Journal of Biomolecular NMR, 7 (1996) 173–189
ESCOM

J-Bio NMR 340

^{1}H, ^{15}N and ^{13}C resonance assignments and monomeric structure of the amino-terminal extracellular domain of epithelial cadherin

Michael Overduin[a,*], Kit I. Tong[a], Cyril M. Kay[b] and Mitsuhiko Ikura[a,c,**]

[a]*Division of Molecular and Structural Biology, Ontario Cancer Institute and Department of Medical Biophysics, University of Toronto, 610 University Avenue, Toronto, ON, Canada M5G 2M9*
[b]*Department of Biochemistry and Medical Research Council of Canada, Protein Structure and Function Group, University of Alberta, Edmonton, AB, Canada T6G 2H7*
[c]*Center for Tsukuba Advanced Research Alliance and Institute of Applied Biochemistry, University of Tsukuba, Tsukuba 305, Japan*

Received 19 December 1995
Accepted 15 February 1996

Keywords: Cadherin; Multidimensional NMR; Resonance assignment; Quasi-β-helix; Tertiary structure; Cell adhesion

Summary

E-cadherin is a transmembrane protein that provides Ca^{2+}-dependent cell adhesion to epithelial cells. The large majority of the ^{1}H, ^{15}N, ^{13}C and ^{13}CO resonances of a 146-amino acid polypeptide from epithelial (E-) cadherin have been assigned using multidimensional NMR spectroscopy. The structure of the amino-terminal 100 amino acids, corresponding to the first extracellular repeat of E-cadherin [Overduin et al. (1995) *Science*, **267**, 386–389], has been refined. The monomeric state of this isolated domain is demonstrated by light scattering and sedimentation analysis. Seven β-strands and two short helices were identified by patterns of NOE cross-peaks, vicinal coupling constants and chemical shift indices. A novel structural motif termed a quasi-β-helix found in the crystal structure of a neural (N-) cadherin domain [Shapiro et al. (1995) *Nature*, **374**, 327–337] is characterized in detail for the first time by NMR. Slowly exchanging amides were concentrated in the β-sheet region and quasi-β-helix. The β-barrel fold of the cadherin domain is topologically similar to the immunoglobulin fold. Comparison of this solution structure to the crystallized dimers of the N-terminal pair of E-cadherin domains [Nagar et al. (1996) *Nature*, **380**, 360–364] and of the homologous single domain of N-cadherin reveals a conserved cadherin fold with minor structural differences, which can be accounted for by differences in metal ligation and oligomeric state.

Introduction

Cadherins are Ca^{2+}-dependent cell adhesion molecules (CAMs) that play important roles in the formation and maintenance of solid tissues (Takeichi, 1991). Over 30 members of this superfamily have been identified and are distinguished by the presence of extracellular (cad) repeats containing approximately 110 amino acids. E-cadherin is a tumor invasion suppressor that is lost or inactivated in a variety of human carcinomas (Mareel et al., 1994; Birchmeier, 1995). For example, mutations and allelic deletions of the E-cadherin gene have been found in gastric and endometrial tumor cells (Oda et al., 1994; Risinger et al., 1994). Classic cadherins such as E-cadherin contain five cad repeats, a single transmembrane segment and a cytoplasmic region. The N-terminal cad repeat (cad1) provides the homophilic specificity that directs cadherin molecules of the same type to adhere to one another. The conserved His^{79}-Ala-Val^{81} (HAV) motif and surrounding residues of cad1 domains have been implicated in specific cell adhesion: peptides encompassing the HAV motif inhibit cell association (Blaschuk et al., 1990), and when two nearby residues (Ser^{78} and Ser^{83}) are mutated to the corresponding residues of placental cadhe-

*Present address: Department of Pharmacology, University of Colorado Health Sciences Center, 4200 East Ninth Avenue, Denver, CO 80262, U.S.A.
**To whom correspondence should be addressed.

Abbreviations: cad, extracellular cadherin repeat; CAM, cell adhesion molecule; CSI, chemical shift index; DTT, dithiothreitol; E-cadherin, epithelial cadherin; N-cadherin, neural cadherin; NOE, nuclear Overhauser enhancement; PFG, pulsed field gradient; rmsd, root-mean-square deviation.

rin, the mutant E-cadherin also inherits some of the specificity of P-cadherin (Nose et al., 1990). Ca^{2+} ligation involves the DXNDN (where X is any amino acid) and DAD motifs (Ringwald et al., 1987; Ozawa et al., 1990a) and confers protease resistance to cadherins (Takeichi, 1990). The fifth cad repeat of classic cadherins abuts the cell membrane and is unique in that it contains four conserved cysteine residues involved in cell adhesion (Ozawa et al., 1990b). The highly conserved cytoplasmic domain of the classic cadherins contains binding sites for catenin proteins, which in turn interact with the actin cytoskeleton (Nagafuchi and Takeichi, 1988; Ozawa et al., 1990c).

The solution structure of the cad1 domain of E-cadherin (Ecad1) (Overduin et al., 1995) revealed a fold similar to the immunoglobulin fold, despite the lack of significant amino acid sequence similarity. In the Ecad1 structure the HAV motif is located in the center of a slightly concave and solvent-exposed β-sheet surface, where it could mediate adhesive interactions. Four conserved sequences (Pro^{10}-Glu-Asn-Glu^{13}, Leu^{66}-Asp-Arg-Glu^{69}, Asp^{100}-Gln-Asn-Asp^{103} and Asp^{134}-Ala-Asp^{136}) were predicted to form a Ca^{2+} binding pocket between the first and second cad domains of E-cadherin. Amino acid sequence analysis suggests that similar pockets exist between each tandem pair of cad domains. The structure of the homologous domain (NCD1/Ncad1) of N-cadherin has been determined by X-ray crystallography (Shapiro et al., 1995a). Three different crystal forms of Ncad1 were solved in the presence of Yb^{3+} or UO_2^{2+} and revealed two separate dimer interactions that were used to construct a zipper model of cell–cell adhesion. In this model the parallel Ncad1 dimer (which involved a symmetric β-strand crossover) suggested how two cadherins from the same cell surface might associate, while the weaker antiparallel dimer interaction revealed a potential cell–cell adhesive interaction. Recently a crystal structure of the two N-terminal domains of E-cadherin (Ecad12) bound to Ca^{2+} was solved (Nagar et al., 1996). Three Ca^{2+} ions are bound in the linker between the domains and serve to linearize and rigidify the molecule. A Ca^{2+}-induced dimer was found in the Ecad12 structure that differs significantly from the Ncad1 parallel dimer. The obligate role of Ca^{2+} in cell adhesion was attributed to its ability to help form mechanically stable cadherin dimers that would be expected to firmly bridge the cytoskeletons of adjacent cells.

In this paper, the backbone and side-chain resonance assignments of a 146-residue E-cadherin polypeptide in the presence of Ca^{2+} are presented. The structure was elucidated using NOE, amide exchange, coupling constant and chemical shift index (CSI) information. Additional NOEs and stereospecific assignments have yielded a higher resolution structure than obtained previously (Overduin et al., 1995). Dynamic light scattering and centrifugation sedimentation equilibrium data are presented that confirm the monomeric state of Ecad1. The Ecad1 structure is compared to crystal structures of E-cadherin (Nagar et al., 1996) and N-cadherin (Shapiro et al., 1995a) domains. This paper presents the first detailed NMR data of a cadherin domain, and provides a basis for further studies of the structure and dynamics of members of this CAM superfamily.

Materials and Methods

Sample preparation

The Ecad1 construct includes the first 144 residues of the processed murine E-cadherin protein (Nagafuchi et al., 1987) plus two additional N-terminal amino acids. The amino acid numbering corresponds to that of mature E-cadherin, the precursor sequence of which has been proteolytically cleaved (Ozawa and Kemler, 1990). Recombinant Ecad1 was obtained as described previously (Tong et al., 1994). Uniformly labeled Ecad1 was produced by overexpression in *Escherichia coli* AR58 using M9 minimal media with 0.1% $^{15}NH_4Cl$ (Isotec, Miamisburg, OH) and 0.2% (w/v) of ^{13}C glucose (Isotec) for double labeling, or 0.1% $^{15}NH_4Cl$ and 0.5% (w/v) D-glucose for ^{15}N labeling. The yield was approximately 5 mg per liter of media. The purified protein was exchanged into NMR buffer consisting of either 95% H_2O/5% D_2O or 99.996% D_2O containing 100 mM KCl, 10 mM perdeuterated dithiothreitol (DTT), 20 mM perdeuterated Tris, 50 μM NaN_3 and 10 mM $CaCl_2$, using rinsed Amicon microconcentrators. The sample pH was 7.2, without correction for deuterium isotope effects. The protein concentration of the NMR samples was between 0.1 and 2 mM in volumes of 500 μl. After several weeks of data collection at 23 °C the protein precipitated and the sample needed to be renatured. The protein was denatured by addition of guanidinium hydrochloride and DTT to concentrations of 2 M and 30 mM, respectively, at pH 8.0. Renatured protein was then obtained by gradual dilution back to the original NMR sample conditions. Spectra of freshly prepared and renatured protein were essentially identical, except for slight losses in signal intensity, presumably due to sample loss during the renaturation procedure.

Analytical ultracentrifugation

A Beckman Model E Analytical Ultracentrifuge was used for the sedimentation equilibrium experiments. Measurements of photographic plates were performed on a Nikon model 6C microcomparator at 50× magnification. To determine initial concentrations, first fringe counts were performed on the samples. An average refractive increment of 4.1 fringes mg^{-1} ml^{-1} was assumed (Babul and Steelwagen, 1969). Sample aliquots of 100 μl were loaded into 12-mm double sector, charcoal-filled Epon cells equipped with sapphire windows. The sedimentation equilibrium runs were performed at 5° and 15 000

rpm for approximately 48 h before equilibrium photographs were taken. The sample contained 0.41 mM protein in the NMR buffer described above, except that nonperdeuterated components were used. An assumed value of 0.73 was used for the partial specific volume of the samples. The apparent molecular weight was calculated from the slope after fitting the ln(y) versus r^2 data to a second-degree polynomial equation, using a least-squares technique.

Light scattering

Molecular weight determinations using light scattering techniques were performed on a Dawn F multi-angle laser light-scattering photometer (Wyatt Technology Corporation, Santa Barbara, CA) according to the methodology previously described (Wyatt, 1993). An aliquot of sample was manually injected onto a Pharmacia Superose 12 Gel Filtration column and the scattering intensity and RI signal of the eluting peaks were monitored. Approximately 10-fold dilution was observed upon injection of the sample of 0.41 mM protein. The RI signal was used to calculate the protein concentration, using an assumed value of 0.185 for the sample dn/dc. For each slice across a peak a Debye plot was produced ($R_\Theta/K*c$ versus $\sin^2\Theta/2$), in which the intercept of the extrapolation of scattering intensities to the zero angle and zero concentration yielded the reciprocal molecular mass. The apparent average molecular mass across the whole peak was then calculated from all the individual slices.

NMR spectroscopy

NMR experiments were performed on a four-channel UNITY-plus 500 spectrometer (Varian) with a triple resonance probe, employing an actively shielded z gradient and a pulsed field gradient (PFG) accessory, unless indicated otherwise. Carrier positions used were 119 (^{15}N), 175 (^{13}CO), 54 ($^{13}C^{\alpha}$), 43 ($^{13}C^{\alpha/\beta}$) and 4.792 (^{1}H) ppm. Spectra were recorded in H_2O at 23 °C, unless otherwise noted.

Two homonuclear NOESY (Jeener et al., 1979) experiments were performed in D_2O on an UNITY-600 spectrometer with sweep widths of 16 (F2) and 12 (F1) ppm, 64 transients and 1024 F2 complex points. One of the NOESY spectra was collected with a 50-ms mixing time (τ_m) and 450 F1 complex points and the other with τ_m = 200 ms and 384 F1 complex points. Two TOCSY (Braunschweiler and Ernst, 1983) spectra were obtained in D_2O with sweep widths of 16 (F2) and 12 (F1) ppm and 400 F1 and 1024 F2 complex points. One of the TOCSY experiments was collected with τ_m = 32 ms and 128 transients and the other with τ_m = 52 ms and 96 transients.

A 2D ^{1}H-^{15}N HSQC (Bodenhausen and Ruben, 1980; Bax et al., 1990) spectrum was recorded using the enhanced sensitivity method (Kay et al., 1992). Acquisitions in the F2 (^{1}H) and F1 (^{15}N) dimensions were collected with 1024 and 300 complex points, respectively, and 16 transients. Sweep widths were 16 (F2) and 27 ppm (F1). A ^{15}N-edited NOESY-HMQC (Marion et al., 1989c; Zuiderweg and Fesik, 1989) was recorded on the ^{15}N-labeled sample with τ_m = 100 ms. In total, 1024 F1 (^{1}H), 120 F2 (^{15}N) and 20 F3 (^{1}H) complex points were collected with 16 transients and sweep widths of 16 (F1), 27 (F2), and 47.7 (F3) ppm. A ^{15}N-edited TOCSY-HMQC (Marion et al., 1989a) experiment was recorded with τ_m = 41 ms using the DIPSI-2 mixing sequence (Shaka et al., 1988) and the same parameters as in the ^{15}N-edited NOESY-HMQC, except that 32 transients were acquired with 64 F2 and 32 F3 points. An HNHB (Archer et al., 1991) spectrum was obtained with 1024 F1 (^{1}H), 58 F2 (^{15}N) and 32 F3 (^{1}H) complex points, 32 transients and the same spectral widths as in the ^{15}N-edited NOESY-HMQC.

A constant-time ^{13}C-edited HSQC (Vuister and Bax, 1992) was obtained in D_2O with 1024 F2 (^{1}H) and 256 F1 (^{13}C) complex points and sweep widths of 16 ppm (F2) and 79.6 ppm (F1). A ^{13}C-edited NOESY-HMQC (Ikura et al., 1990) was recorded in D_2O with τ_m = 100 ms, 16 transients and the following number of complex points and spectral widths: ^{1}H (F1) 128/9 ppm, ^{13}C (F2) 32/35.8 ppm, ^{1}H (F3) 832/16 ppm. Two sensitivity-enhanced HCCH-TOCSY (Bax et al., 1990; Kay, 1993) spectra were recorded with τ_m values of 7 and 14 ms, respectively, 16 transients and the following number of complex points and sweep widths: ^{1}H (F1) 128/7 ppm, ^{13}C (F2) 32/23.87 ppm, ^{1}H (F3) 832/16 ppm.

All triple resonance 3D spectra were recorded on the uniformly ^{13}C-/^{15}N-labeled sample in 95% H_2O/5% D_2O with the following numbers of complex points and sweep widths: HNCO (Kay et al., 1990), ^{15}N (F1) 1024/27 ppm, ^{13}CO (F2) 32/11.93 ppm, ^{1}H (F3) 1024/16 ppm (16 transients); HNCACB (Wittekind and Mueller, 1993), $^{13}C^{\alpha/\beta}$ (F1) 18/60.66 ppm, ^{15}N (F2) 48/27 ppm, ^{1}H (F3) 1024/16 ppm (48 transients); CBCA(CO)NH (Grzesiek and Bax, 1992), $^{13}C^{\alpha/\beta}$ (F1) 32/60.66 ppm, ^{15}N (F2) 48/27 ppm, ^{1}H (F3) 1024/16 ppm (24 transients); $(H^{\beta})C^{\beta}(C^{\gamma}C^{\delta})H^{\delta}$ (Yamazaki et al., 1993), $^{13}C^{\beta}$ (F1) 128/60.66 ppm, $^{1}H^{\delta}$ (F3) 1024/16 ppm (736 transients); $(H^{\beta})C^{\beta}(C^{\gamma}C^{\delta}C^{\epsilon})H^{\epsilon}$ (Yamazaki et al., 1993), $^{13}C^{\beta}$ (F1) 128/60.66 ppm, $^{1}H^{\epsilon}$ (F3) 1024/16 ppm (736 transients); H(CCO)NH (Grzesiek et al., 1993), ^{1}H (F1) 128/8 ppm, ^{15}N (F2) 32/32.6 ppm, ^{1}H (F3) 512/8 ppm (24 transients); and simultaneous-acquisition ^{13}C-/^{15}N-edited NOESY-HSQC (Pascal et al., 1994) in H_2O, ^{1}H (F1) 128/10 ppm, ^{13}C/^{15}N (F2) 32/23.87/59.23 ppm, ^{1}H (F3) 832/16 ppm (16 transients with τ_m = 100 ms). The ^{15}N chemical shift in all of the above triple resonance experiments was recorded in a constant-time manner (Bax et al., 1979; Rance et al., 1984; Powers et al., 1991; Grzesiek and Bax, 1992; Palmer et al., 1992). PFGs were employed to suppress spectral artifacts and minimize the water signal (Bax and Pochapsky, 1992). In all amide proton-detected 2D and 3D experiments PFGs

were used to select the coherence pathways for sensitivity enhancement (Kay et al., 1992). Coherent decoupling with WALTZ-16 (Shaka et al., 1983) was used to maintain in-phase heteronuclear magnetization where possible (Farmer et al., 1992; Grzesiek and Bax, 1992). A SEDUCE-1 ^{13}C homonuclear decoupling sequence (McCoy and Mueller, 1992) was used in the HNCACB experiment.

All data sets were processed on Sun Sparc10 and Sparc20 workstations using the programs nmrPipe and nmrDraw (Delaglio, 1993) for 3D NMR data processing and display, respectively. Where necessary, the residual water signal was eliminated using a time-domain deconvolution approach (Marion et al., 1989b). The residual water signal was removed using a low-order polynomial fit of the data (Muhandiram et al., 1993) from the simultaneous-acquisition ^{13}C-/^{15}N-edited NOESY-HSQC spectrum. The apodization function applied depended on the number of data points: a 72° shifted sine bell was used for up to 60 points, a 81° squared sine bell was used for between 60 and 480 points, and a Lorentzian-to-Gaussian filter with nmrPipe parameters G1 = 10 Hz (inverse exponential width), G2 = 20 Hz (Gaussian width) and G3 = 0 (corresponding to a maximum at the first point in the FID) for over 480 points. A single zero-fill was applied in the t_2 and t_3 dimensions of all 3D spectra prior to Fourier transformation, and linear prediction (Barkhuijsen et al., 1985) was used to double the number of planes in the t_1 (or t_2 of HNCACB) dimension of all triple resonance experiments. The programs Capp, Pipp, and Stapp (Garrett et al., 1991) were used for peak picking, spectral analysis and assigning NOEs. Peak lists from 3D spectra were filtered and sequential assignment candidates were obtained using home-written FORTRAN programs (T.S. Harvey, unpublished results).

Coupling constants and stereospecific assignments

The $^3J_{NH^\alpha}$ coupling constants were obtained from an HMQC-J (Kay and Bax, 1990) spectrum obtained on a UNITY-600 spectrometer with 2048 (F2) and 512 (F1) complex points, 96 transients, and sweep widths of 13.33 (F2) and 27 (F1) ppm. The data were processed with a Lorentzian-to-Gaussian filter with the following nmrPipe parameters in F1: G1 = 14 Hz, G2 = 8 Hz and G3 = 0. The $^3J_{NH^\alpha}$ coupling constants were obtained after correcting for distortions arising from finite line widths with spectral simulations (Redfield and Dobson, 1990; Smith et al., 1991) of F1 slices through resolved HMQC-J cross-peaks. Backbone ϕ dihedral angle restraints of $-60° \pm (15°–25°)$ and $-120° \pm (20°–40°)$ were derived from $^3J_{NH^\alpha} \leq 6.0$ and $^3J_{NH^\alpha} \geq 8.0$ Hz, respectively. Valine and leucine residue methyl groups were stereospecifically assigned from the compatibility of rotamer states with NOE intensities (Zuiderweg et al., 1985). Restraints for dihedral angle χ^1 of −60°, 60° and 180° (all with tolerances of ±20° to 30°) were estimated from comparison of ^{13}C-edited NOESY, ^{15}N-edited NOESY, ^{15}N-edited TOCSY, and HNHB cross-peak intensities (Powers et al., 1993).

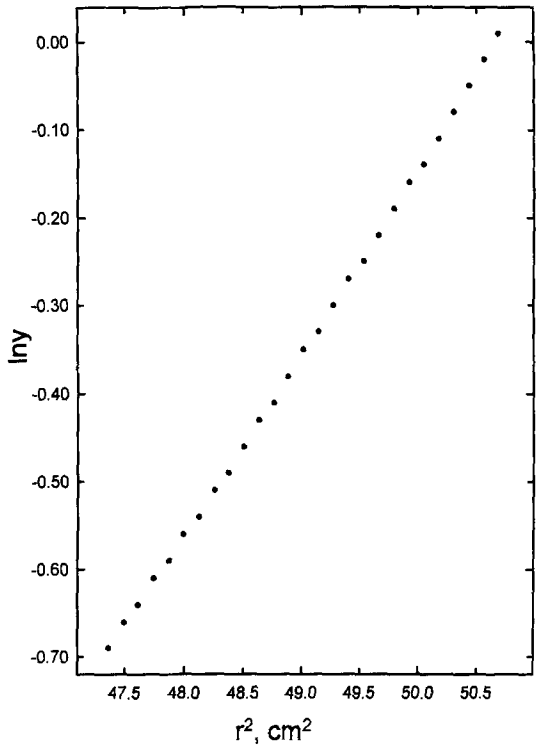

Fig. 1. Sedimentation equilibrium analysis of Ecad1 in the presence of Ca^{2+}. The natural logarithm of the optical density (y) of Ecad1 is plotted as a function of the square of the distance moved in the centrifuge cell (r^2). The monomeric state is supported by the apparent weight-average molecular weight of 14 786 ± 800 Da calculated from the slope of the curve. The loading protein concentration is 0.41 mM.

Hydrogen exchange experiments

The amide proton exchange rates were measured by following the intensity of cross-peaks in a series of sensitivity-enhanced ^1H-^{15}N HSQC (Kay et al., 1992) experiments recorded on a UNITY-600 spectrometer at 23 °C. Four experiments were initiated at 8, 36, 140, and 240 min after dissolving a ^{15}N-labeled Ecad1 sample in NMR buffer, lyophilized from an H_2O NMR buffer. The four experiments were collected with 8, 16, 32 and 64 transients, respectively, and with 64, 128, 128 and 300 F1 complex points, respectively. All four experiments were recorded with 512 F2 complex points and sweep widths of 13.33 and 27 ppm. Pairs of hydrogen-bond restraints were used when the amide proton exhibited slow exchange and could be unambiguously juxtaposed with a CO, or when the hydrogen bond was consistently predicted in initial structures calculated in the absence of the restraint. Hydrogen bonds were defined by restraints defining the H-O distance as 1.9 to 2.5 Å and the N-O distance as 2.9 to 3.5 Å, depending on the longevity of the slowly exchanging NH cross-peak intensity and compatibility with initial structures calculated with looser hydrogen-bond restraints.

Chemical shift indices

CSIs were used to predict ϕ and ψ torsion angles of residues other than proline or glycine (Wishart and Sykes, 1994). In particular, CSIs for the $^1H^\alpha$, $^{13}C^\alpha$ and carbonyl $^{13}C'$ resonances were combined as follows: $CSI^{C^\alpha} + CSI^{C'} - CSI^{H^\alpha}$. A consensus CSI value of +2 or more predicted helical torsion angles of −60° (ϕ) and −40° (ψ), while a consensus CSI value of −2 or less predicted extended torsion angles of −120° (ϕ) and 130° (ψ), all with tolerances of ± 30° or 40°, with the tighter restraint tolerances being supported by larger consensus CSI values and compatibility with initial structures.

Structure calculations

The experimental restraints consisted of interproton distances derived from NOE data, hydrogen-bond restraints and backbone dihedral restraints derived from $^3J_{NH^\alpha}$ coupling constants and CSIs. NOE cross-peaks were classified as strong (< 2.7 Å), medium (< 3.3 Å) or weak (< 5.0 Å) (Williamson et al., 1985; Clore et al., 1986). Standard corrections were added to the upper bounds of restraints involving pseudoatoms for methylene and methyl groups as well as for the H^δ and H^ε protons of tyrosine and phenylalanine (Clore et al., 1986; Wagner et al., 1987). An additional 0.5 Å was added to the upper bounds of NOEs involving methyl groups (Wagner et al., 1987). The position of the Ca^{2+} was inferred from chemical shift perturbations, since no unambiguous restraints to Ca^{2+} could be identified.

Two cis peptide bonds were used in the structure calculations. Pro^{18} displayed strong $d_{\alpha\alpha}(17,18)$ NOEs, which are indicative of a *cis*-Pro (Wüthrich, 1986). In the Ecad12 (Nagar et al., 1996) and Ncad1 (Shapiro et al., 1995b) structures a cis peptide bond between Pro^{46} and Pro^{47} was observed. The lack of Pro^{46} H^α or $^{13}C^\beta$ assignments precluded the identification of NOEs or chemical shifts that would have predicted a cis peptide bond (Scanlon and Norton, 1994). However, a cis peptide bond preceding Pro^{47} was imposed since it was most compatible with a tightly constrained βC-βD loop.

The solution structures were generated and refined using the distance geometry and simulated annealing procedures (Nilges et al., 1988,1991) with X-PLOR v. 2.1 and 3.1 (Brünger, 1992), installed on a four-processor Silicon Graphics Onyx Challenge computer server. An iterative procedure was used in which initial structures were generated with unambiguous restraints. Subsequent refinements included additional restraints with reduced weighting, which were subsequently screened for errors and violations. Ambiguous NOEs (and corresponding restraints used in subsequent calculations) were assigned using the program Stapp (Garrett et al., 1991). During the calculations the temperature was raised to 1000 K and then dropped directly to 300 K. Neither a full Lennard-

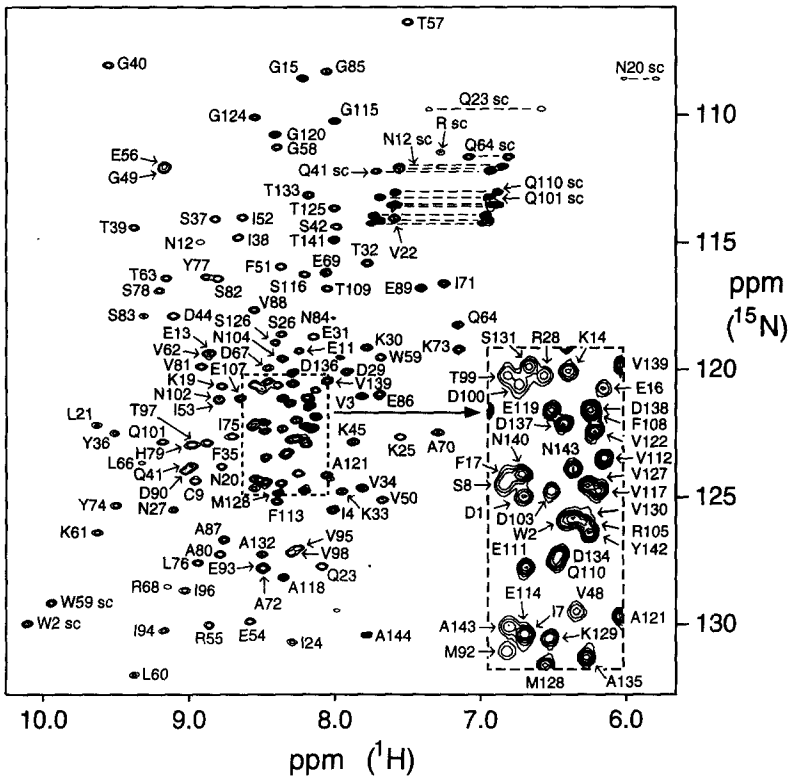

Fig. 2. ^1H-^{15}N HSQC spectrum of Ecad1. Peaks for backbone atom amide protons have been assigned and are labeled with the one-letter code for the amino acid type, followed by the position in the sequence. Tryptophan indole cross-peaks are labeled as 'sc', as are assigned pairs of peaks from side-chain amide protons connected by horizontal lines. The inset is an expanded view of a crowded region delimited by the dotted lines within the spectrum.

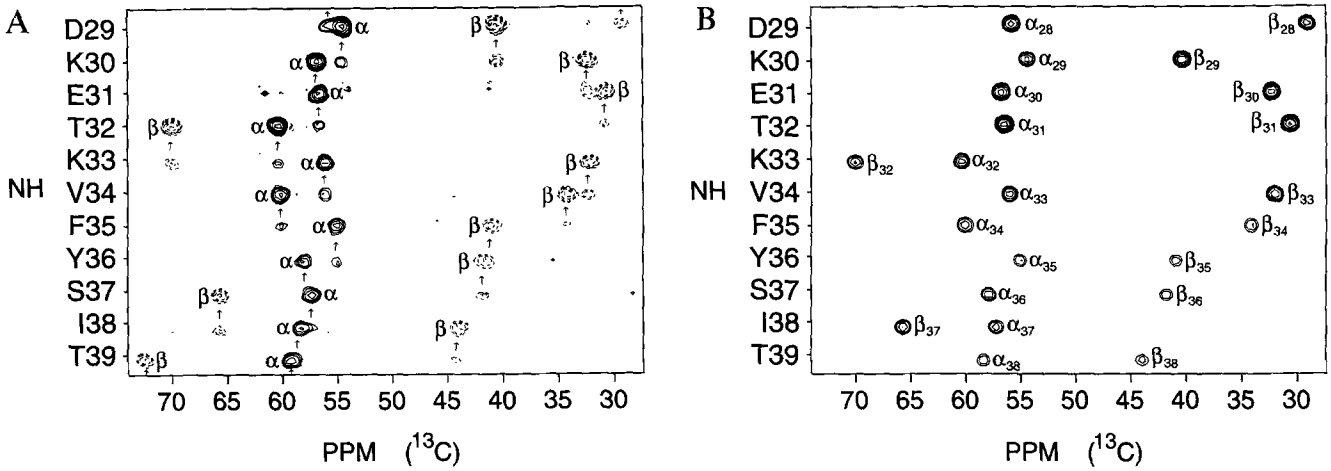

Fig. 3. F1 ($^{13}C^{\alpha/\beta}$) strips for residues 29–39 from (A) the 3D HNCACB and (B) the 3D CBCA(CO)NH spectra of Ecad1. Strips are extracted at the F3/F2 ($^{1}H/^{15}N$) frequencies of the backbone amide of each residue. In each strip of (A), the intraresidue $^{13}C^{\alpha}$ and $^{13}C^{\beta}$ correlations are labeled as 'α' and 'β', respectively. Each strip also contains weaker correlations to the C^{α} and C^{β} atoms of the previous residue, which are indicated by arrows pointing to the intraresidue cross-peaks of the previous residue. Differentiation between C^{α} and C^{β} resonances is aided by their opposite phases: C^{α} resonances are positive (solid lines) while C^{β} resonances are negative (dashed lines). The CBCA(CO)NH spectrum (B) contains only interresidue correlations between the amide proton and the C^{α} and C^{β} atoms of the previous residue, indicated with 'α' and 'β', respectively. Matching of ^{13}C chemical shifts between these two spectra facilitates the sequential assignment of ^{13}C-/^{15}N-labeled proteins.

Jones potential nor electrostatic terms were included. NOE-derived distance restraints were applied with a square-well potential with force constants of 50 kcal mol^{-1} Å^{-2}. The coordinates for the final 20 structures have been deposited in the Protein Data Bank (Chemistry Department, Brookhaven National Laboratory, Upton, NY).

Results

Oligomeric state

Sedimentation equilibrium centrifugation and light-scattering experiments were performed to determine the oligomeric state of Ecad1 in the NMR buffer in the presence of 10 mM $CaCl_2$. From the centrifugation data an apparent average molecular weight of 14 786 ± 800 Da was obtained in a protein concentration range of 0.26 to 0.59 mM, in close agreement with the calculated monomeric molecular mass of 16 134.9 Da. It was not possible to carry out sedimentation equilibrium experiments at higher protein concentrations, because of the large number of fringes encountered at the centrifuge initial loading concentration of 0.41 mM. Higher protein concentrations would result in the larger number of fringes becoming too compressed at the cell bottom to be resolved. The linear relationship between r^2 and the natural logarithm of the optical density indicates that only a single species is present in solution (Fig. 1). Dynamic light scattering (performed on a sample that was diluted approximately 10-fold after being injected into an HPLC column at an initial protein concentration of 0.41 mM) also indicates that Ecad1 is monomeric, with a weight-average molecular mass of 15 254 Da. Approximately the same molecular weight estimates were obtained in the absence of Ca^{2+}. The monomeric state of Ecad1 is further supported by the single set of resonances with narrow sharp line widths seen in the NMR spectra. No significant concentration-

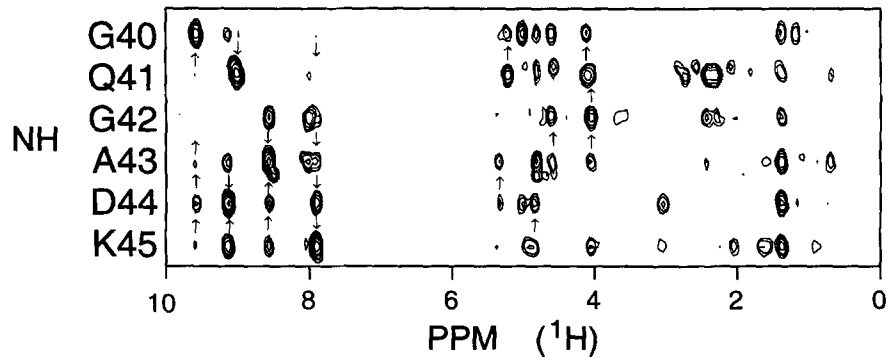

Fig. 4. Strip plots extracted from the $\tau_m = 100$ ms ^{15}N-edited NOESY-HMQC spectrum of Ecad1. Residues Gly^{40} to Lys^{45} are found in the quasi-β-helix. The series of sequential d_{NN} NOE connectivities is shown, as well as novel NOE patterns such as $d_{NN}(40,43)$, $d_{NN}(40,44)$ and $d_{NN}(40,45)$, which are incompatible with a regular α- or 3_{10}-helix. Arrows point toward the intraresidue NOE cross-peaks.

TABLE 1
^{15}N, ^{13}C AND ^{1}H RESONANCE ASSIGNMENTS OF Ecad1

Residue	N	C	C^{α}	C^{β}	Others
M(−2)			55.2 (4.05)	33.5 (2.12*)	C^{γ} 30.8 (2.61*); C^{ε} –
R(−1)		175.1	55.9 (4.38)	30.5 (1.82, 1.75)	C^{γ} 26.7 (1.62*); C^{δ} 43.5 (3.17*)
D1	122.4 (8.46)	175.3	54.4 (4.63)	41.3 (2.72, 2.62)	
W2	122.7 (8.27)	174.3	56.6 (4.73)	29.5 (3.38, 3.19)	$C^{\delta 2}$ 126.7 (7.27); $C^{\varepsilon 3}$ 121.0 (7.56); $C^{\zeta 2}$ 114.2 (7.34); $C^{\zeta 3}$ – (6.92); C^{η} 123.2 (7.00); $N^{\varepsilon 1}$ 130.0 (10.09)
V3	121.3 (7.79)	174.6	61.2 (4.08)	33.9 (1.96)	C^{γ} (0.83*)
I4	125.5 (7.99)		57.4 (4.48)	39.2 (1.77)	C^{γ} – (1.37, 1.13); $C^{\gamma m}$ 16.9 (0.85); C^{δ} 13.3 (0.66)
P5				– (2.30, 2.07)	C^{γ} 27.0 (1.98*); C^{δ} 50.8 (3.79, 3.60)
P6		176.4	62.9 (4.53)	31.7 (2.16, 2.05)	C^{γ} 27.09 (1.79, 1.88); C^{δ} – (3.50)
I7	124.4 (8.46)	174.7	59.8 (4.29)	40.3 (1.79)	C^{γ} – (1.47, 1.05); $C^{\gamma m}$ – (0.62); C^{δ} 13.2 (0.85)
S8	122.2 (8.54)	173.3	56.5 (5.46)	64.8 (3.74*)	
C9	124.4 (8.94)	171.1	55.4 (5.04)	30.0 (3.0 2.37)	
P10		176.9	61.9 (4.76)	32.5 (2.29, 1.89)	C^{γ} – (2.06); C^{δ} – (3.84, 3.74)
E11	119.3 (8.22)	174.8	55.7 (4.21)	29.4 (1.90)	C^{γ} – (2.21, 2.12)
N12	114.9 (8.90)	174.1	53.6 (4.36)	37.4 (3.08, 2.82)	N^{δ} 112.0 (6.83, 7.54)
E13	119.3 (8.84)	175.8	58.2 (4.20)	30.8 (2.16, 1.99)	C^{γ} 37.4 (2.41, 2.20)
K14	120.5 (8.26)	176.0	54.3 (4.68)	33.9 (1.95, 1.71)	C^{γ} 24.3 (1.49*); C^{δ} 28.6 (1.70*); C^{ε} 42.2 (3.03*)
K14[b]			55.5 (4.03)	33.3 (1.89*)	C^{γ} 23.6 (1.43*); C^{δ} 28.8 (1.70*); C^{ε} 41.8 (3.01*)
G15	108.6 (8.19)	172.3	44.3 (3.92, 3.73)		
G15[b]	105.9 (7.78)				
E16	120.8 (8.11)	176.0	55.9 (4.25)	30.2 (1.86, 1.76)	C^{γ} 36.1 (2.21, 2.11)
F17	122.1 (8.52)		56.3 (4.22)	41.1 (2.76, 2.59)	C^{δ} 131.6 (7.12*) C^{ε} 130.6 (7.03*) C^{ζ} 128.8 (6.90)
P18		175.4	62.0 (5.06)	35.6 (2.22, 1.87)	C^{γ} – (1.97, 1.78)
K19	120.6 (8.75)	175.4	54.7 (4.63)	35.3 (1.71, 1.57)	C^{ε} – (3.02*)
N20	123.8 (8.75)	175.4	53.8 (4.57)	38.1 (2.41, 2.28)	N^{δ} 108.6 (5.77, 5.99)
L21	122.1 (9.62)	175.8	55.8 (4.80)	43.8 (1.79, 1.45)	C^{δ} – (0.98*)
V22	114.0 (7.57)	169.8	60.1 (4.45)	33.9 (2.18)	C^{γ} 20.5 (0.83); C^{γ} 19.9 (0.75)
Q23	127.7 (8.07)	174.2	53.5 (3.81)	27.2 (1.12*)	C^{γ} – (0.74, –1.04); N^{ε} 109.8 (6.56, 7.33)
I24	130.7 (8.27)	173.3	60.1 (3.74)	36.6 (1.56)	C^{γ} – (1.14, 1.00); $C^{\gamma m}$ 18.3 (0.65); 12.7 (0.58)
K25	122.6 (7.53)	175.2	54.6 (4.44)	35.4 (1.54, 1.30)	C^{γ} 24.0 (1.08*); C^{δ} 28.65 (1.45*); C^{ε} 41.5 (2.80*)
S26	118.6 (8.34)	175.3	56.7 (4.58)	64.1 (3.63, 3.51)	
N27	125.5 (9.09)	175.9	54.0 (4.67)	38.1 (2.77, 2.81)	
R28	120.6 (8.37)	175.9	55.8 (4.26)	29.2 (1.85, 1.54)	C^{γ} 26.6 (1.64, 1.51); C^{δ} 42.2 (3.12*)
D29	120.1 (7.90)	176.6	54.4 (4.72)	40.4 (2.91, 2.61)	
K30	119.1 (7.76)	177.1	56.8 (4.25)	32.3 (1.84, 1.69)	C^{γ} 24.9 (1.48, 1.42); C^{δ} 28.6 (1.69*); C^{ε} (3.02*)
E31	118.7 (8.12)	175.7	56.5 (4.32)	30.7 (2.09, 1.98)	C^{γ} 35.9 (2.21*)
T32	115.8 (7.76)	171.6	60.2 (4.33)	69.9 (4.01)	C^{γ} – (0.99)
K33	124.8 (7.93)	174.9	56.0 (4.24)	32.1 (1.66, 1.61)	C^{γ} 24.3 (1.16*); C^{δ} 28.7 (1.67*); C^{ε} 41.9 (2.99, 2.81)
V34	124.6 (7.79)	172.7	60.1 (4.26)	34.1 (1.52)	C^{γ} 21.9 (0.55); C^{γ} 22.9 (0.29)
F35	122.8 (8.85)	175.4	55.0 (5.04)	41.0 (2.86, 2.58)	C^{δ} 131.6 (7.04*); C^{ε} 131.6 (7.17*); C^{ζ} 130.3 (7.21)
Y36	122.5 (9.49)	175.7	58.0 (5.66)	41.6 (3.03, 2.76)	C^{δ} 132.7* (7.06*); C^{ε} 117.0* (6.49*)
S37	114.0 (8.80)	172.7	57.3 (4.77)	65.6 (3.95, 3.93)	
I38	114.8 (8.64)	174.6	58.3 (6.09)	43.8 (1.65)	$C^{\gamma m}$ – (0.85); C^{δ} – (0.65)
T39	114.4 (9.35)	172.7	59.2 (4.97)	72.4 (4.57)	C^{γ} 21.4 (1.17)
G40	108.1 (9.53)	174.8	43.2 (5.21, 4.09)		
Q41	123.8 (8.96)	174.9	57.6 (3.98)	28.2 (2.39, 2.29)	C^{γ} 34.0 (2.71, 2.38); N^{ε} 112.2 (6.92, 7.70)
G42	114.4 (7.96)	172.3	45.1 (4.57, 4.76)		
A43	124.3 (8.52)	175.2	51.5 (5.29)	18.8 (1.35)	
D44	117.8 (9.09)	175.1	53.4 (4.80)	40.2 (3.00, 2.88)	
K45	122.8 (7.85)		52.3 (4.88)	32.7 (2.01)	C^{γ} 24.2 (1.35, 1.32); C^{δ} 29.1 (1.58, 1.49); C^{ε} 41.7 (2.97, 2.82)
P46					C^{δ} (3.91, 3.70)
P47		174.0	61.8 (4.69)	32.6 (2.30, 1.90)	C^{γ} 26.6 (1.98); C^{δ} 49.9 (3.51, 3.67)
V48	124.1 (8.23)	179.5	62.8 (3.49)	32.2 (1.91)	C^{γ} 19.4 (0.90); C^{γ} 21.0 (0.81)
G49	112.1 (9.16)	173.9	45.8 (3.84, 3.61)		
V50	125.1 (7.65)	173.9	66.2 (3.02)	31.1 (1.97)	C^{γ} 22.9 (0.73); C^{γ} 22.4 (–0.02)
F51	115.9 (8.35)	174.8	55.3 (5.51)	41.8 (2.72, 2.21)	C^{δ} 131.1 (6.74*); C^{ε} 130.3 (6.11*); C^{ζ} 128.6 (6.45)
I52	114.0 (8.61)	173.0	58.4 (4.73)	42.5 (1.92)	C^{γ} – (1.19, 0.96); $C^{\gamma m}$ – (0.86); C^{δ} 13.20 (0.80)
I53	121.2 (8.77)	173.5	57.7 (5.61)	43.1 (1.91)	C^{γ} – (1.09, 1.22); $C^{\gamma m}$ 16.9 (0.85)
E54	129.9 (8.56)	177.1	56.1 (4.48)	30.1 (2.15*)	C^{γ} 36.7 (2.53, 2.46)
R55	130.0 (8.84)	177.2	59.1 (4.79)	30.6 (2.10, 2.25)	C^{δ} (1.85*); C^{δ} 43.3 (3.32)
E56	112.0 (9.15)	178.0	57.6 (4.82)	30.7 (2.36, 2.20)	C^{γ} 37.9 (2.52, 2.48)

TABLE 1
(continued)

Residue	N	C	C^{α}	C^{β}	Others
T57	106.4 (7.48)	176.0	61.5 (4.74)	70.3 (4.51)	C^{γ} 21.1 (1.41)
G58	111.2 (8.37)	174.3	45.0 (4.02, 2.81)		
W59	119.5 (7.67)	174.0	60.1 (4.40)	28.7 (2.89)	$C^{\delta 2}$ 125.8 (7.13); $C^{\varepsilon 3}$ 121.2 (6.65); $C^{\zeta 2}$ 114.7 (7.37); $C^{\zeta 3}$ 119.2 (7.17); C^{η} 124.5 (7.16); $N^{\varepsilon 1}$ 129.14 (9.93)
L60	132.0 (9.35)	175.5	54.4 (5.04)	43.5 (2.22)	C^{γ} 27.4 (1.63); C^{δ} 25.4 (0.84); C^{δ} 26.5 (0.731)
K61	126.4 (9.61)	174.0	54.2 (5.51)	36.7 (1.43, 1.39)	C^{γ} – (1.70); C^{δ} (1.88); C^{ε} 42.6 (2.99*)
V62	119.3 (8.85)	176.5	57.9 (5.36)	34.7 (2.13)	C^{γ} 20.0 (1.24); C^{γ} 22.9 (1.00)
T63	116.4 (9.14)	172.8	62.9 (4.18)	69.7 (4.48)	C^{γ} – (1.10)
Q64	118.2 (7.13)	170.2	52.1 (3.60)	29.5 (1.92, 1.83)	N^{δ} 111.6 (6.79, 7.06)
P65		176.2	61.9 (4.36)	31.8 (1.98, 1.84)	C^{γ} – (2.02); C^{δ} – (3.25, 2.68)
L66	123.6 (9.30)		53.4 (4.59)	43.4 (1.86, 1.70)	C^{δ} – (1.03); C^{δ} 22.06 (0.59)
D67	119.9 (8.44)	176.5	52.7 (5.08)	42.1 (2.90, 2.64)	
R68	128.5 (9.13)	176.0	58.2 (3.93)	30.4 (1.71, 1.59)	C^{γ} 26.9 (1.63, 1.55); C^{δ} 43.3 (3.05*)
E69	116.2 (8.04)	176.6	58.0 (3.93)	28.8 (2.12, 2.04)	C^{γ} 37.0 (2.41, 2.27)
A70	122.4 (7.27)	177.1	53.5 (4.35)	18.9 (1.51)	
I71	116.6 (7.23)	174.0	61.6 (3.93)	41.1 (1.61)	C^{γ} – (1.37, 0.69); $C^{\gamma m}$ 17.2 (0.07); C^{δ} 13.49 (0.84)
A72	127.8 (8.46)	176.5	52.5 (4.49)	19.5 (1.49)	
K73	119.2 (7.13)	173.2	55.0 (5.25)	35.3 (1.71, 1.59)	C^{γ} – (1.29, 1.13); C^{δ} 29.1 (1.56*); C^{ε} 41.4 (2.82*)
Y74	125.3 (9.48)	173.7	56.4 (4.70)	40.5 (2.80, 2.56)	C^{δ} 133.4 (7.14*); C^{ε} 117.6 (6.65)
I75	122.6 (8.68)	174.1	60.5 (4.68)	38.7 (1.79)	C^{γ} – (1.45, 1.05); $C^{\gamma m}$ 17.8 (0.85); C^{δ} 13.8 (0.69)
L76	127.6 (8.92)	174.4	51.6 (5.14)	45.2 (1.36, 1.33)	C^{γ} – (1.28); C^{δ} 23.9 (0.66); C^{δ} 26.2 (0.29)
Y77	116.3 (8.86)	175.3	56.2 (5.31)	40.7 (2.81, 2.29)	C^{δ} 132.4 (6.84*); C^{ε} 118.3 (6.78*)
S78	116.9 (9.18)	175.4	56.2 (5.45)	65.3 (3.70, 3.57)	
H79	122.9 (8.97)	173.4	54.3 (5.05)	31.8 (3.21, 2.85)	$C^{\delta 2}$ – (6.84); $C^{\varepsilon 1}$ 136.6 (7.77)
A80	127.3 (8.76)	176.1	50.3 (4.96)	20.4 (0.90)	
V81	119.9 (8.90)	174.4	58.5 (5.21)	35.8 (1.90)	C^{γ} 19.4 (0.90); C^{γ} 20.4 (0.83)
S82	116.4 (8.78)	177.3	56.3 (5.11)	64.9 (4.42, 3.87)	
S83	117.9 (9.29)	174.3	61.6 (4.73)	62.5 (4.63, 4.54)	
N84	117.9 (7.99)	175.8	52.5 (3.93)	38.0 (2.87, 2.94)	
G85	108.3 (8.04)	173.5	45.0 (4.21, 3.47)		
E86	121.0 (7.68)	175.5	55.2 (4.31)	30.1 (2.08, 1.93)	C^{γ} 36.4 (2.35*)
A87	126.7 (8.74)	178.8	52.7 (4.69)	18.3 (1.53)	
V88	117.6 (8.53)	174.5	61.1 (4.30)	31.8 (2.04)	C^{γ} 21.2 (0.85); C^{γ} 19.4 (0.61)
E89	116.8 (7.39)	175.8	53.7 (4.52)	33.0 (2.04, 1.79)	
D90	124.0 (9.01)		53.3 (4.93)	39.2 (2.80, 2.56)	
P91		176.2	63.2 (4.76)	31.9 (2.20, 2.37)	C^{γ} – (1.98, 1.90); C^{δ} – (4.00, 3.63)
M92	124.6 (8.53)	173.9	53.8 (4.71)	34.9 (2.12, 1.75)	C^{γ} – (2.29)
E93	127.8 (8.48)	174.4	56.1 (3.90)	31.1 (1.71, 1.70)	C^{γ} – (1.84)
I94	130.2 (9.14)	173.2	59.8 (4.09)	39.6 (1.32)	C^{γ} – (1.62, 0.67); $C^{\gamma m}$ 16.7 (0.47); C^{δ} 14.2 (0.53)
V95	127.0 (8.23)	175.1	60.8 (4.62)	32.8 (1.80)	C^{γ} 21.4 (0.84); C^{γ} – (0.82)
I96	128.7 (9.01)	175.4	59.3 (4.83)	39.8 (1.58)	C^{γ} – (1.01, 0.56); $C^{\gamma m}$ 17.6 (0.66); C^{δ} 13.2 (–0.11)
T97	122.9 (8.95)	172.7	61.2 (4.74)	69.8 (4.06)	C^{γ} – (1.16)
V98	127.2 (8.27)	176.1	61.0 (5.05)	32.2 (2.16)	C^{γ} 21.7 (1.20); C^{γ} 22.5 (1.04)
T99	120.6 (8.53)	173.7	60.1 (4.63)	70.1 (4.36)	C^{γ} 21.5 (1.11)
D100	120.5 (8.44)	176.4	53.9 (4.59)	41.8 (2.61, 2.65)	
Q101	122.8 (9.16)	175.6	56.2 (4.38)	29.2 (2.13, 2.10)	C^{γ} 33.3 (2.42*); N^{ε} 113.2 (6.92, 7.67)
N102	121.1 (8.82)	174.7	53.2 (4.66)	38.7 (2.86, 2.74)	
D103	122.3 (8.34)	175.4	53.9 (4.58)	40.7 (2.64, 2.68)	
N104	119.6 (8.34)	174.5	52.8 (4.69)	38.6 (2.78, 2.71)	
R105	122.7 (8.20)		53.8 (4.61)	29.8 (1.86, 1.75)	C^{γ} 26.4 (1.67*); C^{δ} 43.5 (3.21*)
P106		176.5	63.0 (4.40)	31.7 (2.26, 1.83)	C^{γ} 27.1 (2.01*); C^{δ} 50.5 (3.79, 3.65)
E107	121.1 (8.62)	176.0	56.6 (4.16)	29.8 (1.90, 1.86)	C^{γ} 35.9 (2.20, 2.10)
F108	121.1 (8.16)	175.3	57.1 (4.75)	39.2 (3.21, 3.04)	C^{δ} 131.6 (7.26*); C^{ε} 131.1 (7.34*); C^{ζ} 129.6 (7.29)
T109	116.8 (8.03)	173.6	61.4 (4.28)	69.6 (4.16)	C^{γ} 21.3 (1.18)
Q110	123.3 (8.32)	175.3	55.5 (4.31)	29.3 (2.08, 1.97)	C^{γ} 33.5 (2.33*); N^{ε} 113.0 (6.86, 7.57)
E111	123.4 (8.45)	175.8	56.3 (4.25)	29.9 (2.02, 1.91)	C^{γ} 35.9 (2.22, 2.15)
V112	121.8 (8.11)	175.3	61.8 (4.08)	32.7 (1.98)	C^{γ} 20.7 (0.86, 0.84)
F113	125.2 (8.37)	175.3	57.2 (4.68)	39.5 (3.11, 3.02)	C^{δ} 131.6 (7.26*); C^{ε} 131.1 (7.34*); C^{ζ} 129.6 (7.29)
E114	124.4 (8.44)	176.2	56.4 (4.23)	30.2 (1.98, 1.87)	C^{γ} – (2.19, 2.21)
G115	110.2 (7.98)	173.7	45.0 (3.91*)		
S116	116.2 (8.19)	174.0	57.9 (4.48)	63.6 (3.84, 3.80)	

TABLE 1
(continued)

Residue	N	C	C^{α}	C^{β}	Others
V117	122.3 (8.13)	175.6	62.0 (4.11)	32.5 (2.07)	C^{γ} 20.8 (0.93); C^{γ} – (0.93)
A118	128.2 (8.33)	177.4	52.1 (4.29)	19.0 (1.36)	
E119	121.1 (8.34)	176.7	56.6 (4.25)	30.0 (2.06, 1.96)	C^{γ} 35.9 (2.28*)
G120	110.8 (8.39)	173.3	44.9 (3.92*)		
A121	124.1 (8.04)	177.2	52.0 (4.34)	19.2 (1.35)	
V122	121.4 (8.15)		59.5 (4.41)	32.3 (2.08)	C^{γ} 20.9 (0.97); C^{γ} – (0.95)
P123		177.2	63.3 (4.42)	31.8 (2.32, 1.96)	C^{γ} 27.1 (2.08, 2.00); C^{δ} 50.8 (3.90, 3.71)
G124	110.1 (8.53)	174.2	45.0 (3.99*)		
G124[b]	111.1 (8.54)		45.3 (4.00)		
T125	113.6 (8.00)	174.3	61.5 (4.41)	69.6 (4.27)	C^{γ} 21.6 (1.21)
T125[b]	113.9 (8.13)		61.3 (4.38)	69.3	
S126	118.9 (8.39)	174.3	58.0 (4.51)	63.5 (3.86)	
V127	122.2 (8.18)	175.6	61.9 (4.13)	32.5 (2.09)	C^{γ} 20.8 (0.93); 19.1 (0.93)
M128	124.8 (8.37)	175.5	55.1 (4.48)	32.6 (2.04, 2.00)	C^{γ} 31.6 (2.59, 2.53)
K129	124.4 (8.35)	176.0	55.7 (4.38)	32.9 (1.82, 1.75)	C^{γ} 24.2 (1.42*); C^{δ} 28.8 (1.68*); C^{ϵ} 42.4 (2.99*)
V130	122.7 (8.24)	173.7	61.8 (4.15)	32.7 (2.07)	C^{γ} 20.8 (0.93); C^{γ} – (0.93)
S131	120.5 (8.43)	173.8	57.6 (4.58)	63.6 (3.84)	
A132	127.2 (8.48)	177.6	52.4 (4.41)	19.2 (1.43)	
T133	113.1 (8.16)	174.1	61.3 (4.37)	69.5 (4.30)	C^{γ} 21.5 (1.20)
D134	123.2 (8.30)	175.6	54.1 (4.60)	40.8 (2.71, 2.64)	
A135	124.7 (8.18)	177.2	52.3 (4.30)	19.2 (1.38)	
D136	120.1 (8.27)	175.7	54.1 (4.59)	40.9 (2.68, 2.64)	
D137	121.3 (8.28)	175.8	54.1 (4.61)	40.9 (2.72, 2.64)	
D138	121.1 (8.17)	176.2	54.1 (4.61)	40.9 (2.63, 2.73)	
V139	120.4 (8.03)	175.9	62.3 (4.06)	32.1 (2.13)	C^{γ} 20.8 (0.93); C^{γ} – (0.93)
N140	122.0 (8.47)	175.1	53.1 (4.73)	38.7 (2.76*)	
T141	114.9 (7.98)	173.9	61.7 (4.25)	69.2 (4.15)	C^{γ} 21.0 (1.11)
Y142	122.9 (8.17)	175.0	57.6 (4.58)	38.3 (3.07, 2.97)	C^{δ} 133.0 (7.12*); C^{ϵ} 117.8 (6.82*)
N143	122.0 (8.24)	173.1	52.8 (4.68)	38.8 (2.74, 2.63)	
A144	130.4 (7.76)	167.2	53.5 (4.07)	19.8 (1.32)	

The chemical shifts were obtained at 23 °C and pH 7.2. In each column, ^{15}N and ^{13}C chemical shifts are listed first and the corresponding ^{1}H shifts follow in parentheses. The chemical shift reference used for ^{1}H and ^{13}C is the sodium salt of 3-(trimethylsilyl)propionate, corrected for pH effects (Wishart and Sykes, 1994). ^{15}N chemical shifts are reported relative to external liquid NH_3. In cases where only one resonance has been identified for diastereotopic protons in methylene groups or diastereotopic methyl groups in isopropyl groups, an asterisk indicates certainty that the diastereotopic protons or methyl groups have identical chemical shifts. The accuracy is ± 0.01 ppm for ^{1}H chemical shifts, ± 0.1 ppm for ^{15}N chemical shifts, ± 0.1 ppm for $^{13}C^{\alpha}$ and $^{13}C^{\beta}$, and ± 0.2 ppm for other ^{13}C chemical shifts.

dependent changes were detected in 1D spectra collected from NMR samples of 0.1 to 2 mM protein in the presence of Ca^{2+} (data not shown). Thus, no evidence was found for the dimerization of Ecad1 in the absence or presence of Ca^{2+}.

Resonance assignments

The assignment of resonances of the 146 amino acids of Ecad1 began with the identification of 173 ^{1}H-^{15}N cross-peaks (see Fig. 2). We expect 135 main-chain amide ^{1}H peaks (146 residues minus the N-terminal methionine and 10 proline residues), 13 pairs of cross-peaks from side-chain NH_2 groups, and additional side-chain NH cross-peaks from 2 tryptophan, 7 arginine and 12 lysine residues. The ^{1}H-^{15}N cross-peak scatter, i.e., a preponderance of downfield chemical shifts, suggested a largely β-sheet protein. Although most are resolved, the ^{1}H-^{15}N cross-peaks of the following pairs of residues are overlapped: Glu^{13} and Val^{62}, Asp^{138} and Phe^{108}, Ile^{53} and Asp^{102}, His^{79} and Thr^{97}, and Ile^{7} and Glu^{114}. These overlaps were readily apparent in triple resonance experiments correlating the amide ^{15}N and ^{1}H with ^{13}C resonances. The dipeptide sequences Lys^{14}-Gly^{15} and Gly^{124}-Thr^{125} yielded second, weaker sets of ^{1}H-^{15}N cross-peaks. The former dipeptide is found near a *cis*-proline and the latter is in the disordered C-terminal region next to a proline residue. Thus, the duplicate cross-peaks are likely due to additional, minor conformational states of these sequences. Most of the high-intensity cross-peaks in the crowded random-coil portion of the ^{1}H-^{15}N cross-peak scatter (Fig. 2, inset) are derived from the 40 disordered C-terminal residues.

Virtually complete sequential connectivities were obtained from the CBCA(CO)NH and HNCACB spectra. The CBCA(CO)NH correlates the amide proton and nitrogen with interresidue α and β carbons, while the HNCACB also provides corresponding intraresidue correlations. Sequential assignments of residues 29 to 39,

which encompass part of the βB-βC loop and the βC strand, are shown in Fig. 3. The cross-peaks in these data correlate $^{13}C^{\alpha}$ and $^{13}C^{\beta}$ of residues (i − 1) and i with amide $^{1}H(i)$ and $^{15}N(i)$. The sequential assignments were confirmed by interresidue correlations between side-chain protons of residue i and the amide ^{1}H and ^{15}N of residue (i + 1) in H(CCO)NH spectra. These interresidue correlations were matched to intraresidue correlations observed in ^{15}N-edited TOCSY-HMQC and HNHB spectra. Sequential assignments were also confirmed and extended with NOE data. For example, the sequential assignment of Pro^{5} and Pro^{46}, which are next to proline residues, was based on NOE data due to the lack of scalar correlations to a succeeding amide proton.

Amino acid types of alanine, glycine, serine and threonine residues could be assigned from the $^{13}C^{\alpha}$ and $^{13}C^{\beta}$ chemical shifts observed in CBCA(CO)NH and HNCACB spectra. Several aromatic residues could be identified by connecting the aromatic ring spin systems seen in homonuclear TOCSY spectra to the main-chain assignments through intraresidue H^{α}/H^{β}-H^{δ} NOE cross-peaks found in homonuclear NOESY spectra. The aromatic assignments were confirmed by C^{β}-H^{δ} and C^{β}-H^{ε} cross-peaks observed in $(H^{\beta})C^{\beta}(C^{\gamma}C^{\delta})H^{\delta}$ and $(H^{\beta})C^{\beta}(C^{\gamma}C^{\delta}C^{\varepsilon})H^{\varepsilon}$

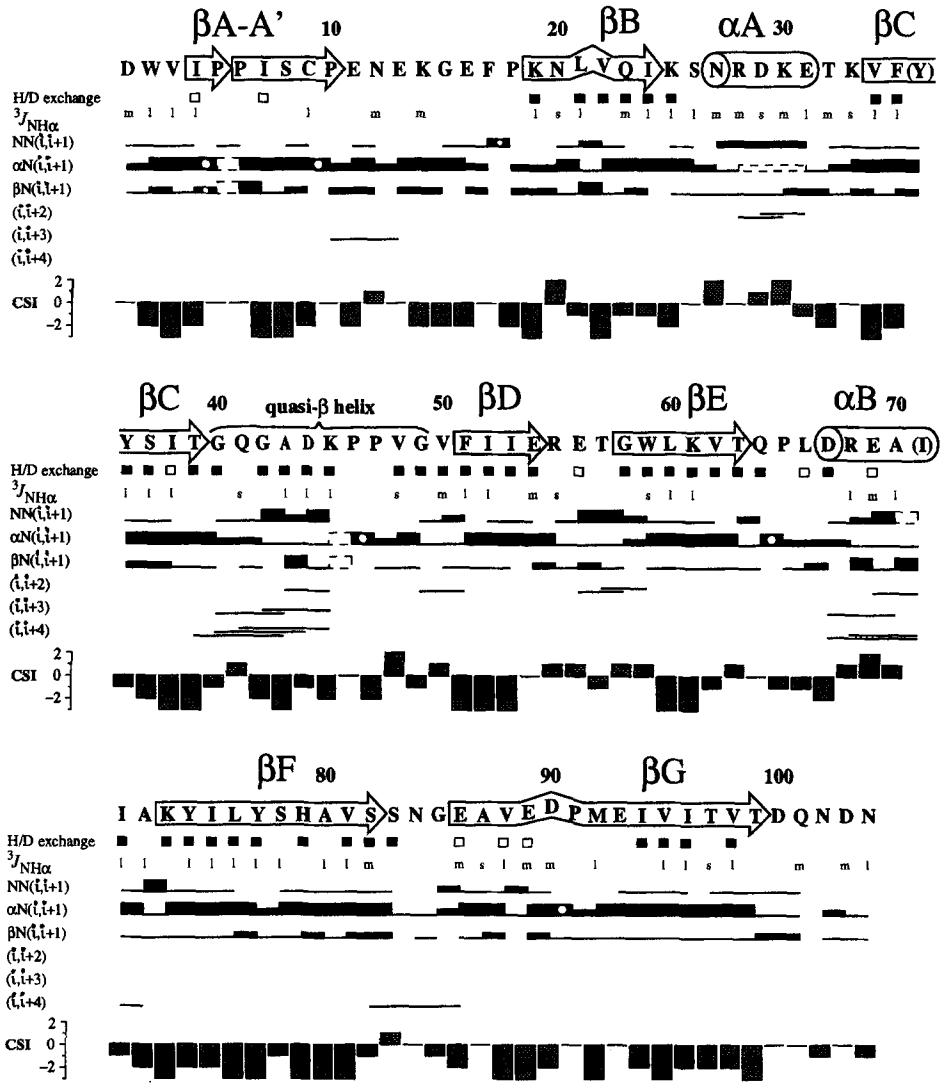

Fig. 5. Sequential connectivities, backbone amide proton exchange rates, $^{3}J_{HN\alpha}$ coupling constants and CSIs for Ecad1. The elements of secondary structure are depicted as arrows for β-strands and as cylinders for helices and are labeled. The sequence positions are indicated by numbers, followed by letters representing the one-letter code for the amino acid sequence. Beneath the sequence, amide proton exchange data are shown with full boxes indicating residues with slowly exchanging backbone NHs (> 38 min) and empty boxes indicating intermediate exchange (between 8 and 38 min). Large (> 8 Hz), medium (6–8 Hz) and small (< 6 Hz) $^{3}J_{NH\alpha}$ coupling constants are indicated by 'l', 'm', and 's', respectively. The rows labeled αN(i,i + 1), NN(i,i + 1) and βN(i,i + 1) summarize sequential NOE cross-peaks. The medium-range NOE cross-peaks between H^{α}, H^{β} and amide protons are consolidated into the (i,i + 2), (i,i + 3) and (i,i + 4) rows. High, medium and low bars represent strong, medium and weak sequential NOE cross-peaks, respectively. Empty boxes with dashed lines indicate ambiguous NOEs, while boxes with a white circle indicate NOEs involving proline H^{δ} instead of amide protons. Consensus CSIs ($CSI^{C^{\alpha}} + CSI^{C} - CSI^{H^{\alpha}}$) (Wishart and Sykes, 1994) are shown at the bottom. Residues for which insufficient chemical shift information is available to calculate a consensus index were assigned a value of zero.

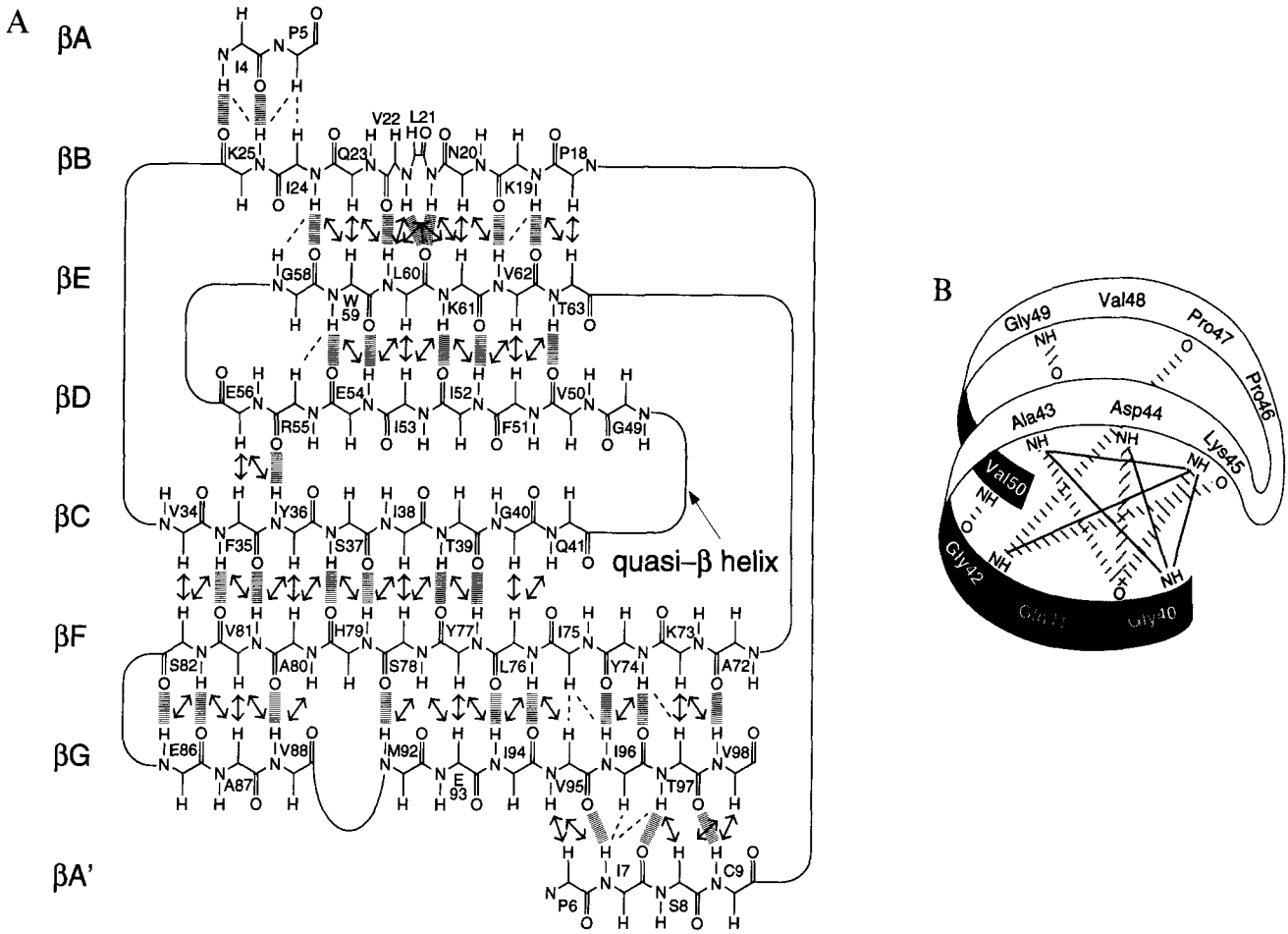

Fig. 6. Topology of the β-sheet and quasi-β-helix of EcadI. (A) The arrangement of the β-strands was deduced from interstrand NOEs between backbone protons and from amide proton exchange data. The backbone of the polypeptide is represented by a chemical bond drawing, with the sequence numbers of the amino acids indicated. Hydrogen bonds (hashed lines), long-range NOEs (arrows), and ambiguous NOEs (dashed lines) are shown. The β-strands are connected by lines and their notation is given on the right. The position of the quasi-β-helix is indicated. (B) The topology of the quasi-β-helix is shown schematically. Hydrogen bonds (hashed lines) and medium- or long-range NOEs between amide protons (solid lines) are shown. Only those backbone amide groups or carbonyl oxygens involved in NMR restraints are depicted. The cross-peaks for the NOEs can be seen in Fig. 4, and the view is similar to that depicted in Fig. 8.

spectra, respectively. The assignment of the side-chain resonances was completed with the aid of HCCH-TOCSY spectra. The most upfield proton resonance (−1.04 ppm) was assigned to a Gln^{23} H^{γ} that lies next to the aromatic ring of Trp^{61}. Main-chain ^{13}CO assignments were obtained from HNCO spectra. In total, assignments were obtained for 95.3% of the ^{1}H, 98.5% of the main-chain ^{15}N, and 88.0% of the NMR-observable ^{13}C resonances (Table 1).

Assignment of NOE cross-peaks

Four complementary sources of NOE data were acquired. The ^{15}N-edited NOESY-HMQC and simultaneous-acquisition ^{13}C-/^{15}N-edited NOESY-HSQC supplied the majority of NOEs involving backbone amide protons. Most of the NOEs involving side-chain protons were obtained from the ^{13}C-edited NOESY-HMQC spectrum collected in D_2O. Many NOEs to aromatic protons, resolved H^{α} protons and methyl groups were resolved in the 2D NOESY spectra collected in D_2O. Selected strips from the ^{15}N-edited NOESY-HMQC (Fig. 4) illustrate the method by which sequential NOEs were assigned. The two N-terminal and 40 C-terminal residues appear to be structurally disordered, since they do not exhibit detectable long- or medium-range NOEs and their narrow and intense resonances suggest conformational flexibility.

Secondary structure

Ten secondary structure elements were identified within residues 1–104 from NOE connectivity patterns, vicinal coupling constants between amide and α-protons ($^{3}J_{NH^{\alpha}}$), amide proton exchange rates and CSIs (Fig. 5). The start and end points of the elements of secondary structure were defined from the hydrogen-bonding patterns and the Kabsch and Sander algorithm (Kabsch and Sander, 1979). Seven β-strands (denoted βA through βG) are characterized by strong sequential NOE cross-peaks between H^{α} and NH protons ($d_{\alpha N}$) and weak d_{NN} NOEs between NH protons of adjacent residues. Moreover,

large $^3J_{NH^\alpha}$ values of these residues were consistent with ϕ torsion angles typical of β-strands. The interruption of the extended conformation in the βB and βG strands by β-bulges and in the βA-βA' junction was apparent from deviations in the secondary structure indicators. For example, weak $d_{\alpha N}(21,22)$ and medium $d_{NN}(21,22)$ NOE cross-peaks (Fig. 5) are found at the βB bulge. The identification of a novel three-residue β-bulge in βG, which contains a proline, was based primarily on the pattern of interstrand NOEs (Fig. 6A). The presence of a pair of five-residue helices (αA and αB) was revealed by strong or medium $d_{NN}(i,i+1)$ NOE cross-peaks, weak $d_{\alpha N}(i,i+1)$ cross-peaks, and small $^3J_{NH^\alpha}$ values. The αB helix contains a $d_{\alpha N}(i,i+4)$ NOE characteristic of an α-helix, although irregularities in the backbone conformation are suggested by large $^3J_{NH^\alpha}$ coupling constants from residues Arg^{68} and Ala^{70}. Backbone chemical shifts are correlated with secondary structure (Spera and Bax, 1991; Wishart et al., 1991; Wishart and Sykes, 1994) and generally confirm the location of the β-strands and helices in Ecad1 (Fig. 5), although the short helices are dominated by end effects. Deviations from the chemical shifts characteristic of β-strands are also found at the β-bulges.

A novel secondary structure motif, termed a quasi-β-helix (Shapiro et al., 1995a), is present in the βC-βD loop (Fig. 6B). The unusual conformation of this loop is reflected by a high density of short-range NOEs, including medium $d_{NN}(i,i+4)$ and weak $d_{NN}(i,i+5)$ NOEs (Figs. 4 and 6B), which are not found in regular helices. Furthermore, the $^3J_{NH^\alpha}$ coupling constants found in this region are rather larger than typically found in a helical backbone conformation. The residues in the quasi-β-helix have chemical shifts similar to those found in β-strands. Amide protons with very slow exchange rates (>4 h at 23 °C) are found in this region (Fig. 5), indicating that they form hydrogen bonds to stabilize the quasi-β-helix in solution.

The β-sheet topology was elucidated from long-range NOEs and the pattern of solvent-protected backbone amide protons (Fig. 6A). All β-strands are antiparallel to one another, except for βA/βG. The strongest pairings are found within the βBED and βCFG sheets. The βC-βD pairing has been improved through the identification of additional NOEs since the original structure determination (Overduin et al., 1995). The two hydrogen bonds connecting the βA and βB strands appeared in the calculated structures and are supported by the slow exchange of backbone amide protons of Ile^4 and Lys^{25}. However, these two hydrogen bonds were not included in the structure calculations since they are supported only by side-chain–side-chain NOEs.

Tertiary structure

The calculation of the structure of residues 1 to 104 employed a total of 1915 NOE restraints, 59 pairs of hydrogen-bond restraints and 155 dihedral angles. The addition of 337 more NOEs and 68 more dihedral angle restraints contributed to the lower root-mean-square deviations (rmsd) and energies of the current structures compared to the earlier generation (Overduin et al., 1995). The 20 structures with the lowest NOE energies were selected from the 77 that were calculated. Superposition of the backbone atoms of secondary structure elements of the 20 best structures (Fig. 7A) yields average rmsd values of 0.48 Å for backbone atoms and 0.85 Å for nonhydrogen atoms of residues found in regular secondary structure elements (Table 2). All of the (ϕ,ψ) angles of residues 1 to 100 of the 20 structures fall within the allowed region of the Ramachandran plot, with 76.4% lying in the energetically most favorable region as defined by the Procheck program (Morris et al., 1992).

TABLE 2
STRUCTURAL STATISTICS OF THE 20 FINAL STRUCTURES OF Ecad1

Parameter	<SA>[a]
Rmsd's from experimental restraints[b]	
Rms distance deviations (Å)	
All (1915)	0.0104 ± 0.0006
Interproton distances (1917)	
Intraresidue (546)	0.0053 ± 0.0024
Interresidue sequential ($\|i-j\|=1$) (654)	0.0064 ± 0.0004
Interresidue short-range ($1<\|i-j\|\leq 5$) (169)	0.0066 ± 0.0040
Interresidue long-range ($\|i-j\|>5$) (548)	0.0118 ± 0.0005
Hydrogen-bond restraints (118)[c]	0.0287 ± 0.0003
Rms dihedral deviations (155)	0.5198 ± 0.0011
Deviations from idealized geometry[d]	
bonds (Å)	0.0078 ± 0.0001
angles (°)	2.0393 ± 0.0011
impropers (°)	1.0405 ± 0.0001
Energetic statistics (kcal mol^{-1})[e]	
E_{noe}	11.1 ± 1.1
E_{cdih}	2.6 ± 0.7
$E_{L\text{-}J}$[c]	−200.4 ± 19.6
Average atomic rmsd's (Å)[f]	
Residues 1 to 100	0.49[g], 0.84[h]
Residues in β-strands or helices	0.61[g], 0.98[h]

[a] <SA> represents the mean value for the 20 final simulated annealing superimposed structures.

[b] The rmsd values from the experimental and covalent geometric restraints used for the X-PLOR structure calculations are listed. The number of terms for the various restraints is given in parentheses. None of the structures exhibited distance violations greater than 0.21 Å or dihedral angle violations greater than 5.04°.

[c] For each hydrogen bond there are two restraints: $r_{NH\text{-}O}=1.9$–2.5 Å and $r_{NH\text{-}O}=2.9$–3.5 Å.

[d] Idealized geometries and Lennard-Jones van der Waals energies are based on CHARMm parameters (Brooks et al., 1983).

[e] Total distance restraint energy as calculated with a square-well NOE and torsion angle force constants of 50 kcal mol^{-1} Å^{-2} and 200 kcal mol^{-1} rad^{-2}, respectively.

[f] The average atomic rmsd values from the mean structure for C^α, C' and N atoms and nonhydrogen atoms are given for selected residues.

[g] Backbone only.

[h] All atoms.

The cad fold consists of a β-barrel including two helical segments. The barrel is composed of two β-sheets (βABED and βCFGA'), which are bridged on one end by a Pro-Pro junction between the βA and βA' strands and on the other end by a hydrogen bond between Tyr^{36} and Arg^{55} (Fig. 6A). Within the barrel, 20 conserved uncharged residues are buried that comprise the hydrophobic core (Fig. 7B). Trp^2 is the most N-terminal residue to contribute to the core, with its aromatic ring contacting Ile^{24}, Ser^{26}, Val^{34} and Tyr^{36}. The C-terminal β-strand, βG, is anchored into the core through the side chains of Ile^{94}, Ile^{96} and Val^{98}. These large hydrophobic residues delimit the ends of the structural domain. The linker region that connects to the second cad domain contains the DQNDN sequence and is oriented near the AB loop and the αB helix. Together, these elements form a negatively charged pocket consisting of conserved acidic residues, capable of binding Ca^{2+} ions (Fig. 7C).

The quasi-β-helix (Shapiro et al., 1995a) is a novel structural element found in the loop between the βC and βD strands of cad1 domains. Although it appears roughly helical in tertiary structure, its secondary structure is more reminiscent of an extended structure. Four hydrogen bonds are found within the wide, right-handed loop consisting of residues 40–45 (Fig. 6B). A second, left-handed loop (residues 47–50) follows a hairpin turn executed by a pair of proline residues and is held against the first loop by three hydrogen bonds. The quasi-β-helix is structurally well defined, due to the high density of hydrogen bonds and NOEs (Fig. 8A). The residues in the quasi-β-helix each have unique, nonrepeating roles. In particular, Gly^{40} and Gly^{42} are required in the right-handed loop since bulky side chains would interfere with the formation of their five hydrogen bonds. The hairpin turn requires a cis peptide bond that can best be provided by a proline residue. The charged side chains of Asp^{44} and Lys^{45} are directed into the solvent, while the more hydrophobic Ala^{43} and Val^{50} side chains point toward the protein (Fig. 8B).

Discussion and Conclusions

The majority of ^1H, ^{13}C and ^{15}N resonances of the N-terminal domain of E-cadherin have been assigned, evidence for its monomeric state has been presented, and its solution structure has been determined. This information provides a basis for examining the adhesive and Ca^{2+}-binding mechanisms of cadherins in solution. The seven-stranded β-barrel structure of Ecad1 is very similar to the crystal structure of the homologous N-terminal domain of N-cadherin (Ncad1) (Shapiro et al., 1995a). The amino acid sequences of the Ncad1 and Ecad1 domains share 41% identity. Superposition of the Ncad1 and Ecad12 (Nagar et al., 1996) dimer structures on the Ecad1 monomer structure (Fig. 9) yields a rmsd value of 1.4 Å in both cases, for 55 C^α positions within the well-defined β-sheet region (residues 7–10, 19–23, 34–54, 58–63, 72–81, and 92–99). The most significant differences between the structures are the conformations of the βA strand and the βB/βC loop. A total of 14 unambiguous long-range NOEs were identified between dispersed aromatic ring resonances of Trp^2 and side-chain protons of Ile^{24}, Ser^{26}, Val^{34}, and Tyr^{36}. These NOEs indicate that the βA strand and the βB/βC loop are in close proximity in the Ecad1 monomer (Fig. 9). This interaction also brings the βA and βB strands into an antiparallel alignment, in which a pair of hydrogen bonds forms between Ile^4 and Lys^{25} (Fig. 6A). In the crystallized Ncad1 dimer the βA strand instead crosses over to another molecule, burying Trp^2 in the neighboring hydrophobic core. No hydrogen bonds are formed in the Ncad1 dimer that are consistent with the slow amide proton exchange rates found for the residues corresponding to the Ecad1 Ile^4 and Lys^{25} residues. In the Ecad12 crystal structure, Trp^2 is disordered and does not appear to be positioned to play a role in the dimer formed in the presence of Ca^{2+}. Thus, conformational heterogeneity may be a feature of the extreme N-terminal cad1 residues.

The cadherin fold is topologically similar to the immunoglobulin fold (Overduin et al., 1995; Shapiro et al., 1995b), which is found in a large number of CAMs in the immune system (Bork et al., 1994). The immunoglobulin fold also contains a core of seven β-strands in the same antiparallel topology. Due to the preponderance of β-hairpins this fold may be particularly easy to form (Richardson, 1977), which is a definite asset in the extracellular environment. Also, β-barrels are ideal 'stacking units' for presenting the distal ends of CAMs out of the glycocalyx and render them available for cell–cell adhesion. Antiparallel packing of the first and last β-strands next to each other allows adhesive forces to travel straight through each domain without twisting the extracellular region. Like in the immunoglobulin fold, the βCFG face mediates adhesive interactions (Overduin et al., 1995). The His^{79} and Val^{81} residues, which have been implicated in adhesion (Blaschuk et al., 1990; Nose et al., 1990), are solvent-exposed in the βF strand (Fig. 7C). The nearby Ser^{78} is buried while Ser^{83} is exposed in the βF-βG loop. These exposed residues are in a position to engage in adhesive interactions.

A novel structural motif, the quasi-β-helix, is characterized by an extensive network of hydrogen bonds, a roughly helical appearance and an unusual backbone conformation. Its presence is likely to be highly sequence-dependent, requiring conformations accessible only to glycine and proline residues, including one *cis*-proline. Based on cadherin sequence comparisons, the conserved quasi-β-helix sequence is **Gly**-Pro-**Gly**-hydrophobic-**Asp**-polar-Pro-*cis*-**Pro**-Xxx-Gly (where Xxx is any residue and bold-type residues are absolutely conserved). This motif

is found in most cad1 repeats, but not in other cad repeats. A functional significance for this motif in cadherin adhesion was inferred from its involvement in protein–protein interactions in the crystal lattice of Ncad1 (Shapiro et al., 1995a).

The Ca^{2+}-binding region is located at the C-terminal end of the molecule. Ca^{2+} binding induces chemical shift perturbations in residues including Glu^{11}, Asp^{67}, Glu^{69}, and Asp^{100}, Asp^{103}, Asp^{134} and Asp^{136}, which are found in the PENE, LDRE and DXNDN and DAD motifs (Overduin et al., 1995). Continuous chemical shift movements with only slight line broadening were observed in the backbone ^{1}H-^{15}N cross-peaks of the residues in these Ca^{2+}-binding motifs during Ca^{2+} titrations, consistent with a relatively low K_d of 1.6×10^{-4} (Tong et al., 1994). The Glu^{11} and Asp^{67} residues were also found to ligate Yb^{3+} and UO_2^{2+} in the Ncad1 structures, suggesting that Ca^{2+} is bound in a similar fashion by Ncad1. However, slight structural differences are evident. In Ncad1 the LDRE sequence in the βE-βF loop is extended and the DXNDN sequence following βG is disordered, while in Ecad1 these two elements are helical and structured, respectively. These differences can be attributed to experimental differences, since Ecad1 was studied in solution at pH 7.2 with Ca^{2+} while Ncad1 was crystallized at pH 4 or 5 with Yb^{3+} and UO_2^{2+}. Furthermore, in the Ecad12 crystal structure it is evident that two cad domains are required to form an integral binding site that binds three Ca^{2+} ions (Nagar et al., 1996). In Ecad12 the βE-βF loop is also helical and the linker following the βG strand is highly structured and intimately involved in Ca^{2+} ligation.

Unlike Ecad1, Ncad1 was reported to dimerize in a

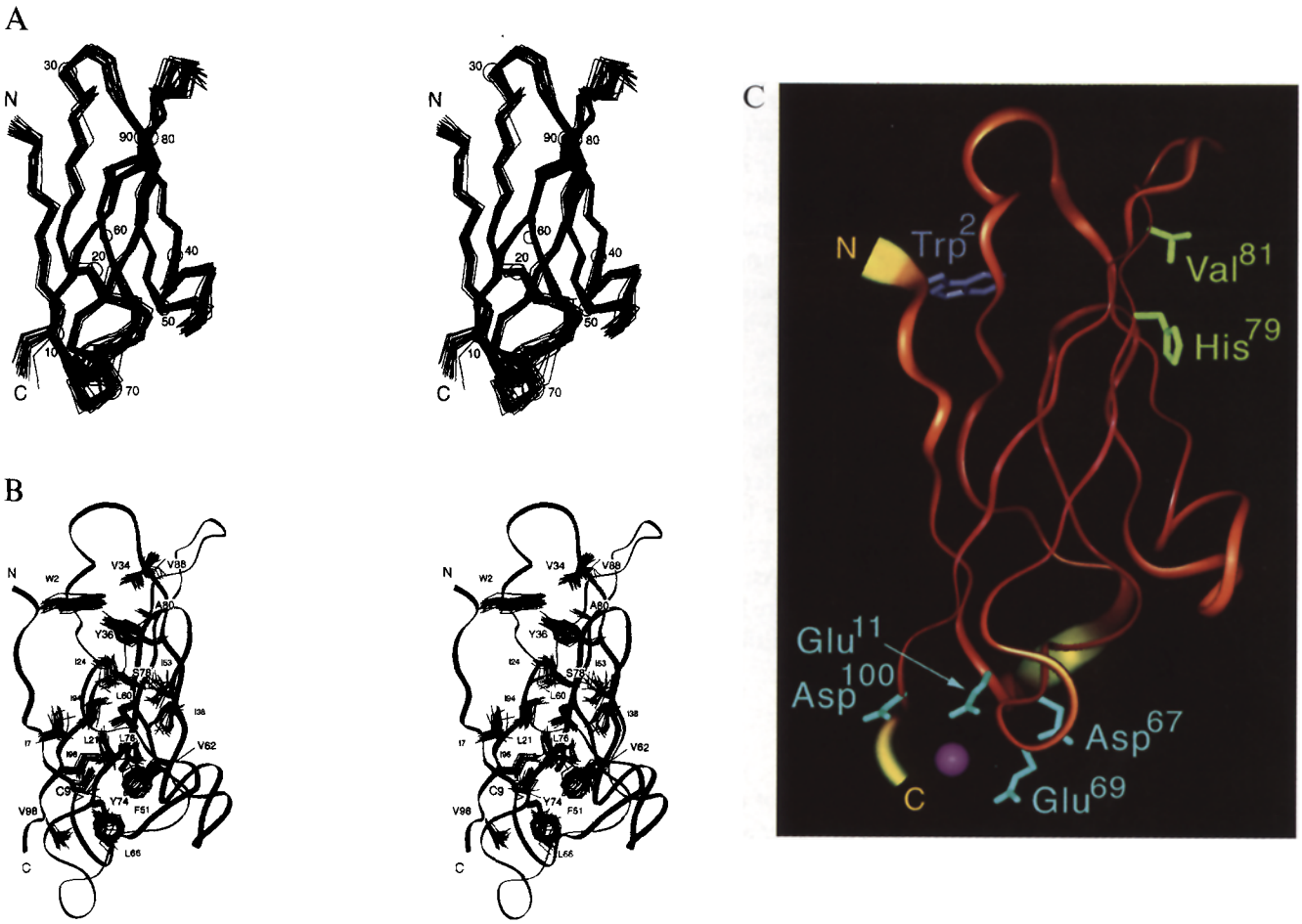

Fig. 7. Tertiary structure of Ecad1. (A) Stereoview of the superimposed C^{α} atoms of the 20 best structures calculated in X-PLOR. The following residues, which are located in secondary structure elements, were used for the alignment of the structures: Ile^{4}-Pro^{10}, Lys^{19}-Ile^{24}, Asn^{27}-Glu^{31}, Val^{34}-Thr^{39}, Phe^{51}-Glu^{54}, Gly^{58}-Thr^{63}, Asp^{67}-Ile^{71}, Lys^{73}-Ser^{82}, and Glu^{86}-Thr^{99}. The N- and C-termini are indicated with 'N' and 'C', respectively, and the C^{α} atom of every 10th residue is circled and labeled with its residue number. (B) Superposition of hydrophobic core residues. The side-chain heavy atoms of 20 residues contributing to the hydrophobic core are shown in a stereoview similar to that in (A). The buried residues are labeled with their type, size indicating depth. The backbone is depicted as a ribbon, with the N- and C-termini indicated. (C) Diagram of the structure closest to the average structure from the same view as in (A). The C^{α} trace is depicted as a red ribbon that becomes wider and more yellow at higher C^{α} rmsd values. The side-chain heavy atoms of two residues implicated in adhesion (His^{79} and Val^{81}) are shown in green, and the side chain of Trp^{2}, the position of which differs between the Ncad1 dimer and Ecad1 monomer structures, is shown in purple. Residues positioned to ligate Ca^{2+} (Glu^{11}, Asp^{67}, Glu^{69} and Asp^{100}) are shown with a magenta sphere, indicating the vicinity of bound Ca^{2+}. 'N' and 'C' denote the N- and C-terminal ends of the peptide chain. The picture was produced with the program InsightII (Biosym Technologies, San Diego, CA).

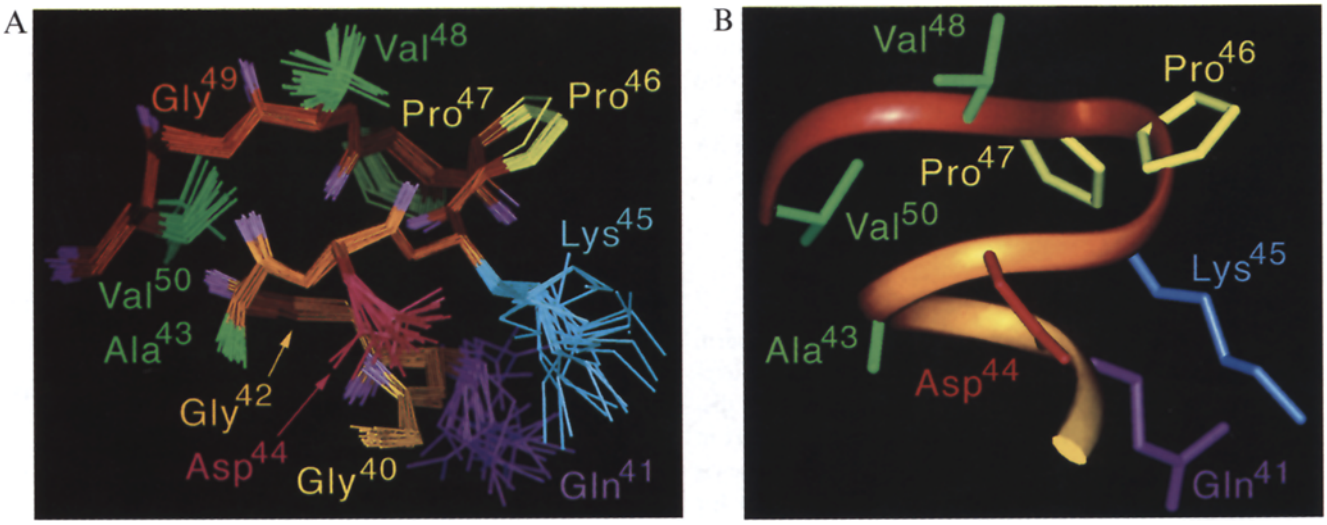

Fig. 8. Tertiary structure of the quasi-β-helix. (A) The backbone atoms of residues 40–50 of the 20 structures are superimposed and shown in orange. The non-hydrogen atoms are colored yellow (Pro), green (Val and Ala), blue (Lys), purple (Gln) and red (Asp). The carbonyl oxygens are shown in lavender. (B) The side-chain heavy atoms of residues 40–50 for the structure closest to the average, with the backbone depicted as a yellow/red ribbon. Side-chain heavy atoms are colored as in (A).

Ca^{2+}-independent fashion in solution (Shapiro et al., 1995a). The dimerization of Ncad1 (containing residues 1 to 108) reportedly involves the intermolecular packing of Trp^2 residues. The studies of Ecad12 (containing residues 1 to 224) demonstrate that the dimerization of E-cadherin requires Ca^{2+}, a pair of tandem cadherin repeats

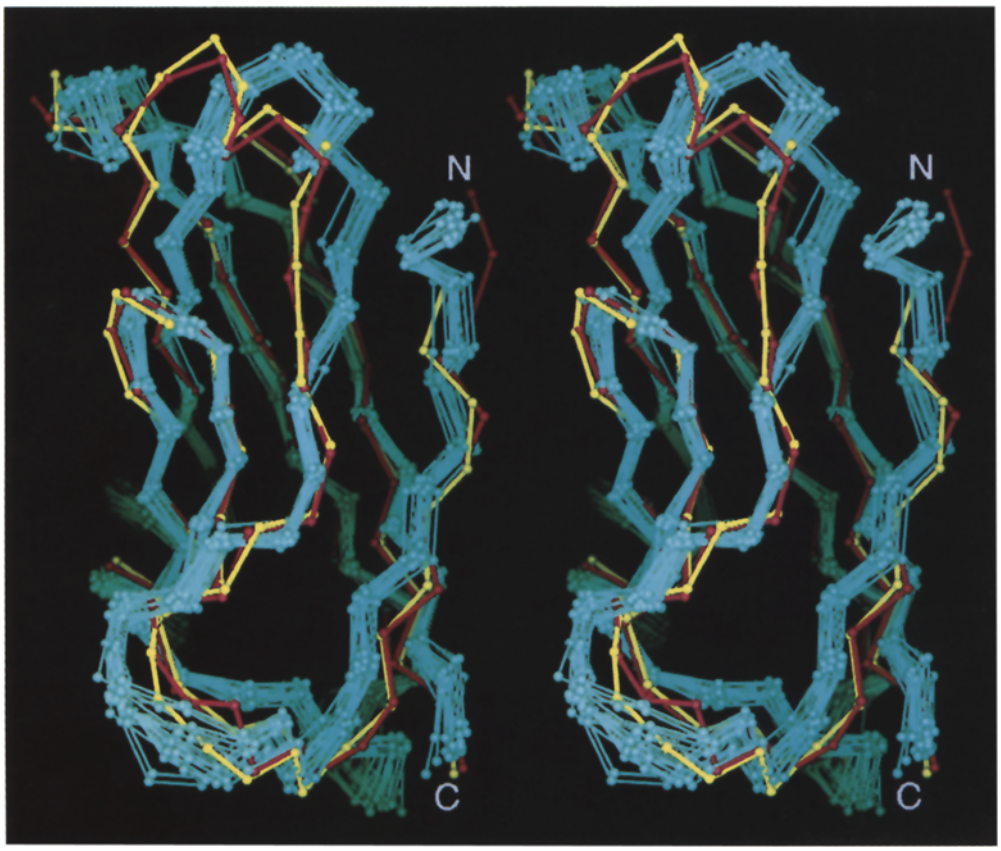

Fig. 9. Superposition of the Ecad1 and Ncad1 structures. The C^{α} atoms of Ncad1, the first domain of Ecad12 and the 20 Ecad1 structures are traced in red, yellow and blue, respectively. The 38 C^{α} atoms of common β-strand residues were superimposed. The stereoview is rotated about the vertical axis by approximately 180° compared to the view seen in Fig. 7, showing the position of the βA strand and the βB-βC loop, which includes the αA helix.

and millimolar protein concentrations (Nagar et al., 1996). The monomeric state of Ecad1 (containing residues 1 to 144) is likely due to truncation of most of the second domain. The minor differences between these structures, involving the Ca^{2+}-binding site and the position of the βA strand and Trp^2, can be accounted for by differences in the metal ligation and oligomeric state.

Acknowledgements

We thank Masatoshi Takeichi for the E-cadherin cDNA clone and continued discussions and encouragement, Jim Rini and Bhushan Nagar for stimulating discussions, Leslie D. Hicks for expert technical support in light scattering and analytical ultracentrifugation, Lewis Kay for providing pulse sequences, Frank Delaglio for providing nmrPipe and nmrDraw, and Dan Garrett for supplying the programs Pipp, Capp, and Stapp. We also thank Tim Harvey and Stefan Bagby for their assistance during the initial stage of this project. This work was supported by a grant to M.I. from the National Cancer Institute of Canada (NCIC). M.O. is supported by a postdoctoral fellowship from NCIC. M.I. is a recipient of a Medical Research Council of Canada Scholarship.

References

Archer, S.J., Ikura, M., Sporn, M.B., Torchia, D.A. and Bax, A. (1991) *J. Magn. Reson.*, **95**, 636–641.

Babul, J. and Steelwagen, E. (1969) *Anal. Biochem.*, **28**, 216–221.

Barkhuijsen, H., De Beer, R., Bovee, W.M.M.J. and Van Ormondt, D. (1985) *J. Magn. Reson.*, **61**, 465–481.

Bax, A., Mehlkopf, A.F. and Smidt, J. (1979) *J. Magn. Reson.*, **35**, 167–169.

Bax, A., Clore, G.M. and Gronenborn, A.M. (1990) *J. Magn. Reson.*, **88**, 425–431.

Bax, A. and Pochapsky, S. (1992) *J. Magn. Reson.*, **99**, 638–643.

Birchmeier, W. (1995) *Bioessays*, **17**, 97–99.

Blaschuk, O.W., Sullivan, R., David, S. and Pouliot, Y. (1990) *Dev. Biol.*, **139**, 227–229.

Bodenhausen, G. and Ruben, D.J. (1980) *Chem. Phys. Lett.*, **69**, 185–189.

Bork, P., Holm, L. and Sanders, C.J. (1994) *J. Mol. Biol.*, **242**, 309–320.

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

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

Brünger, A.T. (1992) *X-PLOR Manual*, v. 3.1, Yale University Press, New Haven, CT.

Clore, G.M., Nilges, M., Sukumaran, D.K., Brünger, A.T., Karplus, M. and Gronenborn, A.M. (1986) *EMBO J.*, **5**, 2729–2735.

Delaglio, F. (1993) *NMRPipe System of Software*, National Institutes of Health, Bethesda, MD.

Farmer, B.T., Venters, R.A., Spicer, L.D., Wittekind, M.G. and Müller, L. (1992) *J. Biomol. NMR*, **2**, 195–202.

Garrett, D.S., Powers, R., Gronenborn, A.M. and Clore, G.M. (1991) *J. Magn. Reson.*, **95**, 214–220.

Grzesiek, S. and Bax, A. (1992) *J. Magn. Reson.*, **96**, 432–440.

Grzesiek, S., Anglister, J. and Bax, A. (1993) *J. Magn. Reson. Ser. B*, **101**, 114–119.

Ikura, M., Kay, L.E., Tschudin, R. and Bax, A. (1990) *J. Magn. Reson.*, **86**, 204–209.

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

Kabsch, W. and Sander, C. (1979) *Biopolymers*, **22**, 2577–2637.

Kay, L.E. and Bax, A. (1990) *J. Magn. Reson.*, **86**, 110–126.

Kay, L.E., Ikura, M., Tschudin, R. and Bax, A. (1990) *J. Magn. Reson.*, **89**, 496–514.

Kay, L.E., Keifer, P. and Saarinen, T. (1992) *J. Am. Chem. Soc.*, **114**, 10663–10665.

Kay, L.E. (1993) *J. Am. Chem. Soc.*, **115**, 2055–2057.

Mareel, M., Bracke, M. and Van Roy, F. (1994) *Mol. Biol. Rep.*, **19**, 45–67.

Marion, D., Driscoll, P.C., Kay, L.E., Wingfield, P.T., Bax, A., Gronenborn, A.M. and Clore, G.M. (1989a) *Biochemistry*, **28**, 6150–6156.

Marion, D., Ikura, M. and Bax, A. (1989b) *J. Magn. Reson.*, **84**, 425–430.

Marion, D., Kay, L.E., Sparks, S.W., Torchia, D.A. and Bax, A. (1989c) *J. Am. Chem. Soc.*, **111**, 1515–1517.

McCoy, M.A. and Mueller, L. (1992) *J. Magn. Reson.*, **98**, 674–679.

Morris, A.L., MacArthur, M.W., Hutchinson, E.G. and Thornton, J.M. (1992) *Proteins*, **12**, 345–364.

Muhandiram, D.R., Farrow, N.A., Xu, G.-Y., Smallcombe, S.H. and Kay, L.E. (1993) *J. Magn. Reson. Ser. B*, **102**, 317–321.

Nagafuchi, A., Shirayoshi, Y., Okazaki, K., Yasuda, K. and Takeichi, M. (1987) *Nature*, **329**, 341–343.

Nagafuchi, A. and Takeichi, M. (1988) *EMBO J.*, **7**, 3679–3684.

Nagar, B., Overduin, M., Ikura, M. and Rini, J. (1996) *Nature*, **380**, 360–364.

Nilges, M., Gronenborn, A.M., Brünger, A.T. and Clore, G.M. (1988) *Protein Eng.*, **2**, 27–38.

Nilges, M., Kuszewski, J.T. and Brünger, A.T. (1991) In *Computational Aspects of the Study of Biological Macromolecules by Nuclear Magnetic Resonance Spectroscopy* (Eds, Hoch, J.C., Poulsen, F.M. and Redfield, C.), Plenum Press, New York, NY, pp. 451–455.

Nose, A., Tsuji, K. and Takeichi, M. (1990) *Cell*, **61**, 147–155.

Oda, T., Kanai, Y., Oyama, T., Yoshiura, K., Shimoyama, Y., Birchmeier, W., Sugimura, T. and Hirohashi, S. (1994) *Proc. Natl. Acad. Sci. USA*, **91**, 1858–1862.

Overduin, M., Harvey, T.S., Bagby, S., Tong, K.I., Yau, P., Takeichi, M. and Ikura, M. (1995) *Science*, **267**, 386–389.

Ozawa, M. and Kemler, R. (1990) *J. Cell Biol.*, **111**, 1645–1650.

Ozawa, M., Engel, J. and Kemler, R. (1990a) *Cell*, **63**, 1033–1038.

Ozawa, M., Hoschutzky, H., Herrenknecht, K. and Kemler, R. (1990b) *Mech. Dev.*, **33**, 49–56.

Ozawa, M., Ringwald, M. and Kemler, R. (1990c) *Proc. Natl. Acad. Sci. USA*, **87**, 4246–4250.

Palmer, A.G., Fairbrother, W.J., Cavanagh, J., Wright, P.E. and Rance, M. (1992) *J. Biomol. NMR*, **2**, 103–108.

Pascal, S., Muhandiram, T., Yamazaki, T., Forman-Kay, J.D. and Kay, L.E. (1994) *J. Magn. Reson. Ser. B*, **101**, 197–201.

Powers, R., Gronenborn, A.M., Clore, G.M. and Bax, A. (1991) *J. Magn. Reson.*, **94**, 209–213.

Powers, R., Garrett, D.S., March, C.J., Frieden, E.A., Gronenborn, A.M. and Clore, G.M. (1993) *Biochemistry*, **32**, 6744–6762.

Rance, M., Wagner, G., Sørensen, O.W., Wüthrich, K. and Ernst, R.R. (1984) *J. Magn. Reson.*, **59**, 250–261.

Redfield, C. and Dobson, C.M. (1990) *Biochemistry*, **29**, 7201–7214.

Richardson, J. (1977) *Nature*, **268**, 495–500.

Ringwald, M., Schuh, R., Vestweber, D., Eistetter, H., Lottspeich, F., Engel, J., Dolz, R., Jahnig, F., Epplen, J., Mayer, S., Müller, C. and Kemler, R. (1987) *EMBO J.*, **6**, 3647–3653.

Risinger, J.I., Berchuck, A., Kohler, M.F. and Boyd, J. (1994) *Nat. Genet.*, **7**, 98–102.

Scanlon, M.J. and Norton, R.S. (1994) *Protein Sci.*, **3**, 1121–1124.

Shaka, A.J., Keeler, J., Frenkiel, T. and Freeman, R. (1983) *J. Magn. Reson.*, **52**, 335–338.

Shaka, A.J., Lee, C.J. and Pines, A. (1988) *J. Magn. Reson.*, **77**, 274–293.

Shapiro, L., Fannon, A.M., Kwong, P.D., Thompson, A., Lehmann, M.S., Grubel, G., Legrand, J.F., Als-Nielsen, J., Colman, D.R. and Hendrickson, W.A. (1995a) *Nature*, **374**, 327–337.

Shapiro, L., Kwong, P.D., Fannon, A.M., Colman, D.R. and Hendrickson, W.A. (1995b) *Proc. Natl. Acad. Sci. USA*, **92**, 6793–6797.

Smith, L.J., Sutcliffe, M.J., Redfield, C. and Dobson, C.M. (1991) *Biochemistry*, **30**, 986–996.

Spera, S. and Bax, A. (1991) *J. Am. Chem. Soc.*, **113**, 5490–5492.

Takeichi, M. (1990) *Annu. Rev. Biochem.*, **59**, 237–252.

Takeichi, M. (1991) *Science*, **251**, 1451–1455.

Tong, K.I., Yau, P., Overduin, M., Bagby, S., Porumb, T., Takeichi, M. and Ikura, M. (1994) *FEBS Lett.*, **352**, 318–322.

Vuister, G.W. and Bax, A. (1992) *J. Magn. Reson.*, **98**, 428–435.

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

Williamson, M.P., Havel, T.F. and Wüthrich, K. (1985) *J. Mol. Biol.*, **182**, 295–315.

Wishart, D.S., Sykes, B.D. and Richards, F.M. (1991) *J. Mol. Biol.*, **222**, 311–333.

Wishart, D.S. and Sykes, B.D. (1994) *Methods Enzymol.*, **239**, 363–392.

Wittekind, M. and Mueller, L. (1993) *J. Magn. Reson.*, **101**, 201–205.

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

Wyatt, P.J. (1993) *Anal. Chim. Acta*, **272**, 1–40.

Yamazaki, T., Forman-Kay, J.D. and Kay, L.E. (1993) *J. Am. Chem. Soc.*, **115**, 11054–11055.

Zuiderweg, E.R.P., Boelens, R. and Kaptein, R. (1985) *Biopolymers*, **24**, 601–610.

Zuiderweg, E.R.P. and Fesik, S.W. (1989) *Biochemistry*, **28**, 2387–2391.