Journal of Biomolecular NMR, 3 (1993) 613–626
ESCOM

J-Bio NMR 153

^{1}H NMR studies of the mercuric ion binding protein MerP: Sequential assignment, secondary structure and global fold of oxidized MerP

Per-Olof Eriksson[a,*] and Lena Sahlman[b]

[a]*Department of Physical Chemistry and* [b]*Department of Biochemistry, University of Umeå, S-901 87 Umeå, Sweden*

Received 5 July 1993
Accepted 23 August 1993

Keywords: NMR; Protein; MerP; Bacterial mercuric ion resistance; Sequential assignment; Secondary structure; Global fold

SUMMARY

The oxidized form of the mercuric ion binding protein MerP has been studied by two-dimensional NMR. MerP, which is a periplasmic water-soluble protein with 72 amino acids, is involved in the detoxification of mercuric ions in bacteria with resistance against mercury. The mercuric ions in the periplasmic space are first scavenged by the MerP protein, then transported into the cytoplasm by the membrane-bound transport protein MerT, and finally reduced to elementary (nontoxic) mercury by the enzyme mercuric reductase. In this work, the ^{1}H NMR spectrum of oxidized MerP (closed disulfide bridge) has been assigned by using homonuclear 2D NMR techniques. The secondary structure and global fold have been inferred from the nuclear Overhauser effect (NOE) data. The secondary structure comprises four β-strands and two α-helices, in the order $\beta_1\alpha_1\beta_2\beta_3\alpha_2\beta_4$. The protein folds into an antiparallel β-sheet, $\beta_2\beta_3\beta_1\beta_4$, with the two antiparallel helices on one side of the sheet. The folding topology is similar to that of acylphosphatase, the activation domain of porcine pancreatic procarboxypeptidase B, the DNA-binding domain of bovine papillomavirus-1 E2 and the RNA-binding domains of the U1 snRNP A and hnRNP C proteins. However, there is no structural similarity between MerP and other bacterial periplasmic binding proteins.

INTRODUCTION

Gram-negative bacteria that are resistant to mercuric ions require three and, in some cases, four proteins for detoxification of Hg^{2+}. The enzyme mercuric reductase is located in the cytoplasm and reduces Hg^{2+} to Hg^0, using NADPH (for a review, see Summers, 1986). Two or three proteins are required to transport the mercuric ions into the cell. There are two membrane proteins, MerT and MerC, where at least MerT is necessary for Hg^{2+} resistance (Hamlett et al., 1992). Some bacteria lack MerC, so it is not clear what function MerC has in transport. Finally, a periplasmic

*To whom correspondence should be addressed.

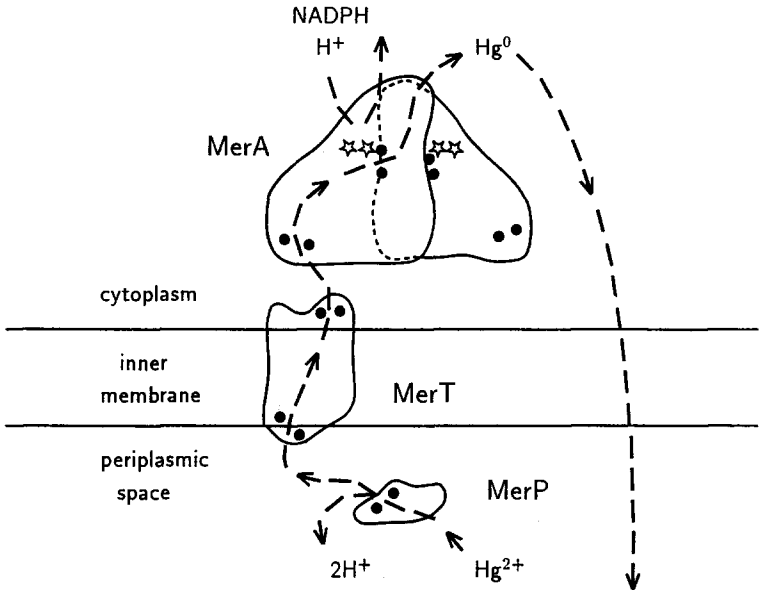

Fig. 1. Model for bacterial mercuric resistance (redrawn after Brown, 1985) showing the locations of the *merA*, *merP* and *merT* gene products relative to the inner membrane. The cysteine pairs, to which the mercuric ion is suggested to bind, are represented by solid circles. The mercuric reductase protein (MerA) is drawn as a dimer. The stars represent the redox active cysteine pair at the active site of MerA.

protein MerP is involved, whose function it is to bind mercuric ions as soon as they diffuse across the outer membrane (Brown, 1985) (Fig. 1). Of the four proteins, mercuric reductase has been thoroughly studied (for a review, see Williams, 1992). Of the other proteins only MerP has been studied at the protein level (Sahlman and Jonsson, 1992). MerP is a small water-soluble protein, containing 72 amino acids after a 19-amino acid long signal sequence has been cleaved off. The amino acid sequence has been deduced from the DNA sequence of the gene *merP* (Barrineau et al., 1984; Misra et al. 1984). The crystal structure of MerP is currently unknown. MerP contains two cysteines, close together in the sequence. Recently, the gene for the MerP protein from the *mer* operon of transposon Tn21 has been cloned into an overproducing bacterial strain and the protein has been purified (Sahlman and Jonsson, 1992). It was shown by in vitro studies that MerP specifically binds one Hg^{2+} per molecule, even in the presence of an excess of thiol compound (cysteine) (Sahlman and Jonsson, 1992). The protein can be purified in two different forms, depending upon whether cysteine is included as a reducing agent in the buffers or not. One form has no reactive thiol groups (oxidized, disulfide bridge) and the other has two (reduced, dithiols). The oxidized form does not bind Hg^{2+} specifically. This indicates that it is the two cysteines in MerP that bind Hg^{2+} (Sahlman and Jonsson, 1992). The size and water solubility of MerP make

Abbreviations: COSY, correlated spectroscopy; DSS, 2,2-dimethyl-2-silapentene-5-sulphonate; DQ, double quantum; DQF, double-quantum filtered; DTT, dithiotreitol; FT, Fourier transform; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; NOESY, NOE spectroscopy; ppm, parts per million; R-COSY, relayed COSY; SCUBA, stimulated cross peaks under bleached alphas; TOCSY, total correlation spectroscopy; TPPI, time proportional phase incrementation. The standard three-letter code for amino acid residues is used in the text and in the table, whereas the one-letter code is used in the figures.

it an excellent candidate for structure determination by NMR. We report here on the nearly complete sequential assignment of the ^1H NMR spectra of oxidized MerP. The secondary structure elements and the global fold are inferred from nuclear Overhauser effect (NOE) data.

MATERIALS AND METHODS

Protein purification and sample preparation

The MerP protein was obtained as previously described (Sahlman and Jonsson, 1992). The solution of oxidized protein was dialyzed against water for 24 h, with three changes of water, and lyophilized. Lyophilized protein samples were stored at 4 °C until used. Protein samples for NMR measurements were prepared by dissolving lyophilized protein in a buffer solution containing 50 mM deuterated acetic acid in H_2O/D_2O 90:10 (vol/vol). The pH was adjusted to 4.9 (uncorrected pH meter reading) by addition of 0.5 M HCl. For preparation of D_2O samples, the lyophilized protein was dissolved in a buffer prepared from 99.9% D_2O. The protein samples were sterile filtered before depositing into sterile 5-mm NMR tubes. The protein concentration of the NMR samples was between 3 and 6 mM.

NMR spectroscopy

Two-dimensional ^1H NMR spectra were recorded on a Bruker AM 500 spectrometer, equipped with an Aspect 3000 computer and digital phase-shifting hardware. The probe temperature was controlled with a Bruker BVT 1000 unit. COSY (Aue et al., 1976; Marion and Wüthrich, 1983), DQF-COSY (Rance et al., 1983), NOESY (Jeener et al., 1979; Kumar et al., 1980), relayed COSY (R-COSY) (Eich et al., 1982; Wagner, 1983; Bax and Drobny, 1985), TOCSY (Braunschweiler and Ernst, 1983; Bax and Davis, 1985) and double-quantum (DQ2D) (Rance et al., 1985) spectra were acquired at different temperatures (15 °C, 28 °C, and 37 °C). Between 512 and 800 experiments were collected, with an evolution time (t_1) increment of 90 μs. The TPPI technique (Drobny et al., 1979; Bodenhausen et al., 1980; Marion and Wüthrich, 1983) was used for frequency sign discrimination in the indirect dimension. For DQ2D spectra, the preparation time was 50 ms, and 1024 experiments were collected with a t_1 increment of 45 μs. The isotropic mixing in the TOCSY experiment was accomplished with a 'clean' MLEV-17 sequence enclosed by 2.5-ms trim pulses (Griesinger et al., 1988). The r.f. field strength was approximately $\gamma B_1/2\pi = 10$ kHz in the TOCSY experiments. Mixing times (τ_m) between 75 and 85 ms were used for TOCSY and between 200 and 300 ms for NOESY. In R-COSY a mixing time of 26 ms was used (Bax and Drobny, 1985). For NOESY and TOCSY spectra, the signal was acquired in pure phase, with the dispersive component in the real quadrature channel, resulting in minimal offset and tilt of the baseline (Marion and Bax, 1988). The spectral width was 5555 Hz, and the acquisition time in the t_2 dimension was 184 ms. Chemical shifts were referenced to internal DSS. The carrier was placed on the H_2O or the residual HDO resonance. The water resonance was suppressed by low-power presaturation during the recycle delay (2.0 s). A SCUBA preparation sequence (Brown et al., 1988), of total length between 50 and 100 ms, was used to allow the detection of cross peaks from signals under the water peak. Solvent peak recovery during the mixing time in the NOESY experiments was reduced by a centrally placed composite 180° pulse (Brown et al., 1988). TOCSY, DQF-COSY and DQ2D spectra of D_2O samples were recorded without presaturation.

NMR spectra were processed off-line using the FELIX program, version 1.0 (Hare Research,

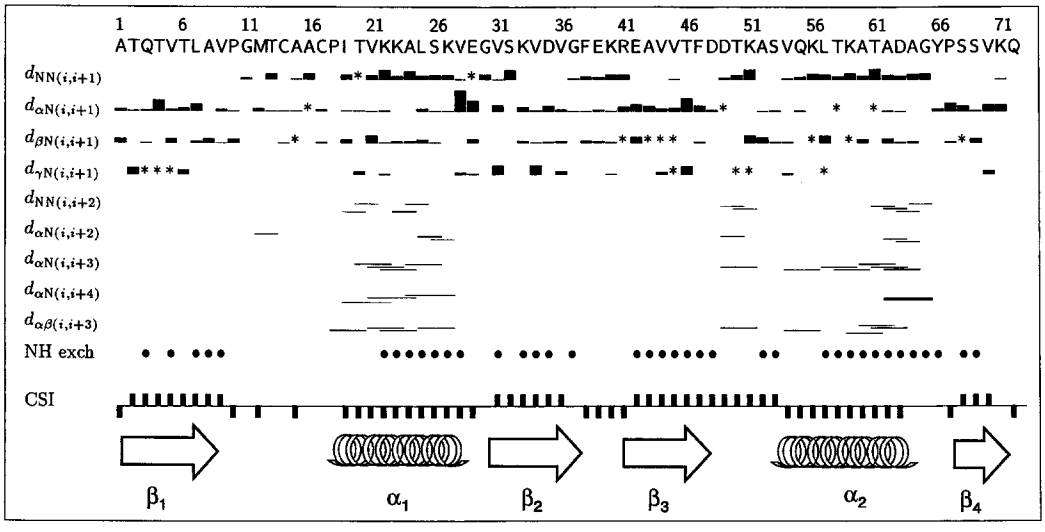

Fig. 2. Sequential and medium-range NOE connectivities, for oxidized MerP, observed in a NOESY spectrum acquired at 37 °C, pH 4.9, in H_2O, with a mixing time of 250 ms. The thickness of the bars and lines is proportional to the relative strength of the NOESY peak. $d_{\alpha,\delta}(i,i+1)$ connectivities for prolines are included with the $d_{\alpha N}(i,i+1)$ connectivities. Stars indicate unobservability of expected cross peaks due to spectrum overlap. The amplitudes of the $d_{\alpha\delta}(i,i+1)$ and $d_{\alpha\beta}(i,i+3)$ connectivities were measured in a NOESY spectrum acquired in D_2O, with a mixing time of 300 ms, and scaled appropriately. Residues with slower amide proton exchange rates are indicated by filled circles. The value of the chemical shift index (CSI) (Wishart et al., 1992) is indicated with a bar, signifying positive or negative displacement of the $C^{\alpha}H$ proton chemical shift from the reference values. The suggested secondary structure elements are indicated as arrows for β-strands and coils for α-helices.

Inc.), on a Silicon Graphics Iris 4D/25 workstation. The residual water signal was removed either by a low-frequency time-domain deconvolution (Marion et al., 1989) or by a polynomial fit of the time-domain data (Callaghan et al., 1984). The time-domain signals in either dimension were apodized with a phase-shifted skewed sinebell window function. The diagonal in the phase-sensitive COSY spectra was removed to produce a 'simple'-COSY (S.COSY) spectrum (Pelczer, 1991). The time-domain signals were zero-filled and Fourier-transformed to give a 2048 × 2048 real matrix resulting in a digital resolution of 2.7 Hz/point. Polynomial baseline correction was applied in the ω_2 dimension before the second FT. The first point in the t_1 dimension was linear predicted to improve the baseline (Marion and Bax, 1989). The t_1 noise in NOESY and TOCSY spectra was reduced using the approach of Manoleras and Norton (1992). Slowly exchanging amide protons were identified from the remaining NH–$C^{\alpha}H$ cross peaks in a 1.5-h COSY spectrum at 37 °C and pH 4.7, started 1 h after the protein was dissolved in D_2O.

RESULTS

Sequence-specific assignment

The ^1H NMR spectrum of oxidized MerP was assigned by homonuclear 2D NMR, according to the sequential assignment strategy, as described by Wüthrich (1986). A large fraction of the fingerprint ($C^{\alpha}H$–NH) cross peaks and the corresponding side chain spin systems were identified

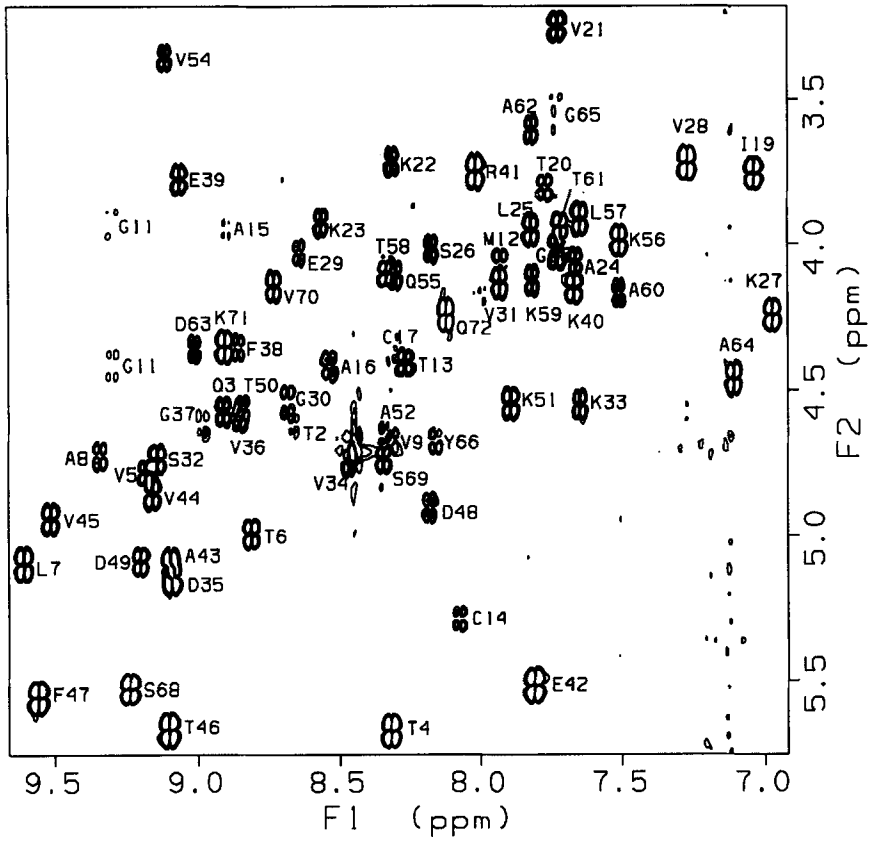

Fig. 3. Fingerprint region of a phase-sensitive COSY spectrum of oxidized MerP in H_2O at pH 4.9 and 37 °C.

with respect to type of amino acid at an early stage of the analysis. Extensive use was made of relayed TOCSY connectivities between amide and side chain protons (Chazin and Wright, 1987). For sequence-specific assignments sequential NOE contacts of type $C^{\alpha}H(i) - NH(i + 1)$ ($d_{\alpha N}(i,i + 1)$) and $NH(i) - NH(i + 1)$ ($d_{NN}(i,i + 1)$) were used (Wüthrich, 1986). Most sequential assignment steps were confirmed by the observation of $C^{\beta}H(i) - NH(i + 1)$ ($d_{\beta N}(i,i + 1)$) NOE contacts. The differential temperature dependence of the chemical shift was frequently used to resolve spectral overlap, in particular in the amide proton region.

The relative amplitudes of the observed sequential and medium-range NOESY cross peaks are summarized in Fig. 2. In Fig. 3 the 'fingerprint' region of a phase-sensitive COSY spectrum of MerP in H_2O at 37 °C is shown. In Fig. 4 some of the sequential assignment steps, using a 300-ms NOESY spectrum acquired in H_2O, are shown. The sequence-specific resonance assignments are summarized in Table 1.

Secondary structure and global fold

Conclusions about the secondary structure of oxidized MerP were drawn from observed sequential and medium-range NOE connectivities (Wüthrich, 1986). The existence of four β-strands was inferred from the observation of strong sequential $d_{\alpha N}(i,i + 1)$ NOESY cross peaks

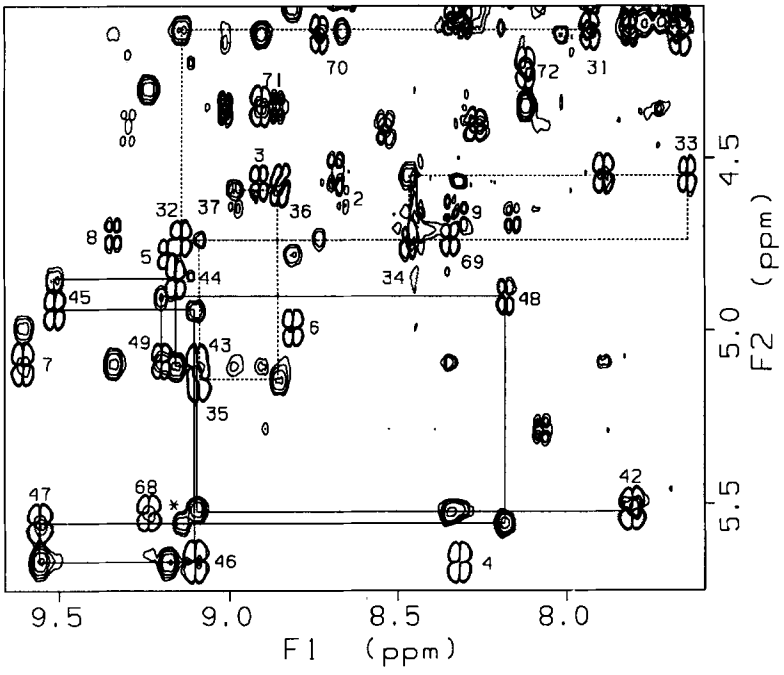

Fig. 4. (a) Superposition of the fingerprint region of a NOESY (τ_m = 250 ms) and a COSY spectrum of oxidized MerP in H_2O. COSY cross peaks are distinguished from the NOESY peaks by their characteristic four-lobe cross-peak pattern. The assignments of two β-strands (β_2: from Val[31] to Phe[38], dotted line; β_3: from Glu[42] to Asp[49], solid line), using sequential $d_{\alpha N}(i,i+1)$ NOESY connectivities and intraresidue COSY $C^{\alpha}H$–NH cross peaks are outlined. Note the absence of a $d_{\alpha N}(i,i+1)$ peak between Ser[32] and Lys[33] and the presence of a cross-strand $d_{\alpha N}(i,i+1)(47,32)$ connectivity (marked with a star), due to the β-bulge at Ser[32] (Fig. 6).

(Fig. 2), extending from Thr[2] to Val[9] (β_1), from Val[31] to Phe[38] (β_2), from Glu[42] to Asp[48] (β_3), and from Ser[68] to Lys[71] (β_4). No $d_{\alpha N}(i,i+1)$ cross peak is observed between Ser[32] and Lys[33] due to a 'β-bulge' at Ser[32] (see below). There is a consistent 'low-field shift' of the $C^{\alpha}H$ protons for all four β-strands, as expressed by the positive values of the chemical shift index (CSI) (Wishart et al., 1992) (Fig. 2), corroborating the existence of β-strand secondary structures. Slow amide proton exchange rates were observed throughout strand β_3 and for some of the residues in strands β_1, β_2, and β_4 (Fig. 2).

The arrangement of the four β-strands into a β-sheet was inferred from the observation of a number of interstrand NOE connectivities (Fig. 5). The cross-peak pattern is consistent with an antiparallel β-sheet where the β-strands are arranged in the order $\beta_4\beta_1\beta_3\beta_2$ (Fig. 5). The existence of slowly exchanging amide protons is well correlated with the expected pattern of hydrogen bonds in the β-sheet. So for example, in strand β_1, residues Thr[2], Thr[4], and Thr[6] are exposed to the solvent and display fast amide exchange rates, while the amide protons of Gln[3], Val[5], and Leu[7] are protected from exchange by hydrogen bonds to strand β_3. The β-sheet structure is amphipatic and shows well-defined polar and apolar sides (Fig. 6a). In particular, in the central part of the sheet the apolar side comprises the side chains of one leucine, one alanine, and five valines while the opposite side comprises a mixture of polar and apolar side chains. The 'structural hydrophobic moment' (Eisenberg et al., 1989) is 10.4 for the entire β-sheet (excluding Lys[71] and Gln[72]).

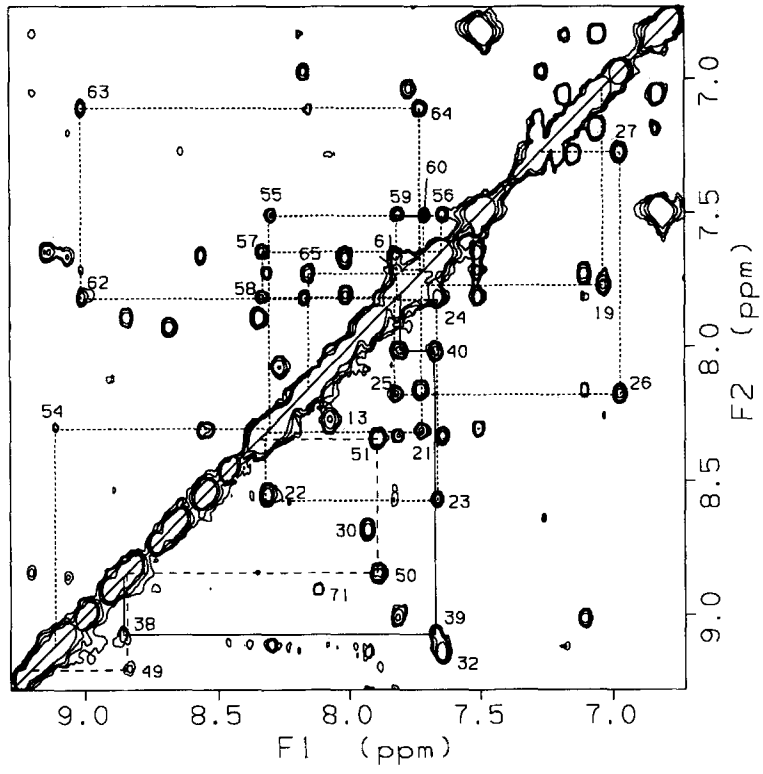

Fig. 4. (b) The amide region of a NOESY spectrum (τ_m = 250 ms) of oxidized MerP in H_2O. Assignments of two main α-helical regions, from Ile[19] to Val[28] and from Val[54] to Tyr[66], using sequential $d_{NN}(i,i+1)$ NOESY connectivities are shown (dotted lines). The labels of the cross peaks refer to the first residue of the pair connected by the cross peak. The helical-like sequence from Asp[49] to Ala[52] is marked with a dashed line and the loop from Phe[38] to Glu[42] with a solid line. The $d_{NN}(i,i+1)(32,33)$ cross peak arising from the β-bulge at Ser[32] (Fig. 5) is indicated.

At the N-terminal end of strand β_2, Ser[32] forms a β-bulge, as evident from a characteristic pattern of strong interstrand NOE contacts (Fig. 5), the disruption of the sequence of $d_{\alpha N}(i,i+1)$ connectivities at Ser[32] (Figs. 2 and 4a), and the presence of a strong $d_{NN}(i,i+1)$ connectivity from Ser[32] to Lys[33] (Fig. 4b). The β-bulge is located at the edge of the β-sheet, as often observed for other proteins (Creighton, 1983).

The existence of two α-helical segments was inferred from continuous stretches of $d_{NN}(i,i+1)$ NOESY cross peaks, and by sets of weaker d(i,i + 2), d(i,i + 3), d(i,i + 4) cross peaks (Figs. 2 and 4b). The first helix (α_1) extends from Ile[19] to approximately Val[28] and the second helix (α_2) from Val[54] to approximately Ala[64]. The next two residues, Gly[65] and Tyr[66], are probably involved in an irregular structure, as indicated by the strong $d_{\alpha N}(62,66)$ cross peak and by a helix-sheet contact between Tyr[66]($C^{\delta}H$) and Val[9]($C^{\gamma}H$) (see below). Helix α_1 is bordered by Pro[18] and Gly[30], residues known to promote the interruption of an α-helical structure. Finally, at the end of helix α_2 we have Gly[65] and Pro[67].

Slow amide proton exchange rates were observed throughout the helical regions, except for the first three N-terminal residues of each helix. This is in line with the expected absence of intraheli-

TABLE 1
PROTON RESONANCE ASSIGNMENTS FOR MerP (OXIDIZED FORM)[a]

Residue	NH	$C^{\alpha}H$	$C^{\beta}H$	$C^{\gamma}H$	Others
Ala^{1}	–	4.13	1.57		
Thr^{2}	8.66	4.63	3.85	1.01	
Gln^{3}	8.91	4.57	1.57, 0.69	2.14, 2.00	
Thr^{4}	8.32	5.67	3.73	1.05	
Val^{5}	9.18	4.78	2.20	1.04, 1.04	
Thr^{6}	8.81	4.99	4.06	1.02	
Leu^{7}	9.61	5.10	1.74, 1.25	1.64	$C^{\delta}H_3$ 0.74, 0.64
Ala^{8}	9.34	4.72	1.41		
Val^{9}	8.31	4.67	2.18	0.86, 0.72	
Pro^{10}	–	4.20	2.36, 1.97	2.01, 2.01	$C^{\delta}H_2$ 4.03, 3.43
Gly^{11}	9.30	3.92, 4.41			
Met^{12}	7.93	4.06	2.49, 2.49	1.74, 1.60	
Thr^{13}	8.26	4.40	4.40	1.13	
Cys^{14}	8.08	5.28	3.48, 3.68		
Ala^{15}	8.90	3.94	1.44		
Ala^{16}	8.53	4.41	1.44		
Cys^{17}	8.27	4.37	4.02, 4.02		
Pro^{18}	–	4.37	2.10, 1.64	–,–	$C^{\delta}H_2$ 3.85, 3.57
Ile^{19}	7.04	3.76	2.10	1.52, 1.29	$C^{\delta}H_3$ 0.88, $C^{\gamma}H_3$ 0.94
Thr^{20}	7.77	3.80	4.11	1.39	
Val^{21}	7.72	3.25	1.89	0.74, 0.51	
Lys^{22}	8.31	3.71	2.03, 1.85	1.34, 1.24	$C^{\delta}H_2$ 1.71, 1.63, $C^{\varepsilon}H_2$ 3.01, 3.08
Lys^{23}	8.56	3.92	1.91, 1.84	1.44, 1.40	$C^{\delta}H_2$ 1.67, 1.67, $C^{\varepsilon}H_2$ 2.93, 2.93
Ala^{24}	7.66	4.06	1.34		
Leu^{25}	7.82	3.95	1.91, 1.16	1.55	$C^{\delta}H_3$ 0.83, 0.55
Ser^{26}	8.17	4.02	3.87, 3.96		
Lys^{27}	6.97	4.24	1.94, 1.82	1.59, 1.56	$C^{\delta}H_2$, 1.70, 1.68, $C^{\varepsilon}H_2$ 2.99, 2.99
Val^{28}	7.27	3.72	2.21	0.98, 0.83	
Glu^{29}	8.64	4.03	2.07, 2.01	2.37, 2.30	
Gly^{30}	8.68	3.74, 4.54			
Val^{31}	7.93	4.13	2.27	0.81, 0.81	
Ser^{32}	9.14	4.73	3.84, 3.72		
Lys^{33}	7.65	4.55	1.77, –	–,–	$C^{\delta}H_2$ –,–, $C^{\varepsilon}H_2$ 2.97, 2.97
Val^{34}	8.47	4.74	1.91	0.87, 0.82	
Asp^{35}	9.09	5.14	2.70, 2.46		
Val^{36}	8.85	4.60	2.04	0.95, 0.88	
Gly^{37}	8.99	3.56, 4.61			
Phe^{38}	8.86	4.35	3.26, 3.09		2,6H 7.20, 3,5H 7.28, 4H 7.14
Glu^{39}	9.06	3.78	2.01, 2.01	2.35, 2.28	
Lys^{40}	7.66	4.15	1.33, 1.46	–,–	$C^{\delta}H_2$ –,–, $C^{\varepsilon}H_2$ 2.92, 2.92
Arg^{41}	8.01	3.75	1.95, 2.10	1.55	$C^{\delta}H_2$ 3.10, 3.03, $N^{\varepsilon}H$ 7.06
Glu^{42}	7.81	5.53	1.74, 1.69	2.20, 2.04	
Ala^{43}	9.10	5.10	0.98		
Val^{44}	9.16	4.86	2.05	0.98, 0.83	
Val^{45}	9.51	4.94	2.05	1.10, 0.95	
Thr^{46}	9.10	5.67	3.92	1.08	
Phe^{47}	9.56	5.55	2.81, 2.65		2,6H 6.83, 3,5H 7.06, 4H 7.17

TABLE 1
(continued)

Residue	NH	$C^{\alpha}H$	$C^{\beta}H$	$C^{\gamma}H$	Others
Asp^{48}	8.18	4.90	2.91, 2.70		
Asp^{49}	9.20	5.09	3.46, 2.40		
Thr^{50}	8.84	4.56	4.35	1.32	
Lys^{51}	7.89	4.54	1.70, –	–, –	$C^{\delta}H_2$ –, –, $C^{\varepsilon}H_2$ 2.95, 2.95
Ala^{52}	8.34	4.65	1.36		
Ser^{53}	6.36	4.83	4.22, 3.94		
Val^{54}	9.11	3.35	2.07	1.15, 0.98	
Gln^{55}	8.30	4.10	1.95, 2.05	2.46, 2.41	
Lys^{56}	7.50	3.99	1.83, 1.79	1.58, 1.49	$C^{\delta}H_2$ 1.70, 1.69, $C^{\varepsilon}H_2$ 3.03, 3.03
Leu^{57}	7.65	3.91	1.99, 1.15	1.68	$C^{\delta}H_3$ 0.47, 0.18
Thr^{58}	8.33	4.10	4.29	1.16	
Lys^{59}	7.81	4.12	1.85, 1.92	–, –	$C^{\delta}H_2$ –, –, $C^{\varepsilon}H_2$ 3.02, 3.02
Ala^{60}	7.51	4.16	1.41		
Thr^{61}	7.72	3.94	4.63	1.19	
Ala^{62}	7.82	3.60	1.65		
Asp^{63}	9.01	4.36	2.75, 2.68		
Ala^{64}	7.11	4.46	1.53		
Gly^{65}	7.73	3.44, 4.02			
Tyr^{66}	8.16	4.67	2.62, 2.33		2,6H 6.50, 3,5H 6.78
Pro^{67}	–	4.30	2.36, 1.88	2.07, 2.07	$C^{\delta}H_2$ 3.63, 3.48
Ser^{68}	9.24	5.53	3.86, 3.77		
Ser^{69}	8.34	4.73	3.87, 3.87		
Val^{70}	8.73	4.15	1.97	0.98, 0.97	
Lys^{71}	8.90	4.35	1.76, 1.55	1.44, 1.40	$C^{\delta}H_2$ 1.62, 1.62, $C^{\varepsilon}H_2$ 2.94, 2.94
Gln^{72}	8.11	4.24	2.13, 1.96	2.33, 2.33	

[a] Shown are the ^1H chemical shifts in ppm relative to internal DSS at pH 4.9 and 37 °C.

cal hydrogen bonds for the first three residues of an α-helix. Both helices display polar and apolar sides, as evident from a 'helical wheel' presentation of the two helices (Fig. 6b). The structural hydrophobic moment (Eisenberg et al., 1989) is 3.6 for helix α_1 (from Ile^{19} to Lys^{27}) and 5.7 for α_2 (from Val^{54} to Ala^{64}).

Since the β-strands are evenly distributed throughout the primary sequence (Fig. 2), arrangement of the four strands into a β-sheet broadly defines the global fold of the protein. For simple topological reasons the two helices would be expected to have an antiparallel orientation. A few observed interhelix NOE connectivities, as for example between Val^{21} ($C^{\gamma}H_3$) and Tyr^{66} ($C^{\delta}H$), between Val^{28} ($C^{\gamma}H_3$) and Ala^{60} (NH), and between Val^{28} ($C^{\beta}H$) and Leu^{57} ($C^{\delta}H$), confirm the antiparallel orientation and demonstrate that the two helices are located on the same side of the β-sheet.

Helix-sheet NOE connectivities are exclusively observed between residues at the apolar side of the α-helices and the apolar side of the β-sheet. So for example NOE connectivities are observed between Val^{21} ($C^{\gamma}H_3$) and Val^{9} (NH) and Phe^{38} ($C^{\delta}H$), between Ala^{52} ($C^{\beta}H_3$) and Phe^{47} ($C^{\beta}H$), between Leu^{57} ($C^{\gamma}H$) and Val^{45} ($C^{\alpha}H$), and between Thr^{61} ($C^{\gamma}H$) and Phe^{38} ($C^{\beta}H$). A NOESY cross peak is also observed between Tyr^{66} ($C^{\delta}H$) and Val^{9}. Since Tyr^{66} resides on the polar side of the

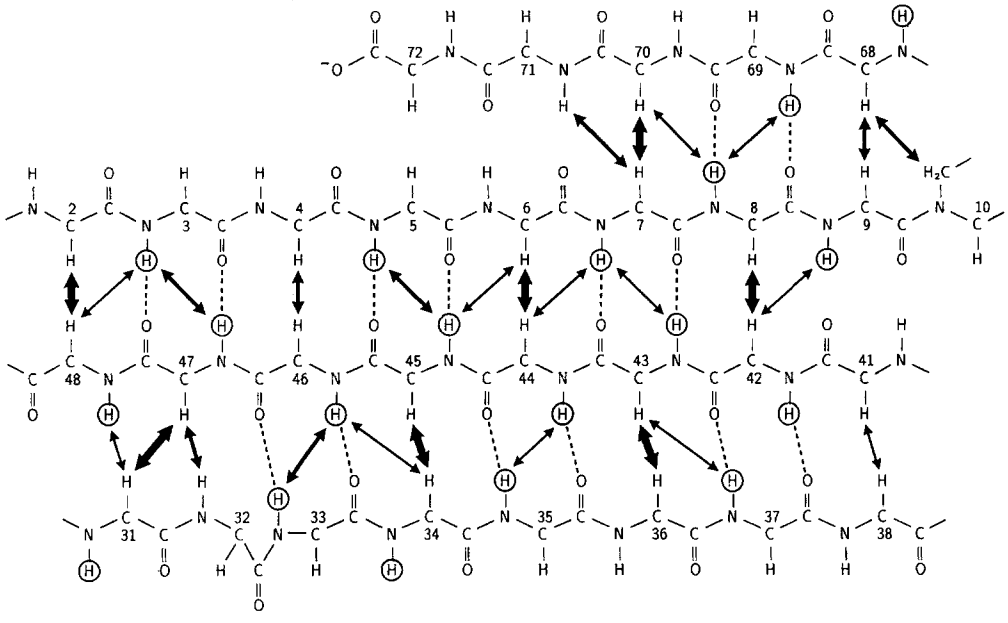

Fig. 5. The four-stranded antiparallel β-sheet in oxidized MerP. The observed interstrand NOE connectivities are indicated with arrows where the thickness semiquantitatively represents the amplitude of the NOESY cross peak. Cross peaks between $C^{\alpha}H$ protons were observed in a 300-ms NOESY spectrum acquired in D_2O, while the cross peaks involving amide protons were observed in a 250-ms NOESY spectrum acquired in H_2O. Slowly exchanging amide protons are marked by circles, and hydrogen bonds implied by slow exchange rates and interstrand NOE connectivities are indicated by dashed lines. The β-bulge at residue Ser^{32} is indicated.

helix (Fig. 6b), this gives a further indication that the regular α-helical structure ends at Ala^{64}. In order to accommodate all observed helix-sheet, and helix-helix NOE contacts it is suggested that the relative orientations of the helices around their long axes are as depicted in Fig. 6b. The long-range NOE connectivities discussed above broadly define the location of the helices relative to the β-sheet. A model for the global fold of oxidized MerP is thus outlined in Fig. 7.

The present data suggest that the region between Gly^{11} and Cys^{17} encompassing the proposed binding site of mercuric ions in reduced MerP (Cys^{14}, Cys^{17}) (Sahlman and Jonsson, 1992), consists of a more-or-less unstructured loop at the surface of the protein. The amide proton exchange rate is fast throughout this region (Fig. 2), indicating the absence of hydrogen bonds and exposure to solvent. A NOE contact between Ala^{16} ($C^{\alpha}H$) and Phe^{38} ($C^{\epsilon}H$) was observed, which indicates that the peptide chain folds back towards the edge of the β-sheet. Also an NOE connectivity is observed between Met^{12} ($C^{\beta}H$) and Val^{21} ($C^{\delta}H$).

The amide proton of Ser^{53} and one of the $C^{\beta}H$ protons of Gln^{3} display large upfield shifts (Table 1). This can be rationalized by the global fold model from the close proximity of Gln^{3} and Ser^{53} to the aromatic side chain of Phe^{47} (Fig. 7).

Indirect confirmation of the global fold model is provided by the observed chemical shift differences between oxidized and reduced MerP (Eriksson and Sahlman, unpublished results). The difference is largest around the site of the cysteine pair (Cys^{14}–Cys^{17}). However, a substantial shift difference is also observed for residues Val^{34} to Arg^{41}, and in particular for the loop connect-

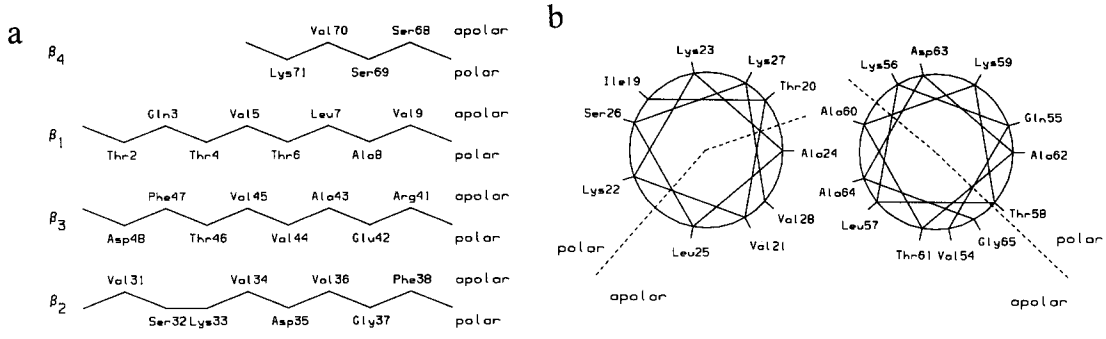

Fig. 6. (a) The four-stranded β-sheet of oxidized MerP, illustrating the well-defined polar and apolar sides of the sheet. Residues with side chains directed upwards are drawn above the lines and vice versa. (b) 'Helical wheel' representation of the two main α-helical regions of oxidized MerP, illustrating the polar and apolar sides of the helices. The helices are shown with a relative orientation as suggested by the observed interhelix and helix-sheet NOE connectivities (see text for details).

ing β-strands β_2 and β_3 (Glu^{39}–Arg^{41}). This is consistent with the global fold model (Fig. 7) which suggests proximity of the cystein-containing loop (Gly^{11}–Pro^{18}) and the loop between β_2 and β_3.

A more well defined structure model of MerP requires the acquisition of a larger set of NOE data. This is currently under way in our laboratory, and in a forthcoming publication the full assignment of the reduced form of MerP will be given, as well as a more detailed description of the 3D structure.

DISCUSSION

We have used homonuclear 2D NMR to study the oxidized form of the mercuric binding protein MerP. The small size of the protein (7474 Da) and the possibility of obtaining high-quality NMR spectra from solutions of the protein facilitated the analysis. Nearly complete sequential assignment of the ^1H NMR spectrum was obtained, and the secondary structure and the global

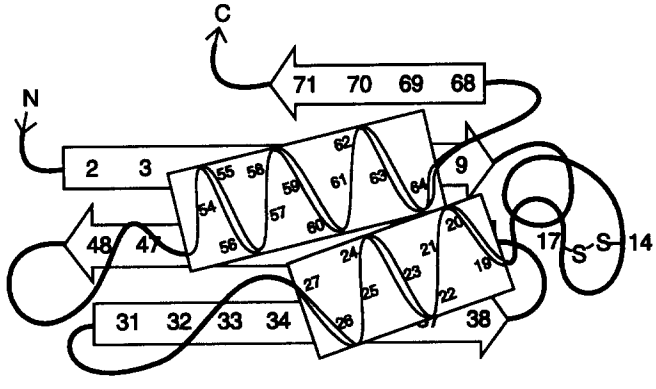

Fig. 7. Global fold of the protein MerP. The structure of the β-sheet is well defined from a number of interstrand NOE connectivities (Fig. 5). The orientations of the α-helices are approximate, deduced from a small number of helix-sheet NOE connectivities. The structures of the connecting loops between the secondary structure elements are largely tentative.

fold of the protein were deduced from the NOE data. The global fold implies that MerP belongs to a group of proteins having a βαβ–βαβ structure. One side of the protein is made up of the four antiparallel β-strands in the order $\beta_4\beta_1\beta_3\beta_2$ while the other side consists of two antiparallel α-helices. This type of structure has been referred to as an open-faced β-sheet sandwich (Richardson, 1981).

The present data indicate an α-helical content of between 30 and 35% for MerP. In a previous study the CD spectrum of MerP was recorded and the amount of α-helical structure was estimated to be around 80% (Sahlman and Jonsson, 1992). However, the software provided with the CD spectrometer used a very small databank, mainly comprising proteins with only α-helices.

MerP is a periplasmic binding protein, and the structures of seven different other periplasmic binding proteins have been determined by X-ray crystallography (for a review see Quiocho, 1990). These include binding proteins for L-arabinose, D-galactose/glucose, D-maltose, leucine/isoleucine/valine, leucine, sulfate, and phosphate. In contrast to MerP, all of these proteins are involved in the transport of nutrients into the bacterial cell. The molecular masses range from 33.6 to 40.5 kDa. The structures of these seven proteins are all similar. The structure of the arabinose-binding protein from *E. coli* has been refined to 1.7 Å. There are two similar globular domains where the ligand binds in a cleft formed between the two domains. Each domain contains six strands of parallel β-sheet and four α-helices where the β-sheet structure is surrounded by the α-helices (Quiocho and Vyas, 1984).

The mercuric ion binding protein MerP that we have studied is much smaller, 7474 Da. The global fold of MerP suggests that the ligand binds at the surface of the molecule, i.e., to the unstructured loop between the first β-sheet and the first α-helix, rather than in a cleft. Thus, there is no structural relationship between MerP and any previously characterized periplasmic binding protein. Since the transport system of mercuric ions lacks some essential features common to other bacterial binding-proteins transport systems, e.g., an ATPase-like membrane-associated protein, it is perhaps not surprising to find that the periplasmic binding protein is different as well.

Surprisingly enough, MerP structurally resembles other recently described protein structures, namely the RNA-binding domain of the U1 small nuclear ribonucleoprotein A (Nagai et al., 1990), the DNA-binding domain of bovine papillomavirus-1 E2 (BPV E2) (Hegde et al., 1992), the RNA-binding domain of hnRNP C (Wittekind et al., 1992a), the activation domain of pancreatic procarboxypeptidase (Vendrell et al., 1991; Coll et al., 1991), acylphosphatase (Pastore et al., 1992), and the ribosomal protein S6 from *Thermus thermophilus* (Lindahl, Svensson, Liljas, Sedelnikova, Eliseikina, Fomenkova, Nevskaya, Nikonov, Garber and Amons, unpublished results). These proteins all have four β-strands forming an antiparallel β-sheet at one side of the protein, but vary in the loops between the strands and the positioning of the α-helices. Two subunits of the DNA binding domain of BPV E2 come together to form a β-barrel by using the antiparallel β-sheets from each subunit. MerP most closely resembles acylphosphatase, the activation domain of pancreatic procarboxypeptidase, and ribosomal protein S6.

Two other, slightly different, proteins belong to the same general group. The first one is the phosphocarrier protein HPr (Wittekind et al. 1990, 1992b). The antiparallel β-sheet in this protein contains the strands in the order $\beta_1\beta_4\beta_2\beta_3$. If the amino acid sequence is read from the C-terminus towards the N-terminus, the order of the β-sheet strands is the same as for MerP. The second one is the C-terminal domain of the ribosomal protein L7/L12 (Leijonmarck and Liljas, 1987). This domain contains an α-helix instead of a β-strand, making it a βαα–βαβ structure. The β-sheets

form an antiparallel domain, with two α-helices on the other side of the protein. Therefore it can be said to be an open-face β-sandwich as well (Richardson, 1981).

It is interesting to note that a group of proteins with such diverse functions should have a similar structure. In most of the proteins discussed above, the β-sheet region is involved in substrate or ligand interaction. In the case of MerP, the ligand Hg^{2+} binds to two cysteine residues located on a loop between strand β_1 and helix α_1. However, binding of the ligand is only one of three main functions proposed for periplasmic binding proteins. The other two are the ability to recognize the inner membrane transport proteins and to transfer the ligand to these. It has also been suggested that binding of the ligand introduces conformational change in the binding protein, and that only this form is able to interact with the membrane bound transport proteins (for a review see Ames et al., 1990). It has not been shown that this model applies to MerP, but if so it seems likely that the β-sheets or α-helices of MerP may be involved in either of these features. It is here interesting to note the existence of the very hydrophobic interior of the MerP protein, between the central part of the β-sheet and the α-helices (Figs. 6 and 7). It is tempting to speculate that a small conformational change, caused by the binding of mercuric ions, may lead to exposure of the hydrophobic surfaces, allowing hydrophobic interaction with the membrane-bound transport protein. We are presently studying the reduced (mercuric ion binding) form of the MerP protein, with and without added Hg^{2+}, in order to detect what, if any, structural changes take place on binding of mercuric ions.

ACKNOWLEDGEMENTS

We thank Ms. Eleonore Granström Skärfstad for technical assistance, Prof. Ray Norton for supplying a copy of the t_1 noise reduction routine, and Dr. Antoinette Killian and Prof. Anders Liljas for fruitful discussions. This work was supported by grants from the Swedish Natural Science Research Council (K-KU 6483-301 to P.O.E. and K-KU 8966-304 to L.S.) and from the Magnus Bergwall foundation (P.O.E.). The NMR spectrometer was purchased with grants from the Knut and Alice Wallenberg foundation and the Swedish Natural Science Research Council.

REFERENCES

Ames, G.F.-L., Mimura, C.S. and Shyamala, V. (1990) *FEMS Microbiol. Rev.* **75**, 429–446.

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

Barrineau, P., Gilbert, P., Jackson, W.J., Jones, C.S., Summers, A.O. and Wisdom, S. (1984) *J. Mol. Appl. Genet.*, **2**, 601–619.

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

Bax, A. and Drobny, G. (1985) *J. Magn. Reson.*, **61**, 306–320.

Bodenhausen, G., Vold, R.L. and Vold, R.R. (1980) *J. Magn. Reson.*, **37**, 93–106.

Braunschweiler, L., Bodenhausen, G. and Ernst, R.R. (1983) *Molec. Phys..*, **48**, 535–560.

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

Brown, N.L. (1985) *Trends Biochem. Sci.*, **10**, 400–403.

Brown, S.C., Weber, P.L. and Mueller, L. (1988) *J. Magn. Reson.*, **77**, 166–169.

Callaghan, P.T., MacKay, A.L., Pauls, K.P., Söderman, O. and Bloom, M. (1984) *J. Magn. Reson.*, **56**, 101–109.

Chazin, W.J. and Wright, P.E. (1987) *Biopolymers*, **26**, 973–977.

Coll, M., Guasch, A., Avilés, F.X. and Huber, R. (1991) *EMBO J.*, **191**, 1–9.

Creighton, T.E. (1983) *Proteins, Structures and Molecular Properties*, W.H. Freeman and Company, New York, p. 230.

Drobny, G., Pines, A., Sinton, S., Weitekamp, D. and Wemmer, D. (1979) *Faraday Div. Chem. Soc. Symp.*, **13**, 49–55.
Eich, G., Bodenhausen, G. and Ernst, R.R. (1982) *J. Am. Chem. Soc.*, **104**, 3732–3733.
Eisenberg, D., Wesson, M. and Wilcox, W. (1989) In *Prediction of Protein Structure and the Principles of Protein Conformation* (Ed., Fasman, G.D.) Plenum Press, New York, pp. 635–646.
Griesinger, C., Otting, G., Wüthrich, K. and Ernst, R.R. (1988) *J. Am. Chem. Soc.*, **110**, 7870–7872.
Gross, K.H. and Kalbitzer, H.R. (1988) *J. Magn. Reson.*, **76**, 87–99.
Hamlett, N.V., Landale, E.C., Davis, B.H. and Summers, A.O. (1992) *J. Bacteriol.*, **174**, 6377–6385.
Hegde, R.S., Grossman, S.R., Laimins, L.A. and Sigler, P.B. (1992) *Nature*, **359**, 505–512.
Jeener, J., Meier, B.H., Backmann, P. and Ernst, R.R. (1979) *J. Chem. Phys.*, **71**, 4546–4553.
Kumar, A., Ernst, R.R. and Wüthrich, K. (1980) *Biochem. Biophys. Res. Commun.*, **95**, 1–6.
Leijonmarck, M. and Liljas, A. (1987) *J. Mol. Biol.*, **195**, 555–580.
Manoleras, N. and Norton, R. (1992) *J. Biomol. NMR*, **2**, 485–494.
Marion, D. and Bax, A. (1988) *J. Magn. Reson.*, **79**, 352–356.
Marion, D. and Bax, A. (1989) *J. Magn. Reson.*, **83**, 205–211.
Marion, D., Ikura, M. and Bax, A. (1989) *J. Magn. Reson.*, **84**, 425–430.
Marion, D. and Wüthrich, K. (1983) *Biochem. Biophys. Res. Comm.*, **113**, 967–974.
Misra, T.K., Brown, N.L., Fritzinger, D.C., Pridmore, R.D., Barnes, W.M., Haberstroh, L. and Silver, S. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 5975–5979.
Nagai, K., Oubridge, C., Jessen, T.H., Li, J. and Evans, P.R. (1990) *Nature*, **348**, 515–520.
Pastore, A., Saudek, V., Ramponni, G. and Williams, R.J.P. (1992) *J. Mol. Biol.*, **224**, 427–440.
Pelczer, I. (1991) *J. Am. Chem. Soc.*, **113**, 3211–3212.
Quiocho, F.A. (1990) *Phil. Trans. R. Soc. Lond. B.*, **326**, 341–351.
Quiocho, F.A. and Vyas, N.K. (1984) *Nature*, **310**, 381–386.
Rance, M., Sørensen, O.W., Bodenhausen, G., Wagner, G., Ernst, R.R. and Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun.*, **117**, 479–485.
Rance, M., Sørensen, O.W., Leupin, W., Kogler, H., Wüthrich, K. and Ernst, R.R. (1985) *J. Magn. Reson.*, **61**, 67–80.
Richardson, J. (1981) *Adv. Protein Chem.*, **34**, 167–339.
Sahlman, L. and Jonsson, B.-H. (1992) *Eur. J. Biochem.*, **205**, 375–381.
Summers, A.O. (1986) *Annu. Rev. Microbiol.*, **40**, 607–634.
Vendrell, J., Billeter, M., Wider, G., Avilés, F.X. and Wüthrich, K. (1991) *EMBO J.*, **10**, 11–15.
Wagner, G. (1983) *J. Magn. Reson.*, **55**, 151–156.
Wishart, D.S., Sykes, B.D. and Richards, F.M. (1992) *Biochemistry*, **31**, 1647–1651.
Williams Jr., C.H. (1992) In *Chemistry and Biochemistry of Flavoenzymes* (Ed., Muller, F.) Vol. 3, CRC Press, Boca Raton, FL, pp. 121–211.
Wittekind, M., Görlach, M., Friedrichs, M.S., Dreyfuss, G. and Müller, L. (1992a) *Biochemistry*, **31**, 6254–6265.
Wittekind, M., Rajagopal, P., Branchini, B.R., Reizer, J., Saier, M.H. and Klevit, R. (1992b) *Protein Sci.*, **1**, 1363–1376.
Wittekind, M., Reizer, J. and Klevit, R.E. (1990) *Biochemistry*, **29**, 7191–7200.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.