Rational Design for the Development of Epidermal Growth Factor Receptor Antagonists

E. J. J. van Zoelen, A. E. G. Lenferink, R. H. Kramer and M. L. M. van de Poll
Department of Cell Biology, University of Nijmegen, Nijmegen, The Netherlands

SUMMARY

Epidermal growth factor (EGF) and transforming growth factor-α (TGFα) bind with similar high affinity to the human EGF receptor. Using a domain-exchange strategy we have shown that the C-terminal linear region of these molecules is involved in high affinity receptor binding. By further single amino acid substitution in this linear C-terminal region, a putative interaction site of these ligands with their receptor has been identified. This identification of a receptor binding domain in EGF/TGFα provides an important initial step in the development of EGF receptor antagonists with significant clinical potential.

Introduction

Polypeptide growth factors play a central role in the control of cellular proliferation, and appear to be involved in a number of pathological processes including cancer, atherosclerosis, psoriasis and autoimmune diseases. Many transformed cells are able to produce growth factors themselves, and also express receptors for these growth factors on their surface. The resulting so-called autocrine growth stimulation may contribute to the autonomous proliferation of tumor cells. Alternatively, tumor cell-derived growth factors may stimulate the proliferation of surrounding non-transformed cells in a paracrine fashion11. It would be a very attractive possibility to be able to block the activity of these tumor-derived growth factors by constructing growth factor receptor antagonists. Such antagonists are defined as growth factor related molecules which are able to bind growth factor receptors without activating them, thereby acting as competitive inhibitors for naturally occurring growth factors. Other possibilities for blocking growth factor activity in vivo, e.g. by using antibodies against growth factors and their receptors or so-called soluble receptors, seem less feasible. Antibodies may induce a secondary immune response, and soluble receptors generally have a relatively low binding affinity for their ligand molecule. In the case of cytokines acting on hematopoietic cells, both naturally occurring and artificially designed antagonists have been characterized14,21. For polypeptide growth factors acting through receptors with intrinsic tyrosine kinase activity, antagonists have not yet been designed.

Epidermal growth factor (EGF) is most likely the best characterized polypeptide growth factor. It consists of a single amino acid chain of 53 amino acids (see Fig. 1) characterized by six cysteine residues, which form three disulfide bridges that are essential for biological activity5. EGF binds with high affinity to a single chain transmembrane 170 kDa receptor, which is present in most cell types studied so far. The extracellular part of the EGF receptor (EGF-R) consists of four distinct domains, of which domains II and IV are rich in cysteines, whereas domain III con-
tains the ligand binding domain. The intracellular part of the molecule contains an intrinsic tyrosine kinase domain, which is activated upon ligand binding. Evidence has been presented that ligand binding results in receptor dimerization, which leads to transphosphorylation of receptor molecules on tyrosine residues. These so-called autophosphorylated tyrosine residues subsequently act as anchoring sites for specific substrate molecules which contain a so-called SH2-domain. Binding and activation of these substrate molecules by tyrosine phosphorylation subsequently results in the generation of a variety of intracellular second messengers essential for growth stimulation, including activation of protein kinase C, p21 Ras and MAP-kinase, and induction of the c-fos and c-myc oncogenes.

EGF is not the only growth factor which is able to activate the EGF-R. Additional growth factors include transforming growth factor-α (TGFα), amphiregulin, heparin-binding epidermal growth factor and betacellulin. Furthermore a number of viral proteins have been characterized, including vaccinia growth factor, shope fibroma growth factor and myxoma growth factor, which also interact with the EGF receptor. All of these growth factors have a number of structural features in common, including the spacing of six cysteine residues similar to that in EGF. The best characterized of these molecules is TGFα, a 50 amino acid polypeptide (see Fig. 1) which plays an important role in embryogenesis and tumorigenesis. Many tumor cells coexpress TGFα and the EGF-R, which provides the possibilities of an autocrine loop. Almost all experimental data indicate that TGFα binds with similar high affinity to the same binding site on the receptor as EGF. Structural analysis using two-dimensional NMR techniques has shown that EGF and TGFα have a very similar conformation in solution. Nevertheless, TGFα has only 42% amino acid homology with EGF, most of which can be explained by their common structural requirements (see Fig. 1). It is therefore presently un-

---

**Fig. 1.** Primary structure of human EGF (top) and human TGFα (bottom). In EGF the cysteines and glycines conserved in all EGF-like growth factors, and essential for maintaining the proper three-dimensional structure, have been indicated in gray. The shaded amino acids have been shown to have only limited freedom of mutation, and are most likely involved in receptor binding. In TGFα the amino acids conserved with EGF are indicated in gray.
clear which amino acids in EGF and TGFα are directly involved in high affinity receptor binding.

The observation that TGFα is involved in the autocrine growth stimulation of many tumor cells, while in addition secreted TGFα can also stimulate proliferation of surrounding normal cells, is the reason why it would be of interest to develop an EGF-R antagonist through the analysis of constructed mutants of EGF and TGFα. Many studies have been performed on the structure-function relationship of EGF and TGFα, mainly using site-directed mutagenesis. These studies have primarily focused on those amino acids which the members of the EGF family have in common. The amino acids which have been shown essential for EGF binding to the receptor through this approach, are indicated in Fig. 1. These experiments have not resulted, however, in the direct identification of specific domains in these growth factor molecules, essential for either receptor binding or for receptor activation. Also, studies using synthetic peptides which correspond to partial sequences of EGF and TGFα, or antibodies directed against these growth factors have not provided conclusive evidence in this respect. For the development of an EGF receptor antagonist, it will be essential to discriminate between receptor binding and receptor activating domains in these molecules. In our approach we have used, therefore, a domain-exchange strategy between EGF and TGFα with the aim to identify domains involved in high affinity receptor binding.

Studies Using Chimeric EGF/TGFα Growth Factors

EGF and TGFα may bind with similar affinity to the human EGF receptor, but it has been established that in contrast to TGFα, EGF binds with only very low affinity to the chicken EGF receptor. Lax and coworkers have used this observation to construct chimeric human/chicken EGF receptors which allowed them to demonstrate that domain III in the extracellular region of the EGF-R contains the ligand binding site. We have used the difference in binding affinity of the chicken

<table>
<thead>
<tr>
<th>Structure</th>
<th>EGF</th>
<th>TGFα</th>
<th>E3T</th>
<th>T3E</th>
<th>E4T</th>
<th>T4E</th>
<th>T3E4T</th>
<th>E3T4E</th>
<th>E6T</th>
<th>T6E</th>
<th>T4E6T</th>
<th>E4T6E</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conserved Cysteines</td>
<td>C C C C CC C</td>
<td>C C C C CC C</td>
<td>C C C C CC C</td>
<td>C C C C CC C</td>
<td>C C C C CC C</td>
<td>C C C C CC C</td>
<td>C C C C CC C</td>
<td>C C C C CC C</td>
<td>C C C C CC C</td>
<td>C C C C CC C</td>
<td>C C C C CC C</td>
<td>C C C C CC C</td>
</tr>
<tr>
<td>Growth Factor Chimeras</td>
<td>++</td>
<td>+++</td>
<td>+++</td>
<td>+</td>
<td>+++</td>
<td>+</td>
<td>+++</td>
<td>+</td>
<td>+++</td>
<td>+</td>
<td>+++</td>
<td>+</td>
</tr>
</tbody>
</table>

Fig. 2. Survey of EGF/TGFα domain exchange mutants. Domains exchanged were bordered by the indicated conserved cysteine residues (C). Growth factor chimeras with high affinity for the chicken EGF receptor are listed as (+++), with low affinity as (+).
Fig. 3. Mutants of EGF/TGFα in the linear C-terminal region. The conserved 6th cysteine (C42 in EGF) and leucine (L47 in EGF) have been indicated in diamonds, EGF sequences in circles, and TGFα sequences in squares. The sequence N-terminal of the 6th cysteine is indicated by either E (EGF sequence) or T (TGFα sequence). High affinity of the mutants for the chicken EGF receptor is indicated by (+++), intermediate affinity by (++), and low affinity by (+).

EGF  E-CQYRDLLKWWELR  +
TGFα  T-CEHADLLA  +++
E6T  E-CEHADLLA  +++
T6E  T-CQYRDLLKWWELR  +
E6ET  E-CQYRDLLA  ++
E6TE  E-CEHADLLKWWELR  +++
T6ET  T-CQYRDLLA  +++
T6TE  T-CEHADLLKWWELR  +++
Q43E  E-CEYRDLLKWWELR  +
Y44H  E-CQHRDLKWWELR  +
R45A  E-CQYADLLKWWELR  +++
EGF48  E-CQYRDLLK  +
EGF49  E-CQYRDLLW  +
EGF50  E-CQYRDLLKW  ++

Fig. 4. Conservation of the putative EGF binding domain in the human (HER), mouse (MER) and chicken (CER) EGF receptors. The putative ligand binding domain identified by Wu et al.25 corresponds to amino acids 352–366 (horizontal dashed line) from the so-called domain III of the human EGF receptor. Domains II and IV correspond to cysteine-rich domains in the extracellular region of the receptor, TM (spanning the dotted area) stands for transmembrane domains, and TK for tyrosine kinase domain of the receptor. Amino acids in gray in MER and CER are not conserved in HER. In addition, the charge of amino acid side chains has been indicated (data from Avivi et al.1).
EGF receptor for EGF and TGFα as a starting point for analyzing receptor binding domains in EGF and TGFα. By exchanging domains between EGF and TGFα we have studied which domain of TGFα, when introduced into EGF, would restore the high affinity for the chicken EGF receptor. This domain in TGFα might then be considered as an important binding domain for the chicken EGF receptor. Since EGF and TGFα are structurally very similar, and the human and chicken EGF receptor differ only in a restricted number of amino acids, the binding domain identified in this way would be most likely of more general importance for the binding of EGF-like growth factors to their receptor molecules.

Figure 2 gives a survey of the EGF/TGFα chimeric constructs that we have made16. In all constructs the domains exchanged were bordered by the conserved cysteine residues in EGF and TGFα. Since the cysteine bridges in these molecules allow the formation of well defined loops with specific functions, this approach conserves the identity of these looped structures. In the nomenclature used in Fig. 2, a growth factor such as E3T represents a chimeric growth factor with EGF sequences N-terminal of the third cysteine, and TGFα sequences C-terminal of the third cysteine. Based on binding studies on NIH3T3 cells transfected with the chicken EGF receptor, we could establish that all chimeras produced had very similar binding affinity for the human EGF receptor. When tested on NIH3T3 cells transfected with the chicken EGF receptor, however, large differences in affinity were observed, as shown in Fig. 2. In agreement with literature data it was observed that TGFα has a high binding affinity, whereas EGF has a low binding affinity. All chimeras tested behaved either similar to TGFα, or similar to EGF. The data in Fig. 2 indicate that all chimeras with TGFα sequences C-terminal of the sixth cysteine have a high binding affinity for the chicken EGF receptor, while all chimeras with EGF sequences in this domain have low binding affinity for this receptor. It can therefore be concluded from these data that the linear C-terminal region of TGFα determines its high binding affinity for the chicken EGF receptor, strongly indicating that this domain is directly involved in ligand-receptor interaction.

The chosen domain-exchange approach has a number of distinct advantages over studies in which single amino acids are altered by site-directed mutagenesis8,14. Firstly, by exchanging domains between two structurally very related molecules, it can be expected that the chimeras obtained will have a very similar three-dimensional structure. In the case of a single amino acid substitution, activity may be lost if, as a result of mutation, the growth factor is unable to fold into an active conformation. Secondly, studies using site-directed mutagenesis have mainly focused on amino acids which are conserved between the members of the EGF family, including the amino acids required for proper protein folding. A domain-exchange strategy, however, focuses on the amino acids which are different between - in our experiments - EGF and TGFα. Thirdly, our conclusions regarding the role of the linear C-terminal region could be based on a relatively small number of mutants (see Fig. 2). The results, however, constitute a basis for more detailed studies on the role of individual amino acids in this domain. In contrast, in site-directed mutagenesis studies, a large number of mutations generally have to be made before a final conclusion can be drawn, particularly because individual amino acids can be mutated into a variety of other amino acids. A fourth and major difference between these approaches results from the observation that most single amino acid substitutions either have no effect on binding activity, or result in a loss or reduction of activity. Since a reduction in biological activity can have many different reasons, including problems with protein expression, processing and folding, it is difficult to draw final conclusions only based on a reduction in activity. In our domain-exchange strategy, EGF has low binding affinity for the chicken EGF receptor, but this affinity is enhanced by the introduction of TGFα sequences. Such an enhanced activity can be interpreted in a more straightforward way than a general reduction in activity. Essential for our approach is that a situation can be created in which the two proteins compared (e.g. EGF and TGFα) have different biological activities; in our experiments binding affinity for the chicken EGF receptor.

Based on the above conclusions, we have investigated in depth which individual amino acids in the linear C-terminal region of TGFα are involved in the high affinity binding to the chicken EGF receptor23. Figure 3 shows the individual amino acids in the C-terminal linear domain of TGFα and EGF. In agreement with the data in Fig. 2, it is shown that TGFα and E6T have a high binding affinity for the chicken EGF receptor, while EGF and T6E have a low binding affinity. The linear C-terminal region of both EGF and TGFα contains a highly conserved leucine residue (Leu 47 in EGF) which is known to be essential for biological activity (see Fig. 1). Additional exchange mutants were therefore made in the linear C-terminal domain on both sites of this leucine residue, in generating the chimeras E6ET, E6TE, T6ET and T6TE (for structure see Fig. 3). Interestingly, three of these mutants had similar high binding affinity for the chicken EGF receptor as TGFα. One, E6ET, had an intermediate affinity, but none of the mutants had an affinity as low as EGF. This shows that elements both N-terminal and C-terminal of the conserved leucine are important for high affinity receptor binding of TGFα. In order to analyze the region between the 6th cysteine and Leu 47, single amino acids were exchanged between EGF and TGFα. Figure 3 shows that the EGF mutants Q43E and Y44H have a low affinity for the chicken EGF receptor, similar to that of EGF, while R45A has a high affinity similar to that of TGFα. This shows that a single amino acid exchange in EGF can confer high binding affinity to the chicken EGF receptor. Similar studies concentrating on the region C-terminal of Leu 47, showed that deletion
of the last four or five amino acids of EGF, resulting in EGF49 and EGF48, does not affect receptor affinity. However, the affinity for the chicken EGF receptor of EGF50 is intermediate between that of EGF and of TGFa.

In general terms, these results can be interpreted by assuming that positively charged amino acids in the C-terminal linear domain of EGF hamper the binding of EGF to the chicken EGF receptor, which results in low binding affinity. In general, EGF is a more acidic protein than TGFa, in view of its isoelectric point of 4.5 against 5.8 for TGFa. However, the linear C-terminal region of EGF is much more basic than that of TGFa, with an apparent isoelectric point of 7.5 against 3.5 for TGFa. In the case of EGF, this is mainly due to the presence of Arg 45, Lys 48 and Arg 53, which are not conserved in TGFa. In EGF, the most important charged residue for impeding receptor binding seems to be Arg 45, since in all mutants analyzed substitution by a neutral amino acid (R45A) is sufficient to confer high affinity binding. When Arg 45 is maintained, low affinity binding is observed particularly when one of the two most C-terminal amino acids contains a positive charge. This explains why T6ET has high binding affinity, while EGF48 and EGF49 have low affinity. The effects of a positive charge close to the C-terminal end seem of less importance than the contribution of Arg 45, since E6ET and EGF50 only have intermediate binding affinity for the chicken EGF receptor.

Discussion

The presented data show that the C-terminal linear region of EGF and TGFa, and in particular Arg 45 in EGF, is involved in high affinity binding to the EGF receptor. A plausible explanation for this observation could be that positively charged amino acids in the ligand binding domain of the chicken EGF receptor cause repulsion of the interacting ligand. Such amino acids should then be absent in this domain in the human EGF receptor. The putative ligand binding site in the EGF receptor has been identified by cross-linking with anti-receptor antibodies which compete with EGF for receptor binding. Figure 4 shows the amino acid sequence of this part of the human EGF receptor located in extracellular domain III, in comparison with corresponding sequences in the mouse and chicken EGF receptors (data taken from ref. 1). In the putative ligand binding domain, indicated by the horizontal dashed line, the chicken EGF receptor differs in four amino acids from the human receptor, but does not have a higher content of positively charged amino acids. In addition, experimental data have shown that introduction of these four amino acids of the chicken EGF receptor into the human EGF receptor, does not result in altered binding affinity for EGF. However, immediately next to the C-terminal of the putative ligand binding domain the chicken EGF receptor contains two doublets of lysine residues which is not conserved in the human EGF receptor. One positively charged arginine is present on position 367 of the mouse EGF receptor, and it is known that human EGF binds with high affinity to the mouse receptors. The positively charged amino acid Lys 368 is most likely responsible, therefore, for the low affinity binding of EGF to the chicken EGF receptor, but mutation studies will have to be performed to directly prove this hypothesis.

The present data indicate that the linear C-terminal region of EGF and TGFa, and in particular the region around Arg 45, is directly involved in receptor binding. This is in agreement with recent NMR data, showing that the linear C-terminal region of TGFa is specifically immobilized upon binding to the EGF receptor. This does not exclude the possibility, that other receptor binding domains are present in EGF-like growth factors, which do not discriminate between human and chicken EGF receptors. Mutational analysis has indicated that particularly the amino acids around the second cysteine of EGF (Tyr 13, Leu 15, His 16; see Fig. 1) have only very limited freedom for mutation, while additional NMR studies have indicated that the C-terminal linear region folds back to interact with these amino acids. It is most likely that these two regions together form a multidomain site for interaction with the receptor. In a recent study, chimeric growth factors have been made combining domains of mouse EGF and human TGFa. Using a similar approach as described here, evidence has been presented in favor of the involvement of the looped structure 21–30 in EGF in receptor binding. From our mutational studies we could confirm that a small part of this domain may be required for proper protein folding needed for high affinity receptor binding (data not shown).

In further studies we are presently comparing the sequence of EGF with that of an EGF-like molecule encoded by the Drosophila Notch gene. This EGF-like repeat in Notch has a cysteine spacing similar to that of EGF, but it does not interact with the EGF receptor. If all the amino acids from EGF which are essential for binding can be introduced into the Notch-repeat, an EGF receptor-binding protein should be obtained. If receptor-activating domains are also absent in Notch, an antagonist can be constructed in this way. If not, additional point mutations in this chimeric protein outside the characterized receptor binding domains may be necessary in order to obtain a functional antagonist.

In recent years it has become clear that the EGF receptor is a member of a receptor family, which also includes erb-B2 (neu), erb-B3 and erb-B4. For the latter two receptors ligands have been characterized which have been designated heregulins. Interestingly, chimeric EGF-heregulin proteins have been designed which interact with both the EGF receptor and erb-B3/B4. Studies on activation of these receptors have shown that ligands for these receptors, including EGF, may induce the formation of homo-
and heterodimeric receptor complexes. It is presently unclear if formation of an EGF receptor homodimer requires one or two EGF molecules, but recent data favour the first possibility. Based on this observation, it has been hypothesized that EGF may contain two distinct binding sites, one for each receptor. This concept has to be worked out in more detail, but it provides the possibility to generate EGF receptor antagonists by mutating EGF in such a way, that it can still bind with high affinity to a first receptor, but is unable to dimerize the second. In this way, the formation of functionally signaling receptor complexes will be prevented. For such an approach, the present experimental data could also form a useful starting point.

Acknowledgements

We would like to thank the Dutch Cancer Society (KWF) for their support.

References

13 Hoyt DW, Harkins RN, Debanne MT, O’Connor-McCourt M, Sykes BD (1994) Interaction of transforming growth factor alpha with the epidermal growth factor receptor: binding kinetics and differential mobility within the bound TGFa. Biochemistry 33: 15283–15292

Key words: EGF – TGFa – Structure-function relationship – Receptor-ligand interaction – Site-directed mutagenesis

Everardus J. J. van Zoelen, Department of Cell Biology, University of Nijmegen, Toernooiveld 1, 6525 ED Nijmegen, The Netherlands, Phone: (31)24-3 65 27 07, Fax: (31)24-3 65 29 38