Sykes, B. D., & Hull, W. E. (1978) *Methods Enzymol. 49*, 270–295.
Sykes, B. D., & Weiner, J. H. (1980) in *Magnetic Resonance in Biology* (Cohen, J. S., Ed.) Vol. 1, pp 171–196, Wiley, New York.
Szer, W., & Shugar, D. (1963) *Acta Biochim. Pol. 10*, 219–231.
Wempen, I., Duschinsky, L., Kaplan, L., & Fox, J. J. (1961) *J. Am. Chem. Soc. 83*, 4755–4766.
Yaniv. M., & Gros, F. (1969) *J. Mol. Biol. 44*, 1–15.

Sequence-Specific ^{1}H NMR Assignments and Structural Characterization of Bovine Seminal Fluid Protein PDC-109 Domain b†

Keith L. Constantine,‡ Vasudevan Ramesh,‡,§ László Bányai,∥ Maria Trexler,∥ László Patthy,∥ and Miguel Llinás*,‡

Department of Chemistry, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, and Institute of Enzymology, Biological Research Center, 1502 Budapest, Hungary

Received August 17, 1990; Revised Manuscript Received November 8, 1990

ABSTRACT: Sequence-specific resonance assignments for the isolated second or b domain of the bovine seminal fluid protein PDC-109 have been obtained from analysis of two-dimensional ^{1}H NMR experiments recorded at 500 MHz. These assignments include the identification of all aromatic and most aliphatic amino acid resonances. Stereospecific assignment of resonances stemming from the Val^2 $CH_3^{\gamma,\gamma'}$ groups and from seven $CH^{\beta,\beta'}$ geminal pairs has been accomplished by analysis of $^3J_{\alpha\beta}$ coupling constants in conjunction with patterns of cross-peak intensities observed in two-dimensional nuclear Overhauser effect (NOESY) spectra. Analysis of NOESY and $^3J_{\alpha NH}$ data reveals a small antiparallel β-sheet involving stretches containing residues 25–28 and 39–42, a *cis*-proline residue (Pro^4), antiparallel strands consisting of residues 1–3, 5–7, and 10–13, and an aromatic cluster composed of Tyr^7, Trp^{26}, and Tyr^{33}. The results of distance geometry and restrained molecular dynamics calculations indicate that the global fold of the PDC-109 b domain, a type II module related to those found in fibronectin, is somewhat different from that predicted by modeling the structure on the basis of homology between type II and kringle units. A shallow depression in the molecular surface which presents a solvent-exposed hydrophobic area—a potential ligand-binding site—is identified in the NMR-based models.

PDC-109 is a major protein component of bovine seminal plasma (Esch et al., 1983). Analysis of the primary structure revealed that PDC-109 consists of two homologous domains, a and b, of 38 and 41 amino acid residues, respectively (Esch et al., 1983). The domains are characterized by a double cystine polypeptide loop pattern distinguished by disulfide bridges connecting Cys residue pairs 1–3 and 2–4. The PDC-109 modules are members of a family of domains collectively known as type II structures (Petersen et al., 1989). Two collinear type II repeats are present in both of the chains of fibronectin (Petersen et al., 1983; Skorstengaard et al., 1984; Kornblihtt et al., 1985). Type II domains are also found as single units in blood coagulation factor XII (McMullen & Fujikawa, 1985) and insulin-like growth factor receptor II (Morgan et al., 1987), as two tandem domains in the bovine seminal plasma protein BSP-A3 (Seidah et al., 1987), and as three head-to-tail repeats in the 72- and 92-kDa type IV collagenases (Collier et al., 1988; Wilhelm et al., 1989). In addition, a remote homology has been identified between the ~40 amino acid, two-disulfide type II domains and the ~80 amino acid, three-disulfide kringle structures found in blood clotting and fibrinolytic proteins (Patthy et al., 1984). In the absence of experimental data, this homology was used to generate molecular models of various type II domains from the X-ray crystallographic structure of bovine prothrombin kringle 1 (Holland et al., 1987).

The physiological roles of PDC-109 are under active investigation. PDC-109 has been observed to stimulate the release of gonadotropins when added to cultures of pituitary cells (Manjunath & Sairam, 1987), and there is evidence that sperm motility is affected by PDC-109 binding (Manjunath et al., 1987). PDC-109 has recently been shown to interact with apolipoprotein A-I in vitro (Manjunath et al., 1989). Fibronectin-like antigens in human seminal plasma (Vuento et al., 1980) are likely to be closely related to PDC-109. The functionally better characterized type II domains are those found in fibronectin, which are unique to the gelatin (denatured collagen) binding region. Thus, the type II domains are believed to be involved in fibronectin/collagen interactions (Yamada, 1989). It is also noteworthy that PDC-109 can be readily purified on a gelatin-agarose column (Manjunath et al., 1987), directly demonstrating gelatin binding mediated by type II domains.

Structural studies of isolated type II domains are crucial if the physiological roles of these modules are to be understood in molecular detail. Here we report the results of two-dimensional (2D)[1] ^{1}H NMR studies of a fragment of PDC-109,

† This research was supported by the U.S. Public Health Service, NIH Grant HL-29409, and by Grants OTKA 323, Tt 174/1986, and AKA 1-3-86-224 of the Hungarian Academy of Sciences. A grant from the Pittsburgh Supercomputing Center through the NIH Division of Research, Cooperative Agreement U41RR04154, is gratefully acknowledged. K.L.C. is a recipient of an NIH predoctoral training grant.
‡ Carnegie Mellon University.
§ Current address: Department of Biochemistry, University of Leicester, Leicester LE1 7RH, U.K.
∥ Institute of Enzymology.

0006-2960/91/0430-1663$02.50/0

consisting of the C-terminal 45 residues, which contains the second type II domain (PDC-109/b). To our knowledge, this is the only isolated, intact type II domain so far obtained from any source. An aqueous saline solution containing 10% acetone/H_2O (v/v) was found to provide suitable resolution with minimal conformational perturbation. By use of the sequential assignment procedure (Wüthrich, 1986), the backbone and most side-chain resonances of the residues in PDC-109/b have been assigned. Our results demonstrate that PDC-109/b adopts a folded, globular conformation. Details of local and secondary structural features were revealed via a qualitative estimation of cross-peak intensities in 2D nuclear Overhauser effect (NOESY) spectra and of ^3J coupling constants in high-resolution phase-sensitive chemical shift correlated (COSY) spectra. A number of dihedral angle and interproton distance estimates were incorporated as constraints in a distance geometry/restrained molecular dynamics protocol to generate 16 independent molecular models of PDC-109/b. A PDC-109/b structure was also produced in a manner independent of the NMR data by modeling based on the 2.8-Å X-ray structure of bovine prothrombin kringle 1 (Park & Tulinsky, 1986). The modeling reveals that the global fold of PDC-109/b is somewhat different from that expected on the basis of the homology to kringle domains (Patthy et al., 1984; Holland et al., 1987). A surface depression, suggestive of a potential ligand-binding site, appears to be reasonably well-defined in most NMR-derived models.

Materials and Methods

Sample Preparation. The PDC-109/b samples used belonged to a batch described elsewhere (Bányai et al., 1990). For NMR experiments, PDC-109/b samples were dissolved to concentrations of 2 mM in D_2O/[$^2H_6$]acetone (9:1 v/v) or H_2O/[$^2H_6$]acetone (9:1 v/v) solvent mixtures containing 50 mM NaCl. (These solvent systems will henceforth be referred to as D_2O and H_2O, respectively.) Prior to preparation of the D_2O sample, all exchange-labile protons were replaced with deuterons by allowing the sample to equilibrate at least 30 min at room temperature in D_2O followed by lyophilization. For the H_2O (D_2O) sample, the pH* was adjusted to 6.8 (6.5). As an internal chemical shift reference, *p*-dioxane was added and was assumed to resonate 3.77 ppm to low-field relative to the sodium (trimethylsilyl)[2,2,3,3-$^2H_4$]propionate signal (De Marco, 1977). For experiments in H_2O, the [$^2H_6$]acetone signal served as the field frequency lock.

NMR Spectroscopy. ^{1}H NMR spectra were recorded with a Bruker WM-300 spectrometer equipped with an Aspect 2000 computer and with a Bruker AM-500 spectrometer equipped with an Aspect 3000 computer. All 2D experiments were recorded at 500 MHz in the phase-sensitive mode with quadrature detection, except for RELAYED-COSY spectra, which were recorded in magnitude mode. Typical spectral widths in ω_2 were 5400 Hz in D_2O spectra and 7800 Hz in H_2O spectra, with the spectral widths in ω_1 set to half these values. Generally, 512 or 768 equally spaced t_1 experiments were collected and Fourier transformed to yield 2048 and 1024 real data points in ω_2 and in ω_1, respectively. For the phase-sensitive spectra this afforded digital resolutions in ω_2 of 2.7 Hz/point (D_2O) or 3.9 Hz/point (H_2O). For improved estimation of ^3J coupling constants, COSY spectra were also processed to yield, after zero filling, digital resolutions in ω_2 of 1.4 Hz/point (D_2O) and 2.0 Hz/point (H_2O).

The 2D NMR experiments were recorded with standard pulse sequences and phase-cycling schemes. The H_2O or residual HDO resonance was suppressed by preacquisition continuous low-power irradiation, with decoupler/transmitter phase coherence. Experiments acquired at 27 °C included COSY (Marion & Wüthrich, 1983), DQF-COSY (Rance et al., 1983), RELAYED-COSY (Wagner, 1983), total correlation (TOCSY) (Braunschweiler & Ernst, 1983; Bax & Davis, 1985), and NOESY (Macura & Ernst, 1980; Kumar et al., 1980; Bodenhausen et al., 1984) spectra. For RELAYED-COSY experiments, delay times of 25 and 30 ms were employed. Mixing times (τ_m) of 56 and 84 ms were used for TOCSY spectra recorded in D_2O and H_2O, respectively. NOESY spectra were recorded with mixing times of 60, 120, 200, and 300 ms, with a 10% random variation of τ_m to quench zero-quantum coherences. In order to shift the solvent resonance and to help resolve degeneracies, additional spectra were also recorded at 35 and 45 °C. COSY and RELAYED-COSY spectra were processed with unshifted sine-bell weighting in both ω_1 and ω_2, NOESY with 45° shifted sine-bell squared filtering in both dimensions, and TOCSY spectra with Gaussian windows.

Sequence-specific resonance assignments were obtained from analysis of 2D spectra recorded at 500 MHz following well-established procedures (Wüthrich, 1986; Chazin et al., 1988; Chazin & Wright, 1988; Englander & Wand, 1987). Initially, spin systems were identified according to type by examining 2D scalar correlated spectra (COSY, RELAYED-COSY, TOCSY) recorded in D_2O. Generally, this allowed for the identification of the unique Gly, Ala, Val, Ile, Leu, and Thr spin systems, as well as the aromatic ring spin systems. The AMX spin systems (Ser, Cys, Asp, Asn, and the CH^{α}–$CH^{\beta,\beta'}$ fragments of the aromatic residues) were distinguished from the remaining long side chain spin systems (Glu, Met, Pro, Arg, and Lys). NOESY spectra recorded in D_2O were then examined in order to connect the aromatic rings to their backbone fragments. Analysis of scalar-correlated spectra recorded in H_2O allowed for the observation of exchangeable protons and facilitated the identification of spin systems. Observation of NH–NH, CH^{α}–NH, and CH^{β}–NH cross-peaks in NOESY spectra recorded on samples dissolved in H_2O allowed for unambiguous sequence-specific assignment of most residues.

Stereospecific assignment of the Val^2 $CH_3^{\gamma,\gamma'}$ groups and several $CH^{\beta,\beta'}$ groups was accomplished according to established procedures (Zuiderweg et al., 1985; Wagner et al., 1987). The $CH^{\beta,\beta'}$ assignments are based on observed patterns of $^3J_{\alpha\beta}/^3J_{\alpha\beta'}$ splittings and, most crucially, intraresidue CH^{α}–$CH^{\beta,\beta'}$ and NH–$CH^{\beta,\beta'}$ NOE intensities. Stereospecific assignment also allows the corresponding χ^1 angle to be constrained to 60°, −60°, or 180°. In cases where stereospecific assignment cannot be achieved, it is still often possible to confirm or exclude $\chi^1 = 60°$ on the basis of $^3J_{\alpha\beta}$ values alone. Backbone ϕ dihedral angle constraints were also obtained, with $-60° > \phi > -180°$ whenever $^3J_{\alpha NH} > 8$ Hz.

A set of approximate upper bound interproton distance constraints was derived from the NOESY spectra via qualitative comparison of cross-peak intensities at the various mixing times to sequential CH^{α}–NH cross-peaks (3.6 Å > $d_{\alpha N}$ > 2.2 Å) in H_2O spectra or to intraresidue CH^{α}–CH^{β} cross-peaks

[1] Abbreviations: ABNR, adopted-basis Newton–Raphson; COSY, phase-sensitive two-dimensional chemical shift correlated spectroscopy; d_{lm}, distance or dipolar (Overhauser) connectivity between protons l and m; DQF-COSY, phase-sensitive double quantum filtered COSY; NOESY, phase-sensitive two-dimensional nuclear Overhauser effect spectroscopy; PDC-109/b, PDC-109 domain b; pH*, glass electrode pH reading uncorrected for solvent or isotope effects; RELAYED-COSY, absolute value relayed COSY; RMSD, root-mean-squares difference; τ_m, mixing time; TOCSY, phase-sensitive two-dimensional total correlation spectroscopy; 2D, two dimensional.

(2.9 Å > $d_{\alpha\beta}$ > 2.2 Å) in D_2O spectra. Since interproton distance constraints are treated as upper bounds only, an effort was made to err on the side of overestimation. Upper bound distances were generally set between 2.5 and 5.0 Å, with a 0.5-Å correction added for methyl groups.

Molecular Modeling. Distance geometry calculations were performed on a VAX 11/780 computer, and restrained energy minimization and restrained molecular dynamics calculations were performed on the Cray Y-MP of the Pittsburgh Supercomputing Center. Sixteen initial conformations were computed with the metric matrix distance geometry program DISGEO (Havel & Wüthrich, 1984; Havel, 1986), implementing appropriate corrections for pseudoatoms (Wüthrich et al., 1983). These structures were optimized by the molecular mechanics/dynamics program CHARMM (Brooks et al., 1983), with harmonic and half-harmonic potential terms applied to enforce dihedral angle and upper distance bound constraints, respectively. DISGEO pseudoatom constraints were taken to the appropriate heavy atoms for the CHARMM calculations. The DISGEO structures were subjected to 2000 steps of adopted-basis Newton–Raphson (ABNR) restrained energy minimization, restrained molecular dynamics using a simulated annealing protocol, and a final 1200-step restrained ABNR minimization. The vacuum molecular dynamics simulations were carried out in three stages as follows: (1) simulation temperature fixed at 1000 K with interproton distance constraint force constants incremented from 2 to 40 kcal mol^{-1} Å^{-2} in 2 kcal mol^{-1} Å^{-2} increments, 0.2 ps per increment; (2) force constants fixed at 40 kcal mol^{-1} Å^{-2} with the temperature lowered to 300 K in 100 K steps, 0.2 ps per temperature increment; (3) 3 ps of restrained dynamics with the temperature fixed at 300 K and the force constants fixed at 40 kcal mol^{-1} Å^{-2}. Integration of Newton's equations of motion employed a time step = 1 fs. Dihedral angle constraint force constants were held fixed for all calculations at 100 kcal mol^{-1} rad^{-2}, except for a constraint designed to hold Pro^4 in the cis conformation, which was set to 200 kcal mol^{-1} rad^{-2}. For the restrained energy minimizations, interproton distance constraint force constants were set to the relatively mild value of 10 kcal mol^{-1} Å^{-2} to help offset potential effects of distance underestimation. The CHARMM Version 19 parameter and topology files were modified to incorporate explicit α-protons in addition to explicit polar protons. The standard two-well potential term for the ω peptide bond dihedral angles was replaced by a single-well (trans) term. All calculations incorporated a distance-dependent dielectric constant numerically equal to the internuclear separation in angstroms. The resulting structures were analyzed in terms of constraint violations, internal energies, and root-mean-square differences (RMSDs) between structures.

The proposed homology between kringle structures and the type II domains (Patthy et al., 1984) motivated us to model PDC-109/b from the prothrombin kringle 1 coordinates available at 2.8-Å resolution (Park & Tulinsky, 1986).[2] The 45-residue PDC-109/b fragment was aligned with a 57-residue fragment of prothrombin kringle 1 as follows:

```
kringle 1: S G I E C Q L W R S R Y P H K P E I N S T T H P G A D L
PDC-109/b: D Y A K C V F P                             F I Y G G K

kringle 1: R E N F C R N P D G S I T G P W C Y T T     S P T L R
PDC-109/b: K Y E T C   T K I G S M W M S W C S L S P N Y D K D R A

kringle 1: R E E C
PDC-109/b: W K Y C
```

[2] Holland et al. (1987) have derived models of type II domains using a 3.8-Å crystallographic structure of prothrombin kringle 1. Their modeling was accomplished with interactive molecular graphics, whereas our approach also relies on energy minimization and molecular dynamics.

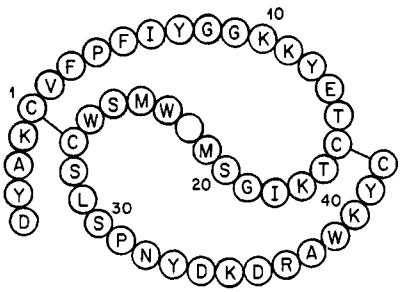

FIGURE 1: Primary structure of the PDC-109 b-domain (PDC-109/b). Residues are labeled according to the standard one-letter amino acid code and numbered by reference to the first half-cystine, which has four residues preceding it. The sequence lacks His and Gln residues. A deletion at position 22 has been introduced to facilitate alignment with homologous type II sequences (Petersen et al., 1989).

Before residue substitutions were made, deletions were closed, and a space for the P-N-Y insert was made, by subjecting the aligned kringle sequence to a 1600-step ABNR energy minimization using CHARMM with suitable harmonic constraint terms (Thewes et al., 1990). The three-residue insert was fitted into the resulting gap, and residue substitutions were made with the molecular graphics program HYDRA (Hubbard, 1986); this was followed by an additional 1200 steps of ABNR minimization. Eight picoseconds of unconstrained molecular dynamics at 300 K was then executed, and the average structure over the last 6 ps was calculated. A final 1200 steps of unconstrained ABNR minimization was performed to generate a model of PDC-109/b based solely on the proposed homology to kringle domains.

Routine inspection of the derived PDC-109/b models was facilitated by use of the HYDRA program (Hubbard, 1986), which is supported by an Evans & Sutherland PS300 graphics workstation. This system was also used to produce Figure 10.

Results and Discussion

Preliminary Characterization. The primary structure of PDC-109/b is illustrated by a planar projection in Figure 1. Initially, ^{1}H NMR spectra of this fragment were acquired at 300 MHz with samples dissolved in pure ^2H_2O at pH* 7.0, 22 °C. No slowly exchanging amide or indole ring NH proton resonances are observed after 30 min of sample preparation, reflecting rapid ^{1}H–^{2}H exchange, which indicates a relatively high conformational flexibility. At neutral pH* the spectrum of the protein appeared broad even at a minimum NMR sample concentration of 1 mM, probably due to aggregation (Manjunath & Sairam, 1987). Upon the addition of small amounts of NaCl (50 mM) and [$^2H_6$]acetone (10% v/v), there was an appreciable sharpening of the resonances compared to spectra recorded in pure ^2H_2O. Only minimal perturbations of the chemical shifts of all resonances were observed, including those appearing far removed from typical random coil positions. This demonstrated the existence of a folded, stable conformation in the mixed solvent which is very similar to or identical with the native structure. Therefore, all subsequent experiments were carried out with samples dissolved in the [$^2H_6$]acetone/NaCl-containing media which, for brevity, are referred to as D_2O and H_2O for the ^2H_2O and ^1H_2O solutions, respectively.

Resonance Assignments. PDC-109/b contains five Tyr, three Trp, and two Phe residues. Figure 2 shows the aromatic region of a RELAYED-COSY spectrum recorded on a PDC-109/b sample in D_2O; most ring proton connectivities are evident in this display. For clarity, the sequentially assigned positions (see below) are indicated in Figure 2. All Tyr ring spin systems were located and subsequently linked to the backbone spin systems by the observation of strong NOESY

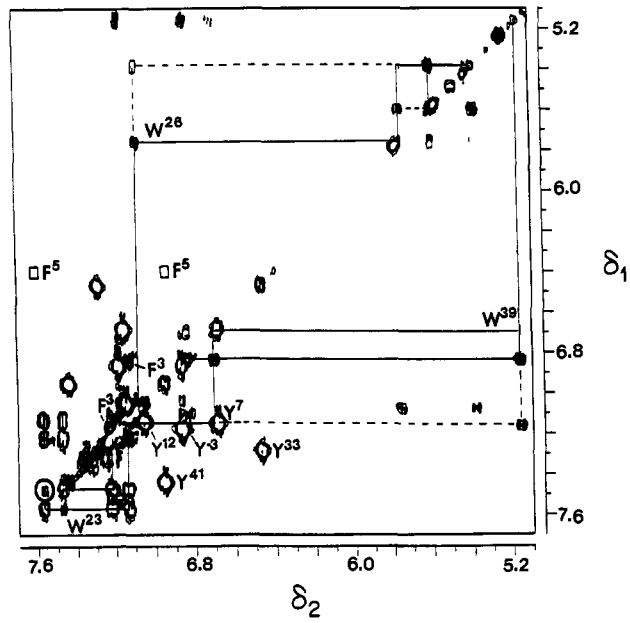

FIGURE 2: 500-MHz ^{1}H NMR RELAYED-COSY spectrum of PDC-109/b. Cross-peaks involving Tyr and Phe aromatic resonances are labeled. Boxes indicate positions of Phe^5 cross-peaks observed in TOCSY and NOESY spectra (not shown). Direct (three-bond) Trp connectivities are indicated by the solid line; nonoverlapping relayed (four bond) connectivities are indicated by dashed lines. A double relay (five bond) cross-peak connecting Trp^{23} H4 and H7 is shown circled. The spectrum was recorded at 27 °C on a sample dissolved in [$^2H_6$]acetone/D_2O (1:9 v/v) to a protein concentration of ~2 mM containing 50 mM NaCl, pH* ~6.5.

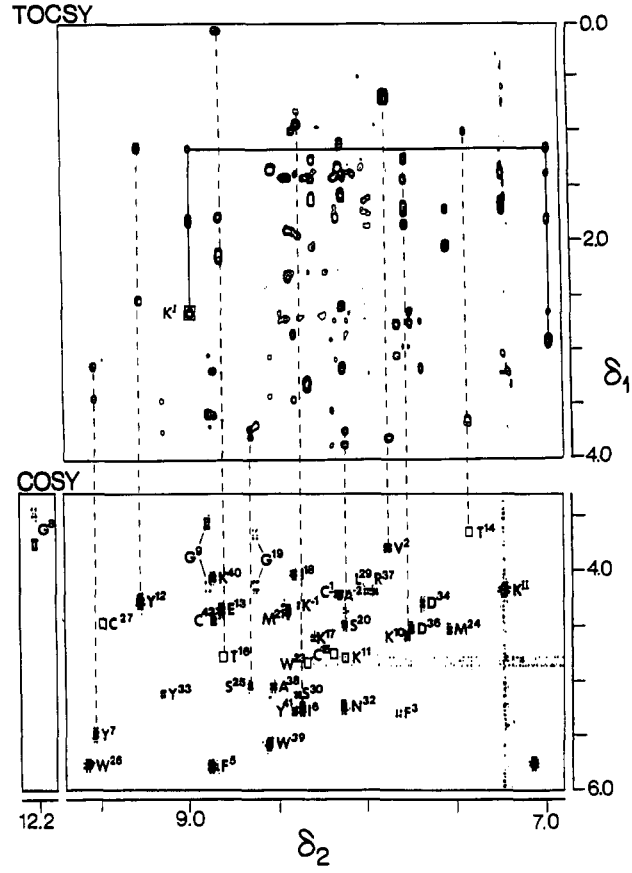

FIGURE 3: 500-MHz ^{1}H NMR combined COSY/TOCSY spectra of PDC-109/b: amide–aliphatic connectivites. NH–CH$^\alpha$ (fingerprint) cross-peaks are labeled in the COSY spectrum (lower panel), including those observed in other COSY and TOCSY spectra (τ_m = 84 ms) recorded at higher temperatures (boxed). Selected NH–aliphatic connectivities are indicated by dashed lines linking the COSY and TOCSY cross-peaks. The Lys-I spin system (apparently a buried Lys side chain) is traced by the solid line in the TOCSY spectrum. Recorded at pH* ~6.8 in 1:9 (v/v) [$^2H_6$]acetone/^1H_2O; other conditions are the same as for Figure 2.

cross-peaks connecting the CH$^{\beta,\beta'}$ resonances to the degenerate aromatic H2 and H6 (H2,6) signals. The Trp ring spin systems were identified and connected to the backbone. The resonances from the H2,6 and H3,5 protons of one of the Phe rings (Phe^5) are broad; therefore, it was only possible to connect one Phe ring (Phe^3) to the CH$^{\beta,\beta'}$ resonances directly. Identification of the second Phe (Phe^5) backbone spin system was obtained in the course of the sequential assignment process.

The identification of resonances stemming from the unique Gly, Thr, Ala, Leu, Ile, Val, and Arg spin systems was achieved by analysis of scalar correlated experiments recorded in H_2O and D_2O (Wüthrich, 1986; Chazin et al., 1988). Since Leu, Val, and Arg residues each occur singly in PDC-109/b (Figure 1), the corresponding spin systems were unambiguously assigned to positions 29, 2, and 37, respectively. The NH and CH$^\alpha$ resonances of Leu^{29} are broad and difficult to discern at 27 °C. Weak cross-peaks are observed between the Leu^{29} NH and CH$_3^{\delta,\delta'}$ groups in the TOCSY spectrum recorded at this temperature (Figure 3); cross-peaks between the CH$^\alpha$ and CH$_3^{\delta,\delta'}$ groups are only observed in a TOCSY spectrum recorded at 45 °C (not shown). The short (AMX in D_2O) and remaining long side chain spin systems were distinguished by the absence or presence, respectively, of cross-peaks due to long-range coupling past the CH$^{\beta,\beta'}$ resonances observed in TOCSY spectra. Three of the four Ser spin systems were tentatively distinguished from the remaining unidentified AMX spin systems on the basis of CH$^\beta$ resonances shifted low-field of 3.6 ppm. These identifications were subsequently confirmed by sequential analysis. The identified spin systems (and their sequentially assigned positions, see below) are indicated in Figure 3, which shows the amide–aliphatic regions of combined COSY/TOCSY spectra.

After spin systems were identified and classified, sequential assignment analysis was undertaken. The most numerous sequential connectivities found in PDC-109/b are those linking the CH$^\alpha$ resonance of residue i with the NH resonance of residue i+1 ($d_{\alpha N}$ connectivities). Short (≅2.5 Å) CH$^\alpha$–NH interproton distances are found primarily in regions of extended strand or β-sheet (Wüthrich, 1986). Figure 4 shows part of the fingerprint region of a NOESY spectrum (τ_m = 200 ms) of PDC-109/b. Several pathways of CH$^\alpha$–NH connectivities are traced out. Cys^1, Val^2, and Phe^3 are linked by strong CH$^\alpha$–NH cross-peaks. An uninterrupted stretch of medium-to-strong CH$^\alpha$–NH NOEs connects Pro^4 through Gly^8. Also traced out explicitly in Figure 4 are sequential connectivities linking Ala^{38} through Cys^{42} and Met^{24} through Cys^{27}. Many of the chemical shift degeneracies apparent in this spectrum were resolved in spectra recorded at 35 °C (not shown). In addition to the $d_{\alpha N}$ connectivities, NOESY cross-peaks between the CH$^{\beta,\beta'}$ resonances of residue i and the NH resonance of residue i+1 ($d_{\beta N}$ connectivities) and between the NH resonances of adjacent residues (d_{NN} connectivities) were used to obtain sequential assignments. Figure 5 summarizes the observed sequential NOE connectivities. Two regions of the sequence—Lys^{-1} through Asp^{-4} and Ser^{20} through Trp^{23}—yield a sparse number of sequential NOEs. This may reflect irregular structural features and/or local conformational flexibility in these regions. The data afforded unambiguous sequential assignments for most residues, as reported in Table I. The assignments of Asp^{-4} and Lys^{-1} remain tentative due to the limited connectivity data. The case of Lys^{35} is more

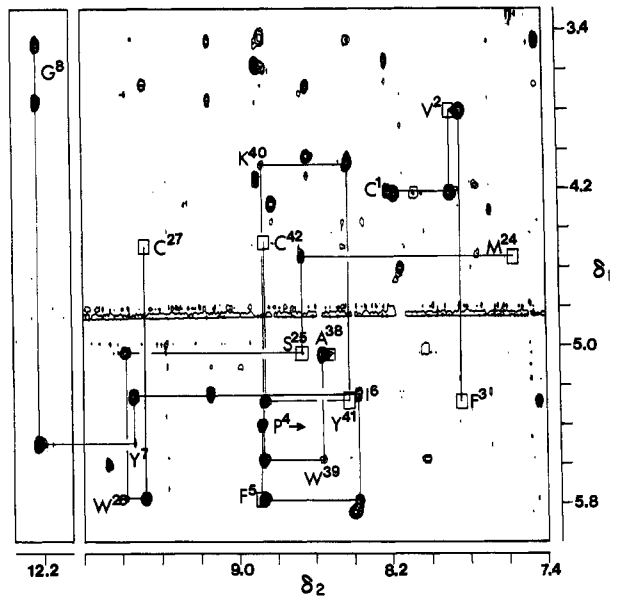

FIGURE 4: 500-MHz ^{1}H NMR phase-sensitive NOESY spectrum of PDC-109/b: fingerprint region. Sequential pathways are traced by solid lines. Intraresidue CH$^\alpha$–NH cross-peaks (corresponding to COSY connectivities) are labeled; those that are too weak to be observed in the NOESY spectrum, but readily observed in COSY spectra, are boxed. Also labeled is the sequential CH$^\alpha$–NH cross-peak that identifies the position of the CH$^\alpha$ resonance of Pro^4 (indicated by the solid arrow). Recorded with τ_m = 200 ms; other conditions are the same as for Figure 3.

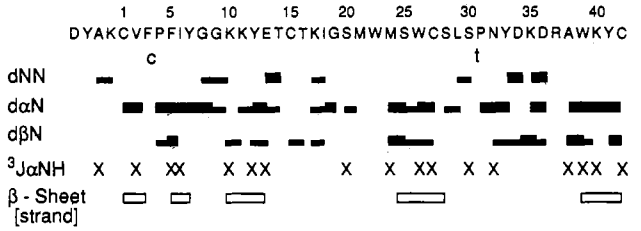

FIGURE 5: Sequential (i, i+1) ^{1}H NMR NOE connectivities of PDC-109/b. The sequence of PDC-109/b (one-letter amino acid code) is listed on top, taking into account the deletion indicated in Figure 1. Strong NOEs are distinguished by thick solid horizontal lines; medium and weak NOEs are denoted by thin solid lines. The t and c labels directly below the sequence denote CH$^\alpha$(i)–Pro CH$^\delta$(i+1) and CH$^\alpha$(i)–Pro CH$^\alpha$(i+1) connectivities characteristic of *trans*- and *cis*-prolines, respectively. $^3J_{\alpha NH}$ values > 8 Hz are indicated by X. Regions of extended conformation, typical of extended strand or β-sheet, are denoted by open bars.

complex—two spin systems (Lys-I and Lys-II) yield connectivities that suggest that either could be Lys^{35}. Lys-I yields a strong NH–NH NOESY cross-peak to Asp^{36}, while Lys-II yields CH$^\alpha$–NH and CH$^\beta$–NH connectivities to Asp^{36}. In addition, Lys-I appears to be buried, perhaps forming an internal salt bridge, as we have been able to observe and assign the corresponding NH$_3^\zeta$ resonance (Figure 3). The difficulty in assigning Lys^{35} may reflect local chemical or conformational heterogencity. Data involving the three ambiguously assigned residues were excluded from the structural analysis.

We have obtained stereospecific assignments for the Val^2 CH$_3^{\gamma,\gamma'}$ groups and the CH$^{\beta,\beta'}$ protons of Phe^3, Pro^4, Pro^{31}, Tyr^{33}, Asp^{34}, Asp^{36}, and Tyr^{41}, as reported in Table I. Stereospecific assignments for the Pro residues are based on CH$^\alpha$–CH$^{\beta,\beta'}$ NOE intensities, since CH$^\alpha$ is always closer to CH$^{\beta_3}$ than it is to CH$^{\beta_2}$. The magnitude of the Val^2 $^3J_{\alpha\beta}$ is > 10 Hz, indicating that the predominant rotamer about the C$^\alpha$–C$^\beta$ bond is g$^+$ (Zuiderweg et al., 1985). A strong Val^2 intraresidue NOE between the NH and the CH$_3^\gamma$ protons at

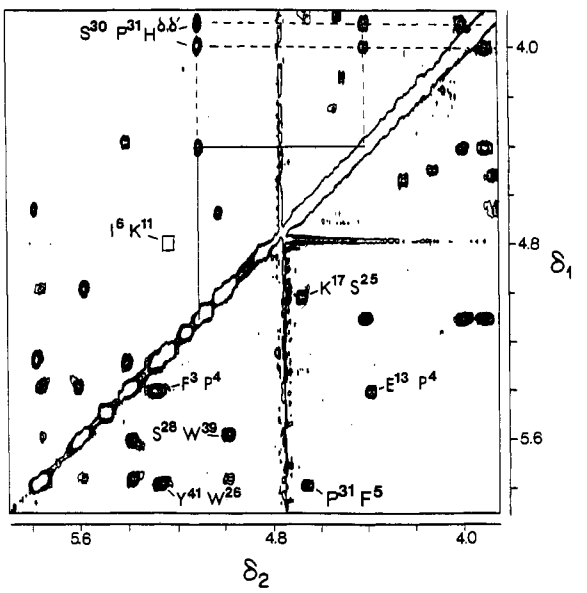

FIGURE 6: 500-MHz ^{1}H NMR phase-sensitive NOESY spectrum of PDC-109/b: low-field aliphatic connectivities. All labels refer to the CH$^\alpha$–CH$^\alpha$ connectivity between Res^2 (δ_2) and Res^1 (δ_1), unless otherwise specified. The box denotes a cross-peak observed at 35 °C. The Ser^{30} intraresidue CH$^\alpha$–CH$^\beta$ connectivity is indicated by solid lines, while sequential Ser^{30}–Pro^{31} connectivities are marked with dashed lines. Recorded with τ_m = 300 ms; other conditions are the same as for Figure 2.

0.69 ppm identifies the latter as CH$_3^{\gamma_1}$. [Nomenclature of prochiral amino acid protons follows the recommendations of the IUPAC–IUB Commission (1970).] For the remaining residues, stereospecific assignments of CH$^{\beta,\beta'}$ resonances were derived according to the following criteria (Wagner et al., 1987). If both $^3J_{\alpha\beta}$ and $^3J_{\alpha\beta'}$ are <5 Hz, then the predominant χ^1 rotamer angle is ~60°. One strong and one weak intraresidue NH–CH$^\beta$ NOE identifies CH$^{\beta_3}$ and CH$^{\beta_2}$, respectively. If $^3J_{\alpha\beta}$ is <5 Hz and $^3J_{\alpha\beta'}$ is >10 Hz, then χ^1 is near 180° or −60°. Strong intraresidue NH–CH$^\beta$ and NH–CH$^{\beta'}$ NOEs then indicate that χ^1 ~ 180°, CH$^\beta$ = CH$^{\beta_2}$, and CH$^{\beta'}$ = CH$^{\beta_3}$. Otherwise, a strong NH–CH$^{\beta'}$ NOE and a weak NH–CH$^\beta$ NOE identifies CH$^{\beta_2}$ and CH$^{\beta_3}$, respectively, and indicate χ^1 ~ −60°.

Secondary and Local Structural Features. Valuable information regarding secondary structure is afforded by the sequential assignment procedure in concert with consideration of measured $^3J_{\alpha NH}$ values (Figure 5). Strong $d_{\alpha N}$ connectivities in conjunction with weak d_{NN} connectivities, weak intraresidue CH$^\alpha$–NH NOEs, and large (>8 Hz) $^3J_{\alpha NH}$ values are diagnostic of β-sheet or extended strand. The data summarized in Figure 5 indicate that residues 1–7 (exclusive of Pro^4, which, as discussed below, is cis), 10–13, 25–28, and 39–42 are characterized by these types of structural features. Conclusive evidence for a β-sheet topology was obtained by identifying NOESY cross-peaks arising from interstrand proximities. Figure 6 shows a portion of a NOESY spectrum recorded in D_2O. We note the strong CH$^\alpha$–CH$^\alpha$ NOESY cross-peaks between the CH$^\alpha$ protons of Ser^{28} and Trp^{39}, and between Tyr^{41} and Trp^{26}. The combined data (including additional interstrand NOESY connectivities) clearly define an antiparallel β-sheet involving residues 25–28 and 39–42 (Figure 7). It is not clear whether or not Leu^{29} and Ala^{38} form part of the β-sheet, as the NOE connectivity data are sparse in this region, in agreement with the previously noted weakness of the Leu^{29} backbone proton resonances. In addition, the Ser^{28} NH resonance has not been located; therefore, the expected $d_{\alpha N}$ connectivity between Cys^{27} and Ser^{28} is not observed. This

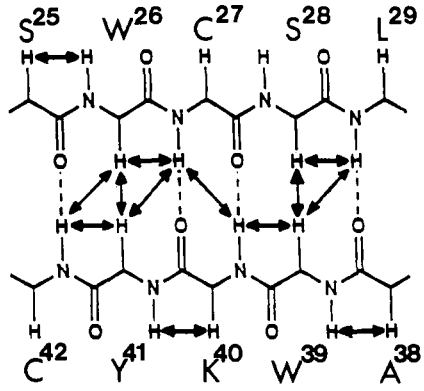

FIGURE 7: Antiparallel β-sheet within PDC-109/b. The arrows indicate sequential and long-range NOEs which define the conformation. Dashed lines indicate likely hydrogen bonds.

β-sheet topology is shifted one residue from that found in the model produced by Holland et al. (1987) and in our kringle-based model, which position Ser^{28} directly opposite to Lys^{40} instead of Trp^{39}. Also observed (Figure 6) are NOEs connecting the CH^{α} proton resonances of residue pairs Ile^6/Lys^{11} and Pro^4/Glu^{13}. These connectivities define antiparallel interstrand contacts; however, they are not sufficient to unambiguously define a standard antiparallel β-sheet. A type I or I′ β-turn involving Tyr^7 through Lys^{10} is indicated by NH–NH NOEs between residue pairs Gly^8/Gly^9, Gly^9/Lys^{10}, and Tyr^7/Lys^{10}. Furthermore, a turn in the vicinity of Tyr^{33} through Asp^{36} is suggested by an Overhauser connectivity between Tyr^{33} H^{α} and Asp^{36} H^{β_2}. Strong NOESY cross-peaks between the CH^{α} protons of residue pairs Lys^{17}/Ser^{25} and Phe^5/Pro^{31} are also observed.

The NOESY experiments provide additional local conformational information. The strong cross-peak between the CH^{α} resonances of Phe^3 and Pro^4 (Figure 6) demonstrates that Pro^4 assumes the cis conformation (Wüthrich, 1986). This section of the sequence is highly conserved in all type II domains (Petersen et al., 1989), suggesting that a *cis*-Pro^4 may represent a common structural feature. The strong cross-peaks between the CH^{α} resonance of Ser^{30} and the $CH^{\delta,\delta'}$ resonances of Pro^{31} demonstrate that Pro^{31} adopts the more usual trans conformation. No conclusive evidence of cis/trans isomerization was found for either proline.

A number of χ^1 dihedral angle constraints have also been obtained. Concerning the residues with stereospecifically assigned $CH^{\beta,\beta'}$ resonances, we find $\chi^1 \sim -60°$ for Phe^3 and $\chi^1 \cong 180°$ for Tyr^{33}, Asp^{34}, Asp^{36}, and Tyr^{41}. In addition, the χ^1 angles of both Ser^{25} and Ser^{30} are $\sim 60°$, since for both residues $^3J_{\alpha\beta}$ and $^3J_{\alpha\beta'}$ are <5 Hz and all CH^{α}–$CH^{\beta,\beta'}$ NOEs are strong. For an additional nine residues, $\chi^1 \sim 60°$ can be excluded on the basis of the simultaneous occurrence of $^3J_{\alpha\beta} > 10$ Hz and $^3J_{\alpha\beta'} < 5$ Hz.

Molecular Modeling. A number of long-range interactions (i.e., between amino acids separated by at least four sequential residues) are observed in the NOESY spectra of PDC-109/b. Figure 8 shows the aromatic region of a NOESY spectrum recorded in D_2O, which reveals aromatic–aromatic and CH^{α}–aromatic long-range connectivities (compare Figures 8 and 2). NOESY cross-peaks that delineate proximities between the rings of Tyr^7, Trp^{26}, Tyr^{33}, and Trp^{39} are explicitly labeled. The proximity between the rings of Tyr^7 and Tyr^{33} is particularly noteworthy, as these rings are over 20 Å apart in our kringle-based model and are located on opposite sides of the protein in the similar model proposed by Holland et al. (1987). Additional connectivities are observed, including NOEs between the Trp^{26} CH^{α} and the Trp^{39} H4 and Tyr^{41}

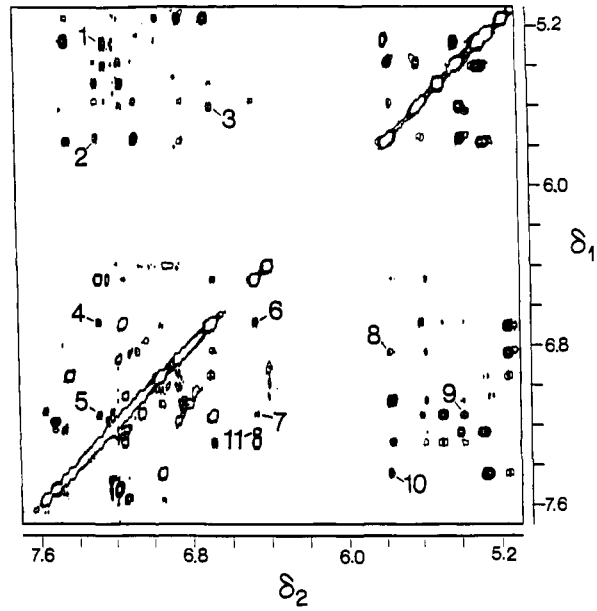

FIGURE 8: 500-MHz ^1H NMR phase-sensitive NOESY spectrum of PDC-109/b: aromatic region. Selected cross-peaks are labeled with Arabic numerals, with the following assigned δ_2–δ_1 correlations: 1, $F^3(H2,6)$–$F^3(CH^{\alpha})$; 2, $Y^{33}(H2,6)$–$W^{26}(H5)$; 3, $Y^7(H3,5)$–$W^{26}(H7)$; 4, $Y^{33}(H2,6)$–$Y^7(H3,5)$; 5, $Y^{33}(H2,6)$–$Y^7(H2,6)$; 6, $Y^{33}(H3,5)$–$Y^7(H3,5)$; 7, $Y^{33}(H3,5)$–$Y^7(H2,6)$; 8, $W^{26}(CH^{\alpha})$–$W^{39}(H^4)$; 9, $W^{26}(H6)$–$Y^7(H2,6)$; 10, $W^{26}(CH^{\alpha})$–$Y^{41}(H2,6)$; 11, $Y^{33}(H3,5)$–$W^{39}(H2)$. Experimental conditions are the same as for Figure 6.

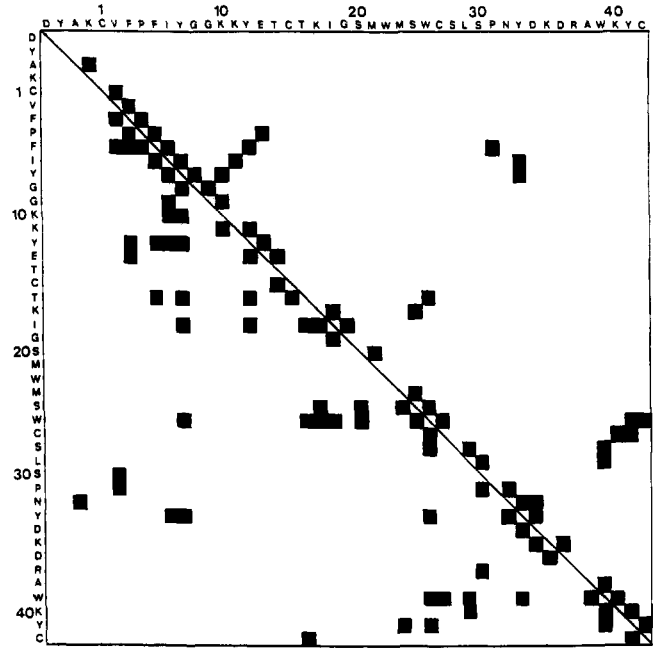

FIGURE 9: Interresidue NOE map for PDC-109/b. The off-diagonal squares in the upper right indicate NOEs observed between backbone protons (NH and CH^{α}); those in the lower left indicate NOEs involving at least one side-chain proton. Amino acid sequences are given along each axis.

H2,6 ring protons. Long-range connectivities are also present in other spectral regions. (Some of these were discussed in the previous section in the context of secondary structural features.) The observed NOEs are summarized in Figure 9, which displays an interresidue NOE connectivity map for PDC-109/b.

An analysis of the NOESY spectra afforded 81 long-range, 19 medium-range, and 62 sequential and intraresidue distance constraints. These data, along with constraints defining the expected β-sheet H-bonds between the polar backbone groups

Table I: ^{1}H NMR Resonance Assignments of PDC-109/b

residue[b]	chemical shift (ppm)[a] NH	CH$^\alpha$	CH$^\beta$	other
Asp^{-4}*		4.17	2.77	
Tyr^{-3}		4.70	3.14, 3.01	H2,6 7.18; H3,5 6.87
Ala^{-1}	8.16	4.25	1.34	
Lys^{-1}*	8.39	4.33	2.29, 1.39	CH$^\gamma$ 2.00
Cys^1	8.17	4.22	3.09, 2.53	
Val^2	7.89	3.82	1.36	CH$_3^{\gamma 2}$ 0.74; CH$_3^{\gamma 1}$ 0.69
Phe^3	7.84	5.29	3.06, 2.77 (st)[c]	H2,6 7.23; H3,5 7.13; H4 6.85
Pro^4		5.40	2.37, 1.92 (st)	CH$^\gamma$ 2.61, 2.06; CH$^\delta$ 3.83, 3.65
Phe^5	8.89	5.79	3.70, 3.05	H2,6 7.62; H3,5 6.94; H4 6.41
Ile^6	8.37	5.25	1.96	CH$_3^\gamma$ 0.96; CH$^\gamma$ 1.69, 1.24; CH$_3^\delta$ 0.84
Tyr^7	9.55	5.49	3.46, 3.15	H2,6 7.15; H3,5 6.69
Gly^8	12.22	3.77, 3.48		
Gly^9	8.90	4.17, 3.59		
Lys^{10}	7.78	4.63	1.86, 1.75	CH$^\gamma$ 1.45, 1.26; CH$^\delta$ 1.59; CH$^\epsilon$ 2.97
Lys^{11}	8.13	4.78	1.62, 1.58	CH$^\gamma$ 1.14; CH$^\delta$ 1.41; CH$^\epsilon$ 2.91
Tyr^{12}	9.28	4.30	2.54, 1.16	H2,6 7.05; H3,5 7.14
Glu^{13}	8.83	4.38	1.79, 2.19	CH$^\gamma$ 2.11*
Thr^{14}	7.43	3.67	4.21	CH$_3^\gamma$ 1.02
Cys^{15}	8.20	4.75	3.56, 2.90	
Thr^{16}	8.82	4.78	3.65	CH$_3^\gamma$ 0.09
Lys^{17}	8.31	4.67	2.06, 1.63	CH$^\gamma$ 1.43, 1.27
Ile^{18}	8.41	4.05	1.93	CH$_3^\gamma$ 1.01; CH$^\gamma$ 1.64, 1.39; CH$_3^\delta$ 0.94
Gly^{19}	8.63	4.16, 3.69		
Ser^{20}	8.14	4.50	3.88, 3.76	
Met^{21}	8.45	4.39	1.92, 1.40	CH$^\gamma$ 2.19, 2.07; CH$_3^\epsilon$ 1.72*
Trp^{23}	8.34	4.83	3.34, 3.28	NH1 10.19; H2 7.19; H4 7.57 H5 7.13; H6 7.22; H7 7.48
Met^{24}	7.55	4.55	2.06, 1.72	CH$^\gamma$ 2.38; CH$_3^\epsilon$ 1.93*
Ser^{25}	8.66	5.02	3.80, 3.74	
Trp^{26}	9.58	5.78	3.68, 2.97	NH1 9.66; H2 7.10; H4 7.08 H5 5.77; H6 5.39; H7 5.60
Cys^{27}	9.49	4.48	2.28, 2.01	
Ser^{28}		4.99	3.45, 3.09	
Leu^{29}	8.02	4.20*	1.74, 1.34	CH$^\gamma$ 1.25; CH$_3^\delta$ 0.70, 0.53
Ser^{30}	8.40	5.12	4.41, 3.64	
Pro^{31}		4.66	2.73, 2.02 (st)	CH$^\gamma$ 2.22, 2.11; CH$^\delta$ 4.00, 3.91
Asn^{32}	8.15	5.24	3.19, 2.60	NH$^\delta$ 7.70, 7.19
Tyr^{33}	9.15	5.09	3.76, 3.48 (st)	H2,6 7.29; H3,5 6.46
Asp^{34}	7.69	4.32	3.18, 2.74 (st)	
Lys^{35}[d]				
Lys-I	8.99	2.66	1.85, 1.81	CH$^\gamma$ 1.17; CH$^\delta$ 1.40, 1.15; CH$^\epsilon$ 2.94, 2.90; NH$_3^\zeta$ 6.98
Lys-II	7.24	4.20	1.72, 1.63	CH$^\gamma$ 1.38; CH$^\delta$ 1.68; CH$^\epsilon$ 3.03
Asp^{36}	7.75	4.54	2.75, 2.64 (st)	
Arg^{37}	7.98	4.21	1.88, 1.76	CH$^\gamma$ 1.62; CH$^\delta$ 3.23, 3.08; NH$^\epsilon$ 5.84
Ala^{38}	8.54	5.01	1.35	
Trp^{39}	8.57	5.58	3.44, 3.06	NH1 9.81; H2 7.24; H4 6.84 H5 5.16; H6 6.71; H7 7.17
Lys^{40}	8.88	4.09	1.84, 1.60	CH$^\gamma$ 1.14; CH$^\delta$ 1.41, 1.35
Tyr^{41}	8.43	5.28	2.86, 3.46 (st)	H2,6 7.44; H3,5 6.95
Cys^{42}	8.87	4.47	3.60, 3.19	

[a] The chemical shift values (±0.01 ppm) were derived from 2D ^{1}H NMR spectra recorded at pH* 6.8 (H_2O) or 6.5 (D_2O), 27 °C. [b] An asterisk indicates tentative assignment. [c] st denotes stereospecific assignment of H$^\beta$ protons, with the first shift value for H$^{\beta 3}$ and the second for H$^{\beta 2}$ (IUPAC–IUB Commission, 1970). [d] Lys-I and Lys-II give rise to NOESY connectivities that suggest that either could be assigned to Lys^{35}.

of Cys^{27} and Lys^{40} (two constraints), the disulfide bridges (Figure 1), and the estimated ranges of ϕ and χ^1 angles, were used to generate 16 initial structures with the distance geometry algorithm DISGEO (Havel & Wüthrich, 1984; Havel, 1986). Although the backbone amide groups of Cys^{27} and Lys^{40} do not exhibit retarded H-exchange with the solvent, the observed NOEs (Figure 7) and the low-field chemical shifts of these NH resonances, $\Delta\delta$ = 1.18 and 0.47 ppm, respectively, relative to typical random coil positions (Bundi & Wüthrich, 1979), strongly suggest that they are involved in interstrand H-bonding.

The DISGEO computations successfully completed substructure embedding, complete structure embedding, and minimization of chirality and constraint error functions (Havel et al., 1983). Two classes of PDC-109/b conformers were produced: a major one (class A) composed of ten structures and a minor one (class B) composed of six structures. The global backbone fold of the minor class is a pseudo mirror image of the major class. Pairwise RMSDs between backbone atoms of the class A and class B conformers are ~7.5–8.2 Å; upon reflection of the coordinates of either class, this value drops to ~2.0–2.5 Å. The average pairwise RMSD between the backbone atoms of the class A structures is 1.70 Å; the largest RMSD between any two class A structures is 2.02 Å, and the smallest is 1.18 Å. The corresponding RMSD values for the class B structures are 1.78, 1.99, and 1.54 Å, respectively. These results are similar to those obtained by Carr et al. (1990), who found major and minor mirror image conformers (with average pairwise backbone atom RMSDs of 2.0 and 2.5 Å, respectively) for the solution structure of a zinc-finger domain. Such a result is indicative of a low-resolution structure determination. The average RMSDs between the class A and class B structures and the kringle-based model are 7.65 and 9.48 Å, respectively, demonstrating that the

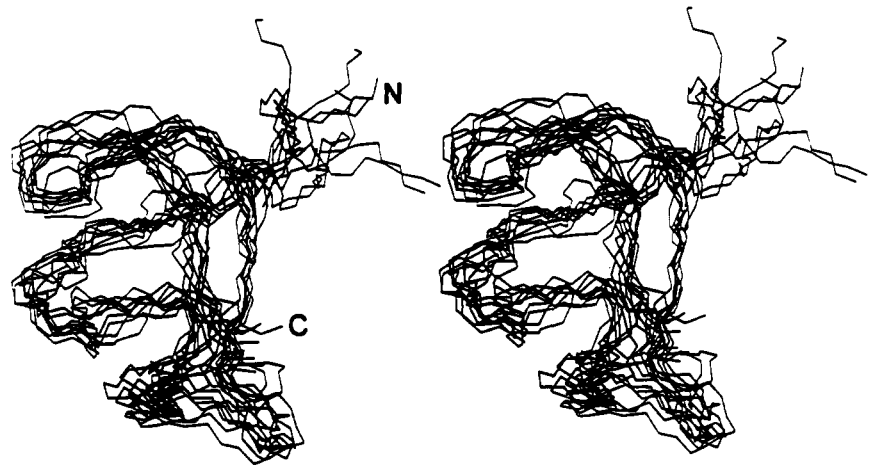

FIGURE 10: Stereoview of PDC-109/b: preliminary backbone structures. The figure shows superimposed models A1–A10 (N, C^{α}, and C′ atoms). The amino and carboxyl termini are indicated by N and C, respectively.

kringle-based model is somewhat inconsistent with the NMR information.

The DISGEO structures were optimized by restrained energy minimization and restrained molecular dynamics, as described under Materials and Methods. For the structures produced by restrained molecular dynamics (prior to final restrained energy minimization), the average backbone atom RMSD between the class A conformers is 2.21 Å, and between the class B conformers, it is 2.53 Å. Class A structures have an average of 5.4 ± 1.4 constraint violations > 1.0 Å, while for class B structures the corresponding value is 7.5 ± 1.1 violations. The increase in the average RMSD between structures stems from conformational relaxation during the restrained molecular dynamics runs. This may reflect the recent observation (Metzler et al., 1989) that metrix matrix distance geometry programs do not explore a completely random sample of the conformation space consistent with the constraints, especially in regions where the number of constraints is limited.

Each structure was subjected to a final 1200-step restrained energy minimization using very mild NOE force constants of 10 kcal mol^{-1} $Å^{-2}$. The resulting structures will henceforth be referred to as A1–A10 and B1–B6 for the class A and class B structures, respectively. The average backbone atom RMSD value between structures is 2.12 Å for A1–A10 and 2.51 Å for B1–B6. The corresponding all-atom values are 3.63 and 3.98 Å, respectively. If the N-terminal peptide tail consisting of residues −4 to −1 is deleted, the average backbone atom RMSD between the class A conformers becomes 1.84 Å. Using a rather mild NOE force constant results in an increase in the number of constraint violations for all structures and represents a compromise between satisfying distance constraints and avoiding highly strained conformations. For the A1–A10 structures, the average number of constraint violations > 1.0 Å is 18.6 ± 2.9, while for B1–B6 the value is 25.5 ± 4.5 violations. These violations are reasonable in view of the overall conformational flexibility of PDC-109/b. The use of a mild force constant helps to compensate for underestimated average distances, which can be caused by spin diffusion (Olejniczak et al., 1986; Neuhaus & Williamson, 1989) and/or internal dynamics (Kessler et al., 1988; Neuhaus & Williamson, 1989; Constantine et al. 1990). The relative total internal energies and relative constraint energies for structures A1–A10 and B1–B6 are given in Table II. The A1–A10 structures have total internal energies that are, on average, 484 kcal mol^{-1} more stable than the B1–B6 structures. The constraint energies of the A1–A10 structures are, on average, 178 kcal mol^{-1} lower than those of B1–B6. These data, along with the fact that the A1–A10 conformers have lower RMSD values among themselves and to the kringle-based model than do the B1–B6 conformers, suggest that the global fold of the class A structures is closer to the correct conformation. Nevertheless, we cannot unambiguously rule out either class. Therefore, the remainder of this discussion will focus on general structural features common to both classes.

Table II: Relative Total and Constraint Energies of PDC-109/b Models

structure	relative energy (kcal/mol) total[a]	constraint[b]
A1	355	93
A2	93	0
A3	333	76
A4	0	18
A5	113	35
A6	8	5
A7	366	131
A8	477	131
A9	402	139
A10	262	115
B1	990	426
B2	649	192
B3	849	299
B4	725	225
B5	596	230
B6	542	139

[a] Relative to the total internal energy of structure A4, which has the lowest total internal energy of any structure. [b] Relative to the constraint energy of structure A2, which has the lowest constraint energy of any structure.

Global Fold. The current level of resolution of the NMR-based models is adequate to describe the approximate global fold of PDC-109/b and certain other overall structural features. Figure 10 shows a stereoview of the superimposed main chains of the 10 class A structures. All class A structures exhibit the same overall folding pattern; the global fold of the class B structures is the mirror image of that shown in Figure 10.

In order to examine the effect of the two H-bond constraints, they were removed, and structure A6 (Table II) was subjected to 1200 additional steps of restrained ABNR energy minimization. The backbone atom RMSD between the resulting structure and model A6 is 0.22 Å, indicating that the H-bond constraints do not significantly influence the global fold. The β-sheet involving residues 25–28 and 39–42 lies roughly perpendicular to the antiparallel strands involving residues 1–3, 5–7, and 10–13. The N-terminal polypeptide tail and the loop

containing Trp[23] (Figure 10, bottom) are poorly defined, a likely reflection of high mobility in these regions.

Several observations regarding side-chain conformations can be made. The Trp[23] side chain extends into the solvent, consistent with the observation that its ring proton resonances are sharp and their chemical shifts (Table I) are extremely close to typical random coil values (Bundi & Wüthrich, 1979). It is also noteworthy that in the NMR-derived models the disulfide bridges tend to be near the surface of the protein. This differs from the disulfide arrangement found in kringle domains, since in the latter the corresponding disulfides are buried (Park & Tulinsky, 1986).

Key residues in the hydrophobic core include Tyr[12], Phe[5], and Thr[16], which are highly conserved in type II domain homologues (Petersen et al., 1989). The CH_3^{γ} resonance of Thr[16] is shifted high field to 0.09 ppm (Table I), indicating ring current shielding. NOEs (not shown) are observed between this methyl group and the Tyr[7] H3,5, Phe[5] H4, Tyr[12] H2,6, Trp[26] NH1, Trp[26] H2, and Trp[26] H7 protons, suggesting that Thr[16] plays an important structural role. Indeed, this residue is found buried and interacting closely with aromatic residues in all NMR-based models. The side chains of Tyr[7], Trp[26], and Tyr[33] form an aromatic cluster. These three aromatic rings are also highly conserved (Petersen et al., 1989). The proximity of Gly[8] to this cluster explains the extreme low-field chemical shift (12.22 ppm) of its NH resonance. In our low-resolution models, the Tyr[7], Trp[26], and Tyr[33] side chains are located in a shallow depression or groove, situated on the backside of the molecule in the perspective presented by Figure 10. This feature, which is roughly similar in all structures, presents a solvent-exposed hydrophobic surface and may define a collagen-binding site. It is likely that this region contributes to the aggregation tendency of PDC-109/b in aqueous solution. This is consistent with the effect of acetone, since disaggregation probably results from the disruption of intermolecular hydrophobic interactions.

Conclusions

We have derived sequential resonance assignments for the ^{1}H NMR spectrum of PDC-109/b, the isolated second type II domain of PDC-109, including stereospecific identification of several prochiral groups. Qualitative analysis of secondary structural features reveals an antiparallel β-sheet pairing the Ser[25]–Ser[28] and Trp[39]–Cys[42] stretches. A second region of antiparallel strands involving residues 1–3, 5–7, and 10–13 has been identified; however, it may be distorted from typical antiparallel β-sheet conformation. The antiparallel strands lie roughly perpendicular to the antiparallel β-sheet mentioned above. These results concur with the secondary structural features predicted on the basis of proposed homology between type II and kringle domains (Patthy et al. 1984; Holland et al., 1987) and also with secondary structure predicted by the Chou–Fasman routine (Odermatt & Engle, 1989). However, molecular modeling of the tertiary structure of PDC-109/b demonstrates that the overall fold differs from that predicted by homology (Holland et al., 1987). This outcome raises some interesting questions regarding the proposed evolutionary connection between type II and kringle domains (Patthy et al., 1984; Holland et al., 1987; Holland & Blake, 1987), which postulates that these units evolved from a common ancestral protein. The evidence presented here demonstrates that the proposed evolutionary relationships do not afford accurate predictive criteria at the level of tertiary folding. The conservation of structural features is considerably less extensive than previously supposed (Holland & Blake, 1987). The major regions of secondary structure are retained, but their spatial arrangement via the interconnecting turns and loops is not. This is understandable considering the rather large number of insertions, deletions, and mutations separating type II and kringle sequences and the low (~23%) sequence homology.

Our NMR-derived models reveal that PDC-109/b adopts a globular conformation consisting mostly of irregular loops and turns. It exhibits a depression in the molecular surface, structured in part by the exposed edges of partially buried aromatic rings, which may define a ligand-binding site. If this suggestion proves correct, then a rough similarity to kringle domains will have been demonstrated, since it is known that the binding sites of kringles are also defined mostly by aromatic residues (Motta et al., 1987; Ramesh et al., 1987; Tulinsky et al., 1988; Thewes et al., 1980). The binding site hypothesis is consistent with the higher affinity of PDC-109 for gelatin relative to collagen (Manjunath & Sairam, 1987), since hydrophobic sites may be exposed upon collagen melting. In addition, PDC-109/b affinity for phenethylamine and the neopentyl groups of alkyl-Superose has recently been demonstrated (Bányai et al., 1990).

The observed two-line AA'BB' spectra of the ring proton resonances of all five Tyr residues indicate that they are all flipping at a rate that is fast on the chemical shift time scale. This information, along with the observed fast ^{1}H–^{2}H exchange kinetics of all labile protons, further suggests that PDC-109/b is highly flexible. Enhanced flexibility may reflect a general feature of functional significance for protein-binding proteins, as mobility may allow conformational adjustments that optimize intermolecular interactions to occur and thus may facilitate the postulated requirement of a "dynamic matching" of surfaces (Lumry & Gregory, 1986). These considerations may be important for type II domain binding to gelatin, since even native collagen is endowed with high dynamic mobility (Torchia & VanderHart, 1976).

In order to improve the resolution and rigorously establish the correct mirror image of PDC-109/b, both the accuracy and the quantity of the NOE data must be enhanced. Work along these lines, involving the preparation of more concentrated samples and the development of relaxation matrix based refinement protocols, is currently underway. The results of these studies will be reported in due course. This work should be well worth the effort, as knowledge of the conformation of PDC-109/b will serve as a valuable reference for understanding the roles played by type II domains in the context of the mosaic proteins (Patthy, 1985; Holland & Blake, 1987) in which they occur.

Added in Proof

While this paper was in press, we have initiated a new series of structure calculations based on an expanded set (341) of NOESY constraints. The new structures, computed via three different approaches (distance geometry, variable target function, and restrained molecular dynamics), confirm that the models shown in Figure 10 are basically correct, in both overall folding and mirror image characteristics (class A). In the process of analysis of the data we have discovered that 12 NOESY cross-peaks were misidentified. This is the main cause for the constraint violations discussed in the paper. The constraints in question, which affect mainly the relative orientation of the strand containing residues 39–42, are noted in the supplementary material.

Acknowledgments

We thank Prof. A. Tulinsky for providing the prothrombin kringle 1 coordinates and Prof. C. L. Brooks for advice regarding the molecular modeling.

Supplementary Material Available

Two tables of dihedral angle constraints and distance constraints (10 pages). Ordering information is given on any current masthead page.

References

Bânyai, L., Trexler, M., Koncz, S., Gyenes, M., Sipos, G., & Patthy, L. (1990) *Eur. J. Biochem. 193*, 801–806.

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

Bodenhausen, G., Kolger, H., & Ernst, R. R. (1984) *J. Magn. Reson. 58*, 370–388.

Braunschweiler, L., & Ernst, R. R. (1983) *J. Magn. Reson. 53*, 521–528.

Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., & Karplus, M. (1983) *J. Comp. Chem. 4*, 187–217.

Bundi, A., & Wüthrich, K. (1979) *Biopolymers 18*, 285–297.

Carr, M. D., Pastore, A., Gausepohl, H., Frank, R., & Roesch, P. (1990) *Eur. J. Biochem. 188*, 455–461.

Chazin, W. J., & Wright, P. E. (1988) *J. Mol. Biol. 202*, 623–636.

Chazin, W. J., Rance, M., & Wright, P. E. (1988) *J. Mol. Biol. 202*, 603–622.

Collier, I. E., Wilhelm, S. M., Eisen, A. Z., Marmer, B. L., Grant, G. A., Seltzer, J. L., Kronberger, A., He, C., Bauer, E. A., & Goldberg, G. I. (1988) *J. Biol. Chem. 263*, 6579–6587.

Constantine, K. L., DeMarco, A., Madrid, M., Brooks, C. L., & Llinâs, M. (1990) *Biopolymers 30*, 239–256.

DeMarco, A. (1977) *J. Magn. Reson. 26*, 527–528.

Englander, S. W., & Wand, A. J. (1987) *Biochemistry 26*, 5953–5958.

Esch, F. S., Ling, N. C., Böhlen, P., Ying, S. Y., & Guillemin, R. (1983) *Biochem. Biophys. Res. Commun. 113*, 861–867.

Havel, T. F. (1986) *DISGEO*, Quantum Chemistry Program Exchange 507, Indiana University.

Havel, T. F., & Wüthrich, K. (1984) *Bull. Math. Biol. 46*, 673–698.

Havel, T. F., Kuntz, I. D., & Crippen, G. M. (1983) *Bull. Math. Biol. 45*, 665–720.

Holland, S. K., & Blake, C. C. F. (1987) *Biosystems 20*, 181–206.

Holland, S. K., Harlos, K., & Blake, C. C. F. (1987) *EMBO J. 6*, 1875–1880.

Hubbard, R. (1986) *HYDRA: Havard York Drawing Program*, University of York, Heslington, York.

IUPAC–IUB Commission on Biochemical Nomenclature (1970) *J. Mol. Biol. 52*, 1–17.

Kessler, H., Griesinger, C., Lautz, J., Müller, A., van Gunsteren, W. F., & Berendsen, H. J. C. (1988) *J. Am. Chem. Soc. 110*, 3393–3396.

Kornblihtt, A. R., Umezawa, K., Vibe-Pedersen, K., & Baralle, F. E. (1985) *EMBO J. 4*, 1755–1759.

Kumar, A., Ernst, R. R., & Wüthrich, K. (1980) *Biophys. Biochem. Res. Commun. 95*, 1–6.

Lumry, R., & Gregory, R. B. (1986) in *The Fluctuating Enzyme* (Welch, G. R., Ed.) pp 1–190, John Wiley & Sons, New York.

Macura, S., & Ernst, R. R. (1980) *Mol. Phys. 41*, 95–117.

Manjunath, P., & Sairam, M. R. (1987) *Biochem. J. 241*, 658–692.

Manjunath, P., Sairam, M. R., & Uma, J. (1987) *Biosci. Rep. 7*, 231–238.

Manjunath, P., Marcel, Y. L., Uma, J., Seidah, N. G., Chretien, M., & Chapdelaine, A. (1989) *J. Biol. Chem. 264*, 16853–16857.

Marion, D., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967–974.

McMullen, B. A., & Fujikawa, K. (1985) *J. Biol. Chem. 260*, 5328–5341.

Metzler, W. J., Hare, D. R., & Pardi, A. (1989) *Biochemistry 28*, 7045–7052.

Morgan, D. O., Edman, J. C., Standring, D. N., Fried, V. A., Smith, M. C., Roth, R. A., & Rutter, W. J. (1987) *Nature (London) 329*, 301–307.

Motta, A., Laursen, R. A., Llinâs, M., Tulinsky, A., & Park, C. H. (1987) *Biochemistry 26*, 3827–3836.

Neuhaus, D., & Williamson, M. P. (1989) *The Nuclear Overhauser Effect in Structural and Conformational Analysis*, VCH Publishers, New York.

Odermatt, E., & Engle, J. (1989) in *Fibronectin* (Mosher, D. F., Ed.) pp 25–45, Academic Press, New York.

Olejniczak, E. T., Gampe, R. T., & Fesik, S. W. (1986) *J. Magn. Reson. 67*, 28–41.

Park, C. H., & Tulinsky, A. (1986) *Biochemistry 25*, 3977–3982.

Patthy, L. (1985) *Cell 41*, 657–663.

Patthy, L., Trexler, M., Vali, Z., Bânyai, L., & Vâradi, A. (1984) *FEBS Lett. 171*, 131–136.

Petersen, T. E., Thogersen, H. C., Skorstengaard, K., Vibe-Pedersen, K., Sahl, P., Sottrup-Jensen, L., & Magnusson, S. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 137–141.

Petersen, T. E., Skorstengaard, K., & Vibe-Pedersen, K. (1989) in *Fibronectin* (Mosher, D. F., Ed.) pp 1–24, Academic Press, New York.

Ramesh, V., Petros, A. M., Llinâs, M., Tulinsky, A., & Park, C. H. (1987) *J. Mol. Biol. 198*, 481–498.

Rance, M., Sorensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 117*, 479–485.

Seidah, N. G., Manjunath, P., Rochemont, J., Sairman, M. R., & Chretien, M. (1987) *Biochem. J. 243*, 195–203.

Skorstengaard, K., Thorgersen, H. C., & Petersen, T. E. (1984) *Eur. J. Biochem. 140*, 235–243.

Thewes, T., Constantine, K., Byeon, I. L., & Llinâs, M. (1990) *J. Biol. Chem. 265*, 3906–3915.

Torchia, D., & VanderHart, D. (1976) *J. Mol. Biol. 104*, 315–321.

Tulinsky, A., Park, C. H., Mao, B., & Llinâs, M. (1988) *Proteins: Struct., Funct., Genet. 3*, 85–96.

Vuento, M., Salonen, E., Kaskimes, A., & Stenman, V. (1980) *Hoppe-Seyler's Z. Physiol. Chem. B 361*, 1453–1456.

Wagner, G. (1983) *J. Magn. Reson. 55*, 151–156.

Wagner, G., Braun, W., Havel, T. F., Schaumann, T., Go, N., & Wüthrich, K. (1987) *J. Mol. Biol. 196*, 611–639.

Wilhelm, S. M., Collier, I. E., Marmer, B. L., Eisen, A. Z., Grant, G. A., & Goldberg, G. I. (1989) *J. Biol. Chem. 264*, 17213–17221.

Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, John Wiley & Sons, New York.

Wüthrich, K., Billiter, M., & Braun, W. (1983) *J. Mol. Biol. 169*, 949–961.

Yamada, K. M. (1989) in *Fibronectin* (Mosher, D. F., Ed.) pp 47–121, Academic Press, New York.

Zuiderweg, E. R. P., Boelens, R., & Kaptein, R. (1985) *Biopolymers 24*, 601–611.