Marshall, A. G. (1978) *Biophysical Chemistry: Principles, Techniques, and Applications*, pp 111–130, John Wiley & Sons, New York.
Mieyal, J. J., Starke, D. W., Gravina, S. A., Dothey, C., & Chung, J. (1991) *Biochemistry 30*, 6088–6097.
Papayannopoulos, I. A., Gan, Z., Wells, W. W., & Biemann, K. (1989) *Biochem. Biophys. Res. Commun. 159*, 1448–1454.
Park, E. M. & Thomas, J. A. (1989) *Arch. Biochem. Biophys.*, *272*, 25–31.
Segel, I., & Johnson, M. (1963) *Anal. Biochem. 5*, 330–337.
Stryer, L., Holmgren, A., & Reichard, P. (1967) *Biochemistry 6*, 1016–1020.
Torchinsky, Y. M. (1981) *Sulfur in Proteins*, pp 133–142, Pergamon Press, New York.
Watts, D. C., & Rabin, B. R. (1962) *Biochem. J. 85*, 507–516.
Yang, Y., & Wells, W. W. (1990a) *J. Biol. Chem. 265*, 589–593.
Yang, Y., & Wells, W. W. (1990b) *FASEB J. 4*, A2118.
Ziegler, D. M. (1985) *Annu. Rev. Biochem. 54*, 305–329.

Structure–Function Relationships in Human Epidermal Growth Factor Studied by Site-Directed Mutagenesis and ^{1}H NMR†

Ulrich Hommel,‡ Timothy J. Dudgeon,‡,§ Anthony Fallon,§ R. Mark Edwards,§ and Iain D. Campbell*,‡
Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, U.K., and British Biotechnology Ltd., Thames Court, Watlington Road, Oxford OX4 5LY, U.K.

Received April 8, 1991; Revised Manuscript Received June 12, 1991

ABSTRACT: In order to elucidate the mechanism of interaction between human epidermal growth factor (EGF) and its receptor, selected variants of EGF, differing by single amino acid substitutions, have been made by site-directed mutagenesis. The receptor affinity of these mutants was determined by a receptor binding competition assay, and the effects of the substitution on the structure of the protein were assessed by ^{1}H nuclear magnetic resonance techniques. Various substitutions of Arg-41 resulted in substantial reduction in receptor affinity of EGF whereas change of Tyr-13 did not affect binding to the receptor. The ^{1}H resonances of all nonexchangeable protons of the Tyr-13 → Leu, Arg-41 → His, and Leu-47 → Glu variants were assigned and compared in order to assess the structural integrity of these mutants, which possess very different spectral and biological properties. In the case of the Leu-47 → Glu mutant, only minor localized spectral changes were observed, confirming that the tertiary structure of the protein is preserved upon mutation. In contrast, for both the Arg-41 → His and Tyr-13 → Leu variants, significant and strikingly similar spectra changes were observed for many residues located far away from the mutated residues. This implies that similar structural alterations have taken place in both proteins, an idea further supported by hydrogen-exchange experiments where the exchange rates of hydrogen-bonded amide protons for both the Tyr-13 → Leu and the Arg-41 → His mutants were found to be about 4 times faster than in the wild-type protein. Nuclear Overhauser enhancements, indicating specific structural features of hEGF, were measured for all mutants to ascertain that gross structural changes have not occurred. A structural model is proposed, to explain the observed effects brought about by the mutations Arg-41 → His and Tyr-13 → Leu.

Human epidermal growth factor is a 53 amino acid protein that stimulates mitogenesis by activating the tyrosine kinase domain of a cell-surface receptor (Carpenter et al., 1978; Ushiro & Cohen, 1980; Ullrich & Schlessinger, 1990). Several other proteins are homologous to EGF,[1] and they are believed to activate the EGF receptor in a way similar to that of EGF by ligand-induced oligomerization of the receptor (Yarden & Schlessinger, 1987). Human TGFα, for example, competes with EGF for the EGF receptor and has properties similar to those of EGF (Derynck et al., 1984; DeLarco & Todaro, 1980). Three members of the Pox virus family also possess EGF homologous sequences, and in some cases these proteins have been shown to bind to the EGF receptor (Lin et al., 1988; Upton et al., 1987; Twardzik et al., 1985). Cell growth promoting activities have also been found for other proteins containing EGF-like sequences such as amphiregulin (Shoyab et al., 1989), heparin binding growth factor (Higashiyama et al., 1991), and fragments of laminin (Panayotou et al., 1989).

Considerable progress has been made recently in understanding the basis for the recognition of EGF by its receptor. The three-dimensional structure of mouse EGF and that of a biologically active derivative, human EGF(1–48), have been determined by a combination of high-resolution ^{1}H NMR and computational techniques (Cooke et al., 1987; Montelione et al., 1987). In addition, the secondary structure of rat EGF (Mayo et al., 1989) and various three-dimensional structures of human TGFα have also been reported (Kohda et al., 1989; Kline et al., 1990; Harvey et al., 1991). This structural knowledge and the information available from the comparison of amino acid sequences of proteins with growth factor activity

† Financial support by EMBO to U.H. is gratefully acknowledged. This is a contribution from the Oxford Centre for Molecular Sciences, which is supported by the Science and Engineering Research Council and the Medical Research Council.
* Author to whom correspondence should be addressed.
‡ University of Oxford.
§ British Biotechnology Ltd.

[1] Abbreviations: ACN, acetonitrile; COSY, correlated spectroscopy; DQF-COSY, double-quantum filtered COSY; hEGF, human epidermal growth factor; NMR, nuclear magnetic resonance; NOESY, nuclear Overhauser spectroscopy; hTGF, human transforming growth factor; RP-HPLC, reversed-phase high-performance liquid chromatography; SDS–PAGE, sodium dodecyl sulfate–polyacrylamide gel electrophoresis; TFA, trifluoroacetic acid.

0006-2960/91/0430-8891$02.50/0

led to a proposal for residues important for receptor recognition (Campbell et al., 1990).

Our aim now is to test the contribution of each of these residues to receptor binding by its replacement in a site-directed mutagenesis study. The results of such studies, however, generally depend on the assumption that the structural integrity of the variant proteins has not been affected (Knowles, 1987). This assumption is not always warranted, and structural studies should be undertaken to assess altered activities of variant proteins. At present, the only methods that give information about specific changes in mutant proteins are X-ray crystallography (Alber et al., 1988) and high-resolution ^{1}H NMR in solution (Folkers et al., 1989; Wilde et al., 1988).

We have previously combined one-dimensional ^{1}H NMR techniques and site-directed mutagenesis in our approach to understanding structure–function relationships in hEGF (Dudgeon et al., 1990). This revealed that substitutions of residue Leu-47 have a large effect on the ability of EGF to bind to its receptor but little effect on the EGF structure: a result that has also been found by others (Moy et al., 1989). Substitutions of Arg-41 and Tyr-13 have indicated the importance of these residues for receptor binding, but no sequence-specific structural information is yet available for either of these variants (Defeo-Jones et al., 1988; Defeo-Jones et al., 1989; Engler et al., 1990).

In this paper, we describe the biological and structural effects that accompany substitutions of Arg-41 and Tyr-13 in hEGF. Dramatic alterations of the receptor binding affinity were observed for all Arg-41 variants, and concomitant changes in the spectral properties for some of them questioned the preservation of their native structure. A detailed characterization of these mutants by analysis of independent parameters sensitive to the structural integrity of proteins is presented.

Experimental Procedures

Mutagenesis and Cloning. EGF mutants were produced either by site-directed mutagenesis on M13mp18 or M13mp19 templates (Kunkel et al., 1987) or by cassette mutagenesis. Oligonucleotides were designed with degeneracy at the codon to be changed, so that multiple amino acid substitutions at one site could be achieved with a single oligonucleotide primer or oligonucleotide pair. For the expression of hEGF, a synthetic hEGF gene was fused to the *Saccharomyces cerevisiae* pre-pro-α-factor leader sequence and cloned into the *Bgl*II expression site of pLF1 or the *Bgl*II–*Bam*HI expression site of pTD4 as described previously (Dudgeon et al., 1990). Mutants of the hEGF gene were verified by DNA sequencing and appropriate fragments cloned into these vectors.

Expression and Purification. EGF is expressed with use of the yeast α-factor secretion pathway that has been described previously (Brake et al., 1984; Dudgeon et al., 1990). Plasmids were transformed into *S. cerevisiae* strain MD50 (a/α *leu2/leu2 pep4-3/+ his3/+*), with leucine auxotrophy as a selectable marker. Transformants were grown at 30 °C in liquid culture containing 0.67% yeast nitrogen base without amino acids (Difco), 1% glucose, and selected amino acids. Cultures were harvested after 2–3 days. Wild-type EGF is produced predominantly as the 1–52 form of the molecule, lacking the C-terminal arginine, but smaller amounts of the 1–51 species and of both the 1–52 and 1–51 methionine sulfoxide forms are also produced (George-Nascimento et al., 1988; T. J. Dudgeon, unpublished results).

EGF and its variants were purified from the yeast cultures by a three-stage purification scheme. First, EGF was concentrated by adsorption of the cell-free medium to a Zetaprep 60-SP cation-exchange cartridge (Cuno Inc.) at pH 3.5 followed by elution with 20 mM Bis-Tris, pH 6.5. In the second step, fractions containing protein were further purified by anion-exchange HPLC with linear gradients running from 100% buffer A (20 mM Bis-Tris, pH 6.5) to 80% buffer B (buffer A + 0.5 M NaCl). Finally, EGF was separated from minor impurities of its oxidized forms on RP-HPLC columns with linear gradients running from 20 to 38% buffer B (ACN + 0.1% TFA) in buffer A (H_2O + 0.1% TFA). The purified protein was homogeneous as determined by RP-HPLC and SDS–PAGE. The concentration of purified protein was determined spectrophotometrically by use of an absorbance calculated from the contributions of tryptophans and tyrosines (optical density in a 1-cm light path at 280 nm; wild type, $A_{0.1\%}$ = 2.89; Tyr-13 → Leu, $A_{0.1\%}$ = 2.67).

Receptor Binding Assays. Radioreceptor assays were performed by use of monkey Vero cell line. Cells were grown to confluence in 24-well plates containing DMEM + 10% FCS. The growth medium was removed, and EGF samples in DMEM + 20 mM HEPES + 1 mg/mL BSA were added to wells, together with 20 nCi of ^{125}I-labeled murine EGF (Amersham), and incubated at room temperature for 1 h. Cells were washed three times in phosphate-buffered saline, to remove unbound label, and then dissolved in 1 M NaOH. Each sample was counted with a γ counter for 400 s.

NMR. The concentration of the dissolved proteins was between 0.4 and 1 mM, and the pH was adjusted to pH 2.9 (uncorrected meter readings). ^{1}H NMR spectra were acquired at 30 °C on either a Bruker AM 600, Bruker AM 500, or a hybrid 500-MHz spectrometer assembled in Oxford. All two-dimensional spectra were acquired in the pure-phase absorption mode (Redfield & Kunz, 1975; Marion & Wüthrich, 1983; States et al., 1982). NOESY spectra (Jeener et al., 1979) and HOHAHA spectra (Davis & Bax, 1985) were acquired with mixing times of 200 m and 50 ms, respectively. In H_2O solutions, solvent suppression was achieved with the jump–return sequence both for NOESY and the HOHAHA experiments (Bax et al., 1987). In D_2O solutions, the residual OH resonance was suppressed by solvent saturation. COSY spectra (Aue et al., 1976) and double-quantum filtered COSY spectra (Rance et al., 1983) were recorded with the standard pulse sequences. Typically, 512–800 increments were recorded with 64–96 scans. Spectra were generally processed on a SUN computer with programs provided in the FTNMR and FELIX software package (Hare Research, Inc.). Gaussian and squared sine-bell window functions were used as resolution enhancement functions in F_2 and F_1, respectively.

For the hydrogen-exchange experiments, COSY spectra were recorded after dissolving the lyophilized protein in D_2O (99.96%) and rapidly adjusting to pH 2.9. Six to eight COSY spectra were then acquired with 256–400 increments and 16 scans, giving a total time per spectrum of 2.5–3.5 h. The absolute values of the intensities of each component of the NH/CαH cross-peaks were extracted and averaged, giving the time-dependent variable $S(t)$. The rate of exchange was then determined from a linear regression analysis of ln ($S(t)$) versus the elapsed time t.

Results

Receptor Affinities. The affinities of the EGF mutants for the EGF receptor were measured by radioreceptor assay, and the data are summarized in Table I. A decrease in the affinity of a mutant will be reflected by a higher concentration of protein required to displace the ^{125}I-labeled murine EGF from the receptor, and hence of a higher IC_{50}. It can be seen that all substitutions of Arg-41 resulted in a substantial increase

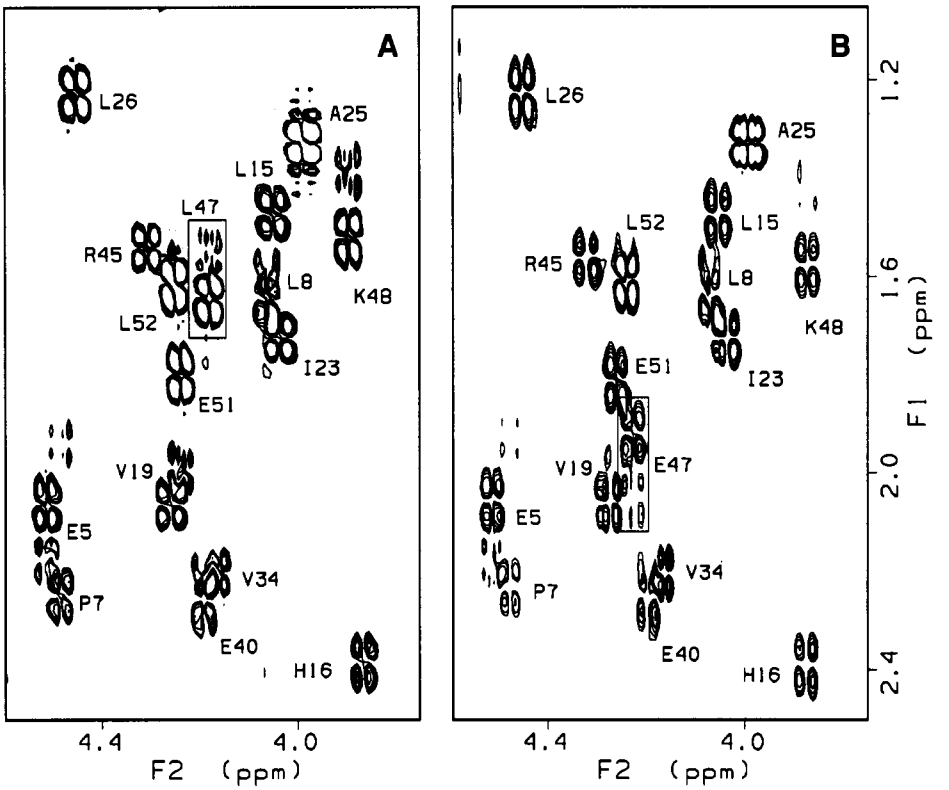

FIGURE 1: Spectral comparison of the wild-type hEGF and the Leu-47 → Glu variant of hEGF. DQF-COSY spectra of (A) wild-type hEGF and (B) the Leu-47 → Glu variant recorded in D_2O are shown. The CαH/CβH cross-peaks of the original residue, Leu-47, and the introduced residue, Glu-47, are marked by a box. Note the close resemblance of the two spectra for most resonances.

Table I: Receptor Affinities of EGF Mutants

EGF(1–52)	IC_{50} (ng/mL)	relative affinity (% wild type)
wild type	10.7 ± 1.1	100
Tyr-13 → Leu	5.3 ± 0.7	202
Arg-41 → His	5460 ± 960	0.2
Arg-41 → Leu	8440 ± 985	0.1
Arg-41 → Gln	2670 ± 326	0.4

(250–800-fold) in the IC_{50}. Mitogenic assays were also performed (data not shown), and it was found that lower mitogenic activity of mutants correlated with their reduced receptor affinities.

Assignment of the ^1H NMR Spectra of the EGF Variants Arg-41 → His, Tyr-13 → Leu, and Leu-47 → Glu. Sequence-specific structural information deduced from ^1H NMR spectroscopy requires the assignment of all the observable proton resonances of the protein (Wüthrich, 1989). In homologous proteins such as variants with a single mutation, this is a straightforward task unless the chemical environment of several protons has changed significantly. This is illustrated in Figure 1 for the Leu-47 → Glu variant, where many spin systems could quite easily be assigned by a simple comparison with the known resonance assignments for the wild-type protein (Cooke et al., 1990). Only minor changes of the CαH resonances close in sequence to the mutated Leu-47 were observed.

In contrast, significant spectral changes in the two-dimensional spectra of Arg-41 → His and Tyr-13 → Leu were observed, and this restricted a straightforward assignment of some residues. These assignments were achieved by analyzing through-space connectivities gathered from NOESY spectra acquired in water. Figure 2A shows the amide region of the NOESY spectrum for the Arg-41 → His mutant, with the sequential connectivities from residues Ser-9 to Cys-14 and those from Glu-40 to Gln-43 depicted. These resonances showed the largest deviations from their wild-type chemical shifts and could not have been assigned without the sequential information from strong NH_i–NH_{i+1} NOEs.

Of particular interest was the assignment of the new spin systems arising from the introduced residues, namely His-41, Leu-13, and Glu-47. As shown for the Leu-47 → Glu variant in Figure 1, the CαH/CβH cross-peaks arising from the introduced glutamic acid are the only obvious change in the spectrum compared with that of the wild-type protein, and hence assignment was straightforward. In contrast, the identification of the new spin system in the Arg-41 → His variant was not possible from D_2O spectra alone because of chemical shift degeneracy between the resonances for the CαH of Glu-24 and the CβH of His-41 (Figure 2B). However, its amide resonance was identified by the connectivities to the amide groups of Glu-40, Cys-42, and Gln-43 (Figure 2A). Further sequential connectivities could be identified from this amide proton to the side-chain protons of Glu 40, connectivities to which were also seen from the Cδ proton of His-41 (Figure 2B). For the introduced Leu-13 in the Tyr-13 → Leu variant, a new $A_3A'_3MPTX$ spin system was identified by COSY and NOESY cross-peaks and its assignment was confirmed by through-space connectivities to the methylene protons of Tyr-29. A table listing all the proton chemical shifts for new spin systems for the Arg-41 → His, Tyr-13 → Leu, and Leu-47 → Glu variants of hEGF, along with chemical shifts, which differ by more than the experimental error (±0.03 ppm) from their wild-type values, is given as supplementary material.

Secondary Shift. The secondary shift is defined as the difference of the chemical shift of a proton resonance observed in the native protein minus its value found in a random coil peptide (Campbell et al., 1975; Bundi & Wüthrich, 1979; Pardi et al., 1983). This parameter is known to be characteristic for the EGF-folding topology (Cooke et al., 1990), and thus it should be a sensitive probe for the structural integrity

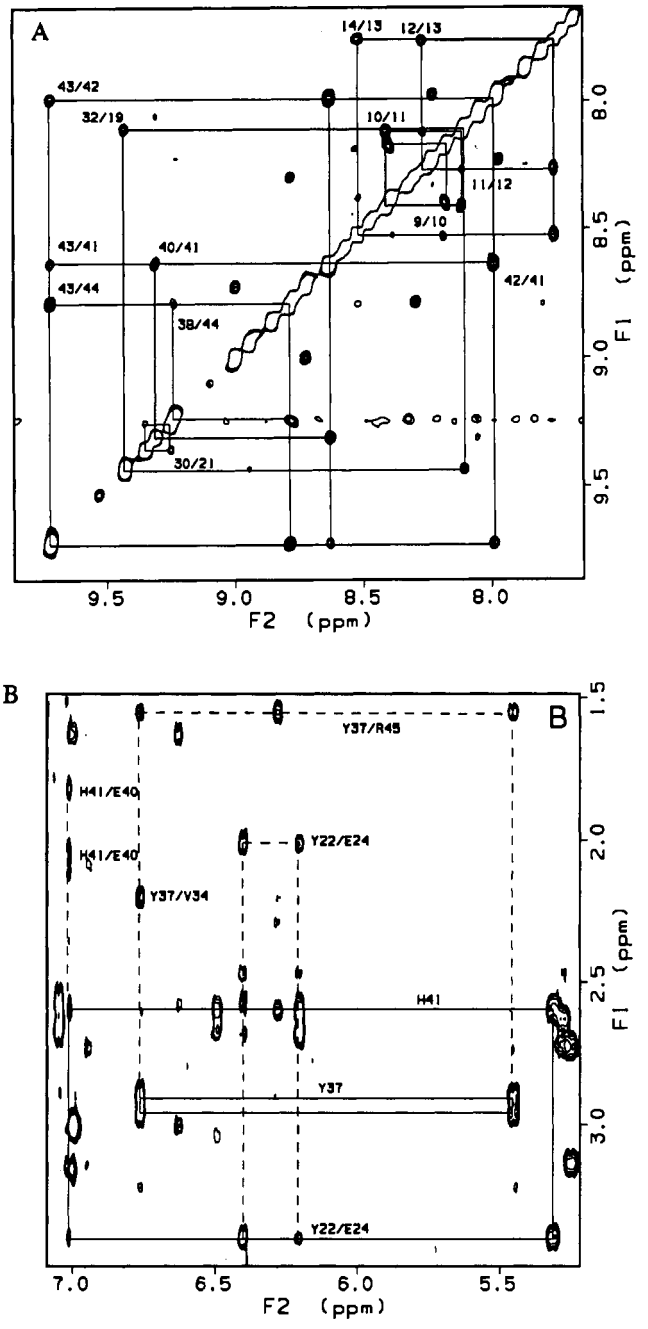

FIGURE 2: Sequential connectivities of residue His-41 in the Arg-41 → His variant. Shown are NOESY spectra acquired (A) in H_2O and (B) in D_2O at a mixing time of 200 ms. In sequential connectivities (A) between the amide protons of residues 9 and 14 as well as 40–44 and (B) from side-chain protons of the preceding residue Glu-40 to the Cδ proton of His-41 are depicted. In addition, medium- and long-range interactions, characteristic for the three-dimensional structure of EGF, can be seen in both panels.

of the variant proteins. In Figure 3A the secondary shifts for the CαH resonances are plotted as a function of the amino acid sequence for the variants Tyr-13 → Leu, Arg-41 → His, and Leu-47 → Glu. As found previously for the wild-type protein, all three variants show characteristic secondary shifts. They follow quite nicely the pattern of secondary structure known from its three-dimensional structure, which is also depicted in this figure.

Small variations between the wild-type and the mutant proteins will, however, be suppressed in such plots. Therefore, the differences between the CαH resonances of the different variants and the wild-type protein are presented in Figure 3B. It can be seen that the chemical shifts for the Leu-47 → Glu variant are very similar to those of the wild-type protein while both the Arg-41 → His and Tyr-13 → Leu variants show significant changes. These changes comprise not only residues close to the introduced spin system but also residues further apart in sequence. For instance, changes around Tyr-13 in the Arg-41 → His mutant and similar changes at Arg-41 in the Tyr-13 → Leu mutant are a characteristic feature. Similarly, both mutants show the same effects on the chemical shifts of residues Cys-20, Tyr-22, Ala-30, and Cys-31 in the major β-sheet. A detailed discussion of possible reasons will follow below.

Parameters Indicating Tertiary Structure. The observed chemical shift differences for both the Tyr-13 → Leu and Arg-41 → His mutants, occurring at both structurally and functionally important parts of the protein, bring into question the three-dimensional integrity of the EGF molecule after mutation, although the overall shift patterns are clearly similar in all the variants (Figure 3A). Therefore, the NOESY spectra of all mutants were qualitatively analyzed to obtain information on specific long-range interactions characteristic for the wild-type protein.

Some of these interactions in the Arg-41 → His variant have already been shown. Typical NOEs between the amide proton resonances of residues on opposite strands in the two β-sheets are shown in Figure 2A. A number of characteristic medium- and long-range connectivities for side-chain protons representing the spatial proximity of Val-34 and Tyr-37, Tyr-22 and Glu-24, and Tyr-37 and Arg-45 are depicted in Figure 2B. For the Tyr-13 → Leu variant, spectra were only acquired in D_2O, but again many NOEs characteristic of the EGF structure were found. In Figure 4, for example, typical CαH–CαH NOEs are shown for residues on opposite strands of the β-sheets: in particular, between Gly-18 and Cys-33, Cys-20 and Cys-31, and Tyr-37 and Arg-45.

Long-range NOEs, which give rise to specific contacts between the amino- and the carboxy-terminal domains of EGF, are found in all three variants. This interface is composed of residues Tyr-13–His-16, Tyr-37, an Arg-41–Gln-43. In all three variants, characteristic NOEs are found between the Cδ protons of Leu-15 and the Cα and Cγ protons of Gln-43, as well as between the Cδ and Cε protons of His-16 and the Cε protons of Tyr-37. In contrast to these general similarities between observed and expected NOEs in the variant proteins, one characteristic NOE has so far not been observed in the Arg-41 → His variant, that between the Cα protons of His-41 and Cys-14.

Parameters Indicating Flexibility. The solvent accessibility of backbone atoms in proteins is directly related to their internal flexibility and hence is a characteristic property of the protein. This accessibility is readily assessed by NMR by measuring hydrogen-exchange rates of amide protons (Wagner & Wüthrich, 1986) and can reveal differences between homologous proteins (Wagner et al., 1984). To this end, amide-exchange rates have been measured for residues within the secondary structural elements of wild-type EGF and the variants Tyr-13 → Leu and Arg-41 → His. Eight residues are involved in the hydrogen-bonding network of the major the minor β-sheet of EGF, and each of these residues has amide protons that exchange slowly after dissolving the protein in D_2O. These protons give rise to cross-peaks, at their CαH chemical shift in COSY spectra, where they can be unambiguously assigned.

Figure 5 shows the first in a series of COSY spectra that contain most of the slowly exchanging residues for the wild-

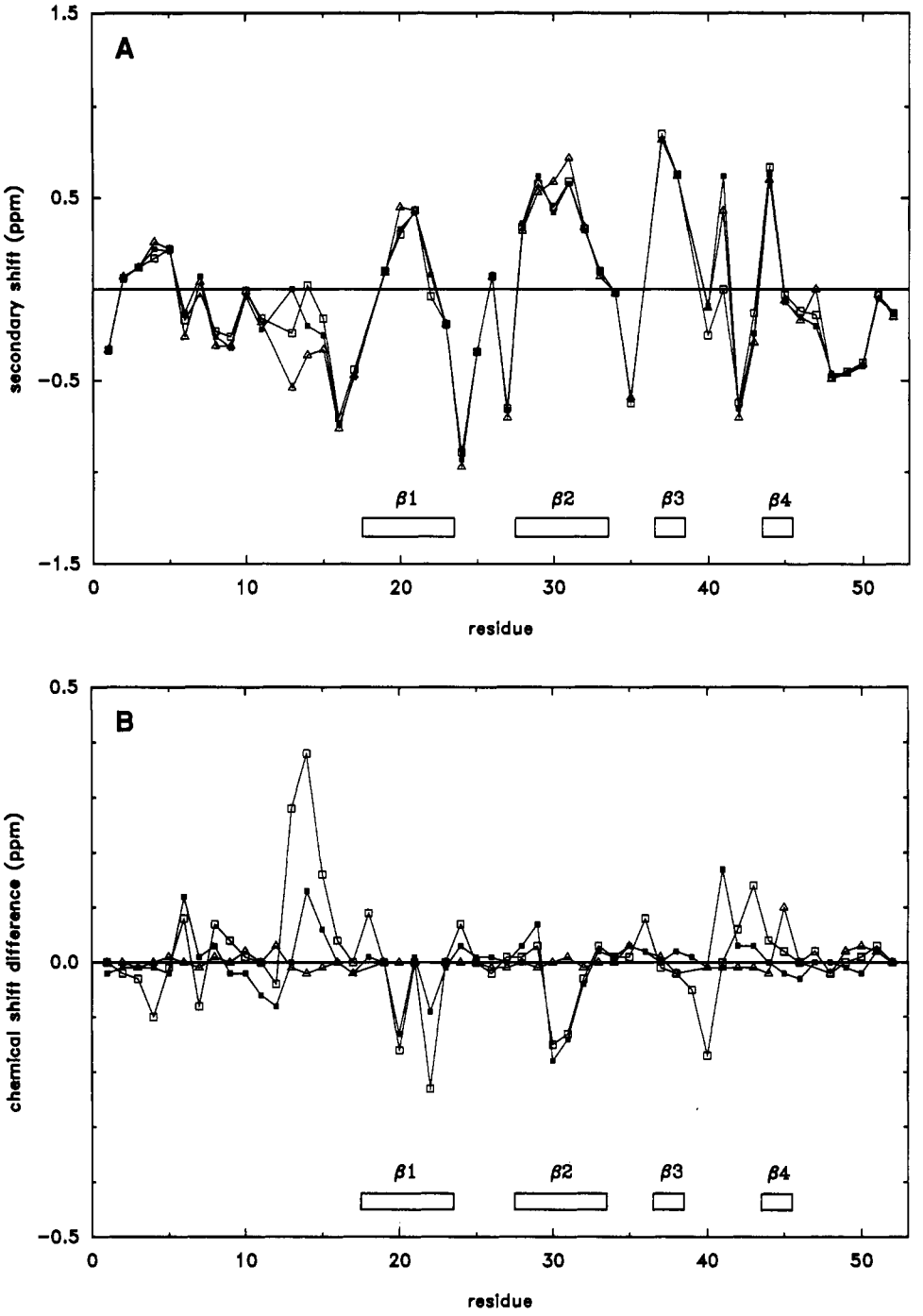

FIGURE 3: (A) Secondary shifts of the different hEGF variants. The chemical shifts for the Cα protons of (Δ) Leu-47 → Glu, (■) Tyr-13 → Leu, and (□) Arg-41 → His were corrected for their random coil values according to Bundi and Wüthrich (1979). The secondary shifts of the glycine residues at position 12, 18, 36, and 39 were not plotted. Positive values represent downfield-shifted resonances compared to the corresponding random coil value. Secondary structural elements are indicated by bars. Shown is the strong correlation between the secondary shifts and the occurrence of the major (β_1, β_2) and the minor (β_3, β_4) β-sheet position for all mutants. (B) Chemical shift differences of the different hEGF variants. The chemical shift difference for a given residue was calculated by subtracting the chemical shift found in the wild-type protein from the value of this residue in the variant hEGF. Major differences are observed around residues 13 and 41 as well as in the major β-sheet for both the Tyr-13 → Leu and Arg-41 → His variants. The differences found in the Leu-47 → Glu are small compared with those of the Tyr-13 → Leu and Arg-41 → His variants.

type and the Tyr-13 → Leu and Arg-41 → His variants. It is obvious from this figure that both variants show the same pattern of slowly exchanging amide residues as the wild-type protein. This observation confirms that the main secondary structural elements of EGF are preserved upon mutation; this is also indicated by long-range NOEs observed across the β-sheets (Figure 2A and 4). To obtain information about the solvent-exchange rates of particular residues, up to eight COSY spectra were acquired and the time dependence of the peak intensities was measured. The corresponding rate con-

Table II: Hydrogen-Exchange Rates[a] for Slowly Exchanging Amide Protons in hEGF and Variants

residue	wild type (10^{-6} s^{-1})	Tyr-13 → Leu (10^{-6} s^{-1})	Arg-41 → His (10^{-6} s^{-1})
Val-19	4 ± 0.5	18 ± 2	19 ± 1
Met-21	6 ± 0.9	27 ± 2	20 ± 2
Ile-23	4 ± 0.5	10 ± 1	11 ± 1
Val-34	13 ± 1	32 ± 7	27 ± 0.3
Ile-38	13 ± 1	38 ± 5	36 ± 0.6

[a] The experimental conditions were pH 2.9 and 30 °C.

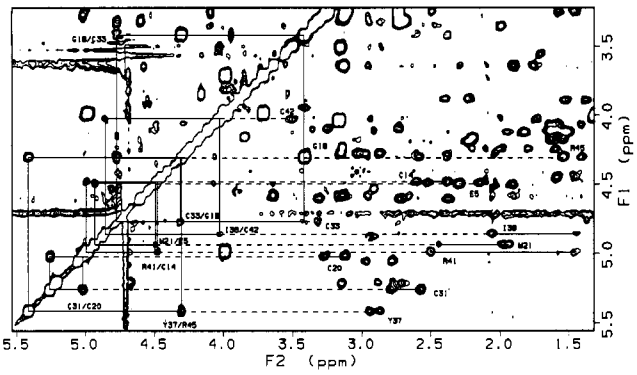

FIGURE 4: Long-range connectivities in hEGF. A part of a NOESY spectrum of the Tyr-13 → Leu variant acquired in D_2O at a mixing time of 200 ms is shown. Some spin systems of residues that give rise to characteristic long-range CαH–CαH NOEs are marked. These NOEs are occurring near the diagonal and are depicted by a box.

stants were obtained by linear regression of the data and are summarized in Table II.

Discussion

The relationship between the structure and function of EGF and its homologue TGFα has attracted wide interest (Carpenter & Cohen, 1990). Knowledge of the three-dimensional structures of human and mouse EGF, as well as of human TGFα, together with amino acid sequence comparisons led to a classification of two different kinds of important residues in EGF (Campbell et al., 1990). Beside residues that seem to play their major role in stabilizing the structural integrity of the protein (Cys-6, Cys-14, Cys-20, Cys-31, Cys-33, Cys-42, Gly-18, Gly-36, Gly-39, and Tyr-37) there are others that might be responsible for specific receptor recognition. With the exception of Leu-47, all these residues (Tyr-13, Leu-15, His-16, Arg-41, and Gln-43) are on one face of the molecule and thus are likely to be directly involved in a specific interaction between EGF and its receptor.

In the present report, the contribution of residues Arg-41 and Tyr-13 toward receptor binding of hEGF has been assessed through substitution by selected residues. To distinguish between local and global effects upon mutation, we have combined information gathered from biological assays with a complete assignment of all nonexchangeable protons of these variants. This allowed comparison of the structure of these proteins using three independent data sets: (i) chemical shifts of proton resonances, (ii) hydrogen-exchange rates for those protons involved in the hydrogen bond network of the β-sheets, and (iii) nuclear Overhauser enhancements characteristic of the EGF-folding topology.

The secondary structure of hEGF is schematically depicted in Figure 6. The dominant structural motif is an antiparallel β-sheet running from residue 18 to 23 and 28 to 33. A second minor antiparallel β-sheet is formed by residues 37–38 and 44–45. Residues located in the β-sheets of EGF have large downfield-shifted CαH resonances (Cooke et al., 1990). The same was found for all three variants, indicating that the formations of both the major and minor β-sheet have not been greatly affected by the mutations. This was further confirmed by the observation of long-range NOEs and slowly exchanging amide protons characteristic for the formation of the major and the minor antiparallel β-sheet in all three variants.

Complete analysis of the chemical shifts of the CαH protons revealed that only small changes were observed for the Leu-47 → Glu variant while substantial localized changes of some chemical shifts could be observed for both the Tyr-13 → Leu and the Arg-41 → His mutants. These changes affected residues Cys-20, Tyr-22, Ala-30, and Cys-31, which play a dominant role in building the structural scaffold of EGF (see also Figure 6). Interpretation of the changes in terms of structure are thus of considerable interest. However, care has to be taken since in both mutants an aromatic ring system has been either introduced or replaced. For instance, some of the observed changes seen in the Tyr-13 → Leu mutant probably reflect the contribution of the aromatic tyrosine ring in the wild-type protein. On the other hand, the introduction of a histidine ring in the Arg-41 → His variant is expected to have about half the ring current effect produced by tyrosines and phenylalanines (Perkins, 1982). It is therefore very unlikely that the replaced histidine is directly causing any secondary shifts as far away as, for instance, Tyr-22 of Cys-20 (15 and 12 Å, respectively).

As shown in Figure 3B, there is a striking similarity between the patterns of spectral changes observed for the Tyr-13 → Leu and Arg-41 → His mutants. This supports the idea of a common structural basis for the observed effects. Inspection of the native EGF structure reveals that Tyr-13 and Arg-41 form intimate contacts across the interface of EGF (Cooke et al., 1987; Hommel et al., unpublished results). Tyr-13 is also in contact with residues of the major β-sheet, and the spectral changes observed there are readily explained by the removal of tyrosyl side chain in the Tyr-13 → Leu variant.

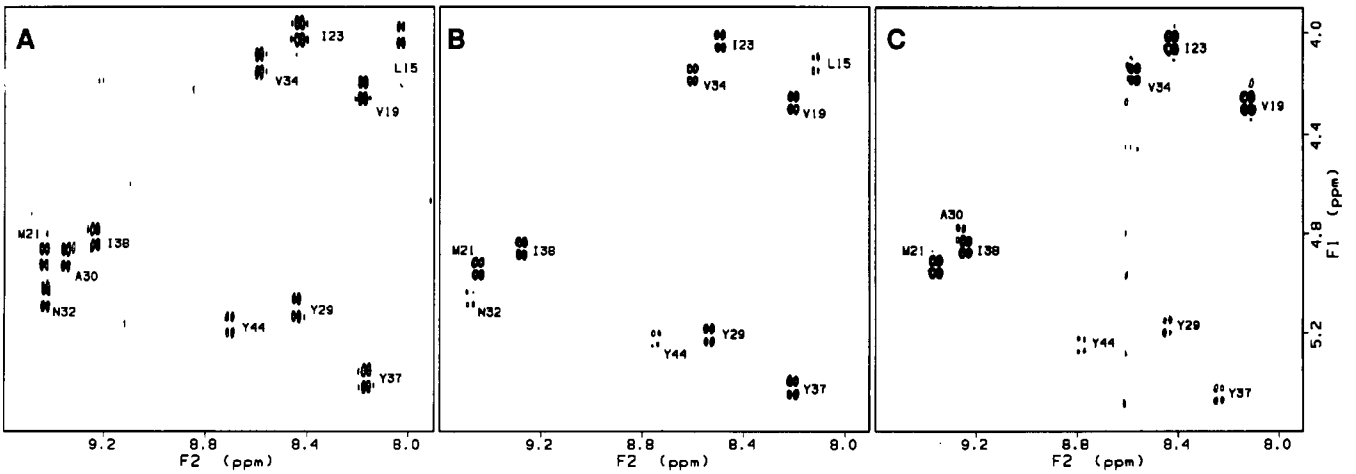

FIGURE 5: Exchange of hydrogen-bonded protons with solvent. The exchange rates of the slow exchanging amide protons of the major and the minor β-sheet with solvent deuterons were followed by COSY spectra. As an example, the first in a series of COSY spectra is given for (A) wild type, (B) Tyr-13 → Leu variant, and (C) Arg-41 → His variant.

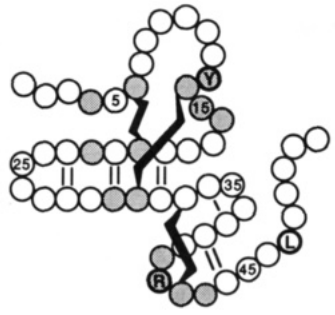

FIGURE 6: Schematic representation of hEGF. The secondary structure of hEGF is dominated by two antiparallel β-sheets. Hydrogen bonds contributing to these secondary structural elements are indicated by double lines. Cystine bridges 6–20, 14–31, and 33–42 are indicated by bars. The position of the amino acid substitutions (Tyr-13, Arg-41, and Leu-47) are highlighted. Residues that experience large chemical shift changes upon mutation of Tyr-13 and Arg-41 are shaded.

The observation of the same pattern of spectral changes in the Arg-41 → His variant suggest that subtle structural rearrangements of Tyr-13 must have taken place in this variant. This can be well understood considering the hydrophobic interactions made by the native residues at this interface. It is interesting to note in this context that substitutions of Arg-41 to the more flexible and less bulky residue glutamine showed far less dramatic spectral changes (data not shown). Experimental evidence for a disturbed interface structure in the Arg-41 → His variant also comes from the fact that the EGF-characteristic NOE across the interdomain interface between the Cα protons of Cys-14 and His-41 is not observed.

The scenario outlined above is further supported by the results from the hydrogen-exchange experiments. Here, a 4-fold increase of the respective exchange rates is observed, indicating a decrease in the stability of the structural scaffold of EGF. This means that the substitutions of Tyr-13 by leucine and of Arg-41 by histidine have the same impact not only on the chemical shifts of the same residues in the major β-sheet but also on the exchange rates of these residues. As has been shown earlier, there is a large hydrophobic patch on top of the major β-sheet comprising residues Tyr-13, Tyr-22, and Tyr-29 (Cooke et al., 1987; Montelione et al., 1987). As a result of efforts to improve the resolution of the wild-type structure, we have recently found that the hydroxyl group of Tyr-13 is near enough to the carbonyl group of Tyr-29 to form a hydrogen bond. Thus, the suggested relay of conformational change brought about by substitution of residue 41 represents a likely structural explanation for the observed effects.

In contrast to the effects of substitutions of Arg-41, the Tyr-13 → Leu mutant was found to have little effect on receptor affinity. This was unexpected on the basis of the conservation of an aromatic ring at this position in all proteins known to bind to the EGF receptor. Others have found that alterations of the residue corresponding to Tyr-13 in hTGFα and myxoma growth factor cause a reduction in their receptor affinity (Defeo-Jones et al., 1988; Lin et al., 1991). While we cannot exclude the possibility that the large decrease in Arg-41 → His receptor binding is partly due to the rearrangement of Tyr-13 in this mutant, the fact that other substitutions of Arg-41 do not cause large spectral changes argues against this. Similarly, nonconservative changes at Tyr-13 might influence the conformation of Arg-41 and thereby affect the receptor affinity. This could explain the reported variabilities in the results of mutagenesis data at this residue.

Another implication of the presented data is that substitutions of Arg-41 can have impact on residues as far away as the major β-sheet. Others are currently studying the mutational effects of substitution of residues in the major β-sheet of hEGF, and in several cases reduced receptor binding affinities have been reported (Engler et al., 1988; Campion et al., 1990). Since it is conceivable that the observed changes could work also in reverse, i.e., mutations in the main β-sheet could affect Arg-41, a structural characterization of any hEGF variant is desirable.

This study on the structure–function relationships of various mutants of hEGF by different ^{1}H NMR techniques has revealed some limitations of the approach. A wide range of structural parameters including NOEs, hydrogen-exchange rates, and chemical shifts were acquired to explore the structural integrity of the variant proteins. Although each of these individual parameters has evident limitations, they can yield complementary and persuasive evidence when combined.

It has been stressed by others that chemical shift information is sensitive to the most subtle changes in the chemical environment of residues in mutant proteins (Folkers et al., 1989; Gao et al., 1990). However, the relationship between chemical shifts and the tertiary structure of a protein are only poorly understood, and hence these shifts are generally difficult to interpret in structural terms (Redfield & Dobson, 1990). In our study we have presented evidence that, in a small protein like EGF, spectral changes produced by mutation can be interpreted as long as complementary mutations are available. This was demonstrated by the similar pattern of chemical shift differences in the major β-sheet of the Arg-41 → His and Tyr-13 → Leu variants and further supported by their similarly increased hydrogen-exchange rates.

In contrast to the sensitivity of chemical shift to protein conformation, qualitative interpretation of the acquired NOEs did not reveal significant structural changes between the different variants. This difference in information is not contradictory since the NOE between a pair of atoms in a protein is known to be only a loose measure of the interproton distance (Wüthrich, 1989; Clore & Gronenborn, 1989). The use of NOE-derived distance constraints has further limitations for a comparison of proteins. Factors such as protein expression levels, protein solubility, internal flexibility, and overall protein stability can vary in mutant proteins and therefore impair the NOE-based calculation of the protein conformation. In EGF, all of these limitations exist as it is neither very globular nor suited to high protein concentrations. Unsuccessful attempts to crystallize this molecule support the idea of a flexible protein, and we have evidence from NMR relaxation measurements that regions of the polypeptide are rather mobile (Hommel et al., unpublished results).

In summary, we have demonstrated that Arg-41 is a very important residue, involved both in receptor binding and in the structural integrity of EGF. As part of a highly conserved domain–domain interface, its contribution to the formation of this interface is of great interest. We are currently studying the contribution made by Arg-41 and other conserved residues in the wild-type molecule in more detail by calculating new EGF structures from a larger number of structural restraints. These studies should provide a deeper understanding of the effects observed in this report.

Acknowledgments

We are grateful to Drs. Robert Cooke and Paul Driscoll for expert experimental advice.

Supplementary Material Available

A table listing proton resonances for the Arg-41 → His, Tyr-13 → Leu, and Leu-47 → Glu variants of hEGF (2

pages). Ordering information is given on any current masthead page.

References

Alber, T., Bell, J. A., Sun, D.-P., Nicholson, H., Wozniak, J. A., Cook, S., & Matthews, B. W. (1988) *Science 239*, 631–635.

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

Bax, A., Sklenar, V., Clore, G. M., & Gronenborn, A. M. (1987) *J. Am. Chem. Soc. 109*, 6511–6513.

Brake, A. J., Merryweather, J. P., Coit, D. G., Heberlein, V. A., Masiarz, F. R., Mullenbach, G. T., Urdea, P., Valenzuela, P., & Barr, P. J. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 4642–4646.

Bundi, A., & Wüthrich, K. (1979) *Biopolymers 18*, 285–298.

Campbell, I. D., Dobson, C. M., & Williams, R. J. P. (1975) *Proc. R. Soc. London A 345*, 23–40.

Campbell, I. D., Baron, M., Cooke, R. M., Dudgeon, T. J., Fallon, A. F., Harvey, T. S., & Tappin, M. J. (1990) *Biochem. Pharmacol. 40*, 35–40.

Campion, S. C., Matsunami, R. K., Engler, D. A., & Niyogi, S. K. (1990) *Biochemistry 29*, 9988–9993.

Carpenter, G., & Cohen, S. (1990) *J. Biol. Chem. 265*, 7709–7712.

Carpenter, G., King, L., & Cohen, S. (1978) *Nature 276*, 409–410.

Clore, G. M., & Gronenborn, A. M. (1989) *Crit. Rev. Biochem. Mol. Biol. 24*, 479–564.

Cooke, R. M., Wilkinson, A. J., Baron, M., Pastore, A., Tappin, M. J., Campbell, I. D., & Sheard, B. (1987) *Nature 327*, 339–341.

Cooke, R. M., Tappin, M. J., Campbell, I. D., Khoda, D., Miyake, T., Fuwa, T., Miyazawa, T., & Ingaki, F. (1990) *Eur. J. Biochem. 193*, 807–817.

Davis, D. G., & Bax, A. (1985) *J. Am. Chem. Soc. 107*, 2821–2822.

Defeo-Jones, D., Tai, J. Y., Wegrzyn, R. J., Vuocolo, G. A., Baker, A. E., Payne, L. S., Garsky, V. M., Oliff, A., & Reimen, M. W. (1988) *Mol. Cell. Biol. 8*, 2999–3007.

Defeo-Jones, D., Tai, J. Y., Vuocolo, G. A., Wegrzyn, R. J., Schofield, T. L., Riemen, M. W., & Oliff, A. (1989) *Mol. Cell. Biol. 9*, 4083–4086.

DeLarco, J. E., & Todaro, G. J. (1980) *J. Cell. Physiol. 102*, 267–277.

Derynck, R., Roberts, A. B., Winkler, M. E., Chen, E. Y., & Goeddel, D. V. (1984) *Cell 38*, 287–297.

Dudgeon, T. J., Baron, M., Cooke, R. M., Fallon, A., & Campbell, I. D. (1990) *FEBS Lett. 261*, 392–396.

Engler, D. A., Matsunami, R. K. M., Campion, S. R., Stringer, C. D., Stevens, A., & Niyogi, S. K. (1988) *J. Biol. Chem. 263*, 12384–12390.

Engler, D. A., Montelione, G. T., & Niyogi, S. K. (1990) *FEBS Lett. 271*, 47–50.

Folkers, P. J. M., Clore, G. M., Driscoll, P. C., Dodt, J., Köhler, S., & Gronenborn, A. M. (1989) *Biochemistry 28*, 2601–2617.

Gao, Y., Boyd, J., Williams, R. J. P., & Pielak, G. J. (1990) *Biochemistry 29*, 6994–7003.

George-Nascimento, C., Gynes, A., Halloran, M., Merryweather, J., Vanezuela, P., Steimer, K. S., Masiarz, F. R., & Randolph, A. (1988) *Biochemistry 27*, 797–802.

Harvey, T. S., Wilkinson, A. J., Tappin, M. J., Cooke, R. M., & Campbell, I. D. (1991) *Eur. J. Biochem. 198*, 555–562.

Higashiyama, S., Abraham, J. A., Miller, J., Fiddes, J. C., & Klagsbrun, M. (1991) *Science 251*, 936–939.

Jeener, J., Meier, B. H., Bachmann, P., & Ernst, R. R. (1976) *J. Chem. Phys. 71*, 4546–4533.

Kohda, D., Shimada, I., Miyake, T., Fuwa, T., & Inagaki, F. (1989) *Biochemistry 28*, 953–958.

Kline, T. P., Brown, F. K., Brown, S. C., Jeffs, P. W., Kopple, K. D., & Mueller, L. (1990) *Biochemistry 29*, 7805–7813.

Kunkel, T. A., Roberts, J. D., & Zakour, R. A. (1987) *Methods Enzymol. 154*, 367–382.

Knowles, J. R. (1987) *Science 236*, 1252–1258.

Lin, Y., Caporaso, G., Chang, P., Ke, X., & Tam, J. P. (1988) *Biochemistry 27*, 5640–5645.

Lin, Y., Ke, X., & Tam, J. P. (1991) *Biochemistry 30*, 3310–3314.

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

Mayo, K. H., Cavalli, R. C., Peters, A. R., Boelens, R., & Kaptein, R. (1989) *Biochem. J. 257*, 197–205.

Montelione, G. T., Wüthrich, K., Nice, E. C., Burgess, A. W., & Sheraga, H. A. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 5226–5230.

Moy, F.-J., Sheraga, H. A., Liu, J.-F., Wu, R., & Montelione, G. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 9836–9840.

Panayotou, G., End, P., Aumailley, M., Timpl, R., & Engel, J. (1989) *Cell 56*, 93–101.

Pardi, A., Wagner, G., & Wüthrich, K. (1983) *Eur. J. Biochem. 137*, 445–454.

Perkins, S. J. (1982) *Biol. Magn. Reson. 4*, 193–336.

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

Redfield, A. G., & Kunz, S. D. (1975) *J. Magn. Reson. 19*, 250–254.

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

Shoyab, M., Plowman, G. D., McDonald, V. L., Bradley, J. G., & Todaro, G. J. (1989) *Science 243*, 1074–1076.

States, D. J., Haberkorn, R. A., & Ruben, D. J. (1982) *J. Magn. Reson. 48*, 286–292.

Twardzik, D. R., Brown, J. P., Ranchalis, J. E., Todaro, G. J., & Moss, B. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 5300–5304.

Ullrich, A., & Schlessinger, J. (1990) *Cell 61*, 203–212.

Upton, C., Macen, J. L., & McFadden, G. (1987) *J. Virol. 61*, 1271–1275.

Ushiro, H., & Cohen, S. (1980) *J. Biol. Chem. 255*, 8363–8365.

Wagner, G., & Wüthrich, K. (1986) *Methods Enzymol. 131*, 307–326.

Wagner, G., Stassinopoulou, C. I., & Wüthrich, K. (1984) *Eur. J. Biochem. 145*, 431–436.

Wilde, J. A., Bolton, P. H., Dell'Acqua, M., Hibler, D. W., Pourmotabbed, T., & Gerlt, J. A. (1988) *Biochemistry 27*, 4127–4132.

Wüthrich, K. (1989) *Science 243*, 45–50.

Yarden, Y., & Schlessinger, J. (1987) *Biochemistry 26*, 1443–1451.