Protein Science (1998), 7:123–131. Cambridge University Press. Printed in the USA.

Folded conformations of antigenic peptides from riboflavin carrier protein in aqueous hexafluoroacetone

SURAJIT BHATTACHARJYA,[1] SATISH K. AWASTHI,[1] P. RADHAKANTHA ADIGA,[2] AND PADMANABHAN BALARAM[1]

[1]Molecular Biophysics Unit, Indian Institute of Science, Bangalore, 560-012, India
[2]Center for Reproductive Biology and Molecular Endocrinology, Indian Institute of Science, Bangalore, 560-012, India

(RECEIVED May 19, 1997; ACCEPTED September 17, 1997)

Abstract

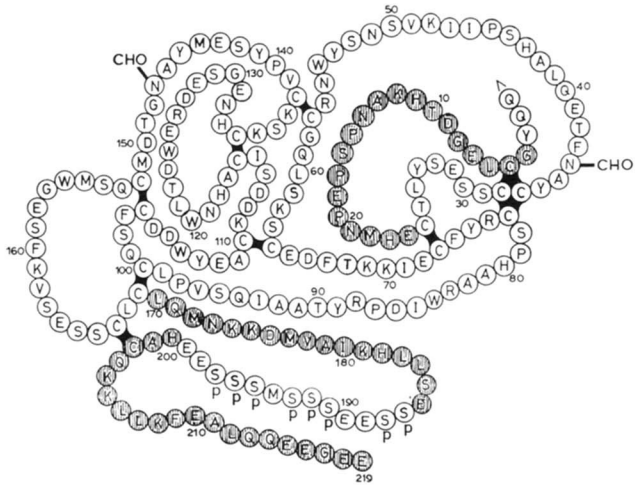
Riboflavin carrier protein (RCP) plays an important role in transporting vitamin B_2 across placental membranes, a process critical for maintenance of pregnancy. Association of the vitamin with the carrier protein ensures optimal bioavailability, facilitating transport. The conformations of three antigenic peptide fragments encompassing residues 4–23 (N21), 170–186 (R18), and 200–219 (Y21) from RCP, which have earlier been studied as potential leads toward a synthetic peptide-based contraceptive vaccine, have been investigated using CD and NMR spectroscopy in aqueous solution and in the presence of the structure-stabilizing cosolvent hexafluoroacetone trihydrate (HFA). In aqueous solution at pH 3.0, all three peptides are largely unstructured, with limited helical population for the peptides R18 and Y21. The percentage of helicity estimated from CD experiments is 10% for both the peptides. A dramatic structural transition from an unstructured state to a helical state is achieved with addition of HFA, as evidenced by intensification of CD bands at 222 nm and 208 nm for Y21 and R18. The structural transition is completed at 50% HFA (v/v) with 40% and 35% helicity for R18 and Y21, respectively. No structural change is evident for the peptide N21, even in the presence of HFA. NMR analysis of the three peptides in 50% HFA confirms a helical conformation of R18 and Y21, as is evident from upfield shifts of $C^{\alpha}H$ resonances and the presence of many sequential NH/NH NOEs with many medium-range NOEs. The helical conformation is well established at the center of the sequence, with substantial fraying at the termini for both the peptides. An extended conformation is suggested for the N21 peptide from NMR studies. The helical region of both the peptides (R18, Y21) comprises the core epitopic sequence recognized by the respective monoclonal antibodies. These results shed some light on the issue of structure and folding of antigenic peptides.

Keywords: antigenic peptides; hexafluoroacetone; NMR; protein folding; RCP

The conformational behavior of peptide fragments from protein sequence has been examined frequently as model systems to understand the mechanism of protein folding (Dyson et al., 1992a, 1992b; Cox et al., 1993; Kemmink & Creighton, 1993, 1995; Waltho et al., 1993; Yang et al., 1995). The observation of local conformational preferences in short linear peptides in solution, in the absence of medium- and long-range interactions found in proteins, has been correlated with the initiation of the folding process. Structure formation in short peptides that comprise antigenic segments in protein is of particular interest in establishing conformational determinants for antigenicity. Antibodies generated against short chemically synthesized peptide fragments from proteins are often able to recognize cognate sequences in the folded native state (Lerner, 1982). This observation has stimulated a great deal of interest in the development of peptide vaccines (Shinnick et al., 1983). Correlation of native state protein crystal structure with epitope mapping studies had initially suggested the surface-exposed and more disordered regions are likely to be more potent antigenic determinants (Tainer et al., 1984, 1985). Seminal studies from different groups have subsequently established well-ordered conformations for many such antigenic peptides in solution (Dyson et al., 1985, 1988, 1990; McInnes et al., 1993). A structured backbone conformation has also been demonstrated by NMR and crystallographic studies of peptides bound to monoclonal Fab fragments (Anglister et al., 1988; Stanfield et al., 1990; Jin et al., 1992; Scherf et al., 1992; Tsang et al., 1992). The incorporation of specific stereochemical constraints into antigenic peptides that stabilize "native like" conformational states appears to be a viable tool for enhancing immunological characteristics (Kaumaya et al., 1990, 1992; Tuchscherer et al., 1992; Gurunath et al., 1995).

Chicken riboflavin carrier protein (RCP) (Fig. 1) is a 219-residue phosphoglycoprotein with nine disulfide bonds (Murthy & Adiga, 1977; White & Merrill, 1988). The primary structure and the disulfide linkages are determined by chemical methods (Hamazume et al., 1984, 1987). RCP plays a vital role in the transportation of vitamin B_2 (riboflavin) across the placental membrane, in

Reprint requests to: Padmanabhan Balaram, Molecular Biophysics Unit, Indian Institute of Science, Bangalore 560-012; e-mail: pb@mbu.iisc.ernet.in.



N 21: GCLEGDTHKANPSPEPNMH*Y* (4-23)
R 18: *Y*LQMNKKDMVAIKHLLSE (170-186)
Y21: *Y*HACQKKLLKFEALQQEEGEE(200-219)

Fig. 1. The covalent structure of the chicken egg white riboflavin carrier protein (RCP) showing the location of the disulfide bonds (reproduced from Hamazume et al., 1984). The three peptide segments that are synthesized are hatched. The N-terminal fragment (N21) encompasses residues 4–23, terminated with a Tyr residue at the C terminus (GCLEGDTHKAN PSPEPNMHEY). An 18-residue fragment (R18) and a 21-residue fragment (Y21) from the C terminus of the protein encompass residues 170–186 (YLQMNKKDMVAIKHLLSE) and 200–219 (YHACQKKLLKF EALQQEEGEE), respectively, with a Tyr at the N termini. Terminal Tyr residues were incorporated to help measure peptide concentrations accurately and permit conjugation to carrier proteins for production of antibodies.

a process critical for embryonic development and maintenance of pregnancy (Murthy & Adiga, 1982; Adiga et al., 1988; Adiga, 1994). Antibodies specific to either RCP or reduced carboxymethylated protein are able to terminate pregnancy in rodents and monkey (Karande et al., 1991). Therefore, RCP provides a potential target for a contraceptive vaccine. In a systematic approach, we have studied several peptide fragments of RCP to look for potent antigenic fragments. Two 21-residue peptides, N21 and Y21 from the N (4–24) and C (200–219) termini, respectively, of the RCP (Fig. 1), are highly immunogenic (Gurunath et al., 1995; Beena et al., 1996). The antipeptide antibodies raised against both the peptides are able to bind RCP in vitro and in vivo, resulting in pregnancy termination in mice (Beena et al., 1996). Another 18-residue peptide fragment, R18 (residues 170–186) is also capable of eliciting antibodies, which crossreact with the carboxymethylated derivative of RCP (unpubl. results). However, polyclonal antibodies raised against R18 do not or marginally crossreact with native RCP and are unable to terminate pregnancy in mice, establishing their inability to neutralize the folded protein in vitro and in vivo. Peptides Y21 and R18 have an N-terminal Tyr residue that is absent in the native protein sequence, introduced specifically for conjugation purposes. In N21, this Try residue is placed at the C terminus.

In this report, we describe the conformational characteristics of these three antigenic peptides by CD and NMR in water and in aqueous hexafluoroacetone hydrate (HFA). It is well known that fluoroalcohols such as 2,2,2 trifluoroethanol (TFE) can stabilize helical conformations in peptides (Nelson & Kallenbach, 1984; Sonnichsen et al., 1992; Kemmink & Creighton, 1995; Bolin et al., 1996; Rajan & Balaram, 1996) in aqueous solution, although stabilization of other secondary structures are also documented, albeit to a lesser extent (Blanco et al., 1994). We have recently observed that HFA can be used as a structure-stabilizing cosolvent analogous to TFE (Rajan et al., 1997). Our results suggest that all three peptides are predominantly unstructured, with limited population of helical conformations in the Y21 and R18 peptides in aqueous solution. A sharp transition is observed to a more helical conformation in Y21 and R18 peptides in the presence of HFA, whereas no structural change is apparent in the N21 peptide, with addition of HFA. The helical conformation is well defined in the middle of the sequence, with significant fraying at the N-terminal end. The structured region in both the peptides encompasses the core epitopic region recognized by monoclonal antibodies.

Results

CD

Figure 2 shows CD studies of the three peptides in aqueous solution and in HFA/H_2O mixture at low pH. In water, the peptides are predominantly unstructured, as evident from a single intense CD band at ~200 nm. A shallow minima at 222 nm could be interpreted as a low population of helical conformation (10%) in Y21 and R18, which is absent in N21. NMR studies indeed indicate a conformational exchange of Y21 in water (see below). A dramatic stabilization of helical structure is achieved at 50% HFA in Y21

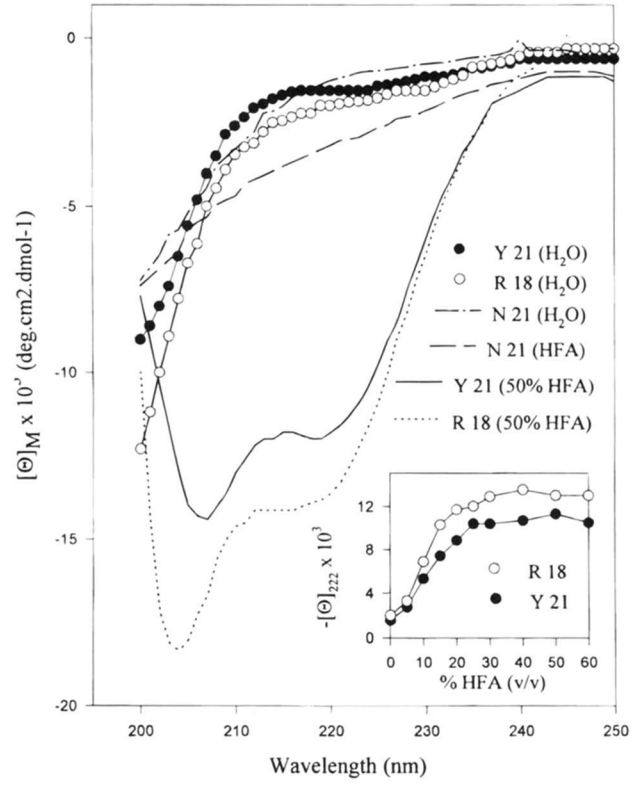


Fig. 2. Far-UV CD spectra of the peptides (N21, Y21, and R18) in aqueous solution and in 50% HFA (v/v) at pH 3.0, 298 K. The peptide concentrations were 60 μM. A pathlength of 1 mm was used to obtain the spectra. The structural transition of the R18 and Y21 peptides at different concentrations of HFA are shown in the inset.

and R18 peptides, as apparent from intensification of 222-nm and 208-nm CD bands, characteristics of helical conformations (Fig. 2). No such structural change is observed for the N21 fragment. The net helical content of the R18 and Y21 peptides is estimated from the ellipticity values at 222 nm to be 40% and 35%, respectively. A CD titration in aqueous HFA solutions indicates a sharp structural transition with a midpoint ~10–12% HFA (Fig. 2, inset). The transition is essentially completed by 50% HFA (v/v). We note that in 50% HFA/water, the mole ratio of the two solvents is ~1:10.

NMR studies

Structural characterization of the Y21 peptide has been performed in water and in a 50% HFA/water (v/v) solution. Severe spectral overlap in the $C^{\alpha}H$ region and occurrence of many identical amino acids (three L, five E, three K, and three Q) and degenerate spin systems precludes detailed structural analysis in water. However, resonances from some of the residues (H2, A3, C4, F11, A13, G19) can be identified either due to their unique spin systems or from the observations of sequential NOEs. The spin system of the three leucine and three lysine residues were also clear in the TOCSY spectra (data not shown). The chemical shift, coupling constant, and NOE data in water are therefore used, qualitatively. The chemical shifts of the $C^{\alpha}H$ proton of the peptides and proteins are highly conformation sensitive, an upfield or downfield shift in $C^{\alpha}H$ resonance from the random coil values are suggestive of either a helical or β-sheet conformation, respectively (Wishart et al., 1991). The extent of deviation from the random coil shift is related to the population of conformers present in solution. The chemical shift of the $C^{\alpha}H$ protons of the three leucine, three lysines, F11, and A13 show an upfield shift of 0.03–0.04 ppm compared with their random coil values (Wüthrich, 1986), indicating a small population of helical conformations. Many of the coupling constant values obtained from resolution-enhanced 1D NMR spectra are of intermediate range (5.5–7.0 Hz), suggesting a conformational exchange between an extended conformation and a helical conformation. The presence of some weak NH/NH NOEs between the amide protons presumably indicates a population of multiple turns or "nascent helical" conformations (Dyson et al., 1988) of the peptide in water. Taken together, NMR and CD results are suggestive of largely unstructured peptide in water with limited segments showing some evidence for population of helical conformation. Because peptides R18 and N21 yielded a CD spectra in H_2O similar to that of Y21, no NMR studies were performed under these conditions.

Conformation of the peptides in aqueous HFA

Sequence-specific assignments

All the NMR experiments of the peptides (Y21, R18) are done at 50% HFA because CD spectra indicate that a structural transition from an unstructured state to a helical state is complete at this solvent composition. Unlike water, the NMR spectra of the peptides in 50% HFA are well dispersed. Complete resonance assignments were achieved using TOCSY, double quantum filtered COSY, and NOESY spectra using standard procedures (Wüthrich, 1986). The intraresidue spin systems are initially identified in TOCSY spectra. In Y21, all five glutamic acids, three glutamines, three lysines, and three leucines are well resolved in the fingerprint region. Spectra are also recorded at different temperatures to resolve some of the overlaps, for instance, H2 resonance merged with water at 315 K, but can be seen at 303 K. The sequence-specific assignments to individual amino acids were made with the help of short-range sequential NOEs ($C^{\alpha}H_i$ to NH_{i+1} and NH_i/NH_{i+1}) and from many medium-range NOEs ($C^{\alpha}H_i$ to NH_{i+3}, $C^{\alpha}H_i$ to $C^{\beta}H_{i+3}$). Figures 3 and 4 show NH/NH and $C^{\alpha}H$/NH NOE connectivities of the two peptides. The NMR spectra of the N21 peptide in 50% HFA are poorly dispersed, precluding sequence-specific assignments for all the residues. The NOESY spectra mainly consists of strong sequential $C^{\alpha}H$/NH NOEs, characteris-

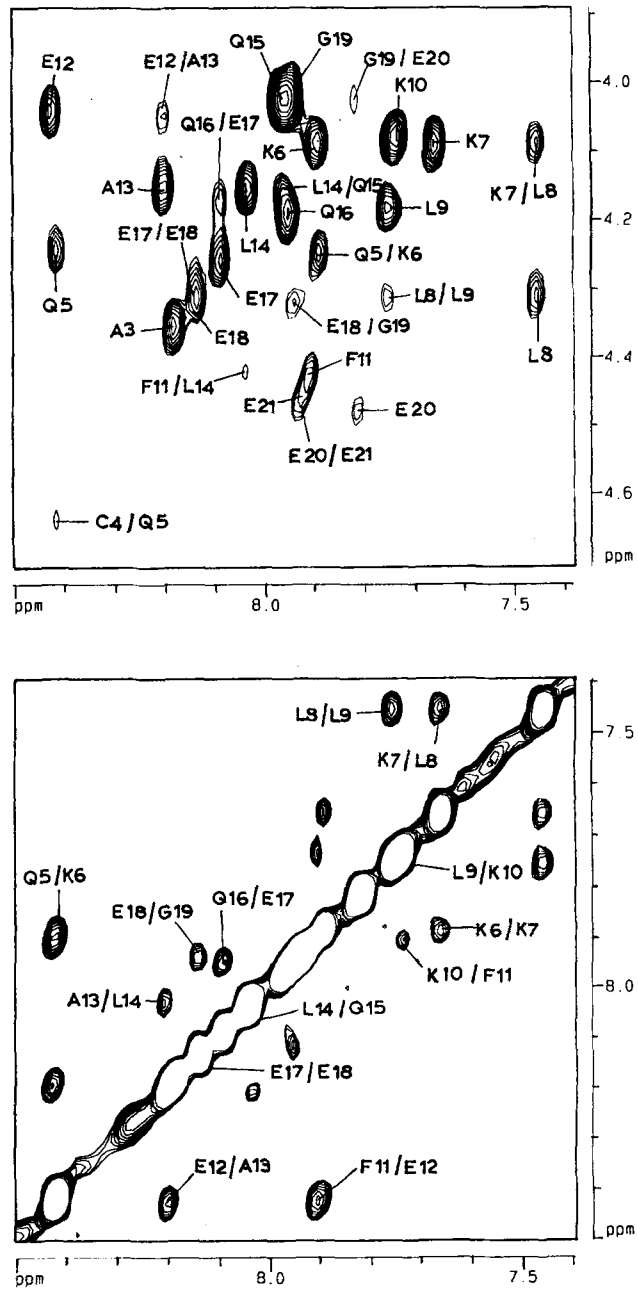


Fig. 3. Partial NOESY spectra of the peptide Y21 in 50% HFA (v/v) at pH 3.0, 310 K, indicating sequential and intraresidues $C^{\alpha}H$/NH NOEs (top) and NH/NH NOEs (bottom). A mixing time of 200 ms was used for the NOESY experiments.

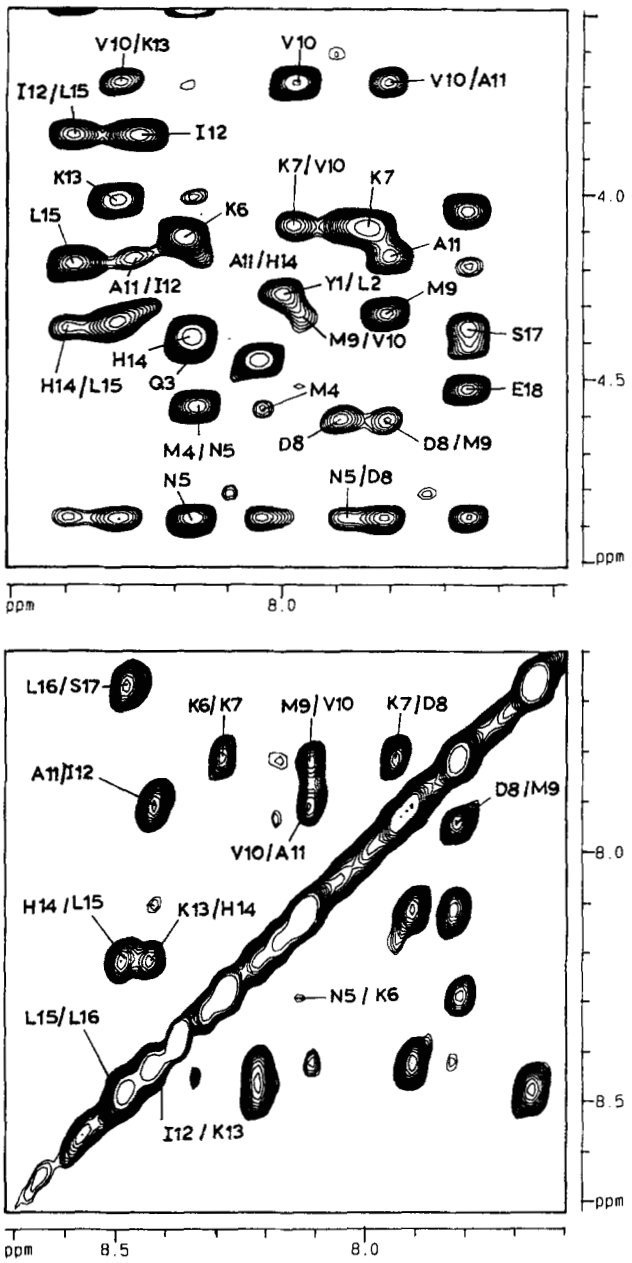


Fig. 4. Partial NOESY spectra of the peptide R18 in 50% HFA (v/v) at pH 3.0, 298 K, indicating $C^{\alpha}H/NH$ NOEs (top) and NH/NH NOEs (bottom). A mixing time of 200 ms was used for the NOESY experiments.

tic of an extended conformation. The presence of a centrally located $(X\text{-}Pro)_3$ segment in the sequence may stabilize formation of extended, polyproline-like helical structures in this peptide (Matsushima et al., 1990). The absence of α-helix induction even in HFA is consistent with the strong tendency of this sequence to favor a largely open extended conformation.

Chemical shifts and coupling constants

As mentioned above, chemical shifts of the $C^{\alpha}H$ protons are exquisitely sensitive to the conformation of the peptides and proteins (Wishart et al., 1991). In a stable helical conformation, the chemical shifts of the $C^{\alpha}H$ protons show a significant upfield shift from the specified random coil values (Wüthrich, 1986), whereas a downfield shift is characteristic of well-defined β-sheet structure (Perkins & Wüthrich, 1979; Dalgarno et al., 1983; Pardi et al., 1983; Wishart et al., 1991). Most of the $C^{\alpha}H$ protons of Y21 and R18 peptides experience a remarkable upfield shift (Fig. 5). In both the peptides, the chemical shifts of the central residues are more perturbed compared with the end residues. The chemical shifts of the end residues are either close to the random coil value, or moved downfield. This indicates that both peptides adopt a stable helical conformation encompassing the middle segment with flexible N and C termini. The conformational conclusion based on the chemical shifts is consistent with the NOE and coupling constants information (see below). The percentage of helicity estimated from the deviation of $C^{\alpha}H$ chemical shifts using the method given by Rizo et al. (1993) suggests that the R18 peptide is more helical (45%) compared with Y21 (35%), in close agreement with CD estimates.

The scalar coupling constant ($^3J_{HN\alpha}$) is another parameter to determine secondary structure, which is related directly to the dihedral angle ϕ of peptide backbone. The coupling constants can be obtained for some of the residues from 1D spectra. The coupling constants are less than 5 Hz for the central residues, whereas coupling constants are $\geq$7 Hz for the end residues. The low J values are indicative of a helical structure with extended ends possessing larger J. Residues for which J values could not be determined due to spectral overlap of the δ $C^{\alpha}H$ deviation from random coil values provide an estimate of Φ (Wishart et al., 1991).

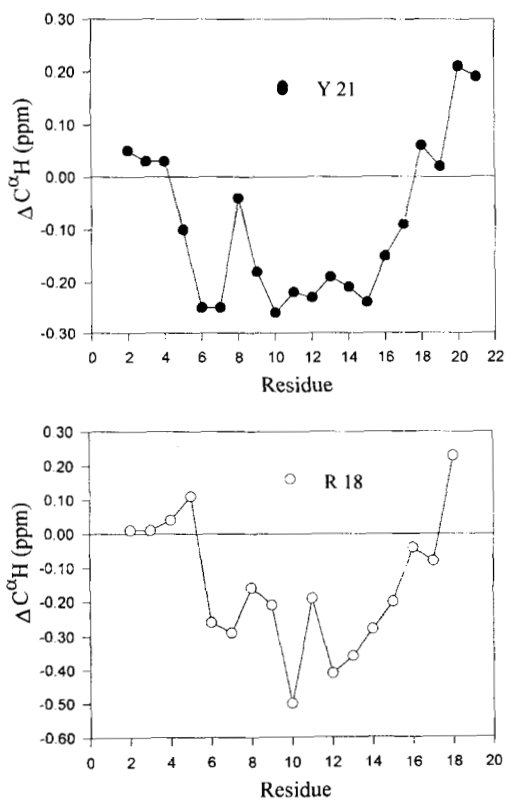


Fig. 5. $C^{\alpha}H$ chemical-shift deviations from the random coil values of the peptides Y21 and R18 in 50% (v/v) HFA, 310 K, pH 3.0.

For residues where both J and δ $C^{\alpha}H$ were determined, good correlations were observed for Φ values.

NOEs

The presence of strong, sequential (i, $i + 1$), NH/NH, $C^{\beta}H$/NH NOEs and weak $C^{\alpha}H$/NH NOEs, along with strong intraresidue $C^{\alpha}H$/NH NOEs are indicative of a significant population of conformations in the helical region of ϕ,ψ space (Dyson et al., 1988; Merutka et al., 1993). However, observation of medium-range NOEs $C^{\alpha}H$/NH (i, $i + 3$), $C^{\alpha}H/C^{\beta}H$ (i, $i + 3$) are indeed indicative of a folded helical conformation. The NOESY spectra (Figs. 3, 4) of Y21 and R18 peptides show strong sequential NH/NH NOEs and weak $C^{\alpha}H$/NH NOEs, along with many medium-range NOEs establishing a helical conformation of the peptides. The sequential NH/NH NOEs in Y21 start from Q5 and continue up to E18, with some interruptions due to overlap. The NH/NH NOEs in R18 also begin at the N-terminal end, with a weak NOE between N5/K6, followed by strong NOEs up to L16. All the medium-range NOEs cannot be observed due to spectral overlap. A summary of the observed NOEs is given in Figures 6 and 7. The helical conformation is not uniformly propagated throughout the sequence. The observation of strong sequential $C^{\alpha}H$/NH NOEs and absence of NH/NH NOEs in the first four residues at the N terminus and the last two residues at the C terminus indicate an extended conformation at the ends. The fraying at ends of sequences is seen frequently in many peptide helices in solution (Dyson et al., 1988, 1992a; Waltho et al., 1993). The helical conformation in the Y21 peptide spans from residue 5 to 18 and the helix is terminated by Gly 19. It is well known that glycine acts as a helix terminator in protein helices, adopting a positive ϕ angle (Aurora et al., 1994). Secondary structure prediction using the RCP sequence by the Chou–Fasman method also indicates a helix termination at the Gly 19 position (Fig. 8). The helical conformation in R18 peptide prevails from Q5 to L16. Both the peptides appears to favor α-helical conformation over 3_{10}-helical structure, because the characteristic $C^{\alpha}H_i/NH_{i+2}$ NOEs (Wüthrich, 1986; Merutka et al., 1993) expected in the latter are absent. The observation of some strong $C^{\alpha}H_i/C^{\beta}H_{i+3}$ NOEs is also suggestive of α-helical conformations of both the peptides ($d_{\alpha\beta}^{i,i+3}$ 2.5–4.2 Å for α-helix and 3.1–5.1 Å for 3_{10}-helix) (Wüthrich, 1986).

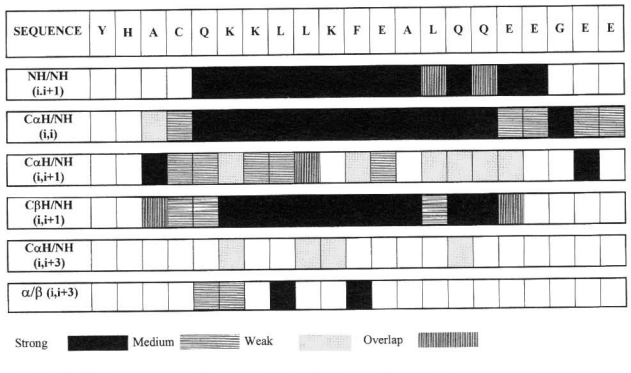


Fig. 6. NOE summary for peptide Y21 applied in the sequential assignment and in the secondary structure analysis. The amino acid sequence is shown at the top. The intensities of the NOEs are categorized as either strong, medium, or weak, and are marked accordingly by different shades.

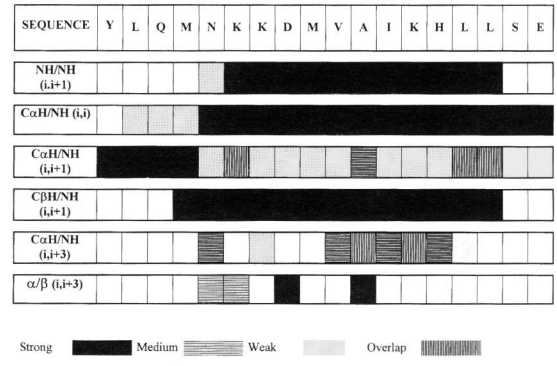


Fig. 7. NOE summary for peptide R18 applied in the sequential assignment and in the secondary structure analysis. The amino acid sequence is shown at the top. The intensities of the NOEs are categorized as either strong, medium, or weak, and are marked accordingly by different shades.

Structure determination

The structure of the peptides has been calculated using constrained molecular dynamics (MD) simulation procedures, imposing restraints from NOE-driven distances and dihedral angles. Figure 9 shows a superposition of backbone atoms of 20 simulated structures for peptides Y21 and R18. The conformation is well defined at the center of the sequences, with appreciable flexibility at the termini. The helical structure of R18 peptide is more rigid and the end residues do not show much conformational variability compared with Y21. The mean ϕ,ψ values calculated over 20 structures do not show any Ramachandran outliners (Ramachandran et al., 1963). The backbone dihedral angles for the ordered segments are well clustered around the allowed region of right-

N 21	G-C-L-E-G-D-T-H-K-A-N-P-S-P-E-P-N-M-H
CF-Pred	B-B-B-X-X-X-T-T-X-X-X-T-T-T-H-H-H-H-H
GOR-Pred	T-T-X-T-T-T-T-T-T-X-X-X-X-X-X-X-X-X-H
Y 21	H-A-C-Q-K-K-L-L-K-F-E-A-L-Q-Q-E-E-G-E-E
CF-Pred	H-H-H-H-H-H-H-H-H-H-H-H-H-H-H-H-X-X-X
GOR-Pred	H-H-H-H-H-H-H-H-H-H-H-H-H-H-H-H-H-H-H
R 18	L-Q-M-N-K-K-D-M-V-A-I-K-H-L-L-S-E
CF-Pred	H-H-H-H-H-H-H-H-B-B-B-B-B-B-B-H-H
GOR-Pred	H-H-H-H-H-H-H-H-H-H-H-H-H-H-H-X-X

Fig. 8. Secondary structure prediction of the three RCP peptides (N21, R18, and Y21) by Chou–Fasman and Garnier–Osguthrope–Robson (GOR) methods. Helical, β-sheet, and turns conformations are designated H, B, and T, respectively. Residues for which no prediction can be made are indicated as X.

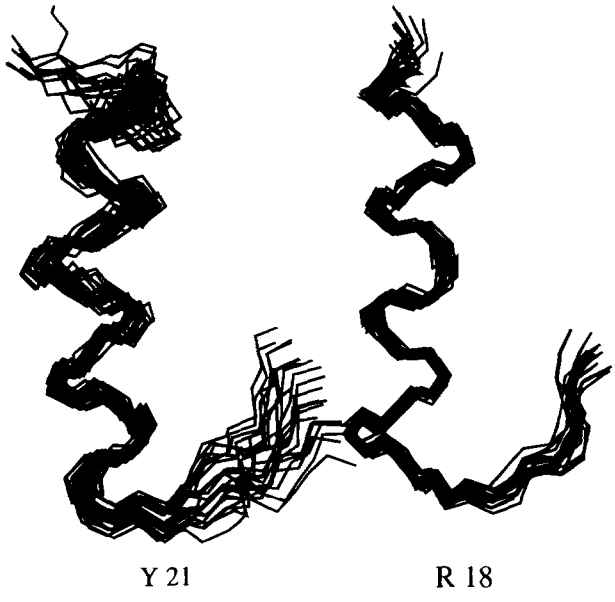


Fig. 9. Superposition of the backbone atoms of 20 constrained MD-simulated structures of the peptides Y21 and R18. The structures superpose closely at the central region of the peptides with an RMSD of 0.8 Å and 0.4 Å for Y21 and R18, respectively.

handed helical conformations in the Ramachandran map, whereas the other residues favor extended conformations.

Discussion

HFA as a structure stabilizer in peptides

Oligopeptides in aqueous solution generally exist as an ensemble of extended conformations due to the strong solvating tendency of water. Synthetic peptide fragments corresponding to structured segments in proteins frequently do not retain the "native"-like conformations in aqueous solution. TFE has been used widely as a stabilizer of helical conformations in aqueous solution of short peptides (Nelson & Kallenbach, 1984; Sonnichsen et al., 1992; Rajan & Balaram, 1996). Helix induction by TFE has been shown to correlate well with theoretically calculated helix propensities in short peptide sequences (Luidens et al., 1996). It has been suggested that TFE induces secondary structure formation by virtue of its hydrophobic character due to the presence of the trifluoromethyl (CF_3) group, which permits the fluoroalcohol to selectively solvate peptides even in the presence of a large excess of water. Further, the inability of the hydroxyl group of TFE to accept hydrogen bonds precludes its insertion into intramolecular peptide hydrogen bonds (Cammers-Goodwin et al., 1996; Rajan & Balaram, 1996; Rajan et al., 1997). These characteristics of TFE are indeed magnified in HFA hydrate, which is a significantly more potent stabilizer of helical peptides. HFA is a potentially amphipathic molecule, with a "teflon-like" hydrophobic face and a polar hydrophilic water-soluble face (Bhattacharjya & Balaram, 1997; Rajan et al., 1997). HFA is an excellent and economical NMR solvent, because it lacks nonexchangeable hydrogens. The results presented in this paper provide a clear demonstration of the ability of HFA to stabilize helical conformations in aqueous solutions in potentially helicogenic sequences. The conformational propensities of the sequences determined using the prediction methods of Chou and Fasman (1978) and Garnier et al. (1978) indicate a highly helical conformation of Y21 and R18 and multiple turn structures for the N21 peptide. Interestingly, termination of helical conformations obtained by NMR correlates well with the prediction algorithms. It has been shown in the case of ribonuclease S peptide that the helix stop signal remains intact in TFE-stabilized helical conformations (Nelson & Kallenbach, 1984).

Correlation of antigenicity and peptide structures

Antigenic peptide fragments have been shown to have an innate tendency to adopt a folded conformation either in solution or in the bound state with antibodies and B cell receptors (Dyson et al., 1988; Zilber et al., 1990). A proper understanding of the structural aspects of peptide antigens may be useful in designing rational approaches enhancing antigenicity by modifying peptide structures (Kaumaya et al., 1990; Tuchscherer et al., 1992). Our interest in the antigenic peptide fragments from RCP derives from the role of this protein in maintenance of pregnancy. Polyclonal antibodies specific for the N- (N21) and C-terminal (Y21) segments of the protein are able to crossreact with native protein in vitro and in vivo, and are efficient in terminating pregnancy in monkeys and mice (Beena et al., 1996). The antipeptide antibodies raised against the R18 peptide, close to the C terminus, do not appreciably interact with RCP in vitro, but bind to the reduced carboxymethylated protein (unpubl. results). Y21 and R18 in aqueous solution are largely unstructured. The NMR study of Y21 in water shows evidence of conformational exchange between extended conformations and a low population of folded helical conformations. The weakly helical conformations observed in water are dramatically stabilized in 50% HFA. The helix is extended from (5–18) in Y21 and (5–16) in R18. The CD and NMR results suggest that the R18 peptide is more helical compared with Y21. The helical structure is also well-defined in R18 as indicated by smaller variations in backbone dihedral angles and low backbone RMSD.

How are the structures of RCP peptides obtained in an HFA/H_2O mixture relevant to antigenic function? Epitope mapping experiments in the case of RCP using overlapping octapeptide fragments identified the $Q^5KKLLFE^{11}$ sequence as the core epitopic segment in Y21 (Beena et al., 1996). A similar analysis for R18 peptide recognized $N^5KKDMVAIK^{13}$ sequence as the principal immunodominant site (S. Lobo & P.R. Adiga, unpubl. results). The core epitopic regions in both the peptides adopt folded helical conformations in HFA solution. In the case of peptide N21, there is no evidence for α-helical folding even in 50% HFA. The presence of a $(Pro\text{-}X)_3$ segment is undoubtedly a strong structural determinant in this sequence, favoring an extended polyproline-like conformation. Interestingly, the core epitope mapped by using overlapping peptides corresponds to the Pro-rich segment (T.K. Beena & P.R. Adiga, unpubl. results). We have earlier demonstrated that the introduction of different helix-stabilizing factors (salt bridges, conformationally constrained amino acids) in the antigenic peptide Y21 indeed enhanced its antigenic potential (Gurunath et al., 1995). Conformational studies of other antigenic peptides also reveal that the sequence encompassing the antibody-binding region is more ordered compared with the rest of the sequence. Indeed, in the case of an immunogenic peptide fragment from myohemerythrin, Dyson et al. (1988) had shown that the peptide adopts a transient helical structure (nascent helix) in aque-

ous solution, which becomes stabilized in aqueous TFE solution. Only the part of the peptide that exhibits nascent helical conformations forms a stable helical structure in TFE, the N-terminal half of the peptide remains extended in water and TFE. This study also pointed out that the helical part of myohemerythrin fragment encompasses the epitope for a number of monoclonal antibodies, which recognize the intact protein. The other representative examples that suggest a well-ordered conformations of the antigenic sequences are a peptide from antigenic domain of Herpes simplex virus glucoprotein D-1 (Williamson et al., 1986), peptides from torpedo acetylcholine receptor (Chung et al., 1989, 1991), peptide from B-subunit of cholera toxin (Anglister et al., 1988; Zilber et al., 1990), the principle neutralizing determinant of HIV virus (Zvi et al., 1992), and peptide antigens from the receptor binding domain of *Pseudomonas aeruginosa* (McInnes et al., 1993).

Subsequent to the completion of the studies described in this paper, a crystal structure of RCP at 2.5-Å resolution has been reported (Monaco, 1997). Although coordinates and dihedral angles are not yet available, the structural description reveals that segments 176–183 (corresponds to 8–15 in R18) and 202–210 (corresponds to 5–13 in Y21) form helices. The NMR results on both peptides Y21 and R18 are in good agreement, but reveal somewhat more extended helical segments compared with the crystal structure. Interestingly, for the segment corresponding to peptide N21, no secondary structure assignment is provided in the description of the crystal structure (Monaco, 1997).

Implications for protein folding

The conformational characteristics of many peptide fragments from proteins are probed in aqueous solution or in mixtures of fluoroalcohols/water to understand protein folding pathways. The framework model of protein folding suggests formation of local secondary structures to nucleate folding, which are subsequently further stabilized or rearranged during the formation of tertiary structure (Ptitsyn & Finkelstein, 1980; Kim & Baldwin, 1982). Observation of transient or stable structures in isolated peptides without any long-range interaction supports initiation of folding processes determined by short-range interaction, justifying the framework model (Dyson et al., 1985, 1988; Osterhout et al., 1989). Peptide fragments comprising the entire sequences of proteins have indeed been investigated to establish sequences with stable secondary structures (Dyson et al., 1992a; Yang et al., 1994, 1995; Kemmink & Creighton, 1995; Munoz et al., 1995; Doak et al., 1996). In an alternative mechanism, the so-called "hydrophobic collapse" model (Dill et al., 1993), the burial of the hydrophobic residues from water to form a collapsed state is the primary event, followed by formation of stable secondary structure. The distinction between the two processes is often difficult due to the fast process of folding. However, formation of such nonspecific hydrophobic clusters are also shown in folding studies of some proteins (Lumb & Kim, 1994; Agashe et al., 1995) and in isolated peptides (Kemmink et al., 1993). The evidence for transient secondary structures in peptide sequences in water and their subsequent stabilization by HFA and other fluoroalcohols presumably indicates that both the mechanisms play an important role in the folding process. The formation of stable secondary structures in water is difficult, as a result of the competing ability of water to form hydrogen bonds with the peptide backbone. Structure stabilization achieved by fluoroalcohols can be viewed as a consequence of their "desiccating properties," which result in a drying out of the peptide backbone, providing hydrophobic surroundings to form stable secondary structure. Solvents like HFA, which can exhibit amphipathically, forming a hydrophobic "teflon-like" face and a hydrophilic hydrogen bonding face, are particularly effective in providing a hydrophobic milieu in which peptide folding is facilitated (Bhattacharjya & Balaram, 1997; Rajan et al., 1997). Similar phenomena are presumably important during the early stages of protein folding. Short peptide sequences can form transient secondary structure, which exists in conformational exchange with largely unstructured states. Detection of these transient structures are most often experimentally difficult. However, these transiently formed conformations can be stabilized by hydrophobic collapse, which effectively drives out water, facilitating formation of stable secondary structures.

Materials and methods

Ultrosyn KA resin, all Fmoc side-chain protected pentafluorophenyl active esters, DMF, piperidine, 1-hydroxybenztriazole, and trifluoroacetic acid (TFA) were purchased from Pharmacia Biochrom (Sweden). The HFA (a covlent hydrate of the ketone, a gem diol, hexafluoropropan-2,2-diol) was obtained from Aldrich Chemical Co.

Peptide synthesis

Peptides were synthesized on a LKB-Biolynx 4175 semi-automated peptide synthesier using Fmoc chemistry. Side-chain protecting groups were cleaved by treatment with 95% TFA for 3 h. The purity of the peptides was checked in an analytical HPLC column. Peptides were characterized by amino acid sequencing in the case of Y21 and N21, and by electrospray or MALDI-TOF mass spectrometry in the case of Y21 and R18.

CD

All CD spectra were recorded on a JASCO-J-500A spectropolarimeter using a cell pathlength of 1 mm. Peptide concentrations were determined by tyrosine absorbance at 275 nm (ϵ_{275} = 1,340 M^{-1} cm^{-1}). Peptide concentrations for CD measurements ranged from 40 to 70 μM. The pH of the samples was adjusted to a minimum of 3 after addition of HFA. Mean residue ellipticity was plotted as a function of wavelength or HFA concentration. The percentage of helicity was estimated from θ_{222}, taking 31,580 $deg \cdot cm^2 \cdot dmol^{-1}$ as 100% helix (Sonnichsen et al., 1992).

NMR spectroscopy

^{1}H-NMR spectra were recorded on a Bruker AMX 400 spectrometer. The NMR spectra were recorded either in 90% H_2O/10% D_2O or in 40% H_2O/10% D_2O/50% HFA. The pH of all samples was adjusted to 3 by addition of 0.1 M HCl. A set of temperature-dependent 1D experiments was done to fix the temperature for the 2D experiments. Two-dimensional double-quantum filtered COSY, TOCSY, and NOESY (Cavanagh et al., 1996) experiments were performed at different temperatures (298 K, 303 K, and 315 K) for complete resonance assignment. NOESY experiments were done at 200-ms and 300-ms mixing times. A 70-ms mixing time was used for the TOCSY experiments. The residual water was suppressed by presaturating the water resonance by a 55-db pulse in a recycle delay of 1.5 s. Chemical shifts are with reference to sodium

3-(trimethylsilyl)-propionate-d4 (TSP). All the 2D data were acquired at 1 K × 512 data points, with 40–72 transients using a 5500–5000 Hz spectral width. Two-dimensional data were zero filled to 1 K × 1 K data points; the FIDs were multiplied by sine $\pi/4$ function prior to Fourier transformation.

Determination of NMR constraints

Depending on the crosspeak intensities, the NOEs were classified into three different distance categories: strong, 2.0–3.0 Å; medium, 2.5–3.5 Å; and weak, 3.5–4.5 Å. In case of overlap, pseudoatoms were used, allowing 1-Å relaxation in distance criteria. A total number of 135 and 163 NOEs were used for structure determination of Y21 and R18, respectively. All the backbone to backbone, backbone to side-chain, and side-chain to side-chain NOEs were considered. The dihedral angles (ϕ) used as constraints were calculated from $^3J_{NH\alpha}$ using a modified Karplus-type equation (Pardi et al., 1984) and an empirical relationship given by Wishart et al. (1991) relating the chemical shift deviation of the $C^\alpha H$ proton and the dihedral angle ϕ, $\Delta\delta = -0.013\phi - 1.20$, where $\Delta\delta$ is the deviation of chemical shift of $C^\alpha H$ from its random coil value. The spectral overlap precluded determination of J for all the residues. In such cases, we have calculated ϕ using the above relationship. ϕ values obtained using both methods vary only by $\pm 10°$.

Structure calculation

All the computations were performed on an Iris Silicon Graphics workstation using the INSIGHT II molecular modeling program. CVFF force field was used for energy minimization and MD simulations were performed with the Discover module. A biharmonic skewed potential function was used in the total energy expression to include distance constraints. All the peptide bonds were maintained in a *trans* geometry throughout. A force constant of 10–20 kcal/mol was used on all constraints as a penalty function for violation of distance criteria. A few cycles of steepest-descent energy minimization were done without any constraints to relax short contacts from the initial structures. Different starting structures were used from extended to helical conformations for structure calculation with imposed constraints. The first and last residues were fixed in extended conformations. The constrained MD simulations were performed at 300 K for 200 ps with a 25-ps equilibration period. All the constraints were not imposed together. The short-range distance constraints were used to generate an initial set of structures, which were further refined by medium- and long-range NOE distances. The geometry of the structures and the restraint violations were examined in each step of simulation. A set of 50 structures were generated, of which 20 structures showed low RMS deviation from the average structure. The "stereochemical goodness" of the 20 different structures was evaluated by calculating the ϕ,ψ dihedral angles.

Acknowledgments

This work is supported in part by a grant from the Department of Biotechnology, India (PRA). Dr. David King (Berkeley) and Dr. Paul Humphery (Kratos Analytical, UK) are gratefully acknowledged for the mass spectral characterizations. The NMR experiments and computational studies were performed in Sophisticated Instruments Facility (SIF), Indian Institute of Science, Bangalore.

References

Adiga PR. 1994. In: *Vitamin receptors: Vitamins as ligands in cell communication*. New York: Cambridge University Press.

Adiga PR, Karande AA, Visweswariah SS, Sheshagiri PB. 1988. In: *Perspectives in primate reproductive biology*. New Delhi: Wiley Eastern. pp 129–140.

Agashe VR, Shastry MCR, Udgaonkar JB. 1995. Initial hydrophobic collapse in the folding of barstar. *Nature 377*:754–757.

Anglister J, Jacob C, Assulin O, Ast G, Pinker R, Arnon R. 1988. NMR study of the complex between a synthetic peptide derived from the B subunit of cholera toxin and three monoclonal antibodies aganist it. *Biochemistry 27*:717–724.

Aurora R, Srinivasan R, Rose GD. 1994. Rules for α-helix termination by glycine. *Science 264*:1126–1130.

Beena TK, Karande AA, Adiga PR. 1996. Antigenic determinants at the carboxy terminus of chicken egg white riboflavin carrier protein (RCP): Epitope mapping and antibody-mediated pregnancy in rodents. *Vaccines 14*:307–312.

Bhattacharjya S, Balaram P. 1997. Hexafluoroacetone hydrate as a structure modifier in proteins: Characterization of a molten globule state of hen egg-white lysozyme. *Protein Sci 6*:1065–1073.

Blanco FJ, Jimenez MA, Pineda A, Rico M, Santoro J, Nieto JL. 1994. NMR solution structure of the isolated N-terminal fragment of protein-G B_1 domain: Evidence of trifluoroethanol induced native-like β-hairpin formation. *Biochemistry 33*:6004–6014.

Bolin KA, Pitkeathly M, Mirankar A, Smith LJ, Dobson CM. 1996. Insight into a random coil conformation and an isolated helix: Structural and dynamical characterization of the C-helix peptide from hen lysozyme. *J Mol Biol 261*:443–453.

Cavanagh J, Fairbrother WJ, Palmer AG III, Skelton NJ. 1996. In: *Protein NMR spectroscopy: Principle & practice*. San Diego, California: Academic Press.

Chou PY, Fasman GD. 1978. Empirical predictions of protein conformations. *Annu Rev Biochem 47*:251–276.

Chung MT, Demange P, Marraud M, Tsikirais V, Sakarellos C, Papdouli I, Kokla A, Tzartos SJ. 1991. Two-dimensional ^{1}H NMR study of antigen antibody interactions: Binding of synthetic decapeptides to an anti acetylcholine receptor monoclonal antibody. *Biopolymers 31*:769–776.

Chung MT, Marraud M, Hadjidakis I, Bairaktari E, Sakarellos C, Kokla A, Tzartos SJ. 1989. Two-dimensional ^{1}H NMR study of synthetic peptides containing the main immunogenic region of the torpedo acetyl choline receptor. *Biopolymers 28*:465–478.

Cox J, Evans PA, Packman LC, Williams DH, Woolfson DN. 1993. Dissecting the structure of a partially folded protein: Circular dichroism and nuclear magnetic resonance studies of peptides fron ubiquitin. *J Mol Biol 234*:483–492.

Dalgarno DC, Levine BA, Williams RJP. 1983. Structural information from NMR secondary chemical shifts of peptides α C-H protons in proteins. *Biosci Rep 3*:443–452.

Dill KA, Fiebig K, Chan HS. 1993. Cooperativity in protein-folding kinetics. *Proc Natl Acad Sci USA 90*:1942–1946.

Doak DG, Mulvey D, Kawaguchi K, Villalain J, Campbell ID. 1996. Structural studies of synthetic peptides dissected from the voltage-gated sodium channel. *J Mol Biol 258*:672–687.

Dyson HJ, Cross KJ, Houghten RA, Wilson IA, Wright PE, Lerner RA. 1985. The immunodominant site of a synthetic immunogen has a conformational preference in water for a type-II reverse turn. *Nature 318*:480–483.

Dyson HJ, Merutka G, Waltho JP, Lerner RA, Wright PE. 1992a. Folding of peptides fragments comprising the complete sequence of proteins: Models for initiation of protein folding I. Myohemerythrin. *J Mol Biol 226*:795–817.

Dyson HJ, Rance M, Houghten RA, Wright PE, Lerner RA. 1988. Folding of immunogenic peptide fragments of proteins in water solution. *J Mol Biol 201*:201–217.

Dyson HJ, Satterthwait AC, Lerner RA, Wright PE. 1990. Conformational preference of synthetic peptides derived from the immunodominant site of the circumsporozoite protein of *Plasmodium falciparum* by ^{1}H NMR. *Biochemistry 29*:7828–7837.

Dyson HJ, Sayre JR, Merutka G, Shin H, Lerner RA, Wright PE. 1992b. Folding of peptides fragments comprising the complete sequence of proteins: Models for initiation of protein folding II. Plastocyanin. *J Mol Biol 226*:819–835.

Garnier J, Osguthrope DJ, Robson B. 1978. Analysis of the accuracy and implications of simple models for predicting the secondary structure of globular proteins. *J Mol Biol 120*:97–120.

Goodwin CA, Allen TJ, Oslick SL, McClure KF, Lee JH, Kemp DS. 1996. Mechanism of stabilization of helical conformations of polypeptides by water containing trifluoroethanol. *J Am Chem Soc 118*:3082–3090.

Gurunath R, Beena TK, Adiga PR, Balaram P. 1995. Enhancing peptide antigenicity by helix stabilization. *FEBS Lett 361*:176–178.

Hamazume Y, Mega T, Ikenaka T. 1984. Characterization of hen egg white- and yolk-riboflavin binding proteins and amino acid sequence of egg white-riboflavin binding protein. *J Biochem 95*:1633–1641.

Hamazume Y, Mega T, Ikenaka T. 1987. Position of disulfide bonds in riboflavin-binding protein of hen egg white. *J Biochem 101*:217–223.

Jin L, Fendly BM, Well JA. 1992. High resolution functional analysis of antibodies–antigen interactions. *J Mol Biol 226*:851–865.

Karande AA, Kuzhandai V, Adiga PR. 1991. A monoclonal antibody recognizing the C-terminal region of chicken egg white riboflavin carrier protein terminates early pregnancy in mice. *Mol Immunol 28*:476–480.

Kaumaya PTP, Berndt K, Heindorn D, Trewhella J, Kezdy FJ, Goldberg E. 1990. Synthesis and biophysical characterization of engineered topographic immunogenic determinants with $\alpha\alpha$ topology. *Biochemistry 29*:13–23.

Kaumaya PTP, VanBuskrik A, Goldberg E, Pierce SK. 1992. Design and immunological properties of topographic determinants of a protein antigen LDH-C_4 as vaccines. *J Biol Chem 267*:6338–6346.

Kemmink J, Creighton TE. 1993. Conformations of peptides representing the entire sequence of bovine pancreatic trypsin inhibitor (BPTI) and their roles in folding. *J Mol Biol 234*:861–878.

Kemmink J, Creighton TE. 1995. Effects of trifluoroethanol on the conformation of peptides representing the entire sequence of bovine pancreatic trypsin inhibitor. *Biochemistry 34*:12630–12635.

Kemmink J, Mierlo C, Scheek RM, Creighton TE. 1993. Local structure due to an aromatic–amide interaction observed by ^{1}H-nuclear magnetic resonance spectroscopy in peptides related to the N terminus of bovine pancreatic trypsin inhibitor. *J Mol Biol 230*:312–322.

Kim PS, Baldwin RL. 1982. Specific intermediates in the folding reactions of small proteins and the mechanism of protein folding. *Annu Rev Biochem 51*:459–489.

Lerner RA. 1982. Trapping the immunological repertoire to produce antibodies of predetermined specificity. *Nature 299*:592–596.

Luidens MK, Figge J, Breese K, Vajda S. 1996. Predicted and trifluoroethanol induced α-helicity of polypeptides. *Biopolymers 39*:367–376.

Lumb KJ, Kim PS. 1994. Formation of a hydrophobic cluster in a denatured bovine pancreatic trypsin inhibitor. *J Mol Biol 236*:412–420.

Matsushima N, Creutz CE, Kretsinger RH. 1990. Polyproline, β-turn helices. Novel secondary structures proposed for the tandem repeats within rhodopsin, synaptophysin, synexin, gliadin, RNA polymerase III, hordein and gluten. *Proteins Struct Funct Genet 7*:125–155.

McInnes C, Sonnichsen FD, Kay CM, Hodges RS, Sykes BD. 1993. NMR solution structure and flexibility of a peptide antigen representing the receptor binding domain of *Pseudomonas aeruginosa*. *Biochemistry 32*:13432–13440.

Merutka G, Morikis D, Bruschweiler R, Wright PE. 1993. NMR evidences for multiple conformations in a highly helical model peptide. *Biochemistry 32*:13089–13097.

Monaco HL. 1997. Crystal structure of chicken riboflavin-binding protein. *EMBO J*. Forthcoming.

Munoz V, Serrano L, Jimenez MA, Rico M. 1995. Structural analysis of peptides encompassing all α-helices of three α/β parallel proteins: che-y, flavodoxin and p21-ras: Implications for α-helix stability and the folding of α/β parallel proteins. *J Mol Biol 247*:648–669.

Murthy CVR, Adiga PR. 1982. Pregnancy suppression by active immunization aganist gestation specific riboflavin carrier protein. *Science 216*:191–193.

Murthy US, Adiga PR. 1977. Riboflavin binding protein of hen's egg: Purification and radioimmunoassay. *Ind J Biochem Biophys 14*:118–124.

Nelson JW, Kallenbach NR. 1984. Persistence of the α-helix stop signal in the S-peptide in trifluoroethanol solution. *Biochemistry 28*:5256–5261.

Osterhout J, Baldwin RL, York E, Stewart JM, Dyson HJ, Wright PE. 1989. ^{1}H NMR studies of the solution conformations of an analogue of the C-peptide of ribonuclease A. *Biochemistry 28*:7059–7064.

Pardi A, Billeter M, Wüthrich K. 1984. Calibration of the angular dependence of the amide proton-C^α proton coupling constants, ^{3}JNHα, in a globular protein. Use of ^{3}JNHα for identification of helical secondary structure. *J Mol Biol 180*:741–751.

Pardi A, Wagner G, Wüthrich K. 1983. Protein conformation and proton nuclear-magnetic resonance chemical shifts. *Eur J Biochem 137*:445–454.

Perkins SJ, Wüthrich K. 1979. Ring current effects in the conformation dependent NMR chemical shifts of aliphatic protons in the basic pancreatic trypsin inhibitor. *Biochim Biophys Acta 576*:409–423.

Ptitsyn G, Finkelstein AV. 1980. Similarities of protein topologies: Evolutionary divergence, functional convergence or principles of folding? *Quart Rev Biophys 13*:339–386.

Rajan R, Awasthi SK, Bhattacharjya S, Balaram P. 1997. "Teflon coated peptides:" Hexafluoroacetone trihydrate as a structure stabilizer for peptides. *Biopolymers*. Forthcoming.

Rajan R, Balaram P. 1996. A model for interaction of trifluoroethanol with peptides and proteins. *Int J Peptide Protein Res 48*:328–336.

Ramachandran GN, Ramakrishan C, Sasisekharan V. 1963. Stereochemistry of polypeptide chain configurations. *J Mol Biol 7*:95–99.

Rizo J, Blanco FJ, Kobe B, Bruch MD, Gierasch LM. 1993. Conformational behavior of *Escherichia coli* omp A signal peptides in membrane mimetic environment. *Biochemistry 32*:4881–4894.

Scherf T, Hiller R, Naider F, Levitt M, Anglister J. 1992. Induced peptide conformation in different antibody complexes: Molecular modelling of the three-dimensional structure of peptide–antibody complexes using NMR-driven distance restraints. *Biochemistry 31*:6884–6897.

Shinnick TM, Sutcliffe JG, Green N, Lerner RA. 1983. Synthetic peptide immunogens as vaccines. *Annu Rev Microbiol 37*:425–446.

Sonnichsen FD, Van Eyk JE, Hodges RS, Sykes BD. 1992. Effect of trifluoroethanol on protein secondary structure: An NMR and CD study using a synthetic actin peptide. *Biochemistry 31*:8790–8798

Stanfield RL, Fieser TM, Lerner RA, Wilson IA. 1990. Crystal structures of an antibody to a peptide and its complex with peptide antigen at 2.8 Å. *Science 248*:712–719.

Tainer JA, Getzoff ED, Alexander H, Houghten RA, Olson AJ, Lerner RA, Hendrickson WA. 1984. The reactivity of anti-peptide antibodies is a function of the atomic mobility of sites in a protein. *Nature 312*:127–134.

Tainer JA, Getzoff ED, Paterson Y, Olson AJ, Lerner RA. 1985. The atomic mobility component of protein antigenicity. *Annu Rev Immunol 3*:501–535.

Tsang P, Rance M, Fieser TM, Ostresh JM, Houghten RA, Lerner RA, Wright PE. 1992. Conformational and dynamics of an Fab′-bound peptide by isotope-edited NMR spectroscopy. *Biochemistry 31*:3862–3871.

Tuchscherer G, Servis C, Corrandin G, Blum U, Rivier J, Mutter M. 1992. Total chemical synthesis, characterization, and immunological properties of an MHC class I model using the TASP concept for protein de novo design. *Protein Sci 1*:1377–1386.

Waltho JP, Feher VA, Merutka G, Dyson HJ, Wright PE. 1993. Peptide models of protein folding initiation sites. I. Secondary structure formation by peptides corresponding to the G- and H- helices of myoglobin. *Biochemistry 32*:6337–6347.

White HB, Merrill AH. 1988. Riboflavin binding proteins. *Annu Rev Nutr 8*:278–279.

Williamson MP, Hall MJ, Honda BK. 1986. ^{1}H NMR assignment and secondary structure of a Herpes simplex virus glycoprotein D-1 antigenic domain. *Eur J Biochem 158*:527–536.

Wishart DS, Sykes BD, Richards FM. 1991. Relationship between nuclear magnetic resonance chemical shift and protein secondary structure. *J Mol Biol 222*:311–333.

Wüthrich K. 1986. *NMR of proteins and nucleic acids*. New York: Wiley.

Yang JJ, Pitkeathly M, Radford SE. 1994. Far-UV circular dichroism reveals a conformational switch in a peptide fragment from the β-sheet of hen lysozyme. *Biochemistry 33*:7345–7353.

Yang JJ, Buck M, Pitkeathly M, Kotik M, Haynie DT, Dobson CM, Radford SE. 1995. Conformational properties of four peptides spanning the sequence of hen lysozyme. *J Mol Biol 252*:483–491.

Zilber B, Scherf T, Levitt M, Anglister J. 1990. NMR-derived model for a peptide–antibody complex. *Biochemistry 29*:10032–10041.

Zvi A, Hiller R, Anglister J. 1992. Solution conformation of a peptide corresponding to the principal neutralizing determinant of HIV-1_{IIIB}: A two-dimensional NMR study. *Biochemistry 31*:6972–6979.