ELSEVIER

Biochimica et Biophysica Acta 1371 (1998) 241–253

BIOCHIMICA ET BIOPHYSICA ACTA
BBA

Solution and solid state conformation of the human EGF receptor transmembrane region

Alan C. Rigby, Chris W.M. Grant, Gary S. Shaw *

Department of Biochemistry and McLaughlin Macromolecular Structure Facility, The University of Western Ontario, London, Ontario, Canada N6A 5C1

Received 26 November 1997; accepted 2 February 1998

Abstract

The epidermal growth factor receptor (EGFR) is a member of the tyrosine kinase family of signalling cell surface molecules. Signalling by this protein is mediated through binding of epidermal growth factor to its extracellular region ultimately leading to phosphorylation of several residues on the intracellular portion of the receptor. The only means of communication between the intracellular and extracellular domains is via the transmembrane region of the protein. In this work we describe the first structural studies of a 34-residue synthetic peptide (hEGFRp), representative of the human EGFR transmembrane region, using two-dimensional and ^{2}H wideline NMR and CD spectroscopies. In water the peptide demonstrated a lack of regular secondary structure and existed as oligomers. Addition of the lipomimetic solvent, trifluoroethanol (TFE), led to the production of monomeric structured species. Analysis of NMR spectra of the hEGFRp indicated that an α-helix was present between residues M626 and R647. This observation was reinforced by solid state ^{2}H NMR studies in lipid bilayers which showed typical 'Pake' spectra indicating axially symmetric motion. The helical region in hEGFRp commences four residues later than predicted via hydrophobicity profiles, and extends to include several charged arginine residues which would lie on the cytosolic side of the membrane. These observations provide the first evidence that the transmembrane α-helical region in EGFR may not only traverse the membrane but may continue to the cytosolic region near T654, an important phosphorylation site.

Keywords: Cytosolic region; Membrane; Tyrosine kinase; Phosphorylation site; 'Pake' spectrum

1. Introduction

Many growth factors and hormones mediate biological responses through binding and activation of cell-surface receptors. One important membrane receptor is the epidermal growth factor receptor (EGFR), a member of the tyrosine kinase receptor superfamily. This 170 kDa glycoprotein is activated via extracellular binding of the epidermal growth factor (EGF). The architecture of the EGFR is a structural blueprint for many other receptor tyrosine kinases [1–5]. It possesses a glycosylated extracellular ligand binding domain of approximately 620 residues which is responsible for binding EGF [6]. A 540 residue intracellular cytoplasmic domain contains a consensus sequence typical of the tyrosine kinase gene family [7–9] including an ATP binding site, situated near the transmembrane–cytosolic interface

Abbreviations: EGFR, epidermal growth factor receptor; hEGFRp, human epidermal growth factor receptor peptide, residues 622–654; POPC, 1-palmitoyl-2-oleoylphosphatidylcholine; TFE, trifluoroethanol; CSI, chemical shift index

* Corresponding author.

0005-2736/98/$19.00
PII S0005-2736(98)00020-0

of EGFR. The extra and intracellular domains of the receptor are anchored to the plasma membrane by a single hydrophobic transmembrane spanning region [9]. Allosteric regulation of ligand binding, protein tyrosine kinase activity and ultimately, attenuation of the EGFR is controlled by phosphorylation of T654 [7,10].

Type I transmembrane regions, such as the one in the EGFR, are recognized as single hydrophobic stretches of amino acids that traverse the hydrophobic membrane environment. These regions provide the only means of communication between the cytoplasmic and extracellular domains of the receptor. Consequently, it has been suggested that receptor transmembrane regions may possess unique structural and/or chemical characteristics which contribute to their mechanism of action [2]. Observations from X-ray crystallographic analyses of multi-segment spanning membrane proteins such as bacteriorhodopsin, cytochrome oxidase and the photosynthetic reaction centre show that transmembrane regions are usually α-helical and range from 18–36 amino acids in length [1,2]. A recent investigation of glycophorin in detergent micelles has demonstrated a dimeric arrangement of 23-residue α-helices [11]. Similar information is not available for single segment transmembrane spanning proteins such as EGFR. This has made prediction and modelling of single transmembrane region proteins an area of intense interest with several algorithms available [12–15]. These approaches are very useful for prediction of the presence and approximate location of transmembrane regions based on the amino acid sequence. However, direct experimental evidence for the α-helical nature of this region, especially at the membrane/fluid interface is less common. Given the importance and spatial proximity of the phosphorylation site (T654) and the ATP binding site to this transmembrane region, structural information of these regions would be valuable.

One approach to obtain three dimensional structure and dynamics information of receptor proteins is to dissect them into modules or domains. This procedure has been used successfully to determine the three dimensional structures of several intracellular portions of receptors including the SH2 and SH3 domains from $pp60^{c\text{-}src}$ [16], Syk kinase [17] and Ab1 [18,19]. Previous studies for transmembrane regions have included NMR structure determinations of the putative transmembrane regions of the *neu* oncogene protein [20], bacteriorodopsin [21], the S_4 section of the sodium channel protein [22] and magainin [23] using lipomimetic solvent environments. The bulk of these studies have used peptides ranging from 18–23 residues. While this peptide length is likely adequate to span the lipid bilayer, it allows no structural information to be obtained for the extra or intracellular regions adjacent to the membrane.

A potentially powerful approach for obtaining structural and dynamics information about the membrane spanning region in an archetypical receptor such as the EGFR is to combine high resolution and wideline ^{2}H NMR methods employing synthetic peptides which include the proposed transmembrane regions. Other solid state NMR methods such as ^{15}N chemical shift tensor [24] or rotational resonance [25] can also be used. Typically, the secondary structure of the peptide can be determined using established two dimensional NMR techniques in a lipomimetic environment. This information may then be used to aid interpretation of solid state NMR data obtained for the peptide in a lipid bilayer employing strategically placed ^{2}H probes within the peptide. The latter method allows orientation and mobility of the deuterated peptide to be determined. These approaches have been used elegantly by Shon et al. [26] and Nambudripad et al. [27] in studies of the Pf1 coat protein in detergent micelles, phospholipid bilayers and virus particles. In the present work we have used this approach to study a synthetic peptide indicative of the transmembrane portion of the human epidermal growth factor receptor comprising residues S621 to T654 (hEGFRp), including ^{2}H labels at the methyl positions of residues A623, M644 and V650. This 34-residue peptide was designed to include approximately 8–10 residues expected to reside on the cytoplasmic side of the membrane allowing an estimation of the secondary structure of this region for a type I receptor. The results of our work were also compared to several predictive methods for the α-helical transmembrane region of the EGFR.

2. Materials and methods

Deuterated L-methionine ([d_3]-Met) and perdeuterated L-alanine ([d_4]-Ala) were obtained from Isotec

(Miamisburg, OH). Perdeuterated L-valine ([d_8]-Val) was purchased from Cambridge Isotope Laboratories (Woburn, MA). Deuterium oxide (99.9 at.% ^{2}H), and deuterated 2,2,2-trifluoroethyl alcohol-d_3 (99.5 at.% ^{2}H) were purchased from Isotec.

hEGFRp was synthesized using FMOC chemistry by Chiron Mimotopes (Clayton, Australia) and stored as a freeze dried powder. The peptide sequence was;

$$\overset{621}{\text{K}}\text{-I-A-T-G-M-V-G-A-}\overset{630}{\text{L}}\text{-L-L-L-L-V-V-A-L-G-}$$

$$\overset{640}{\text{I}}\text{-G-L-F-M-R-R-R-H-I-}\overset{650}{\text{V}}\text{-R-K-R-T.}$$

Three deuterated amino acid analogues were incorporated into the peptide at positions [d_4]-A623, [d_3]-M644, and [d_8]-V650. The synthesis of these FMOC deuterated derivatives is described elsewhere [28]. Peptide purification was accomplished by standard reversed phase HPLC analysis using a Vydac protein C4 column (0.46×25 cm) and a 25 min linear gradient from 0.1% phosphoric acid/H_2O to 0.1% phosphoric acid/CH_3CN. Elution was monitored at 214 nm with a flow rate of 1.5 ml/min. The purity of the peptide was verified using analytical HPLC with a Zorbax SB-C18 reversed phase column where hEGFRp eluted as a single peak. hEGFRp was verified by amino acid analysis and electrospray mass spectrometry; MW_{obs} 4014.4, MW_{calc} 4015.9. This hEGFRp was used in all NMR and CD spectroscopic studies.

2.1. Circular dichroism spectroscopy

CD experiments were performed using an OLIS On-Line Instrument (On-Line Instrument Systems, Bogart, GA) equipped with a thermostatically controlled cell holder stabilized by a circulating water bath. The cell holder temperature was maintained at 23°C. The spectropolarimeter was interfaced to an IBM AT personal computer for data collection. The sample was placed into a cell with a path length of 0.01 cm, and spectra were recorded between 195 and 250 nm. The spectropolarimeter was routinely calibrated before data collection with D-10-camphorsulfonic acid (CSA), at 192.5 and 290.5 nm.

CD experiments were carried out for a variety of TFE:H_2O conditions. Initial studies utilized a stock solution of approximately 764 μM hEGFRp in 300 μl of TFE. This stock solution was used to create a series of dilutions, which varied from 100% TFE to 20% TFE. Similar experiments were performed using H_2O as the initial solvent and encompassed a solvent range of 100% H_2O to 20% H_2O. The concentrations of the actual samples where spectra were acquired ranged from 23.2 ± 3.3 μM for the 100% TFE sample to 23.9 ± 3.1 μM for the 40% H_2O sample as determined from amino acid analyses performed in triplicate. During data acquisition, the total sample absorbance was maintained less than 1 to maximize the signal to noise ratio [29]. In all experiments a blank spectrum for the same percentage of TFE/H_2O was acquired and used for baseline correction. The percentage of α-helix in hEGFRp was determined using equation,

$$[\theta]_\lambda = \left(f_H - \frac{ik}{N}\right)[\theta]_{H\lambda_0}$$

where $[\theta]_\lambda$ is the observed helicity at 222 nm, f_H is the fraction of the peptide that is helical, i represents the number of helical segments, k is the wavelength dependent constant of 2.6 at 220 nm, N is the total number of residues within the peptide, and $[\theta]_{H\lambda 0}$ is the ellipticity for an α-helix of infinite length [30].

2.2. NMR spectroscopy

Purified hEGFRp was dissolved in a variety of solvent mixtures of H_2O, D_2O and TFE-d_3. Initial experiments used hEGFRp (2.1 mg) dissolved in 500 μl 99.9% D_2O at a pH of 7.09 (uncorrected meter reading). To optimize NMR spectra the peptide was also dissolved in various ratios of TFE and H_2O, and the pH was altered from acidic conditions at pH 2.3 to basic conditions at pH 9.1 and vice versa via the addition of NaO^{2}H or ^{2}HCl. For two dimensional experiments, 2.1 mg of hEGFRp was dissolved in 375 μl TFE and 125 μl H_2O added. The solution was adjusted to a final pH of 6.16. For ^{2}H NMR studies, lipid bilayers containing hEGFRp were prepared as previously described [28]. Briefly, the peptide was dissolved in TFE and the appropriate amount of POPC added to a peptide:POPC ratio of 3:50. The solvent was removed and the residue rehydrated in 30 mM HEPES containing 20 mM NaCl and 5 mM EDTA in deuterium-depleted water at pH 7.2.

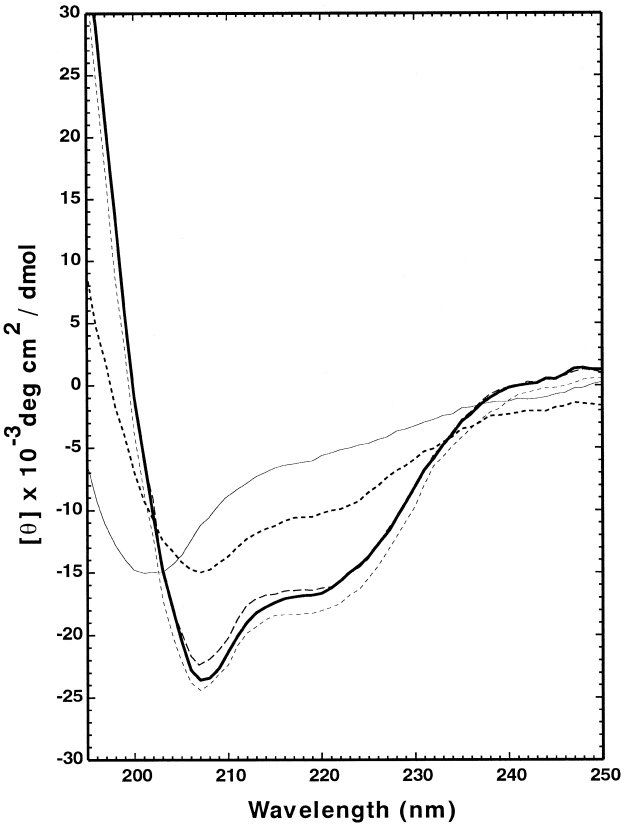

Fig. 1. Circular dichroism spectra of hEGFRp at a variety of H_2O/TFE concentrations at 25°C, pH 7.09. hEGFRp was initially dissolved in TFE to a concentration of 764 μM and then diluted to final concentrations of 23–24 μM with either H_2O or TFE to the desired H_2O:TFE ratio. The spectra represent the following ratios of H_2O:TFE; 100% H_2O (————), 60% H_2O/40% TFE (----), 40% H_2O:60% TFE (– – –), 20% H_2O:80% TFE (------) and 100% TFE (————).

^{1}H NMR spectra were acquired on a 500 MHz Varian Unity spectrometer. All spectra were referenced to internal DSS at 0.0 ppm. Typical acquisition parameters for one dimensional spectra were, acquisition time, 2.0 s; spectral window, 6000 Hz; 90° pulse width, 11.5 μs and 1024 transients. Phase-sensitive DQF-COSY [31], TOCSY [32,33] and NOESY spectra were collected using 64 transients and 512 complex data points in F_2 and 256 increments in F_1 yielding acquisition times of 0.171 s in t_2 and 0.043 s in t_1. TOCSY spectra were collected for 30 and 50 ms mixing times. NOESY spectra [34] were recorded in the hypercomplex mode [35] for mixing times of 150, 300, and 400 ms. In all experiments, residual HDO was suppressed with a weak 2.0 s presaturation pulse. Spectra were processed using the program Vnmr (Varian) on a Silicon Graphics XS24 workstation using either sine-bell or $\pi/8$ shifted sine-bell weighting functions, with baseline correction in the F_2 dimension.

^{2}H NMR spectra were acquired at 76.7 MHz using the quadrupolar echo sequence without composite pulses. Typically, 5.2 ms 90° pulses were used with 30 μs delays between pulses and prior to acquisition. A 100 ms recycle delay was used. The spectral window was 100 kHz and data was processed using 100 Hz line broadening.

3. Results

The 34-residue hEGFRp comprised residues I622-T654 from human EGFR with residue S121 replaced by K. It included the proposed membrane spanning region, residues I622-M644, and also T654, the site of phosphorylation by protein kinase C. The N-terminal residue, K621, was biotinylated for use in future studies. For ^{2}H NMR studies, deuterated amino acids were incorporated at positions A623, M644 and V650 using analogues deuterated at the methyl groups in

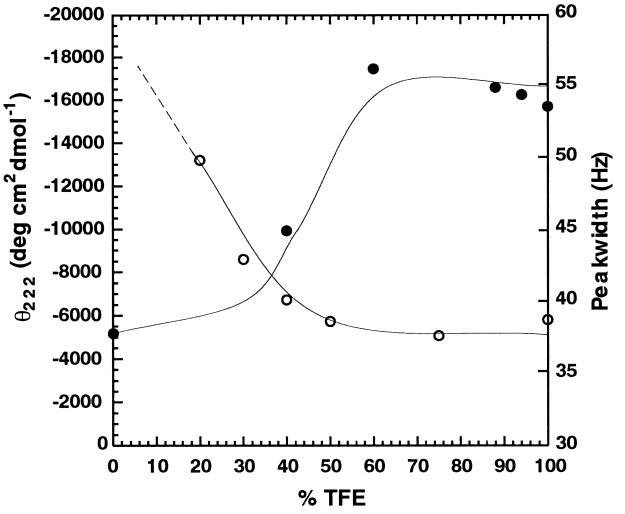

Fig. 2. Effect of TFE on θ_{222} of circular dichroism spectra (●) and on peakwidths of selected resonances in ^{1}H NMR spectra (○). The plot shows the observed θ_{222} for a given TFE concentration for hEGFRp at 25°C, pH 7.09 (left). Peakwidths were measured as a function of % TFE using the aliphatic methyl region of the spectra centred about 0.96 ppm for a 1 mM hEGFRp sample in D_2O, pH 7.09, 25°C (right).

each amino acid. Based on hydropathy analysis, this should place residues A623 and M644 near the lipid bilayer surfaces and residue V650 beyond the surface. An important aspect of this work was to define the secondary structure of the peptide in the bilayer and in the cytosolic region adjacent to the membrane. This can not be determined by conventional ^{2}H NMR spectroscopy alone. As a stepping stone towards these studies, the conformation of the hEGFRp was initially determined using high resolution NMR techniques in a lipomimetic solvent.

3.1. Production of a monomeric hEGFRp species

In solution, hydrophobic membrane-spanning peptides such as the hEGFRp tend to be insoluble or aggregate—the aggregates frequently being composed of unfolded forms of the peptide or multimeric β-sheet like structures in which the secondary structure likely does not represent the native fold [36,37]. Our initial studies of hEGFRp focused on the production of a monomeric form of the peptide in solution and in lipid bilayers. This step was essential since it was important to establish a method which preserved the native peptide secondary structure while producing a molecular co-dispersion with lipid. The secondary structure of hEGFRp was determined in solution so this information could be used in solid state analyses [28].

Although hEGFRp contains 20 hydrophobic residues (of 34 total), it was found to be water soluble over a broad pH range (2.3–9.1). High resolution ^{1}H NMR spectra of 1 mM solutions of the peptide (data not shown) were characterized by broad linewidths, independent of the pH and affected little by temperature. hEGFRp was subsequently found to dissolve rapidly in the organic solvent trifluoroethanol (TFE) which has been pioneered by Gullick

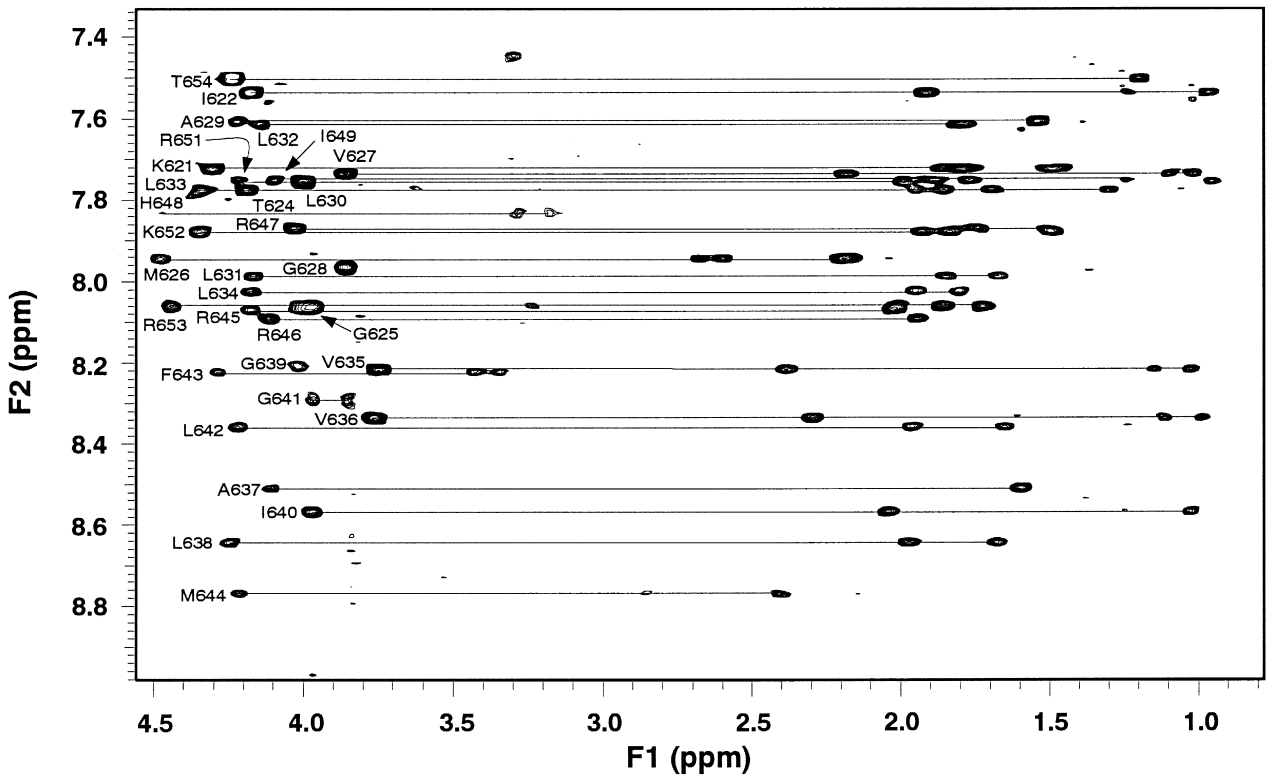

Fig. 3. Region of the 500 MHz TOCSY spectrum showing correlations between NH protons and other intraresidue protons for a mixing time of 50 ms. Data was acquired using 1.0 mM hEGFRp in 75% TFE:25% H_2O, pH 6.16 at 37°C. Spin systems are identified near the NH α correlation for each residue and through bond connectivities indicated by lines. Some spin systems cannot be identified at this contour level.

et al. [20], Mulvet et al. [22], and Norwood et al. [38] to study hydrophobic peptides including those derived from other membrane proteins. Unlike the aqueous samples, it provided ^{1}H NMR spectra with narrow linewidths (7–10 Hz) similar to those obtained for other 34-residue peptides [39]. Conveniently, TFE also dissolved the lipids of interest for solid state work as has been previously shown by Opella et al. [40]. These qualities made TFE the medium of choice for characterization of the solution structure of the transmembrane hEGFRp and preparation and characterization of the peptide in lipid bilayers.

To better quantify the ^{1}H NMR observations described above, the secondary structure of hEGFRp was studied by CD spectroscopy in lipomimetic environments comprised of various ratios of TFE and H_2O. The resulting CD spectra for these experiments are shown in Fig. 1. The spectrum of hEGFRp in an aqueous environment showed a negative band at 200 nm indicative of an unstructured peptide in aqueous solution. As the proportion of TFE was increased the

Table 1
^{1}H chemical shift data for the hEGFRp[a]

Residue	NH	αCH	βCH	γCH	Others
Biotin					4.35, 4.56, 2.96, 3.23
K621	7.73	4.36	1.95, 1.80	1.51	ε H:2.38
I622	7.54	4.17	1.92	1.53	γ' 1.22, δ 0.96
A623(d-4)	7.88				
T624	7.78	4.17	4.34		γ 1.29
G625	8.06	3.97, 4.01			
M626	7.95	4.47	2.19	2.60, 2.68	
V627	7.74	3.86	2.19		γ 1.09, γ'' 1.00
G628	7.97	3.86			
A629	7.60	4.18	1.54		
L630	7.75	4.00	1.99, 1.85	1.78	δ 1.10, δ' 0.94
L631	7.99	4.17	1.85, 1.82	1.66	δ 1.02, δ' 0.93
L632	7.61	4.14	1.83, 1.68	1.78	δ 1.12, δ' 0.93
L633	7.77	4.30	1.92, 1.70	1.54	δ 1.30, δ' 0.95
L634	8.02	4.21	1.96, 1.83	1.65	δ 0.98, δ' 0.93
V635	8.22	3.75	2.39		γ 1.14, γ' 1.02
V636	8.34	3.76	2.29		γ 1.12, γ' 0.99
A637	8.51	4.18	1.60		
L638	8.65	4.24	1.97, 1.68	1.88	δ 1.15, δ' 0.93
G639	8.21	4.01, 3.90			
I640	8.57	3.97	2.04	1.85, 1.22	γ 1.02, γ' 0.88
G641	8.29	3.97, 3.86			
L642	8.36	4.17	1.94, 1.85	1.59	γ 1.14, γ' 0.93
F643	8.23	4.28	3.43, 3.35		δ 7.28, ε 7.42, ζ 7.34
M644(d-3)	8.78	4.26	2.40, 2.14		
R645	8.07	4.17	2.01, 1.85	1.71	δ 3.20
R646	8.09	4.12	1.94, 1.93	1.82, 1.70	δ 3.15
R647	7.87	4.02	1.94, 1.83	1.72, 1.50	
H648	7.83	4.47	3.28, 3.17		^{2}H 7.65, ^{4}H 7.42
I649	7.75	4.08	1.97	1.24, 1.10	γ 0.94
V650(d-8)	8.05				
R651	7.75	4.17	1.87, 1.75	1.67, 1.54	
K652	7.88	4.34	1.92, 1.83	1.49	δ 1.77, 1.72
R653	8.06	4.44	1.99, 1.85	1.71	
T654	7.52	4.18	4.24		γ 1.19

[a]Referenced to DSS at 0 ppm in 75% TFE:25% H_2O at 37°C.
[b]Deuterated amino acid residues are [d_4]-A623, [d_3]-M644 and [d_8]-V650.

band intensities at 208 and 222 nm increased in magnitude and the spectrum of hEGFRp in 100% TFE was characteristic of a peptide with a significant proportion of α-helix present. It was also apparent that the CD spectra at 60–100% TFE were very similar having $\theta_{222} = -16485 \pm 740$ deg cm^2 dm^{-1}. These results are summarized in Fig. 2 by a smooth transition from minimum magnitude of θ_{222} at 100% H_2O to a maximum amplitude at 60–100% TFE. An isodichroic point near 203 nm was observed consistent with a two-state coil-helix transition for the peptide as the amount of TFE was increased [41].

Experiments involving similar solvent ratios of TFE and H_2O were also performed using ^{1}H NMR spectroscopy although at about 50-fold higher concentrations than CD spectra. An initial 1 mM hEGFRp sample was prepared in TFE, titrated with H_2O, and peakwidths were measured as a function of % TFE (Fig. 2). At low TFE concentrations all resonances were very broad. Due to this excessive broadening of the peaks, the cluster of resonances in the methyl region was measured as a function of % TFE, rather than any individual peak. At the higher TFE concentrations (> 50%), the spectral lines exhibited minimal

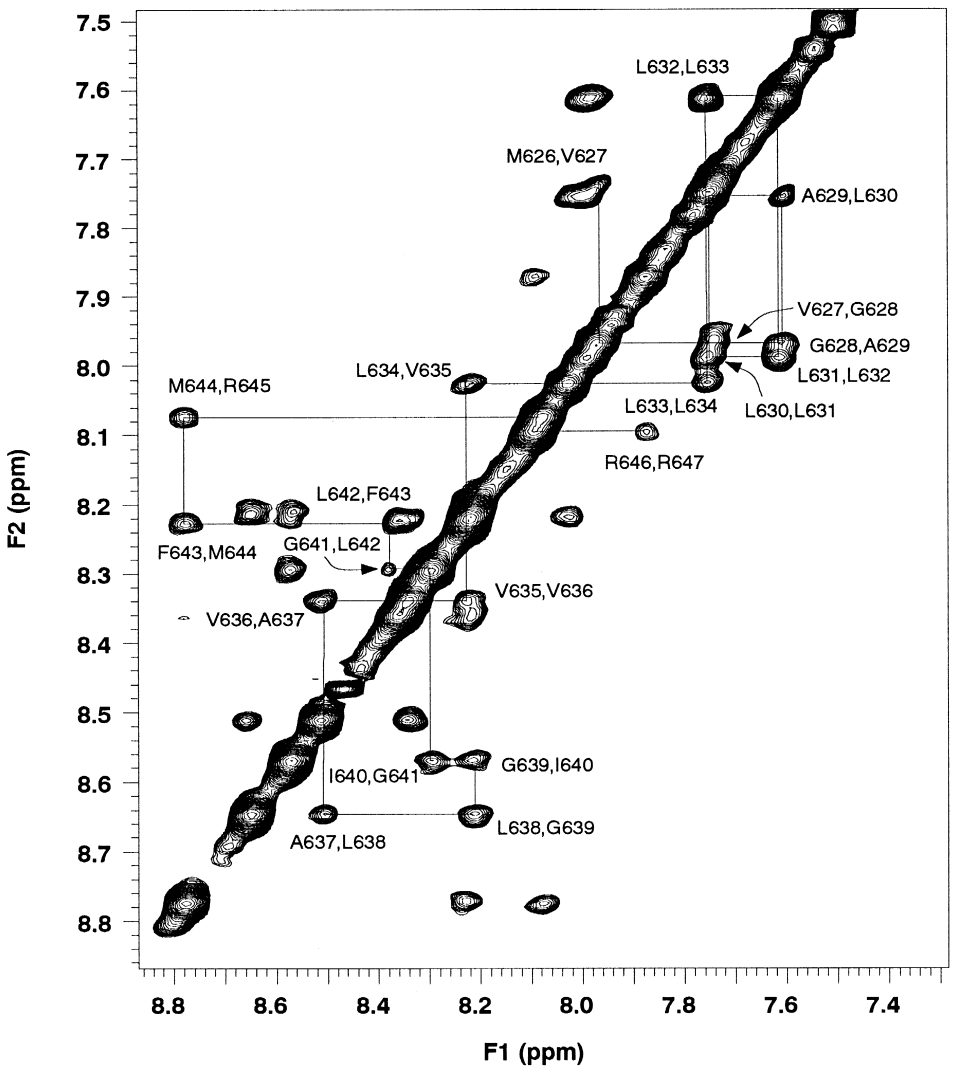

Fig. 4. NH,NH region from the 150 ms NOESY spectrum of 1.0 mM hEGFRp in 75% TFE:25% H_2O, pH 6.16 and 37°C. Sequential $d_{NN}(i,i+1)$ connectivities are traced from M626 to R647. For regions where peaks are congested, spectra were acquired at 25°C to shift several resonances and facilitate correlations.

shifting and had approximately the same peakwidth (Fig. 2) of 37.6 ± 1 Hz. It is clear from this data that a decrease in peakwidth for hEGFRp parallels an increase in magnitude of θ_{222} over similar aqueous TFE combinations.

3.2. Structure of hEGFRp in lipomimetic media

Both CD and ^{1}H NMR titration experiments showed that aqueous solutions containing > 60% TFE were an optimal membrane-mimetic medium where the hEGFRp was soluble. This combination possessed a high non-polar content to mimic a lipid environment and some aqueous component to mimic the extracellular/cytosolic environments. This was important because the entire hEGFRp sequence was not expected to lie in the membrane but rather 8–10 residues were expected to reside in the aqueous cytoplasm. Consequently, the hEGFRp was prepared in an aqueous 75% TFE (v/v) medium at pH 6.16 for further ^{1}H NMR characterization.

The sequential assignment of the hEGFRp was accomplished using a variety of standard two-dimensional NMR techniques [42,43]. In the 'finger print' region of the DQF–COSY spectrum, 29 out of a possible 32 residues were observed. Crosspeaks expected for A623 and V650 were absent due to deuteration at the α carbon site. A TOCSY spectrum, Fig. 3, allowed the identification of intraresidue connectivities, and proved valuable for residues which had extended sidechains such as I622, V627, V635, and V636. The sequential assignment of the hEGFRp used spin system assignments from TOCSY spectra and a 300 ms NOESY spectrum. Table 1 provides the chemical shifts of the ^{1}H resonances assigned for the hEGFRp.

The secondary structure in the hEGFR peptide was determined from the identification of a large number of $d_{NN}(i,i+1)$, $d_{\alpha N}(i,i+3)$ and $d_{\alpha\beta}(i,i+3)$ connectivities. Fig. 4 shows the d_{NN} region from the 150 ms NOESY spectrum of hEGFRp indicating a contiguous stretch of nOes from M626 to R647. The data shows many well resolved resonances between 8.0–8.8 ppm with many intense NH–NH connectivities. Several closely spaced NH resonances between 7.9–8.0 ppm made analysis of connectivities in this region difficult at 37°C. However, at 25°C some of these overlapping peaks were well resolved allowing unambiguous identification. In addition, stretches of $d_{\alpha N}(i,i+3)$ and $d_{\alpha\beta}(i,i+3)$ connectivities were identified between residues G625 and R647, Fig. 5. Many of these residues also exhibited $^3J_{HN\alpha} < 6$ Hz consistent with an α-helical structure present in this region. Use of the CSI for αCH resonances showed a systematic pattern of upfield shifted resonances relative to random coil values also indicative of α-helix formation [44]. The exception to this was the small sequence, G–I–G, whose αCH chemical shifts showed no preference for regular secondary structure although characteristic $d_{\alpha N}(i,i+3)$ crosspeaks were observed.

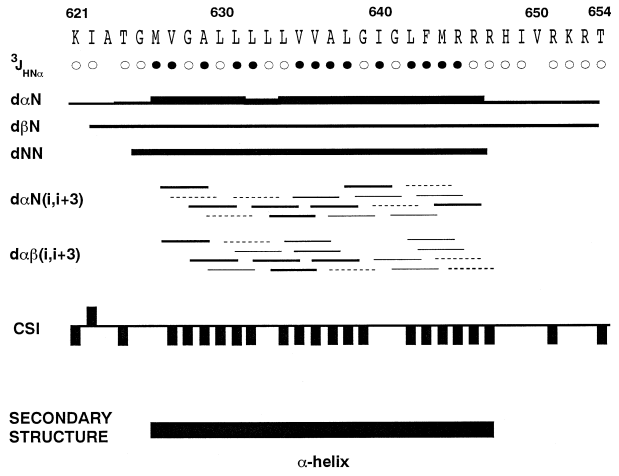

Fig. 5. Amino acid sequence and summary of NOE data for the hEGFR peptide taken from 150 ms NOESY spectra. The thickness of the various lines corresponds to the approximate intensity of the measured NOE. NOEs that can not be unequivocally identified due to resonance degeneracy are indicated with dashed lines. Backbone coupling constants ($^3J_{HN\alpha}$) were measured from DQF–COSY spectra for resolved resonances. Coupling constants are indicated for < 6 Hz (•) and 6–8 Hz (◦). No coupling constants > 8 Hz were observed. CSI refers to the chemical shift index [44] for αCH resonances, except A623 and V650 which are α-deuterated.

3.3. ^{2}H NMR analysis of hEGFRp in lipid bilayers

With an α-helical secondary structure between M626 and R647 defined in hEGFRp, it was now important to apply this information to the placement of hEGFRp in a synthetic membrane environment. The hEGFRp was incorporated into POPC lipid bilayers and the resulting assemblies studied by ^{2}H NMR spectroscopy, Fig. 6. A first method utilized

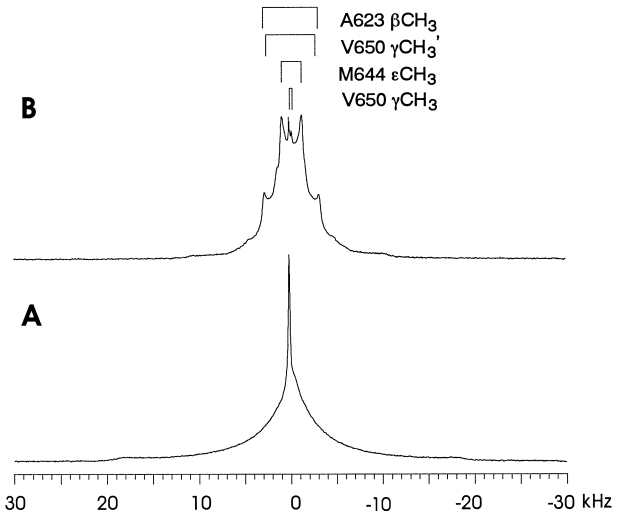

Fig. 6. Regions of the ^{2}H wideline NMR spectra of hEGFRp in POPC including 30 mM HEPES buffer, 20 mM NaCl and 5 mM EDTA at 30°C and pH 7.2. The peptide was assembled into unsonicated bilayers of POPC at a peptide/lipid molar ratio of 3:50. Spectrum A shows the ^{2}H NMR spectrum obtained from a sample prepared by detergent dialysis as described in the text. Spectrum B shows the ^{2}H NMR spectrum obtained from a sample prepared from hydration of a thin film generated by removal of solvent from peptide/lipid solutions in TFE. In spectrum B, the quadrapolar splittings for A623 βCH_3 (6.2 kHz), V650 γCH_3 (5.6 kHz), M644 εCH_3 (3.2 kHz) and Val γCH_3 (1.2 kHz) are indicated above the Pake pattern as determined from labelling of the individual residues [28].

traditional aqueous detergent dialysis and yielded a spectrum (Fig. 6A) which was broad and featureless. In a second method, using an approach similar to preparation of high resolution NMR samples, hEGFRp was dissolved first in TFE and then POPC added to the solution. A similar method has been previously described by Opella et al. [40]. In contrast to the detergent dialysis method, removal of the solvent and subsequent rehydration yielded a more conventional 'Pake' spectrum (Fig. 6B) representative of axially symmetric motion of the hEGFRp about an axis near to or perpendicular to the bilayer. Similar ^{2}H spectra have been shown for gramicidin [45,46] and the fd coat protein [47] containing deuterated methyl groups. The ^{2}H NMR spectrum of hEGFRp resulted from deuterium labels at A623 (αCD, βCD_3), M644 (εCD_3) and V650 (αCD, βCD and $\gamma\gamma' CD_3$). As indicated in Fig. 6B the quadrupolar splittings for the deuterated methyl groups for A623, M644 and V650 were intense and well resolved. Unambiguous assignment of these Pake doublets were made using a series of monodeuterated peptides [28]. Under these conditions, the largest observed splitting was 6.2 kHz arising from the A623 β-CH_3. The εCH_3 for M644 had a splitting of 3.2 kHz and the non-degenerate γCH_3 groups for V650 had splittings of 5.6 and 1.2 kHz.

4. Discussion

The observation that hEGFRp is soluble in aqueous solution is interesting given the hydrophobicity of the peptide. This solubility likely arises from the high density of positively charged residues near the C-terminus especially at pH < 7. Under these conditions, hEGFRp exists in a highly aggregated form as judged by NMR spectroscopy which yielded broad featureless resonances. This is supported by the increased linewidths in hEGFRp (> 20 Hz) compared to monomeric peptides (8–10 Hz) of a similar size [39]. hEGFRp is mostly unstructured in aqueous solution, having a CD negative intensity at θ_{200}, little evidence for α-helix ($\theta_{222} > -5000$ deg cm^2 dm^{-1}) or β-sheet ($\theta_{215} > -8000$ deg cm^2 dm^{-1}), and no NH or αCH resonances downfield of 9 ppm or 5 ppm, respectively. Consistent with these results the lack of quadrupolar splitting in the ^{2}H NMR spectrum of hEGFRp in POPC (when prepared from aqueous solutions) is indicative of non-axially symmetric motion arising from a disordered or aggregated form of the peptide. These observations are in agreement with other studies which show many hydrophobic peptides form aggregated complexes in aqueous solution [36].

4.1. hEGFRp is helical in solution and bilayers

In a lipomimetic environment comprised of aqueous TFE, hEGFRp forms a monomeric, α-helical structure as determined from CD and NMR data. TFE was the solvent of choice for these studies because it is known to stabilize the formation of α-helices in peptides which have the propensity to form an α-helix [48–50] such as transmembrane regions and it has

recently been demonstrated to be selective in that it will not induce or impose secondary structural restraints on regions that are normally unstructured in proteins [51–53]. For hEGFRp the percentage of α-helix plateaued at approximately 60% TFE. Since no further increase in θ_{222} was observed (within experimental error) at TFE concentrations greater than 60%, these findings suggest that TFE does not induce additional α-helix formation in regions of the peptide that are unstructured or would not normally form a helical segment [49,51–53]. Analysis of the CD data shows approximately 19 of the 34 residues form a contiguous α-helix in hEGFRp in this lipomimetic environment. Under similar conditions NMR data shows an α-helix comprising 22 residues and spanning positions M626 to R647 in hEGFRp.

In POPC bilayers (prepared from TFE solutions) the ^{2}H NMR spectrum of hEGFRp exhibits a typical 'Pake' pattern. The nature of this pattern is characteristic of a peptide adopting an α-helix in bilayers and undergoing axially symmetric rotation as has been observed in gramicidin [45,46], fd coat protein [47] and Pf1 protein [26]. This observation indicates that the α-helical nature of hEGFRp in 60% aqueous TFE is representative of the secondary structure found in bilayers. The magnitude of the quadrupolar coupling constant for A623 βCD_3 (6.2 kHz) is larger than expected for mobile alanine residues undergoing isotropic motion as observed in ^{2}H NMR studies of the fd helical coat protein near the N-terminus [54]. However, it is also greater than that expected for a stable helical conformation oriented perpendicular to the bilayer where an A623 $\alpha C-\beta CD_3$ angle of about 56° with the helical long axis would occur ($\Delta\nu \approx 1.4$ kHz). This would indicate that although helical, hEGFRp may adopt an orientation near A623 some 10° from the bilayer normal in a fashion more reminiscent of that found in the *neu* transmembrane region [24]. Alternatively, it is possible that the conformation of A623 in lipid bilayers is an average of multiple conformations in agreement with solution NMR data which indicates this region is not clearly α-helical.

4.2. Location of the EGFR transmembrane helix

The α-helix length (26 residues) and location (M626–R647) in hEGFRp were compared with structural predictions for the α-helix membrane spanning region of EGFR (Table 2) and general observations for type I membrane spanning protein sequences. In general, the length of the α-helix observed from NMR data (26 residues) is longer than most predictive methods (18–23 residues) which predict only the membrane spanning region. This indicates that some α-helix must be found outside the membrane environment.

Table 2
Transmembrane regions of EGFR

Method	Positions	No. of residues
This work	M626–R647	22
Jones et al. [15]	I619–F643	25
Persson and Argos [14][a]	S621–I640	20
Rost et al. [12,13][b]	I622–M644	23
	T624–G641	18
Hydropathy	I622–M644	23
TmBASE [67][c]	G625–M644	20
SOSUI[d]	S621–M644	24
DAS [68][e]	G625–L642	18

[a] http://www.embl-heidelberg.de/sspred/ssp_sin.html.
[b] http://www.embl-heidelberg.de/predictprotein/predictprotein.html.
[c] http://ulrec3.unil.ch/software/TMPRED_form.html.
[d] http://www.tuat.ac.jp/~mitaki/adv_sosui.
[e] http://www.biokemi/su/se/~server/DAS.

Examination of our data and Table 2 reveals the location of the transmembrane region of hEGFRp. Both NMR analysis and a majority of predictive methods indicate the region between M626 and M644 is helical and spans the membrane. A majority of the predictive methods indicate M644 is the terminating helical residue. In agreement with this, a preference for phenylalanine residues (F643) has been noted [14,55–57] at the interior membrane interface of type I transmembrane proteins. High resolution NMR observations show that hEGFRp is α-helical in this region satisfying hydrogen bonding requirements for the predicted portion of the transmembrane helix.

Most predictive methods indicate the α-helical transmembrane region in the EGFR includes I622 at the N-terminus. This is earlier in the sequence than identified by NMR methods (M626) and shows some ambiguity exists in this region. This supports observations from the present ^{2}H NMR studies which showed axially symmetric motion, as would be ob-

served for an α-helix, but were inconclusive for orientation. In agreement with this, previous studies from computational conformational analysis of EGFR and related peptides show a non-helical preference for regions A624–G625 and V627–G628 [58,59]. Further, analysis of type I membrane spanning sequences suggests that P620 and S621 (K621 in hEGFRp) lie outside the membrane while threonine (T624) is usually not found in α-helical membrane spanning regions but is more characteristic of an N-capping residue [56]. In addition, EGFR does not have any acidic amino acids near this region which may also be important for helix dipole stabilization [60], although this has been noted for other type I membrane spanning proteins [57].

Alternatively, the α-helix in hEGFRp may be truncated at the N-terminus compared to that predicted because the length of peptide does not provide sufficient $i,i+4$ hydrogen-bonding to maintain the α-helix for residues K621–T624. Similar observations have been made in high resolution studies of the *neu* protein and magainin where the first and last 3–4 residues are not helical [22,23]. As has been pointed out by Reithmeier [57], an α-helical transmembrane segment would need to extend for one complete turn outside the membrane to satisfy all hydrogen bonding requirements.

4.3. Evidence for a cytoplasmic helix

The 34-residue length of hEGFRp is longer than would typically traverse a bilayer allowing a structural examination of the region at the membrane–cytosol interface. While a majority of the transmembrane predictive methods provide M644 as the transmembrane helix terminating residue, it is interesting that residues M644–R647 in hEGFRp are α-helical in a lipomimetic medium. Based on transmembrane spanning predictions (Table 2) this region would lie to the cytosolic side of the membrane. The three sequential arginine residues at positions R645, R646 and R647 are consistent with the positive inside rule for type I transmembrane sequences acting as a stop transfer signal for membrane insertion [2,56,61,62]. In addition, the position of the arginine residues in the helix agrees with a preference for basic residues near the C-cap of an α-helix [60] acting to neutralize the helix dipole [63–65].

In POPC bilayers, wideline ^{2}H NMR spectra indicate the εCH_3 group of M644 has a very small quadrupolar coupling constant (3.2 kHz) indicating its motion is not isotropic. This is consistent with the solution NMR results indicating an α-helix in this region. The magnitude of the coupling constant likely results from reorientation about the Cα–Cβ, Cβ–Cγ and Cγ–S bonds yielding a more mobile side chain and is similar to that noted in bacteriorhodopsin for a surface exposed methionine [66]. Together, these results support methods predicting M644 has a helical backbone conformation and resides at the membrane/fluid interface.

Solution NMR data indicates the α-helical structure of hEGFRp ends shortly before V650 at position R647. Supporting this, wideline ^{2}H NMR results for V650 provide coupling constants of 5.6 and 1.2 kHz for the two non-degenerate γCD_3 groups indicating that backbone and sidechain motion is not isotropic as would be expected from rapid reorientation of this residue in solution. However, there is no indication of a large coupling constant similar to those observed for interior membrane spanning residues of gramicidin (26–27 kHz) for valines suggested to be in a fixed rotameric state on a rigid backbone frame [45,46]. Rather, the δCD_3 splittings for V650 are more reminiscent of observations for V1 in gramicidin where it was suggested that 'wobbling' of the peptide about an axis perpendicular to the bilayer occurs in addition to three-site flipping of the γCD_3 groups between a preferred conformation at $\chi_1 = 180°$ and conformers at $\chi_1 = \pm 60°$ [45,46]. In this respect it would be useful to have dynamics information about the backbone region of V650 to identify whether an α-helix is extended past R647 or whether the observed quadrapole coupling is a result of backbone and sidechain motional averaging as observed in solution NMR data.

Acknowledgements

We would like to thank Kathy Barber for her help in preparing this manuscript and Dr. Lana Lee (University of Windsor) for her help with CD experiments. This research was supported by a grant from the Medical Research Council of Canada (GSS) and an Ontario Graduate Studentship awarded to ACR.

Funding for the NMR spectrometer in the McLaughlin Macromolecular Structure Facility was made possible through grants from the Medical Research Council of Canada and the Academic Development Fund of The University of Western Ontario and generous gifts from the R. Samuel McLaughlin Foundation and London Life Insurance of Canada.

References

[1] V.T. Marchesi, Adv. Exp. Med. Biol. 205 (1986) 107–120.
[2] C.J. Brandl, R.B. Deber, L.C. Hsu, G.A. Wooley, X.K. Young, C.M. Deber, Biopolymers 27 (1988) 1171–1182.
[3] Y. Yarden, A. Ullrich, Biochemistry 27 (1988) 3113–3119.
[4] M.D. Hollenberg, FASEB J. 5 (1991) 178–186.
[5] W.J. Fantl, D.E. Johnson, L.T. Williams, Annu. Rev. Biochem. 62 (1993) 453–481.
[6] A.C. King, P. Cuatrecasas, J. Biol. Chem. 257 (1982) 3053–3060.
[7] T. Hunter, J.A. Cooper, Annu. Rev. Biochem. 54 (1985) 897–930.
[8] A. Ullrich, J.S. Coussens, T.J. Hayflick, A. Gray, A.W. Tam, J. Lee, Y. Yarde, T.A. Libermann, J. Schlessinger, J. Downward, E.L. Mayes, N. Whittle, M.D. Waterfield, P.H. Seeburg, Nature 309 (1984) 418–425.
[9] C.D. Carpenter, H.A. Ingraham, C. Cochet, G.M. Walton, C.S. Lazar, J.M. Sowadski, M.G. Rosenfeld, G.N. Gill, J. Biol. Chem. 266 (1991) 5750–5755.
[10] T. Hunter, N. Ling, J.A. Cooper, Nature 314 (1984) 480–483.
[11] K.R. MacKenzie, J.H. Prestegard, D.R. Engelman, Science 276 (1997) 131–133.
[12] B. Rost, R. Casadio, P. Fariselli, C. Sander, Protein Sci. 4 (1995) 521–533.
[13] B. Rost, P. Fariselli, R. Casadio, Protein Sci. 5 (1996) 1704–1718.
[14] B. Persson, P. Argos, J. Mol. Biol. 237 (1994) 182–192.
[15] D.T. Jones, W.R. Taylor, J.M. Thornton, Biochemistry 33 (1994) 3038–3049.
[16] R.X. Xu, J.M. Word, D.G. Davis, M.J. Rink, D.H. Willard, R.T. Gampe Jr., Biochemistry 34 (1995) 2107–2121.
[17] S.S. Narula, R.W. Yuan, S.E. Adams, O.M. Green, J. Green, T.B. Philips, L.D. Zydowsky, M.C. Botfield, M. Hatada, E.R. Laird, M.J. Zoller, J.L. Karas, D.C. Dalgarno, Structure 3 (1995) 1061–1073.
[18] M. Overduin, B. Mayer, C.B. Rios, D. Baltimore, D. Cowburn, Proc. Natl. Acad. Sci. U.S.A. 89 (1992) 11673–11677.
[19] Y.Q. Gosser, J. Zheng, M. Overduin, B.J. Mayer, D. Cowburn, Structure 3 (1995) 1075–1086.
[20] W.J. Gullick, A.C. Bottomley, F.J. Lofts, D.G. Doak, D. Mulvey, R. Newman, M.J. Crumpton, M.J.E. Sternberg, I.D. Campbell, EMBO J. 11 (1992) 43–48.
[21] A.S. Arseniev, I.V. Maslennikov, V.F. Bystrov, A.T. Kozhich, V.T. Ivanov, Y.A. Ovchinnikov, FEBS Lett. 231 (1988) 81–88.
[22] D. Mulvey, G.F. King, R.M. Cooke, D.G. Doak, T.S. Harvey, I.D. Campbell, FEBS Lett. 257 (1989) 113–117.
[23] J. Gesell, M. Zasloff, S.J. Opella, J. Biomol. NMR 9 (1997) 127–135.
[24] S.O. Smith, B.J. Bormann, Proc. Natl. Acad. Sci. U.S.A. 92 (1995) 488–491.
[25] A.E. McDermott, F. Creuzet, R. Griffin, L.E. Zawadzke, Q.Z. Ye, C.T. Walsh, Biochemistry 29 (1990) 5767.
[26] K.J. Shon, Y. Kim, L.A. Colnago, S.J. Opella, Science 252 (1991) 1303–1305.
[27] R. Nambudripad, W. Stark, S.J. Opella, L. Makowski, Science 252 (1991) 1305–1308.
[28] A.C. Rigby, K.R. Barber, G.S. Shaw, C.W.M. Grant, Biochemistry 35 (1996) 12591–12601.
[29] R.W. Woody, Methods Enzymol. 246 (1995) 34–71.
[30] C.T. Chang, C.-S.C. Wu, J.T. Yang, Anal. Biochem. 91 (1978) 13–31.
[31] M. Rance, O.W. Sorensen, G. Bodenhausen, G. Wagner, R.R. Ernst, K. Wuthrich, Biochem. Biophys. Res. Commun. 117 (1983) 479–485.
[32] L. Braunschweiler, R.R. Ernst, J. Magn. Reson. 53 (1983) 521–528.
[33] A. Bax, D.G. Davies, J. Magn. Reson. 65 (1985) 335–360.
[34] J. Jeener, B.H. Meier, P. Bachmann, R.R. Ernst, J. Chem. Phys. 71 (1979) 4546–4553.
[35] D.J. States, R.A. Haberkorn, D.J. Ruben, J. Magn. Reson. 48 (1982) 286–292.
[36] J.A. Killian, T.P. Trouard, D.V. Greathouse, V. Chupin, G. Lindblom, FEBS Lett. 348 (1994) 161–165.
[37] G.D. Henry, B.D. Sykes, Methods Enzymol. 239 (1994) 515–535.
[38] T.J. Norwood, D.A. Crawford, M.E. Stevenson, P.C. Driscoll, I.D. Campbell, Biochemistry 31 (1992) 6282.
[39] G.S. Shaw, R.S. Hodges, B.D. Sykes, Biopolymers 32 (1992) 391–397.
[40] S.J. Opella, Y. Kim, P. McDonnell, Methods Enzymol. 239 (1994) 536–560.
[41] N. Greenfield, G.D. Fasman, Biochemistry 8 (1969) 4108–4116.
[42] K. Wuthrich, G. Wider, G. Wagner, W. Braun, J. Mol. Biol. 155 (1982) 311–319.
[43] M. Billeter, W. Braun, K. Wuthrich, J. Mol. Biol. 155 (1982) 321–346.
[44] D.S. Wishart, B.D. Sykes, Methods Enzymol. 239 (1994) 363–392.
[45] K.C. Lee, T.A. Cross, Biophys. J. 66 (1994) 1380–1387.
[46] K.C. Lee, S. Huo, T.A. Cross, Biochemistry 34 (1995) 857–867.
[47] L.A. Colnago, K.G. Valentine, S.J. Opella, Biochemistry 26 (1987) 847–854.
[48] J.W. Nelson, N.R. Kallenbach, Biochemistry 28 (1989) 5256–5261.
[49] S.R. Lehnman, J.L. Tuls, M. Lund, Biochemistry 29 (1990) 5590–5596.

[50] M.G. Zagorski, C.J. Barrow, Biochemistry 31 (1992) 5621–5631.
[51] A.L. Breeze, T.S. Harvey, R. Bazzo, I.D. Campbell, Biochemistry 30 (1991) .
[52] S.I. Segawa, T. Fukuno, K. Fujiwara, Y. Noda, Biopolymers 31 (1991) 497–509.
[53] F.D. Sonnichsen, J.E. Van Eyk, R.S. Hodges, B.D. Sykes, Biochemistry 31 (1992) 8790–8798.
[54] K.G. Valentine, D.M. Schneider, G.C. Leo, L.A. Colnago, S.J. Opella, Biophys. J. 49 (1986) 36–38.
[55] S. Li, C.M. Deber, J. Biol. Chem. 268 (1993) 22975–22978.
[56] C. Landolt-Marticoena, K.A. Williams, C.M. Deber, R.A.F. Reithmeier, J. Mol. Biol. 229 (1993) 602–608.
[57] R.A.F. Reithmeier, Curr. Opin. Struct. Biol. 5 (1995) 491–500.
[58] P.W. Brandt-Rauf, R. Monaco, M.R. Pincus, J. Protein Chem. 13 (1994) 227–231.
[59] P.W. Brandt-Rauf, M.R. Pincus, R. Monaco, J. Protein Chem. 14 (1995) 33–40.
[60] J.S. Richardson, D.C. Richardson, Science 240 (1988) 1648–1652.
[61] R.E. Dalbey, TIBS 15 (1990) 253–257.
[62] G. von Heijne, J. Mol. Biol. 225 (1992) 487–494.
[63] D.E. Blagdon, M. Goodman, Biopolymers 14 (1975) 241–245.
[64] K.R. Shoemaker, P.S. Kim, D.N. Brems, S. Marqusee, E.J. York, I.M. Chaiken, J.M. Stewart, R.L. Baldwin, Proc. Natl. Acad. Sci. U.S.A. 82 (1985) 2349.
[65] W.G.J. Hol, P.T. van Duijnen, H.J.C. Berendsen, Nature 278 (1978) 443–447.
[66] M.P. Keniry, H.S. Gutowsky, E. Oldfield, Nature 307 (1984) 383–386.
[67] K. Hofmann, W. Stoffel, Biol. Hoppe-Seyler 347 (1993) 166.
[68] M. Cserzo, E. Wallin, I. Simon, G. von Heijne, A. Elofsson, Protein Eng. 10 (1997) 673–676.