Mary Koeppe Luidens
Department of Medicine
Albany Medical College
Albany, NY 12208

James Figge
Departments of Medicine
and Pathology
Albany Medical Center
Albany, NY 12208
and Department of
Biomedical Sciences
State University of New York
Albany, NY 12201

Kimberly Breese
Department of Medicine
Albany Medical College
Albany, NY 12208

Sandor Vajda*
Department of
Biomedical Engineering
and BioMolecular Engineering
Research Center
Boston University
Boston, MA 02215

Predicted and Trifluoroethanol-Induced α-Helicity of Polypeptides

The α-helix stabilizing solvent 2,2,2-trifluoroethanol (TFE) is frequently used as a medium for determining the average α-helicity of polypeptides by CD spectroscopy. CD spectra measured in solutions containing 10, 15, 20, 50, and 90% (vol/vol) TFE are presented for 5 peptides that were selected to demonstrate possible variations in the effect of TFE concentration on the secondary structure. The analysis is extended to 6 further peptides whose CD spectra as measured in TFE are documented in the literature. The observed α-helicity at a high TFE concentration is compared with the α-helicity determined by a structure prediction method that combines conformational filtering [*S. Vajda, (1993)* Journal of Molecular Biology, *Vol. 229, pp. 125–145*], *and a Monte Carlo simulation* [*J. Figge et al. (1993)* Protein Science, *Vol. 2, pp. 155–164*]. *For the set of 11 peptides we find a correlation of 0.84 between the predicted* $[\Theta]_{222}$ *values and the corresponding values observed by CD spectroscopy in a high concentration of TFE* ($p < 0.01$). *Although we generally find a good correlation at high TFE concentration between observed and predicted α-helicity, there are several peptides that do not follow the predicted behavior. An analysis of the TFE titration curves in one such case revealed that TFE*

Received June 23, 1995; accepted December 21, 1995.

* To whom correspondence should be addressed.

Biopolymers, Vol. 39, 367–376 (1996)

CCC 0006-3525/96/030367-10

can induce a sharp transition from a partial β-sheet conformation to an α-helical conformation as the TFE concentration is increased above a critical value. © *1996 John Wiley & Sons, Inc.*

INTRODUCTION

Conformational analysis of short linear peptides is frequently carried out in structure-forming solvents such as 2,2,2-trifluoroethanol (TFE). TFE, a hydrophilic and hydrogen-bonding solvent, is known to stabilize marginally stable α-helices.[1-5] Some amount of TFE is frequently called for when determining the average α-helicity of polypeptides using CD spectroscopy. Even in TFE, α-helices are formed only by fragments that have some intrinsic α-helical propensity,[6-8] and hence the method provides useful structural information.[9,10] However, as emphasized in a number of recent papers, α-helicity measured in TFE must be regarded with caution,[11-14] because it may not reflect the α-helicity that would be obtained in other situations, such as in a particular protein microenvironment.[11]

Information on the effect of TFE can be obtained by generating titration curves, i.e., performing CD measurements in TFE/water mixtures with increasing concentrations of TFE. At low TFE concentrations, the CD spectra of randomly coiled or partially structured peptides show little α-helicity. However, the spectra of such peptides, and even the spectra of some peptides in β-sheet conformation, generally exhibit a transition as TFE concentration is increased above a critical value.[13,15,16] Both the sharpness of the transition and the critical TFE concentration are sequence dependent. For some peptides low TFE concentrations (10–20%) yield a sharp transition to maximum α-helicity, while for others the addition of a small amount of TFE does not substantially increase the α-helicity above that observed in water, and the maximum α-helicity is attained only at a relatively high TFE concentration.

The major goal of the present work is to compare the α-helicity of short peptides observed in high TFE concentrations with that determined by a structure prediction method. CD spectra measured at 10, 15, 20, 50, and 90% TFE are presented for 5 peptides that were selected to demonstrate possible variations in the effect of TFE concentration on the secondary structure.[12,17] The analysis is extended to 6 additional peptides whose CD spectra as measured in TFE are documented in the literature. The observed α-helicity at high TFE concentration is compared with predicted α-helicity, which is determined by a combination of conformational filtering[18] and a Monte Carlo simulation,[17] and expressed as mean residue ellipticity at 222 nm ($[\Theta]_{222}$).

As is the case for all statistical methods of secondary structure prediction, conformational filtering is trained on a set of known protein structures. However, using nmr data, we have shown that the parameters extracted from the statistical analysis of crystallized proteins are transferable to polypeptides, and yield reasonable predictions of conformational preferences in aqueous solution,[18] i.e., in a completely different environment. As will be shown, conformational filtering combined with a Monte Carlo simulation predicts the α-helicity of short linear peptides with reasonable accuracy; the correlation between predicted $[\Theta]_{222}$ values and those measured at high TFE concentrations is $r = 0.84$ for the 11 peptides considered in this paper. Although we generally find good correlation at high TFE concentration between observed and predicted α-helicity, we present examples of peptides that do not follow the predicted behavior. An analysis of the TFE titration curves in one such case reveals that TFE can induce a sharp transition from a partial β-sheet conformation to an α-helical conformation as the concentration of TFE is increased above a critical value.

MATERIALS AND METHODS

Peptide Sequences

Five peptides were selected for analysis by CD spectroscopy (Table I). The first peptide (Rb3) contains a potent helical seed sequence linked to a segment of the papilloma virus 18 E7 protein sequence that binds to the retinoblastoma (Rb) protein.[17] Due to the acetyl and amide blocking groups, this peptide has no protonatable end groups and thus might be expected to show some differences from the other peptides in the table. The next three peptides are from proposed transcriptional activation regions of steroid hormone receptors:[12] a conserved region of the human estrogen receptor (hER),[19] the tau2 region of the human glucocorticoid receptor (hGR),[20] and the homologous tau2 region of the human mineralocorticoid receptor (hMR). The fifth peptide (hMR-pro) has a proline substituted for the isoleucine which occurs at position 11 of the hMR peptide.

Peptide Synthesis and Purification

The synthesis of the peptides has been described previously.[12,17] To recapitulate, peptides were synthesized

Table I Peptide Sequences[a]

Peptide	Sequence
1. Rb3	acetyl—A E T A A A D L L C H E Q L S—amide
2. hER	(538) NH_2—D L L L E M L D A H R L H A P T—COOH
3. hGR	(529) NH_2—T P T L V S L L E V I E P E V L—COOH
4. hMR	(735) NH_2—T P S P V M V L E N I E P E I V—COOH
5. hMR-pro	NH_2—T P S P V M V L E N P E P E I V—COOH

[a] The position of the initial amino acid in the corresponding full-length protein is indicated in parentheses.

The full-length proteins are as follows: hER, human estrogen receptor; hGR, human glucocorticoid receptor; hMR, human mineralocorticoid receptor.[12] Rb3 contains a potent helical seed sequence linked to a segment of the papilloma virus 18 E7 protein sequence that binds to the retinoblastoma (Rb) protein.[17]

using solid-phase methods.[21] Each peptide was purified to greater than 95% homogeneity using preparative reverse phase high performance liquid chromatography (RP-HPLC). Purity was verified by analytical RP-HPLC using a linear gradient of acetonitrile in water containing 0.1% trifluoroacetic acid on a Vydac C-18 column. The amino acid composition of each peptide was confirmed with quantitative amino acid analysis and plasma desorption mass spectrometry.[22] The latter yielded the following masses for the [M+H] + species of each peptide:[12,17] Rb3, 1614.3 ± 1.0; hER, 1846.0 ± 1.0; hGR, 1752.7 ± 1.0; hMR, 1767.6 ± 1.0; and hMR-pro, 1752.4 ± 1.0. The corresponding calculated monoisotopic and isotopically averaged masses were Rb3, 1612.8 and 1613.8; hER, 1845.0 and 1846.2; hGR, 1752.0 and 1753.1; hMR, 1766.9 and 1768.1; and hMR-pro, 1750.9 and 1752.0, respectively. The linear sequence of each peptide (except for Rb3, which had a blocked N-terminus) was confirmed with automated gas-phase sequencing. Peptides were stored under argon prior to use to avoid oxidation.

Stock Solutions

Purified samples of the peptides were dissolved in Milli-Q purified water (hER, hMR-pro, Rb3) or 90% (v/v) trifluoroethanol (TFE; nmr grade: Aldrich) in water (hGR, hMR) at a concentration of 1 mg/mL, and then aliquots of these stock were diluted with the appropriate amounts of TFE and water to obtain the desired final peptide concentration in 10–90% (v/v) TFE in water. Concentrations of standard solutions for each peptide were determined by quantitative amino acid analysis.

CD Measurements and Analysis

CD spectra were recorded at 20°C on a Jasco J-720 spectropolarimeter using 0.50–1.00 mm path-length cells. The instrument was calibrated with (+)-10-camphorsulfonic acid. CD data were recorded from 260.0 to 182.0 nm at 0.2 nm decrements with a scan speed of 20 nm/min, a 1.0 nm band width, and an averaging time of 0.5 s. Five scans of each sample and solvent were taken and the averaged, baseline-corrected spectra were smoothed with a Savitzky-Golay filter.[23] Data are expressed as mean residue ellipticity, $[\Theta]$ (deg cm^2 dmol^{-1}), calculated using the number of peptide amide bonds (calculated as $N - 1$, where N represents the number of amino acid residues in the peptide). Concentration independence of the mean residue ellipticity at 222 nm was established for peptides hER, hMR, hMR-pro, and Rb3, implying that secondary structure is not dependent on peptide aggregation.[11,12,17] In contrast, the mean residue ellipticity at 222 nm for the hGR peptide increased as a function of peptide concentration, from −14,300 deg cm^2 dmol^{-1} (at a peptide concentration of 30 μM) to −12,150 deg cm^2 dmol^{-1} (at 300 μM).[12]

Analysis of Secondary Structure

The CD spectra of five peptides (Table I) were analyzed for secondary structure using singular-value decomposition combined with a variable selection algorithm[24,25] using a 22-protein basis set. The criteria for acceptable analysis of secondary structures are as described by Johnson:[26] (1) The sum of secondary structural elements should be from 0.96 to 1.05. For the Rb3 peptide, the upper limit had to be expanded to 1.06 in order to obtain a solution. (2) The proportion of each secondary structure should be greater than or equal to −0.05. (3) The rms error should be less than or equal to 0.22.

Prediction of $[\Theta]_{222}$ Values

The $[\Theta]_{222}$ value of a polypeptide is predicted in two stages. The goal of stage I is to estimate the probability of the α-helical conformation for each residue of a given sequence. In stage II we use the estimated probabilities to generate 100,000 conformations for the peptide, and calculate the probabilities that the chain contains α-helices of r residues long, where $r = 6, 7, 8, \ldots, N$, and N is the total number of residues in the polypeptide. The predicted $[\Theta]_{222}$ value for the peptide is then calculated from these probabilities by an empirical equation.[12,27]

Stage I: Conformational Filtering. Conformational filtering estimates the probabilities of 16 regions of the (ϕ, ψ) map[28] for each amino acid residue of the given sequence. In this paper we use only the probability of region A, defined by $-110° < \phi < -40°$ and $-90° < \psi < -10°$, which includes α-helical conformations. To give a brief description of the procedure,[18] we consider a sequence of N residues. Our goal is to determine the probabilities p_i^m, $i = 1, \ldots, 16$, of the 16 conformational states for the mth residue. Conformational filtering is based primarily on the simultaneous use of forward and backward Markov models in which the probability of a state for a residue depends only on the state of the nearest neighbors of that residue. Extracting doublet data from high resolution x-ray structures of 91 proteins, we established transitional probabilities for the 400 residue pairs.[18] For example, $p(\mathrm{Ala}_i^{m+1}/\mathrm{Gly}_j^m)$ denotes the normalized frequency of an Ala residue in state i when it is preceeded by a Gly residue in state j. More generally, we compile 400, 16 times 16 forward Markov matrices with elements of the form $T_{ij}^m = p(R_i^{m+1}/R_j^m)$. The same information is arranged in form of backward Markov matrices $\tilde{T}_{ij}^m = p(R_i^{m-1}/R_j^m)$. These matrices are the input data for predicting the probability distributions for the residues of a particular sequence by repeated application of the forward and backward transition formulas $\mathbf{p}^{m+1} = \mathbf{T}^m\mathbf{p}^m$ and $\mathbf{p}^{m-1} = \tilde{\mathbf{T}}^m\mathbf{p}^m$ to obtain forward and backward probabilities for each residue. If the starting probability $\mathbf{p}^1$ is an eigenvector of the matrix $\mathbf{T}_2\mathbf{T}_3 \ldots \mathbf{T}_n\tilde{\mathbf{T}}_{n-1} \ldots \tilde{\mathbf{T}}_2\tilde{\mathbf{T}}_1$, then the solution is stable, the probabilities do not change when repeating the forward or backward iteration.[18] The final probabilities in this nearest-neighbor part of the algorithm are obtained by averaging the forward and backward estimates. Notice that in spite of using only nearest-neighbor interactions, changes in probabilities of any residue progress along the entire chain. We also included a heuristic procedure to further improve the estimates by extending the analysis to overlapping tripeptide and tetrapeptide fragments.[18]

Stage II: Calculation of Predicted $[\Theta]_{222}$ Value by Monte Carlo Simulation. A Monte Carlo simulation was employed to estimate the probabilities that a given polypeptide of length N contains α-helical segments of length r, where $r = 6, 7, 8, \ldots, N$. Notice that hexamers have been found to be the shortest peptides whose conformation is substantially affected by the TFE.[5-7] In the first part of the simulation, 100,000 representative conformations were generated for each polypeptide using the position-specific probabilities from stage I above. In the second part of the simulation, each of the 100,000 conformations was scanned from position 2 to position $(N - 1)$ to find all independent segments of length i, where $i = 4, 5, 6, \ldots (N - 2)$, such that every residue within the segment was assigned to Zimmerman region A. These i residues represent the internal residues contained within a predicted α-helical segment of length r, where $r = i + 2$. That is, the entire α-helical segment of length r is defined to contain i internal residues, each in Zimmerman conformation A, plus two end residues (with unspecified conformation). The number of properly aligned amide bonds within the α-helical segment of length r, by definition, will then be j, where $j = r - 1 = i + 1$. The value of j can range from 5 to $(N - 1)$. The CD signal arises from the properly aligned amide bonds. Therefore, the parameter j was used to organize the output from the Monte Carlo simulation. For each allowed value of j we estimated the probability $w(j)$ of the peptide containing an α-helical segment with j amide bonds as the ratio of such conformations among the 100,000 structures generated, resulting in a discrete probability distribution function with values $w(5), w(6), \ldots, w(N - 1)$.

The predicted $[\Theta]_{222}$ value was calculated for each polypeptide using the probability distribution function $w(j)$ and an empirical relationship between CD parameters per residue and the α-helix length that was proposed by Chen et al.,[27] and extended by Scholtz et al.,[29] to take into account the effect of the temperature. Each contribution from an α-helical fragment with j amide bonds was multiplied by the factor $j/(N - 1)$, where the denominator is a normalization factor accounting for the total number of amide bonds in the polypeptide. The value of $[\Theta]_{222}$ was calculated as

$$[\Theta]_{222} = \sum_{j=5}^{(N-1)} j/(N-1)*w(j)*H(j) \quad (1)$$

In this equation, $H(j)$ represents the mean residue ellipticity in deg cm^2 dmol^{-1} of a complete α-helix of $j + 1$ residues, and is calculated by the empirical formula

$$H(j) = -40000*[1 - 1.53/(j+1)] + 100*T \quad (2)$$

given by Scholtz et al.[29] In this formula, the constant, $-40{,}000$, represents the mean residue ellipticity at 222 nm of an infinite α-helix; 1.53 is the

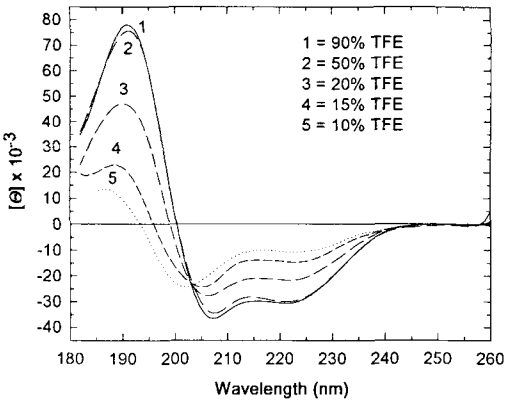

FIGURE 1 CD spectra of the Rb3 peptide (16.7 μM) in 10–90% (vol/vol) TFE at 20°C. Units for [Θ] are deg $cm^2\ dmol^{-1}$.

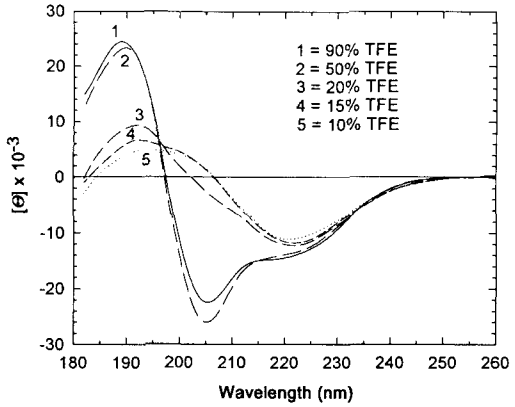

FIGURE 3 CD spectra of the hGR peptide (35.6 μM) in 10–90% (vol/vol) TFE at 20°C. Units for [Θ] are deg $cm^2\ dmol^{-1}$.

value of the chain-length dependence parameter (i.e., x in the equation of Scholtz et al.[29]), as estimated by Luidens et al.[12] from an analysis of a highly helical 17 residue peptide;[30] and T is the temperature in degree Celsius.

RESULTS

Figures 1–5 show the far-uv CD spectra from 260 to 182 nm for the 5 peptides listed in Table I, dissolved in solutions containing 10 to 90% (vol/vol) TFE concentrations with water at 20°C.

The Rb3 peptide (Figure 1) in 90% TFE shows features characteristic of an α-helical structure, with a positive maximum near 191 nm, and double negative maxima near 208 and 222 nm. These features appear as the TFE concentration is increased. For example, the positive maximum shifts from 186.8 to 190.8 nm as the TFE percentage increases from 10 to 90%. The position of the first negative maximum shifts from 202.0 to 207.2 nm, and the value of $[\Theta]_{222}$ shifts from −10,700 to −30,600 deg $cm^2\ dmol^{-1}$ as the TFE concentration increases from 10 to 90%. Note the occurrence of an isodichroic point at approximately 203 nm.

The hER peptide (Figure 2) also shows features of α-helical structure that appear as the TFE concentration is increased, with the magnitude of the positive and negative maxima for hER being about one-half of those seen for RB3. The positive maximum shifts from 188.8 to 191.2 nm as the TFE percentage increases from 10 to 90%. The position of the first negative maximum shifts from 204.0 to 206.8 nm, while the value of $[\Theta]_{222}$ shifts from −6,700 to −16,700 deg $cm^2\ dmol^{-1}$ as the TFE

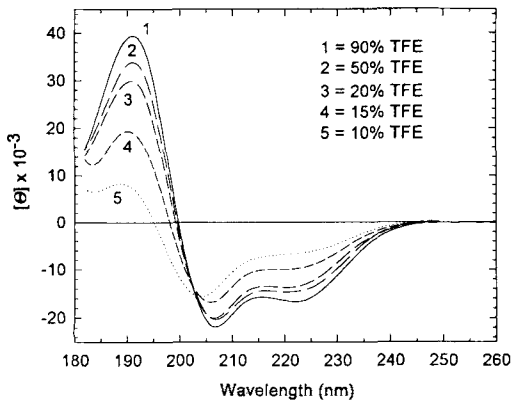

FIGURE 2 CD spectra of the hER peptide (37.6 μM) in 10–90% (vol/vol) TFE at 20°C. Units for [Θ] are deg $cm^2\ dmol^{-1}$.

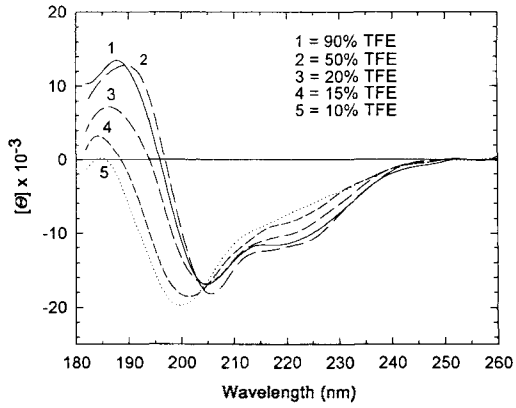

FIGURE 4 CD spectra of the hMR peptide (10.9 μM) in 10–90% (vol/vol) TFE at 20°C. Units for [Θ] are deg $cm^2\ dmol^{-1}$.

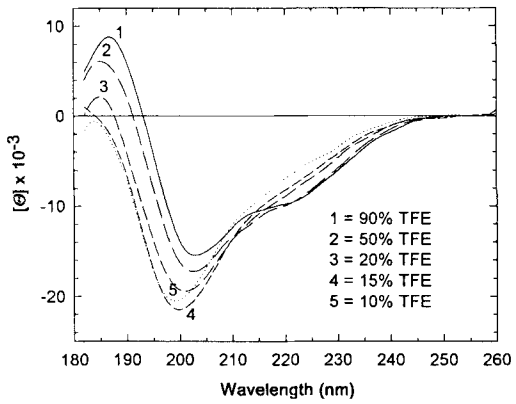

FIGURE 5 CD spectra of the hMR-pro peptide (31.9 μM) in 10–90% (vol/vol) TFE at 20°C. Units for $[\Theta]$ are deg cm^2 dmol^{-1}.

concentration goes from 10 to 90%. There is also an isodichroic point near 203 nm.

The hGR peptide (Figure 3) shows a markedly different pattern. In 20% or less TFE the CD curves are characteristic of β-sheet plus undefined structure. In 50% TFE and 90% TFE the spectrum undergoes a major change and features suggestive of α-helical structure are seen. The positive maximum is located at 189.0 nm in 90% TFE, with the negative maximum at 205.4 nm, and the value of $[\Theta]_{222}$ is $-14{,}300$ deg cm^2 dmol^{-1}.

The hMR peptide (Figure 4) shows a pattern reminiscent of those seen with Rb3 and hER, with features of partial α-helical structure appearing as the percent TFE increases from 10 to 50%. The positive maximum shifts from 184.8 to 189.6 nm as the TFE percentage increases from 10 to 50%. The negative maximum shifts from 199.8 to 205.6 nm, and the value of $[\Theta]_{222}$ shifts from $-6{,}900$ to $-11{,}600$ deg cm^2 dmol^{-1} as the TFE concentration goes from 10 to 50%. Paradoxically, when the TFE concentration is raised from 50 to 90%, there is a slight loss of α-helical character. There is also an isodichroic point near 203 nm (allowing for some noise).

The hMR-pro peptide (Figure 5) shows only marginal α-helical character. The positive maximum shifts only to 186.8 nm as the TFE percentage increases to 90%. The negative maximum shifts from 199.0 to 202.8 nm. The value of $[\Theta]_{222}$ shifts from -6200 to -9500 deg cm^2 dmol^{-1} as the TFE concentration goes from 10 to 90%. Note the complete lack of an isodichroic point near 203 nm.

The CD spectra of the 5 peptides listed in Table I were analyzed for secondary structure using the decomposition procedure of Manavalan and Johnson.[25] Results are shown in Table II. The simplest available measure related to α-helicity is $[\Theta]_{222}$. Although the value of $[\Theta]_{222}$ is affected by the nonhelical components of the secondary structure,[25] at least at high TFE concentration there is very good correlation between the percent α-helicity calculated by the deconvolution and $[\Theta]_{222}$ (Figure 6). Given this correlation, we have elected to use $[\Theta]_{222}$ as the parameter to compare observed and predicted α-helicity. The results of this comparison are given in Figure 7.

As shown in Figure 7, in addition to the five peptides listed in Table I, we studied the six peptide sequences shown in Table III. The CD spectra of these peptides, observed at a high TFE concentration (70–90%), are published.[10,31,32] Therefore, it is possible to extract relevant $[\Theta]_{222}$ values from the literature for these peptides. In addition, relatively complete sets of TFE titration curves have been published for the peptide analogues of the four myohemerythrin helices[31] and for the S peptide of ribonuclease A.[10,32]

Figure 7 shows that there is a good correlation between the predicted and experimental values of $[\Theta]_{222}$ for the 11 peptides considered in this study ($r = 0.84$; $p < 0.01$). Seven of the peptides fall very close to the line of identify (Figure 7: no. 1 [RB3], no. 4 [hMR], no. 6 [MYOH-N], no. 7 [MYOH-A], no. 9 [MYOH-C], no. 10 [MYOH-D], and no. 11 [S-peptide]). However, in the case of four peptides (Figure 7: no. 2 [hER], no. 3 [hGR], no. 5 [hMR-pro], and no. 8 [MYOH-B]) the predicted α-helicity differs substantially from that observed in TFE.

DISCUSSION

Our goal is to establish a relationship between the α-helicity observed at high TFE concentration and the predicted α-helicity (Figure 7). The measures of α-helicity used here are (1) percentage α-helicity from the decomposition of the spectrum, and (2) mean residue ellipticity at 222 nm, $[\Theta]_{222}$. As shown in Figure 6, there is a strong correlation between the α-helicity calculated by spectral decomposition and $[\Theta]_{222}$ at high TFE concentration. Therefore, we elected to use $[\Theta]_{222}$ as our parameter of α-helicity, since this value can also be extracted from literature data. As shown in Figure 7, there is generally a good correlation between measured and predicted $[\Theta]_{222}$ values, where the predicted values are derived from known protein struc-

Table II Secondary Structural Content (%) of Peptides in TFE Solutions Determined by the Singular Value Decomposition and Variable Selection Method of Manavalan and Johnson[25]

Peptide	% TFE	α-Helix	Parallel β-Sheet	Antiparallel	Turn	Other
Rb3	10	34	0	2	37	31
	15	48	0	2	29	23
	20	62	0	6	23	5
	50	76	0	0	23	0
	90	84	0	0	22	0
hER	10	24	0	23	28	28
	15	34	0	19	25	20
	20	46	0	12	26	14
	50	48	0	11	26	13
	90	52	0	8	28	11
hGR	10	16	26	0	8	51
	15	18	10	10	14	55
	20	26	13	0	15	49
	50	50	0	4	35	12
	90	48	0	6	29	20
hMR	10	23	6	2	33	38
	15	31	4	0	29	37
	20	34	0	2	37	32
	50	41	0	5	30	28
	90	34	0	11	27	30
hMR-pro	10	19	3	16	28	33
	15	23	1	14	30	32
	20	24	1	13	28	32
	50	32	3	3	25	38
	90	30	0	7	30	37

tures using conformational filtering and a Monte Carlo simulation (see Materials and Methods).

Examination of Figure 7 reveals that some peptides behave substantially differently than their predicted behavior. The poorly predicted peptides fall into two groups. There are three points that lie significantly below the line of identity in Figure 7 (peptides no. 3 [hGR], no. 5 [hMR-pro], and no. 8 [MYOH-B]), indicating that the α-helicity of each of these peptides is substantially underpredicted. In contrast, the α-helicity of peptide no. 2 (hER) is substantially overpredicted. This peptide is known to be especially stable as an α-helix, not only in TFE, but also in 2 m*M* sodium dodecyl sulfate, a solvent system known to promote β-structure.[12]

Further insights into the folding behavior of substantially underpredicted peptides can sometimes be obtained from an evaluation of the TFE titration curves. For example, peptide no. 3 (hGR) forms a partial β-sheet and undefined structure at low TFE concentrations (10–20%, Table II and Figure 3). As the concentration of TFE is increased to 50%, there is a sharp transition to a 50% α-helical structure, and the β-sheet component is completely lost (Table II and Figure 3). Peptide no. 4 (hMR) does not display this type of behavior even though this peptide is homologous to hGR (and 9 of 16 residues are identical). In fact, the helicity of hMR is well predicted (Figure 7). Therefore, the specific amino acid residues that differ between these homologous peptides appear to have an effect on the conformational preferences of these two species in TFE-containing solutions.

Other authors[16,33] have also documented that high concentrations of TFE can induce a transition from β-structure to an α-helical conformation in certain polypeptides. Fan et al.[16] studied monellin, which in the native state is composed of an A-chain that is entirely in the β-conformation, and a B-chain that contains both α- and β-structure. CD and preliminary two-dimensional nmr assignments showed that a segment of the monellin A chain undergoes a structural reorganization from β-sheet to α-helix in 50% TFE.[16] In this alcohol-denatured state, the conformation of the A-chain is

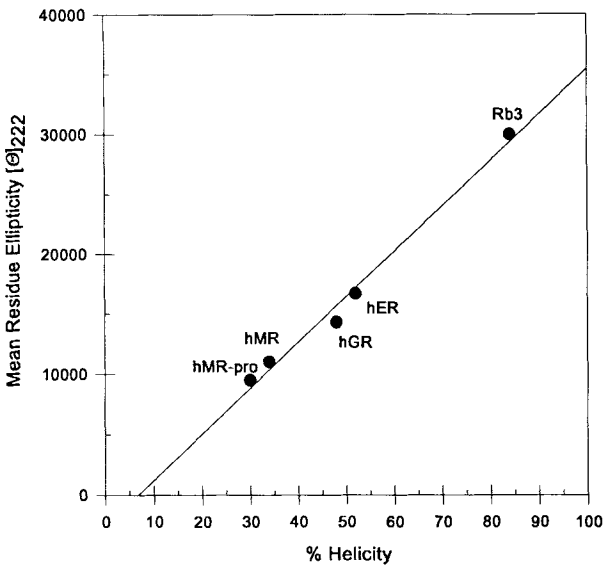

FIGURE 6 Absolute value of the mean residue ellipticity at 222 nm, $[\Theta]_{222}$, in 90% TFE vs percent α-helicity for peptides 1–5. Percent α-helicity is taken from Table II for 90% TFE. Units for $[\Theta]_{222}$ are deg cm^2 dmol^{-1}.

clearly influenced by the environmental context in which the residues are found, in keeping with the observations of Zhong and Johnson.[11] Barrow et al.[33] showed that the β-(1-42)amyloid peptide undergoes a transition from β-structure to α-helical structure with increasing concentrations of TFE. The behavior of this peptide is very similar to the behavior documented for the hGR peptide in the present study.

For peptide no. 5 (hMR-pro), some characteristic features of coil–helix transition are missing. For example, there is no isodichroic point near 203 nm (Figure 5), suggesting this peptide does not have a strong tendency to form an α-helix. In the case of peptide no. 8 (MYOH-B), substantial α-helix formation is seen only at high TFE concentration.[31] In contrast, peptide no. 9 (MYOH-C) also had substantial α-helical content only at high TFE concentration,[31] but was well predicted (Figure 7).

The mechanism by which TFE induces α-helical structure is not known with certainty. It has been proposed that the induction of α-helices by TFE is due to weakened hydrogen bonding between the solvent and the peptide.[34] Under such circumstances, TFE is thought to stabilize secondary structure because intramolecular hydrogen bonding within the polypeptide chain is favored at the expense of intermolecular hydrogen bonds between the solvent and the peptide.[34] This mechanism would be expected to stabilize α-helices, and the observed α-helicity should then reflect the intrinsic α-helical propensity of the specific peptide sequence. The correlation documented in this report ($r = 0.84$) between observed and predicted α-helicity is in accord with such a model.

Nevertheless, as pointed out by Jasanoff and Fersht,[13] the above model is not fully supported by experimental data. For example, this model cannot easily explain our data for the hGR peptide (Figure 3), which does not behave as predicted by our algorithm. In particular, the model does not readily explain why TFE stabilizes α-helices in peptides that are already in, or have a preference for the β-sheet conformation, as demonstrated by our hGR peptide, and also by Fan et al.,[16] Barrow et al.[33] and Wei and Fasman.[34] In addressing this issue, the latter authors have also proposed that TFE acts as a denaturant that destabilizes hydrophobic regions of proteins.[16,34] Since the β-sheet structure involves the alignment of noncontiguous regions of the polypeptide chain and is thought to be stabilized by packing of nonpolar residues in the interior of the sheet, it is proposed that TFE can denature such a β-sheet structure.[34] This model could explain the results obtained in the present study for the hGR peptide. One could argue that at low TFE concentrations, the intrinsic preference of the peptide to form β-structure is manifest due to the stabilization

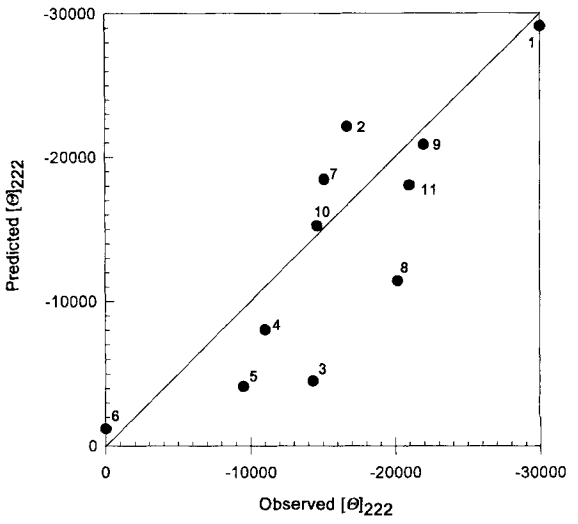

FIGURE 7 $[\Theta]_{222}$ predicted vs $[\Theta]_{222}$ observed for peptides 1–11 in high TFE. Units for $[\Theta]_{222}$ are deg cm^2 dmol^{-1}. Peptides are identified by number as given in Tables I and III: no. 1, Rb3; no. 2, hER; no. 3, hGR; no. 4, hMR; no. 5, hMR-pro; no. 6, MYOH-N; no. 7, MYOH-A; no. 8, MYOH-B; no. 9, MYOH-C; no. 10, MYOH-D; no. 11, S-peptide.

Table III Peptide Sequences Taken From the Literature[a]

Peptide	Sequence
6. MYOH-N term	(1) NH_2—G W E I P E P Y V W D E S F R V F Y—COOH
	(18) NH_2—Y E Q L D E E H K K I F K G I F D C I R
7. MYOH-A	D—COOH
	(40) NH_2—S A P N L A T L V K V T T N H F T H E E A
8. MYOH-B	M M D—COOH
	(63) NH_2—E V V P H K K M H K D F L E K I G G L—
9. MYOH-C	COOH
10. MYOH-D	(93) NH_2—A K N V D Y C K E W L V N H I K—COOH
	(1) NH_2—K E T A A A K F E R Q H M D S S T S A—
11. S-peptide	COOH

[a] The full-length proteins are as follows: MYOH, myohemerythrin—N-terminus and A-, B-, C-, and D-helices[31]; S-peptide, ribonuclease A.[10,32] The position of the initial amino acid in the corresponding protein is indicated in parentheses.

of intramolecular hydrogen bonding. However, at higher TFE concentrations, the hydrophobic interior of the β-sheet structure would be disrupted and the peptide would then refold as an α-helix.

The above model gives some possible insights into the activity of TFE but it is still not clear whether TFE interacts directly with the peptide chain. Jasanoff and Fersht[13] have proposed that the effects of increasing concentrations of TFE in titration experiments are due to TFE/H_2O exchange at peptide binding sites. Such a model implies that TFE binds or interacts directly with the peptide, and as TFE concentrations are increased, H_2O dissociates and TFE binds to sites on the peptide.[13] It is further proposed that the binding of TFE increases the probability that individual residues will go into the α-helical conformation. At present, there is no experimental data to prove that TFE actually binds to peptides as proposed.[35] Therefore, while some possible models of TFE action can be offered, further experiments will be needed to sort out the exact molecular mechanisms involved. Any future models must be able to explain the complex behavior of peptides such as the hGR peptide reported in this study.

In conclusion, high concentrations of TFE provide a suitable solvent to characterize the α-helical propensities of short peptides, and a good correlation with the predicted α-helicity, based upon the average behavior of peptide segments in proteins, can be obtained. However, one must use caution, since in some cases the α-helicity observed at high TFE concentration can differ substantially from the predicted value. Therefore, in addition to performing CD measurements at high concentrations of TFE, it is generally useful to predict the α-helical content by the combined conformational filtering/Monte Carlo method described herein. Finally, the use of TFE titration curves can help uncover unexpected folding behavior in some cases where a discrepancy exists between predicted and observed α-helicity.

This study was funded in part by NIH grant DK02031 (MKL), by a Grant from the Lucille P. Markey Charitable Trust (JF), and by NIH grant AI30535 (SV).

We thank Dr. Richard J. Ridge (Protein and Nucleic Acid Chemistry Center, Woods Hole Oceanographic Institution, Woods Hole, MA) for peptide syntheses and quantitative amino acid analyses. We are indebted to Dr. Robert MacColl (New York State Department of Health, Wadsworth Center for Laboratories and Research, Albany, NY) for use of the J-720 spectropolarimeter.

REFERENCES

1. Goodman, M. & Listowsky, I. (1962) *J. Am. Chem. Soc.* **84,** 3770–3771.
2. Goodman, M., Listowsky, I., Masuda, Y. & Boardman, F. (1963) *Biopolymers* **1,** 33–42.
3. Goodman, M. & Rosen, I. G. (1964) *Biopolymers* **2,** 537–559.
4. Goodman, M., Langsam, M. & Rosen, I. G. (1966) *Biopolymers* **4,** 305–319.
5. Goodman, M., Verdini, A. S., Toniolo, C., Phillips, W. D. & Bovey, F. A. (1969) *Proc. Nat. Acad. Sci.* **64,** 444–450.
6. Goodman, M., Naider, F. & Toniolo, C. (1971) *Biopolymers* **10,** 1719–1730.
7. Goodman, M., Naider, F. & Rupp, R. (1971) *Bioorg. Chem.* **1,** 310–328.

8. Goodman, M. & Saltman, R. P. (1981) *Biopolymers* **20,** 1929–1948.
9. Pallai, P. V., Mabilia, M., Goodman, M., Vale, W. & Rivier, J. (1983) *Proc. Nat. Acad. Sci.* **80,** 6770–6774.
10. Nelson, J. W. & Kallenbach, N. R. (1986) *Proteins Struct. Funct. Genet.* **1,** 211–217.
11. Zhong, L. & Johnson, W. C., Jr. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 4462–4465.
12. Luidens, M. K., Aks, C. S., Zhu, Q-L., Smith, T. F., MacColl, R. & Figge, J. (1993) *Peptide Res.* **6,** 134–139.
13. Jasanoff, A. & Fersht, A. R. (1994) *Biochemistry* **33,** 2129–2135.
14. Waterhous, D. V. & Johnson, W. C., Jr. (1994) *Biochemistry* **33,** 2121–2128.
15. Liebes, L. F., Zand, R. & Phillips, W. D. (1975) *Biochim. Biophys. Acta* **405,** 27–39.
16. Fan, P., Bracken, C. & Baum, J. (1993) *Biochemistry* **32,** 1573–1582.
17. Figge, J., Breese, K., Vajda, S., Zhu, Q., Eisele, L., Andersen, T. T., MacColl, R., Friedrich, T. & Smith, T. F. (1993) *Protein Sci.* **2,** 155–164.
18. Vajda, S. (1993) *J. Mol. Biol.* **229,** 125–145.
19. Danielian, P. S., White, R., Lees, J. A. & Parker, M. G. (1992) *EMBO J.* **11,** 1025–1033.
20. Hollenberg, S. M. & Evans, R. M. (1988) *Cell* **55,** 899–906.
21. Kent, S. B. (1988) *Ann. Rev. Biochem.* **57,** 957–989.
22. Sundqvist, B., Roepstorff, P., Fohlman, J., Hedin, A., Hakansson, P., Kamensky, I., Lindberg, M., Salehpour, M. & Sawe, G. (1984) *Science* **226,** 696–698.
23. Savitzky, A. & Golay, M. J. E. (1964) *Anal. Chem.* **36,** 1627–1639.
24. Hennessey, J. P., Jr. & Johnson, W. C., Jr. (1981) *Biochemistry* **20,** 1085–1094.
25. Manavalan, P. & Johnson, W. C., Jr. (1987) *Anal. Biochem.* **167,** 76–85.
26. Johnson, W. C., Jr. (1990) *Proteins Struct. Funct. Genet.* **7,** 205–214.
27. Chen, Y. J., Yang, J. T. & Chau, K. H. (1974) *Biochemistry* **13,** 3350–3359.
28. Zimmerman, S. S. & Scheraga, N. A. (1977) *Proc. Natl. Acad. Sci. USA* **84,** 8075–8079.
29. Scholtz, J. M., Qian, H., York, E. J., Stewart, J. M. & Baldwin, R. L. (1991) *Biopolymers* **31,** 1463–1470.
30. Merutka, G., Shalongo, W. & Stellwagen, E. (1991) *Biochemistry* **30,** 4245–4248.
31. Dyson, H. J., Merutka, G., Waltho, J. P., Lerner, R. A. & Wright, P. E. (1992) *J. Mol. Biol.* **226,** 795–817.
32. Nelson, J. W. & Kallenbach, N. R. (1989) *Biochemistry* **28,** 5256–5261.
33. Barrow, C. J., Yasuda, A., Kenny, P. T. M. & Zagorski, M. G. (1992) *J. Mol. Biol.* **225,** 1075–1093.
34. Wei, J. & Fasman, G. D. (1995) *Biochemistry* **34,** 6408–6415.
35. Storrs, R. W., Truckses, D. & Wemmer, D. E. (1992) *Biopolymers* **32,** 1695–1702.