Various chimeras of the ErbB1-specific ligands epidermal growth factor (EGF) and transforming growth factor-α (TGFα) display an enlarged repertoire as activators of ErbB2/ErbB3 heterodimers. Mutational analysis indicated that particularly residues in the N terminus and B-loop region of these ligands are involved in the broadened receptor specificity. In order to understand the receptor specificity of T1E, a chimeric ligand constructed by the introduction of the linear N-terminal region of TGFα into EGF, we determined in this study the solution structure and dynamics of T1E by multidimensional NMR analysis. Subsequently, we studied the structural characteristics of T1E binding to both ErbB1 and ErbB3 by superposition modeling of its structure on the known crystal structures of ErbB3 and liganded ErbB1 complexes. The results show that the overall structure of T1E in solution is very similar to that of native EGF and TGFα but that its N terminus shows an extended structure that is appropriately positioned to form a triple β-sheet with the large antiparallel β-sheet in the B-loop region. This conformational effect of the N terminus together with the overall flexibility of T1E, as determined by 15N NMR relaxation analysis, may be a facilitative property for its broad receptor specificity.

The structural superposition models indicate that hydrophobic and electrostatic interactions of the N terminus and B-loop of T1E are particularly important for its binding to ErbB3.

The ErbB signaling network is composed of the ErbB1 (or EGFRI, ErbB2 (HER2/Neu), ErbB3 (HER3), and ErbB4 (HER4) tyrosine kinase receptors. Upon ligand binding, these receptors dimerize into a variety of homodimeric and heterodimeric receptor complexes whereby the intrinsic kinases become activated, which results in a cascade of second messengers and a diversity of subsequent downstream signaling (1, 2). ErbB receptors play an important role in growth and differentiation of cells, whereas overexpression of both receptors and ligands has been found in several human cancers (3).

As a consequence, the ErbB signaling network is increasingly used as a therapeutic target for the development of anti-tumor drugs (4, 5). Recent crystallographic studies have boosted this field of research by providing a wealth of information on the structure of ErbB-ligand complexes (6, 7).

A critical step in ErbB receptor signaling is the binding of EGF-like ligands to the extracellular domain of the receptor. More than a dozen soluble ligands have been identified that can be categorized into three distinct groups. A first group, composed of epidermal growth factor (EGF),1 transforming growth factor-α (TGFα), and amphiregulin, binds specifically to ErbB1. A second group consists of neuregulin (NRG) with its multiple isoforms, which have specific affinity for ErbB3 and ErbB4. A third group binds to both ErbB1 and ErbB4 and is composed of β-cellulin, heparin-binding EGF, and epiregulin (8, 9). For ErbB2, no soluble ligand has been identified, but it forms the preferred dimerization partner for all other members of the ErbB family.

Despite having distinct receptor binding specificity, all ErbB growth factors have an EGF-like domain as a common motif, which is defined by three disulfide bridges that generate three looped regions, designated the A-, B-, and C-loop, in addition to linear N and C termini. A single hinge residue between the fourth and fifth cysteine divides the EGF-like ligand into an N-and C-terminal half, each with a two-stranded antiparallel β-sheet.

For several ligands, including EGF, TGFα, NRG-1α, and betacellulin, solution structures have been determined by NMR analysis under a variety of experimental conditions (10–15). Recently, also crystal structures have been reported of EGF (16), ErbB3 in unliganded form (17), and ErbB1 in complex with EGF (7) or TGFα (6). The last two structures established the relative orientation of the four domains in the extracellular part of the receptor and confirmed earlier observations that both EGF and TGFα bind a single ErbB1 receptor molecule, which subsequently dimerizes to form a 2:2 complex. Moreover, these studies indicated that specific residues in the B-loop region of EGF and TGFα are involved in binding to the first extracellular domain (domain I) of the ErbB1 receptor, whereas several residues in the A-loop, C-loop, and C terminus are in close contact with the third extracellular domain (domain III). Unlike EGF, also residues in the N terminus of TGFα directly contact the receptor in domain I.

Site-directed mutagenesis and phage display studies have

1 The abbreviations used are: EGF, epidermal growth factor; TGFα, transforming growth factor α; NRG, neuregulin; Mut, methanol utilization; HSQC, heteronuclear single-quantum coherence; NOE, nuclear Overhauser effect; TOCSY, total correlation spectroscopy; NOESY, nuclear Overhauser effect enhancement spectroscopy; r.m.s., root mean square.

Published, JBC Papers in Press, July 16, 2003, DOI 10.1074/jbc.M305603200

From the Departments of §Cell Biology and ¶Biophysical Chemistry, University of Nijmegen, Toernooiveld 1, 6525 ED Nijmegen, The Netherlands

Miriam Wingens‡§, Tine Walma¶, Hugo van Ingen‡, Catelijne Stortelers‡,
Jeroen E. M. van Leeuwen‡, Everardus J. J. van Zoelen‡, and Geerten W. Vuister¶

Structural Analysis of an Epidermal Growth Factor/Transforming
Growth Factor-α Chimera with Unique ErbB Binding Specificity*

Received for publication, May 28, 2003, and in revised form, July 14, 2003

‡ To whom correspondence should be addressed. Tel.: 31-24-3652522; Fax: 31-24-3652999; E-mail: m.wingens@sci.kun.nl.
identified amino acids that have only limited freedom of mutation and are thought to be involved directly in receptor binding (18). For EGF and TGFα binding to ErbB1, these highly conserved residues are located around the second and sixth cysteine and in the C terminus (19, 20). In contrast, for NRG binding to ErbB3, hydrophobic and charged residues located in the N terminus and B-loop appear of primary importance (21, 23). Supposing that all EGF-like ligands bind their respective ErbB receptor in a similar orientation, it appears that domain I of ErbB3 contributes mostly to ligand specificity, whereas for ErbB1 this is located in domain III (24, 25). In order to study ligand binding specificity, we have previously used a domain exchange strategy between EGF and TGFα and shown that replacing the N-terminal linear region of EGF with that from TGFα results in a chimera, designated here as T1E, with unique ErbB binding specificity. T1E not only maintains high affinity for ErbB1, the natural receptor for both EGF and TGFα, but has in addition gained the ability to bind to ErbB2/ErbB3 heterodimers with nearly identical affinity as NRG-1 (26). Subsequent mutation analysis showed that the TGFα residues His 5 and Phe 6 in the N terminus and the EGF residue Leu 8 at the tip of the B-loop are essential for the high affinity of T1E for ErbB2/ErbB3 heterodimers. Replacement of Leu 8 for the Glu residue present at the equivalent position in TGFα strongly impaired binding to ErbB2/ErbB3 heterodimers without affecting binding to ErbB1 (26). These results indicated that the combination of residues in the N terminus of TGFα and in the B-loop region of EGF mediate the enhanced binding affinity for ErbB2/ErbB3 heterodimers.

High affinity binding requires a proper three-dimensional structure of the ligand in combination with specific residues in the receptor binding domain. In order to understand the altered receptor specificity of T1E compared with EGF and TGFα, we have studied structure and dynamics of T1E by two- and three-dimensional heteronuclear NMR techniques. Subsequently, we modeled the interaction of T1E with both the ErbB1 and ErbB3 receptors, on the basis of the recently published crystal structures of the extracellular regions of ErbB3 (17) and the EGF-ErbB1 and TGFα-ErbB1 complexes (6, 7). Using these models, the specific roles of residues in the N terminus and B-loop of T1E in ErbB2/ErbB3 interaction were analyzed.

EXPERIMENTAL PROCEDURES

Construction of T1E Expression Vector—To produce sufficient amounts of T1E for NMR analysis, we used the P. pastoris glycosylation system (Invitrogen). Thereto, T1E was cloned into the P. pastoris expression vector pPICzAα. A gene construct encoding the human TGFα/EGF chimera T1E, previously described by Stortelers et al. (26), was used as a template for the amplification of the T1E fragment by PCR. The primers used for PCR introduced an XhoI and a SalI restriction site permitting directional cloning of the amplified DNA in frame with the α-factor leader sequence in the pPICzAα expression vector. A stop codon preceding the SalI site prevented expression of the His 5 tag. The amplified fragment was first cloned into PCR-2.1 TOPO® cloning vector (Invitrogen). The recombinant TOPO vector was digested with XhoI/SalI to generate a 189-bp fragment, which was subsequently introduced into the pPICzAα expression vector using its XhoI/SalI sites. The correctness of the constructs was confirmed by DNA sequencing. Approximately 7 μg of the DNA construct were linearized with BstXI prior to transformation of P. pastoris X33 (Mut + ) and KM71H (MutS) cells. The electroporation method of the EasyShot® expression kit (version F; Invitrogen) was used for transfection. After 30 min electroporation and 24 h incubation, the transformed cells were plated on methanol utilization (Mut) phenotype. Southern blotting was used to confirm the presence of multiple integrations in the genome of selected high zeocin-resistant colonies. In small scale expression screens, the expression levels showed to be proportional to the number of integrations. A Mut + clone with multiple integrations was selected for large scale expression and 15N labeling.

Expression of Recombinant 15N-labeled T1E in P. pastoris—To produce 15N-labeled T1E, we slightly modified the protocol of Wiles et al. (27). In brief, 25 ml of BMG culture medium (0.1 M potassium phosphate (pH 6.0), 0.34% (w/v) yeast nitrogen base without amino acids or ammonium sulfate, 1% (v/v) glycerol, 0.00004% (w/v) biotin, and 1% (w/v) ammonium sulfate) were inoculated with a fresh colony of the selected Mut + clone and grown overnight while shaking at 30 °C to midlogarithmic phase. The culture was scaled up to 500-ml batches of BMG medium with ammonium sulfate but containing 0.1% (w/v) 15N-labeled ammonium chloride as its sole nitrogen source (99% 15N; Campro Scientific, Veenendaal, The Netherlands) in 2-liter baffled flasks. The culture temperature, the pH, and the agitation rate were controlled automatically. The cells were transferred to an induction medium (0.1 M potassium phosphate (pH 6.0), 0.34% (w/v) yeast nitrogen base without amino acids or ammonium sulfate, 0.5% (v/v) methanol, 0.00004% (w/v) biotin, and 1% (w/v) 15N-labeled ammonium chloride) and grown for 9 days with additional supplies of methanol every 24 h. The minimal amount of nitrogen source needed for optimal protein production was determined to be 0.1% in the growth medium and 1% in the induction medium.

Protein Purification—The P. pastoris culture supernatant was subjected to cation exchange chromatography (SP-Toyopearl 550C, Tosoh Corp., Tokyo, Japan) and eluted with a linear salt gradient (0.2–1.0 M NaCl in 0.05 M NaAc). The fractions were tested for binding affinity to 125I-mouse EGF binding competition assay and subsequently described by van de Poll et al. (28) and analyzed in parallel by nondenaturing SDS-PAGE and Western blot analysis using a polyclonal antibody (Ab-3) raised against recombinant wild type human EGF (Onco gene Science Inc., Cambridge, MA). Specific antibody binding was detected by a goat anti-rabbit antibody linked to horseradish peroxidase and visualized by enhanced chemiluminescence. Fractions positive for ErbB1 binding and displaying only a single band on Western blot were collected, dialyzed against 0.2 M HAc, and lyophilized. After resolubilization in 0.5 M HAc, the protein was finally purified by reverse phase high pressure liquid chromatography on a Deltapak C 18 column (Waters Associates, Milford, CT). Elution was carried out using a linear gradient of 20–40% acetonitrile (v/v) in 0.1% trifluoroacetic acid. Underivatized laser desorption-ionization-time-of-flight mass spectrometry, the purified protein was identified as 15N-labeled T1E(1–54) lacking the C-terminal arginine. The fractions corresponding to the main peak were collected, and the volume was reduced using speedvac centrifugation and subsequently dialyzed against double distilled water to remove remaining traces of organic solvents and finally against 50 mM phosphate buffer, pH 6.3. We thus obtained ~13 mg of uniformly 15N-labeled T1E/ liter of culture medium. The labeled material was biologically indistinguishable from the E. coli-derived T1E described previously (26).

Data Analysis and Structure Determination—The NMR sample was prepared to contain 0.8 mM protein in 85% H 2 O, 15% D 2 O (v/v), in 50 mM sodium phosphate (pH 6.3) with 20 μg/ml Pefabloc (Roche Applied Science). NMR spectra were recorded at 298 K on Varian Unity Inova spectrometers operating at 800- and 600-MHz 1H resonance frequencies. The following multidimensional experiments were recorded and analyzed: 1H HSQC, HNCA, HNHB, two-dimensional TOCSY, two-dimensional NOEY, 1H NOEY, 1H NOESY, and 1H-filtered TOCSY (for both the aromatic and the aliphatic region). All spectra were processed on a Silicon Graphics work station using NMRPipe software (29). Backbone and side chain resonances were assigned using XEASY (30). Nuclear Overhauser effect (NOE) peaks from the two- and three-dimensional NOESY spectra were classified either as very weak, weak, medium, or strong with an upper distance cutoff of 0.7, 5.0, 5.5, and 6.0 Å, respectively. All NOEs were performed using the program XPLOR 3.851 (31). Dihedral angle restraints were derived from J(1H,1H)-coupling constants, which were measured by a three-dimensional HNHA experiment (32). Subsequently, additional ϕ and χ angle restraints were predicted using TALOS (33), resulting in a total of 98 dihedral angle restraints. To improve the accuracy of all EFG-II templates, the structures were refined in water using a restrained molecular dynamics protocol under a CHARMM22 force field (34). The resulting ensemble of 36 structures was evaluated with MOLMOL (35), PROCHECK-NMR (36), and WHAT IF (37). Figures were produced using the programs RasTop and MOLMOL. Structural data for comparison with the T1E structure were taken from complexes of TGFα (14), HAP for the NMR structure of NRG-1a (13), LILJ for the crystal structure of EGF (16), and the Protein Data Bank files available on the World Wide Web at www.rcsb.org/wwpdb/structures/egf (10) for the NMR structure of EGF. The crystal structures of ErbB3 and the

Solution Structure and Dynamics of an EGF/TGFα Chimera
Solution Structure and Dynamics of an EGF/TGFα Chimera

Fig. 1. Solution structure of the EGF/TGFα chimera T1E. a, stereo views of the ensemble of 36 structures of T1E in two orientations, superimposed using the backbone atoms of all residues (pairwise r.m.s. deviation $2.43 \pm 0.63$ Å). Residues colored dark red are based on TGFα; gray and blue (β-sheet)-colored residues are EGF-based. b, sequence alignment of T1E with TGFα, EGF, and NRG-1β. The arrows depict residues that mediate enhanced binding affinity for ErbB2-ErbB3 heterodimers. For clarity, the ω-loop of NRG-1β, which is known to be dispensable for...
TABLE I
Structural statistics for the NMR structure of the EGF/TGFα chimera T1E

<table>
<thead>
<tr>
<th>Structural statistics</th>
<th>T1E</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total protein</td>
<td>2.43 ± 0.63</td>
</tr>
<tr>
<td>Bond lengths</td>
<td>0.973</td>
</tr>
<tr>
<td>Omega angle restraints</td>
<td>1.133</td>
</tr>
<tr>
<td>Inside/outside distribution</td>
<td>1.502</td>
</tr>
<tr>
<td>Average deviation (Z-scores, null deviation = 1)</td>
<td>1.61 ± 0.44</td>
</tr>
<tr>
<td>Average deviation (Z-scores, null deviation = 0)</td>
<td>0.44 ± 0.44</td>
</tr>
<tr>
<td>Coarse packing quality</td>
<td>-2.525</td>
</tr>
<tr>
<td>Ramachandran plot appearance</td>
<td>-2.097</td>
</tr>
<tr>
<td>χ₁-ψ2 rotamer normality</td>
<td>-2.470</td>
</tr>
</tbody>
</table>

TABLE I
Structural statistics for the NMR structure of the EGF/TGFα chimera T1E

<table>
<thead>
<tr>
<th>Structural statistics</th>
<th>T1E</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total protein</td>
<td>2.43 ± 0.63</td>
</tr>
<tr>
<td>Bond lengths</td>
<td>0.973</td>
</tr>
<tr>
<td>Omega angle restraints</td>
<td>1.133</td>
</tr>
<tr>
<td>Inside/outside distribution</td>
<td>1.502</td>
</tr>
<tr>
<td>Average deviation (Z-scores, null deviation = 1)</td>
<td>1.61 ± 0.44</td>
</tr>
<tr>
<td>Average deviation (Z-scores, null deviation = 0)</td>
<td>0.44 ± 0.44</td>
</tr>
<tr>
<td>Coarse packing quality</td>
<td>-2.525</td>
</tr>
<tr>
<td>Ramachandran plot appearance</td>
<td>-2.097</td>
</tr>
<tr>
<td>χ₁-ψ2 rotamer normality</td>
<td>-2.470</td>
</tr>
</tbody>
</table>

Values based on WHAT-CHECK output (53).

RESULTS

Structure Calculation and Quality—Using a series of homonuclear and heteronuclear NMR experiments, ¹H and ¹⁵N resonance assignments were obtained for T1E using the Wüthrich sequential assignment protocol (42). The proton resonance assignments were mainly derived from two-dimensional TOCOSY and two-dimensional NOESY spectra, whereas a three-dimensional ¹⁵N NOESY spectrum was used to identify the sequential ¹⁵N resonances of backbone amide groups and to verify the obtained proton resonance assignments.

From the NOE spectra, a set of 660 NOEs was collected, composed of 309 intrarresidual, 194 sequential, 67 medium range, and 90 long range NOEs. This resulted in an average of 12.2 distance restraints per residue. Additional constraints included nine hydrogen bonding restraints, three constraints from the disulfide bridges, and 98 φ and ψ angle restraints. We calculated an ensemble of 36 high resolution NMR structures of T1E. A superposition of this ensemble in two different orientations is shown in stereo view in Fig. 1a. The structural statistics are listed in Table I. All accepted structures contained neither distance violations greater than 0.5 Å nor angle violations greater than 5° from experimental data. Analysis of the Ramachandran plot showed that for the 36-structure ensemble 82.9% of the residues were found in the most favored regions, 16.4% in the additionally allowed regions, and 0.7% in generously allowed regions. No residues were found in the disallowed regions of the Ramachandran plot.

The data presented in Fig. 1a show that the protein contains two major stretches of secondary structure, consistent with the EGF fold. A large anti-parallel β-sheet is formed by residues Val¹⁻Ile²⁵ (β₁) and Lys³⁰⁻Asn³⁴ (β₂), and a smaller anti-par-

Solution Structure and Dynamics of an EGF/TGFα Chimera 39117

EBB3 binding, is not shown. Color coding is as in a, c, T1E structure ensemble superimposed for B-loop residues (pairwise r.m.s. deviation 1.27 ± 0.4 Å) and C-loop residues (pairwise r.m.s. deviation 0.91 ± 0.29 Å), respectively. d) ribbon presentation of the T1E structure color-coded according to residue-specific S² values. Residues with S² = 0–0.7 are displayed in pink, S² = 0.7–0.8 in violet, and S² = 0.8–1.0 in light blue. No data are available for residues colored gray. The A-, B-, and C-loop as well as N and C termini are indicated, and the cysteine bridges are depicted by gray lines, e) ribbon presentation of the T1E structure color-coded according to residue-specific Rα values. Residues requiring a chemical exchange contribution are shown in green, f) superposition of the lowest energy structure of T1E onto EGF and TGFα in native and complexed forms. T1E is shown in green, native EGF in light blue (molecule A from 1JL9), native TGFα in orange (1YUG, mean structure), EGF complexed to ErbB1 in blue (from 1IVO.pdb), and TGFα complexed to ErbB1 in red (from 1MOX.pdb).
allel β-sheet is formed by residues Tyr39–Ile40 (β1) and Tyr46–Arg47 (β2) (see Fig. 1b). These two structural elements are connected by three disulfide bridges, located between Cys8 and Cys22, Cys16 and Cys33, and Cys35 and Cys44. In Fig. 1c, the TIE structures in the ensemble were superimposed for B-loop residues (left panel) and C-loop residues (right panel), revealing good convergence to a single fold in the regions containing the large and small β-sheets (in blue), with a global pairwise backbone r.m.s. deviation of 1.27 ± 0.4 Å and 0.91 ± 0.29 Å for B- and C-loop, respectively. The relative orientation of the N-terminal and C-terminal halves is determined primarily by 25 long range distance constraints derived from NOEs between N-terminal half residues (Leu17, His18, Gly20, Val21, or Met23) and C-terminal half residues (Asn34, Val36, Arg43, Cys44, or Gln45). Due to the elongated shape of the molecule, these NOEs cluster in a limited region at the base of the large β-sheet, and consequently no structural information is available connecting the B-loop region directly to the C-loop region. As a result, the pairwise r.m.s. deviation for the entire protein is inflated to 2.43 ± 0.63 Å.

Notably, the N terminus adopts an extended conformation that approximately follows the β1-strand, although it is much less ordered than the large β-sheet in the B-loop itself (Fig. 1a). Its conformation is determined, however, by several unambiguous, long range NOEs between Val1 and Leu28, Val2 and Leu28, Ser3 and Leu28, Phe5 and Met23, and Asn6 and Tyr24. In 4 of 36 structures in the ensemble, this resulted in a triple β-sheet formation of the N terminus with the antiparallel β-sheet in the B-loop, based upon criteria used by PROCHECK-NMR.

Relaxation and Dynamics Parameters—Despite the high degree of structural similarity, most members of the EGF family of polypeptide growth factors show marked differences in receptor binding specificities. Since ligand binding may be influenced by intrinsic dynamical properties of the protein, we characterized the backbone dynamics of TIE.

### Table II

<table>
<thead>
<tr>
<th>TGFα residues</th>
<th>ErbB1 residues</th>
<th>Corresponding residues in ErbB3</th>
</tr>
</thead>
<tbody>
<tr>
<td>His4</td>
<td>Tyr101</td>
<td>Tyr104</td>
</tr>
<tr>
<td>Phe5</td>
<td>Tyr45</td>
<td>Leu45</td>
</tr>
<tr>
<td>Arg22</td>
<td>Tyr45</td>
<td>Leu45</td>
</tr>
<tr>
<td>Glu26</td>
<td>Tyr101</td>
<td>Tyr104</td>
</tr>
<tr>
<td>Glu27</td>
<td>Arg125</td>
<td>Tyr129</td>
</tr>
<tr>
<td>Lys29</td>
<td>Glu29</td>
<td>Asp23</td>
</tr>
<tr>
<td>EGF residues</td>
<td>ErbB1 residues</td>
<td>Corresponding residues in ErbB3</td>
</tr>
<tr>
<td>Met21</td>
<td>Tyr45</td>
<td>Leu45</td>
</tr>
<tr>
<td>Ile23</td>
<td>Leu14</td>
<td>Leu14</td>
</tr>
<tr>
<td>Leu26</td>
<td>Leu26</td>
<td>Met26</td>
</tr>
<tr>
<td>Lys28</td>
<td>Glu29</td>
<td>Asp23</td>
</tr>
</tbody>
</table>

![Fig. 2. Dynamics of the EGF/TGFα chimera TIE. Shown is a plot of the 15N relaxation data (R1, R1', and NOE enhancement), and subsequent Modelfree analysis data (S2, generalized order parameter, τc, internal correlation time, and Rex, chemical exchange contribution) of TIE as a function of residue number. The β-sheets in the secondary structure of TIE and location of cysteine bridges are indicated at the top.](image-url)
The nice dispersion in the \(^{15}\)N HSQC spectra allowed the measurement of \(^{1}H\)-\(^{15}\)N NOE values for all but one nonproline residue. Data for His\(^{12}\) did not yield reliable fits in the relaxation analysis. As shown in Fig. 2, \(^{1}H\)-\(^{15}\)N NOE values averaged between 0.7 and 0.8 for residues 15–23 and 34–47, indicating relatively little backbone motions on a picosecond to nanosecond time scale. Adjacent regions, such as the tip of the B-loop, had \(^{1}H\)-\(^{15}\)N NOE values between 0.6 and 0.7, indicating more backbone mobility. The N- and C-terminal ends of T1E revealed extensive local flexibility on a fast time scale indicated by NOE values ranging from 0.6 down to ~0.755.

To further analyze the backbone internal motions, the Model-free approach was used (28, 29). Backbone amide order parameters \(S_2\) were determined for 46 residues (Fig. 2). These \(S_2\) values were grouped into three classes, which are visualized in Fig. 1d using a color-coding scheme. The local flexibility of the N- and C-terminal ends is reflected by \(S_2\) values of 0.2–0.7 and \(\tau_2\) values of >500 ps. Interestingly, the residues in the N and C termini, as well as His\(^{18}\), Gly\(^{20}\), Ala\(^{27}\), Leu\(^{28}\), Gly\(^{41}\), and Trp\(^{51}\), could only be fitted using a model incorporating an intermediate time scale motional parameter, although for the nonterminal residues the fit yielded relatively large error bars on the accompanying \(\tau_2\) values. In Fig. 1e, a ribbon presentation of the lowest energy T1E structure is shown with residues that required a contribution of the transverse relaxation rate \(R_\text{ex}\) in green. Non-zero values of \(R_\text{ex}\) were found for several residues mainly in the A-loop (residues Ser\(^{11}\)-Cys\(^{16}\)) and to a lesser extent also in the B-loop and C-loop region, which is indicative for chemical exchange processes within the millisecond to microsecond time range (Fig. 1e). Overall, the structure of T1E is relatively flexible throughout the entire amino acid sequence, with enhanced backbone flexibility and internal motions in the N and C termini and chemical exchange contributions in A-, B-, and C-loop regions.

**Structural Superposition of T1E with Related Structures**

Since the chimera T1E is composed of parts of the EGF and TGF\(\alpha\) sequences, we compared the T1E structure with the parental molecules by superposition. We superimposed the lowest energy T1E structure onto native EGF (molecule A from crystal dimer 1JL9) and native TGF\(\alpha\) (mean NMR structure 1YUG) as well as onto crystal structures of EGF and TGF\(\alpha\) in the receptor-bound state, as obtained from Protein Data Bank entries 1IVO and 1MOX, respectively (Fig. 1f). Differences in relative orientations of the N- and C-terminal halves resulted in relatively high overall r.m.s. deviation values of 5.83, 4.22, 4.78, and 3.62 Å for superposition onto native EGF, native TGF\(\alpha\), complexed EGF, and complexed TGF\(\alpha\), respectively. However, the B-loop region of T1E superimposed very well onto the B-loop region of both receptor-bound EGF and TGF\(\alpha\), with r.m.s. deviation values of only 0.92 and 0.86 Å, respectively.

**A Comparison of T1E Binding to ErbB1 Domain I and ErbB3 Domain I—T1E is unique in its properties since it can bind both to ErbB1 and ErbB3. Since the amino acid requirements for ligand binding to ErbB1 and ErbB3 are very much different, the three-dimensional structure of T1E provides the opportunity to compare the structural requirements for binding of a single ligand to two distinct receptor molecules. Domain exchange and phage display studies of T1E (26, 43) indicated that residues His\(^{4}\) and Phe\(^{5}\) in the B-loop of ErbB1 are highly involved in binding the N-terminal and B-loop region of T1E that confer specificity toward ErbB2/ErbB3 complexes (26). The structural superposition modeling was composed of a three-step procedure. We first superimposed T1E onto complexed TGF\(\alpha\). The structure of complexed TGF\(\alpha\) was preferred over that of complexed EGF, because only the N terminus of TGF\(\alpha\) is known to be involved in receptor binding (6). In the second step, the complete domain I of ErbB3 (with the homology-modeled N-terminal region; see “Experimental Procedures”) was superimposed on ErbB1 domain I obtained from the crystal structure of TGF\(\alpha\)-ErbB1. Finally, an overlay of these two sets was made based on the orientation of complexed TGF\(\alpha\) in both superpositions, thereby assuming that T1E binds ErbB1 in a similar manner as TGF\(\alpha\) does and that binding to ErbB3 domain I is similar to binding to ErbB1 domain I (Fig. 3a) compares the hydrophobic character of the receptor surface residues in these structural superposition models of T1E-ErbB1 domain I and T1E-ErbB3 domain I. Of particular interest are the residues juxtaposed to the crucial N-terminal residues His\(^{4}\) and Phe\(^{5}\). In the T1E-ErbB1 complex, His\(^{4}\) is in close proximity to Ser\(^{99}\) and Tyr\(^{101}\), allowing for stacking interaction with the latter. This interaction appears to be preserved in the T1E-ErbB3 complex, where the homologous Tyr\(^{104}\) replaces Tyr\(^{101}\), whereas Leu\(^{102}\) occupies the position of Ser\(^{99}\). Similarly as in TGF\(\alpha\), T1E-Phe\(^{5}\) is oriented toward Leu\(^{69}\) and Tyr\(^{45}\) in the T1E-ErbB1 complex. This residue faces a much more hydrophobic environment in the T1E-ErbB3 complex, formed by Met\(^{72}\) and Leu\(^{85}\). Furthermore, ErbB1 contains a large Phe residue at position 20. In ErbB3, two small residues, Gly\(^{48}\) and Ala\(^{52}\), are located in the corresponding region, making the hydrophobic cavity in ErbB3 larger and better available for large hydrophobic N-terminal ligand residues. These differences in hydrophobicity between ErbB1 domain I and ErbB3 domain I strongly suggest a role for specific hydrophobic contacts in ligand–ErbB receptor binding selectivity. In addition to specific hydrophobic contacts, we compared the contribution of electrostatic interactions in receptor binding specificity. Fig. 3b displays the electrostatic potentials for the ErbB1 domain I and ErbB3 domain I with T1E. The ligand binding interface of the ErbB3 domain I shows a larger surface with electronegative potential than the corresponding surface of ErbB1 domain I. Residues jointly responsible for this electronegative patch on ErbB3 domain I are Glu\(^{145}\), Asp\(^{146}\), and Glu\(^{131}\). Furthermore, the positive electrostatic potential resulting from Arg\(^{122}\) and Lys\(^{87}\) in ErbB1 are absent in ErbB3. In Fig. 3c, the receptor binding interface of the ligands is shown. T1E, EGF, and TGF\(\alpha\) are displayed as a surface representation with electrostatic potentials colored as in Fig. 3b, together with the backbone trace of ErbB1 domain I. The models in Fig. 3c
Fig. 3. Comparison of the T1E/ErbB1 domain I and T1E/ErbB3 domain I complexes obtained by structural superposition. a, hydrophobic interactions of T1E N terminus with domain I of ErbB1 and ErbB3. The backbone trace of T1E is colored green with selected side chains His4 and Phe5 shown in ball-and-stick representation (blue). Domain I of both receptors is shown in a space-filling representation and colored according to hydrophobicity (52), ranging from yellow (less hydrophobic) to red (more hydrophobic). b, Electrostatic potential plots of ligand-binding interfaces of ErbB1 domain I and ErbB3 domain I in surface representation. Relative orientation of T1E is shown by the green backbone trace. Backbones of residues His4, Phe5, and Leu28 are shown in yellow. Backbone traces of acidic and basic residues of T1E are colored
are rotated 180° around the vertical axis with respect to Fig. 3b. In the TGFα-ErbB1 complex, Glu27 at the B-loop tip of TGFα interacts with Arg22 of ErbB1 (6). In comparison, this negatively charged Glu27 of TGFα would probably be located in an unfavorably negatively charged environment on ErbB3, in particular close to Glu131. In line with these considerations, previous work showed that in EGF/TGFα chimeras, Glu27 prevents interaction with ErbB2-ErbB3 complexes, whereas the corresponding Leu in EGF is favorable to this interaction (26). Furthermore, a comparison of the electrostatic potentials of the three ligands (Fig. 3c) reveals that Arg22 in TGFα is responsible for a large positively charged patch, whereas EGF is more negatively charged on this side of the molecule as a result of the presence of Glu3 and Glu24 in this region. T1E lacks Glu3, but Glu26 at the end of the β1-strand is also responsible for a small negative patch on T1E. Together, these structural superposition models show large differences in electrostatic potentials of both ErbB receptors and EGF-like ligands, indicating that electrostatic attraction and repulsion play a major role in ErbB receptor specificity of ligand binding.

DISCUSSION

EGF, TGFα, and the chimera T1E all bind to the ErbB1 receptor with high affinity. Of these three ligands, T1E is unique because it is also able to interact with ErbB2-ErbB3 heterodimers (26). Ligand binding results from a proper conformation in combination with the presence of specific residues at positions that directly interact with the receptor. In the present study, we have determined the solution structure and quantified the backbone dynamics of T1E. To explain the binding selectivity of T1E, we have constructed models of T1E-ErbB1 domain I and T1E-ErbB3 domain I, using structural superposition. The present data indicate that the additional affinity of T1E for ErbB2-ErbB3 heterodimers most probably stems from the interplay of three different factors: (i) the conformational effect of the N terminus, (ii) the dynamic properties of T1E, and (iii) the presence of specific residues that form part of the binding epitope for domain I of ErbB3. The structural superposition models of T1E-ErbB1 domain I and T1E-ErbB3 domain I strongly suggest that both hydrophobic interactions and electrostatic attraction/repulsion may be prominent aspects in determining ErbB ligand binding selectivity. In addition, our data provide structural support for the observation that His8 and Phe8 in the N terminus and Leu28 in the B-loop region are critical residues in mediating ErbB3 binding affinity.

The overall structural fold of T1E in solution is very similar to the solution structures of other members of the EGF family, such as EGF and TGFα (10, 14): an N-terminal half with a flexible loop region (A-loop) and a large antiparallel β-sheet (forming the B-loop) and a C-terminal half that contains a minor antiparallel β-sheet (forming the C-loop). The N- and the C-terminal halves of EGF-like ligands are connected by a hinge residue, Asn18 in the case of T1E. In comparison with EGF and TGFα, the N terminus of T1E displays φ,ψ-angles consistent with an extended conformation. In 4 of 36 structures in the ensemble, the N terminus explicitly forms a triple β-sheet with the B-loop. The observation that NRG-1α, which is a natural ligand for the ErbB2-ErbB3 heterodimer, also contains a triple β-sheet in this structurally homogeneous region (13) suggests that stabilization of the N-terminal region into an extended structure could be a requirement for EGF-like growth factors to bind ErbB3. Although the N terminus shows enhanced backbone flexibility, receptor binding of suitably folded T1E may drive the equilibrium of T1E into the receptor-bound, triple β-sheet-containing structure.

Whereas the secondary structure and global fold of EGF-like structures are quite similar, the overall r.m.s. deviations of the superpositions are relatively high (Fig. 1f). This may represent real differences between these structures or, alternatively, suggest that some of the structures are less well defined, either too loose or too tight. In our study, we used a CHARMM22 water refinement protocol (34) to generate the structures of the T1E ensemble. This resulted in reliable Z-scores for the local geometry (Table I). Furthermore, the Ramachandran Z-score, which is a good indicator of the overall quality of the ensembles, yielded highly acceptable values for NMR structure ensembles. Hence, we consider our structure to be reliable in terms of a representation of our experimental data.

Although uncertainty in NMR data does not necessarily indicate motion between the N- and C-terminal halves, the presence of hinge-bending motions has previously been suggested for ErbB ligands (14, 44–47). The present observation that the B- and C-loop regions of the T1E ensemble superimpose nicely (see Table I) suggests that such motions could indeed be at play. In addition, many residues in T1E (Asp13, Gly14, Tyr15, Cys16, Val21, Glu42, Arg43, Gln45, and Tyr46) at the interface between the N- and C-terminal halves exhibit nitrogen-15 exchange line broadening (Fig. 1e), similarly as observed previously for TGFα, which is consistent with the existence of hinge-bending motions. Furthermore, superposition of TGFα structures in bound and native state revealed a change in relative orientation of the N- and C-terminal halves, which could indicate that bending of the hinge is a prerequisite for proper receptor binding (6).

Dynamics may play a significant role in broad receptor binding specificity of T1E. From the backbone relaxation data of T1E and the subsequent Modelfree analysis, it appears that T1E is a relatively flexible molecule. The protein exhibits significant backbone motions in the picosecond to nanosecond time scale, and ~40% of the residues require a chemical exchange contribution. The residues with the largest Rα values are located in the A-loop (Figs. 1e and 2). Furthermore, the data show Rα values for groups of residues that are separated in sequence but close together in space (Asp1 and Cys2 in the N terminus; Met23, Tyr24, and Ile25 in the β1-strand; and Tyr31 and Ala32 in the β2-strand) that may be indicative of synchronous motion of these domains. It has been proposed that flexibility may correlate with broadened receptor-ligand binding specificity (48), and the observed flexibility of T1E may thus be a facilitative property for its broad receptor specificity. T1E can adopt multiple conformations, thus facilitating binding to domains I and III of both ErbB1 and ErbB3 receptors by an induced fit mechanism. In this view, the conformation of T1E in which the N terminus and the B-loop form a triple β-sheet could particularly be favorable for binding to domain I of ErbB3. If only 10% of the total available T1E is in the correct conformation to bind ErbB3, as indicated by the percentage of structures showing a triple β-sheet, this could explain the
relatively low affinity of T1E for ErbB3 receptors alone and the need for stabilization by ErbB2 (26). A role for dynamics in the ligand-receptor recognition process has previously also been suggested for NRG-1α (49), where flexibility was observed in regions that are clearly important for receptor binding, such as both the N- and C-terminal regions (23). Our relaxation data, together with those described for other ErB ligands, indeed suggest that dynamic flexibility contributes to broad specificity of the ErB receptor recognition process.

To elucidate the role of specific residues in N terminus and B-loop of T1E for ErbB3 receptor binding, we made structural superposition models of T1E-ErbB1 domain I and T1E-ErbB3 domain I. From these models, we propose that the N terminus of T1E is involved in binding to ErbB3 by hydrophobic packing. His4 and Phe5, two residues in the N terminus of T1E that were earlier found to be important for ErbB2/ErbB3 binding specificity (26), are in proximity to a hydrophobic pocket in ErbB3 (Fig. 3a). The observation that this pocket is more hydrophobic and larger than the corresponding region in ErbB1 suggests that it favors binding of large hydrophobic residues such as His4 and Phe5. Moreover, the preformed extended conformation of the N terminus in T1E could facilitate surface presentation of His4 and Phe5, enhancing the formation of the required intermolecular contacts. In addition, these large hydrophobic residues in the N terminus of T1E could also be involved in compensating for less favorable acidic residues in the B-loop region. Unfavorable electrostatic interactions between Glu26 in T1E and Glu27 in ErbB3 could potentially be masked by His4 and Phe5 in the N terminus. This hypothesis is suggested for NRG-1 ligand-receptor recognition process has previously also been acknowledged for Cancer Research, Victoria, Australia) for providing the three-dimensional coordinates of the crystal structure of the first three domains of the extracellular portion of EGFRII complexed with TGFα (Protein Data Bank accession number 1CAK).

REFERENCES

42. Wuthrich, K. (1986) NMR of Proteins and Nucleic Acids, John Wiley & Sons,
Solution Structure and Dynamics of an EGF/TGFα Chimera