen bonding interactions introduced with the methoxy group were blocking or replacing those necessary for the unique folding behavior. Raising the pH to 9.0 did in fact alter the folding behavior to something more nearly approaching that of the wild-type peptide, although not totally recovering the switch activity. This introduces a third factor where the physico-chemical properties of tryptophan are optimal, and alterations—even those that strengthen some of the forces involved—reduce the overall efficiency of the system.

Finally, if the principal interactions establishing cooperativity of refolding were those between the aromatic quadrupole and the conserved basic residues, then inverting the charges on both the quadrupole and the side chains concerned should reproduce to a large extent the behavior of the native peptide. This is not the case. The conformational activity of pentafluorophenylalanine LAV is unique and retains none of the characteristics typical of the switch peptide. Stability for the β-fold is so compromised that it is simply not present in aqueous solution, whereas in contrast to the native peptide, helical structure is already established. It does not dominate the apolar conformation, though, and after a certain point in the titration β-structure appears and eventually equals that of helix. There is clearly no $\beta \rightarrow$ helix conversion and no cooperativity. However, since introduction of the five fluorine atoms has the effect of rendering the side chain markedly more hydrophobic than wild-type Phe, the conformational behavior of pentafluorophenylalanine-LAV may be confounded by the tendency to exclude this residue from solvent. On the other hand, this does not explain the loss of β-structure, as a number of hydrophobic peptides adopt the β-hairpin conformation in aqueous solution.

Tryptophan in nature is used sparingly, and is the residue least frequently found in proteins. It is also the only amino acid to have just one codon (UGG), thus minimizing the chances of a tryptophan appearing at random within a protein sequence. This suggests that having a tryptophan residue anywhere but where it is absolutely necessary can be dangerous. The charged-to-alanine experiments previously carried out on the LAV peptide certainly support the idea that it is quite a powerful source of interactive forces, as the best interpretation of the results was that it was involved with multiple side chains. The data presented here on modifications of the aromatic ring suggest that the ability of the tryptophan side chain to promote cooperative refolding in switch sequences is based not on any single one of its physico-chemical properties, but on a critical balance between them. The relative strength of the tryptophan quadrupole and absolute dipole moment, the size of the aromatic ring system, and the presence of a hydrogen bond acceptor together constitute a unique and highly potent system. In the LAV peptide, at least, these are used to the fullest extent and any alteration adversely affects stability. In view of these findings, the role of tryptophan residues wherever they occur in proteins might bear closer examination.

This research received in-house support from the German Cancer Research Center. We thank the group of Dr. Wolfgang Wenzel at the Forschungszentrum Karlsruhe for use of the TURBOMOLE software, Prof. François Diederich for helpful discussions and his suggestions regarding the methoxytryptophan derivative, and Dr. Rüdiger Pipkorn for his expert peptide synthesis.

REFERENCES

1. Reed, J., and V. Kinzel. 1991. A conformational switch is associated with receptor affinity in peptides derived from the CD4—binding domain of gp120 from HIV I. *Biochemistry*. 30:4521–4528.
2. Reed, J., and V. Kinzel. 1993. Primary structure elements responsible for the conformational switch in the envelope glycoprotein gp120 from human immunodeficiency virus type 1: LPCR is a motif governing folding. *Proc. Natl. Acad. Sci. USA*. 90:6761–6765.
3. von Stosch, A. G., V. Kinzel, R. Pipkorn, and J. Reed. 1995. Investigation of the structural components governing the polarity-dependent refolding of a cd4-binding peptide from gp120. *J. Mol. Biol.* 250:507–513.
4. von Stosch, A. G., M. A. Jiménez, V. Kinzel, and J. Reed. 1995. Solvent polarity-dependent structural refolding: a CD and NMR study of a 15-residue peptide. *Proteins Struct. Funct. Gen.* 23:196–203.
5. von Stosch, A. G., C.-W. von der Lieth, and J. Reed. 1999. Molecular dynamics study of the proposed β-hairpin form of the switch domain from HIV 1 gp120 alone and complexed with an inhibitor of cd4 binding. *Proteins*. 34:197–205.
6. Ahlrichs, R., M. Bär, M. Häser, H. Horn, and C. Kölmel. 1989. Electronic structure calculations on workstation computers: the program system Turbomole. *Chem. Phys. Lett.* 162:165–169.
7. Treutler, O., and R. Ahlrichs. 1995. Efficient molecular numerical integration schemes. *J. Chem. Phys.* 102:346–354.

8. van Arnim, M., and R. Ahlrichs. 1998. Performance of parallel TURBOMOLE for density functional calculations. *J. Comput. Chem.*19:1746–1757.

9. Eichkorn, K., O. Treutler, H. Öhm, M. Häser, and R. Ahlrichs. 1995. Auxiliary basis sets to approximate Coulomb potentials. *Chem. Phys. Lett.* 240:283–289.

10. Eichkorn, K., F. Weigend, O. Treutler, and R. Ahlrichs. 1997. Auxiliary basis sets for main row atoms and transition metals and their use to approximate Coulomb potentials. *Theor. Chem. Acc.* 97:119−124.

11. Schäfer, A., C. Huber, and R. Ahlrichs. 1994. Fully optimized contracted Gaussian basis sets of triple ζ-valence quality for atoms Li to Kr. *J. Chem. Phys.* 100:5829–5835.

12. Schäfer, A., H. Horn, and R. Ahlrichs. 1992. Fully optimized contracted Gaussian basis sets for atoms Li to Kr. *J. Chem. Phys.* 97:2571–2577.

13. Guha, R., M. T. Howard, G. R. Hutchison, P. Murray-Rust, H. Rzepa, C. Steinbeck, J. K. Wegner, and E. L. Willighagen. 2006. The blue obelisk—interoperability in chemical informatics. *J. Chem. Inf. Model.* 46:991–998.

14. Reed, J., and T. A. Reed. 1997. A set of constructed type spectra for the practical estimation of peptide secondary structure from circular dichroism. *Anal. Biochem.* 254:34–40.

15. Mecozzi, S., J. A. P. West, and D. A. Dougherty. 1996. Cation-π interactions in aromatics of biological and medicinal interest: electrostatic potential surfaces as a useful qualitative guide. *Proc. Natl. Acad. Sci. USA.* 93:10566–10571.

16. Ma, J. C., and D. A. Dougherty. 1997. The cation-π interaction. *Chem. Rev.* 97:1303–1324.

17. Meyer, E. A., R. K. Castellano, and F. Diederich. 2003. Interactions with aromatic rings in chemical and biological recognition. *Angew. Chem. Int. Ed.* 42:1210–1250.

18. Quiñonero, D., C. Garau, C. Rotger, A. Frontera, P. Ballester, A. Costa, and P. M. Deyà. 2002. Anion-π interactions: do they exist? *Angew. Chem. Int. Ed.* 41:3389–3392.

19. Imamura, T., K. Konishi, and K. Konishi. 2006. Interaction of indole derivatives and tryptophan peptides with interfaces of sodium dodecyl sulfate micelles. *J. Pept. Sci.* 12:403–411.

20. Katragadda, M., and J. D. Lambris. 2006. Expression of compstatin in *Escherichia coli*: incorporation of unnatural amino acids enhances its activity. *Protein Expr. Purif.* 47:289–295.

21. DeLano, W. L. 2002. The PyMOL User's Manual. DeLano Scientific, Palo Alto, CA.