The EMBO Journal vol.13 no.15 pp.3517–3523, 1994

Solution structure of the epidermal growth factor-like domain of heregulin-α, a ligand for p180^{erbB-4}

Koji Nagata[1,2], Daisuke Kohda[1], Hideki Hatanaka[1], Saori Ichikawa[1], Satoru Matsuda[3], Tadashi Yamamoto[3], Akinori Suzuki[2] and Fuyuhiko Inagaki[1,4]

[1]Department of Molecular Physiology, Tokyo Metropolitan Institute of Medical Science, Honkomagome 3-18-22, Bunkyo-ku, Tokyo 113, [2]Department of Agricultural Chemistry, Faculty of Agriculture, University of Tokyo, Yayoi 1-1-1, Bunkyo-ku, Tokyo 113 and [3]Department of Oncology, Institute of Medical Science, University of Tokyo, Shiroganedai 4-6-1, Minato-ku, Tokyo 108, Japan
[4]Corresponding author

Communicated by J.Schlessinger

p185^{erbB-2} and p180^{erbB-4} are epidermal growth factor (EGF) receptor-like tyrosine kinases, whose co-expression is observed in many breast carcinomas. Heregulins (HRGs), which contain an immunoglobulin unit and an EGF-like domain, bind to p180^{erbB-4} and activate p180^{erbB-4} and p185^{erbB-2} through transphosphorylation or receptor heterodimerization. The EGF-like domain is sufficient for the activation. Despite the sequence similarity, no cross activity is seen between the p180^{erbB-4} ligands (HRGs) and the p170^{erbB-1} ligands [EGF and transforming growth factor (TGF)-α]. To investigate the structural basis of receptor specificity, we have determined the solution structure of the EGF-like domain of HRG-α by two-dimensional ^{1}H nuclear magnetic resonance spectroscopy and simulated annealing calculations. Though its main-chain fold is similar to those of EGF and TGF-α, distinctive structural features are observed on the molecular surface including an ionic cluster and hydrophobic patches, which afford HRG-α the specific affinity for p180^{erbB-4}. The structure should provide a basis for the structure–activity relationship of HRGs and for the design of drugs which prevent progression of breast cancer.

Key words: epidermal growth factor-like domain/heregulin/nuclear magnetic resonance structure/p180^{erbB-4}/receptor recognition

Introduction

p185^{erbB-2} (HER-2 or Neu) (Bargmann *et al.*, 1986; Yamamoto *et al.*, 1986) and p180^{erbB-4} (HER-4) (Plowman *et al.*, 1993a) are receptor-type tyrosine kinases that are similar to p170^{erbB-1} [the epidermal growth factor (EGF) receptor or HER-1] (Ullrich *et al.*, 1984) and p160^{erbB-3} (HER-3) (Plowman *et al.*, 1990). Simultaneous overexpression of these *erbB* family tyrosine kinases is often observed in human cancer cells (Plowman *et al.*, 1993b). In particular, overexpression of p185^{erbB-2} is known to correlate with a poor prognosis in breast, ovarian and endometrial cancers and non-small cell lung adenocarcinoma (Holmes *et al.*, 1992). Because of the involvement of p185^{erbB-2} in human cancer and its similarity to the EGF receptor, the isolation of the ligand for the receptor has been a major focus. In 1992, the purification and cloning of the specific activators for p185^{erbB-2} were reported: (i) heregulin (HRG)-α, a 45 kDa glycoprotein isolated from the conditioned medium of MDA-MB-231 human breast carcinoma cells and its related molecules (HRG-β1, β2, β3) (Holmes *et al.*, 1992) and (ii) Neu differentiation factor (NDF), a rat homolog of HRG-α isolated from the conditioned medium of Rat1-EJ *ras*-transformed fibroblasts (Wen *et al.*, 1992). The HRG molecule is a mosaic glycoprotein containing an immunoglobulin unit and an EGF-like domain. The EGF-like domain has been demonstrated to be sufficient to activate p185^{erbB-2} (Holmes *et al.*, 1992). It has, however, been shown recently that HRG binds not to p185^{erbB-2} but to a homologous receptor, p180^{erbB-4}, and activates p185^{erbB-2} indirectly through transphosphorylation or receptor heterodimerization (Peles *et al.*, 1993; Plowman *et al.*, 1993b). No ligands for p185^{erbB-2} have yet been identified. The HRG family proteins have also been shown to play important roles in development, regeneration and tumor formation in the nervous system including stimulation of acetylcholine receptor synthesis (ARIA) (Falls *et al.*, 1993) and mitogenesis for Schwann cells (glial growth factors, GGFs) (Marchionni *et al.*, 1993).

The EGF-like domain, which is defined by six cysteine residues characteristically spaced over a sequence of 35–40 amino acid residues, is shared by many functionally diverse proteins including growth factors [e.g. EGF (Carpenter and Cohen, 1979, 1990), TGF-α (Derynck *et al.*, 1984), heparin-binding EGF-like growth factor (HB-EGF; Higashiyama *et al.*, 1991), amphiregulin (AR; Shoyab *et al.*, 1989) and vaccinia virus growth factor (VGF; Blomquist *et al.*, 1984)], cell adhesion molecules (e.g. laminin) and plasma proteins (e.g. protein C). The EGF-like domains of HRG-α and βs share sequence identity with the EGF family growth factors [HB-EGF (45% within the region between the first and sixth cysteine residues), TGF-α (32%) and EGF (27%)] (Figure 1). In spite of the sequence similarity, the EGF-like domains of HRGs and the EGF family growth factors are biologically distinct; the EGF-like domains of HRGs bind specifically to p180^{erbB-4} but not to p170^{erbB-1}, whereas HB-EGF, TGF-α, EGF, AR and VGF bind to p170^{erbB-1} but not to p180^{erbB-4} (Holmes *et al.*, 1992; Wen *et al.*, 1992; Plowman *et al.*, 1993b). In order to elucidate the receptor specificity on a structural basis, we have determined the high-resolution solution structure of the EGF-like domain of HRG-α at pH 6.0 at 37°C by two-dimensional ^{1}H nuclear magnetic resonance (NMR) spectroscopy and simulated

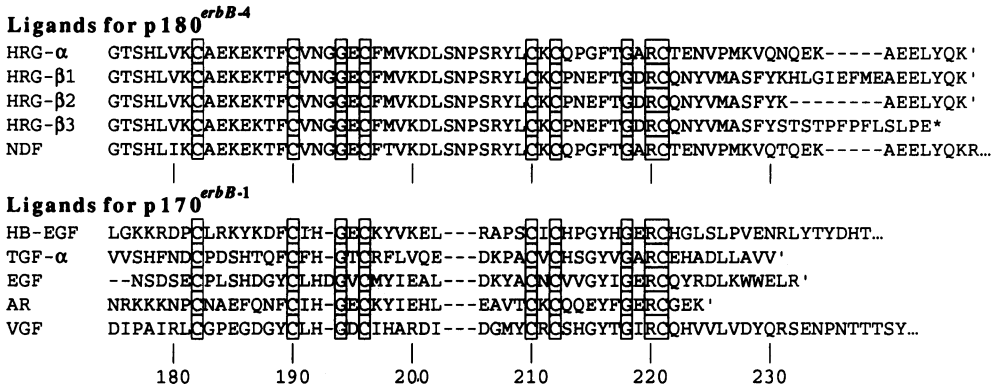

Fig. 1. Amino acid sequences of the EGF-like domain of HRGs and its comparison with the EGF family growth factors. Numbering of residues is based on the proHRG sequence so that the sequences begin with amino acid 175 of the proforms. Strictly conserved amino acids are boxed. Dashes represent gaps in the sequence and are inserted for best alignment. Apostrophes, proposed or known COOH-terminal processing; asterisks, positions of the COOH-terminus of proHRG-β3.

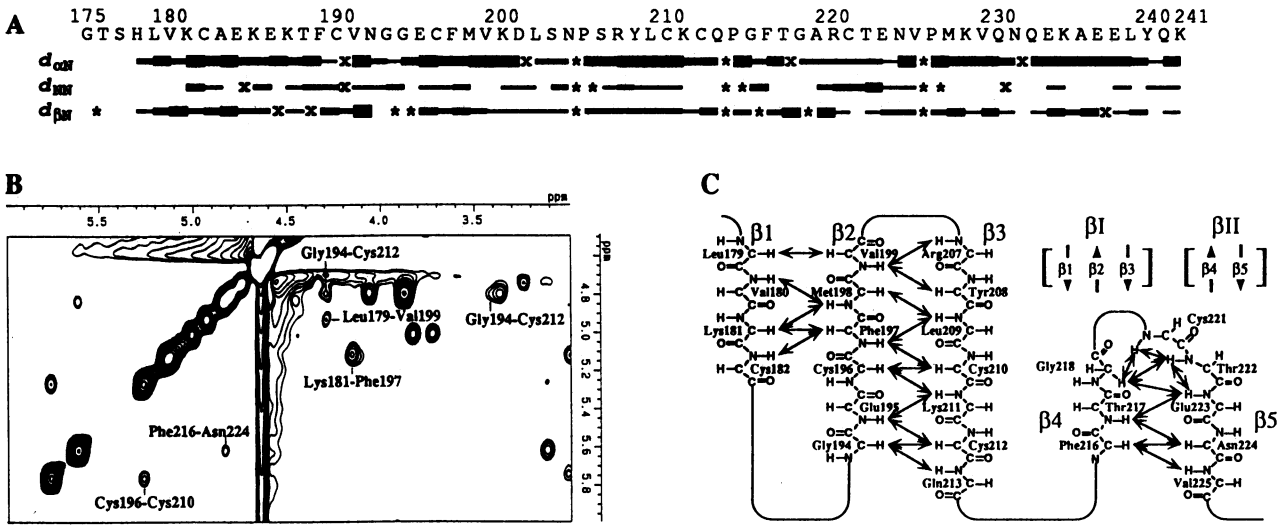

Fig. 2. Sequential assignments and secondary structure of the EGF-like domain of HRG-α. (**A**) Sequential NOE connectivities. The height of the bars indicates the approximate intensity of the NOESY cross-peaks recorded with a mixing time of 75 ms. *, undefined connectivity; ×, NOE connectivity which is not clearly observed due to overlapping with other NOE peaks. (**B**) $C^{\alpha}H-C^{\alpha}H$ region of the NOESY spectrum in D_2O (75 ms mixing time). Interstrand NOEs between α protons are labeled with the assignments. (**C**) Secondary structure of the EGF-like domain of HRG-α identified with interstrand backbone NOE connectivities. The β-strand β3 (residues 207–213) extends over the N- and C-terminal subdomains (residues 175–211 and 212–241, respectively).

annealing calculations, and compared it with the reported structures of EGF and TGF-α.

Results

Secondary structure

The EGF-like domain of HRG-α [HRG-α (175–241)] was synthesized by the solid-phase peptide synthesis based on Boc chemistry (K.Nagata, D.Kohda, S.Matsuda, T.Yamamoto, A.Suzuki and F.Inagaki, in preparation). Two-dimensional ^1H NMR spectra were recorded at 600 MHz at pH 6.0 and 37°C. The resonances were assigned to individual protons in a sequence-specific manner using the sequential assignment method (Figure 2A) (Wüthrich, 1986). Six NOE crosspeaks between α protons were observed indicating the presence of antiparallel β-sheets (Figure 2B). The secondary structure of the EGF-like domain of HRG-α was determined from the analysis of the interstrand NOEs (Figure 2C). The EGF-like domain of HRG-α is composed of five β-strands (termed β1−β5) which form two antiparallel β-sheets (termed βI and βII). The β-sheet in the N-terminal region, βI, is composed of three β-strands: β1 (residues Leu179−Cys182), β2 (residues Gly194−Val199) and β3 (residues Arg207−Gln213); the β-sheet in the C-terminal region, βII, is composed of two short β-strands: β4 (residues Phe216−Thr217) and β5 (residues Glu223−Asn224).

Tertiary structure

A total of 534 distance constraints which included 239 intraresidual, 138 sequential ($|i-j|=1$), 41 short-range ($2 \leq |i-j| \leq 5$) and 116 long-range ($|i-j| \geq 6$) constraints were derived from the assigned NOE crosspeaks measured with a mixing time of 75 ms. Dihedral angle constraints including 15 φ and 7 χ1 were obtained. No constraints for hydrogen bonds were obtained because slowly exchanging amide protons were not detected at pH 6.0 at 37°C. Based on the structural analyses of cystine-containing proteolytic fragments, the three disulfide bonds were assigned to Cys182−Cys196, Cys190−Cys210 and Cys212−Cys221

Table I. Structural statistics

	<SA>[a]	$(SA)_r$[b]
R.m.s.d.s from experimental		
distance constraints (Å) (537)	0.084 ± 0.002	0.079
Number of distance constraint violations >0.5 Å	1–3[c] (maximum 0.72 Å)	1[c] (0.51 Å)
F_{NOE} (kcal/mol)[d]	187.2 ± 8.5	167.8
F_{repel} (kcal/mol)[d]	118.5 ± 3.6	106.0
E_{L-J} (kcal/mol)[e]	−171.4 ± 14.4	−131.9
R.m.s.d.s from idealized geometry		
Bonds (Å) (1047)	0.007 ± 0.0003	0.006
Angles (degrees) (1893)	2.05 ± 0.02	2.00
Impropers (degrees) (408)[f]	1.13 ± 0.02	1.09

[a]<SA> are the 10 refined simulated annealing structures.
[b]$(SA)_r$ is the restrained minimized mean structure, where the mean structure was obtained by averaging the coordinates of the individual <SA> structures best-fitted to each other.
[c]The set of upper-bound distance constraints obtained is tighter than in the usually used three-level classification (strong ≤2.8 Å, medium ≤3.5 Å and weak ≤5.0 Å). Hence, a few violations >0.5 Å remained, but posed no problems.
[d]The value of the square-well NOE potential, F_{NOE}, is calculated with a force constant of 50 kcal/mol/Å^2. The value of F_{repel} is calculated with a force constant of 4 kcal/mol/Å^4 with the van der Waals radii scaled by a factor of 0.8 of the standard values used in the CHARMm empirical function.
[e]E_{L-J} is the Lennard–Jones van der Waals energy calculated with the CHARMm empirical energy function, which was not included in the simulated annealing calculations.
[f]The improper torsion term is used to maintain the planar geometry and chirality.

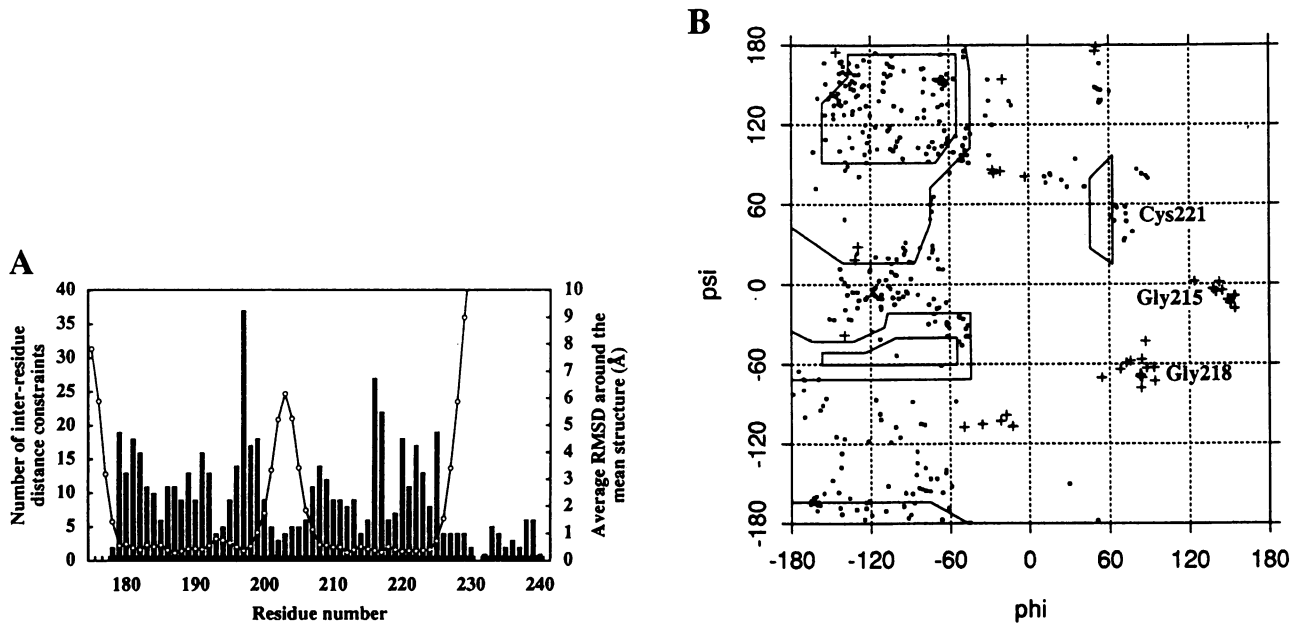

Fig. 3. (**A**) Number of NOE constraints and r.m.s.d.s for each residue. The number of inter-residue distance constraints (columns) and the average values of the main-chain (N, C^α, C′) r.m.s.d.s (circles) were plotted as a function of residue number. (**B**) Ramachandran plot for the well-defined regions (residues 179–199 and 207–225) of the final 10 structures. +, glycine residues. The residual violations are located mainly in the irregular helix region (around $\phi = -100°$, $\psi = 0°$) and loop regions.

(K.Nagata, D.Kohda, S.Matsuda, T.Yamamoto, A.Suzuki and F.Inagaki, in preparation), the same as those for EGF and TGF-α; the three distance constraints of the disulfide bonds were thus obtained. The three-dimensional structures were calculated with X-PLOR (Brünger, 1990) using the simulated annealing protocol (YASAP) on the 559 distance and dihedral angle constraints. A total of 100 calculations were carried out, and a final set of 10 structures was selected on the basis of agreement with the experimental constraints and van der Waals energy, with the cut-off taken at $F_{NOE} + F_{repel} < 325.13$ kcal/mol (Table I). The number of inter-residue distance constraints and average root-mean-square deviations (r.m.s.d.s) around the mean structure for each residue (Figure 3A) and the Ramachandran plot for the 10 structures (Figure 3B) are shown. The structure was well-defined except for the N-terminal region (residues Gly175–His178), central loop region (residues Lys200–Ser206) and the C-terminal flanking region outside the EGF-like structural unit (residues Pro226– Lys241) (Figure 3A). The r.m.s.d.s between the final 10 structures and the mean structure are 0.58 ± 0.11 Å for the backbone heavy atoms (N, C^α, C′)

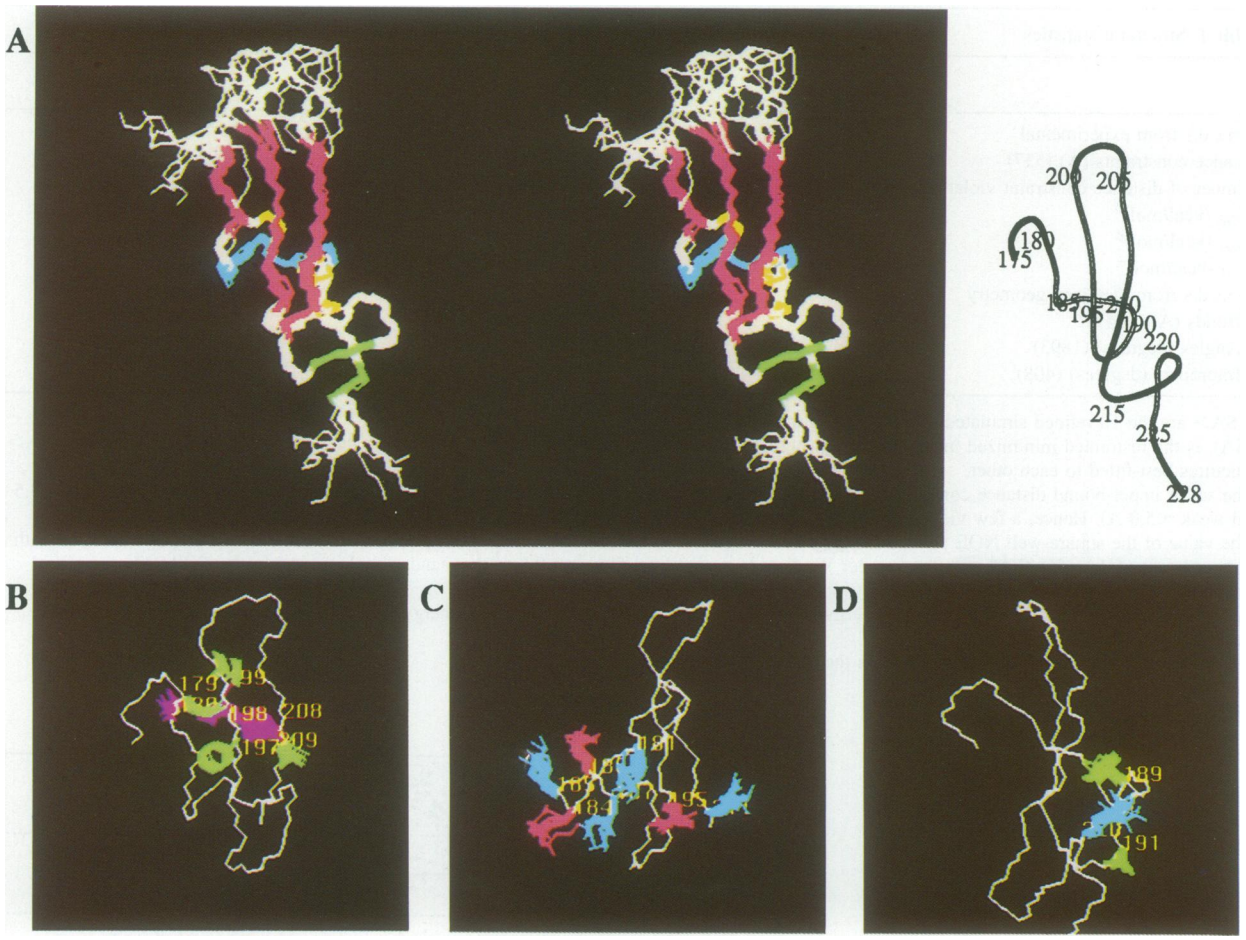

Fig. 4. Solution structure of the EGF-like domain of HRG-α (residues 175–228). The 10 converged structures are superimposed. The C-terminal flanking region outside the EGF-like domain (Val229−Lys241) is highly disordered and is not shown. (**A**) Main-chain atoms (N, C^{α}, C′; stereo view). The β-strands β1, β2 and β3 in the N-terminal β-sheet βI are shown in red, the irregular helix in blue and the β-strands β4 and β5 in the C-terminal β-sheet βII in green. (**B**) Hydrophobic patches on both sides of the N-terminal β-sheet βI. Averaged main chain of the N-terminal subdomain and overlaid hydrophobic side chains are shown. The hydrophobic side chains projecting above and below the sheet are colored green and purple, respectively. (**C**) Ionic cluster on the irregular helix and βI viewed from the left side of (A). Averaged main chain of the N-terminal subdomain and overlaid charged side chains are shown. Acidic and basic residues are colored in red and blue, respectively. (**D**) Hydrophobic pocket in the subdomain interface viewed from the right side of (A). Averaged main chain and overlaid side chains are shown. The side chain of Arg220 is colored blue; those of Phe189 and Val191 are green. All the amino acid residues with side chains are labeled with sequence numbers.

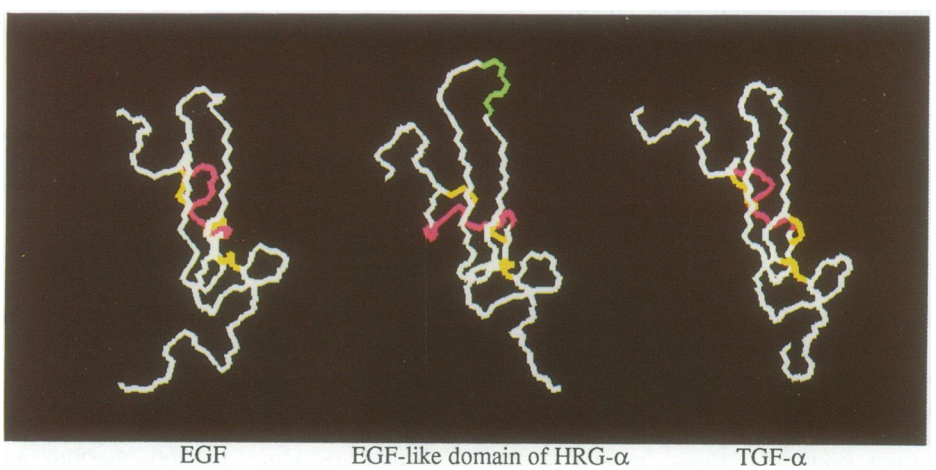

Fig. 5. Comparison of the main-chain structure of the EGF-like domain of HRG-α with those of EGF (Brookhaven Protein Data Bank entry 1EPG; Kohda and Inagaki, 1992b) and TGF-α (Brookhaven Protein Data Bank entry 4TGF; Kline *et al.*, 1990). Main-chain atoms (N, C^{α}, C′) and disulfide bonds of the mean structure are shown. The main chains are shown in white; disulfide bonds in yellow. The irregular helices (residues Glu184−Phe189) are colored red. The three amino acid insertion (Ser203-Asn204-Pro205) in the EGF-like domain of HRG-α is colored green.

Fig. 6. Proposed residues involved in the p180^{erbB-4} recognition of the EGF-like domain of HRG-α. Averaged main chain and overlaid side chains considered to be important for the binding of HRG-α to p180^{erbB-4} are shown. The side chains of residues proposed to be involved in the p180^{erbB-4} recognition of HRG-α are colored; acidic, basic and hydrophobic residues are shown in red, blue and green, respectively. The main chain is shown in white. All the amino acid residues with side chains are labeled with sequence numbers.

and 1.29 ± 0.31 Å for all non-hydrogen atoms except for the less well-defined regions.

The tertiary structure of the EGF-like domain of HRG-α (Figure 4A) is characterized by two subdomains: the N-terminal subdomain (residues Gly175–Lys211) and the C-terminal subdomain (residues Cys212–Lys228). The N-terminal subdomain consists of a triple-stranded anti-parallel β-sheet (βI) involving peptide segments Leu179–Cys182 (β1), Gly194–Val199 (β2) and Arg207–Gln213 (β3). The triple-stranded β-sheet βI is stabilized by hydrophobic clusters on both sides of the sheet including the Leu179, Phe197, Val199 and Leu209 side chains on one side and the Val180, Met198 and Tyr208 side chains on the other side (Figure 4B). Peptide segment Glu184–Phe189 adopts an irregular helix which is attached to βI by the Cys182–Cys196 and Cys190–Cys210 disulfide bonds. The side chains of residues on the irregular helix (Glu184-Lys185-Glu186-Lys187) and βI (Lys181, Glu195 and Lys211) form an ionic cluster (Figure 4C). The loop between β2 and β3 (residues Lys200–Ser206), where HRG-α has a three amino acid insertion not found in EGF and TGF-α, takes a flexible structure. In the flexible loop, a putative Asn glycosylation site (Asn204-Pro205-Ser206) is found, but the glycosylation is not required for the EGF-like domain of HRGs to activate p180^{erbB-4} or p185^{erbB-2}, as demonstrated by bioassay (Holmes *et al.*, 1992 and data not shown). The C-terminal subdomain forms a small anti-parallel double hairpin structure involving peptide segments Cys212–Gln213 (part of β3), Phe216–Thr217 (β4) and Glu223–Asn224 (β5). Residues Gln213–Phe216 between β3 and β4 form a distorted type II β-turn. Peptide segment Gly218–Thr222 forms a multiple-bend structure which looks like a single turn of a left-handed helix. The unusual bent structure, with positive φ values for the residues Gly215, Gly218 and Cys221 (Figure 3B), is stabilized by the Cys212–Cys221 disulfide bond. Though the C-termini of EGF and TGF-α are the residues 232 and 228, respectively, it is proposed that HRG-α might have a longer C-terminal tail terminating at residue 241 (Holmes *et al.*, 1992). But the C-terminal flanking region lying outside the EGF-like structural unit (residues Val229–Lys241) is diverse in sequence among HRGs (Figure 1); furthermore, it is not uniquely defined, with only a few long-range NOEs (Figure 3A), suggesting that the region is not required for HRG-α to bind to p180^{erbB-4}. The flexibility of the C-terminal flanking region is probably required for the processing of the membrane-anchored precursor of HRG-α (proHRG-α) (Holmes *et al.*, 1992; Wen *et al.*, 1992).

Discussion

Structure comparison with EGF and TGF-α

In spite of the low sequence similarity (27 and 32% to EGF and TGF-α, respectively), the overall main-chain fold of the EGF-like domain of HRG-α is similar to those of EGF/TGF-α previously reported (Figure 5; Kline *et al.*, 1990; Harvey *et al.*, 1991; Hommel *et al.*, 1992; Kohda and Inagaki, 1992a,b; Montelione *et al.*, 1992; Moy *et al.*, 1993). When compared with EGF/TGF-α, r.m.s.d.s are 1.74 and 1.44 Å, respectively, for the main-chain atoms (N, C^{α}, C′) excluding the divergent and less well-defined regions. Nine residues are thoroughly conserved in the EGF-like domains of HRGs and other EGF family growth factors: the six cysteine residues, Gly194, Gly218 and Arg220. The linkages of the three disulfide bonds in the EGF-like domain of HRG-α are just the same as those in EGF/TGF-α (K.Nagata *et al.*, in preparation) and the conserved glycine residues in both HRGs and EGF/TGF-α take positive φ values (Figure 3B). These results indicate that the conserved cysteine and glycine residues are essential to construct the characteristic backbone fold of the EGF-like structural unit. The conserved Arg220 is located in the subdomain interface with its side chain

lying in a hydrophobic pocket formed by Phe189 and Val191 (Figure 4D), as is the case for EGF and TGF-α (Kohda and Inagaki, 1992b; Hommel *et al.*, 1992; Moy *et al.*, 1993). In EGF, Arg220 is essential for the p170^{erbB-1} recognition; the substitution of Lys, Gln, His, Leu or Ala for Arg220 drastically reduced the affinity for p170^{erbB-1}, and yet did not significantly alter the native protein conformation (Engler *et al.*, 1990; Hommel *et al.*, 1991). Thus Arg220 in HRG-α is probably required functionally to confer to the molecule an affinity for p180^{erbB-4} rather than structurally to define the inter-subdomain orientation.

Despite the similarity of the overall main-chain fold in the EGF-like domains of HRG-α and EGF/TGF-α, some structural features are uniquely observed in HRG-α. First, the N-terminal β-strand β1 in HRG-α is well-defined (Figure 4A), instead of being disordered or less well-defined as in EGF/TGF-α (Kline *et al.*, 1990; Harvey *et al.*, 1991; Hommel *et al.*, 1992; Kohda and Inagaki, 1992a,b; Montelione *et al.*, 1992; Moy *et al.*, 1993). The strand β1 in HRG-α makes tight hydrophobic interactions with β2 among the successive hydrophobic residues Leu179– Val180 on β1 and the residues Phe197, Met198 and Val199 on β2 (Figure 4B), whereas such interactions are weak in EGF/TGF-α because their N-terminal peptide segment contains only one hydrophobic residue (Phe179 in TGF-α) or none (in EGF). Second, the orientation of the irregular helix relative to the N-terminal β-sheet βI is different between HRG-α and EGF/TGF-α. The helix in HRG-α runs perpendicularly across the bottom of the antiparallel β-strands, β2 and β3, while the helix in EGF/TGF-α runs obliquely across the β-strands (Figure 5). In EGF/TGF-α, the His186 side chains on the helix interacts with Tyr198/Phe198 and Tyr208/Pro208 side chains on the β-sheet to form a solvent inaccessible pocket (Harvey *et al.*, 1991; Kohda and Inagaki, 1992b; Montelione *et al.*, 1992). However, such interactions are not present in HRG-α. Instead, an ionic cluster is formed among the residues on the irregular helix (Glu184-Lys185-Glu186-Lys187) and on βI (Lys181, Glu195 and Lys211) (Figure 4C). Third, the helix observed in the C-terminal region of EGF (residues Leu226–Glu230) at pH 2.9 (Hommel *et al.*, 1992) is not found in HRG-α at pH 6.0. The corresponding region of HRG-α at pH 6.0 does not have a well-defined conformation. Such a difference may be derived from the difference in pH. The main chain and side chain orientation of the EGF molecule in the C-terminal region has been shown to change with changes between acidic and neutral pH values; the helix observed at pH 2.0 is disturbed at pH 6.8 (Kohda and Inagaki, 1992b).

Determinants of HRG-α essential for binding to p180^{erbB-4}

Three regions have been proposed for the receptor (p170^{erbB-1})-binding surface in EGF based on the mutagenic studies: (i) the face of the N-terminal β-sheet βI on the opposite side of the irregular helix (Ile199; Campion *et al.*, 1990; Koide *et al.*, 1992a), (ii) the inter-subdomain cleft (Arg220; Engler *et al.*, 1990; Hommel *et al.*, 1991) and (iii) the C-terminus (Leu226; Moy *et al.*, 1989; Dudgeon *et al.*, 1990). Because of the structural similarity between the EGF-like domain of HRG-α and EGF/TGF-α and the sequence similarity between p180^{erbB-4} and p170^{erbB-1}, HRG-α is considered to have a similar receptor (p180^{erbB-4})-recognition surface to that of EGF/TGF-α. Thus, we propose that the residues which are strictly conserved in HRGs but not conserved in EGF/TGF-α that are located in the above three regions should be involved in p180^{erbB-4} recognition (Figure 6). Leu226 in EGF/TGF-α is changed to Pro and Met in HRG-α and βs, respectively. This residue is conserved in all the p170^{erbB-1} ligands except AR, which terminates at 224. The EGF analogs in which Leu226 is substituted by Ser, Val, Glu, Asp and Ala possess a reduced affinity for p170^{erbB-1} by 1/7 to 1/53 compared with the wild type (Moy *et al.*, 1989; Dudgeon *et al.*, 1990). Therefore, the alteration of Leu226 to Pro or Met in HRGs should result in a marked reduction in the affinity for p170^{erbB-1}. Pro or Met at 226, together with Val225, which is thoroughly conserved in HRGs, may be essential for the specific binding of HRGs to p180^{erbB-4} (Figure 6). The hydrophobic patch (Leu179, Phe197, Val199 and Leu209; Figure 4B) and the adjacent ionic cluster (Lys181, Glu195 and Lys211; Figure 4C) on the N-terminal β-sheet βI (Figure 6) form a characteristic surface of HRG-α, which is probably important for specific binding to p180^{erbB-4}. The hydrophobic patch is more extensive than those of EGF and TGF-α. Since the side chain of the residue at 209 is required to be small for the binding of EGF to p170^{erbB-1} (Koide *et al.*, 1992b), Leu209 in HRGs reduces the affinity to p170^{erbB-1} and may enhance the affinity to p180^{erbB-4}. The ionic cluster on βI including residues Lys181, Glu195 and Lys211 is absent in the high-affinity ligands for p170^{erbB-1} and may be advantageous for the binding to p180^{erbB-4}. The distribution of potential receptor-recognition sites around the whole EGF-like domain of HRG-α (Figure 6) implies that the ligand might interact with the p180^{erbB-4} receptor at multiple sites. It is true in the case of insulin–insulin receptor where two of the three extracellular domains of the receptor (domains I and III) are involved in insulin binding (Schumacher *et al.*, 1993).

In the present study we have determined the solution structure of the EGF-like domain of HRG-α and showed that its main-chain fold is similar to those of EGF and TGF-α. We propose here that the receptor specificity between the EGF-like domain of HRG-α and EGF/TGF-α may be the result of different exposed patches on the potential receptor binding surface. The structure should provide a basis for detailed and precise discussion on the structure–activity relationship of the HRGs and the EGF family proteins and gives a clue to design of drugs which prevent progression of breast cancer.

Materials and methods

NMR measurements

The EGF-like domain of HRG-α [HRG-α (175–241)] was synthesized by solid-phase synthesis based on the Boc chemistry. The three disulfide bonds were assigned by structural analyses of cystine-containing proteolytic fragments (K.Nagata, D.Kohda, S.Matsuda, T.Yamamoto, A.Suzuki and F.Inagaki, in preparation). The synthetic HRG-α (175–241) stimulated tyrosine phosphorylation of p185^{erbB-2} at 0.3 nM on MDA-MB-453 human breast carcinoma cells as reported (Holmes *et al.*, 1992). The synthetic HRG-α (175–241) was dissolved at a concentration of 4 or 2 mM in 90% H_2O/10% D_2O or D_2O (pH 6.0, direct meter reading), respectively. ^1H NMR spectra were measured at 600 MHz on a JEOL JNM-α600 spectrometer at 37°C. DQF-COSY (Rance *et al.*, 1983), PE-

COSY (Müller, 1987), TOCSY (50 ms mixing time) with a modified DIPSI-2 pulse sequence (Cavanagh and Rance, 1992) and NOESY (75 or 150 ms mixing time; Jeener *et al.*, 1979; Macura *et al.*, 1981) were recorded in the phase-sensitive mode (States *et al.*, 1982). Water resonance was suppressed by DANTE pulse (Zuiderweg *et al.*, 1986). Two-dimensional spectra were recorded using a data size of 512 (t_1)×1024 (t_2) (512×2048 for PE-COSY) with a spectral width of 7000 Hz. After zero-filling once in the t_2 and twice in the t_1 dimension, 2048×2048 real data matrices were finally obtained and digital resolution was 3.4 Hz/point in both dimensions (512×4096 real data matrix and 1.7 Hz/point digital resolution in the F_2 dimension for PE-COSY).

Structure calculations

Interproton distance constraints were derived from NOE crosspeak intensities (peak height) in the NOESY spectra recorded with a mixing time of 75 ms according to the method of Hatanaka *et al.* (1994). Crosspeaks in the NOESY spectra were picked using a homemade C program and edited with Felix (Biosym Technologies, Inc., San Diego, CA). The peak intensities were translated into distances on the basis of the relation of NOE intensity $\propto$ (distance)$^{-6}$ and a standard distance of sequential $d_{\alpha N}$ in β-sheet = 2.2 Å (Wüthrich, 1986). The upper-bound distance constraints were the calculated distance plus 0.5 Å. The lower-bound constraints were set to 1.8 Å. The distances involving methylene and methyl protons and ring protons of phenylalanine and tyrosine were referred to as single $(<r^{-6}>)^{-1/6}$ average distances so that no corrections for center averaging were made (Clore *et al.*, 1986). Dihedral angle constraints were obtained based on the analysis of DQF-COSY, PE-COSY and/or NOESY spectra (Wagner *et al.*, 1988). The three-dimensional structures were calculated by the simulated annealing method with X-PLOR (Molecular Simulations, Inc., Walthan, MA) using the distance and dihedral angle constraints. A final set of 10 converged structures was selected from 100 calculations on the basis of agreement with the experimental data and van der Waals energy. A mean structure was obtained by averaging the coordinates of the structures that were superimposed in advance to the second-best converged structure and then minimizing under the constraints (Clore *et al.*, 1986).

Acknowledgements

We are grateful to Drs Chieko Kitada, Akira Isogai, Hiromichi Nagasawa and Hiroshi Kataoka for helpful discussion and Hiroyuki Fukuda for technical advice. This study was performed through Special Coordination Funds of the Science and Technology Agency of the Japanese Government. This study was also supported by a grant from the Human Frontier Science Program.

References

Bargmann,C.I., Hung,M.-C. and Weinberg,R.A. (1986) *Nature*, **319**, 226–230.

Blomquist,M.D., Hunt,L.T. and Baker,W.C. (1984) *Proc. Natl Acad. Sci. USA*, **81**, 7363–7367.

Brünger,A.T. (1990) *X-PLOR Software Manual Version 2.1*. Yale University Press, New Haven, USA.

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

Carpenter,G. and Cohen,S. (1979) *Annu. Rev. Biochem.*, **48**, 193–216.

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

Cavanagh,J. and Rance,M. (1992) *J. Magn. Reson.*, **96**, 670–678.

Clore,G.M., Brünger,A.T., Karplus,M. and Gronenborn,A.M. (1986) *J. Mol. Biol.*, **191**, 523–551.

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

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

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

Falls,D.L., Rosen,K.M., Corfas,G., Lane,W.S. and Fischbach,G.D. (1993) *Cell*, **72**, 801–815.

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

Hatanaka,H., Oka,M., Kohda,D., Tate,S., Suda,A., Tamiya,N. and Inagaki,F. (1994) *J. Mol. Biol.*, **240**, 155–166.

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

Holmes,W.E. *et al.* (1992) *Science*, **256**, 1205–1210.

Hommel,U., Dudgeon,T.J., Fallon,A., Edwards,R.M. and Campbell,I.D. (1991) *Biochemistry*, **30**, 8891–8898.

Hommel,U., Harvey,T.S., Driscoll,P.C. and Campbell,I.D. (1992) *J. Mol. Biol.*, **227**, 271–282.

Jeener,J., Meier,B.H., Bachmann,P. and Ernst,R.R. (1979) *J. Chem. Phys.*, **71**, 4546–4553.

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

Kohda,D. and Inagaki,F. (1992a) *Biochemistry*, **31**, 677–685.

Kohda,D. and Inagaki,F. (1992b) *Biochemistry*, **31**, 11928–11939.

Kohda,D., Hatanaka,H., Odaka,M., Mandiyan,V., Ullrich,A., Schlessinger,J. and Inagaki,F. (1993) *Cell*, **72**, 953–960.

Koide,H. *et al.* (1992a) *Biochim. Biophys. Acta*, **1120**, 257–261.

Koide,H. *et al.* (1992b) *FEBS Lett.*, **302**, 39–42.

Macura,S., Huang,Y., Suter,D. and Ernst,R.R. (1981) *J. Magn. Reson.*, **43**, 259–281.

Marchionni,M.A. *et al.* (1993) *Nature*, **362**, 312–318.

Montelione,G.T., Wüthrich,K., Burgess,A.W., Nice,E.C., Wagner,G., Gibson,K.D. and Scheraga,H.A. (1992) *Biochemistry*, **31**, 236–249.

Moy,F.J., Scheraga,H.A., Liu,J.F., Wu,R. and Montelione,G.T. (1989) *Proc. Natl Acad. Sci. USA.*, **86**, 9836–9840.

Moy,F.J., Li,Y.-C., Rauenbuehler,P., Winkler,M.E., Scheraga,H.A. and Montelione,G.T. (1993) *Biochemistry*, **32**, 7334–7353.

Müller,L. (1987) *J. Magn. Reson.*, **72**, 191–196.

Peles,E., Ben-Levy,R., Tzahar,E., Liu,N., Wen,D. and Yarden,Y. (1993) *EMBO J.*, **12**, 961–971.

Plowman,G.D., Whitney,G.S., Neubauer,M.G., Green,J.M., McDonald, V.L., Todaro,G.J. and Shoyab,M. (1990) *Proc. Natl Acad. Sci. USA*, **87**, 4905–4909.

Plowman,G.D., Culouscou,J.-M., Whitney,G.S., Green,J.M., Carlton,G.W., Foy,L., Neubauer,M.G. and Shoyab,M. (1993a) *Proc. Natl Acad. Sci. USA*, **90**, 1746–1750.

Plowman,G.D., Green,J.M., Culouscou,J.-M., Carlton,G.W., Rothwell,V.M. and Buckley,S. (1993b) *Nature*, **366**, 473–475.

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

Schumacher,R., Soos,M.A., Schlessinger,J., Brandenburg,D., Siddle,K. and Ullrich,A. (1993) *J. Biol. Chem.*, **268**, 1087–1094.

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

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

Ullrich,A. *et al.* (1984) *Nature*, **309**, 418–425.

Wagner,G., Braun,W., Havel,T.F., Schaumann,T., Go,N. and Wüthrich,K. (1988) *J. Mol. Biol.*, **196**, 611–639.

Wen,D. *et al.* (1992) *Cell*, **69**, 559–572.

Wüthrich,K. (1986) *NMR of Proteins and Nucleic Acids*. John Wiley and Sons, New York, USA.

Yamamoto,T., Ikawa,S., Akiyama,T., Semba,K., Nomura,N., Miyajima,N., Saito,T. and Toyoshima,K. (1986) *Nature*, **319**, 230–234.

Zuiderweg,E.P.R., Hallenga,K. and Olejniczak,E.T. (1986) *J. Magn. Reson.*, **70**, 336–343.

Received on March 16, 1994; revised on May 16, 1994