Usman, N., Ogilvie, K. K., Jiang, M. Y., & Cedergren, R. L. (1987) *J. Am. Chem. Soc. 109*, 7845-7854.
Varani, G., Wimberly, B., & Tinoco, I., Jr. (1989) *Biochemistry 28*, 7760-7772.
Werntges, H., Steger, G., Riesner, D., & Fritz, H. J. (1986) *Nucleic Acids Res. 14*, 3773-3790.
Westhof, E., Dumas, P., & Moras, D. (1985) *J. Mol. Biol. 184*, 119-145.
Wickstrom, E. L., Bacon, T. A., Gonzalez, A., Freeman, D. L., Lyman, G. H., & Wickstrom, E. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 1028-1032.
Williamson, J. R., Raghuraman, M. K., & Cech, T. R. (1989) *Cell 59*, 871-880.
Woese, C. R., Winker, S., & Gutell, R. R. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 8467-8471.
Yager, T. D., & von Hippel, P. H. (1991) *Biochemistry 30*, 1097-1118.
Zuker, M. (1989) *Science 244*, 48-52.
Zuker, M., & Stiegler, P. (1981) *Nucleic Acids Res. 9*, 133-148.
Zuker, M., Jaeger, J. A., & Turner, D. H. (1991) *Nucleic Acids Res. 19*, 2707-2714.

Cell Adhesion Promoting Peptide GVKGDKGNPGWPGAP from the Collagen Type IV Triple Helix: Cis/Trans Proline-Induced Multiple ^{1}H NMR Conformations and Evidence for a KG/PG Multiple Turn Repeat Motif in the All-Trans Proline State†

Kevin H. Mayo* and Dennisse Parra-Diaz
Program in Molecular Pharmacology and Structural Biology, Department of Pharmacology, Jefferson Cancer Institute, Thomas Jefferson University, Life Sciences Building, Philadelphia, Pennsylvania 19107
James B. McCarthy and Mary Chelberg
Department of Lab Medicine and Pathology, University of Minnesota School of Medicine, Minneapolis, Minnesota 55455
Received January 9, 1991; Revised Manuscript Received May 31, 1991

ABSTRACT: Peptide GVKGDKGNPGWPGAPY (called peptide IV-H1), derived from the protein sequence of human collagen type IV, triple-helix domain residues 1263-1277, represents an RGD-independent, cell-specific, adhesion, spreading, and motility promoting domain in type IV collagen. In this study, peptide IV-H1 has been investigated by ^{1}H NMR (500 MHz) spectroscopy. Cis-trans proline isomerization at each of the three proline residues gives rise to a number of slowly exchanging (500-MHz NMR time scale) conformation states. At least five such states are observed, for example, for the well-resolved A14 βH_3 group, and K3, which is six residues sequentially removed from the nearest proline, i.e., P9, shows two sets. The presence of more than two sets of resonances for residues sequentially proximal to a proline, e.g., A14-*cis*-P15 and A14-*trans*-P15, and more than one set for a residue sequentially well-removed from a proline, e.g., K3, indicates long range conformation interactions and the presence of preferred structure in this short linear peptide. Many resonances belonging to these multiple species have been assigned by using mono-proline-substituted analogues. Conformational (isomer) state-specific 2D ^{1}H NMR assignments for the combination of cis and trans proline states have been made via analysis of COSY-type, HOHAHA, and NOESY spectra. Peptide IV-H1 in the all-trans proline state ttt exists in relatively well-defined conformation populations showing numerous short- and long-range NOEs and long-lived backbone amide protons and reduced backbone NH temperature coefficients, suggesting hydrogen-bonding, and structurally informative $^3J_{\alpha N}$ coupling constants. The NMR data indicate significant β-turn populations centered at K3-G4, K5-G6, P9-G10, and P12-G13, and a C-terminal γ-turn within the A14-P15-Y16 sequence. These NMR data are supported by circular dichroic studies which indicate the presence of 52% β-turn, 10% helix, and 38% random coil structural populations. Since equally spaced KG and PG residues are found on both sides of peptide IV-H1 in the native collagen type IV sequence, this multiple turn repeat motif may continue through a longer segment of the protein. Synthetic peptide IV-H1 overlapping sequence "walk throughs" indicate that the primary biological activity is localized in the GNPGWPGAP double β-turn domain, which contains the backbone constraining proline residues. This proline-domain conformation may suggest a collagen type IV receptor-specific, metastatic cell adhesion promoting binding domain.

Type IV collagen is a collagenous glycoprotein that forms the major scaffolding of basement membranes (Timpl et al., 1981; Yurchenco & Furthmayr, 1983; Tsilibary & Charonis, 1986) and binds to other components like laminin, entactin, heparan sulfate, etc. (Timpl & Dziadek, 1986; Charonis et al., 1985; Laurie et al., 1986; Fugiwara et al., 1984). Type IV collagen functions in part as an adhesive protein since it promotes cell adhesion and specifically interacts with the surfaces of a variety of cell types (Kurkinen, 1984; Aumailley & Timpl, 1986; Sugrue, 1987; Murray et al., 1979). Members of the integrin superfamily of membrane proteins that bind to multiple extracellular matrix components, such as fibronectin and laminin (Ruoslahti & Pierschbacher, 1987; Ignatius & Reichardt, 1988; Gehlsen et al., 1986), have recently been

† This work was supported by generous research grants from the W. W. Smith Charitable Trust, the Elsa Pardee Foundation, the Leukemia Task Force, and the National Cancer Institute (CA-43924), and benefited from NMR facilities made available to Temple University through Grant RR-04040 (to K.H.M.) from the National Institutes of Health. D.P.-D. is indebted to the Graduate School of Temple University for a Future Faculty Fellows postdoctoral fellowship award. Dr. Carl Burke of Merck Sharp & Dohme Labs, West Point, PA, is gratefully acknowledged for having run and analyzed the CD spectrum of peptide IV-H1.
* Address correspondence to this author.

0006-2960/91/0430-8251$02.50/0

suggested to serve as receptors for binding to type IV collagen (Tomaselli et al., 1987). A number of integrins recognize RGD-containing sequences (Ruoslahti & Pierschbacher, 1987) like those found in type IV collagen; however, in one report (Herbst et al., 1988) the attachment of bovine aortic endothelial cells to type IV collagen was not inhibited by GRGDS, suggesting that at least these cells do *not* use the RGD sequence for cell attachment to type IV collagen. A similar situation exists in laminin, where an RGD is present in the A-chain, but cells have not been shown to use this sequence (Gehlsen et al., 1986). This information would suggest that, for several cell types, binding is *not* primarily mediated by integrins or that these integrins do *not* recognize RGD-containing sequences of the protein.

Type IV collagen has an unusual structure containing a large noncollagenous NC1 domain at the C-terminal end and numerous interruptions along the length of the Gly-X_1-X_2 triple helical sequence (Brazel et al., 1987; Babel & Glanville, 1984; Killen et al., 1988). Such interruptions have been hypothesized to contribute to increased flexibility of type IV collagen when compared to other interstitial collagens (Hofmann et al., 1984; Kilchherr et al., 1985). Type IV collagen is also known to self-assemble, forming an irregular polygonal structure (Yurchenco & Furthmayr, 1984; Yurchenco & Rubin, 1987). Other than participate in cell adhesion and spreading, type IV collagen also promotes cell motility as tested by migration and haptotaxis in modified Boyden chambers (Herbst et al., 1988). Although it is not known what components of the cell surface are important for triggering cell motility in response to type IV collagen, this aspect of cell behavior is of particular significance for the movement of various cells across basement membranes.

Recently, an RGD-independent integrin-binding peptide, GVKGDKGNPGWPGAPY, was derived from the type IV collagen triple helix domain (Chelberg et al., 1990). This peptide, called IV-H1, is known specifically to promote melanoma cell adhesion, spreading, and motility. Due to its biological significance, we undertook an ^{1}H NMR[1] study of this 16-mer. Generally, short linear peptides like IV-H1 exist in solution in an ensemble of highly fluctuating structures whose NMR spectra represent some average conformation. Recently, however, several different secondary structures have been reported for linear peptides in water. Baldwin and coworkers have shown by CD and NMR that both the isolated S-peptide of ribonuclease (residues 1–20) and the shorter C-peptide derived from it are partially helical in water (Kim & Baldwin, 1982, 1984; Bierzynski et al., 1982; Shoemaker et al., 1985). NMR studies have given unequivocal evidence for a highly populated β-turn conformation in the immunodominant domain of an immunogenic peptide from influenza virus hemagglutinin (Dyson et al., 1985, 1986, 1988a,b). There, Dyson et al. (1988a) showed that β-turn formation is strongly dependent on the amino acid sequence and demonstrated that both type II and type VI turns can be surprisingly stable in water solution. Dyson et al. (1988b) also showed that a short synthetic peptide comprised of C-terminal residues 69–86 from myohemerythrin was able to form a nascent helix in water solution. In both of these cases, the peptides were highly immunogenic and induced antibodies reactive with the folded native protein. Other studies on short peptides derived from fibronectin (Reed et al., 1988) and from fibrinogen (Ni et al., 1988, 1989a; Mayo et al., 1990) had similar findings. Recently, a number of other short linear peptides have been shown to form "NMR"-stable structures in water [for example, Bruch et al. (1989); Mammi et al. (1989); Otter et al. (1989)].

Although biologically active amino acid sequences have been derived from type IV collagen, no structural information is available on possible average peptide backbone folding. While knowledge of these sequences is important, it is, afterall, the three-dimensional structure that confers biological activity and proper function. Three pieces of information made peptide IV-H1 interesting from an NMR solution structure determination standpoint: (1) the presence of two PG dipeptide sequences that have a high potential for β-turn structure (Chou & Fasman, 1978; Dyson et al., 1988a); (2) its significant biological activity as mentioned above; and (3) antipeptide antibodies made from this peptide cross-react with parent type IV collagen and inhibit biological activity (Chelberg et al., 1990), suggesting that this short linear peptide does possess some preferred conformation. This paper specifically addresses peptide IV-H1's NMR-derived solution structure(s).

Materials and Methods

Peptide Synthesis. Peptides representing amino acid sequences from human type IV collagen were synthesized with a Beckman 990 peptide synthesizer, either by Dr. Robert Wohlheuter at the Microchemical Facility of the University of Minnesota or by Dr. Bianca Conti-Tronconi (University of Minnesota at St. Paul). The procedures used were based on the Merrifield solid-phase system as described elsewhere (Stewart & Young, 1984). Lyophilized crude peptides were purified by preparative reverse-phase HPLC on a C-18 column, by using an elution gradient of 0–60% acetonitrile with 0.1% trifluoroacetic acid in water. The purity and composition of the peptides was verified by peptide sequencing and reverse-phase HPLC analysis.

NMR Spectroscopy. Freezed-dried samples for NMR measurements were dissolved in either D_2O or in H_2O/D_2O (9:1). Protein concentration was in the range of 10–40 mM. pH was adjusted by adding microliter quantities of NaOD or DCl to the protein sample. NaCl was added in some experiments to a concentration of 50 mM. For most experiments, the temperature was controlled at 288 K. All NMR spectra were acquired on a GN-Omega-500 spectrometer equipped with a Sun-3/160 computer.

For sequential assignments, COSY (Aue et al., 1976; Wider et al., 1984), double-quantum filtered COSY (Piantini et al., 1982; Shaka & Freeman, 1983), and NOESY (Jeener et al., 1979; Wider et al., 1984) experiments were performed. 2D homonuclear magnetization transfer (HOHAHA) spectra, used to identify many spin systems completely, were obtained by spin-locking with a MLEV-17 sequence (Bax & Davis, 1985) with a mixing time of 64 ms. All spectra were acquired in the phase-sensitive mode (States et al., 1982). The water resonance was suppressed by direct irradiation (1 s) at the water frequency during the relaxation delay between scans as well as during the mixing time in NOESY experiments.

The majority of the 2D NMR spectra were collected as 512 or 1024 t_1 experiments, each with 1K or 2K complex data

[1] Abbreviations: NMR, nuclear magnetic resonance; 2D NMR, two-dimensional NMR spectroscopy; COSY, 2D NMR correlated spectroscopy; HOHAHA, 2D NMR homonuclear Hartmann–Hahn spectroscopy; NOE, nuclear Overhauser effect; NOESY, 2D NMR nuclear Overhauser effect spectroscopy; rf, radio frequency; FID, free induction decay; CD, circular dichroism; H1 or PPP, parent peptide GVKGDKGNPGWPGAPY; peptide AAA, alanine tri-proline-substituted peptide PPP; peptide SPP, serine mono-proline(P9)-substituted peptide PPP; peptide PSP, serine mono-proline(P12)-substituted peptide PPP; peptide PPS, serine mono-proline(P15)-substituted peptide PPP; ttt, ttc, tct, ctt, tcc, ctc, cct, and ccc, conformation state specific assignments standing for cis(c)/trans(t) isomer states of P9, P12, and P15, respectively, as described in Table I.

points over a spectral width of 5 kHz in both dimensions with the carrier placed on the water resonance. A total of 64 or 96 scans were generally time averaged per t_1 experiment. The data were processed directly on the Sun-3/160 computer. Data sets were multiplied in both dimensions by a 0°- to 60°-shifted sine-bell or Lorentzian to Gaussian transformation function and generally zero-filled to 1K in the t_1 dimension prior to Fourier transformation.

For identification of long-lived backbone amide protons, COSY (Aue et al., 1976; Wider et al., 1984) experiments were performed at 2 °C and pH 3.5. Peptide IV-H1 was first exchanged and freeze-dried from ^1H_2O solution. ^2H_2O at ice bath temperature was then added to the dried peptide kept in the NMR tube on ice. The pH was adjusted to 3.5 by adding microliter quantities of NaOD or DCl. The magnitude COSY experiment was run for approximately 6 h, collecting 256 time incremented 1K data point FIDs, each with 32 transients. 2D NMR COSY data sets were processed directly on the Sun 3/160 GN-Omega-500 NMR spectrometer. Data sets were first multiplied in both dimensions by a 0°- to 10°-shifted sine-squared bell function and zero-filled to 512 in the t_1 dimension prior to Fourier transformation. Long-lived backbone amide protons were identified by comparing cross-peak positions to sequentially assigned COSY data sets accumulated under the same conditions, but in the presence of mostly ^1H_2O (90%/10% ^2H_2O) where all NH resonances were present. αN cross-peaks that remained during the ^2H_2O COSY accumulation are called "long-lived" in this paper—the larger the cross-peak, the longer the relative life time.

The temperature dependences of backbone NH chemical shifts were followed at pH 6 and 15 °C from 1D NMR and 2D NMR COSY-type spectra where cross-peaks were assigned or checked at various temperatures. Since conditions used for identification of long-lived amides were lower pH and lower temperature, peptide IV-H1 conformation could vary from conditions used for structural analysis, i.e., pH 6 and 15 °C. Additional supportive information on possible hydrogen-bonded amide protons was deemed necessary. NH chemical shift was plotted versus temperature and slopes were derived from these data. For a short linear peptide, a shallower slope generally indicates a relatively more hydrogen-bonded NH proton (Wüthrich, 1986).

Circular Dichroism. CD spectra were measured on a Jasco J-500 automatic recording spectropolarimeter coupled with a data processor. Curves were recorded digitally and fed through the data processor for signal averaging and baseline subtraction. Spectra were recorded at 15 °C in 20 mM phosphate buffer, pH 6.5, over 185–260 nm with a 2-mm path-length quartz cuvette. The peptide concentration (0.04 mg/mL) was determined by measuring the absorbance at 280 nm and by using a calculated extinction coefficient of 7040 M^{-1} cm^{-1}. The scan speed was 5.0 nm/min. Spectra were signal-averaged 16 times, and an equally signal-averaged solvent baseline was subtracted. The program VARSELEC, provided by Dr. W. Curtis Johnson, was used for secondary structure analysis.

Cell Adhesion Assay. Standard protocols for assessing peptide cell adhesion activity have been used (Humphries et al., 1987; McCarthy et al., 1990). Synthetic peptides were either covalently fixed to specialized plastic surfaces or dried directly on various types of surfaces (Chelberg et al., 1990; McCarthy et al., 1990), or a more sensitive approach that coupled peptides to larger carrier proteins like ovalbumin or bovine serum albumin (Humphries et al., 1987) was used. The latter procedure gave better results at lower peptide concentration. Considering our use of a series of peptides, relative comparison of activities was essentially the same using either method. Coupling of these synthetic peptides to larger carrier proteins apparently enhances peptide/cell adhesion activity. When iodinated peptides were used as tracers, coupling efficiencies of 4–5 peptides per carrier protein molecule were achieved.

Highly metastatic mouse melanoma cells were allowed to adhere for 1 h to surfaces coated with increasing levels of petpide or of peptide coupled to larger carrier proteins. For controls, coupled OA or BSA was used to show that the coupling reaction did not create activity per se. Moreover, controls were done where synthetic peptides from other sources not known to promote cell adhesion were employed. Control experiments routinely gave negative results.

Results

Five homologues amino acid sequences of peptide IV-H1 are identified in Figure 1 as peptides PPP, AAA, SPP, PSP, and PPS. Peptide PPP is synonymous with peptide IV-H1, which corresponds to residues 1263–1277 from the triple helix domain of human collagen type IV (Babel & Glanville, 1984; Brazel et al., 1987; Killen et al., 1988). The standard one-letter codes for amino acid residues have been used. These peptides vary only at the three proline positions found in the native-derived fragment, peptide PPP. In sequential order, PPP stands for P9, P12, and P15. Relative to peptide PPP, peptide AAA, therefore, has all alanine residues in place of proline residues at these positions, while the three mono-proline-substituted peptides have one serine residue substituted for one proline residue in each case. Serine was chosen as the substituting residue since it is not naturally present in this sequence; additional alanine residues would have interferred with multiple *cis/trans*-proline-induced A14 resonance assignments, and glycine was not chosen since five such residues were already present. These abbreviated names will be used throughout the text to refer to these peptides. The peptides are labeled from N-terminal residue 1 through residue 16.

Cis/Trans Proline Isomerization and Multiple Conformations. Although peptide IV-H1 is chemically pure (>95%) as judged by HPLC and amino acid sequencing, COSY-type spectra show more cross-peaks than expected for most spin systems. The most apparent case of multiple resonances is seen in the presence of five alanine αH–βH_3 cross-peaks as shown in Figure 2 and labeled with roman numerals. A 1D NMR spectral trace of the same A14 βH_3 region is shown at the top of Figure 2. Chemically, peptide IV-H1 contains only one alanine residue, i.e., A14, and although lysine and proline side-chain resonances overlap with A14 βH_3 resonances in the 1D trace, no such overlap is observed for A14 αH–βH_3 cross-peaks in the 2D contour plot (Figure 2). Moreover, possible ambiguities in this 2D NMR spectral region due to threonine βH–γH_3 cross-peaks cannot occur since these residues are absent from the peptide. The only long-chain hydrophobic residue, i.e., V2, has identifiable γH_3 resonances more upfield-shifted at about 0.9 ppm. Since there is no ambiguity in their belonging to multiple A14 resonances, these five cross-peaks must, therefore, be due to monomolecular conformational heterogeneity or intermolecular aggregation. The alanine resonance population distribution, however, is concentration independent in the range 1–50 mM; therefore, monomolecular conformational heterogeneity is the reason for these various alanine species. Since peptide AAA shows no such multiple ^1H NMR species (data not shown), proline imino acid residues, P9, P12, and P15, are most likely responsible for the appearance of these multiple states. Moreover, cis–trans proline isomerization is known to be slow on an NMR

time scale (Grathwohl & Wüthrich, 1976, 1981); therefore, such NMR long-lived states should be expected to be obseved in peptide IV-H1 if prolines are the reason for this resonance multiplicity.

Since A14 sequentially precedes a proline residue, i.e., P15, cis–trans proline isomerization could readily explain the presence of two of these A14 resonances, but not five as observed in Figure 2. However, if cis/trans isomer states for all three prolines were considered, a maximum of eight species of peptide IV-H1 could be observed. These species would also be expected to show varying populations depending on individual *cis/trans*-proline distributions and conformational influences. Table I lists the various combinations for the XP peptide bonds and gives the predicted fractional population computed as the simple product of individual *cis/trans*-proline fractional populations taken from short model peptides (Grathwohl & Wüthrich, 1976, 1981). On the basis of these predicted values, six of eight possible combined isomer states could be experimentally observed. The last two states, designated cct and ccc, have very low predicted populations and would be expected to fall into the NMR noise level with apparent concentrations in the 0.1–0.2 mM range. The six expected most populated isomer states are, therefore, ttt, ttc, tct, ctt, tcc, and ctc.

Further evidence that these multiple A14 resonances result from proline cis–trans isomerization comes from NMR spectral comparisons of mono-proline substituted analogues of peptide IV-H1 (Figure 3). These sequences have been identified in Figure 1. Double-quantum filtered COSY spectra show that peptide PPS (Figure 3a) shows three not five A14 αH–βH_3 cross-peaks, while peptide PSP (Figure 3b) shows only two cross-peaks, and peptide SPP (Figure 3c) shows the disappearance of only one minor A14 cross-peak. The two A14 αH–βH_3 cross-peaks observed in Figure 3b for peptide PSP can most likely be assigned to P15 cis and trans states. Since the trans state is the more populated one (Grathwohl & Wüthrich, 1976, 1981), these can be tentatively assigned to states ttt and ttc in accord with Table I and as indicated in Figure 3b. In peptide SPP, the one cross-peak that no longer appears relative to peptide PPP (Figure 2; cross-peak IV) is tentatively assigned to state ctt (Table I) since P9 is no longer present in this peptide. As with peptide PSP assignments, the largest cross-peak for peptide SPP can likewise be tentatively assigned to state ttt, and the cross-peak for state ttc can be assigned by shift comparison with peptide PSP. The remaining A14 cross-peaks probably arise from states ctc, tct, and tcc as indicated in Figure 3c. For peptides PSP and SPP, such tentative assignments were relatively straightforward since chemical shifts of remaining cross-peaks varied little from those for peptide PPP (Figure 2). Peptide PPS is more complicated since chemical shifts of remaining cross-peaks varied considerably. This is probably due to the fact that P15 had been substituted and this residue is of course sequentially proximal to A14 and necessary for its conformational integrity. Some assignment insight, however, can still be had. For example, these three cross-peaks can not be assigned to states ttc, tcc, or ctc, since populations of this cis peptide bond would not be observed. Once again the most populated state can be assigned to state ttt, and the other two states, therefore, most probably arise from states ctt and tct. These comparisons and tentative assignments also allow cross-peaks in Figure 2 to be assigned as follows: III with state ttt; IV with state ctt, and V with state ttc. Cross-peaks I and II, by a process of elimination, must be assignable to states tct, tcc, or ctc.

In any event, observation of five of these proline-induced multiple states in A14 resonances implies the presence of relatively long-range conformational interactions. Normally, short linear peptides, like peptide IV-H1, are highly flexible with such peptide fold effects being averaged out. Moreover, cis–trans proline isomer states are generally only observed with sequentially neighboring residues. In the present case, however, P9, P12, and P15 are sequentially proximal and probably tend to decrease overall backbone flexibility more than if only one proline residue were present. This in turn may promote some relatively stable backbone folding.

	1				5				10				15
peptide PPP:	G	V	K	G	D	K	G	N	P	G	W	P	G A P Y
peptide AAA:	G	V	K	G	D	K	G	N	A	G	W	A	G A A Y
peptide SPP:	G	V	K	G	D	K	G	N	S	G	W	P	G A P Y
peptide PSP:	G	V	K	G	D	K	G	N	P	G	W	S	G A P Y
peptide PPS:	G	V	K	G	D	K	G	N	P	G	W	P	G A S Y

FIGURE 1: Amino acid sequences for peptide PPP and its analogues. The amino acid sequence of peptide PPP is shown by one-letter codes. The sequence is derived from the protein sequence of human collagen type IV, triple-helix domain residues 1263–1277 as discussed in the text. Analogues to this sequence, where proline residues have been substituted by alanines or serines, are also given using one-letter codes.

Conformation Isomer-State-Specific Sequential Resonance Assignments. Normally a peptide of 16 residues would give about 100 proton NMR resonances. With the possibility of observing eight relatively long-lived conformation states in peptide IV-H1, however, this number increases to about 800, equivalent to a peptide the size of some 128 residues. Of the problems that confront us regarding such state-specific resonance assignments, resonance overlap and varying cis–trans isomer state populations are major concerns. For this complex situation of conformation-state-specific resonance assignments, an alternative approach to resonance assignments had to be used. This approach involves the use of proline-substituted homologues of peptide IV-H1 (see Figure 1) as mentioned above, relative resonance intensity comparisons with already assigned states, grouping of partially assigned conformational states into more complete state assignments via intrastate NOE associations, and state-specific αN fingerprint NOESY/COSY sequential resonance assignments. In some cases, varying solution conditions, i.e., pH and temperature, had to done. Only through a combination of these could more complete conformation-state-specific assignments be made. Moreover, due to population differences and experimental limitations, like signal-to-noise and resonance overlap, etc., not all states could be observed and/or completely assigned. To simplify the conformation-state-specific assignment process, states are designated for the combination of *cis/trans*-proline isomers from which they arise. A three-letter code is used where each position in the code stands for the cis/trans isomer of each proline position. These states are identified by lower case letters as listed in Table I for the eight possible states. In each triplet, "t" and "c" stand for *trans*- and *cis*-proline isomers, respectively, for P9, P12, and P15 in that order.

For conformation-state-specific sequential resonance assignments, the best place to start is at chemically, although perhaps not conformationally, unique residues. For peptide IV-H1, these are V2, D5, N8, W11, Y16, and, as already mentioned, A14. COSY spectra of the downfield aromatic resonance region for peptide PPP allows spin system groupings for W11 and Y16 ring proton resonances (data not shown). For W11 and Y16, two aromatic ring proton spin system subsets can be traced out. One set displays significantly stronger cross-peaks and conventional NMR intensities than

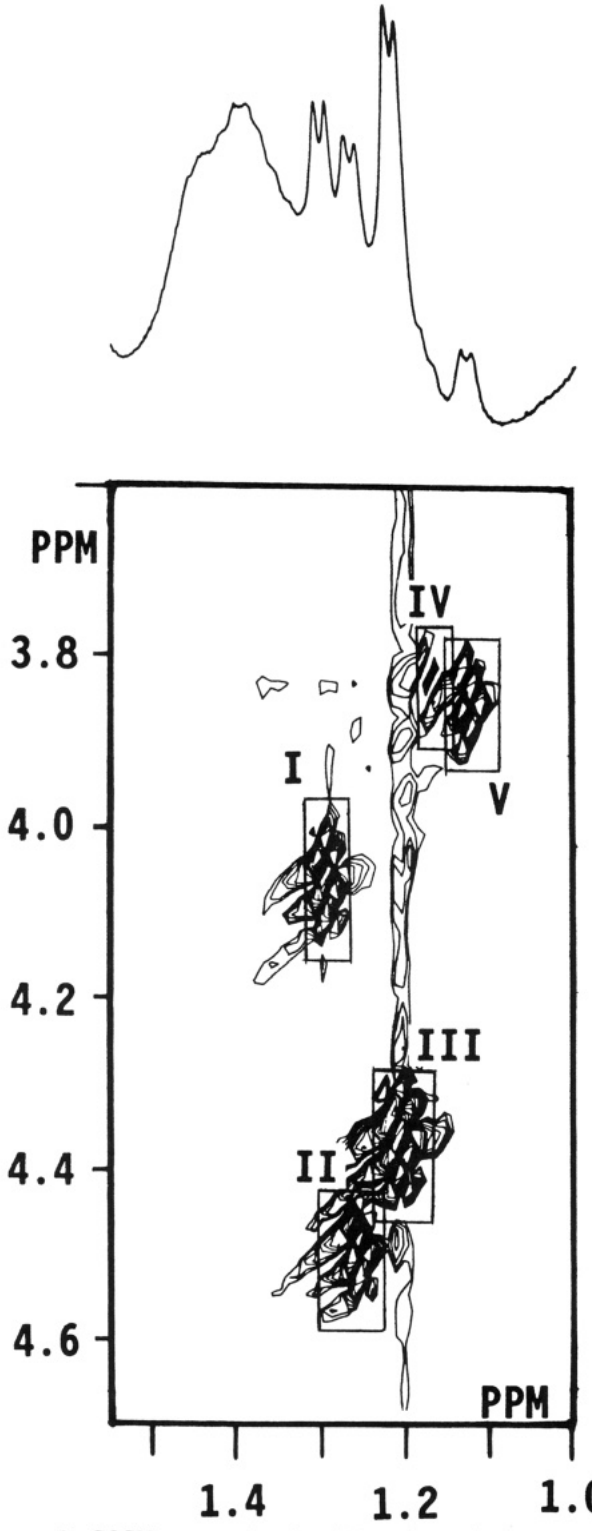


FIGURE 2: COSY contour plots for A14 conformation states. The upfield A14 αCH–βCH cross-peak region from 2D ^{1}H NMR phase-sensitive COSY contour plots is shown for peptides PPP. Assignments given are discussed in the text. The data set was collected as 512 hypercomplex FIDs containing 2K words and was processed on a Sun-3/160 computer. The data set was zero-filled to 1024 in the t_1 domain. The raw data were then multiplied by a 30°-shifted sine-squared function in t_1 and t_2 prior to Fourier transformation. Peptide concentration was 10 mM in 20 mM NaCl. The solution had a pH of 7 and a temperature of 15 °C. Labeling of resonances is as discussed in the text.

the other. Since both W11 and Y16 are sequentially flanked by proline residues, one might assume that the set with weaker intensity arises from the normally less populated cis isomer state (Grathwohl & Wüthrich, 1976, 1981). Since there exists

Table I: Conformation State Designations for Peptide PPP

1 5 10 15 GCKGDKGNPGWPGAPY I^a I^a I^a peptide bond (c/t)^b				
N8–P9 (10/90)	W11–P12 (35/65)	A14–P15 (10/90)	conf state desig	predicted fract pop.
trans	trans	trans	ttt	0.53
trans	trans	cis	ttc	0.06
trans	cis	trans	tct	0.28
cis	trans	trans	ctt	0.06
trans	cis	cis	tcc	0.03
cis	trans	trans	ctc	0.03
cis	cis	trans	cct	0.007
cis	cis	cis	ccc	0.004

^{a} Residues substituted. ^{b} Cis/trans proline isomer distributions have been taken from Grathwohl and Wüthrich (1981) for short model peptides.

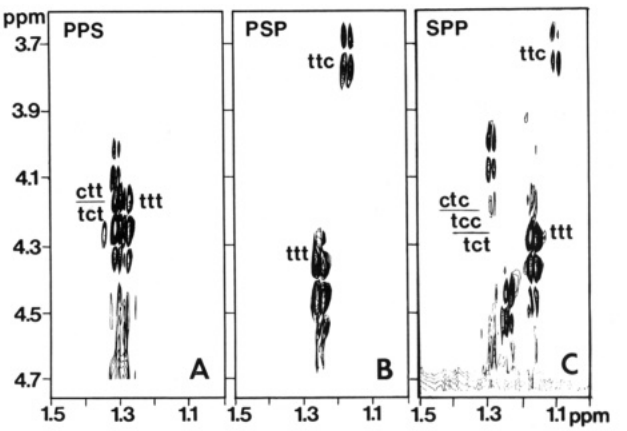


FIGURE 3: A14 region for peptides PPS, PSP, and SPP. The αH–βH$_3$ cross-peak region from double-quantum filtered COSY spectra of peptides PPS, PSP, and SPP is shown. State-specific assignments are labeled as discussed in the text. Data sets were collected as 512 hypercomplex FIDs each containing 2K words and were processed by zero-filling to 1024 in the t_1 domain and multiplying the raw data by a 30°-shifted sine-squared function in t_1 and t_2 prior to Fourier transformation. Peptide concentration varied from about 5 to 10 mM, at a pH of 7 and a temperature of 15 °C. NaCl was added to a concentration of 20 mM.

a high probability that Trp C4 and C2 and Tyr C2/C6 ring proton resonances at short mixing times will give NOEs to their respecitive α- and β-proton resonances (Wüthrich, 1986), it can be seen in Figure 4 that both major and minor W11 and Y16 ring proton resonances show this effect. The W11 C4 and C2 resonances of the major population, identified by roman numeral I in Figure 4, give cross-peaks to a resonance in the αH resonance region and to three resonances in the βH resonance region. From COSY spectral analysis (Figure 4b), the αH and two of the βH region resonances are spin-coupled to each other. These can be tentatively assigned to αH and βH resonances of the W11 spin system for the major population. By similar reasoning, it can be seen that the less intense W11 C2 and C4 ring proton resonances, roman numeral II, also show NOEs to spin-coupled αH–βH$_2$ cross-peaks for the minor population. The Y16 (2,6) ring proton resonances likewise give cross-peaks to Y16 αH and βH$_2$ resonances, as labeled in Figure 4. For Y16, only general spin system groupings could be made at this point.

Assignment to the W11–P12 trans or cis state, for example, can be made by observing the W11 αH resonance line in the NOESY contour plot of Figure 4c. The W11 αH resonance for the major population gives NOEs to one set of nondegenerate proline δH resonances. Only *trans*-proline will give

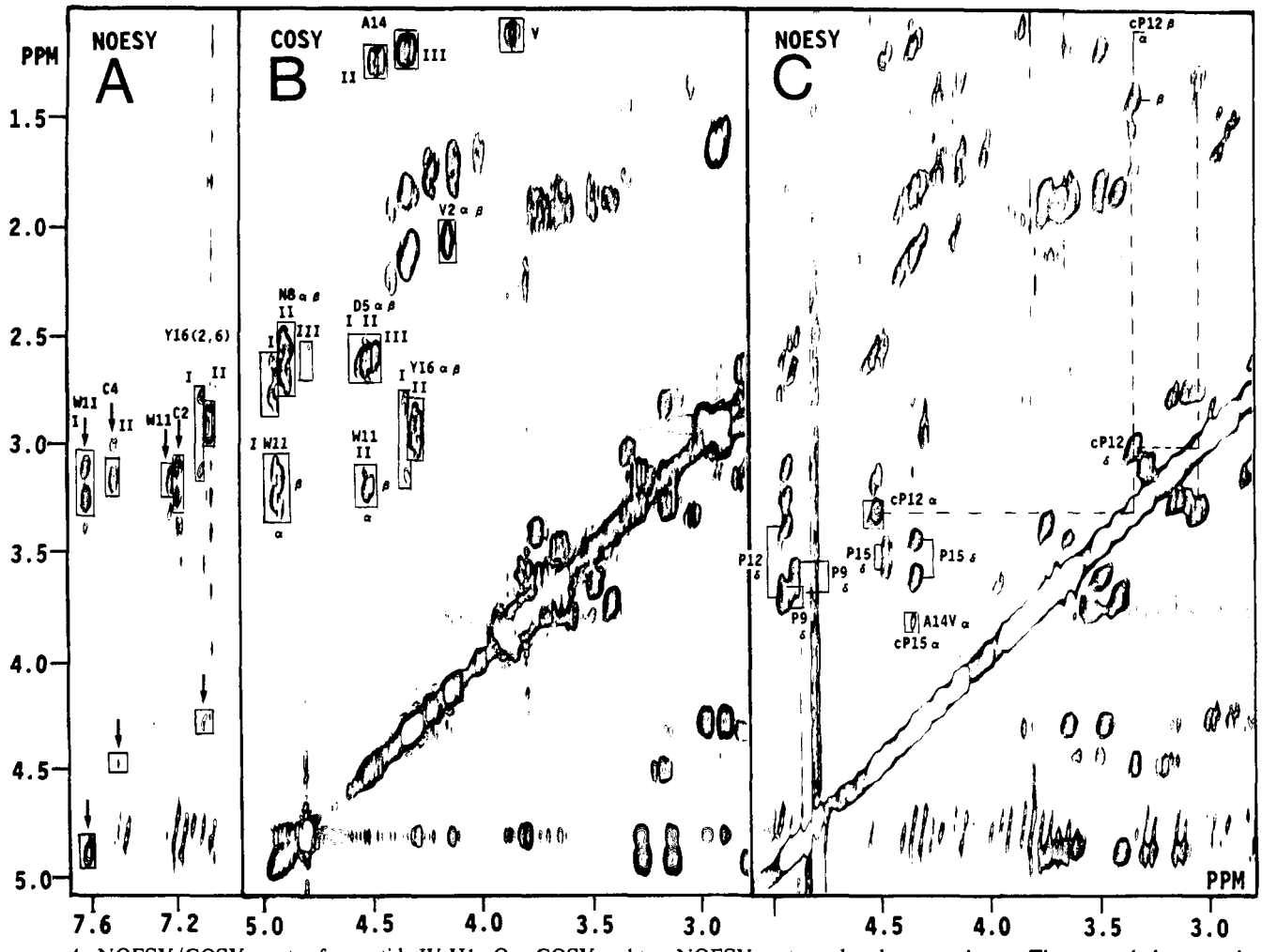


FIGURE 4: NOESY/COSY spectra for peptide IV-H1. One COSY and two NOESY contour plot plates are shown. The two end plates consist of D_2O NOESY contour plots for the aromatic/upfield resonance region (A) and for the αCH–proline upfield region (C). The middle plate (B) is taken from a D_2O magnitude COSY contour plot in order to correlate with the NOESY contour plot plates and to follow assignments discussed in the text. The data sets were collected in ^2H_2O (0.6 mL), with 40 mM peptide IV-H1 (PPP), pH 7, and no additional salt (low ionic strength). A total of 512 hypercomplex FIDs containing 1K words were collected and processed on a Sun-3/160 computer. The mixing time for the NOESY spectra was 0.1 s. The data set was zero-filled to 512 in t_1. The raw data were then multiplied by a 40°-shifted sine-squared function in t_1 and t_2 prior to Fourier transformation. Labeling of resonances is as discussed in the text.

such strong NOEs from its δH resonances to the preceeding N-terminal αH resonance where interproton distances are less than 4 Å (Wüthrich et al., 1984). In the cis state, the $d_{\alpha\delta(i,i+1)}$ distance is greater than 5 Å, and no NOE would be expected. On the other hand, a strong $d_{\alpha\alpha}$ NOE would be expected for the *cis*-proline peptide bond (Wüthrich et al., 1984). This is obseved in the W11 αH minor population. Only on this basis could definite conformation-state-specific cis–trans assignments be made for W11 spin system subsets.

Having identified *trans*-P12 δH and *cis*-P12 αH resonances, the remainder of these proline spin system assignments could be made by tracing connectivities from their δH or αH resonances through to other resonances in the same spin system subset. Proline assignments, in general, were independently made by comparing the proline resonance regions for peptides PPP, PPS, PSP, and SPP. For example, Figure 5 compares NOESY contour plot extracts of the proline γ/δ resonance cross-peak region for peptides SPP and PPS. In the case of proline resonances, chemical shifts varied no more than about 0.05 ppm among mono-proline-substituted peptides, and assignments were generally straightforward as indicated in Figure 5. The previously assigned *trans*- and *cis*-P12 resonance spin systems were thereby confirmed, and those for P9 and P15 were also made.

Although *cis*-proline resonances are generally found more upfield than their trans isomer counterparts (Wüthrich, 1986), the *cis*-P12 spin system is even more highly upfield-shifted than expected. In Figure 4c, the *cis*-P12 spin system has been partially traced out. The nondegenerate βH resonances, for example, are unusually chemically shifted from each other by about 1 ppm; the more upfield βH resonance is most probably highly ring current shifted due to proximity with the W11 aromatic ring system. An average orientation above or below the plane of the W11 ring could account for the magnitude of this shift.

Jumping next to A14 αH resonances that have been identified in Figure 2, it is apparent that the two most downfield A14 αH resonances (labeled II and III in Figure 2) reflect the xxt state by virtue of relatively strong cross-peaks to P15 δH resonances (Figure 4c). The A14, βH_3 resonance V lacks cross-peaks to proline δH resonances but does display one to an αH resonance, probably that of *cis*-P15 as indicated in Figure 4c; this, therefore, can be assigned to the A14–*cis*-P15 state, i.e., state xxc. These observations are consistent with assignments made from analysis of data in Figure 3. The minor A14 αH resonance IV is not observed in Figure 4, probably due to its minimal population. It should also be noted that A14 cross-peak chemical shifts are slightly different from

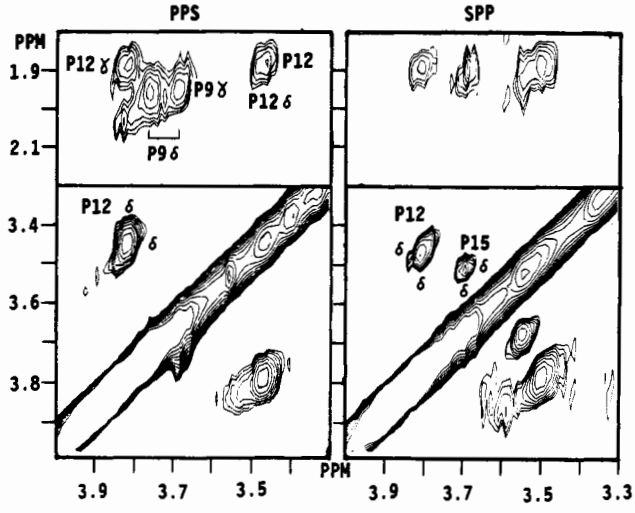


FIGURE 5: Comparison of proline resonance region for peptides PPS and SPP. Two COSY contour plot plates are shown for the δ–δ and γ–δ proline resonance regions for peptides PPS and SPP. Data have been collected and processed as described in the legend to Figure 3. Labeling is as discussed in the text.

those shown in Figures 2 and 3 due to the absence of added NaCl in the Figure 4 data sets. As mentioned earlier, varying solution conditions was necessary to resolve and, therefore, to assign cross-peaks. In this case, A14 in some states is apparently proximal to a titratable carboxylate group, thus accounting for its sensitivity to varying ionic strength.

Only two AMX type residues, D5 and N8, remain unassigned, whereas the COSY spectrum (Figure 4b) indicates that about six AMX spin systems are still unassigned. Multiple D5 spin systems, labeled I, II, and III, were recognized by their pH-dependent chemical shift changes, which reflect a carboxylate side-chain titration with a pK_a value of about 4 (data not shown). This is later confirmed by assignment of the 2D NMR αN NOESY fingerprint region where conformation-state-specific assignments for D5 will be made. This accounted for three of the remaining six AMX spin systems. The two most downfield-shifted resonances, labeled N8 I and II, of the three remaining AMX spin systems give strong cross peaks to P9 δH resonances. Since only N8 sequentially precedes a proline residue, i.e., P9, assignment of these to N8 in the txx state could be made. Another N8 conformation state can be identified with state cxx, as indicated in Figure 4b by roman numeral III. This minor state N8 αH resonance shows a cross-peak to a *cis*-P9 αH resonance (observable in other NOESY data sets).

At present, partial conformation-state-specific assignments have been made for the three proline residues, i.e., P9, P12, and P15, as well as for N8, W11, and A14. General spin system assignments have also been made for D5 and Y16 without regard to specific cis–trans proline isomer states. The link between a partial conformation-state assignment, e.g., xcx, and a complete conformation state assignment, e.g., tct, is not as straightforward. From Figure 3, some tentative conformation-state assignments have been made for A14 resonances. N8 resonances can be assigned to specific conformation states as well. Of the three N8 αH–βH resonance sets, the least populated one (already partially assigned to state cxx) can now be assigned to state ctx since state ccx has the lowest probability of being observed (Table I). Of the two N8–*trans*-P9 spin system subsets, the more populated one, labeled N8 II, can be tentatively assigned to the ttx state, and the other to the tcx state, labeled N8 I in Figure 4b. The N8 ttx state assignment can be confirmed since one set of txx N8 βH resonances shows NOEs to the xtx W11 βH resonances (to be discussed in terms of conformational constraints in the next section).

From analysis of HOHAHA and COSY (Figure 4) spectra, αN cross-peaks can be at least partially identified or grouped in the double-quantum filtered COSY contour plot (Figure 6). Here, V2, K3/K6, D5, N8, W11, A14, and Y16 cross-peaks have been labeled or grouped if multiple resonances were apparent. The remainder of these cross-peaks must belong to glycine, labeled with a G in Figure 6, i.e., for G1, G4, G7, G10, and G13 resonances. It is interesting to note that while most multiple resonances for a given spin system are generally grouped close together, cross-peaks for W11 and especially for A14 are more separated, suggesting a greater conformational influence at these residue positions. Notably, both W11 and A14 are sequentially positioned between the three backbone-constraining proline residues.

The NOESY αN fingerprint contour plot for peptide IV-H1 (Figure 7) allows sequential assignments of multiple states to be made. Since tracing through all species in one plot would be confusing, Figure 7 traces through state ttt only. The V2 αH–NH cross-peak is apparently unique and allows one to trace through the first four residual αH–NH patterns. Regardless of the specific isomer state, an unambiguous αN tracing can be made from G1 αCH to G4 αCH. Multiple resonances for other N-terminal residues G1, K3, and G4 are also generally not apparent. K3, however, in its upfield $\alpha\beta$ domain, does demonstrate two spin system sets beginning at its αH resonance; shift differences in K3 backbone and side-chain resonances are only about 0.005 ppm. Obvious conformational multiplicity among these states appears to begin at residue D5 through to the C-terminus.

The most intense D5 resonance, as seen in the Figure 6 COSY spectrum and here in Figure 7, was assigned to state ttt in accord with expected isomer state populations given in Table I. A clear tracing from D5 to G7 could then be followed (Figure 7). N8 αH resonance positions for states ttx, tcx, and ctt could be identified from data given in Figure 4. Since peptide PPS shows similar multiple N8 spin system patterns, P15 plays no apparent role in N8 multiple conformation states. For this reason and from predicted isomer state populations (Table I), states ttx and tcx have been more completely assigned as states ttt and tct, respectively. Considerable overlap of glycine resonances prohibited a clear tracing from the G7 αH resonance in state ttt to the respective N8 NH resonance cross-peak. W11 NH resonances were easily followed from their respective βH resonances. In this NOESY experiment, three, not two, sets of W11 βH resonances are apparent; these presumably arise from states ttt, tct, and ctt or ttc. The most intense W11 spin system subset was assigned to state ttt; this subset was, in fact, already independently assigned to state xtx (Figure 4). The ttt state W11 NH can be traced to G10 αH, and from G10 NH to P9 αH. Since a *cis*-proline αH is not expected to show an NOE to its C-terminal neighbor NH resonance, this must be a *trans*-P9 αH resonance, thus supporting this state ttt assignment. Lastly, since an A14 βH_3 resonance has already been assigned to state ttt, its A14 NH resonance can be traced to G13 αH, and G13 NH to P12 αH. By the same reasoning used above for assignment to *trans*-P9, this P12 αH resonance must also be from the trans isomer state. Notice that G13 resonances are significantly shifted from other glycine resonances, probably indicating conformational and/or W11 ring current shift effects. Multiple Y16 cross-peaks overlap somewhat in the αH domain (Figure 4). From intensity comparisons and distributions shown in Table

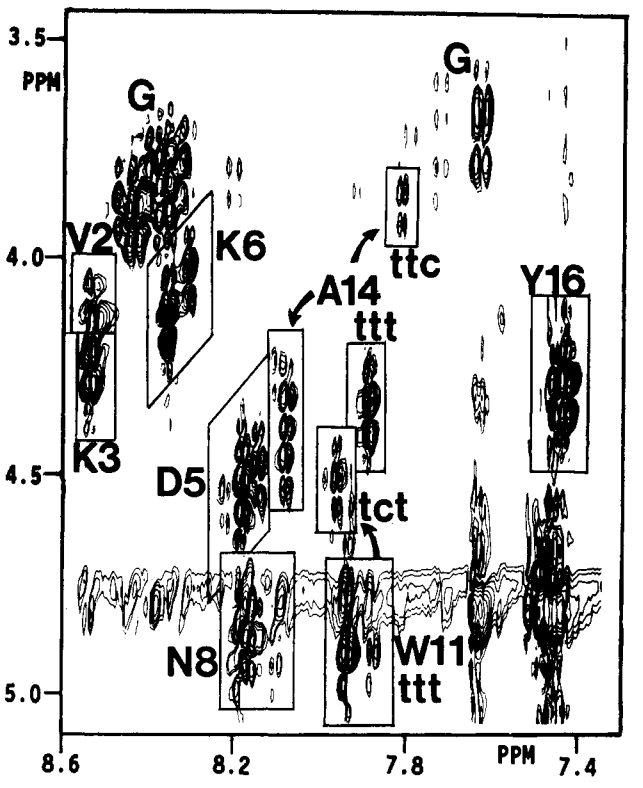


FIGURE 6: Double-quantum filtered COSY of the αN fingerprint region. The αH–NH fingerprint region from a double-quantum filtered COSY spectrum of peptide IV-H1 is shown. Cross-peak labeling and state-specific assignments are discussed in the text. Data sets were collected as 512 hypercomplex FIDs each containing 2K words and were processed by zero-filling to 1024 in the t_1 domain and multiplying the raw data by a 30°-shifted sine-squared function in t_1 and t_2 prior to Fourier transformation. Peptide concentration was 40 mM, at a pH of 6 and a temperature of 15 °C.

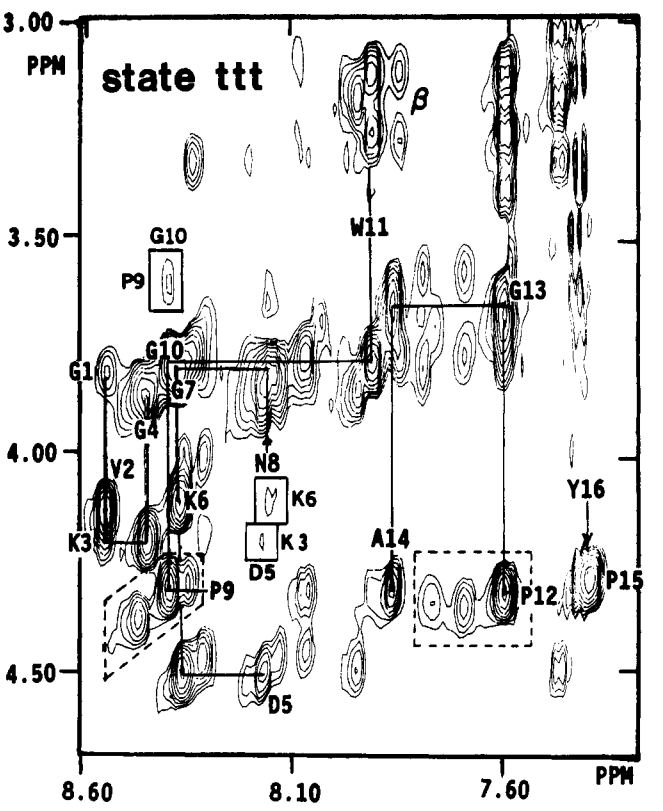


FIGURE 7: NOESY αCH–NH fingerprint region of peptide IV-H1. The αCH–NH resonance region from a NOESY contour plot is shown for sequential resonance assignments. Data were collected in ^1H_2O (0.6 mL) with 40 mM peptide IV-H1 at pH 6 and 15 °C. A total of 512 hypercomplex FIDs containing 1K words were collected and processed as discussed in the legend to Figure 4. The mixing time was 0.4 s. Labeling of resonances is described in the text.

I, the more upfield Y16 NH resonance can be assigned to state ttt (most intense), while the more downfield Y16 NH can be assigned to state ttc due to its sequential proximity to P15 and to the fact that Y16 multiple resonances are not observed in peptide PPS.

As is evident from the above paragraph, the general approach to conformation-state-specific sequential resonance assignments is to trace αN(i,i+1) spin-system subset sequential fragments through to recognizable *trans*-proline, i.e., P9 and P12, αH resonances that, from state to state, are relatively close in chemical shift as indicated in Figure 7 (boxed in with a broken line) For these *trans*-proline assignments, it has been assumed that *cis*-proline αH resonances are considerably more chemically shifted, presummably upfield, and, as mentioned above, that these groups of P9 and P12 αH to C-terminal NH resonance cross-peaks arise from the trans state. Relative overall isomer state populations are then compared with predicted proline isomer distributions given in Table I for possible assignment candidates. For instance, of the presummed P9 αH cross-peaks, none can be due to state cxx; therefore, according to Table I, the only remaining candidates with sufficient populations to be observed are states ttt, ttc, tct, and tcc. State ttt is expected to have the largest population. Of the remaining three possibilities, state tct is predicted to be second most populous and can therefore be assigned to the more downfield of these P9 αH cross-peaks. For greater confidence in these state-specific assignments, mono-proline-substituted analogues are used. Figure 8 gives the same αN NOESY fingerprint contour plot region for peptides SPP and PPS. For peptide SPP, these P9 cross-peaks are absent as are those associated with upfield P9 resonances; multiple S9 cross-peaks, however, are observed in their place. For peptide PPS, only two such P9 αH cross-peaks are found. These must be due to states ttx and tcx. The three P9 αH cross-peaks observed for peptide IV-H1 (PPP) must be due to states ttt, tct, and ttc for several reasons; (1) *cis*-P9 shows no such NOE; (2) P15 must have some infuence at position P9 since one such cross-peak vanished in peptide PPS; and (3) population arguments are such that state ttc should have a larger population than state tcc. Furthermore, for sequential assignments that involve residues N8, W11, and A14, these arguments can be checked by previous partial or full conformation state assignments. The remainder of these conformation-state-specific assignments has been made by using similar arguments. Resonance assignments for all identifiable isomer-specific states are listed in Table II.

Preferred Conformation Populations of State ttt. As indicated above, peptide IV-H1 must exist in solution in some reasonably defined conformational populations. Initial indication that state ttt exists in a preferred conformation(s) comes from the observation that chemical shifts are significantly different from those observed in random coil tetrapeptides (Bundi & Wüthrich, 1979). In fact, when these random coil shifts are compared with those values given in Table II, shift differences in backbone NH resonances for residues N8, G13, A14, and Y16 range from 0.3 to 0.8 ppm; smaller yet substantial NH shift differences in the 0.15–0.3 ppm range are also noted for V2, K3, D5, and W11 NH resonances. Although in short linear peptides such chemical shift differences from random coil positions for αH and βH resonances are normally much less than those for NH resonances, peptide IV-H1 also has many residues showing substantial differences

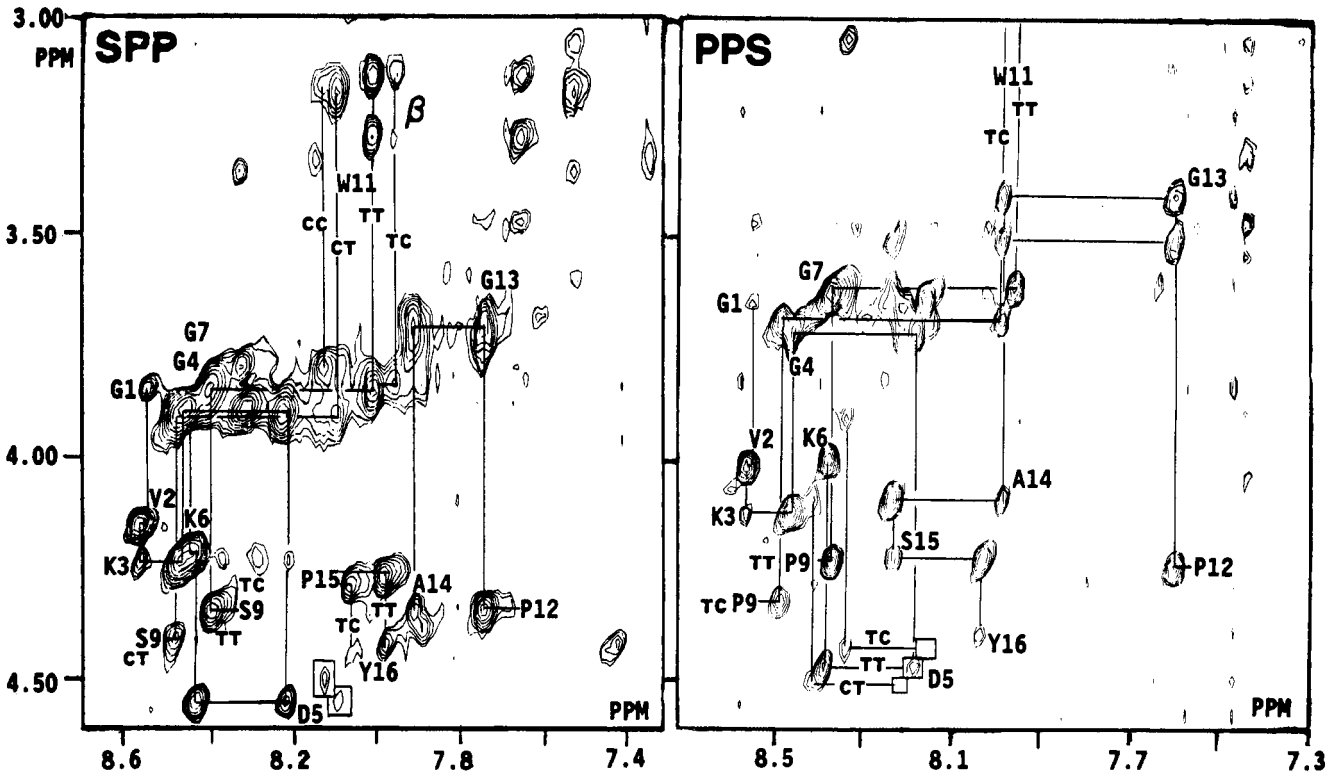


FIGURE 8: NOESY αN fingerprint region of peptides SPP and PPS. The αH–NH NOESY fingerprint region for peptides SPP and PPS are shown for comparison with data in Figure 7. NOESY data have been collected and processed as discussed in the legend to Figure 7, and labeling is as discussed in the text.

for αH and βH resonances in the 0.2–0.3 ppm range: D5, K6, N8, W11, P12, G13, A14, P15, and Y16. Moreover, the signs of these shift differences are both positive and negative, indicating that a systematic error in these comparisons is unlikely. Given the unpredictable nature of chemical shifts, however, such comparisons can perhaps only be used to support the idea that a purely random structure for peptide IV-H1 does not exist. The largest shift differences are observed for residues in the C-terminal domain, which contains the three backbone-constraining proline imino acid residues. This suggests that a more preferred conformation exists within this sequence domain.

A consistent network of interresidue NOEs, some of which are identified in Figures 9 through 12 with others being summarized in Figure 13, allows deduction of preferred peptide backbone folding patterns that limit the number of possible conformations normally available to a peptide the size of IV-H1. In combination with NMR data on long-lived backbone NHs, the temperature dependence of NH chemical shifts, and structurally informative $^3J_{\alpha N}$ coupling constants (Figure 13), these NOE data give evidence for a peptide IV-H1 state ttt solution conformation consistent with the presence of significant populations of multiple β-turns centered at K3–G4, K6–G7, P9–G10, and P12–G13, and a γ-turn centered at P15. Cross-peaks in Figures 9 through 12 have been labeled for sequential assignments identified in the previous section. For proline isomer states tct, ttc, and ctt, three-letter code labels have been used, while for cross-peaks belonging to state ttt, the ttt three-letter code label has been omitted.

Figure 9 shows the NH–NH resonance region from an H_2O NOESY contour plot. Four NH to NH NOEs are apparent in state ttt: two medium NOEs between G10–W11 and G13–A14 NHs, and two weak NOEs between G4–D5 and G7–N8 NHs. Similar NH–NH NOEs are apparent for some of these same residue pairs in other states as labeled in Figure 9. While the observation of $d_{NN(i,i+1)}$ NOEs suggests the presence of some folded structure, it is uninformative as to how much folded structure there is, or what that (those) possible structure(s) is (are). NOESY contour plot extracts shown in Figures 7, 10, 11, and 12 identify additional interresidue NOEs that further limit possible structures. Figure 10 is taken from a H_2O NOESY data set in order to show NH and aromatic residue ring proton through space connectivities to αH and more upfield proton resonances. Figure 11 is a contour plot of the same resonance region, but from a D_2O NOESY data set to remove any ambiguity arising from aromatic and amide proton resonance overlap. In this case, only aromatic ring proton resonance connectivities to upfield proton resonances are observed, as labeled in the figure. Figure 12 indicates βH and proline δH NOE connectivities. Figure 13 summarizes these NOE data, gives $^3J_{\alpha N}$ coupling constants below 6 Hz and above 8 Hz, and identifies possible H-bonded backbone amide protons from long-lived backbone NH resonances and temperature dependencies of NH resonances. $^3J_{\alpha N}$ coupling constants were generally structurally uninformative expect in a few cases; most of these fell in the 6.5–8 Hz range, which by itself is normally structurally uninformative for short linear peptides like peptide IV-H1 (Kessler & Bermel, 1986).

The following subheadings of this section are organized for specific amino acid sequences in order to address NMR parameters that allow deduction of structural elements in that particular sequence. Deduced structures discussed in these subheadings refer to state ttt.

V2, K3, G4, and D5. Starting from the N-terminus, the first apparent turn is evidenced by a d_{NN} NOE between G4 and D5 (Figure 9) and a $d_{\alpha N(i,i+2)}$ NOE between K3 and D5 (Figure 7). NOEs have also been noted between the G4 NH and the V2 γH_3 and K3 βH_2 and γH_2 resonances (Figure 10). A $^3J_{\alpha N}$ coupling constant of 5–6 Hz for K3 is also consistent with a β-turn within this sequence (Wüthrich, 1986). Moreover, a medium-sized V2–K3 $d_{\alpha N}$ NOE (even at short mixing times, i.e., 0.1 s) and a V2 $^3J_{\alpha N}$ coupling constant of

Table II: ^{1}H NMR Sequence Assignments (ppm)[a]

residue (state)	NH	αCH	βCH	others
G1		3.82		
V2	8.60	4.12	2.06	γ = 0.89
K3 (txx)	8.61	4.24	1.68, 1.74	γ = 1.34, 1.37; δ = 1.63; ε = 2.93
(cxx)	8.61	4.23	1.69, 1.73	γ = 1.35; δ = 1.64; ε = 2.94
G4 (xxx)	8.44	(3.87)[b]		
D5 (ttt)	8.27	4.51	2.60	
(ctt)	8.30	4.53	2.60	
(tct)	8.22	4.47	2.59	
K6 (ttt)	8.42	4.12	1.63, 1.73	γ = 1.26; δ = 1.54; ε = 2.88
(ctt)	8.41	4.04	1.60, 1.67	γ = 1.21; δ = (1.54); ε = (2.88)
(tct)	8.31	4.03	1.57, 1.66	γ = 1.26, 1.18; δ = 1.49; ε = 2.85
G7 (ttt)	8.37	3.85		
(tct)	8.32	3.77		
(ctt)	8.36	3.8		
N8 (ttt)	8.18	4.88	2.51, 2.64	
(tct)	8.19	4.93	2.62, 2.75	
(ctt)	8.16	4.77	2.50, 2.65	
P9 (ttt)		4.32	1.88, 2.11	γ = 1.87; δ = 3.58, 3.66
(tct)		4.39	1.86, 2.22	γ = 1.94; δ = 3.72, 3.64
(ttc)		4.31	(2.11, 1.8)[b]	γ = (1.8);[b] δ = (?)
(ctt)		4.58	1.92, 2.14	γ = (?); δ = 3.48, 3.61
G10 (ttt)	8.39	3.79		
(tct)	8.47	3.86		
(ttc)	8.35	3.78		
W11 (ttt)	7.93	4.91	3.13, 3.27	7.62 (4H), 7.21 (2,6H); 7.46 (7H), 7.10 (5H)
(tct)	7.96	4.49	3.17	7.50 (4H), 7.21 (2,6H); 7.48 (7H), 7.10 (5H)
(ctt) (or ttc?)	7.87	4.90	3.13, 3.28	
P12 (ttt)		4.32	2.12, 1.80	γ = 1.83; δ = 3.72, 3.38
(tct)		3.31	0.65, 1.53	γ = 1.38, 1.48; δ = 3.31, 3.01
(ttc)		4.36	(?)	(?)
(ctt)		4.35	(?)	(?)
G13 (ttt)	7.62	3.65, 3.75		
(ttc)	7.71	3.57, 3.77		
(tct)	8.34	3.78		
(ctt)	7.78	3.73		
A14 (ttt)	7.94	4.30	1.14	
(ttc)	7.87	3.81	1.05	
(tct)	8.14	4.42	1.18	
(tcc/cct?)	8.14	4.17	1.14	
(ctt)	8.08	3.81	1.11	
P15 (ttt)		4.31	2.05, 1.80	γ = 1.85; δ = 3.43, 3.6
(ttc)		4.32	2.1, 1.82	γ = (?); δ = (?)
(tct)		4.31	(2.05, 1.80)[b]	γ = 1.85; δ = 3.46, 3.57
Y16 (ttt)	7.47	4.82	2.86, 2.97	7.05 (2H,6H); 6.77 (3H,5H)
(ttc)	7.50	4.35	2.78, 3.12	7.08 (2,6H); 6.74 (3,5H)

[a] Solution conditions are pH 6.0, 288 K, and 10% ^2H_2O/90% ^1H_2O. The HDO resonance is 4.86 ppm downfield from TSP. [b] The chemical shift varies slightly from state to state; resonance overlap prohibits specific assignments.

8–9 Hz suggest an extended chain population for N-terminal residues V2–K3 going into the turn conformation. In this structure, the D5 carboxylate group would be proximal to the N-terminal NH_3^+ and/or the K3 ε-NH_3 group for possible salt-bridge formation and turn stabilization. Furthermore, the V2 NH not only is long-lived at pH 3.5 and 2 °C but also shows one of smallest NH resonance temperature coefficients (0.0033 ppm/deg) at pH 6 and 15 °C (Figure 13), suggesting probable hydrogen-bond formation. This proposed hydrogen bond could be formed among several possibilities, including the D5 backbone carbonyl oxygen or the D5 side-chain carboxylate oxygen. Additionally, the absence of a long-lived backbone amide proton and a relatively large NH temperature coefficient (0.011 ppm/deg) for D5 suggests that a tight β-turn for this sequence does not form, although this is difficult to

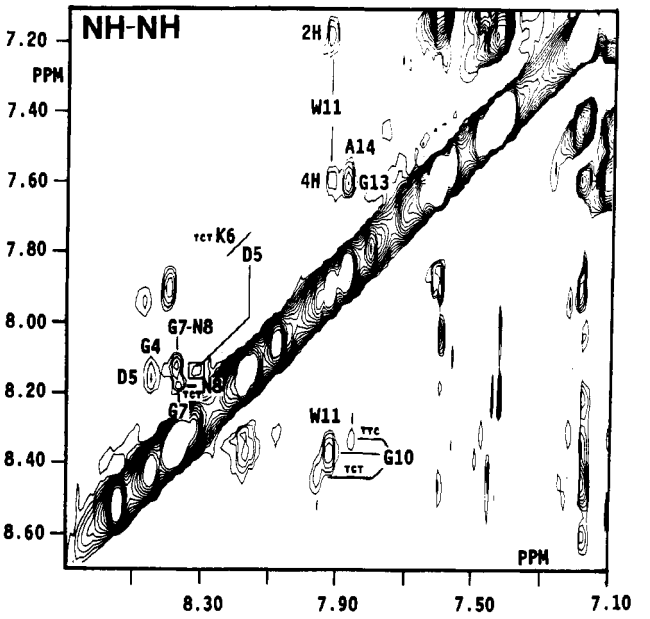


FIGURE 9: NOESY spectrum NH resonance region. A 500-MHz ^{1}H NMR NOESY contour plot of the NH resonance region for peptide IV-H1. The data set was collected in 90% ^1H_2O/10% ^2H_2O (0.6 mL) with 40 mM peptide and no additional salt (low ionic strength) at pH 6 and 15 °C. A total of 512 hypercomplex FIDs containing 1K words were collected and processed on a Sun-3/160 computer. The mixing time was 0.4 s. The data set was zero-filled to 512 in t_1. The raw data were then multiplied by a 40° shifted sine-squared function in t_1 and t_2 prior to Fourier transformation. Labeling of resonances is as discussed in the text.

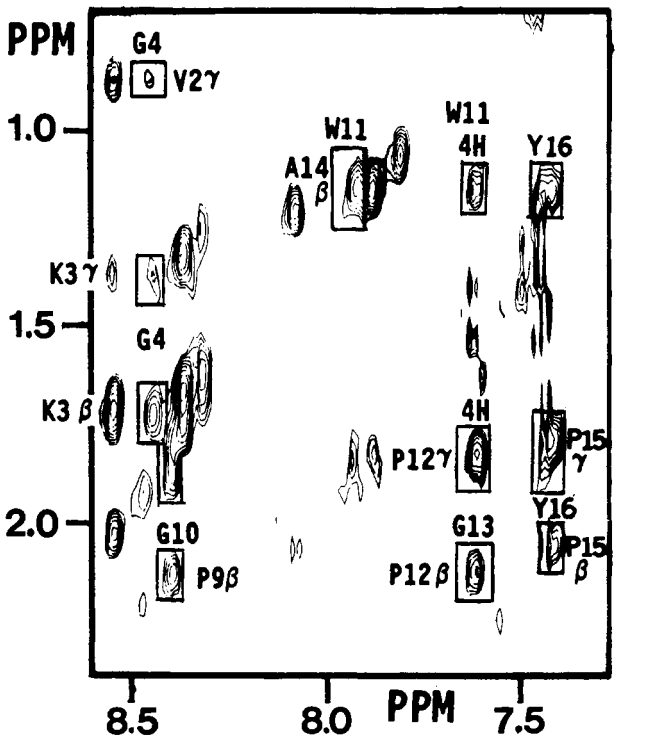


FIGURE 10: NOESY spectrum of NH/upfield region. The NH/upfield region of the NOESY data set discussed in the Figure 9 legend is shown. Boxed in cross-peaks are labeled as discussed in the text.

interprete for a normally highly flexible terminal segment; on the other hand, the presence of a network of NOEs and structurally informative $^3J_{\alpha N}$ coupling constants suggests that flexibility may not be the answer here and that a tight β-turn may in fact not form.

Although these NMR/NOE data give evidence for a significant β-turn population within the V2–D5 sequence, the

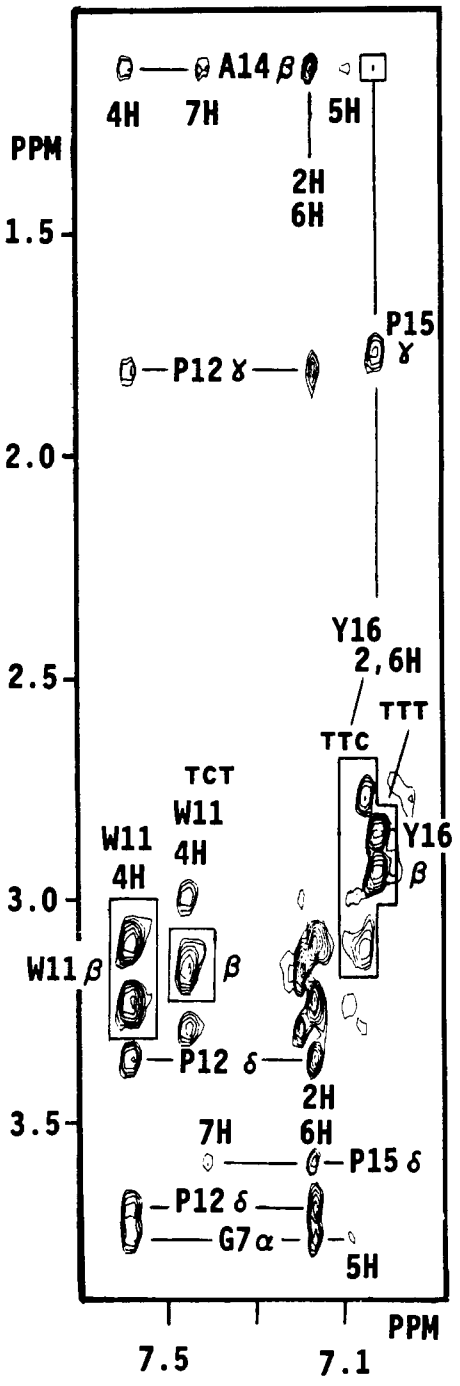


FIGURE 11: D_2O NOESY spectrum of aromatic/upfield region. This NOESY data set was collected as described in the Figure 9 legend. The peptide was dissolved in D_2O in order to exchange all labile amide protons. Connectivities are therefore shown between aromatic ring proton resonances and upfield aliphatic proton resonances.

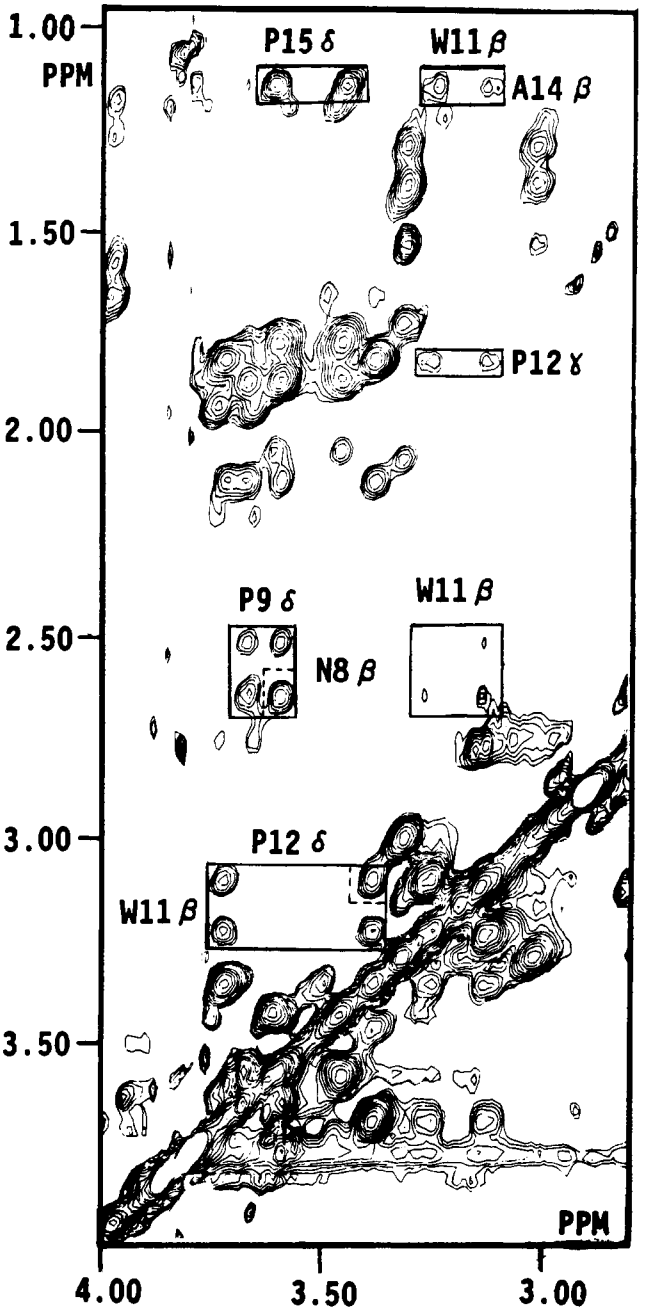


FIGURE 12: D_2O NOESY spectrum of upfield region. A NOESY contour plot is shown for the upfield region of peptide IV-H1 in D_2O. Proline δH and N8, W11, and A14 βH connectivities are boxed in as discussed in the text. Experimental details are given in the legend to Figure 9.

specific type of turn, e.g., I or II, has yet to be addressed. As the major classes of β-turns, type I and type II turns differ largely in the relative orientation of the turn position 2–3 peptide bond (Rose et al., 1985). In type II, the position 3 NH is orientated nearly in the same direction as the position 2 αH. In type I, it points in the other direction. Ambiguity exists in the V2–D5 sequence as to which turn type may be present. A type II conformational population is evidenced by a large $d_{\alpha N}$ NOE between K3 and G4 (Figure 7), observed even at short mixing times, i.e., 0.1 s, indicating a preferred short $d_{\alpha N(i,i+1)}$ distance between K3 and G4 like that required for a type II turn. With short linear peptides, however, this is difficult to resolve on the basis of a $d_{\alpha N}$ NOE alone. In this case, however, the $d_{\alpha N}$ NOE magnitude may be meaningful since other $d_{\alpha N}$ NOEs in the N-terminal domain are much smaller at a mixing time of 0.1 s (Figure 13). Glycine coupling constants could be helpful in establishing the turn type, but most glycine αN cross-peaks overlap and an unambiguous analysis of G4 coupling constants is impossible. An additional network of NOEs among the G4 NH and K3 and V2 side-chain resonances (Figure 10), on the other hand, supports a type I turn conformation. Since the N-terminus seems to adopt a more extended chain conformation as noted above, the V2 and K3 residues would be positioned on the same side of the β-turn, and the only way to explain an NOE between the G4 NH and the V2 γH_3 resonances (not to mention those with the K3 side-chain resonances) is with a turn type I conformational population. This apparent ambiguity may be explained by a recent structure analysis of cyclic pentapeptides as models for reverse turns (Stradley et al., 1990), which

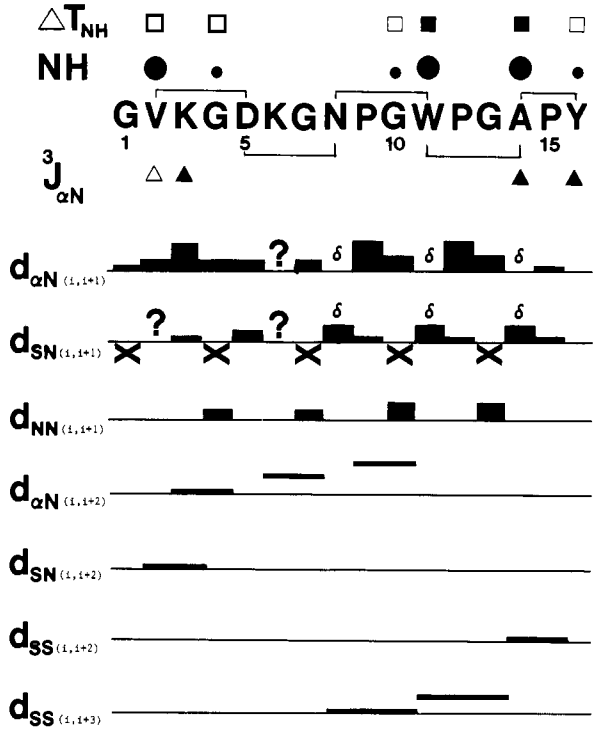


FIGURE 13: Summary of NMR data for peptide IV-H1. The peptide sequence of IV-H1 is shown together with $^3J_{\alpha N}$ coupling constants below 6 Hz (▲) and above 8 Hz (△). Long-lived amide protons are indicated by a filled-in circle for backbone NH positions where NH resonances were present long enough to give a magnitude COSY cross-peak (about 6 h at 2 °C) as discussed under Materials and Methods; a larger circle indicates a relatively longer lived NH by virtue of a larger remaining COSY cross-peak. The temperature dependence of NH resonance chemical shifts are indicated for values of 0.003–0.0035 (□) and 0.0035–0.0042 (■) ppm/deg. The most temperature-dependent NH belonged to D5 with 0.011 ppm/deg. Interresidue NOEs are identified according to Wüthrich (1986). β- and γ-turn positions are indicated above or below the peptide sequence by lines connecting the first and last residue in each turn. "S" in a subscript stands for "side chain". A "?" indicates overlap prevented observation of a possible cross-peak, and an "x" below the line indicates that this NOE definition is inapplicable.

showed that type I and type II turns are in dynamic equilibrium, fast on an NMR time scale. In this respect, both turn types could coexist in equilibrium, exhibiting structural characteristics of both turn populations.

D5, K6, G7, and N8. A reverse turn in this sequence is evidenced by a G7–N8 d_{NN} NOE (Figure 1) and a $d_{\alpha N(i,i+2)}$ NOE between K6 and N8 (Figure 7). For various reasons, insight into a specific turn type can not be had. K6 and G7 NH resonances overlap, such that observation of any possible NOEs between the G7 NH and D5 or K6 resonances cannot be unambiguously made. Moreover, $^3J_{\alpha H}$ coupling constants for G7 cannot be resolved due to overlap with most other glycine resonances, and the K6 $^3J_{\alpha N}$ coupling constant is about 7.5 Hz.

N8, P9, G10, and W11. G10–W11 d_{NN} (Figure 1) and P9–W11 $d_{\alpha N(i,i+2)}$ NOEs (Figure 7), in combination with an observed long-lived backbone W11 NH and a significantly reduced NH temperature coefficient (0.004 ppm/deg) (Figure 13), indicate a significant β-turn population within this sequence. Unlike the first two N-terminal turns identified above, this β-turn shows additional "long-range" NOEs between the βH_2 resonances of N8 and W11 (Figure 12) in turn positions 1 and 4, respectively, supporting the N8–W11 turn assignment. This $d_{\beta\beta(i,i+3)}$ long-range NOE both increases our level of confidence in the presence of this turn as well as indicates the structural preference for a turn involving these residue positions.

The P9 δH resonances show NOEs not only to the N8 αH resonance but also to the N8 βH resonance; these NOEs have been boxed in in Figure 12. At shorter NOESY mixing times, one of these four cross-peaks (boxed in with a dashed line) appears strongest, suggesting that a preferred N8 side-chain conformation is favored in the structure. This was initially suggested by the presence of nondegenerate N8 βH resonances and significant chemical shift differences from random coil positions as mentioned above. The same can be said for the W11 residue both as regards nondegenerate βH resonances and P12 δH–W11 βH and N8 βH–W11 βH relative NOE magnitudes. In fact, stereospecific βH assignments can be made. The N8 βH proton that gives the largest NOE to one of the P9 δH resonances also gives the largest NOE to one of the W11 βH resonances. That W11 βH resonance, in turn, gives the largest NOE to one of the P12 δH resonances. This information suggests that N8 and W11 βH_2 groups in peptide IV-H1 are relatively less mobile than side-chain β-methylene groups in similar short linear peptides.

Although PG residues are normally found in positions 2 and 3 of type II β-turns (Dyson et al., 1988a), a clear NOE between the G10 NH in position 3 of the turn and the P9 δH resonance (Figure 7) is consistent with a turn type I population. Interestingly, although G10 is proposed to be in turn position 3, its NH is both long-lived and exhibits one of the smallest temperature coefficients, suggesting a significant hydrogen-bond population. Consistent with the network of observed NOEs within this domain is the position of the N8 side-chain amide carbonyl oxygen within range of the G10 NH for possible hydrogen-bond formation. This specific hydrogen bond could only occur in a type I β-turn conformation. Alternatively, a range of orientations for the G10 NH between a type I turn and a γ-turn to the N8 carbonyl would also be consistent with these data. In some ways, a peptide bond orientation more toward a γ-turn, or between a β-turn type I and a γ-turn, would best account for the network of G10 NH–P9 NOEs, as well as for possible steric hinderance with the bulky W11 side-chain ring.

W11, P12, G13, and A14. A fourth β-turn is found in this segment. As with the N8–W11 turn, a d_{NN} NOE is observed between turn position 4 (A14) and turn position 3 (G13) (Figure 9), and a $d_{\beta\beta(i,i+3)}$ NOE is observed between W11 and A14 (Figure 12). An A14 $^3J_{\alpha N}$ coupling constant of 5–6 Hz supports the presence of a preferred β-turn conformation within this domain. The only clearly resolved glycine cross-peak, G13, gives individual $^3J_{\alpha N}$ coupling constants of 8–9 and 5–6 Hz, also consistent with this proposed β-turn. In this case, no $d_{\alpha N(i,i+2)}$ NOE could be observed between P12 and A14 due to overlap of A14 and P12 αH resonances. The presence of strong G13 NH to P12 αH and βH NOEs (Figure 10) supports the presence of a type II β-turn. Moreover, PG residues have a high probability of being in positions 2 and 3 of a type II β-turn (Dyson et al., 1988a). Furthermore, steric hinderance of the P12 and W11 ring systems in a type I turn conformation would probably inhibit significant turn I populations for this specific sequence.

A network of NOEs supportive of this type II turn conformation is also observed between W11 ring proton resonances and P12, A14, P15, and G7 resonances as indicated in Figure 11. Since the W11 2H and 6H ring resonances overlap, ambiguity, regarding structure elucidation, arises from their analysis. On the other hand, W11 4H, 5H, and 7H ring resonances are resolved from each other. Both the 4H and

5H resonances give long-range NOEs to the A14 βH_3 resonances, and, surprisingly, to the G7 αH_2 resonances, while the 7H resonance, and presumably the 6H resonance, gives NOEs to the A14 βH_3 and P15 δH resonances. This network of NOEs is consistent with the W11 phenyl ring positioned above the β-turn plane and the imidazole ring positioned above the P12 ring. The D5–N8 turn plane would then be stacked more or less above the W11 ring plane. In this model, the W11 ring would be somewhat sandwiched between two turns. This structure accounts for the W11 4H/5H to G7 αH_2 NOEs and is consistent with the observation that G13 and A14 resonances are more upfield, presumably ring current, shifted than their random coil counterparts (Bundi & Wüthrich, 1979).

A14, P15, and Y16. While it is unusual for terminal positions in short linear peptides to exist in any perferred conformation, peptide IV-H1 seems to be an exception. The Y16 $^3J_{\alpha N}$ coupling constant, like the A14 coupling constant, is 5–6 Hz, and the Y16 NH is relatively long-lived with a reduced temperature coefficient (Figure 13), suggesting the presence of an interresidue hydrogen bond. The Y16 NH gives an expected NOE not only to the P15 αH resonance (Figure 7) but also to the P15 βH and γH resonances and the A14 βH_3 resonance (Figure 10). The Y16 (2,6) ring proton resonance shows NOEs to the P15 β/γH and A14 βH_3 resonances (Figure 11). On the N-terminal side of P15, the A14 βH_3 resonance gives strong NOEs to the P15 δH resonances. Supporting the idea of this preferred C-terminal conformation is the fact that Y16 βH resonances are nondegenerate and highly chemically shifted from their random coil positions. Taken together, these data strongly suggest a highly populated γ-turn conformation from A14 to Y16, with a hydrogen bond between the Y16 NH and the A14 carbonyl oxygen and the Y16 ring stacked on top of the P15 ring and the γ-turn plane.

Conformational Populations in Other States. Although most structural information derived from these NMR data concerns state ttt, some information is found for other proline-induced isomer states tct, ttc, and ctt. At least some of the preferred structure identified for state ttt is preserved in these minor states. The apparent lack of NMR/NOE information in minor states tct, ttc, and ctt is most probably due to lower overall cis/trans proline isomer populations and perhaps lower structural populations in these states.

No interresidue $i,i+2$ or greater NOEs have yet been observed in minor states tct, ttc, or ctt. NH–NH NOEs are, however, observed between G7 and N8 and between G10 and W11 in states tct and ttc (Figure 9). The state tct W11 NH is also relatively long-lived as in state ttt. This information suggests that turns may be preserved in the same positions as noted in state ttt, i.e., D5–N8 and N8–W11. The same may be said for the V2–D5 turn, but near overlap of D5 NH resonances in these states makes it difficult to assess whether these d_{NN} NOEs are still present.

It is interesting to note that the state ttt medium G13–A14 d_{NN} NOE is not obseved in any of these three minor states. State tct and ttc populations should be large enough to allow observation of this NOE if it were to exist, especially since d_{NN} NOEs were observed for other residue positions in these states. This suggests that the P12 and P15 isomer states are critical to the W11–A14 β-turn formation. While it seems apparent why the P12 cis state would disfavor such a turn structure, it is not apparent why P15 cis state would. Perhaps the *cis*-P15 ring causes steric hinderence to turn formation within this sequence.

In state ttc, the C-terminal Y16 residue also seems to exist in some preferred conformation on the basis of the observation that Y16 βH resonances are better separated in chemical shift than they were in state ttt, but it is unclear what that conformation might be.

Circular Dichroism. The far-ultraviolet CD spectrum of peptide IV-H1 is shown in Figure 14. The spectrum gives a strong negative ellipticity peak at 192 nm with a shoulder at about 205 nm. A weak negative band can also be observed at 235 nm. This spectrum differs appreciably from CD spectra of peptides known to assume random coil conformations in solution. Moreover, this CD trace differs from CD spectra of model peptides containing β-turns of various types (Chrisma et al., 1984; Rose et al., 1985). Normally, β-turn structure exhibits a strong negative band at about 190 nm with two positive bands at about 200–205 and 220–225 nm; helical structures generally give a strong positive band at about 195 nm with weaker negative bands at about 210–215 and 220–225 nm. For peptide IV-H1, the strong negative band at 192 nm by itself is consistent with β-turn type structure (Rose et al., 1985). The remainder of the spectrum suggests the presence of additional conformational types. Gaussian analysis of the CD spectrum in Figure 14 suggests the presence of about 52% β-turn, 10% helix, and 38% random coil or other type structure. Furthermore, the weak negative band at 235 nm is typical of γ-turn conformation (Chrisma et al., 1984), consistent with NMR structural data for the A14–P15–Y16 sequence presented above. In short, the CD spectrum of peptide IV-H1 can most reasonably be said to arise primarily from the presence of two or more β-turns; however, the spectral characteristics are not consistent with the majority of β-turn models examined to date.

Biological Activity of "Walk-Through" Peptides. Since peptide IV-H1 consists of a GX_1X_2 sequence repeat motif that, at least in solution, forms a multiple β-turn structure with glycine always in position 3 of a β-turn, one could ask the question if all or only some of these turns or GX_1X_2 repeats are necessary for full peptide IV-H1 cell adhesion activity—the idea being that perhaps only one turn may be necessary for promoting cell adhesion. In order to gain some insight into this question, overlapping sequence "walk-through" peptides were synthesized as shown in Figure 15. Peptide IV-H1, which consists of five GX_1X_2 repeat units, is number 4 on this list. Other peptides differ from it by moving one GX_1X_2 unit at a time along the parent type IV collagen sequence in either the N- or C-terminal direction; as an internal control, the length of each peptide is always conserved at 15 residues or 5 GX_1X_2 triplets.

Cell adhesion activity was assayed for each peptide, as described under Materials and Methods. The results are displayed in bar graph format in Figure 15 as relative percent cell adhesion. Relative refers to cell adhesion activity relative to peptide IV-H1, which was arbitrarily set at 100%. These results indicate that removal of any part of peptide IV-H1 C-terminal nonet that includes the three proline residues reduces cell adhesion activity to near background, while removal of one or two of the N-terminal GX_1X_2 triplets reduces cell adhesion activity only by about half. Although the mechanism of cell adhesion is quite complex, it seems clear that the GNPGWPGAP sequence is absolutely necessary for any significant level of peptide IV-H1 cell adhesion activity. Any one of these GX_1X_2 units alone is not sufficient; all three must be present.

Discussion

Short linear peptides, like peptide IV-H1, generally exist in solution in an ensemble of highly fluctuating structures

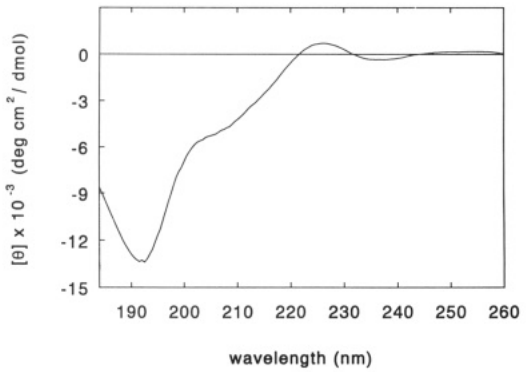
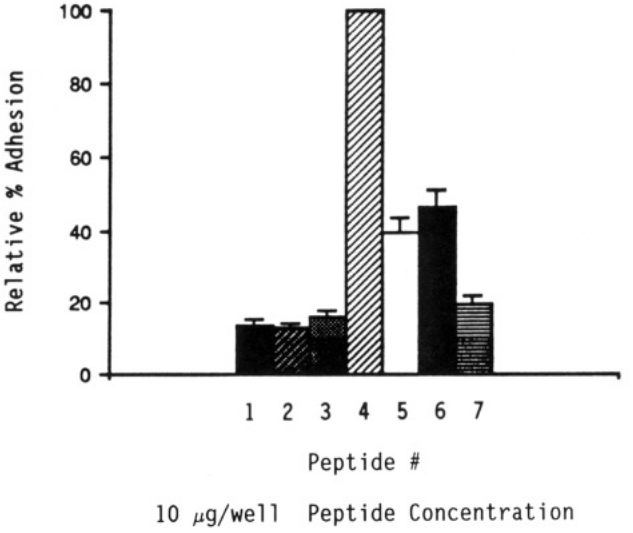


FIGURE 14: Far-ultraviolet CD spectrum of peptide IV-H1. The circular dichroic spectrum of peptide IV-H1 is shown as millidegrees versus wavelength (nm). The peptide concentration is 0.1 mg/mL in 20 mM phosphate buffer, at pH 6.5 and 15 °C. Other experimental variables are given under Materials and Methods.

whose NMR spectral parameters average. Normally the presence of one proline residue, for example in the middle of a peptide chain, would dampen backbone flexibility at its own and, via steric effects, possibly sequentially neighboring peptide bonds, but that would still leave N- and C-terminal segments with considerable degrees of freedom for rapid "NMR time scale" structural interconversion averaging. Two sets of resonances due to cis/trans isomer states will then be observed for residues sequentially neighboring that proline residue. With three proline residues present in one peptide, however, one possibility would be to observe cis/trans states in residues proximal to each respective proline. In other words, residues sequentially well removed from that particular proline would not be sensitive to its cis/trans isomer state. In peptide IV-H1, this is not the case. Each proline undergoes the normally observed, relatively slow, cis/trans peptide bond interconversion, but, at a given residue position in peptide IV-H1, more than the one or two expected conformation states are observed. With three prolines and two isomers, a total of eight states are possible. Some of these have been observed at specific residues, and others have not been observed due to populations that fall into the NMR spectral noise level. For example, five A14, four G13, and three W11 species, to mention a few, have been observed and assigned to isomer-specific conformation states. These observations provide direct evidence that proline cis/trans equilibria are coupled with the overall molecular conformation in peptide IV-H1. Multiple populations indicate that the chemical environment of residues sequentially well removed from any given proline residue is sensitive to that proline residue's isomer state and therefore to that proline's conformation. Peptide IV-H1 being a short linear peptide, therefore, must exist in solution in some defineable conformation(s), and all three proline residues must play a significant role in that definition. A network of NMR/NOE structural constraints indicating the presence of multiple turns has been found for peptide IV-H1. Significant populations of β-turns are centered at K3–G4, K6–G7, P9–G10, and P12–G13. Lastly, a C-terminal γ-turn is centered at the P15 residue. Consistent with these NMR-derived structural domains are CD results that, at the very least, suggest the presence of significant populations of β-turn conformation.

The two N-terminal β-turns are both centered at KG residues, i.e., K3–G4 and K6–G7. The β-turn sequence V2–K3–G4–D5, in particular, exhibits structural features similar to the GRGD sequence in a fibronectin-derived synthetic hexapeptide, GRGDSP (Reed et al., 1989). In that study, two "nested" β-turns were proposed. These turns fit best the type III or type I class. Type I and III turns differ only by ψ,Φ angles of −30° in the i+2 position (Venkatachalam, 1968). For the peptide IV-H1 V2–D5 β-turn, evidence has been given for β-turn type I and type II populations that may be in equilibrium (Stradley et al., 1990). In the hexapeptide, a β-turn is centered at the RG dipeptide, while in peptide IV-H1 the turn is centered at KG residues. In either instance, an aspartic acid residue is at position 4. Peptide IV-H1 from collagen type IV and this hexapeptide from fibronectin also have similar functions in cell adhesion. Both IV-H1 (Chelberg et al., 1990) and RGD (Ruoslahti & Pierschbacher, 1987) peptide sequences bind receptors of the integrin superfamily and promote cell adhesion via this molecular interaction. At present it is evident that the RGD peptide does not inhibit cell attachment to peptide IV-H1 coated substrata, suggesting that different receptors are active in each case (Chelberg et al., 1990). In the case of RGD-containing proteins/peptides (e.g., laminin, fibronectin, fibrinogen, collagen, vitronectin, and von Willebrand factor), cell adhesion specificity is indicated by a number of receptors, integrins (Hynes, 1987), each of which is capable of recognizing only a single RGD-containing protein ligand or a limited number of such ligands.

Collagen IV peptide IV-H1 is the case of a metastatic cell adhesion promoting sequence that does not contain RGD and yet, as we have seen, does contain a KGD sequence with β-turn type I/III structural elements similar to those found in fibrinonectin-derived RGD-containing hexapeptide (Reed et al., 1988). Furthermore, although we have seen in this present



Peptide #	Peptide Sequence
1	GPPGLPGIDGVKGDK
2	GLPGIDGVKGDKGNP
3	GIDGVKGDKGNPGWP
4	GVKGDKGNPGWPGAP
5	GDKGNPGWPGAPGVP
6	GNPGWPGAPGVPGPK
7	GWPGAPGVPGPKGDP

FIGURE 15: Cell adhesion activity of sequences of overlapping IV-H1 peptides. A series of overlapping peptides were synthesized in order to study structure–function relationships in peptide IV-H1. These sequences are given at the bottom of the figure. For cell adhesion activity data, substrata were coated with overlapping peptides from the IV-H1 sequence as designated in the figure. Data are presented as percent relative cell adhesion to that observed on IV-H1 and represent the mean of triplicate determinations.

study that the KGD part of the peptide is not sufficient for activity, its presence is necessary for the full complement of activity. One explanation why this is so and why similar RGD-containing peptide sequences interact specifically with different integrins is that this sequence serves as a shared binding site, with the specificity being generated by a second unique binding site. This idea is consistent with our sequence "walk-through" peptide IV-H1 cell adhesion data where the C-terminal segment containing the three proline residues must be present for any significant activity to be observed. Alternatively, the specificity could reside in the conformation of the RGD/KGD tripeptide, and the role of the surrounding sequences would be to force the tripeptide determinant into an appropriate conformation. This idea would be in agreement with the fact that the RGD sequence can take very different conformations is different proteins to fine tune activity (Ruoslahti & Pieraschbacher, 1987). This β-turn sequence with a cationic group in position 2, a glycine in position 3, and an anionic group in position 4, therefore, may be generally significant to cell adhesion activity.

Preferred solution structures of small proline-containing pentapeptides have been found by NMR and CD analyses (Dyson et al., 1988a). Data indicate that significant β-turn populations are present when a PG sequence occupies turn positions 2 and 3, respectively. For two of the β-turns identified in peptide IV-H1, PG residues are located in these turn positions, i.e., P9–G10, and P12–G13. In the small model peptides, however, differentiation among turn types, i.e., type I or II, could not be made with any certainty. Venkatachalam (1968) has indicated that the PG sequence at turn positions 2 and 3 favors a type II turn. While the peptide IV-H1 P12–G13 β-turn does show type II characteristics, a significant type I β-turn population is indicated in the N8–W11 sequence by the presence of $d_{\delta N(i,i+1)}$ and/or $d_{\gamma N(i,i+1)}$ NOEs between proline and glycine residues. A recent structure analysis of proline-containing cyclic pentapeptides as models for reverse turns (Stradley et al., 1990) showed that type I and II turns are in dynamic equilibrium, fast on the NMR time scale, and that type I turns, with proline in position 2, are present at greater than 50% of the *trans*-proline population. This is consistent with our observation for peptide IV-H1 sequence N8–W11, where a significant population of turn type I has been proposed.

What is also interesting in conjunction with the NMR-derived solution conformation of peptide IV-H1 is the fact that the C-terminal domain from about N8 to Y16 gives relatively larger magnitude NOEs; the largest number of "long-range" NOEs, and most of the longest lived backbone amide protons with reduced temperature coefficients, suggesting not only that this domain is better structurally defined but also that it is less flexible than the N-terminal domain. This is at least probably partly the result of restricted backbone motion imposed by the structurally constraining proline imino acid residues, P9, P12, and P15. Of structure–function interest here is the correlation of this proline-containing sequence with observed cell adhesion activity. Cell adhesion activity studies using a series of overlapping peptides that "walk through" the sequence of peptide IV-H1 have indicated that only peptides containing the nonet GNPGWPGAP showed any significant cell adhesion activity, suggesting that residues and/or structures within the GNPGWPGAP sequence are absolutely necessary for maintaining maximal activity. It may be that a better defined, less flexible structure has the effect of being a more active receptor ligand, thereby imparting greater biological activity.

Of further note is the fact that the PGWPGAP part of the sequence is entirely conserved in human and mouse species (Brazel et al., 1987; Babel & Glanville, 1984; Killen et al., 1988). In particular, tryptophan is rarely found in collagen, and yet it is conserved in at least these two species. This does not imply which residues in this segment are necessary for activity; perhaps all are—some for structure and some for direct receptor interaction. Structurally, this triproline-containing segment gives two well-defined β-turns centered at PG residues. Moreover, this conserved tryptophan residue may be "stacked" onto the W11–A14 β-turn plane as well as being positioned under (respectively) the N8–W11 β-turn plane. In effect, W11 is somewhat sandwiched between both turns, perhaps giving additional stabilization to the structure. This type of structure would further constrain the tryptophan's already normally dampened internal motions and explain why stereospecific W11 βH assignments could be made.

The fact that antipeptide IV-H1 antibodies cross-react with parent type IV collagen (Chelberg et al., 1990) strongly suggests that the peptide IV-H1 sequence is positioned at a surface domain in the parent molecule. This is relevant for two reasons: (1) interaction of peptide IV-H1 sequence with its receptor as discussed above, and (2) hydroxylation of proline residue(s) by proline hydroxylase known to occur in collagens in conjunction with triple helix stabilization and glycosylation. Brahmachari and Ananthanarayanan (1978, 1979) proposed that proline hydroxylase selectively hydroxylates prolyl residues in position 2 of β-turn-containing X_1–PG–X_2 sequences. Moreover, the identity of the preceeding and the residue following the PG site are considered important in the regulation of hydroxylation. In support of this, Bhatnager et al. (1978) found that two model sequences, $(VPGVG)_n$ and $(APGG)_n$ (Rapaka et al., 1978), could inhibit collagen hydroxylation by proline hydroxylase. Furthermore, hydroxyproline formation in collagen leads to enhanced stability of the collagen triple helix (Rosenbloom et al., 1973; Berg & Prockop, 1973) due to hydrogen bonding between the hydroxyproline's hydroxyl group and the peptide backbone (Ramachandran et al., 1973). Regarding peptide IV-H1, the P12 position in native type IV collagen has been found to be hydroxylated (Killen et al., 1988) and P12, at least in peptide IV-H1, does exist in position 2 of a β-turn.

The biopolymer $(GPP)_n$ has been used as a model system for triple-helical collagen structure since its X-ray diffraction pattern shows the main features of that of collagen, with closely similar helical parameters (Traub et al., 1967). $(GPP)_n$ does not exist in a strict β-turn conformation (Yonath & Traub, 1969). However, on one hand, $(GPP)_n$ is not a collagen type sequence repeat unit like $(GP–X_1)_n$ or $(G–X_1X_2)_n$; on the other hand, the short hexadecapeptide IV-H1 is not typical of the large collagen molecular triple helix. A general comparison of the X-ray derived structure of $(GPP)_n$ and peptide IV-H1 indicates that positions 1, 2, and 3 of our W11–A14 β-turn are similar to the GP structure in $(GPP)_n$ but that the fourth residue in this β-turn confers most of the structural disparity. Therefore, some of the structural features in both systems are similar. Perhaps triple-helix formation with peptide IV-H1 modifies the ψ,Φ angles at positions 3 and 4 of these β-turns to induce more similar structures, or maybe this multiple β-turn repeat structure in peptide IV-H1 makes it unique for cell receptor recognition in a structural pattern that is closer to the $(PPG)_n$ structure.

This multiple β-turn structural motif found to exist throughout the peptide IV-H1 sequence may persist through a larger sequence segment of the collagen type IV triple-helix

domain For peptide IV-H1, β-turns were found to be centered at PG and KG residues, with glycine always in position 3 of the turn. If we view a larger segment of the sequence, for example:

where we have extended the IV-H1 sequence by adding several GX_1X_2 repeats on either end of the peptide, the observation can be made that PG and KG units generally repeat at equal intervals. Assuming the same structural motif, brakets above and below this sequence indicate positions of proposed β-turns continuing out from the IV-H1 parent sequence. Only the initial N-terminal turn has DG instead of PG or KG at its center position 2 and 3. DG, however, is also found in this type of turn structure. It may be that type IV collagen contains this turn repeat unit often throughout its structure. We are currently investigating this longer peptide to see if this is indeed the case.

Lastly, although a preferred solution conformation for peptide IV-H1 has been derived from these NMR data, it may or may not be the biologically active structure. For instance, while the C-terminal γ-turn is structurally interesting, it is probably not biologically significant since the GX_1X_2 repeat is prematurely disrupted and tyrosine does not naturally occur at this position. On the other hand, since antipeptide IV-H1 antibodies cross-react with parent type IV collagen and cell receptors (Chelberg et al., 1990), this would suggest that the derived solution conformation is close to the active one. Moreover, proline and glycine resiudes are most often found in turn-type structures, and PG sequences generally exist in β-turn positions 2 and 3. This is consistent with what has been observed in peptide IV-H1. Furthermore, since peptide IV-H1 possesses such a well-defined solution conformation, it seems likely that this structure is at least generally preserved when interacting with its receptor. This remains to be seen, however. In the case of an enzyme–ligand or receptor-binding complex, the "bound" state can be investigated by performing NOESY experiments on the peptide in question in the presence of its receptor (Ni & Scheraga, 1989b,c). These studies are now underway in conjunction with computer modeling studies.

Registry No. Peptide IV-H1, 129015-02-7; peptide AAA, 134938-83-3; peptide SPP, 134938-84-4; peptide PSP, 134938-85-5; peptide PPS, 134938-86-6; GPPGLPGIDGVKGDK, 134938-87-7; GLPGIDGVKGDKGNP, 134938-88-8; GIDGVKGDKGNPGWP, 134938-89-9; GDKGNPGWPGAPGVP, 134938-90-2; GNPGWPGAPGVPGPK, 134938-91-3; GWPGAPGVPGPKGDP, 134938-92-4; GVKGDKGNPGWPGAP, 119953-07-0.

References

Aue, W. P., Bartholdi, E., & Ernst, R. R. (1976) *J. Chem. Phys. 64*, 2229–2246.

Aumailley, M., & Timpl, R. (1986) *J. Cell Biol. 103*, 1569–1575.

Babel, W., & Glanville, R. W. (1984) *Eur. J. Biochem. 143*, 545–556.

Bax, A., & Davis, D. G. (1985) *J. Magn. Reson. 65*, 355–360.

Berg, R. A., & Prockop, D. J. (1973) *Biochem. Biophys. Res. Commun. 52*, 115–120.

Bhatnagar, R. S. Rapaka, R. S., & Urry, D. W. (1978) *FEBS Lett. 95*, 61–64.

Bierzynski, A., Kim, P. S., & Baldwin, R. L. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 2470–2474.

Brahmachari, S. K., & Ananthanarayanan, V. S. (1978) *Curr. Sci. 47*, 107–109.

Brahmachari, S. K., & Ananthanarayanan, V. S. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 5119–5123.

Brazel, D., Oberbäuer, I., Dieringer, H., Babel, W., Glanville, R. W., Deutzmann, R, & Kühn, K. (1987) *Eur. J. Biochem. 168*, 529–536.

Bruch, M. D., McKnight, C. J., & Gierasch, L. M. (1989) *Biochemistry 28*, 8554–8561.

Bundi, A., & Wüthrich, K. (1979) *Biopolymers 18*, 285–298.

Charonis, A. S., Tsilibary, E. C., Yurchenco, P. D., & Furthmayr, H. (1985) *J. Cell Biol. 100*, 1848–1853.

Chelberg, M. K., Mc Carthy, J. B., Skubitz, A. P. N., Furcht, L. T., & Tsilibary, E. C. (1990) *J. Cell Biol. 111*, 261–270.

Chou, P. Y., & Fasman, G. D. (1978) *Adv. Enzymol. Relat. Areas Mol. Biol. 48*, 45–148.

Chrisma, M., Fasman, G. D., Balaram, H., & Balaram, P. (1984) *Int. J. Pept. Protein Res. 23*, 411–419.

deGennes, P. G. (1979) in *Scaling Concepts in Polymers Physics*, Cornell University Press, Ithaca, NY.

Dyson, H. J., Cross, K. J., Houghten, R. A., Wilson, I. A., Wright, P. E., & Lerner, R. A. (1985) *Nature (London) 318*, 480–483.

Dyson, H. J., Cross, K. J., Ostresh, J., Houghten, R. A., Wilson, I. A., Wright, P. E., & Lerner, R. A. (1986) in *Synthetic Peptides as Antigens. CIBA Foundation Symposium 119*, p 5875, John Wiley & Sons, Chichester, U.K.

Dyson, H. J., Rance, M. Houghten, R. A., Lerner, R. A., & Wright, P. E. (1988a) *J. Mol. Biol. 201*, 161–200.

Dyson, H. J., Rance, M., Houghten, R. A., Wright, P. E., & Lerner, R. A. (1988b) *J. Mol. Biol. 201*, 201–217.

Fugiwara, S., Wiedemann, H., Timpl, R., Lustig, A., & Engel, J. (1984) *Eur. J. Biochem. 143*, 145–157.

Gehlsen, K. R., Dillner, L., Engvall, E., & Ruoslhati, E. (1986) *Science 241*, 1228–1229.

Grathwohl, C., & Wüthrich, K. (1976) *Biopolymers 15*, 2025–2041.

Grathwohl, C., & Wüthrich, K. (1981) *Biopolymers 20*, 2623–2633.

Herbst, T. J., McCarthy, J. B., Tsilibary, E. C., & Furcht, L. T. (1988) *J. Cell Biol. 1067*, 1365–1373.

Hofmann, H., Voss, T., Kühn, K., & Engel, J. (1984) *J. Mol. Biol. 172*, 325–343.

Humphries, M. J., Komoriya, A., Akiyama, S. K., Olden, K., & Yamada, K. M. (1987) *J. Biol. Chem. 262*, 6886–6892.

Hynes, R. O. (1987) *Cell 48*, 549–555.

Ignatius, M. J., & Reichardt, L. F. (1988) *Neuron 1*, 713–725.

Jeener, J., Meier, B. H., Bachman, P., & Ernst, R. R. (1979) *J. Chem. Phys. 71*, 4546–4553.

Kessler, H., & Bermel, W. (1986) in *Methods in Stereochemical Analysis 6: Applications of NMR Spectroscopy to Problems in Stereochemistry and Conformational Analysis* (Takenchi, Y., & Marchand, A. P., Eds.) VCH, Deerfield Beach, FL.

Kilchherr, E., Hoffmann, H., Steigemann, W., & Engel, J. (1985) *J. Mol. Biol. 186*, 403–415.

Killen, P. D., Burbelo, P., Sakurai, Y., & Yamada, Y. (1988) *J. Biol. Chem. 263*, 8706–8709.

Kim, P. S., & Baldwin, R. L. (1984) *Nature (London) 307*, 329–334.

Kurkinen, M., Taylor, A., Garrels, J. I., & Hogan, B. L. M. (1984) *J. Biol. Chem. 259*, 5915–5922.

Laurie, G. W., Bing, J. T., Kleinman, H. K., Hassell, J. R., Aumailley, M., Martin, G. R., & Feldman, R. J. (1986) *J. Med. Biol. 189*, 205–216.

Mammi, S., Foffani, M. T., Peggion, E., Galleyrand, J. C., Bali, J. P., Simonetti, M., Göhring, W., Moroder, L., &

Wünsch, E. (1989) *Biochemistry 28*, 7182–7188.
Mayo, K. H., Burke, C., Lindon, J. N., & Kloczewiak, M. (1990) *Biochemistry 29*, 3277–3286.
McCarthy, J. B., Skubitz, A. P. N., Zhao, Q., Yi, X., Mickelson, D. J., Klein, D. J., & Furcht, L. T. (1990) *J. Cell Biol. 110*, 777–787.
Murray, J. C., Stingl, G., Kleinman, H. K., Martin, G. R., & Katz, S. I. (1979) *J. Cell Biol. 80*, 197–202.
Ni, F., Scheraga, H. A., & Lord, S. T. (1988) *Biochemistry 27*, 4481–4490.
Ni, F., Konishi, Y., Frazier, R. B., & Scheraga, H. A. (1989a) *Biochemistry 28*, 3082–3094.
Ni, F., Meinwald, Y. C., Vasquez, M., & Scheraga, H. A. (1989b) *Biochemistry 28*, 3094–3105.
Ni, F., Konishi, Y., Bullock, L. D., Rivetna, M. N., & Scheraga, H. A. (1989c) *Biochemistry 28*, 3106–3119.
Otter, A., Kotovych, G., & Scott, P. G. (1989) *Biochemistry 28*, 8003–8010.
Piantini, U., Sørensen, O. W., & Ernst, R. R. (1982) *J. Am. Chem. Soc. 104*, 6800–6805.
Ramachandran, G. N., Bansal, M., & Bhatnagar, R. S. (1973) *Biochim. Biophys. Acta 322*, 166–171.
Rapaka, R. S., Okamoto, K., & Urry, D. W. (1978) *Int. J. Pept. Protein Res. 20*, 218–237.
Reed, J., Hull, W. E., von der Lieth, C.-W., Kübler, D., Suhai, S., & Kinzel, V. (1988) *Eur. J. Biochem. 178*, 141–154.
Rosenbloom, J., Harsh, M., & Jimenez, S. A. (1973) *Arch. Biochem. Biophys. 158*, 478–484.
Rose, G. D., Gierasch, L. M., & Smith, J. A. (1985) *Adv. Protein Chem. 37*, 1–110.
Ruoslahti, E., & Pierschbacher, M. D. (1987) *Science 238*, 491–497.
Shaka, A. J., & Freeman, R. (1983) *J. Magn. Reson. 51*, 161–169.
Shoemaker, K. R., Kim, P. S., Brems, D. N., Marqusee, S., York, E. J., Chaikin, I. M., Stewart, J. M., & Baldwin, R. L. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 2349–2353.
States, D. J., Haberkorn, R. A., & Ruben, D. J. (1982), *J. Magn. Reson. 48*, 286–293.
Stewart, J. M., & Young, J. D. (1984) *Solid Phase Peptide Synthesis*, 2nd ed., p 135, Pierce Chemical Co., Rockford, IL.
Stradley, S. J., Rizo, J., Bruch, M. D., Stroup, A. N., & Gierasch, L. M. (1990) *Biopolymers 29*, 263–287.
Sugrue, S. P. (1987) *J. Biol. Chem. 262*, 3338–3343.
Timpl, R., & Dziadek, M. (1986) *Int. Rev. Pathol. 29*, 1–112.
Timpl, R., Wiedemann, H., von Velder, V., Furthmayr, H., & Kuhn, K. (1981) *Eur. J. Biochem. 120*, 203–211.
Tomaselli, K. J., Damsky, C. H., & Reichardt, L. F. (1987) *J. Cell Biol. 105*, 2347–2358.
Traub, W., Shmueli, U., Suwalsky, M., & Yonath, A. (1967) in *Conformation of Biopolymers* (Ramachandran, G. N., Ed.) Vol. 2, p 449, Academic Press, New York.
Tsilibary, E. C., & Charonis, A. S. (1986) *J. Cell Biol. 103*, 2467–2473.
Venkatachalam, C. M. (1968) *Biopolymers 6*, 1425–1436.
Wider, G., Macura, S., Anil-Kumar, Ernst, R. R., & Wüthrich, K. (1984) *J. Magn. Reson. 56*, 207–234.
Woody, R. W. (1968) *J. Chem. Phys. 49*, 4797–4803.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley-Interscience, New York.
Wüthrich, K., Billeter, M., & Braun, W. (1984) *J. Mol. Biol. 180*, 715–740.
Yonath, A., & Traub, W. (1969) *J. Mol. Biol. 43*, 461–478.
Yurchenco, P. D., & Furthmayr, H. (1983) *Biochemistry 23*, 1839–1850.
Yurchenco, P. D., & Rubin, G. C. (1987) *J. Cell Biol. 105*, 2559–2568.