NMR Solution Structure of a Cytoplasmic Surface Loop of the Human Red Cell Anion Transporter, Band 3[†]

David Askin, Graham B. Bloomberg, Eric J. Chambers, and Michael J. A. Tanner*

Department of Biochemistry, School of Medical Sciences, Bristol University, Bristol, BS8 1TD, U.K.

Received December 23, 1997; Revised Manuscript Received April 17, 1998

ABSTRACT: The membrane domain of the human red cell anion transport protein, band 3, is too large to be studied by solution nuclear magnetic resonance spectroscopy (NMR), and its amphiphilic nature requires the use of detergents for solubilization. An alternative approach is to divide the protein into smaller (trans-membrane or surface loop) domains for NMR study. We report the structure of a 46-residue synthetic peptide that corresponds to the cytoplasmic surface loop connecting the putative 12th and 13th trans-membrane spans (residues 796−841) in the 14 span model of band 3. This peptide was shown by circular dichroism (CD) to be 38% helical in 30% trifluoroacetic acid. Two regions of helix (one close to the N-terminus of the peptide and one close to the C-terminus of the peptide) were identified by NMR. Long-range nuclear Overhauser effect (NOE) cross-peaks showed the two helices to be in near proximity. The helices were separated by a proline-rich loop that exhibited local order but was mobile with respect to the rest of the peptide. We discuss how the NMR structure of this loop fits the current models of band 3 structure and topology and the results of recent mutagenesis experiments. A cyclic version of this peptide was synthesized and studied by CD, but NMR studies were not possible due to the low solubility of this peptide.

Band 3 is the major integral membrane protein of the red cell. It is responsible for the rapid one-for-one exchange of chloride and bicarbonate ions across the cell membrane, an important step in the elimination of CO_2 from the body. Human band 3 (911 amino acids; ref *1*) possesses two distinct domains that can be isolated and purified following mild tryptic cleavage (*2*): the 40 kDa N-terminal domain (residues 1 to 360), located within the cytoplasm, acts as an anchor for the erythrocyte skeleton and binds several glycolytic enzymes (*3*, *4*); the 55 kDa C-terminal domain (residues 361−911) is membrane bound and is wholly responsible for the ion transport function of band 3 (*5*, *6*). For general reviews see refs *7*−*10*.

Hydropathy analysis, and information obtained from site-directed mutagenesis, proteolytic cleavage, and chemical modification studies, suggests that the membrane domain of human band 3 possesses up to 14 trans-membrane spans (*8*, *10*). However, a more recent study suggests that the protein may contain only 12 membrane-spanning regions (*11*). Though band 3 is relatively easy to purify from red cell membranes and forms 25% of the total membrane protein (1.2×10^6 copies per cell; ref *12*), as with many other integral membrane proteins, it has proved difficult to crystallize. Two-dimensional crystals of band 3 have been obtained, but only low resolution structures (up to 20 Å) have so far been published (*13*−*15*).

It has been shown by circular dichroism (CD)[1] that the membrane domain of band 3 is highly helical (60%), with the trans-membrane segments being almost entirely helical (*16*). Nuclear magnetic resonance spectroscopy (NMR) studies have shown that synthetic peptides corresponding in sequence to the first and second trans-membrane spanning segments take up a predominantly helical structure in organic solvent (*17*). The structure of band 3, like those of most other polytopic membrane proteins with a high helical content, is thought to be based on a tightly packed helical bundle with the loops that connect the helical spans extending outside the membrane. The protein exists as a mixture of dimers and tetramers in the red cell membrane, while the isolated membrane domain exists as a dimer (*18*). The functional significance of the different oligomeric states has not yet been elucidated.

The membrane domain of band 3 is too large to be studied directly by solution NMR, and its amphiphilic nature requires the use of detergents for solubilization. One approach that can be taken is to divide the protein into smaller domains amenable to study in solution by NMR. Polytopic proteins

[†] This work was supported by grants from the Wellcome Trust and is a publication from the Bristol Molecular Recognition Centre.

* Author to whom correspondence should be addressed. Telephone: (+) 44 (0) 117-9288271. Fax: (+) 44 (0) 117-9288274. E-mail: m.tanner@bris.ac.uk.

[‡] Atomic coordinates for the structure of peptide 12/13 determined by NMR have been deposited with the Protein Data Bank, Chemistry Department, Building 555, Brookhaven National Laboratories, P.O. Box 5000, Upton, Long Island, NY 11973-5000. The PDB File Name is 1BH7.

[1] Abbreviations: CD, circular dichroism; DANTE, delays alternating with nutations for tailored excitation; DQF−COSY, double quantum filtered correlation spectroscopy; HOHAHA, homonuclear Hartmann−Hahn spectroscopy; HPLC, high-pressure liquid chromatography; loop 12/13, the cytoplasmic loop that connects the 12th and 13th trans-membrane spans of human band 3; NMR, nuclear magnetic resonance spectroscopy; NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser effect spectroscopy; rmsd, root-mean-square deviation; TFA, trifluoroacetic acid; TFE, trifluoroethanol; TFE-d_3, perdeuterated trifluoroethanol.

S0006-2960(97)03158-9 CCC: $15.00
Published on Web 07/30/1998

can be considered as two distinct domains—the membrane spanning segments and the surface loops. In principle, if the structure of these sections can be determined separately, they can be combined to yield an integrated model for the structure of the protein. Recent results with fragments of integral protein suggest that both the trans-membrane helices (*17*) and surface loops of integral proteins can fold independently (*19*, *20*). We have synthesized peptides corresponding both to putative trans-membrane segments (*17*) and polar surface loops of band 3. Here we present the results of a CD and NMR study of a relatively large surface loop domain of band 3.

The surface loop that we have studied corresponds in sequence to the cytoplasmic loop (residues 796−841; loop 12/13) that connects the 12th and 13th trans-membrane spans of human band 3 in the 14 span model (*8*, *10*, *21*). This loop contains a number of histidine and lysine residues that have been shown to be required for anion transport (*22*). This loop also binds a monoclonal antibody (BRIC 132), and epitope mapping using BRIC 132 has demonstrated that residues 813−824 are located within the cytoplasm (*23*). The cytoplasmic location of this loop has been recently confirmed by N-glycosylation cassette mutagenesis (*11*). We report the results of a study of the structure of a synthetic peptide containing this loop.

MATERIALS AND METHODS

Peptides. The solid-phase synthesis of the peptides used the Fmoc (9-fluoenylmethoxycarbonyl) strategy on a Milligen 9050 automatic peptide synthesizer and a PEG-PS (poly(ethylene glycol) − polystyrene graft; 0.18 mmol/g substitution) resin. The linker used on the resin (PAL; 5-(4-(Fmoc)-aminoethyl-3,5-dimethoxyphenoxy) valeric acid) produced a peptide with a C-terminal amide following release from the resin. Trifluoroacetic acid (TFA)/ethanedithiol/water/phenol/anisol (92:3:2:2:1 v/v/v/w/v) mixture was used to cleave the peptides from the resin support. Following TFA removal by rotary evaporation, the peptides were collected by precipitation with diethyl ether.

The peptides were purified by reverse-phase high-pressure liquid chromatography (HPLC) on a Waters model 510 system using a Vydac C-18 column (20 cm × 22 mm) with an acetonitrile/water/0.1% TFA solvent system and a gradient of 35% to 55% acetonitrile.

The purity of the peptides was assessed by analytical reverse-phase HPLC on an Applied Biosystems model 140A system using a Vydac C-18 column (15 cm × 2.1 mm), and the identity of the peptide was authenticated by mass spectrometry on a VG Quattro triple quadropole electrospray ionization mass spectrometer and by NMR. Analytical reverse-phase HPLC of the purified peptide 12/13 gave a single broad peak. The identity of the peptide was authenticated by mass spectrometry (calcd *m*/*z* 5399.7; found *m*/*z* 5399.4). The identity of the peptide was later also confirmed following assignment of the NMR ^{1}H resonances. The analysis by mass spectrometry of the leading and trailing edges of the broad HPLC peak gave a *m*/*z* value of 5399; therefore, the peptide was chemically pure but heterogeneous in conformation or aggregation state. The purified dicysteine 47-residue peptide, before cyclization, also gave a single broad peak by analytical HPLC. Mass spectrometry confirmed the chemical purity of this peptide which was also heterogeneous in conformation or aggregation state.

Peptides with N- and C-terminal cysteine residues were cyclized by the air oxidation/diffusion method (*24*, *25*). Peptides were dissolved in 25 mM acetic acid at low concentration (0.01 mg/mL) to avoid polymer formation. The solution was adjusted to pH 8.0 with triethylamine and allowed to oxidize in air for 3 days without stirring. Analytical HPLC of the cyclized peptide gave a single peak with no evidence of polymer formation. Analysis by mass spectrometry showed a reduction of 2 mu, indicating that cyclization had occurred. Yields of cyclized product varied from 50% to 75%.

Circular Dichroism. CD spectra of freshly prepared peptide samples were recorded from 185 to 260 nm on a Jobin−Yvon dichrograph (model CD6) under a constant nitrogen flush (9 L/min). The peptides were studied at a concentration of 50 μM in 2 mm path length cuvettes. A temperature range from 25 to 50 °C, a pH range of 3.5 to 7.5, and a trifluoroethanol (TFE) concentration range of 0% to 75% was employed. Ellipticity is reported as mean residue ellipticity (deg cm^2 dmol^{-1}) and the helix content of the samples was estimated from the $[\theta]_{222\ nm}$ minimum (*26*).

Nuclear Magnetic Resonance Spectroscopy. For NMR studies, 5 mM peptide was dissolved in 12 mM phosphate, 30% TFE-d_3 (perdeuterated trifluoroethanol), and the pH*-(apparent, unadjusted for isotope effects) adjusted to 3.5 with small volumes of NaOH (0.05 M). Wilmad 5 mm sample tubes were used. Experiments were performed at temperatures of 25, 30, or 35 °C.

Proton NMR spectra were recorded on a JEOL Alpha 500-MHz spectrometer. A spectral width of 6000 Hz was used. Solvent suppression was by presaturation using the DANTE pulse sequence (*27*), and the transmitter offset was set equal to the irradiated solvent residue. The NMR spectra were referenced to the TFE signal at 3.88 ppm. The time proportional phase incremental method (*28*) was used for phase-sensitive acquisition.

Phase-sensitive double quantum filtered correlation (DQF−COSY) spectra (*29*−*31*) with the number of acquisitions typically set to 64 were used for spin system assignment. Homonuclear Hartmann−Hahn (HOHAHA) spectra (*32*−*34*) were acquired typically with 32 acquisitions and with mixing times of 150 or 300 ms. Phase-sensitive nuclear Overhauser effect (NOESY) spectra (*35*, *31*) were recorded with the number of acquisitions typically set to 64 and with a mixing time of either 75 or 150 ms. NOESY intensities, for use in the structure calculations, were estimated by counting cross-peak contour levels. Three classes of distance range were used: short (0−3.0 nm), medium (0−4.0 nm), and long (0−5.0 nm). The lower limits were explicitly set to zero (*36*).

Amide protons that slowly exchanged with the solvent (due to protection from exchange, for example by H-bonds) were identified. For this, the fully protonated peptide was freeze-dried and then dissolved in 12 mM phosphate, 30% TFE-d_3, 70% D_2O, pH* 3.5. Several one-dimensional and short (5 h) phase-sensitive two-dimensional NOESY spectra were acquired, and the changes in amide peak intensities were observed.

The NH−CαH coupling constants ($^3J_{HN-H\alpha}$) were measured from the DQF−COSY experiment. The NMR data

was zero filled in F2 to give a final digital resolution of 0.37 Hz. Estimates for the dihedral angles (ϕ) were obtained: for $^3J_{HN-H\alpha} < 6.0$ Hz, ϕ was set to $-60° \pm 30°$; for $^3J_{HN-H\alpha} > 8.0$ Hz, ϕ was set to $-120° \pm 40°$; and for $^3J_{HN-H\alpha} > 9.0$ Hz, ϕ was set to $-120° \pm 30°$.

All the spectra were processed with either a 60° shifted sine-bell function or a 45° shifted sine-bell squared function in both F1 and F2 using FELIX (Biosym Technologies) software.

Molecular Modeling. Structure calculations were performed using X-PLOR (*37*). Bond lengths were constrained by the SHAKE algorithm (*38*). Initially only short-range NOESY constraints ($|i-j| \leq 4$) were included. A selection criterion of no NOE violations of >0.05 nm, no dihedral angle violations of >5°, and an arbitrary energy cutoff to remove misfolded structures was applied. Further structure refinement included longer range NOE constraints. The program PROCHECK (*39*) was used to obtain Ramachandran plots for the calculated structures.

RESULTS

The amino acid sequence of the 46-residue peptide (12/13) containing the sequence of the loop between the putative 12th and 13th trans-membrane spans is given below (peptide numbering given above the sequence and residue location in human band 3 protein given below the sequence):

```
     5    10    15    20    25    30    35    40    45
GVTSL SGIQL FDRIL LLFKP PKYHP DVPYV KRVKT WRMHL FTGIQ C-CONH2
   800   805   810   815   820   825   830   835   840
```

For studies of the cyclic version of this peptide, an additional cysteine residue was added on the N-terminal side of glycine 1. We included some of the putative transmembrane helix sequences of the 12th and 13th spans at the N- and C-termini of the peptide. We reasoned that this might help stabilize the peptide structure, for example by helix−helix interactions. The peptide was acetylated at the N-terminus and amidated at the C-terminus to remove main-chain charges not present in the native protein and to prevent any aggregation that might have occurred by end-to-end charge association.

Cyclization. Since the N- and C-termini of polytopic membrane protein loops must return in a similar direction toward the membrane, a cyclic form of the peptide was also examined. Cyclization was carried out by the formation of an intramolecular disulfide bond between cysteine residues introduced at the N- and C-termini of the peptide. Since cyclization would reduce the conformational space available for peptide 12/13, we reasoned that this may also help peptide folding.

CD Spectroscopy

Linear Peptide 12/13. The CD spectrum of the linear peptide in aqueous solution is shown in Figure 1. This spectrum in aqueous buffer was typical of a random coil structure with the sole feature being a fairly intense minimum at 200 nm. Upon addition of TFE, the spectra showed an increasing degree of helicity. Initially, as the TFE concentration was increased, the percentage helicity of the peptide (estimated from the minimum at 222 nm) increased as shown in Figure 1 (inset). At a TFE concentration of 30% the

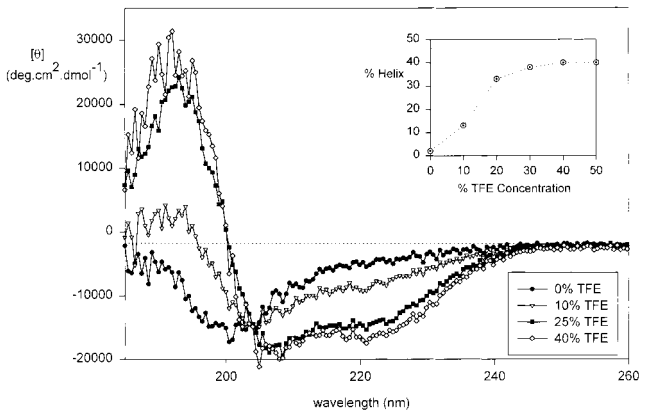

FIGURE 1: Circular dichroism spectra of peptide 12/13 titrated against TFE. Buffer: 12 mM sodium phosphate, pH 3.5. Temperature: 298 K. Peptide concentration: 50 μM.

percentage helicity obtained was 38%, and increasing the concentration of TFE had little effect on the CD spectrum. An isodichroic point was observed at 204 nm, indicating that the transition from extended coil to helix was by a two-state process (*40*). The partial helical structure taken up by peptide 12/13 was not due to aggregation since reducing the peptide concentration did not change the CD spectrum (results not shown). At 30% TFE concentration, increasing the temperature from 25 to 40 °C resulted in a 5% decrease in helicity (as estimated from the 222 nm minimum) with a further 3% decrease at 50 °C. The maximum at 193 nm showed a much larger fall (of 38% from 25 to 50 °C). The minimum at 222 nm is relatively insensitive to α-helix length, whereas the maximum at 193 nm decreases markedly with decreasing α-helix length (*41*). Thus the 8% decrease in the magnitude of the 222 nm minimum due to the decreased α-helix length resulted in a much larger fall in the magnitude of the 192 nm maximum. The effect of pH on the CD spectrum of peptide 12/13 was also examined. The spectra changed little from neutral (pH 7.5) to acidic (pH 3.5) pH (results not shown). Salts (sodium chloride or sodium phosphate) to 100 mM also had little affect on the CD spectra (results not shown).

Cyclic Peptide 12/13. The CD spectrum of cyclic peptide 12/13 in aqueous solution was typical of a random coil structure, similar to that observed for the linear peptide (results not shown). The peptide showed an increase in structure upon TFE addition and this occurred at lower TFE concentrations than observed for the linear peptide. The maximum amplitudes in the CD spectra for the cyclic peptide were reached at a TFE concentration of 25% (helicity = 37%). At 25% TFE concentration, the CD spectrum of the cyclic peptide was very similar to the spectrum of the linear peptide at 30% TFE.

NMR Studies

Cyclic Peptides. We attempted to study the cyclic 12/13 peptide in 25% TFE-d_3, but due to the low solubility of this peptide only poor NMR spectra were obtained which were unsuitable for NMR analysis. In an attempt to increase the solubility of the cyclic peptide we added the four-residue sequence CGRG to the N-terminal glycine of the linear 12/13 peptide. The glycine residues served as spacers (so that the *N*- and C-termini were not too tightly restrained), and

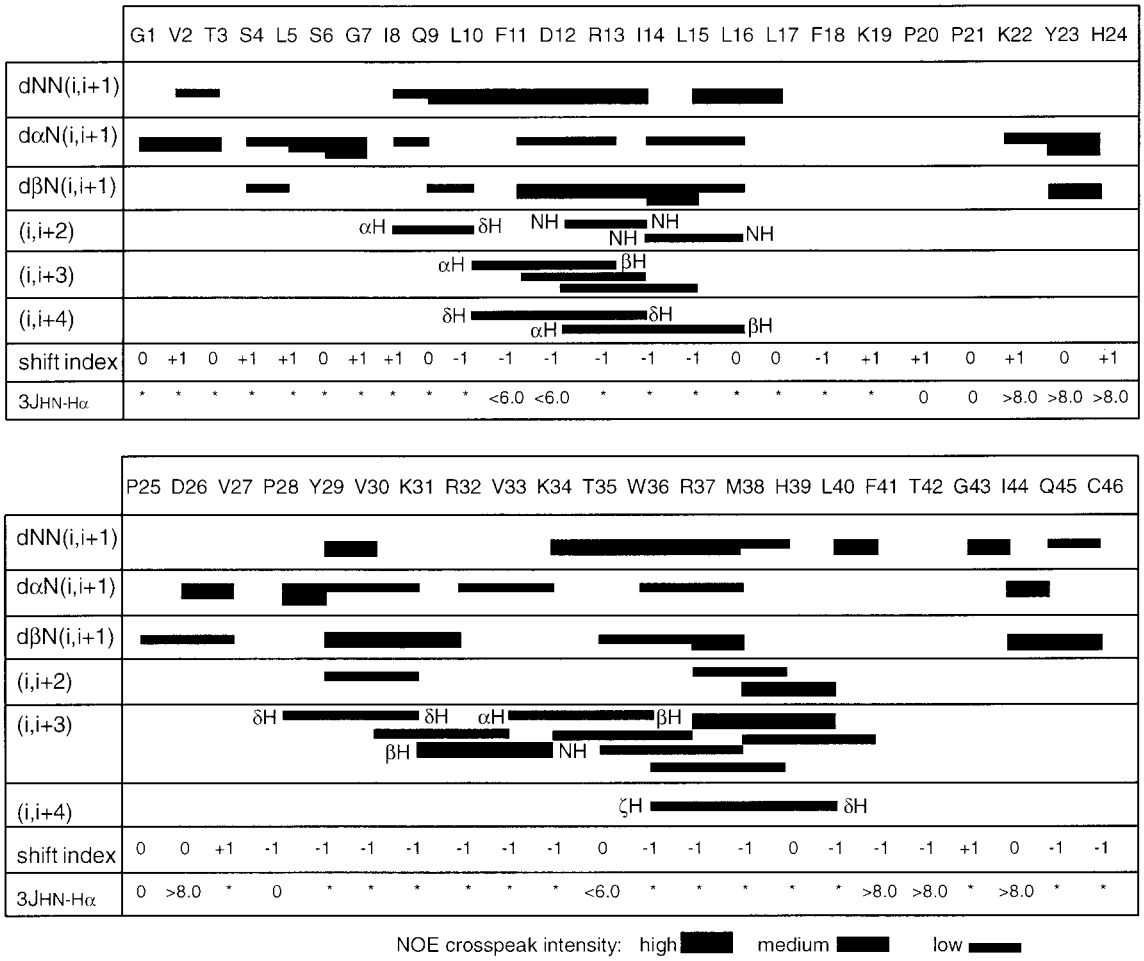

FIGURE 2: The sequential and medium-range NOE cross-peaks, the chemical shift indices (*43*), and the backbone coupling constants [$^3J_{HN-H\alpha}$ (Hz)] of peptide 12/13. All nonsequential NOE cross-peaks are dαN except where stated. Backbone coupling constants between 6 and 8 Hz are marked with an asterisk (*).

the arginine residue was included to increase solubility. This cyclic peptide was also of very low solubility and unsuitable for NMR study. A further cyclic peptide with the arginine replaced with a phosphorylated serine residue was also synthesized. This strategy was also unsuccessful in increasing the solubility of the cyclic peptide.

Linear Peptide 12/13. Peptide 12/13 could be solubilized in aqueous solution to 27 mg/mL (5 mM).

Assignment of the proton resonances of peptide 12/13 in 30% TFE-d_3 was done by the sequential assignment method (*42*) using COSY, HOHAHA, and NOESY spectra. The sequential assignments were supported by the presence of several dαN(i,i+1), dβN(i,i+1) and dNN(i,i+1) cross-peaks as shown in Figure 2. NOE spectra were acquired at three temperatures (25, 30, and 35 °C) and this helped in the assignment of overlapping peaks within crowded areas (with many of the overlapping peaks shifting enough to enable a confident assignment). A shorter peptide, representing residues I8 to G43 of peptide 12/13, was also synthesized and fully assigned. The assignment of this peptide was useful in resolving any ambiguities in the assignment of linear peptide 12/13. As found for linear peptide 12/13, two stretches of helix, one close to the N-terminus and one close to the C-terminus, were also present in this shorter peptide. The amide-proton region of the NOE spectrum of linear peptide 12/13, at 25 °C, using a mixing time of 150 ms, is shown in Supporting Information.

It was possible to fully assign all but one (lysine 19) of the 46 residues (see Supporting Information). The proton resonances of peptide 12/13 in 30% TFE-d_3 are given in Supporting Information. The number of NOE constraints per residue obtained for peptide 12/13 is given in Figure 3. There was a total of 245 intraresidue, 167 sequential, and 68 nonsequential NOE constraints. Two regions of sequence showed nonsequential NOE cross-peaks typical of a helical conformation (for example (i,i+3) connectivities; see Figure 2): I8 to L15 and Y29 to F41. Helix formation in these areas was also indicated by the presence of medium-intensity sequential NN(i,i+1) cross-peaks (extending from Q9 to I14, from L15 to L17, and between Y29 and V30 and from K34 to M38) and an absence of medium or strong sequential αN(i,i+1) cross-peaks. There were also several (i,i+2) cross-peaks in the two helical regions and (i,i+4) cross-peaks between L10 and I14, D12 and L16, and W36 and L40.

There was an absence of strong (nonsequential) NOE cross-peaks for the remaining residues of peptide 12/13. Nine long-range NOE cross-peaks were observed: one each from residue L5 to residues K31, R32, and V33; two each from residue D12 to residues V27 and P28; and two from residue R13 to residue V27 (see Table 1). These were all of weak intensity except for the D12 (Hβ1) to P28 (Hα) cross-peak which was of medium intensity. NOE spectra of linear peptide 12/13, at 25 °C, and with a mixing time of 150 ms,

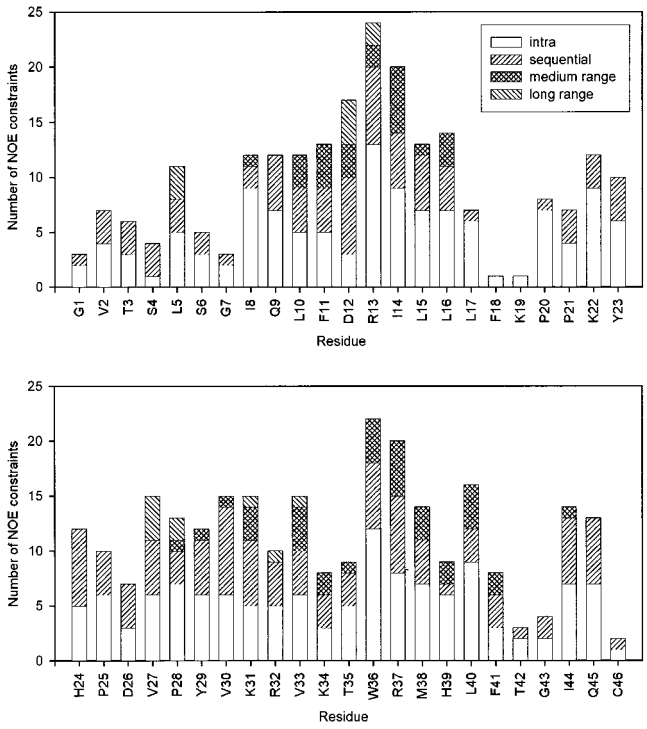

FIGURE 3: The number of intraresidue, sequential, and long-range NOE constraints per residue for peptide 12/13.

Table 1: Long-Range NOE Connectivities

residues	protons
L5 to K31	α to δ*
L5 to R32	α to N
L5 to V33	δ to N
D12 to V27	β1 to α; β2 to α
D12 to P28	β1 to δ; β2 to δ1
R13 to V27	N to α; β1 to α

with the nine long-range NOE cross-peaks highlighted (2 pages) are given in Supporting Information.

Though the coupling constants of most of the amide resonances fell within the 6−8 Hz range, 10 values of $^3J_{HN-H\alpha}$ fell outside this range (see Figure 2) and the corresponding dihedral angle restraints of these were included in the structure calculations.

The amide protons of residues K34 and R37 were protected from exchange with the solvent: the 1D NMR NH peak intensities decreased little over the length of the solvent exchange experiment of 2 days. No other amide-proton resonances remained after the first 5 h.

It has been reported that there is a relationship between protein secondary structure and α-^{1}H chemical shift (*43*). An upfield shift in the α-proton resonance (from the random coil value) is observed for all 20 naturally occurring amino acids when placed in a helical configuration, and a downfield shift occurs when placed in a β-sheet configuration. The chemical shift index (*44*) of each residue of peptide 12/13 is given in Figure 2. The chemical shift indices of residues 10−15 (of −1) indicate helix formation. Both L16 and L17 have indices of value 0 suggesting that helix termination occurs here. This is in agreement with the data from the NOESY experiments with two dNN(i,i+2) (D12 to I14 and I14 to L16), three (i,i+3) cross-peaks (from L10 to R13, F11 to I14 and D12 to L15), and two (i,i+4) cross-peaks (from L10 to I14 and D12 to L16). The presence of a dαβ-(i,i+4) cross-peak and absence of any dαN(i,i+2) cross-peaks (the only (i,i+2) cross-peaks present are between amide protons) confirms this is a short stretch of α-helix. The chemical shift indices of P28 to T42 indicate helix formation (with only T35 and H39 not having indices of −1). This again agrees with the NOE data. There are three dαN(i,i+2) cross-peaks (Y29 to K31, R37 to H39, and M38 to L40), six dαN(i,i+3) cross-peaks (the first being V30 to V33, and the last M38 to F41), one dβN(i,i+3) (K31 to K34) cross-peak, one dαβ(i,i+3) (V33 to W36) cross-peak, and one (i,i+4) cross-peak (W36 to L40). In addition, the only amide protons protected from solvent exchange were in this region (K34 and R37). The presence of three dαN(i,i+2) NOE cross-peaks in the Y29 to F41 region and the absence of any dαN(i,i+4) cross-peaks suggests the presence of a 3_{10} helix. All the chemical shift indices of residues K19 to V27 had values of 0 or +1: no helix occurs here.

From the NOESY spectra it was possible to distinguish the configuration of three of the four proline residues. The NOE connectivity from the β-protons of proline 20 to the δ-protons of proline 21 (the only sequential cross-peak observed between the two residues) indicated that the amide bond of the major P21 isomer was in the trans configuration. There was a strong NOE cross-peak from the α-proton of H24 to the δ-protons of P25: the amide bond of the major P25 isomer was in the trans configuration. The medium intensity NOE cross-peak observed between the α-proton of V27 and the δ-protons of P28 indicated that here, also, the amide bond of the major P28 isomer was in the trans configuration.

Structure calculations were performed using X-PLOR (*37*) with a total of 480 distance restraints from the NOESY spectra, 10 dihedral angle restraints ($^3J_{HN-H\alpha}$) from the DQF−COSY spectra, and four H-bonding distance restraints (two distance restraints for each of the two H-bonds) from the amide exchange experiments. The two H-bonds identified from the exchange experiment fell within a stretch of residues calculated to form a 3_{10} helix. Further refinement was performed including N−O and HN−O distance constraints of 0−3.5 Å and 0−2.5 Å, respectively, for each H-bond.

Using the distance and dihedral angle restraints, fifteen low-energy structures were calculated. All these structures had empirical energies of less than 410 kcal/M and were in good agreement with the experimental restraints with no NOE violations >0.05 nm and no torsion angle violations greater than 5°. It was apparent that the loop exhibited local order and was mobile with respect to the two helices. The 15 structures are shown superimposed (backbones only) in Figure 4. Average structures were determined using X-PLOR: the backbone root-mean-square deviation (rmsd) for residues I8 to L15 is 1.1 Å; for residues L16 to D26 the backbone rmsd is 2.0 Å; and for residues V27 to L40 the backbone rmsd is 1.7 Å. The Ramachandran plots for each of the individual structures showed that nearly all the dihedral angles were in the allowed regions (data not shown). The residues with unfavorable $\phi-\psi$ angles were either in the flexible regions that connect the helices to the loop, or at the N- or C-termini. Figure 5 shows an example of a low-energy structure of peptide 12/13. The residues which

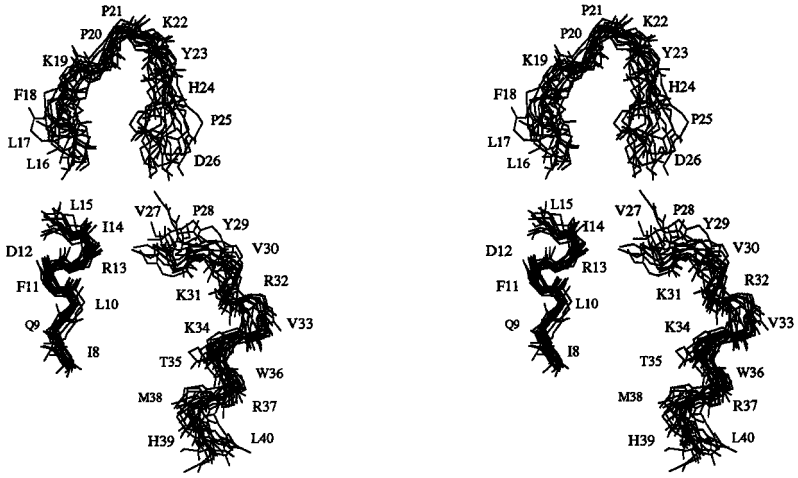

FIGURE 4: Stereoview of the 15 lowest energy structures. Residues I8 to L15, residues L16 to D26, and residues V27 to L40 are each shown superimposed.

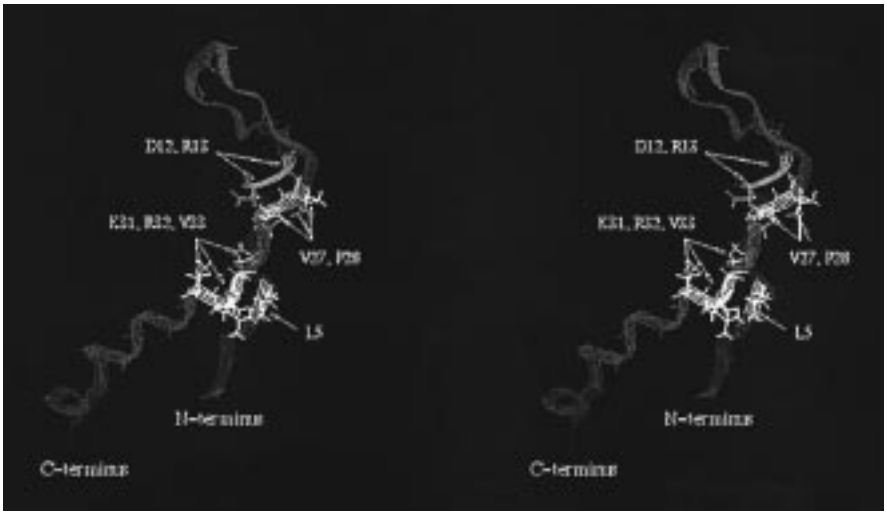

FIGURE 5: Stereoview of a low-energy structure for peptide 12/13. Residues exhibiting long-range NOE cross-peaks are highlighted.

exhibited long-range NOE cross-peaks (see Table 1) are highlighted and labeled.

DISCUSSION

The CD spectrum of peptide 12/13 in aqueous buffer showed features characteristic of an extended coil structure. Addition of TFE to 30% (v/v) increased the helicity of the peptide. Further increasing the TFE concentration had little effect on the CD spectrum, so we studied peptide 12/13 in 30% TFE. This alcohol is widely used as a structure-inducing cosolvent, and there are reports of TFE stabilizing both α-helix and extended *β*-sheet structure (*45*−*51*). The presence of salt (sodium phosphate or sodium chloride to 100 mM), change in pH from 7.5 to 3.5, or temperature from 25 to 40 °C, resulted in only small changes in the CD spectra.

The peptide was studied by NMR in 30% TFE-d_3, 12 mM sodium phosphate, pH 3.5, at temperatures of 25, 30, and 35 °C. This allowed certain features of the structure of peptide 12/13 in 30% TFE to be identified. A stretch of helix occurs at the N-terminus of the peptide (approximately from residue I8 to L15) and at the C-terminus (approximately from residue Y29 to F41). This conclusion is supported by sequential (dNN and dαN) and short-range (i,i+2), (i,i+3), and (i,i+4) NOE cross-peaks, and by amide exchange protection (K34 and R37), and the values (−1) of the chemical shift indices. It was shown by NMR that approximately 21 of the 46 residues (45%) are within a helix; this is a little greater than the estimate by CD of 38% helicity. The NMR data suggest I8 to L15 forms an α-helix, while Y29 to F41 is a 3_{10} helix. The presence of nontypical values for helices of $^3J_{HN-H\alpha}$ (>6.0 Hz except for residues F11, D12, and T35) could indicate that the helices are somewhat distorted, and this could account for some of the discrepancy in the calculated helix content between NMR and CD. In Figure 4 the distortion of the helices can be clearly seen (particularly for I8 to L10 and Y29 to R32). Only two of the amide protons within the helices were protected from exchange with the solvent: it is possible that the isolated peptide in aqueous solution is in rapid exchange between the structured helical/loop conformation identified by NMR and a much looser extended coil conformation. This may explain the high values of $^3J_{HN-H\alpha}$ obtained for the two helices. The presence of a number of long-range NOE cross-peaks (Table 1) between the two stretches of helix indicates the presence of tertiary structure. Figure 6 shows helical net diagrams for the two helices. The helix extending from residue I8 to L15 is represented as an α-helix, and the helix from residue Y29 to F41 is represented as a 3_{10} helix (these

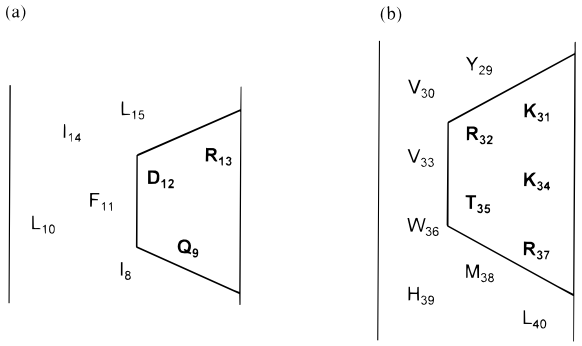

FIGURE 6: Helical net diagrams for the two helices: (a) helix I8 to L15; (b) helix Y29 to L40. Polar residues are given in bold type. As discussed in the text, I8 to L15 is represented as an α-helix, and Y29 to L40 as a 3_{10} helix.

representations are only approximate since these helices are, as discussed earlier, somewhat distorted). It can be seen that the polar residues are clustered, and the amphiphilic nature of these helices may be important for tertiary interactions that maintain the overall structure of this domain in intact band 3. However, it is clear from Figure 5 that in linear peptide 12/13 the two helices do not interact over a substantial part of their length. The structure of these regions of the helices are stabilized by intra-helix interactions. Inter-helical interactions occur between the two N-termini of these helices. There are a number of possible side-chain interactions that may be responsible for the folding of peptide 12/13 into the loop structure. Hydrophobic interactions between the side chains of L5 and V33 may be important. Residue W36 is on the same side of the C-terminal helix as V33; the indole ring of W36 may also take part in this interaction. Residues L800 and V828 (L5 and V33 of peptide 12/13) are conserved for all isoforms of human, mouse, rat, trout, and chicken band 3 so far sequenced. Residue W831 (W36 of peptide 12/13) is conserved for most isoforms and sequences, and where it is not the residue substitution is conservative. This implies that these residues may well participate in important structural interactions. In addition, residues R13 and D26 (residues R808 and D821 of human band 3) may form an ion pair that stabilizes the loop structure. Both these residues are conserved for all isoforms of human, mouse, and rat band 3 (with a conservative substitution of D26 to E for one of the three band 3 isoforms of both rat and mouse).

Since the two helices are amphiphilic, it might have been expected that the hydrophobic sides of the helices would interact more extensively. However, it is likely that the hydrophobic faces of these helices pack against hydrophobic regions present on neighboring cytoplasmic loop domains or the N- or C-termini that extend into the cytoplasm (leaving the polar surfaces exposed to the environment). It would be useful to study the structure taken up by other cytoplasmic loop domains (and the C-terminus) to determine if these also possess hydrophobic regions when folded that could aid in inter-loop packing. However there also remains the possibility that TFE may induce an artificial structure in the peptide.

The two helices are connected by a proline-rich loop that itself exhibits local order, and the short regions connecting the helices to the loop are highly flexible. This flexibility may be necessary for the transport function of band 3 since loop 12/13 may have a role in anion binding (*22*). Rigid body motion of binding loops has been observed in proteins such as eglin c (*52*) and the long neurotoxin LSIII (*53*). It is also possible that this region is more rigid in the native protein when it is packed on the neighboring loops. No dNN sequential NOE cross-peaks were seen between residues L17 and Y29, but strong sequential dαN NOE cross-peaks were present between K22 and Y23, Y23 and H24, D26 and V27, and P28 and Y29. The presence of strong dαN sequential cross-peaks and an absence of dNN sequential cross-peaks often indicates absence of helical structure and the presence of extended β-sheet or β-turn. Most residues in this loop had $^3J_{HN-H\alpha}$ values typical of extended coil, though K22, Y23, H24, and D26 all had $^3J_{HN-H\alpha} > 8.0$ (which can signify the presence of β-structure). However for this region the chemical shift indices showed no clustering of values, suggesting the absence of α- or β-structure. The NOE cross-peaks define this as the region where peptide 12/13 loops back upon itself; the binding site of the monoclonal antibody BRIC 132 has been mapped to this region (K19 to Y29) suggesting that this may represent an exposed region of the protein.

Site-directed mutagenesis studies of Muller-Berger et al. (*22*) on mouse band 3 have shown that both K832 and K835 (residues K19 and K22 of peptide 12/13) are required for anion transport. (The residues are numbered here as for the mouse band 3, and are conserved in both the mouse and human band 3, as well as most other species so far sequenced.) Mutation of either H837 or H852 (residues H24 and H39 of peptide 12/13) also reduced chloride transport. These authors suggested that H-bond formation may occur between the substrate anion (as an acceptor) and three or four donors including K832 (residue K19 for peptide 12/13), K835 (K22), H837 (H24), and H852 (H39). In the 15 lower energy structures the side chains of residues K19, K22, and H24 of peptide 12/13 are in close proximity (Figure 7) suggesting such a H-bond arrangement for anion binding is possible. It was also shown by Muller-Berger et al. (*22*) that the reduction of transport following the mutation of either H839 or H852 (residues H24 and H39 for peptide 12/13), was partially (H24) or fully (H39) recovered by a second site mutation at K558 (the covalent binding site of the anion transport inhibitor H_2DIDS [4,4′-diisothiocyanostilbene-2,2′-disulfonate]). Residue K558 is located in membrane span 5 and is distant in the amino acid sequence from the two histidine residues. The loss of chloride transport and regain following the second site mutation was particularly dramatic for the H39 residue; this may represent the loss and regain of tight packing of H39 against neighboring helices or loops following changes in helix packing. If H24 does represent one of the residues that make up the monovalent anion binding site, the reduction in chloride transport may be due to the loss of one donor site which is not recovered by a change in helix packing following the K558 mutation. In Figure 4, H24 is close to the loop apex of peptide 12/13, whereas H39 is positioned toward the bottom of a helical stretch. Mutation of proline residues in the vicinity of H24 inhibited band 3 mediated chloride transport, but this was not recovered by the second site mutation at K558. These mutations may be directly affecting the geometry of the anion binding site. Muller-Berger et al. (*22*) also showed that K31 and K34 were not required for transport activity. In Figure

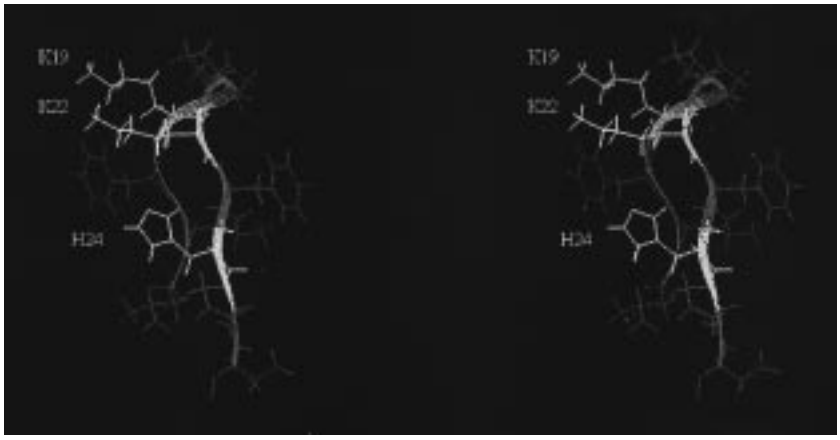

FIGURE 7: Stereoview showing the position of the lysine and histidine residues within the loop apex of peptide 12/13.

4 it can be seen that both these residues are on the same face of the Y29 to F41 helix. It could be that this face of the helix (between K31 to K34, and perhaps for a longer stretch of this helix) is exposed (to the cytoplasm) in intact band 3.

There are a cluster of conserved basic resides in the cytoplasmic loop connecting the putative sixth and seventh trans-membrane spans (loop 6/7), and these have recently been shown to be important for the transport of the divalent anion sulfate in red cells (*54*). In contrast, natural mutations in this loop had little effect on the binding and transport of monovalent anions. It is possible that residues K19, K22, and H24 of peptide 12/13 represent the binding site for monovalent anions (chloride and bicarbonate), while the additional basic residues of loop 6/7 are required for divalent anion binding. If this is the case, then it is possible to speculate on the proximity of the 6/7 and 12/13 loops. Given that the distance between any two oxygen atoms of the tetrahedral sulfate anion is 2.3 Å, the binding sites on the two loops are within this distance plus the length of two H-bonds (approximately 4 Å) i.e., 6.3 Å apart when sulfate is bound.

It has been suggested that band 3 may possess up to 14 trans-membrane spans, though a recent study by Popov et al. (*11*) suggested that only 12 spans may be present. The structure taken up by peptide 12/13 (and in particular the position of the tip of the loop and the alignment taken up by the N- and C-termini) fits better the 12 span model of Popov et al. than the 14 trans-membrane spanning model (*8*, *10*, *21*).

A low-resolution three-dimensional map of the dimeric form of the membrane domain of band 3 has been published (*15*). The map shows a 40 Å thick basal domain (sufficient to span the cell membrane) with an approximate cross sectional area of 1500 Å (large enough to accommodate 12−14 trans-membrane helices). A protrusion measuring 25 × 80 Å extends 40 Å from the surface of the basal domain and into the aqueous medium. This projection can be accounted for by polypeptide located in the large C-terminal cytoplasmic loops (for example, loop domains 10/11 and 12/13), the C-terminal tail, and the cytoplasmic sequence at the N-terminal side of the membrane domain (residues 359−403) which forms part of the membrane domain. When the loop domain of peptide 12/13 (residues 17−28) is in the same plane as the two helices (residues 8−15, and residues 29−41) the distance from I8 to K22 (the apex of the loop domain of peptide 12/13) is approximately 32 Å and the distance from L40 to K22 is approximately 36 Å. Within the loop domain, residues P20 and H24 are approximately 7 Å apart and residues L17 and D26 are approximately 14 Å apart. The structure taken up by peptide 12/13 could clearly be accommodated by the extramembraneous protrusion identified by Wang et al. (*15*), and it is possible for both of the helices identified by NMR to be positioned outside the lipid bilayer. Exposed hydrophobic residues of the 12/13 loop of band 3 may be shielded from the cytoplasm by packing to the other cytoplasmic regions of the membrane domain, which probably constitute the rest of the protrusion. Figure 8 shows how the structure taken up by peptide 12/13 can be positioned within the three-dimensional map of Wang et al. (*15*).

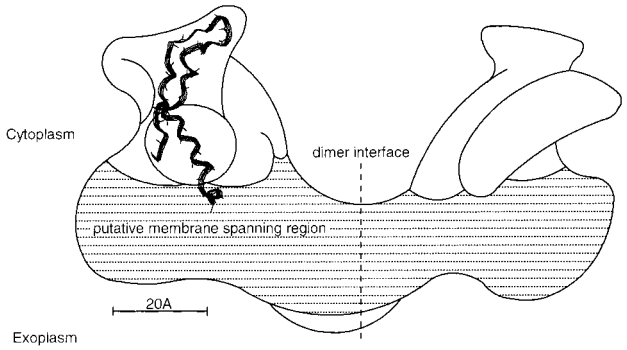

FIGURE 8: A possible position for peptide 12/13 within the three-dimensional map (*15*) of the band 3 dimer. Vertical striped areas represent the protein−protein contacts in the crystals, and horizontal striped areas represent the region within the lipid bilayer. (The representation of the band 3 structure in this figure is an adaptation of the three-dimensional map given in ref *15*.)

The results of a study by NMR of the three cytoplasmic loops of the polytopic membrane protein, rhodopsin, have been recently published (ref *19*; and see ref *20* for an earlier study of the third cytoplasmic loop). Short synthetic peptides corresponding to these loops were shown to take up well-ordered structures in the absence of the hydrophobic trans-membrane helices. Two of the three loop peptides exhibited biological activity in solution (inhibiting the activation of G protein, transducin, by light-activated rhodopsin): the structures taken up by the peptides in solution retain important elements of the activated rhodopsin. Though no biological assay exists to test the structure taken up by peptide 12/13,

the folding of this peptide into a defined loop structure is thus not without precedence. As was found in the rhodopsin study, the primary sequence of loop 12/13 without the full trans-membrane sequences of helices 12 and 13 is sufficient for folding. It was concluded by Yeagle et al. (*19*) that the loops of rhodopsin may contribute to the folding of the membrane protein during synthesis and membrane insertion; this may also be true for loop 12/13 of band 3.

ACKNOWLEDGMENT

We thank Dr. Angelo R. Gargaro for advice and helpful discussion.

SUPPORTING INFORMATION AVAILABLE

The amide-proton region of the NOE spectrum of linear peptide 12/13, at 25 °C, and with a mixing time of 150 ms. The proton resonances of linear peptide 12/13 at 25 °C in 30% TFE-d_3. NOE spectra of linear peptide 12/13, at 25 °C, and with a mixing time of 150 ms, with the nine long-range NOE cross-peaks highlighted (4 pages). Ordering information is given on any current masthead page.

REFERENCES

1. Tanner, M. J. A., Martin, P. G., and High, S. (1988) *Biochem. J. 256*, 703−712.
2. Steck, T. L., Ramos, B., and Strapazon, E. (1976) *Biochemistry 15*, 1154−1161.
3. Low, P. S. (1986) *Biochim. Biophys. Acta 864*, 145−167.
4. Low, P. S., Willardson, B. H., Thevenin, B., Kannan, R., Mahler, E., Geahlen, R. L., and Harrison, M. (1989) in *Anion Transport Proteins of the Red Blood Cell Membrane* (Hamasaki, N., and Jennings, M. L., Eds.), pp 103−118, Elsevier, Amsterdam.
5. Lepke, S., and Passow, H. (1976) *Biochim. Biophys. Acta 455*, 353−370.
6. Grinstein, S., Ship, S., and Rothstein, A. (1979) *Biochim. Biophys. Acta 455*, 353−370.
7. Jennings, M. L. (1989) *Annu. Rev. Biophys. Biophys. Chem. 18*, 397−430.
8. Reithmeier, R. A. F. (1993) *Curr. Opin. Struct. Biol. 3*, 515−523.
9. Tanner, M. J. A. (1993) *Semin. Hematol. 30*, 34−57.
10. Tanner, M. J. A. (1989) *Methods Enzymol. 173*, 423−431.
11. Popov, M., Tam, L. Y., Li, J., and Reithmeier, R. A. F. (1997) *J. Biol. Chem. 272*, 18325−18332.
12. Steck, T. L. (1978) *J. Supramol. Struct. 8*, 311−324.
13. Dolder, M., Walz, T., Hefti, A., and Engel, A. (1993) *J. Mol. Biol. 231*, 119−132.
14. Wang, D. N., Kuhlbrandt, W., Sarabia, V. E., and Reithmeier, R. A. F. (1993) *EMBO J. 12*, 2233−2239.
15. Wang, D. N., Sarabia, V. E., Reithmeier, R. A. F., and Kuhlbrandt, W. (1994) *EMBO J. 13*, 3230−3235.
16. Oikawa, K., Lieberman, D. M., and Reithmeier, R. A. F. (1985) *Biochemistry 24*, 2843−2848.
17. Gargaro, A. R., Bloomberg, G. B., Dempsey, C. E., Murray, M., and Tanner, M. J. A. (1994) *Eur. J. Biochem. 221*, 445−454.
18. Casey, J. R., and Reithmeier, R. A. F. (1991) *J. Biol. Chem. 266*, 15726−15737.
19. Yeagle, P. L., Alderfer, J. L., and Albert, A. D. (1995) *Biochemistry 34*, 14621−14625.
20. Yeagle, P. L., Alderfer, J. L., Salloum, A. C., Ali, L., and Albert, A. D. (1997) *Biochemistry 36*, 3864−3869.
21. Wood, P. G. (1992) in *The Band 3 Proteins: Anion Transporters, Binding Proteins and Senescence Antigens* (Bamberg, E., and Passow, H., Eds.) pp 325−352, Elsevier, Amsterdam.
22. Muller-Berger, S., Karbach, D., Konig, J., Lepke, S., Wood, P. G., Appelhans, H., and Passow, H. (1995) *Biochemistry 34*, 9315−9324.
23. Wainwright, S. D., Mawby, W. J., and Tanner, M. J. A. (1989) *Biochem. J. 272*, 265−268.
24. Boumrah, D., Campbell, M. M., Fenner, S., and Kinsman, R. G. (1991) *Tetrahedron Lett. 32*, 7735−7738.
25. Bloomberg, G. B., Askin, D., and Tanner, M. J. A. (1996) in *Innovation and Perspectives in Solid-Phase Synthesis and Combinational Libraries* (Epton, R., Ed.) pp 321−322, Mayflower Scientific Ltd., Birmingham, England.
26. Chen, Y.-H., Yang, J. T., and Martinez, H. M. (1972) *Biochemistry 11*, 4120−4131.
27. Morris, G. A., and Freeman, R. (1978) *J. Magn. Reson. 29*, 433−462.
28. States, D. J., Haberkorn, R. A., and Ruben, D. J. (1982) *J. Magn. Reson. 48*, 286−292.
29. Piantini, U., Sorensen, O. W., and Ernst, R. R. (1982) *J. Am. Chem. Soc. 104*, 6880−6881.
30. Shaka, A. J., and Freeman, R. (1983) *J. Magn. Reson. 51*, 169−173.
31. Bodenhauser, G., Kogler, H., and Ernst, R. R. (1984) *J. Magn. Reson. 58*, 370−388.
32. Braunschweiler, L., and Ernst, R. R. (1983) *J. Magn. Reson. 53*, 521−528.
33. Bax, A., and Davies, D. G. (1985) *J. Magn. Reson. 65*, 335−360.
34. Rance, M. (1987) *J. Magn. Reson. 74*, 557−564.
35. Kumar, A., Ernst, R. R., and Wuthrich, K. (1980) *Biochem. Biophys. Res. Commun. 95*, 1−6.
36. Hommel, U., Harvey, T. S., Driscoll, P. C., and Campbell, I. D. (1992) *J. Mol. Biol. 227*, 271−282.
37. Brunger, A. T. (1992) *X-PLOR Manual*, Yale University, New Haven, CT.
38. Ryckaert, J.-P., Ciccotti, G., and Berensen, H. J. C. (1977) *J. Comput. Phys. 23*, 327−341.
39. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283−291.
40. Jasanoff, A., and Fersht, A. R. (1994) *Biochemistry 33*, 2129−2135.
41. Woody, R. W. (1985) in *The Peptides, Vol. 7* (Blout, E. R., Bovey, F. A., Goodman, M., and Loton, N., Eds.) pp 15−114, Academic Press, New York.
42. Wuthrich, K. (1986) in *NMR of Proteins and Nucleic Acids*, John Wiley, New York.
43. Wishart, D. S., Sykes, B. D., and Richards, F. M. (1991) *J. Mol. Biol. 222*, 311−333.
44. Wishart, D. S., Sykes, B. D., and Richards, F. M. (1992) *Biochemistry 31*, 1647−1651.
45. Nelson, J. W., and Kallenbach, N. R. (1986) *Proteins: Struct. Funct. Genet. 1*, 211−217.
46. Nelson, J. W., and Kallenbach, N. R. (1989) *Biochemistry 28*, 5256−5261.
47. Segawa, S., Fukuno, T., Fujiwara, K., and Noda, Y. (1991) *Biopolymers 31*, 497−509.
48. Dufour, E., and Haertle, T. (1990) *Protein Eng. 4*, 185−190.
49. Thomas, P. D., and Dill, K. A. (1993) *Protein Sci. 2*, 2050−2065.
50. Blanco, F. J., Jimenez, M. A., Rico, M., Santoro, J., and Nieto, J. L. (1993) *J. Am. Chem. Soc. 115*, 5887−5888.
51. Blanco, F. J., Jimenez, M. A., Pineda, A., Rico, M., Santoro, J., and Nieto, J. L. (1994) *Biochemistry 33*, 6004−6014.
52. Peng, J. W., and Wagner, G. (1992) *Biochemistry 31*, 8571−8586.
53. Connolly, P. J., Stern, A. S., and Hoch, J. C. (1996) *Biochemistry 35*, 418−426.
54. Bruce, L. J., Cope, D. L., Jones, G. K., Schofield, A. E., Burley, M., Povey, S., Unwin, R. J., Wrong, O., and Tanner, M. J. A. (1997) *J. Clin. Invest. 100*, 1693−1707.

BI973158D