Volume 261, number 2, 392–396 FEBS 08176 February 1990

Structure-function analysis of epidermal growth factor: site directed mutagenesis and nuclear magnetic resonance

T.J. Dudgeon[1], R.M. Cooke[1], M. Baron[1], I.D. Campbell[1], R.M. Edwards[2] and A. Fallon[2]

[1]*Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU* and [2]*British Biotechnology Ltd, Watlington Road, Oxford OX4 5LY, UK*

Received 18 December 1989

The role of leucine-47 in determining the structure and activity of human epidermal growth factor was examined using site-directed mutagenesis. Wild type protein and four variants in which Leu^{47} was replaced by valine, glutamate, aspartate and alanine were produced from yeast. ^1H NMR experiments demonstrated that substitution of Leu^{47} had little effect on the protein structure. The observed reduction in receptor binding affinity caused by the substitutions could thus be attributed to perturbation of a residue directly involved in receptor interactions.

Epidermal growth factor, human; Protein NMR; Site directed mutagenesis

1. INTRODUCTION

The 53 amino acid protein epidermal growth factor (EGF) is a powerful mitogen for certain cell lines [1]. It elicits its response by binding to the EGF receptor, an integral membrane protein with intrinsic tyrosine kinase activity [2]. Other proteins which are known to bind to and activate the EGF receptor include transforming growth factor alpha (TGFα) [3] and several poxvirus growth factors [4–6].

Although there is considerable variation in the sequences of EGFs and TGFαs, their three-dimensional structures are quite similar [7], and they display similar affinities for the EGF receptor [8,9]. Consideration of the sequences and structures of EGF and TGFα enabled us to postulate that tyrosine 13, leucine 15, histidine 16, arginine 41, glutamine 43 and leucine 47 might be residues important for interactions with the receptor [7]. This led to a site directed mutagenesis project in our laboratories to test this hypothesis. The aim was to combine receptor binding studies and high resolution NMR so that a distinction could be made between a mutation causing minor local changes in structure and one causing widespread structural disruption.

The importance of leucine 47 (leucine 48 in TGFα) for receptor binding has been demonstrated by several groups [10–12]. In no case, however, has it been shown that the overall structure of the protein was unaffected by the mutation. The present study investigates in more detail the role of leucine 47 in determining the structure and receptor binding affinity of human EGF (hEGF). The wild type form of hEGF(1–52) and four variants in which leucine 47 is replaced have been produced from yeast using the alpha factor secretion pathway [13]. After purification the receptor binding affinity of each was determined and their structural integrity assessed by high-resolution nuclear magnetic resonance (NMR). The results suggest that reductions in activity associated with substitutions at leucine 47 are not caused by widespread structural changes and thus involve a residue required for direct interaction with the receptor.

Correspondence address: I.D. Campbell, Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, UK

2. MATERIALS AND METHODS

2.1. E. coli *and yeast strains*

E. coli strains MC1061, JM103 and RZ1032 were used for plasmid manipulations and site directed mutagenesis. *Saccharomyces cerevisiae* strain MD50a/α leu2-3/leu2-3 leu2-112/leu2-112 pep4-3/+ his3-11/his3-11 his3-15/his3-15 was used for protein expression.

2.2. *Plasmids*

The plasmid pSP46-αF/EGF contains as a *Bgl*II-*Bam*HI restriction fragment a synthetic hEGF gene joined to the yeast α factor leader, so that the codon for the first amino acid of EGF immediately follows the codons for the lysine-arginine cleavage site of the yeast KEX2 protease. The pLF1-αF/EGF expression vector (fig.1) was created by ligating this *Bgl*II-*Bam*HI fragment into the *Bgl*II expression site of pLF1, a derivative of pMA91 [14] in which the ampicillin resistance gene of pBR322 is substituted for the gene from pSP46 which lacks the *Pst*I site.

2.3. *Mutagenesis*

Site directed mutagenesis was performed on a M13mp19 clone containing the *Pst*I-*Bam*HI fragment of pSP46-αF/EGF and a single 20-mer primer with a degenerate mismatch at the position corresponding to leucine 47. The degeneracy in the sequence allowed the

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/90/$3.50

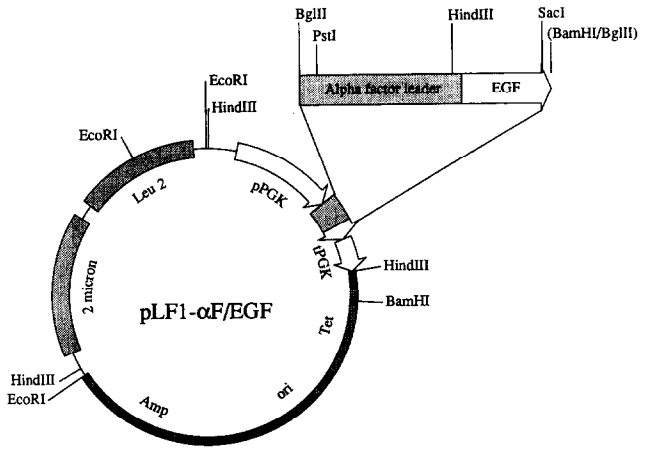

Fig.1. (A) A reversed phase HPLC trace of purified hEGF(1–52) using a Vydac (4.6 × 150 mm) C18 wide pore column. Buffer A = 0.1% trifluoroacetic acid in H_2O. Buffer B = 0.1% trifluoroacetic acid in acetonitrile. The gradient ran from 20% B to 50% B in 25 min, at 1 ml/min, after being held at 20% for 5 min. 10 μg of sample was injected after 1 min and the absorbance was measured at 280 nm. (B) SDS-PAGE, 0.5 μg of wild type and the 4 mutants of hEGF were loaded in tracks B–F; the molecular masses of some marker proteins are indicated in track A; the staining was with Coomassie blue.

substitution of leucine for alanine, valine, aspartic acid, glutamic acid and glycine. Sequencing of 17 plaques resulted in the isolation of all these mutants except glycine. The *PstI-SacI* fragments were then cloned back into pLF1-αF/EGF to produce pTD508, pTD509, pTD510 and pTD511 which produce the glutamic acid, valine, aspartic acid and alanine mutants respectively.

2.4. *Culture conditions and EGF purification*

Yeast transformants, which had been produced using a previously described protocol [15], were grown selectively in media containing per litre: 6.7 g nitrogen base w/o amino acid (Difco) and 10 g glucose, supplemented with selected amino acids. After 2–3 days of growth at 30°C in a rotary incubator the cultures were harvested and the EGF purified from cell free broth by reversed phase and ion exchange HPLC. Typically between 2 and 4 mg EGF was produced per litre of culture.

2.5. *Radioreceptor assay*

The ability of each purified protein to inhibit the binding of ^{125}I labelled murine EGF (Amersham) was measured using a monkey Vero cell line. The protein concentration was determined spectrophotometrically using $E^{280}_{0.1\%} = 2.89$/cm (which was calculated by summing the contributions of the tyrosine and tryptophan residues).

2.6. *NMR*

Samples for NMR studies were prepared by dissolving pure protein in D_2O (99.8% pure). After lyophilisation the protein was resuspended in D_2O and the pH adjusted, where necessary, with NaOD and DCl to pH 2.85–2.95 (uncorrected meter readings). NMR spectra were acquired at 303 K using a Bruker AM600 MHz spectrometer. One-dimensional spectra were collected in the pulsed Fourier mode with 400 transients per spectrum, and with pre-irradiation of the residual solvent resonance.

3. RESULTS AND DISCUSSION

Human EGF and the four leucine 47 mutants were secreted with relatively high efficiency using the yeast alpha factor secretion pathway. Purified hEGF was at least 95% pure as determined by reversed phase HPLC and SDS-PAGE (fig.2). Characterisation by amino acid sequencing and fast atom bombardment mass spectrometry (data not shown) showed clean N-terminal cleavage by the KEX2 protease of the alpha factor secretion pathway. The protein was found to be missing the C-terminal arginine 53, consistent with earlier results which show that this residue is removed by a yeast protease [16]. However, the 1–52 form of EGF is found in vivo [17] and has similar activity to the 1–53 form [9].

NMR was used to investigate the effect of the amino acid substitutions on the structure of the protein. Having already assigned the spectra of the 1–48 and 1–53 forms of hEGF [18,19], it was relatively easy to identify spectral features characteristic of 'correctly folded' hEGF. The downfield regions of one-dimensional spectra of hEGF and the four mutants are shown in fig.3 and assignments of several resonances are indicated.

The structural integrity of each mutant may be assessed from the shifts experienced by resonances. As is evident from fig.3 and other parts of the spectrum (not shown) similar patterns of shifts are observed for resonances in the mutant and wild type spectra, suggesting little perturbation to the folding of the protein. This is further suggested by similar patterns of nuclear Overhauser effects (NOEs) which are also characteristic of the folding of the protein. Backbone amide protons which are slow to exchange with solvent deuterons have been assigned to secondary structural features of hEGF [18,19]. Since the same amide protons are slowly exchanging in each protein (fig.3), similar networks of hydrogen bonds and secondary structure are indicated.

As well as structural information, qualitative details of dynamic behaviour are available from the NMR spectra. The fact that resonances corresponding to hydrogen bonded protons are observed indicates similar lifetimes for the hydrogen bonds and thus a similar degree of flexibility for each protein. More localised information can be obtained from the observation that the 3,5H doublet of tyrosine 29 is broadened; this can be attributed to a ring current interaction with tyrosine 22 and partially restricted motion of the tyrosine 29 ring. This broadening is observed in all spectra suggesting similar mobilities for this region in each protein.

These observations indicate that there is no major change in conformational or dynamic features of the leucine 47 mutants. Some features of the spectra of the mutants, however, differ from that of the wild type. Most differences can be directly related to the substitutions, for example the spectrum of the alanine mutant shows the appearance of a new doublet at 1.39 ppm which may be assigned to the β CH_3 of the new alanine. Two-dimensional spectra (not shown) were used to find the new resonances of residue 47 of the other mutants,

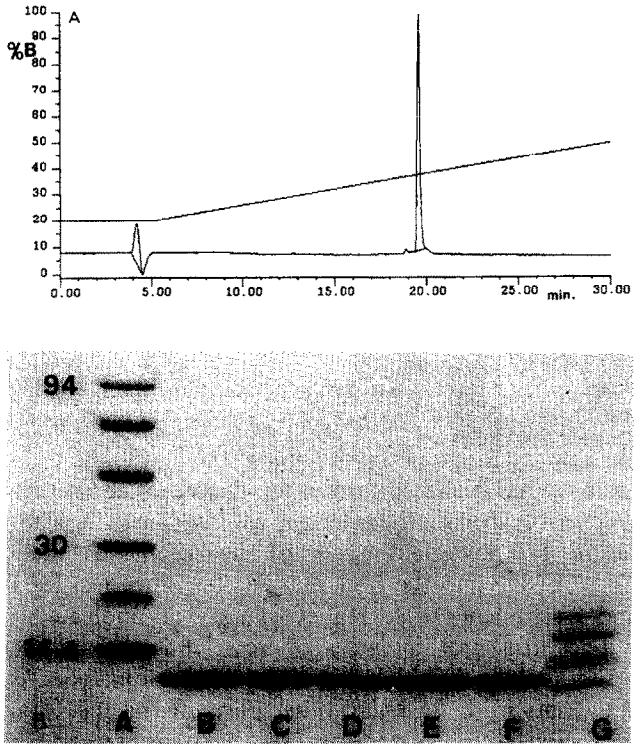

Fig.2. The downfield regions of one-dimensional NMR spectra of hEGF(1–52) and the four leucine 47 mutants. The resonances of slowly exchanging amide protons are marked on the top spectrum, as are some of the aromatic resonances of tyrosine 29, tryptophan 49 and tryptophan 50.

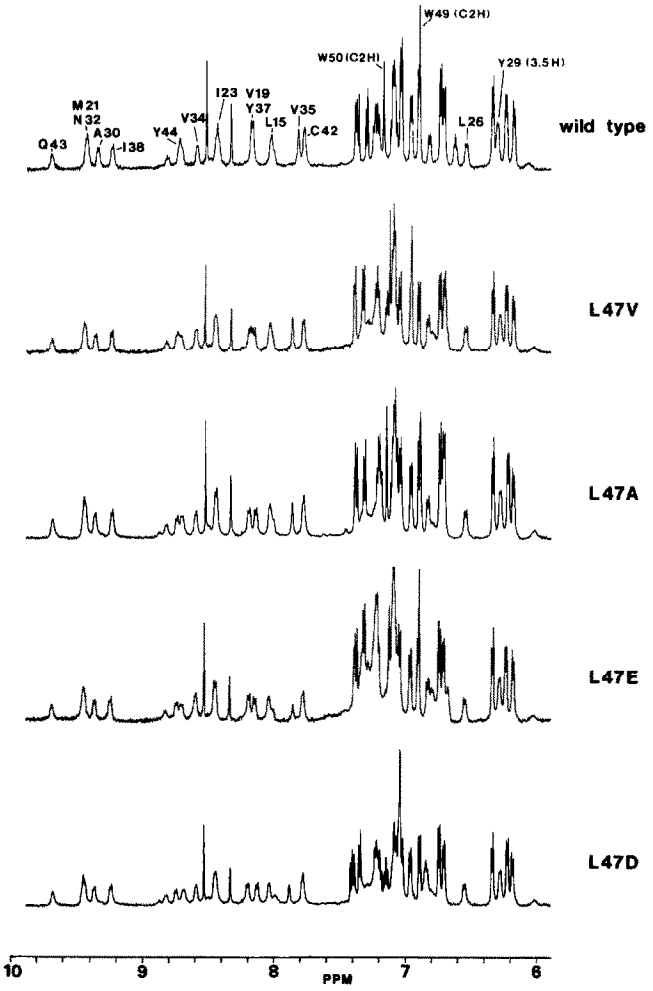

Fig.3. The radioreceptor assay of hEGF(1–52) and the four leucine 47 mutants on monkey Vero cell line. The relative amount of ^{125}I labelled murine EGF which was able to bind is plotted against the concentration of each protein. Lines were fitted to the data following the function, $F=((A-D)/[1+(X/C)^B])+D$, which gives the IC_{50} from parameter C [20]. (Parameter A was set to 100%.) Inset: IC_{50} values obtained from the data.

as well as to confirm the disappearance of resonances of leucine 47. Other changes are observed for the residues following leucine 47, in particular the resonances of tryptophans 49 and 50 can be seen to shift in fig.3. Additionally, the shifts of the methyl resonances of valine 35 vary slightly between the different proteins. These resonances are subject to a large ring current shift from tryptophans 49 and 50 (R.M. Cooke et al., manuscript in preparation). Thus some alteration in the packing of the tryptophans with valine 35 is suggested. It should be stressed, however, that residues after leucine 47 have been shown to be unimportant for receptor binding [9], and so any structural changes in this region are not likely to affect receptor binding.

Receptor binding studies show that the leucine 47 mutants have reduced binding affinity as seen by their reduced ability to compete with ^{125}I labelled murine EGF (which is indicated by their higher IC_{50} values shown in fig.4). Changing leucine 47 to a valine results in around a seven-fold increase in the IC_{50} whereas changes to alanine, aspartic acid or glutamic acid all result in around a further seven-fold increase in IC_{50}. Similar results are seen with human A431 cell line, though with more scatter of the data. These activities are broadly consistent with the data of others [10–12] which show a reduction in activity on substituting the residue corresponding to leucine 47 for another amino acid. For each mutant there is a strong correlation between the decrease in receptor binding and the decrease in mitogenic activity on a human fibroblast cell line (data not shown).

Thus, since no significant structural changes are detected on substituting leucine 47, the considerable reduction in receptor binding affinity on changing this amino acid even to valine strongly suggests that leucine 47 is involved in a highly specific contact with the receptor. Larger changes than that caused by valine result from the presence of alanine, glutamate or aspartate. This suggests that leucine 47 participates in a hydrophobic interaction with the receptor. The substitutions of valine and alanine possibly result in progressive diminution of the Van der Waals interac-

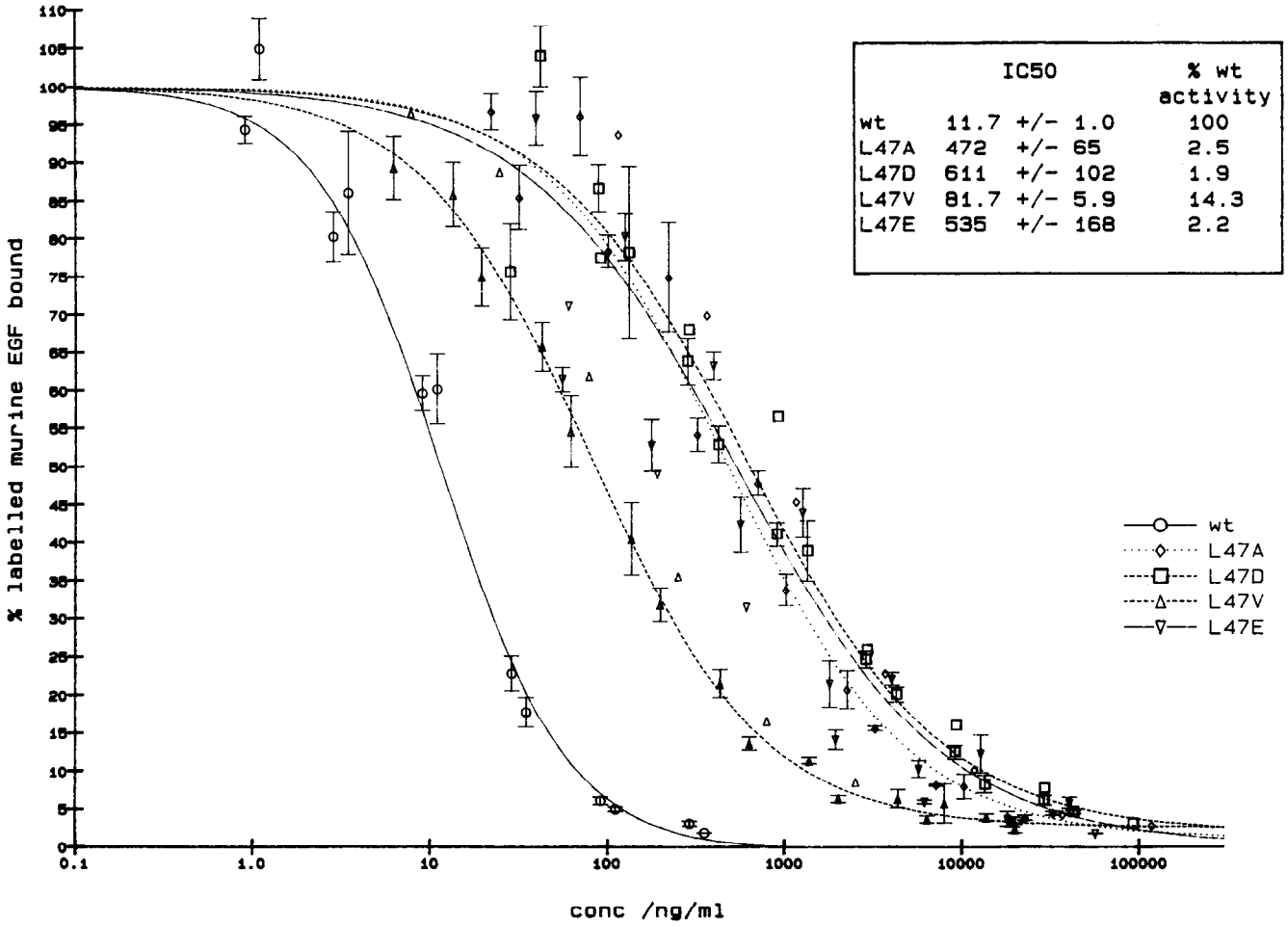

Fig.4. Plasmid map of pLF1-αF/EGF. Amp = ampicillin resistance gene; ori = *E. coli* origin of replication; Tet = tetracycline resistance gene (Amp-ori-Tet region from pBR322); pPGK = yeast phosphoglycerate kinase promoter; tPGK = yeast phosphoglycerate kinase terminator.

ting surface, while the introduction of a negative charge with glutamate or aspartate possibly destabilises the hydrophobic environment.

In conclusion it has been shown, by a combination of mutagenesis, binding assay and NMR techniques, that leucine 47 of hEGF is at the interface between the growth factor and its receptor. Although minimal structural perturbation was detected in the leucine 47 mutants, we have recently carried out preliminary NMR studies on other EGF mutants and have, in one case, observed significant spectral perturbations. This suggests that reductions in activity can sometimes be accompanied by significant changes in protein structure. Using NMR to distinguish between local and global structural changes, thus promises to be a useful additional tool in mapping the surface involved in growth factor-receptor interactions.

Acknowledgements: This is a contribution from the Oxford Centre for Molecular Sciences. We are grateful for the financial support provided by the SERC and Monsanto (for R.M.C.). We also thank Alan and Sue Kingsman for the supply of yeast expression vectors.

REFERENCES

[1] Burgess, A.W. (1989) Br. Med. Bull. 45, 401–424.
[2] Carpenter, G., King, L. and Cohen, S. (1978) Nature 276, 409–410.
[3] Marquardt, H. and Todaro, C.J. (1982) J. Biol. Chem. 257, 5220–5225.
[4] Stroobant, P., Rice, A.P., Gullick, W.J., Cheng, D.J., Kerr, I.M. and Waterfield, M.D. (1985) Cell 42, 383–393.
[5] Upton, C., Macen, J.L. and McFadden, G. (1987) J. Virol. 61, 1271–1275.
[6] Lin, Y.-Z., Caporaso, G., Chang, P.-Y., Ke, X.-H. and Tam, J.P. (1988) Biochemistry 27, 5640–5645.
[7] Campbell, I.D., Cooke, R.M., Baron, M., Harvey, T.S. and Tappin, M.J. (1989) Prog. Growth Factor Res. 1, 13–22.
[8] Nexo, E. and Hansen, H.F. (1985) Biochim. Biophys. Acta 843, 101–106.
[9] Gregory, H., Thomas, C.E., Young, J.A., Willshire, I.R. and Garner, A. (1988) Regul. Pept. 22, 217–226.
[10] Ray, P., Moy, F.J., Montelione, G.T., Liu, J.-F., Narang, S.A., Scheraga, H.A. and Wu, R. (1988) Biochemistry 27, 7289–7295.
[11] Engler, D.A., Matsunami, R.K., Campion, S.R., Stringer, C.D., Stevens, A. and Niyogi, S.K. (1988) J. Biol. Chem. 263, 12384–12390.
[12] Lazar, E., Watanabe, S., Dalton, S. and Sporn, M.B. (1988) Mol. Cell. Biol. 8, 1247–1252.

[13] Brake, A.J., Merryweather, J.P., Coit, D.G., Heberlein, V.A., Masiarz, F.R., Mullenbach, G.T., Urdea, M.S., Valenzuela, P. and Barr, P.J. (1984) Proc. Natl. Acad. Sci. USA 81, 4642–4646.
[14] Mellor, J., Dobson, M.J., Roberts, N.A., Tuite, M.F., Emtage, J.S., White, S., Lowe, P.A., Patel, T., Kingsman, A.J. and Kingsman, S.M. (1983) Gene 24, 1–14.
[15] Beggs, J.D. (1978) Nature 275, 104–109.
[16] George-Nascimento, C., Gynes, A., Halloran, M., Merryweather, J., Vanezuela, P., Steimer, K.S., Masiarz, F.R. and Randolph, A. (1988) Biochemistry 27, 797–802.
[17] Gregory, H. and Preston, B.M. (1977) Int. J. Pept. Prot. Res. 9, 107–118.
[18] Carver, J.A., Cooke, R.M., Esposito, G., Campbell, I.D., Gregory, H. and Sheard, B. (1986) FEBS Lett. 205, 77–81.
[19] Cooke, R.M., Wilkinson, A.J., Baron, M., Pastore, A., Tappin, M.J., Campbell, I.D., Gregory, H. and Sheard, B. (1987) Nature 327, 339–341.
[20] DeLean, A., Munson, P.J. and Rodbard, D. (1978) Am. J. Physiol. 235, E97–107.