Cell, Vol. 110, 763–773, September 20, 2002, Copyright ©2002 by Cell Press

Crystal Structure of a Truncated Epidermal Growth Factor Receptor Extracellular Domain Bound to Transforming Growth Factor α

Thomas P.J. Garrett,[1,2,5] Neil M. McKern,[2,4] Meizhen Lou,[1,2] Thomas C. Elleman,[4] Timothy E. Adams,[2,4] George O. Lovrecz,[2,4] Hong-Jian Zhu,[2,3] Francesca Walker,[2,3] Morry J. Frenkel,[2,4] Peter A. Hoyne,[2,4] Robert N. Jorissen,[2,3] Edouard C. Nice,[2,3] Antony W. Burgess,[2,3,5] and Colin W. Ward[2,4,5]

[1]Walter and Eliza Hall Institute of Medical Research

[2]Cooperative Research Centre for Cellular Growth Factors

[3]Ludwig Institute for Cancer Research
Post Office Royal Melbourne Hospital
Parkville, Victoria 3050
Australia

[4]CSIRO Health Sciences and Nutrition
343 Royal Parade
Parkville, Victoria 3052
Australia

Summary

We report the crystal structure, at 2.5 Å resolution, of a truncated human EGFR ectodomain bound to TGFα. TGFα interacts with both L1 and L2 domains of EGFR, making many main chain contacts with L1 and interacting with L2 via key conserved residues. The results indicate how EGFR family members can bind a family of highly variable ligands. In the 2:2 TGFα:sEGFR501 complex, each ligand interacts with only one receptor molecule. There are two types of dimers in the asymmetric unit: a head-to-head dimer involving contacts between the L1 and L2 domains and a back-to-back dimer dominated by interactions between the CR1 domains of each receptor. Based on sequence conservation, buried surface area, and mutagenesis experiments, the back-to-back dimer is favored to be biologically relevant.

Introduction

The biology and biochemistry of the epidermal growth factor receptor (EGFR) system has intrigued scientists and clinicians for more than twenty-five years. Aberrant signaling by the EGF/EGFR family is associated with a number of cancers including brain, head and neck, pancreatic, and colorectal tumors, making it a significant therapeutic target (reviewed in Yarden, 2001). Furthermore, a constitutively active mutant form of the receptor, in which exons 2–7 (coding for residues 6–273) are deleted, has been observed in many glioblastoma, breast, and ovarian cancers (Moscatello et al., 1995).

Four members of the EGFR family of tyrosine kinases have been identified in vertebrates, namely EGFR (HER1/ErbB1), ErbB2 (Neu/HER2), ErbB3 (HER3), and ErbB4 (HER4). They are capable of forming homo- or heterodimers and possibly higher-order oligomers, following activation by a subset of a dozen potential ligands including EGF, transforming growth factor α (TGFα), amphiregulin, betacellulin, epiregulin, heparin binding EGF, epigen, and the neuregulins, giving rise to a diverse signaling network (see Groenen et al., 1994; Yarden and Sliwkowski, 2001). ErbB2 has no known ligand but seems to be the preferred partner for heterodimerization with other EGFR family members (Sundaresan et al., 1998). ErbB3 has an inactive kinase domain, and hence signals through ErbB family heterodimers (Riese and Stern, 1998).

The human EGFR is a large (1186 residues), monomeric, modular glycoprotein with an extracellular ligand binding domain, a transmembrane region, and an intracellular cytoplasmic tyrosine kinase that is flanked by noncatalytic regulatory regions (Ullrich et al., 1984). The extracellular portion of human EGFR (residues 1–621) consists of four subdomains, L1, CR1, L2, and CR2 (Bajaj et al., 1987; Ward et al., 1995), also referred to as domains I–IV (Lax et al., 1988).

Ligand-induced receptor dimers have been proposed as the primary signaling system for EGFR (Lemmon and Schlessinger, 1994; Moriki et al., 2001) with the stoichiometry of the EGFR dimer complex being 2:2 (Lemmon et al., 1997). Two models for the nature of the 2:2 dimer were described which could not be distinguished. The favored model was ligand-mediated dimerization, where each EGF molecule made contact with both sEGFR molecules. The alternative was receptor-mediated dimerization, where each ligand binds to only one sEGFR molecule, inducing a conformational change that promotes receptor-receptor contacts (Lemmon et al., 1997).

Recently we described a truncated, soluble version of the hEGFR extracellular domain, comprising residues 1–501 (sEGFR501), which binds human EGF or TGFα with 13- to 14-fold higher affinity than the full-length EGFR ectodomain and is capable of forming 2:2 dimers in the presence of EGF (Elleman et al., 2001). In this report we describe the use of sEGFR501 to generate crystals of the human TGFα:sEGFR501 dimer complex and determine the three-dimensional structure to 2.5 Å resolution. Previous reports on crystallization of an EGF:EGFR complex have used the EGFR ectodomain secreted from A431 tumor cells. This soluble receptor consists of residues 1–615 fused to an additional 18 amino acids (Ullrich et al., 1984) and gave crystals that diffracted to 10 Å (Gunther et al., 1990), 6 Å (Weber et al., 1994), or 4.5 Å (Degenhardt et al., 1998).

Results and Discussion

Overall Structure

sEGFR501 is comprised of three structural domains, namely L1, CR1, and L2, plus the first module from the second Cys-rich region CR2. Crystals of TGFα:sEGFR501 contain two molecules of each polypeptide in the asymmetric unit. Crystallographic statistics are given in Table

[5]Correspondence: tgarrett@wehi.edu.au (T.P.J.G.), tony.burgess@ludwig.edu.au (A.W.B.), colin.ward@hsn.csiro.au (C.W.W.)

Table 1. Summary of Crystallographic Data

Data Set	Resolution (Å)	Mean I/s	R_{merge}[a]	Completeness (%) (Multiplicity)	Number of Sites	R_{cullis}[b]	Phasing Power[c]	f.o.m.[d]
Native	2.9	11.1	0.129	96.9 (2.78)				0.31/0.84
$Pt(NO_3)_2$	2.8	11.9	0.095	97.8 (3.85)	4	0.71	0.71	
PIP	2.5	10.8	0.075	90.2 (3.17)	2	0.91	0.91	
$K_2Au(CN)_2$	3.0	9.1	0.091	97.8 (3.43)	4	0.21	2.21	
Refinement Resolution (Å)	Number of Reflections (Free)		Number of Atoms	R_{cryst}[e]	R_{free}[e]	Bonds[f] (Å)	Angles[f] (°)	
20–2.5	48,006 (2379)		8,687	0.237	0.289	0.007	1.50	

PIP, di-μ-iodobis(ethylenediamine)diplatinum nitrate

[a] $R_{merge} = \Sigma_h\Sigma_j |I_{hj} - I_h|/\Sigma_h\Sigma_j I_h$, where I_{hj} is an intensity measurement j of I_h and I_h is the mean of a reflection h.

[b] $R_{cullis} = \Sigma_h||F_{PH} - F_P| - |F_{Hcalc}||/\Sigma_h||F_{PH}| - |F_P||$, where F_{PH}, F_P, and F_{Hcalc} are, respectively, derivative, native, and heavy atom structure factors for centric reflection h.

[c] Phasing power $= \Sigma_h|F_{Hcalc}|/\Sigma_h\epsilon$, where F_{Hcalc} is defined above and ϵ is the lack of closure.

[d] f.o.m. (figure of merit) $= <\cos(\Delta\alpha_h)>$, where $\Delta\alpha_h$ is the error in the phase angle for reflection h. Values are given before and after density modification.

[e] R_{cryst} and R_{free} are defined in the CNS manual (Brunger et al., 1998).

[f] Rms deviation for bond distances and angles.

1. There are two possible dimer interactions: a back-to-back dimer dominated by interactions between the CR1 domains of each receptor (Figures 1A and 1B, see Supplemental Figures S1 and S2 at http://www.cell.com/cgi/content/full/110/6/763/DC1) and a head-to-head dimer involving contacts between the L1 and L2 domains (Figures 1C and 1D). The back-to-back complex is approximately 33 × 78 × 103 Å while the head-to-head complex is 65 × 75 × 128 Å. Each TGFα molecule is clamped between the L1 and L2 domains from the same sEGFR501 molecule and makes contact with only one receptor molecule in the dimer. In the back-to-back dimer, the two ligands are located on opposite sides of the complex with the closest approach 70.9 Å apart (Figure 1A). In the head-to-head dimer, the two ligands are centrally located and are separated by 15 Å (Figure 1D).

We conclude that the back-to-back dimer corresponds to the 2:2 TGFα:sEGFR501 complex that is formed in solution (Elleman et al., 2001) from compari-

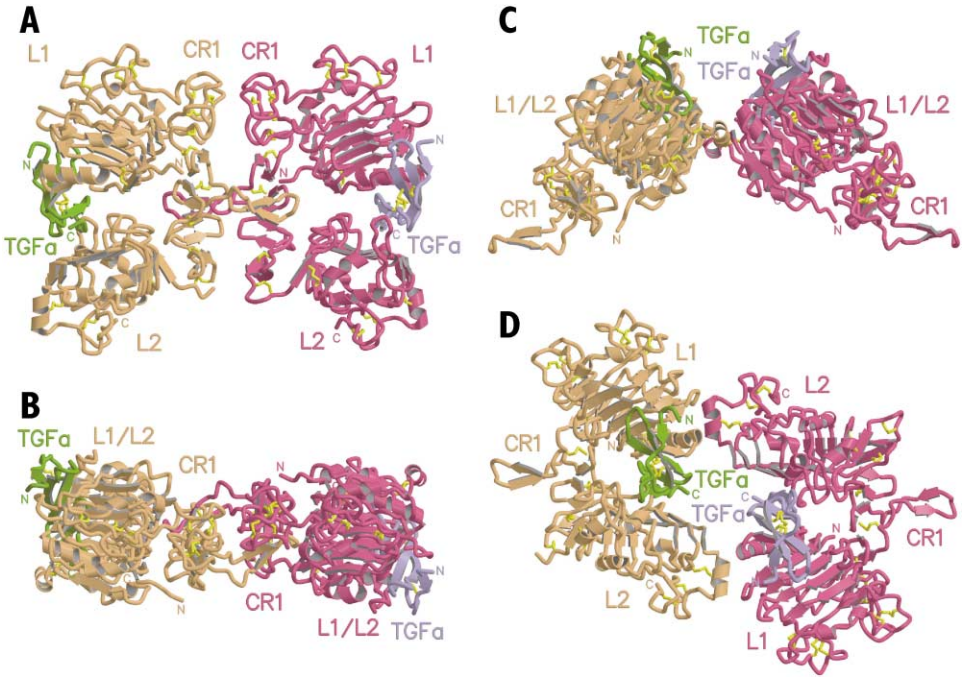

Figure 1. Polypeptide Trace for the 2:2 TGFα:sEGFR501 Complex

(A) Side view of the back-to-back dimer. The sEGFR501 molecules are shown in orange (molecule A) and magenta (molecule B). The two TGFα molecules are colored green and lilac. Disulfide bonds are drawn in yellow.

(B) The back-to-back dimer viewed down the dimer axis.

(C) Side view of the head-to-head dimer.

(D) The head-to-head dimer viewed down the dimer axis.

sons of the amount of buried surface area in the two dimer options, the lack of symmetry in the head-to-head dimer compared to that seen in the back-to-back dimer, the sequence conservation at the dimer interfaces (described later), and the characteristics of the receptors mutated at both interfaces (described later). In the head-to-head dimer, only 510 Å^2 of accessible surface area is buried on each molecule, and this is distributed over two patches 39 Å apart. The residues involved are 21, 24, 25, 28, and 48–51 on both L1s, 471, 473, 474, 476, and 477 on both L2s, plus 32 (molecule A) and 443 and 478 from molecule B. In contrast, in the back-to-back dimer, 1125 Å^2 on each receptor is buried (see later for details). Biologically relevant protein-protein interfaces usually bury more than 700 Å^2 of surface per molecule and often about 1000 Å^2 (Lo Conte et al., 1999), implying that the back-to-back configuration is more likely to be the functional dimer. There is a lack of symmetry at the two L1-L2′ interfaces in the head-to-head dimer that corresponds to a 6 Å translation of the L2′ helix (residues 471–479) relative to L1 (Figure 1D). Such structural ambiguity is not seen in the back-to-back dimer, the noncrystallographic symmetry being very close to a pure 2-fold rotation, implying that this is the functional dimer. It is further supported by experiments where a model of the EGF receptor CR2 domain (Jorissen et al., 2000) was superimposed onto the structure determined here for the first modules of the CR2 domains of the two sEGFR501 molecules. In the back-to-back dimer, the rod-like domains of CR2 project toward each other underneath sEGFR501, consistent with the ability to form disulfide-linked dimers via a Cys mutation three residues upstream of the transmembrane domain when ligand binds to mutant receptors (Sorokin et al., 1994). The same superimposition performed on the head-to-head dimer results in the modeled CR2 domains projecting away from each other and is inconsistent with the Cys mutant data (Sorokin et al., 1994). Finally, the back-to-back dimer has now been seen in another crystal form of the EGFR complex (Ogiso et al., 2002 [this issue of *Cell*]).

Receptor Domain Architecture

The L1, CR1, and L2 domains of EGFR show both structural (Figure 2) and sequence (Figure 3) homology to the first three domains of the type I insulin-like growth factor receptor (Garrett et al., 1998), with the L domains resembling other leucine-rich repeat or solenoid proteins (Ward and Garrett, 2001; Kobe and Kajava, 2001). Like the IGF-1R, the EGFR L domains (1) comprise six turns of a β helix capped at each end by a helix and a disulfide bond; (2) have a conserved Trp (Trp176 in CR1 and Trp492 in CR2) inserted between the fourth and fifth turns of the β helix; (3) contain a large β sheet (second sheet, green in Figure 2) flanked by two shorter ones on either side (blue and yellow) with the face opposite the large β sheet being more irregular; (4) have a stack of conserved Gly residues at positions 39, 63, 85, and 122 in L1 and 343, 379, 404, and 435 in L2 (Figure 3A) at the edge before the second β sheet and a short Asn ladder at the edge after it; and (5) have a loop from the fourth turn of each solenoid that protrudes from the large (second) β sheet.

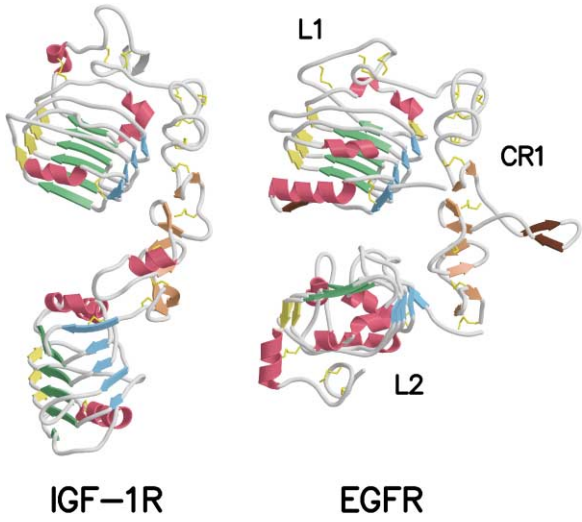

Figure 2. Comparison of sEGFR501 with the First Three Domains of IGF-1R

For clarity the ligand in the TGFα:sEGFR501 complex is not shown. L1 domains are oriented similarly. Helices are indicated by curled, red ribbons and β strands by broad arrows. The blue, green, and yellow β strands depict the three prominent parallel β sheets within the L1 and L2 domains. The β strands in the Cys-rich domains are colored orange, except for those in the CR1 loop, which are colored brown. The black strand forms part of a β sheet with the ligand. The side chains of disulfide-linked cysteine residues are depicted as yellow sticks.

For both L1 and L2 domains of EGFR, the long β strand in the first turn of the solenoid is missing. In L1 it is replaced by a long V-shaped excursion (residues 8–18) that sits over the large β sheet of L1 to form a major part of L1's ligand binding surface (Figure 2). In L2 it is replaced by a loop (residues 316–326) that also contacts the ligand (see Ligand-Receptor Interactions). These two loops correspond to major insertions in the first turn of each of the L1 and L2 domains of EGFR relative to IGF-1R (Figure 3A). A third insertion is seen as an extra loop in the second turn of L2 (residues 351–369), which corresponds to the epitope for LA22, LA58, and LA90, antibodies that prevent ligand binding (Wu et al., 1989). These three loops, together with conserved loops in the fourth turn of the solenoid, all participate in important interactions with the ligand.

The arrangement of the eight disulfide-bonded modules in CR1 is similar to that of IGF-1R (Figures 3A and 3B), although the slightly different orientations result in CR1 of the EGFR being a straight rod, bent at module 6, compared to the curved CR domain of IGF-1R (Figure 2). The CR1 domain of EGFR makes contact with L1 along one side of the solenoid (sheet 1, burying 1375 Å^2 of accessible surface area) and also makes appreciable contact with the L2 domain via modules 6 and 7 (burying 860 Å^2). This is different from the IGF-1R structure, where the L2 domain is rotated away to lie almost perpendicular to the axis of L1 (Figure 2). Thus, the C-terminal region of CR1 may act as a hinge in the ligand-free form of the EGFR as modules 7 and 8 appear somewhat mobile, having some of the largest temperature factors in the structure.

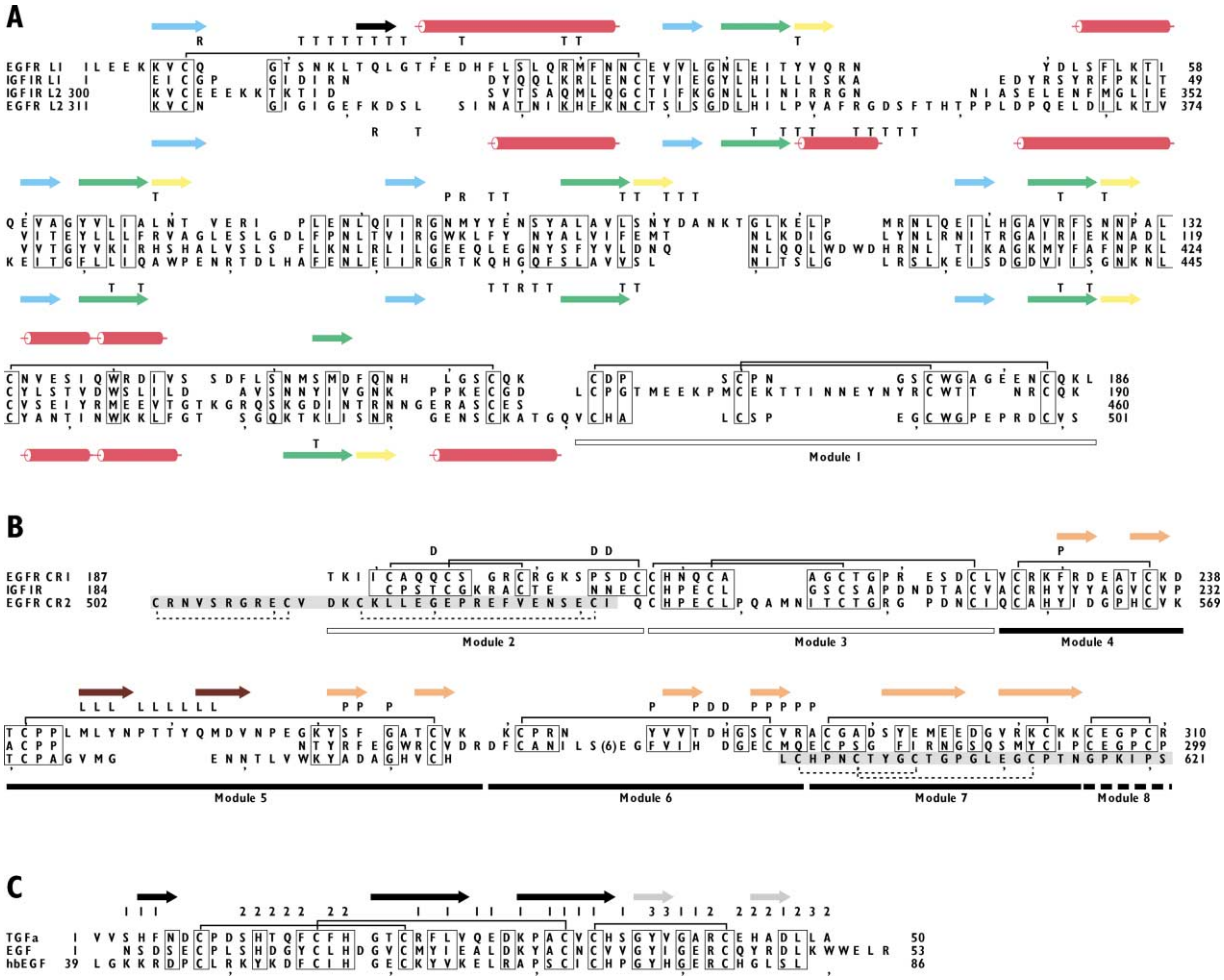

Figure 3. Structure-Based Sequence Alignment of the Human EGFR Ectodomain, Human TGFα, and Related Proteins

(A) The receptor L1 and L2 domains plus the first module of the Cys-rich regions, CR1 and CR2. Positions with conserved physicochemical properties of amino acids are boxed. Disulfide bond connections are shown as solid lines. Secondary structure elements are indicated above and below the sequences as cylinders for α helices and arrows for β strands (color-coding is the same as Figure 2). Residues buried at different protein-protein interfaces are indicated by: T at the TGFα, L1, and L2 interfaces; R at the L1-L2 contacts; and P at the CR1 loop interface.

(B) Modules 2–8 of the receptor Cys-rich region CR1 and modules 2–7 of CR2. Three types of disulfide bonded modules are indicated by bars below the sequences, and residues not conforming to the CR1 pattern are shaded gray. The unfilled bars below parts of the Cys-rich sequences indicate modules with two disulfide bonds (in a Cys 1-3 and 2-4 arrangement), the solid bars indicate modules that contain a single disulfide bond and have a β-finger motif, and the dashed bar indicates residues present in a disulfide-linked bend consisting of only five residues. Disulfide bonds are shown in solid lines, except for those that do not conform to the CR1 pattern, which are indicated as dashed lines. The number in parentheses shows where amino acids have been omitted. Boxed residues and secondary structure elements are as in (A). Residues buried at protein-protein interfaces are indicated as follows: L, CR1 loop residues that are buried; P, residues to which the CR1 loop binds; D, other residues in the dimer interface.

(C) Human TGFα, EGF, and heparin binding EGF. The numbers 1, 2, and 3 indicate ligand residues contacting L1, L2, or both L1 and L2, respectively. The positions of the two sets of β strands in the ligand are indicated by black and gray arrows above the sequences. One of these sets of β strands (black) forms a β sheet with the black β strand of the receptor (see A). Disulfide bond connections are shown as solid lines. The apostrophes above and below the sequences indicate intervals of ten residues. This figure was prepared using ALSCRIPT.

The most striking feature of CR1 is a large ordered loop from module 5 that projects directly away from the ligand binding site. The loop consists of residues 242–259 and contains an antiparallel β-ribbon (Figure 2). This loop is highly conserved within the EGFR family and is not present in IGF-1R (Figure 3B) or other members of the insulin receptor family. It is a major contributor to the receptor-receptor contacts in the back-to-back dimer (Figures 1A and 1B).

Structure of TGFα

More than ten mitogenic peptides form a family of ligands that can bind to members of the EGFR family. However, apart from residues Gly19, Gly40, and the three conserved disulfide bonds that are needed to maintain structure, only Arg42 is conserved throughout the family, and pairwise sequence identities between the ligands are often less than 35%. Three-dimensional structures have been determined by NMR for EGF (Mon-

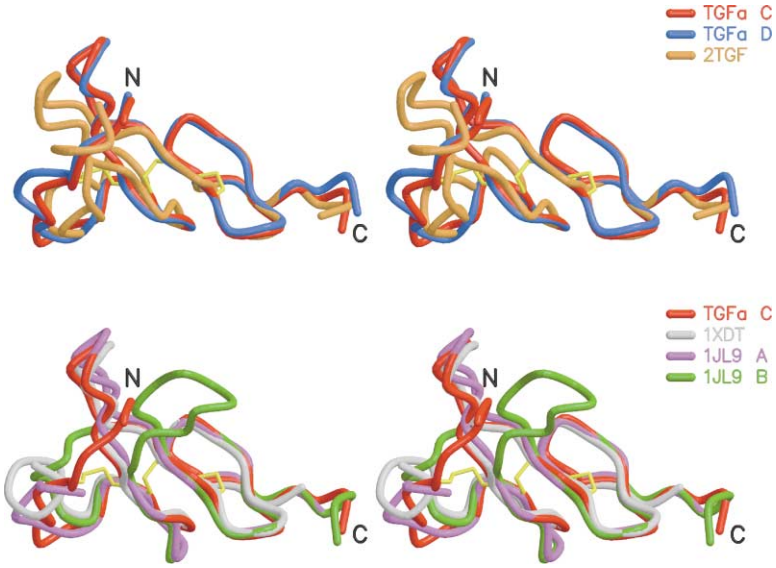

Figure 4. Stereo View of Ligand Cα Traces

Ligands were superimposed using the Cα atoms of core residues (see text). The Cα traces have been rotated 55° about a horizontal axis relative to Figure 5 and are as follows: TGFα molecules C (red with yellow disulfide bonds) and D (blue) from this study; minimized average NMR structure of TGFα (orange) from PDB:2TGF (Harvey et al., 1991); diphtheria toxin bound heparin binding EGF (gray) from PDB:1XDT (Louie et al., 1997); and crystallized human EGF molecule A (purple) and molecule B (green) from PDB:1JL9 (Lu et al., 2001). While the two TGFα molecules from this study (C and D), EGF molecule A of 1JL9, and hbEGF superimpose well from residue 14, the top half of the B loop adopts different conformations in the NMR structure (2TGF) and 1JL9:B (top of each panel). In the uncomplexed form, the N terminus of TGFα is structurally heterogeneous (bottom left of each panel).

telione et al., 1987; Cooke et al., 1987; Kohda and Inagaki, 1992; Barnham et al., 1998), TGFα (Tappin et al., 1989; Harvey et al., 1991; Moy et al., 1993), and heregulin (Nagata et al., 1994; Jacobsen et al., 1996) and by X-ray crystallography for heparin binding EGF (hbEGF) in complex with diphtheria toxin (Louie et al., 1997) and EGF (Lu et al., 2001). These structures show that TGFα and its relatives are relatively flexible molecules built on a small structurally conserved core (Figure 4). In particular, the 14 N-terminal residues, the "B loop," and the extreme C terminus are often quite disordered. From a comparison of the two molecules of EGF in the asymmetric unit, Lu et al. (2001) found that the common structural core comprised only residues 13–21 and 30–47 (equivalent to 15–22 and 31–48 in TGFα; Figure 3C), which encompassed half of the large β-ribbon and a small, C-terminal β-ribbon. The structure of TGFα, seen here in the complex, shows substantially more order, with a third, N-terminal β strand (residues 4–6) aligned with the large β-ribbon (residues 19–33) to form a three-stranded β sheet and an ordered C terminus. The structure of TGFα in the 2:2 complex is triangular or crescent-shaped. The two TGFα molecules in the dimer superimpose well on each other (rmsd 0.70 Å for 44 Cα atoms; Figure 4). They are structurally similar to the human EGF molecule A (rmsd 1.33 Å for 41 Cα atoms) in the EGF crystal structure (Lu et al., 2001) and even more closely to hbEGF (0.66 Å for 34 Cα atoms) in its complex with diphtheria toxin (Louie et al., 1997).

Ligand-Receptor Interactions

In the complex, TGFα interacts with the large β sheets of both the L1 and L2 domains of one receptor molecule (Figures 1 and 5). Relative to IGF-1R, the position of L2 corresponds to a rotation by 105° at the L2/CR1 module7 interface or 122°–130°, relative to L1 of IGF-1R. More than a third of the ligand's accessible surface area is buried by the L1 and L2 domains of the receptor (about 745 Å^2 by L1 and about 785 Å^2 by L2), and over 60% of the ligand's residues make contact with the receptor. The footprint of the ligand on the receptor covers most of the large (second) sheet of each L domain, running from the top left corner to abut the loop in the fourth rung of the solenoid (Figures 1 and 5). The involvement of both the L1 and L2 domains in ligand binding is consistent with the results of chicken-human EGF receptor domain swapping experiments (Lax et al., 1989). The buried surface area on each L domain is similar to that found in antibody-antigen interactions and suggests that each L domain could possibly bind a ligand on its own. Indeed, this has been shown for the L2 domain, where a proteolytic fragment was capable of binding EGF (Kohda et al., 1993), and for the L1-CR1 fragment of ErbB3, which bound heregulin (Singer et al., 2001).

In the contact with L1, the inner curved face of the crescent-shaped TGFα sits across the large sheet and extends to the N-terminal helix of L1 (Figure 5). More than half the buried surface area of L1 comes from a V-shaped loop (residues 8–18) that runs across the large sheet, replacing the first strand of the corresponding sheet in IGF-1R. In the center of this interface, TGFα makes contact with the receptor, primarily via main chain atoms. One strand from the large β sheet of TGFα (residues 29–35) sits edge on to the receptor and aligns with the latter part of the V-shaped loop (residues 15–17). This enables the receptor to contribute part of the V as a fourth parallel β strand to the first and larger of the ligand's two β sheets (Figure 5A). Asn12, which is conserved in all of the EGFR family except ErbB2, makes a side chain to main chain contact with the peptide N atom of Gly40 in TGFα, and the Oγ1 atom of Thr15 from L1 makes a hydrogen bond to Ala41 O of TGFα (Figure 5B). This interface is also characterized by two hydrophobic contacts around Leu14 and Leu17 from L1 and hydrophilic and electrostatic interactions involving His4 and the B loop residues Arg22, Gln26, Glu27, and Lys29 of TGFα with the L1 domain residues Tyr45, Tyr89, Glu90, Tyr101, Arg125, and Asn128 (Figure 5). The location of the N terminus of TGFα near Tyr101 in the complex is consistent with the chemical crosslinking data

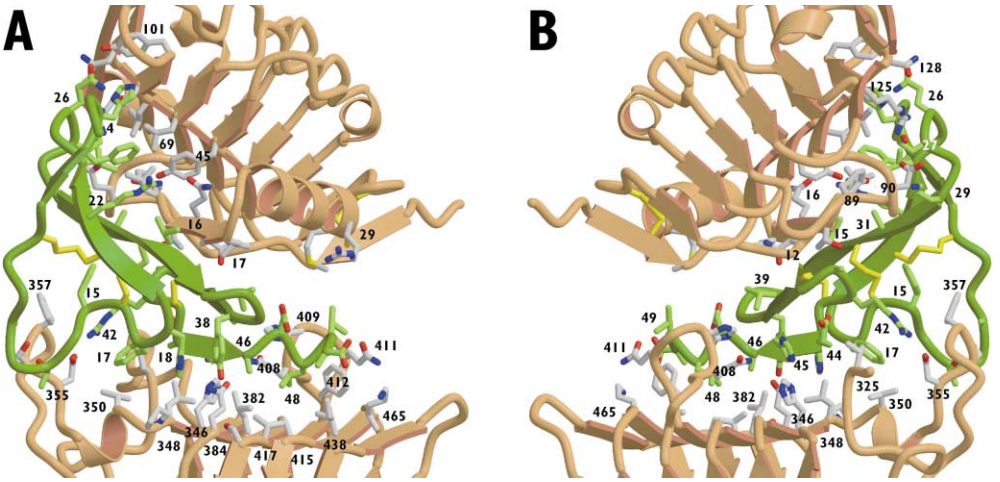

Figure 5. Structure of the Ligand:Receptor Binding Surfaces

(A) Ribbon representation showing the contacts between sEGFR501 (orange with gray side chains) and TGFα (green) viewed from the left in Figure 1A. Residue numbers for TGFα are italicized. O and N atoms are shown in red and blue, respectively.

(B) As in (A) but rotated 180°.

of Woltjer et al. (1992). It should be noted that the lack of conservation in ErbB2 of two key residues in this interface (Arg for Thr/Ser at position 15 and Met for Asn at position 12) would prevent any of the EGF family of ligands from binding to L1 due to steric hindrance and a loss of a hydrogen bond.

The interface between L2 and TGFα is formed mostly from the side chain atoms of both the ligand and receptor. TGFα sits on the flat face (i.e., the large β sheet) of L2, surrounded by three loops (residues 316–326, 352–363, and 405–412) that project out from the plane of the sheet (Figures 5A and 5B). The contact between the ligand and receptor is an alternating series of stripes of hydrophobic and hydrophilic interaction across the interface. These are as follows: (1) Phe15 of TGFα sits against Phe357 of EGFR; (2) the strictly conserved Arg42 of TGFα is sandwiched between Phe15 and Phe17 of the ligand, facilitating the correct orientation and environment to make a salt bridge with the strictly conserved Asp355 of the receptor; (3) Phe17 and the lower part of Glu44 from TGFα interact with Leu325, Leu348, and Val350 from L2; (4) the next hydrophilic region contains four histidines, His18 and His45 of TGFα and His346 and His409 of L2, as well as Tyr38 and Glu44 from TGFα and Gln384 and Gln408 from L2; and (5) there is a hydrophobic pocket in L2 (Leu382, Gln408, His409, Phe412, Val417, Ile438) centered over Ala415, which holds the highly conserved Leu48 of TGFα (Leu47 in EGF), the ligand residue with the largest buried surface. The C terminus of TGFα is sandwiched between domains L1 and L2, with the side chain of Leu49 contacting both L domains. Leu49 may well define the final positioning of the L domains in the complex. Lys465 from L2 is near the C terminus of TGFα and may stabilize the terminal carboxyl group. Lys465 has been chemically crosslinked to residue 45 in a mutant form of mouse EGF (Summerfield et al., 1996). Some carbohydrate nearby could possibly also affect ligand binding.

Data from chemical crosslinking (Woltjer et al., 1992; Summerfield et al., 1996) and site-directed mutagenesis of the ligands (reviewed by Groenen et al., 1994) are consistent with the structure of the complex and highlight the critical importance of Arg42 and Leu48. Binding to the receptor does restrict the choice of ligand side chain in some positions. Residue 24 (TGFα) is always aliphatic (Leu, Ile, or Val) and residue 31 is a small amino acid such as Ala, Ser, or Thr. Both are buried in the L1 interface. The conservation of Arg42 and Leu48 had been observed, but conservation of an aromatic residue at 15 and a hydrophobic residue at 17 are necessary both for ligand binding and for holding Arg42 in the correct orientation. The cluster of His residues in the middle of the L2 interface may play a part in release of the ligand at low pH following endocytosis.

Receptor-Receptor Interactions

Unlike other growth factor receptor complexes, the ligand is not found at the dimer interface in the 2:2 complex of TGFα:sEGFR501. Thus, ligand-induced dimerization of sEGFR501 implies that binding of ligand induces a conformational change in the receptor that promotes receptor-receptor interactions. The most notable feature of the back-to-back dimer is a long loop (residues 242–259) that is specific to the EGFR family and is not found in the CR of IGF-1R (Figures 2 and 3B) or other members of the insulin receptor family. From each receptor, the loop projects out from the fifth module of CR1, across the other CR1 domain to a space between L1, L2, and CR1 domains of the neighboring receptor (Figures 1A and 1B). Contact is made by residues 244–253 of the CR1 loop in, say, molecule A with residues 229–239, 262–278, and 282–288 on the concave face of the CR1 domain of molecule B (Figure 6). The buried surface areas are 480 Å^2 and 330 Å^2, respectively. At specific positions in the CR1 loop, there is remarkable sequence conservation across all ErbB family members. Tyr246 is strictly conserved and is completely buried in the interface. The Oη atom of TyrA246 (receptor molecule A) makes hydrogen bonds with the GlyB264 N and CysB283 O atoms (receptor molecule B), and the phenyl

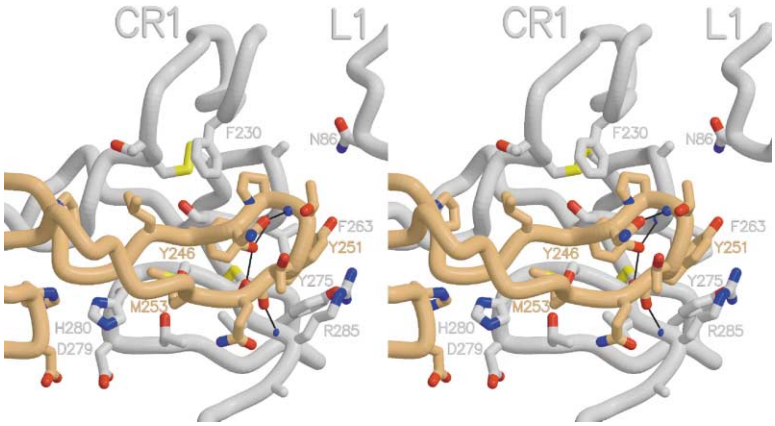

Figure 6. Stereo View of the Contacts between the CR1 Loop of EGFR Molecule A with CR1 of EGFR Molecule B in the Back-to-Back Dimer Interface

CR1 loop of EGFR molecule A is shown in orange, and CR1 of EGFR molecule B is in gray. Interchain hydrogen bonds are drawn in black along with the hydrogen bond from Asn247 of molecule A (not labeled), which appears to stabilize the loop tip conformation. The single letter code and residue number is used for amino acid residues. The dimer axis lies vertically at the left at H280.

ring sits against the Cβ atoms of SerB262 and SerB282 and the face of the following peptides (Figure 6). Residue 251 is strictly conserved as Tyr or Phe and in this interface makes a hydrophobic contact via the benzene ring with the PheB263, GlyB264, TyrB275, and ArgB285. The Oη of TyrA251 is exposed to solvent. Additional hydrophobic contacts are made by ProA248 to PheB230 and AlaB265 and by MetA253 to ThrB278. There is also a hydrogen bond from TyrA251 O to ArgB285 N (Figure 6).

Other conserved residues of the CR1 loop, such as Asn247 and Asn256, do not make contact with the other half of the dimer, but hydrogen bond back onto the main chain and appear to be important for maintaining the loop in the appropriate conformation. There are four positions in the loop (residues 243, 248, 255, and 257) where proline is found in at least one member of the human EGFR family with ErbB3 having as many as three prolines. These prolines would further stabilize the conformation of the loop.

The loop not only touches the CR1 domain of its partner, but also reaches across to contact the L1 and L2 domains of the other receptor molecule (burying a surface area of 40 Å^2 on L1 and 5 Å^2 on L2). AsnB86 touches ThrA249 and, with a slight rearrangement, could form a hydrogen bond between the side chains. Neither residue is conserved in other ErbB receptors, although polar residues predominate at these positions. ThrA250, which is conserved in other ErbB receptors, sits near IleB318 but the reason for the conservation is not apparent. Although these interactions are quite weak, it is possible that the binding of the loop from one receptor may be affected by binding of ligand to the other, as ligand binding may alter the relative positions of the L domains.

Two other regions also participate in the back-to-back dimer contact. One is near the two long loops, where Asp279 and His280 of receptor A make contact across the dimer axis with the corresponding residues from receptor B (Figures 1A and 1B). A second region of contact is near the N-terminal end of the CR1 domain in Cys-rich module 2, where residues 193–195 and 204–205 from molecule A contact 193–194 and 204–205 from molecule B, burying about 225 Å^2.

Functional Characterization of Mutant EGFRs

In order to establish the biological relevance of the two dimers identified in crystals of the TGFα:sEGFR501 complex, mutant receptors designed to probe the two dimer interfaces were analyzed. Single amino acid substitutions Glu21Ala, Arg470Leu, Asn473Asp, Ser474Glu, and Ala477Asp were prepared to test the head-to-head dimer. When transiently expressed in 293 cells, which express low endogenous levels of EGFR ($<1 \times 10^4$ receptors/cell), or when stably expressed (Glu21Ala) in the hemopoietic cell line BaF/3, which does not express EGFR family members (Walker et al., 1998), these mutants showed normal EGF binding, kinase activation, dimerization (Figure 7), and internalization (data not shown). In contrast, mutants of the back-to-back dimer, a CR1 loop deletion (residues Δ242–259) from the full-length receptor and sEGFR501 with multiple substitutions in the CR1 loop (Tyr246Asp, Asn247Ala, Thr249Asp, Tyr251Glu, Gln252Ala, and Met253Asp), were defective. The ΔCR1-loop clones fail to show ligand-induced dimerization and ligand-induced receptor kinase activation and exhibit only low-affinity binding (Figures 7A–7C). Anti-phosphotyrosine Western blotting (Figure 7C) demonstrates the absence of ligand-induced receptor autophosphorylation in the ΔCR1-loop clones. Furthermore, ligand-induced tyrosine phosphorylation of other substrates and activation of MAPK (as determined by immunoblotting with antibodies that recognize phospho-MAPK, data not shown) are missing in the ΔCR1-loop clones. Similarly, the sEGFR501 mutant with multiple CR1 loop substitutions fails to show ligand-induced dimerization (Figure 7D) and exhibits 15-fold lower affinity binding for EGF on BIAcore analysis (500 nM versus 30 nM for sEGFR501).

Conclusion

Ligand-induced dimerization (or oligomerization) of receptors is a common means of signal transduction, and in all cases seen so far the ligand participates directly in the dimerization of receptors. For VEGF/Flt-1 (Wiesmann et al., 1997), nerve growth factor (NGF)/TrkA receptor (Wiesmann et al., 1999), bone morphogenic protein (BMP)/BMP receptor (Kirsch et al., 2000), interferon γ (IFNγ)/IFNγ receptor (Thiel et al., 2000), and tumor necrosis factor (TNF)/TNF receptor (Banner et al., 1993), the ligand is a dimer or trimer before forming the 2:2 complex or 3:3 complex and, in the structures determined, the receptors do not contact each other. In the 2:2 complex of the fibroblast growth factor (FGF)/FGF

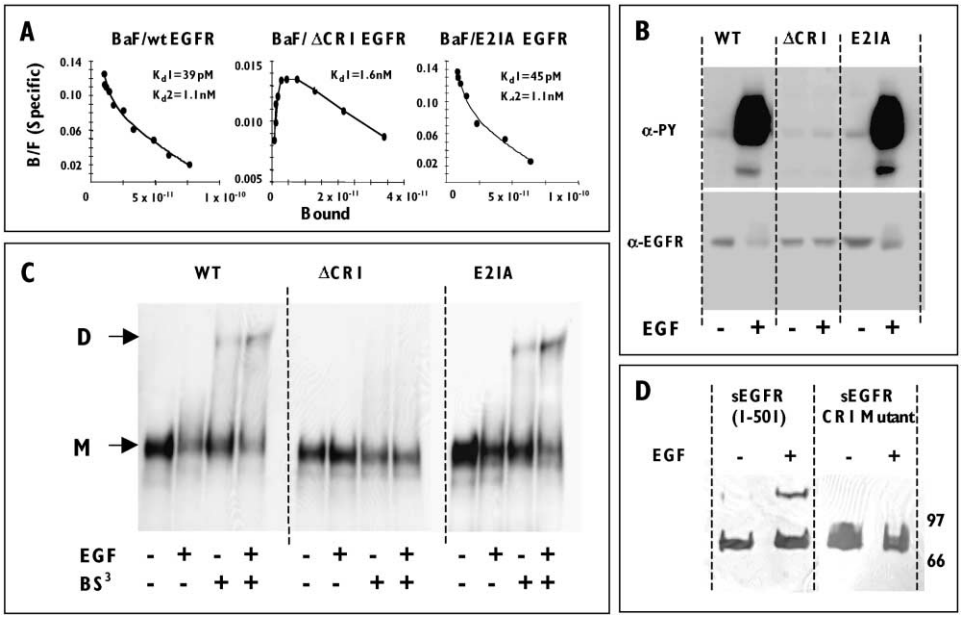

Figure 7. Functional Characterization of EGFR Mutants Expressed in BaF/3 Cells

(A) Ligand binding by wild-type and mutant EGFRs expressed in BaF/3 cells. Scatchard plots of ^{125}I-EGF binding to clones expressing the wt, E21A, or ΔCR1 EGFR were analyzed using the Radlig program to yield estimates of receptor affinity. The three cell lines expressed comparable receptor numbers as assessed by M2 or 528 antibody binding and FACS analysis. Shown are the plots for cold ligand titration assay; identical results were obtained titrating the radiolabeled EGF.

(B) EGF-dependent tyrosine kinase activation. This was determined in total cell lysates by sequential immunoblotting with anti-phosphotyrosine (top) or anti-EGFR (bottom) antibodies. The anti-EGFR antibodies have slightly lower affinity for the hyperphosphorylated form of the EGFR. The results are representative of multiple experiments on at least four independently derived clones for each mutant.

(C) Ligand-induced EGFR dimerization. Crosslinking of the EGFR via the extracellular portion was performed using BS^3 at 37°C to maximize dimer yield. Samples were analyzed by SDS-PAGE on 3%–8% gradient gels and immunoblotting with anti-EGFR antibodies. These data are representative of at least four separate experiments.

(D) Ligand-induced sEGFR501 dimerization. Crosslinking of wild-type and CR1 loop mutant (Tyr246Asp, Asn247Ala, Thr249Asp, Tyr251Glu, Gln252Ala, and Met253Asp) was carried out as described previously (Elleman et al., 2001).

receptor, the ligands do not contact each other but are dimerized by heparin (Plotnikov et al., 2000; Schlessinger et al., 2000; Pellegrini et al., 2000). The FGF receptors do contact each other and the two FGF ligands lie at the dimer interface with a heparin molecule sitting between two FGFs. In the 2:2 complex of granulocyte colony-stimulating factor (GCSF)/GCSF receptor (Aritomi et al., 1999), each ligand binds both receptors but there are no contacts between the two ligands or the two receptor fragments. Finally, in the growth hormone, erythropoietin, and prolactin/receptor complexes, there is only one ligand molecule in the 1:2 complex, and the two receptor molecules make contact with ligand and with each other (de Vos et al., 1992).

The back-to-back TGFα:EGFR complex described here and the EGF:EGFR complex reported by Ogiso et al. (2002 [this issue of *Cell*]) represent new and surprising ways in which receptors and protein ligands interact. EGFR ligands bind at a site remote from the dimer interface to promote and/or stabilize receptor dimerization. A precedent for this has been seen in the rat metabotrophic glutamate receptor where the small ligand, glutamate, binds between two domains of the receptor monomer, causing them to go from an "open" to a "closed" form (Kunishima et al., 2000). Such a mechanism could also occur in the EGFR family where the ligand binds both L1 and L2, fixing the relative orientations of the two domains. Compared to IGF-1R, there is a substantial rearrangement of L domains in EGFR (Figure 2), although a conformational change of such a magnitude would not be necessary. A smaller change in L domain positions upon ligand binding, possibly with hinge motions seen at the CR1 module 5/6, 6/7, and 7/L2 interfaces (relative to IGF-1R), could enable EGFR extracellular domains to form dimers.

Clearly, the structure presented here contributes only part of the information required to improve our understanding of EGF receptor signaling at the cell surface. Truncated forms of the receptor, such as delta 2-7 or with the extracellular region deleted, are constitutively active, indicating that determinants also exist in the transmembrane region and/or the intracellular domain that facilitate receptor association and activation. Recently unligated EGFR dimers have been detected on cell surfaces (Moriki et al., 2001). Ligand binding to these preformed dimers may induce a reorientation of the extracellular domains with a consequential reorientation (and activation) of the kinase domains.

Experimental Procedures

Protein Preparation of sEGFR501

The derivation of stably transfected Lec8 cells expressing sEGFR501 and the subsequent purification and characterization of the secreted ectodomain has been described in detail (Elleman et

al., 2001). Purified sEGFR501 was shown by isoelectric focusing gels to be unstable on storage, the majority of isoforms being transformed into products with less acidic isoelectric points. This change is accompanied by a small mobility increase (estimated at 1–2 kDa) on SDS polyacrylamide gels. Sequence analysis showed that the new product had no N-terminal degradation, suggesting partial or complete loss of the acidic residue-rich C-terminal tag and enterokinase cleavage site. Prolonged storage led to the majority of protein converting to the least acidic isoform of pI ~6.6. The conversion of a fresh preparation of sEGFR501 to a stable, less acidic isoform was more reproducible and rapid if it was subject to limited proteolysis at ambient temperature in Tris-buffered saline (pH 8) for ~180 min with endoproteinase Asp-N (Boehringer-Mannheim) at an enzyme:protein ratio of 1:1000 (w/w). The least-acidic isoform of apparent pI ~6.2 was isolated from the other components by anion exchange chromatography. The digest was bound to three Uno Q2 columns (BioRad) connected in series to a BioLogic HR liquid chromatography instrument in 20 mM ethanolamine/50 mM taurine (pH 8.0) buffer, and the least acidic form was the first product obtained by isocratic elution in the same buffer containing 15 mM lithium acetate. The purified protein was incubated with endoglycosidase F (PNGase-free; Boehringer Mannheim) at a ratio of 10–20 Units/mg protein, followed by rechromatography over Superdex 200 to remove enzyme and low-molecular-weight cleavage products.

Crystallization and Data Collection

sEGFR501 obtained from the above procedures appeared nearly homogeneous on SDS and IEF gels and was used in crystallization trials alone and in combination with several ligands. The best-diffracting crystals were obtained from mixtures containing a 5-fold molar quantity of human TGFα (GroPep receptor grade) compared to sEGFR501. Crystals of sEGFR501 in complex with TGFα were grown in 7% PEG 3350, 20% Trehalose, 10 mM $CdCl_2$, and 100 mM HEPES (pH 7.5) mother liquor and belonged to the space group $P2_1$ ($a = 51.59$, $b = 198.71$, $c = 78.90$ Å, $\beta = 102.03°$). These crystals were cryo-cooled to −170°C in the same mother liquor. Data were recorded on a Rigaku RAXIS VI area detector using a Siemens M18XHF X-ray generator with Yale/MSC mirrors or a Rigaku RU300 generator and AXCO capillary optics. Crystals were also derivatized by soaking in mother liquors containing 1–10 mM heavy atom compounds, and diffraction data were collected as before and statistics are given in Table 1. The resolution limit was defined as where $I/\sigma = 2$ for 50% of the reflections. Notable anisotropy was observed for the diffraction limit of the crystals and in the mosaic spread of diffraction maxima.

Phase Determination and Structure Refinement

Phasing by multiple isomorphic replacement was performed with programs from CCP4 (Collaborative Computational Project Number 4, 1994) and SHARP (De La Fortelle and Bricogne, 1996), and the resulting electron density maps were improved by solvent flattening and histogram matching with DM (Cowtan, 1994). Details are given in Table 1. Density averaging using noncrystallographic symmetry was not of much value as the proteins corresponded to more than three rigid groups. The polypeptide chains for two receptor and two ligand molecules were fitted manually and refined with CNS (Brunger et al., 1998). As the highest-resolution data were collected for the PIP derivative, these data were used for the final stages of refinement. During the refinement, an overall anisotropic temperature factor was applied, with the magnitude of the semiaxes being −18.4, 5.6, and 12.7 $Å^2$. The refined structure contains 1097 amino acids, 14 carbohydrate residues, 7 Pt^{2+}, 11 Cd^{2+}, and 4 Cl^- ions, and 79 water molecules. Clear connected density was present for each polypeptide chain although density was poorer for residues 148–160 and 289–307 in each receptor. No density was found for ligand residues C1 and D1-D2 or for receptor residues A306 and beyond residues A500 and B501.

Construction of N-Terminal Tagged EGF Receptor and Mutants

The polymerase chain reaction (PCR) using a human EGFR cDNA (accession number x00588) (Ullrich et al., 1984) was used to generate EGFR expression constructs containing the EGFR leader sequence (and a small portion of 5′ noncoding sequence, base pair 131–261), followed by the FLAG coding sequences with Hind III and Xho I on its 5′ and 3′ ends, respectively, cloned into a mammalian expression vector pcDNA3 (Invitrogen). This yielded the wild-type N-terminal tagged EGF receptor construct, M2-EGFR. PCR products containing point mutations and CR1-loop deletion were cloned using the wild-type M2-EGFR as a template. The point mutation constructs are E21A, R470L, N473D, S474E, and A477D. The CR1-loop deletion construct contains a replacement of nucleotides 988–1035 by GCC, resulting in CR1-loop residues 244–259 being replaced by a single alanine residue. The sEGFR501 CR1 loop mutant (Tyr246Asp, Asn247Ala, Thr249Asp, Tyr251Glu, Gln252Ala, and Met253Asp) was generated by oligonucleotide-directed in vitro mutagenesis using the USB-T7 Gen kit, transiently expressed, purified, and characterized as described previously (Elleman et al., 2001).

Functional Analysis of EGFR Mutants

NIH3T3 and/or 293 cells (American Type Culture Collection) were grown and transfected as described previously (Walker et al., 1998). Two days after transfection, cells were washed with serum-free medium, starved for 2 hr, and treated with or without EGF (100 ng/ml) for 10 min. Whole-cell lysates were prepared and fractionated by SDS-gel electrophoresis using 4%–20% polyacrylamide gels and Western blots developed using monoclonal antibodies M2 (anti-FLAG, Sigma, Brizzard et al., 1994) or 4G10 (anti-phosphotyrosine, Upstate Biotechnology) as described (Walker et al., 1998).

Characterization of Wild-Type and Mutant EGFR Stably Expressed in BaF/3 Cells.

Expression vectors containing the mutant EGFR constructs were transfected into the IL-3-dependent murine hemopoietic lineage BaF/3, selected, and analyzed as described previously (Walker et al., 1998). Receptor dimerization was mentioned by crosslinking with the water-soluble homobifunctional crosslinker BS^3 (Pierce) at a final concentration of 1.2 mM. Scatchard plots and estimates of affinities and receptor numbers were obtained using the Radlig program (Kell for Windows, BioSoft). Ligand-induced receptor kinase activation was analyzed by Western blotting the cell lysates with the anti-phosphotyrosine monoclonal antibody 4G10 (Upstate Biotechnology).

Acknowledgments

We wish to thank our many friends and colleagues, and in particular Peter Colman, who over the years have made valuable contributions to our attempts to determine the structure of the EGF receptor. In no small way, this work was inspired by the discoveries, insights, and persistence of Stanley Cohen and his colleagues. This work was supported in part by the Australian and Victorian State Governments through CSIRO and the Biomolecular Research Institute and by grants to A.W.B., E.C.N., and F.W. from the NH&MRC of Australia. We would like to thank Xiao-wen Xiao, Louis Lu, Hui-Hua Zhang, Kim Richards, Albert van Donkelaar, Julie Rothaker, Nathan Hall, Suzanne Orchard, and Josephine Iaria for valuable contributions.

Received: April 22, 2002
Revised: August 12, 2002

References

Aritomi, M., Kunishima, N., Okamoto, T., Kuroki, R., Ota, Y., and Morikawa, K. (1999). Atomic structure of the GCSF-receptor complex showing a new cytokine-receptor recognition. Nature *401*, 713–715.

Bajaj, M., Waterfield, M.D., Schlessinger, J., Taylor, W.R., and Blundell, T. (1987). On the tertiary structure of the extracellular domains of the epidermal growth factor and insulin receptors. Biochim. Biophys. Acta *916*, 220–226.

Banner, D.W., D'Arcy, A., Janes, W., Gentz, R., Schoenfeld, H.J., Broger, C., Loetscher, H., and Lesslauer, W. (1993). Crystal structure of the soluble human 55 kd TNF receptor-human TNFβ complex: implications for TNF receptor activation. Cell *73*, 431–445.

Barnham, K.J., Torres, A.M., Alewood, D., Alewood, P.F., Domagala,

T., Nice, E.C., and Norton, R.S. (1998). Role of the 6–20 disulfide bridge in the structure and activity of epidermal growth factor. Protein Sci. *7*, 1738–1749.

Brizzard, B.L., Chubet, R.G., and Vizard, D.L. (1994). Immunoaffinity purification of FLAG epitope-tagged bacterial alkaline phosphatase using a novel monoclonal antibody and peptide elution. Biotechniques *16*, 730–735.

Brunger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jian-Sheng Jiang, J.S., Kuszewski, J., Nilges, M., Navraj, S., et al. (1998). Crystallography & NMR systems: a new software suite for macromolecular structure determination. Acta Crystallogr. D *54*, 905–921.

Cooke, R.M., Wilkinson, A.J., Baron, M., Pastore, A., Tappin, M.J., Campbell, I.D., Gregory, H., and Sheard, B. (1987). The solution structure of human epidermal growth factor. Nature *327*, 339–341.

Cowtan, K. (1994). DM: an automated procedure for phase improvement by density modification. Joint CCP4 and ESF-EACBM Newslett. Prot. Crystallogr. *31*, 34–38.

Degenhardt, M., Weber, W., Eschenburg, S., Dierks, K., Funari, S.S., Rapp, G., and Betzel, C. (1998). Crystallization and preliminary X-ray crystallographic analysis of the EGF receptor ectodomain. Acta Crystallogr. *D54*, 999–1001.

De La Fortelle, E., and Bricogne, G. (1996). Maximum-likelihood heavy-atom parameter refinement for multiple isomorphous replacement and multiwavelength anomalous diffraction methods. Methods Enzymol. *276*, 472–494.

de Vos, A.M., Ultsch, M., and Kossiakoff, A.A. (1992). Human growth hormone and extracellular domain of its receptor: crystal structure of the complex. Science *255*, 306–312.

Elleman, T.C., Domagala, T., McKern, N.M., Nerrie, M., Lonnqvist, B., Adams, T.E., Lewis, J., Lovrecz, G.O., Hoyne, P.A., Richards, K.M., et al. (2001). Identification of a determinant of epidermal growth factor receptor ligand-binding specificity using a truncated, high-affinity form of the ectodomain. Biochemistry *40*, 8930–8939.

Garrett, T.P.J., McKern, N.M., Lou, M., Frenkel, M.J., Bentley, J.D., Lovrecz, G.O., Elleman, T.C., Cosgrove, L.J., and Ward, C.W. (1998). Crystal structure of the first three domains of the type-1 insulin-like growth factor receptor. Nature *394*, 395–399.

Groenen, L.C., Nice, E.C., and Burgess, A.W. (1994). Structure-function relationships for the EGF/TGF-α family of mitogens. Growth Factors *11*, 235–257.

Gunther, N., Betzel, C., and Weber, W. (1990). The secreted form of the epidermal growth factor receptor. J. Biol. Chem. *265*, 22082–22085.

Harvey, T.S., Wilkinson, A.J., Tappin, M.J., Cooke, R.M., and Campbell, I.D. (1991). The solution structure of human transforming growth factor α. Eur. J. Biochem. *198*, 555–562.

Jacobsen, N.E., Abadi, N., Sliwkowski, M.X., Reilly, D., Skelton, N.J., and Fairbrother, W.J. (1996). High-resolution solution structure of the EGF-like domain of heregulin-α. Biochemistry *35*, 3402–3417.

Jorissen, R.N., Epa, V.C., Treutlein, H.R., Garrett, T.P.J., Ward, C.W., and Burgess, A.W. (2000). Characterization of a comparative model of the extracellular domain of the epidermal growth factor receptor. Protein Sci. *9*, 310–324.

Kirsch, T., Sebald, W., and Dreyer, M.K. (2000). Crystal structure of the Bmp-2-Bria ectodomain complex. Nat. Struct. Biol. *7*, 492–496.

Kobe, B., and Kajava, A.V. (2001). The leucine-rich repeat recognition motif. Curr. Opin. Struct. Biol. *11*, 725–732.

Kohda, D., and Inagaki, F. (1992). Three-dimensional nuclear magnetic resonance structures of mouse epidermal growth factor in acidic and physiological pH solutions. Biochemistry *31*, 11928–11939.

Kohda, D., Odaka, M., Lax, I., Kawasaki, H., Suzuki, K., Ullrich, A., Schlessinger, J., and Inagaki, F. (1993). A 40-kDa epidermal growth factor/transforming growth factor α-binding domain produced by limited proteolysis of the extracellular domain of the epidermal growth factor receptor. J. Biol. Chem. *268*, 1976–1981.

Kunishima, N., Shimada, Y., Tsuji, Y., Sato, T., Yamamoto, M., Kumasaka, T., Nakanishi, S., Jingami, H., and Morikawa, K. (2000). Structural basis of glutamate recognition by a dimeric metabotropic glutamate receptor. Nature *407*, 971–977.

Lax, I., Johnson, A., Howk, R., Sap, J., Bellot, F., Winkler, M., Ullrich, A., Vennstrom, B., Schlessinger, J., and Givol, D. (1988). Chicken epidermal growth factor (EGF) receptor: cDNA cloning, expression in mouse cells, and differential binding of EGF and transforming growth factor alpha. Mol. Cell. Biol. *8*, 1970–1978.

Lax, I., Bellot, F., Howk, R., Ullrich, A., Givol, D., and Schlessinger, J. (1989). Functional analysis of the ligand binding site of EGF-receptor utilizing chimeric chicken/human receptor molecules. EMBO J. *8*, 421–427.

Lemmon, M.A., and Schlessinger, J. (1994). Regulation of signal transduction and signal diversity by receptor oligomerization. Trends Biochem. Sci. *19*, 459–463.

Lemmon, M.A., Bu, Z., Ladbury, J.E., Zhou, M., Pinchasi, D., Lax, I., Engelman, D.M., and Schlessinger, J. (1997). Two EGF molecules contribute additively to stabilization of the EGFR dimer. EMBO J. *16*, 281–294.

Lo Conte, L., Chothia, C., and Janin, J. (1999). The atomic structure of protein-protein recognition sites. J. Mol. Biol. *285*, 2177–2198.

Louie, G.V., Yang, W., Bowman, M.E., and Choe, S. (1997). Crystal structure of the complex of diphtheria toxin with an extracellular fragment of its receptor. Mol. Cell *1*, 67–78.

Lu, H.S., Chai, J.J., Li, M., Huang, B.R., He, C.H., and Bi, R.C. (2001). Crystal structure of human epidermal growth factor and its dimerization. J. Biol. Chem. *276*, 34913–34917.

Montelione, G.T., Wuthrich, K., Nice, E.C., Burgess, A.W., and Scheraga, H.A. (1987). Solution structure of murine epidermal growth factor: determination of the polypeptide backbone chain-fold by nuclear magnetic resonance and distance geometry. Proc. Natl. Acad. Sci. USA *84*, 5226–5230.

Moriki, T., Maruyama, H., and Maruyama, I.N. (2001). Activation of preformed EGF receptor dimers by ligand-induced rotation of the transmembrane domain. J. Mol. Biol. *311*, 1011–1026.

Moscatello, D.K., Holgado-Madruga, M., Godwin, A.K., Ramirez, G., Gunn, G., Zoltick, P.W., Biegel, J.A., Hayes, R.L., and Wong, A.J. (1995). Frequent expression of a mutant epidermal growth factor receptor in multiple human tumors. Cancer Res. *55*, 5536–5539.

Moy, F.J., Li, Y.C., Rauenbuehler, P., Winkler, M.E., Scheraga, H.A., and Montelione, G.T. (1993). Solution structure of human type-α transforming growth factor determined by heteronuclear NMR spectroscopy and refined by energy minimization with restraints. Biochemistry *32*, 7334–7353.

Nagata, K., Kohda, D., Hatanaka, H., Ichikawa, S., Matsuda, S., Yamamoto, T., Suzuki, A., and Inagaki, F. (1994). Solution structure of the epidermal growth factor-like domain of heregulin-α, a ligand for p180^{erbB-4}. EMBO J. *13*, 3517–3523.

Ogiso, H., Ishitani, R., Nureki, O., Fukai, S., Yamanaka, M., Kim, J.-H., Saito, K., Sakamoto, A., Inoue, M., Shirouzu, M., and Yokoyama, S. (2002). Crystal structure of the complex of human epidermal growth factor and receptor extracellular domains. Cell *110*, this issue, 775–787.

Pellegrini, L., Burke, D.F., von Delft, F., Mulloy, B., and Blundell, T.L. (2000). Crystal structure of fibroblast growth factor receptor ectodomain bound to ligand and heparin. Nature *407*, 1029–1034.

Plotnikov, A.N., Hubbard, S.R., Schlessinger, J., and Mohammadi, M. (2000). Crystal structures of two FGF-FGFR complexes reveal the determinants of ligand-receptor specificity. Cell *101*, 413–424.

Riese, D.J., II, and Stern, D.F. (1998). Specificity within the EGF family/ErbB receptor family signaling network. Bioessays *20*, 41–48.

Schlessinger, J., Plotnikov, A.N., Ibrahimi, O.A., Eliseenkova, A.V., Yeh, B.K., Yayon, A., Linhardt, R.J., and Mohammadi, M. (2000). Crystal structure of a ternary FGF-FGFR-heparin complex reveals a dual role for heparin in FGFR binding and dimerization. Mol. Cell *6*, 743–750.

Singer, E., Landgraf, R., Horan, T., Slamon, D., and Eisenberg, D. (2001). Identification of a heregulin binding site in HER3 extracellular domain. J. Biol. Chem. *47*, 44266–44274.

Sorokin, A., Lemmon, M.A., Ullrich, A., and Schlessinger, J. (1994).

Stabilization of an active dimeric form of the epidermal growth factor receptor by introduction of an inter-receptor disulfide bond. J. Biol. Chem. *269*, 9752–9759.

Summerfield, A.E., Hudnall, A.K., Lukas, T.J., Guyer, C.A., and Staros, J.V. (1996). Identification of residues of the epidermal growth factor receptor proximal to residue 45 of bound epidermal growth factor. J. Biol. Chem. *271*, 19656–19659.

Sundaresan, S., Roberts, P.E., King, K.L., Sliwkowski, M.X., and Mather, J.P. (1998). Biological response to ErbB ligands in nontransformed cell lines correlates with a specific pattern of receptor expression. Endocrinology *139*, 4756–4764.

Tappin, M.J., Cooke, R.M., Fitton, J.E., and Campbell, I.D. (1989). A high-resolution 1H-NMR study of human transforming growth factor α. Structure and pH-dependent conformational interconversion. Eur. J. Biochem. *179*, 629–637.

Thiel, D.J., Le Du, M.-H., Walter, R.L., D'Arcy, A., Chene, C., Fountoulakis, M., Garotta, G., Winkler, F.K., and Ealick, S.E. (2000). Observation of an unexpected third receptor molecule in the crystal structure of human interferon-gamma receptor complex. Structure Fold Des. *8*, 927–936.

Ullrich, A., Coussens, L., Hayflick, J.S., Dull, T.J., Gray, A., Tam, A.W., Lee, J., Yarden, Y., Libermann, T.A., Schlessinger, J., et al. (1984). Human epidermal growth factor receptor cDNA sequence and aberrant expression of the amplified gene in A431 epidermoid carcinoma cells. Nature *309*, 418–425.

Walker, F., Hibbs, M.L., Zhang, H.H., Gonez, L.J., and Burgess, A.W. (1998). Biochemical characterization of mutant EGF receptors expressed in the hemopoietic cell line BaF/3. Growth Factors *16*, 53–67.

Ward, C.W., and Garrett, T.P.J. (2001). The relationship between the L1 and L2 domains of the insulin and epidermal growth factor receptors and leucine-rich repeat modules. BMC Bioinformatics *2*, 4.

Ward, C.W., Hoyne, P.A., and Flegg, R.H. (1995). Insulin and epidermal growth factor receptors contain the cysteine repeat motif found in the tumor necrosis factor receptor. Proteins *22*, 141–153.

Weber, W., Wenisch, E., Gunther, N., Marnitz, U., Betzel, C., and Righetti, P.G. (1994). Protein microheterogeneity and crystal habits: the case of epidermal growth factor receptor isoforms as isolated in a multicompartment electrolyzer with isoelectric membranes. J. Chromatogr. A *679*, 181–189.

Wiesmann, C., Fuh, G., Christinger, H.W., Eigenbrot, C., Wells, J.A., and de Vos, A.M. (1997). Crystal structure at 1.7 Å resolution of VEGF in complex with domain 2 of the Flt-1 receptor. Cell *91*, 695–704.

Wiesmann, C., Ultsch, M.H., Bass, S.H., and de Vos, A.M. (1999). Crystal structure of nerve growth factor in complex with the ligand-binding domain of the TrkA receptor. Nature *401*, 184–188.

Woltjer, R.L., Lukas, T.J., and Staros, J.V. (1992). Direct identification of residues of the epidermal growth factor receptor in close proximity to the amino terminus of bound epidermal growth factor. Proc. Natl. Acad. Sci. USA *89*, 7801–7805.

Wu, D.G., Wang, L.H., Sato, G.H., West, K.A., Harris, W.R., Crabb, J.W., and Sato, J.D. (1989). Human epidermal growth factor (EGF) receptor sequence recognized by EGF competitive monoclonal antibodies. Evidence for the localization of the EGF-binding site. J. Biol. Chem. *264*, 17469–17475.

Yarden, Y. (2001). The EGFR family and its ligands in human cancer: signalling mechanisms and therapeutic opportunities. Eur. J. Cancer *37*, (Suppl 4), S3–S8.

Yarden, Y., and Sliwkowski, M.X. (2001). Untangling the ErbB signalling network. Nat. Rev. Mol. Cell Biol. *2*, 127–137.

Accession Numbers

Atomic coordinates have been deposited in the Protein Data Bank (1MOX) and are available from the authors for academic purposes.

Note Added in Proof

Recently, the structure of ErbB3 was determined (Cho, H.S., and Leahy, D.J. [2002]. Structure of the extracellular region of HER3 reveals an interdomain tether. Science *297*, 1330–1333), and it shows that a large hinge motion between domains CR1 and L2 may occur upon binding.