Biochem. J. (1986) 239, 13–18 (Printed in Great Britain)

Structural characterization and exposure of aromatic residues in epidermal growth factor from the rat

Kevin H. MAYO,*‖ Paul SCHAUDIES,† C. Richard SAVAGE,† Antonio DE MARCO,‡ and Robert KAPTEIN§
*Department of Chemistry and the †Department of Biochemistry (Medical School), Temple University, Philadelphia, PA 19122, U.S.A., ‡Istituto di Chimica delle Macromolecole, CNR, Via E. Bassini 15/A, 20133 Milano, Italy, and §Physical Chemistry Laboratory, University of Groningen, Nijenborgh 16, 9747 AG Groningen, The Netherlands

Aromatic amino acid residues in epidermal growth factor (EGF) isolated from the rat have been investigated by proton n.m.r. and nuclear Overhauser methods at 500 MHz and by photochemically induced dynamic nuclear polarization (photo-c.i.d.n.p.) experiments at 360 MHz. Rat EGF contains six aromatic residues, i.e. one histidine and five tyrosine residues. pH titration data allow identification of the histidine imidazole ring protons, whereas two-dimensional n.m.r. correlated spectroscopy establishes connectivities between tyrosine ring (2,6) and (3,5) proton resonances. Photo-c.i.d.n.p. data give evidence for solvent exposure of the one histidine and the five tyrosine residues in rat EGF. Nuclear Overhauser experiments and pH titration data suggest proximity relationships among four of the tyrosine residues and the histidine residue. These data indicate the presence of a clustered, aromatic, structural domain on the protein surface and may provide a clue to the understanding of the functional structure of EGF.

INTRODUCTION

Recently, epidermal growth factor (EGF) from the rat has been isolated (Schaudies & Savage, 1986). On a cellular level, rat EGF works in a similar fashion to mouse EGF and human EGF (urogastrone) by binding to a transmembrane receptor molecule and becoming internalized in the cell by receptor-mediated endocytosis (Carpenter & Cohen, 1976, 1979; McKanna *et al.*, 1979; Schaudies & Savage, 1986), ultimately leading to increased protein, RNA and DNA synthesis and cell division.

Although the amino acid sequence is unknown, rat EGF contains one histidine and five tyrosine residues and is similar in composition to mouse EGF. Moreover, rat EGF competes with mouse EGF for binding to the same cellular receptor and elicits the same cellular responses as does mouse EGF; these facts suggest that mouse and rat EGF have similar structural receptor-binding regions. The poor reaction of rat EGF with anti-(mouse EGF) antibodies and vice versa (Schaudies & Savage, 1986), however, suggests that differences in their structures do exist.

Tryptophan, for example, which has been suggested to be involved in receptor binding by mouse and human EGF, is lacking in rat EGF. Mouse EGF and insulin both have surface-exposed regions with a high incidence of hydrophobic aliphatic and aromatic residues. At least with insulin, a region of clustered aromatic residues has been proposed to be critical in the receptor-binding process (Blundell *et al.*, 1971); perhaps this extrapolation could be made to EGF, which also functions as a protein hormone in a similar manner. This present study characterizes the aromatic residues in rat EGF by using n.m.r. and photo-c.i.d.n.p. spectroscopic techniques.

MATERIALS AND METHODS

Rat EGF was purified from rat salivary-gland extracts by using the method of Schaudies & Savage (1986). Purity was checked by amino acid analysis, gel electrophoresis and n.m.r.

Samples for ^1H n.m.r. measurements had been freeze-dried from ^2H_2O solutions and redissolved in ^2H_2O immediately before the experiment. The final protein concentration was about 0.5 mM in 10 mM-potassium phosphate buffer. The p^{2}H was adjusted by adding μl increments of NaO^{2}H or $^2H_3PO_4$ to a 0.5 ml sample. All quoted p^{2}H values are pH-meter readings uncorrected for isotope effects (Kalinichenko, 1976; Bundi & Wüthrich, 1979*a*,*b*).

^1H-n.m.r. spectra were recorded in the Fourier mode on a Bruker WM-500 spectrometer at 303 K. The solvent ^2H signal was used for the field-frequency lock. All chemical shifts are quoted in p.p.m. downfield from DSS.

N.O.e.s were generated by irradiating the desired peak for 0.6 s at a power level sufficient to null *z*-magnetization of the irradiated peak in 0.05 s. A 2 ms delay was introduced before accumulation to reduce transient effects, and a time of 3.5 *s* was allowed between scans to allow for recovery of *z*-magnetization. Difference spectra were obtained by subtracting the f.i.d. of the irradiated peak spectrum from the f.i.d. of a control spectrum that had been irradiated under the same experimental conditions in a region where no resonances occur.

The basic pulse sequence for the two-dimensional homonuclear *J*-correlated (COSY) experiment is (J. Jeener, unpublished work; Aue *et al.*, 1976):

$$(P_1\text{–}t_1\text{–}P_2\text{–}t_2\text{–}\tau)_n$$

where P_1 and P_2 indicate 90° pulses; t_1 is the evolution

Abbreviations used: EGF, epidermal growth factor; DSS, sodium 4,4-dimethyl-4-silapentanesulphonate; n.O.e., nuclear Overhauser enhancement; f.i.d., free induction decay; 2D-n.m.r. COSY, two-dimensional n.m.r. correlated spectroscopy; c.i.d.n.p., chemically induced dynamic nuclear polarization.

‖ To whom correspondence and reprint requests should be sent.

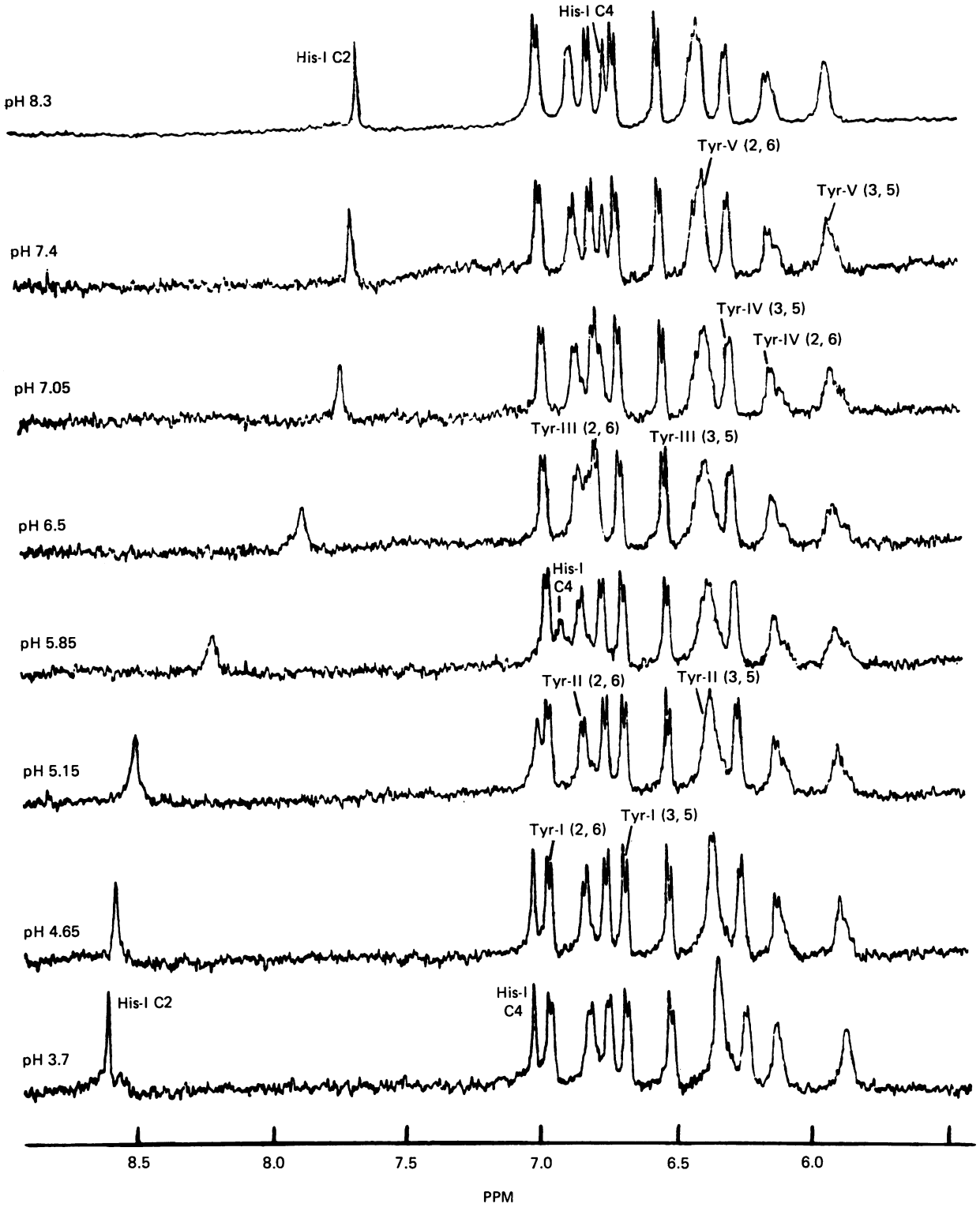

Fig. 1. 500 MHz proton-n.m.r. spectra of the aromatic region in rat EGF

Spectra are shown as a function of pH as indicated in the Figure. These spectra represent the time-averaged accumulation of 256 to 512 scans taken over approx. 12–24 min respectively by using a 90° pulse on a 0.4 ml sample at concentrations of 0.5 mM rat EGF in 10 mM-KH_2PO_4. Assignments shown are discussed in the text.

period between the two pulses during which the various magnetization components are labelled with their characteristic frequencies; t_2 is the time during which the signal is recorded, and τ is the delay between successive acquisitions to allow for spin-lattice relaxation. In this particular experiment, the mixing pulse, P_2, was 45°. Axial peaks were suppressed by a 16-step phase-cycling routine; quadrature detection was used in both dimensions, and the carrier frequency was placed at the centre of the spectrum. Data processing was done on a VAX 11785 computer system with software supplied by Dr. Dennis Hare, Columbia University, New York, NY, U.S.A. Other details are given in the legend to Fig. 2.

The samples for the photo-c.i.d.n.p. experiments contained about 0.2 mM-protein in ^2H_2O; the flavin-dye concentration was 0.4 mM. Photo-c.i.d.n.p. n.m.r. spectra were recorded on a Bruker HX-360 spectrometer, controlled by an Aspect 2000 computer. A Spectra Physics model-171 argon ion laser was employed as the light source. A 0.5 s light pulse (4 W, multi-line) was used, with a 0.05 s delay before the 90° radiofrequency pulse. The 'neat' photo-c.i.d.n.p. response was induced by irradiating the sample with laser light after two saturation pulse trains of 2 s each. Four 'light' scans were accumulated for each spectrum. Although the photo-c.i.d.n.p. enhancement was quite large, it was necessary nevertheless to subtract four 'dark' transients normalized for each 'light' scan in order to remove residual background magnetization. 3-*N*-Carboxymethyllumiflavin was generously given by Dr. F. Muller, University of Wageningen, Wageningen, The Netherlands.

RESULTS

500 MHz n.m.r. spectra of aromatic proton resonances in rat EGF are shown in Fig. 1 as a function of pH. These resonances must account for the ring protons of the one

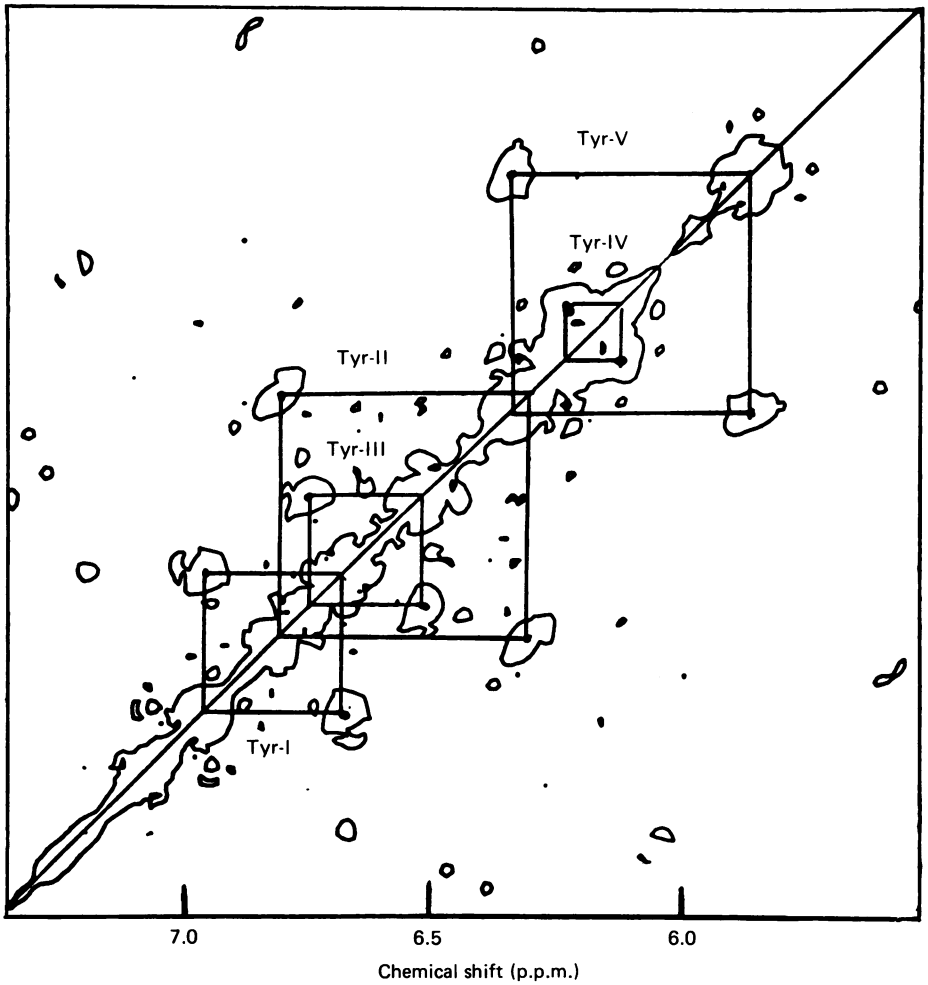

Fig. 2. 500 MHz COSY spectrum (contour plot) of the aromatic proton region in rat EGF

The spectral width was ±2500 Hz. The data set consisted of 2048 points in the t_2 dimension and 512 points in the t_1 dimension; 144 f.i.d.s were accumulated for each value of t_1, with a 1 s delay between acquisitions. A 90° pulse (12 μs) was used for the first pulse, but the second pulse was 6 μs (45°). The resulting data matrix was processed with a phase-shifted sine-bell in both dimensions and was zero-filled to 1024 points in the f_1 dimension. The absolute value mode is used and the data matrix was symmetrized. Sample conditions are as described in Fig. 1 at pH 3.5. Lines connect spin-coupled resonances through cross-peaks. The labels above the connecting lines refer to resonances in Fig. 1 and discussed in the text. The rat EGF concentration in this experiment was approx. 0.5 mM.

histidine and five tyrosine residues in the protein. The His-I C(2) and C(4) proton resonances are singlets of intensity one proton each and are readily identifiable in Fig. 1 by their pH behaviour. The pH-dependent broadening of histidine resonances near the histidine pK_a is probably the result of chemical exchange between acidic and basic forms of the histidine (Patel *et al.*, 1972), where the exchange between protonated and unprotonated forms is no longer fast enough to fulfill the fast exchange condition at 500 MHz (Campbell *et al.*, 1974).

The remaining aromatic proton resonances must be associated with tyrosine ring protons. Most of these resonances have the typical AA′-BB′ splitting pattern of (2,6) and (3,5) proton resonances commonly observed for tyrosine ring protons. The two most upfield resonances seem to have minor signals associated with them; this phenomenon is similar to that observed in mouse EGF (Mayo, 1984, 1985). The reason for this is not yet well understood.

In order to group the tyrosine ring-proton resonances into complete spin systems, 2D-n.m.r. COSY was used (J. Jeener, unpublished work; Aue *et al.*, 1976). Fig. 2 shows a contour plot of such an experiment and gives the (2,6) and (3,5) proton spin–spin connectivities as labelled in Fig. 2 and 1. Assignment of the (2,6) and (3,5) proton resonances in a particular spin system was made on the basis of photo-c.i.d.n.p. results given below.

Photo-c.i.d.n.p. difference spectra are shown in Fig. 3 at two pH values along with a reference spectrum at the same spectrometer frequency (360 MHz). The tyrosine (3,5) proton resonances give large negative (emissive) signals, whereas the intensities of the (2,6) proton resonances depend upon the cross-relaxation rate of a particular tyrosine residue under these experimental conditions (Hore *et al.*, 1982; Zetta *et al.*, 1982). Only a small positive effect, if any, is observed for some of the (2,6) proton resonances. The (3,5) proton resonances can, therefore, be identified in those tyrosine spin systems which demonstrate a photo-c.i.d.n.p. effect. At pH 7.9, all the aromatic residues in rat EGF exhibit a photo-c.i.d.n.p. effect. Tyrosine residues I and III demonstrate the largest effect, whereas tyrosine II exhibits the smallest one. A photo-c.i.d.n.p. effect on His-I is also clearly evident as the normally observed

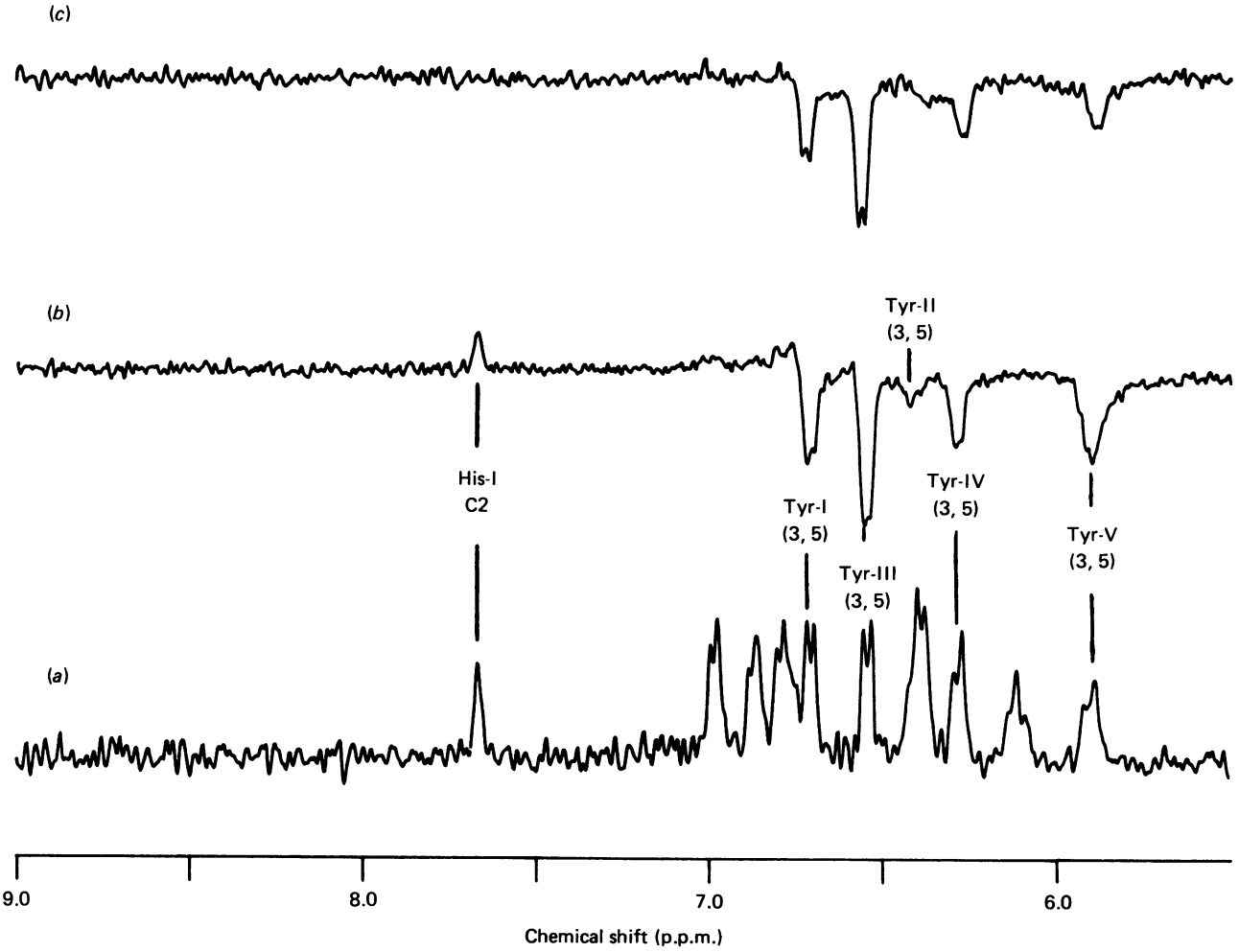

Fig. 3. 360 MHz Photo-c.i.d.n.p. Spectra of rat EGF

(*a*) 360 MHz proton-n.m.r. reference spectrum; conditions were the same as in Fig. 1 at pH 7.9. Shown are photo-c.i.d.n.p. difference spectra at pH 7.9 (*b*) and at pH 3.5 (*c*) Flavin dye (0.4 mM) was added to the sample for the photo-c.i.d.n.p. experiment.

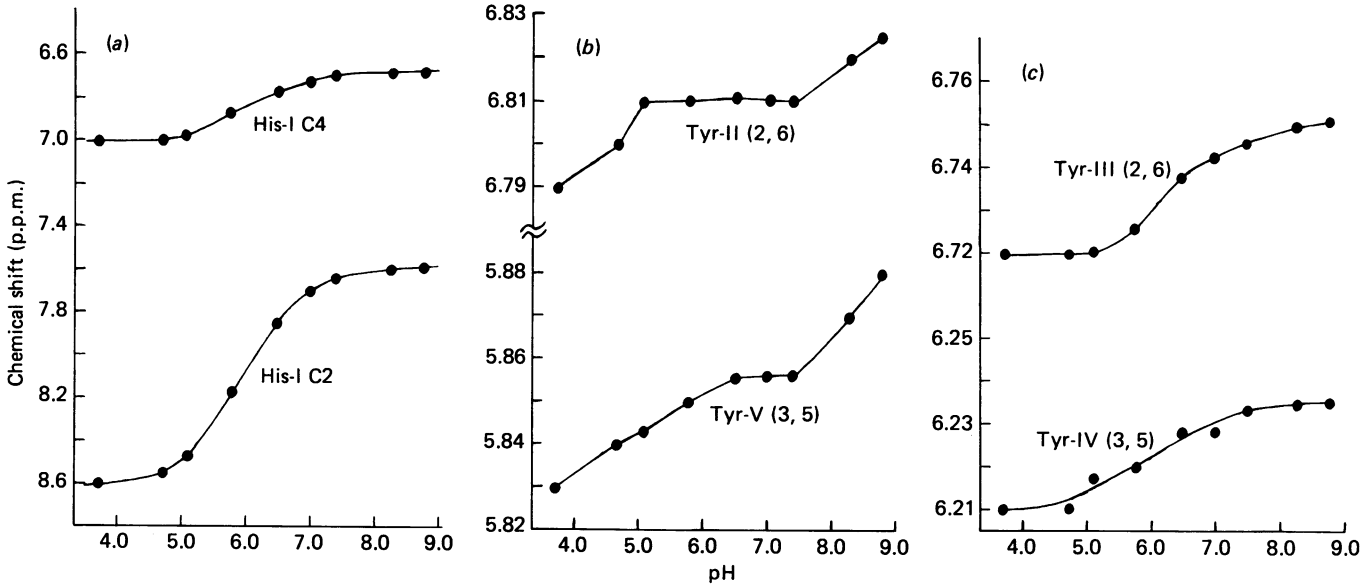

Fig. 4. Chemical shift versus pH for aromatic resonances in rat EGF

Shown are a series of plots of aromatic proton resonance chemical shifts as a function of pH for His-I (*a*), for Tyr-II (2,6) and Tyr-V (3,5) (*b*) and for Tyr-III (2,6) and Tyr-IV (3,5) (*c*). With the exception of His-I data, lines are drawn between points only as a visual aid.

positive (absorptive) signal. Lowering the pH to 3.5 reduces the photo-c.i.d.n.p. effect only slightly on Tyr-IV and -V, whereas no polarization of His-I is observed at low pH. Protonation of the histidine imidazole ring generally reduces the photo-c.i.d.n.p. effect (Kaptein, 1978), so the histidine is probably unobserved for this reason, since the pH is well below its pK_a. These results suggest that all aromatic residues in rat EGF are, to various extents, exposed to the flavin dye and, therefore, most likely accessible to the solvent. The diminution in the Tyr-IV and -V signals also suggests that the protein conformation is slightly more compact at low pH.

The chemical shifts of some of the aromatic proton resonances are plotted as a function of pH in Fig. 4. By using the Henderson–Hasselbach equation (Edsall & Wyman, 1958), a pK_a of 6.0 can be estimated for His-I. This pK_a is expected for a non-associated, solvent-exposed, histidine residue (Lehninger, 1978). Aside from the His-I ring protons, several tyrosine proton resonances are also chemically shifted by varying the pH. Valuable structural information on the proximity of amino acid side chains in a protein can be gained from such pH titration data. For example, the Tyr-III (2,6) and Tyr-IV (3,5) proton resonance pH behaviours reflect that of His-I, whereas the Tyr-II (2,6) proton resonance has inflexion points near the expected pK_as of a glutamic acid or aspartic acid residue, or even the *C*-terminus and probably of a tyrosine residue or the *N*-terminus in rat EGF. The Tyr-V (3,5) proton resonance pH titration, moreover, reflects the pK_a of His-I as well as those observed on Tyr-II. These data would suggest that Tyrs II, III, IV and V are proximal to each other as well as to His-I.

Proximity relationships among tyrosine residues were confirmed by n.O.e. experiments. N.O.e. difference spectra at pH 6.9 and 303 K obtained by subtracting spectra with radiofrequency irradiation on and off selective tyrosine proton resonances for 0.6 s are presented in Figs. 5(*a*)–(*d*). Irradiation of any tyrosine (3,5) proton resonance shows a strong n.O.e. at its respective (2,6) proton resonance position and vice versa. This n.O.e. is expected, since the 2–3 and 5–6 proton pairs are only 0.25 nm (2.5Å) distant from one another. Indirect saturation accounted for some of the apparent peaks in the n.O.e. difference spectra (indicated by an 'X'). Control experiments for indirect saturation were performed by irradiating at frequencies near resonances to observe the magnitude of decoupler spill-over. The remaining labelled resonances in Fig. 5 are the result of real n.O.e.s. Proximity relationships, therefore, do exist among Tyr-II, Tyr-III, Tyr-IV and Tyr-V.

DISCUSSION

In general, proteins fold in a polar medium so as to mask most of their hydrophobic residues from the polar solvent in order to reduce the heat of solvation. In the case at hand, rat EGF is folded in such a way that all of its hydrophobic aromatic residues are solvent-exposed. A protein as small as rat EGF (i.e. 5.2 kDa) would not be expected to have all its hydrophobic residues internalized in the protein core; however, it is rather unusual for a protein to have all its aromatic residues solvent-exposed. In the case of mouse EGF, two of its eight aromatic residues are internalized in the protein (De Marco *et al.*, 1983).

Most of these solvent-exposed aromatic residues are also suggested by the present data to be clustered together on the protein surface. Proximity among Tyrs II, III, IV and V is suggested by n.O.e. data, whereas pH titration data support this and suggest that Tyr-II and Tyr-V are proximal to a glutamic acid or an aspartic acid residue or even the *C*-terminus. Tyr-III, -IV and -V are also proximal to His-I. Assuming the interpretation of

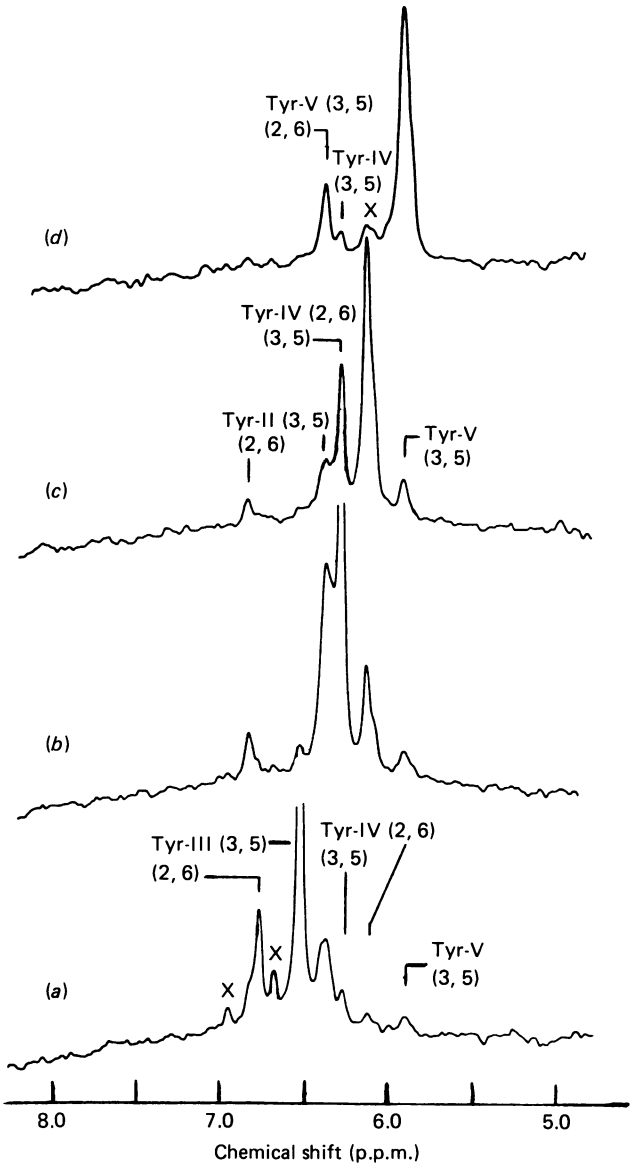

Fig. 5. N.O.e.s. given by aromatic proton resonances in rat EGF

Four n.O.e. difference spectra are shown, each representing the difference between an accumulation of 600 f.i.d.s taken with saturating radiofrequency on-resonance. In each case the radiofrequency irradiation time was 0.6 s. Spectra are for saturation of the Tyr-III (3,5) (*a*); for Tyr-IV (3,5) (*b*); for Tyr-IV (2,6) (*c*), and for Tyr-V (3,5) (*d*) proton resonances. Labelling of resonances is as in Fig. 1 and as described in the text.

these data to be correct, this would mean that on the surface of rat EGF there exists an area in which one histidine and four tyrosine residues are located.

On the basis of the (2,6) and (3,5) tyrosine ring-proton resonance chemical shifts, Tyr-I has resonances near the random-coil positions predicted for tyrosine residues in a fully extended conformation (Bundi & Wüthrich, 1979*a*,*b*). Tyr-IV and Tyr-V are in a highly anisotropic environment, whereas Tyr-II and Tyr-III appear to be in an intermediate state. This anisotropic environment may well be due to the presence of ring-current shifts caused by the proximity of several aromatic residues. Although Tyr-I is suggested by photo-c.i.d.n.p. data to be solvent-exposed, it does not seem to be proximal to these clustered aromatic residues and, therefore, probably exists elsewhere on the protein surface.

Even though the structural characteristic of solvent-exposed clustered aromatic residues seems rather uncommon in proteins, similar structural domains do exist in insulin (Blundell *et al.*, 1971, 1972; Blundell, 1975) and mouse EGF (De Marco *et al.*, 1983; Mayo, 1984). EGF and insulin are both protein hormones whose physiological mechanisms are, in many ways, similar. Both bind a transmembrane cellular receptor triggering a cascade of cellular responses, some similar and some quite different. Since this type of structural domain seems to be observed in at least these protein hormones, it may well be of general significance for the functionality of this class of proteins.

This work was supported by a grant from the National Institutes of Health (U.S.A.) (GM-34662) (to K.H.M.) and from a donation generously given by the Glenmede Trust Fund to Temple University and benefitted from n.m.r. facilities made available to Yale University through Grant CHE-79161210 from the National Science Foundation. This work was also supported by a grant from the North Atlantic Treaty Organization (NATO) Scientific Affairs Division (RG.85/0424) (to K.H.M. and A.D.)

REFERENCES

Aue, W. P., Bartholdi, E. & Ernst, R. R. (1976) J. Chem. Phys. **64**, 2229–2235
Blundell, T. L. (1975) New Sci. **18**, 662–664
Blundell, T. L., Cutfield, J. F., Cutfield, S. M., Dodson, E. J., Dodson, G. G., Hodgkin, D. C., Mercola, D. A., & Vijayan, M. (1971) Nature (London) **231**, 506–511
Blundell, T. L., Dodson, G. G., Hodgkin, D. C. & Mercola, D. A. (1972) Adv. Protein Chem. **26**, 279–402
Bundi, A. & Wüthrich, K. (1979*a*) Biopolymers **18**, 279–285
Bundi, A. & Wüthrich, K. (1979*b*) Biopolymers **18**, 285–297
Campbell, I. D., Lindskog, S., & White, A. I. (1974) J. Mol. Biol. **90**, 469–489
Carpenter, G. & Cohen, S. (1976) J. Cell Biol. **71**, 159–171
Carpenter, G. & Cohen, S. (1979) Biochem. Actions Horm. **5**, 121–156
De Marco, A., Menegatti, E. & Guarneri, M. (1983) FEBS Lett. **159**, 201–206
Edsall, J. T. & Wyman, J. (1958) Biophys. Chem. **1**, 406–476
Hore, P. J., Egmond, M. R., Edzes, H. T. & Kaptein, R. (1982) J. Magn. Reson. **49**, 122–150
Kalinichenko, P. (1976) Stud. Biophys. **58**, 235–240
Kaptein, R. (1978) in NMR Spectroscopy in Molecular Biology (Pullman, B., ed.), pp. 211–229, Reidel Co., Amsterdam
Lehninger, A. L. (1978) Biochemistry, p. 74, Worth Publisher, New York
McKanna, J. A., Haigler, H. T. & Cohen, S. (1979) Proc. Natl. Acad. Sci. U.S.A. **76**, 5689–5693
Mayo, K. H. (1984) Biochemistry **23**, 3960–3973
Mayo, K. H. (1985) Biochemistry **24**, 3783–3794
Patel, D. J., Woodward, C. K. & Bovey, F. A. (1972) Proc. Natl. Acad. Sci. U.S.A. **69**, 599–602
Schaudies, P. & Savage, C. R. (1986) Comp. Biochem. Biophys., in the press
Zetta, L., Kaptein, R. & Hore, P. J. (1982) FEBS Lett. **145**, 277–280

Received 21 April 1986/11 June 1986; accepted 1 July 1986