Multidomain Binding of Transforming Growth Factor α to the Epidermal Growth Factor Receptor†

Audrey Richter,‡ J. Wayne Conlan,§ Michael E. Ward,§ Stephen G. Chamberlin,‡ Peter Alexander,‡ Nigel G. J. Richards,*,∥ and Donna E. Davies*,‡

CRC Unit of Medical Oncology and Department of Microbiology, Southampton University Medical School, Southampton General Hospital, Southampton SO9 4XY, U.K., and Departments of Chemistry and Pharmacology, University of Florida, Gainesville, Florida 32611

Received November 27, 1991; Revised Manuscript Received July 24, 1992

ABSTRACT: Solubilized epidermal growth factor receptor (EGF-R) has been used in an extension of the Geysen epitope mapping protocol in order to provide additional insight into the amino acid residues in human transforming growth factor α (hTGFα) which are critical to recognition and binding. Overlapping heptapeptides which encompassed the 50 amino acid primary sequence of hTGFα were synthesized on a polyethylene solid phase, and the amount of detergent-solubilized EGF-R bound to each peptide was measured using ELISA. EGF-R appeared to bind reproducibly to four heptapeptides cognate to sequences in both the N- and C-domains of hTGFα (residues 22–28, 28–34, 36–42, and 44–50). Visualization of these four regions on three-dimensional solution phase structures of hTGFα, derived from ^{1}H NMR measurements [Kline, T.-P., Brown, F. K., Brown, S. C., Jeffs, P. W., Kopple, K. D., & Mueller, L. (1990) *Biochemistry 29*, 7805–7813], indicated that the peptide segments are located on a *single* face of the protein and suggested the presence of a potential receptor binding cavity. If peptide segments within both the N- and C-domains of hTGFα are involved in binding to EGF-R, then this has direct consequences for possible molecular mechanisms by which receptor activation might take place. For example, the observed conformational flexibility in the six NMR-derived hTGFα structures due to variations in the main-chain torsion angles of Val-33, in combination with the involvement of residues from both domains in the proposed binding cavity, may imply that receptor activation results from interdomain reorientation in the protein ligand. Such a model is consistent with recent investigations of the interaction of EGF-R and its ligands using physical methods, which have indicated changes in the solution conformation of the receptor upon ligand binding [Greenfield, C., Hiles, I., Waterfield, M. D., Federwisch, M., Wollmer, A., Blundell, T. L., & McDonald, N. (1989) *EMBO J. 8*, 4115–4123]. We anticipate that the receptor-binding assay reported in this study might also be more generally applicable in probing the interaction of other biologically important peptides, and proteins, with cellular receptors of similar structure to that of EGF-R.

A number of different proteins, the best studied being epidermal growth factor (EGF)[1] and transforming growth factor α (TGFα), can bind to, and consequently activate, EGF receptors (EGF-R) on the surface of epithelial, mesenchymal, and a range of cancer cells, resulting in the initiation of DNA synthesis (Derynck, 1988; Massagué, 1990; Carpenter & Cohen, 1990; Ullrich & Schlessinger, 1990). Growth-promoting characteristics are associated with other proteins, such as heparin-binding EGF-like growth factor (Higashiyama et al., 1991) and amphiregulin (Shoyab et al., 1989), which contain EGF-like sequences and which presumably also exert their biological effects via EGF-R. All ligands for EGF-R show sequence homology. In particular, there are six conserved cysteine residues which, by participating in three disulfide bonds, give rise to three well-defined loops (A, B, and C) and the formation of two protein domains (Campbell et al., 1989, 1990). The N-terminal domain of hTGFα is generally defined as comprising residues 1–33, while the C-terminal domain is composed of residues 34–50.[2] Importantly, these disulfides impart sufficient conformational rigidity to EGF and TGFα for ^{1}H NMR measurements to be used in defining information concerning the secondary, and tertiary, structural characteristics of these proteins in aqueous solution (Cooke et al., 1987; Montelione et al., 1987; Kohda et al., 1989; Tappin et al., 1989; Kline et al., 1990; Harvey et al., 1991; Hommel et al., 1991).

Several approaches, mainly arising from comparison of the primary sequences of protein ligands able to bind EGF-R (Campbell et al., 1989, 1990), have begun to identify the oligopeptide segments of EGF and TGFα which mediate their binding interaction with EGF-R. However, details of the roles played by several regions of these proteins remain obscure, and in addition, little firm evidence has been obtained upon the molecular mechanisms by which these mitogenic proteins activate EGF-R and thereby initiate the cellular response. For example, the ability of monoclonal, monospecific human

† This work was supported by the Cancer Research Campaign (U.K.). Additional funding was provided by the American Cancer Society through an institutional grant to N.G.J.R.

* To whom correspondence should be addressed.
‡ CRC Unit of Medical Oncology.
§ Department of Microbiology.
∥ Departments of Chemistry and Pharmacology.

[1] Abbreviations: hTGFα, human transforming growth factor α; hEGF, human epidermal growth factor; mEGF, mouse epidermal growth factor; EGF-R, epidermal growth factor receptor; NMR, nuclear magnetic resonance; TFA, trifluoroacetic acid; IUdR, [^{125}I]-5-iodo-2′-deoxyuridine; ABTS, 2,2′-azinobis(3-ethylbenzothiazoline-6-sulfonic acid); DMEM, Dulbecco's minimal essential medium; DMF, dimethylformamide; Ac_2O, acetic anhydride; $(i\text{-Pr})_2NEt$, diisopropylethylamine; PBS, phosphate-buffered saline; SDS, sodium dodecyl sulfate; TMB, 3,3′,5,5′-tetramethylbenzidine; BSA, bovine serum albumin; HEPES, *N*-(2-hydroxyethyl)piperazine-*N*′-2-ethanesulfonic acid.

[2] Amino acid numbers given throughout the text correspond to the actual residue numbers of the ligand cited.

0006-2960/92/0431-9546$03.00/0

EGF (hEGF) antibodies to bind the hEGF/EGF-R complex (Katsuura & Tanaka, 1989) has been interpreted as indicating that hEGF residues 13–20 are exposed on the surface of the complex while peptides from both the B-loop region (amino acids 22–32) *and* from the C-terminal domain (amino acids 33–53) are in contact with the receptor. On the other hand, in a separate study (Hoeprich et al., 1989), monoclonal antibodies directed against the B-loop of TGFα did not appear to inhibit mitogenesis, leading to the claim that this structural element does not contribute to recognition and binding of the growth factor by EGF-R. We note, however, that the interpretation of such work is difficult as only antibodies possessing very high affinity for the B-loop residues could compete with EGF-R in binding TGFα, given that the dissociation constant for the complex between EGF-R and TGFα is approximately 10^{-9}–10^{-10} M (Lax et al., 1991).

A second approach, using recombinant DNA techniques, used to identify the functional residues within EGF and TGFα has been the preparation of mutants of both protein ligands with single or, in some cases, double amino acid substitutions and comparison of their biological activity, assuming that the tertiary structure remains unaffected (Campion et al., 1990; Cooke et al., 1990; Dudgeon et al., 1990; Engler et al., 1990, 1991). In these studies, mutation of specific residues in the C-domain (Defeo-Jones et al., 1988, 1989; Lazar et al., 1988, 1989; Campbell et al., 1990; Hommel et al., 1991) or the B-loop (Engler et al., 1988; Matsunami et al., 1990) of TGFα and/or EGF has resulted in loss of activity. In particular, Leu-48 and Arg-42 of TGFα have clearly been shown as critical in receptor binding, as have the homologous residues in hEGF (Leu-47 and Arg-41) (Engler et al., 1990; Matsunami et al., 1991). The key role of Leu-47 in hEGF activity has been confirmed in independent experiments involving controlled proteolysis of the protein (Gregory et al., 1988).

Measurements of the mitogenic activity of synthetic peptides have also been used to locate receptor-binding regions in TGFα and EGF, with an almost uniform lack of success. Hence, a large series of peptides, which included sequences cognate to the A-, B-, and C-loops of TGFα, failed to induce DNA synthesis (Defeo-Jones et al., 1988; Darlak et al., 1988, Tam et al., 1991). Several groups (Darlak et al., 1988; Eppstein et al., 1989) have also failed to reproduce an earlier claim (Nestor et al., 1985) that the linear peptide corresponding to residues 34–43 of TGFα blocks the mitogenic action of EGF. On the other hand, an analogue of this peptide, modified to increase its membrane affinity, has been reported to block EGF-dependent growth of a mammary carcinoma cell line (Eppstein et al., 1989), although it did not appear to compete with hEGF in binding to EGF-R. Similar studies using the corresponding sequence of mEGF (residues 34–42) have also failed to demonstrate any ability of the peptide to block the mitogenic effect of EGF although this peptide does appear to block the angiogenic effect of this growth factor (Nelson et al., 1991). Other studies involving linear peptides composed of sequences taken from EGF proteins have shown that two peptides cognate to residues 20–31 of mouse EGF (mEGF) (Komoriya et al., 1984; Nelson et al., 1991) and residues 13–32 of hEGF (Katsuura & Tanaka, 1989) are mitogenic, albeit at concentrations 10^4–10^5 times higher than that which is required for intact hEGF. These results are consistent with reports of the activity of linear and cyclic peptides, based on homologous residues in hTGFα from other laboratories (Han et al., 1988; Tam et al., 1991). Only large, bicyclic peptides possessing primary sequences which include segments cognate to the B-loop region appear to be significantly mitogenic (Heath & Merrifield, 1986; Tam et al., 1991), but these are still much less potent than the intact protein.

Using a novel extension of current methods for mapping peptide epitopes (Geysen et al., 1987), we have attempted to define the receptor-binding sites of hTGFα by assaying the ability of EGF-R to bind to immobilized overlapping heptapeptides encompassing the whole sequence of the mature protein. EGF-R lends itself to this approach as solubilized receptor can be prepared in relatively large quantities and its binding to the solid-phase peptides can be determined using a monoclonal antibody which does not block ligand–receptor interaction. In combination with structural information upon hTGFα, derived from ^{1}H NMR measurements and computer modeling (Kline et al., 1990), our results imply the existence of a complex *multidomain* interaction between TGFα and EGF-R. In this paper we present a detailed description of the assay, the evidence which suggests that the interaction between hTGFα and EGF-R is due to interactions with residues in both domains of the growth factor, and a discussion of the implications of a multidomain interaction for the possible molecular mechanism by which TGFα and EGF might activate EGF-R.

MATERIALS AND METHODS

All tissue culture reagents were from Gibco, U.K. The HN5 human squamous cell carcinoma cell line (Cowley et al., 1986) was a gift from Professor B. Gusterson, Institute for Cancer Research, Sutton, U.K. mEGF was a gift from Dr. P. Moore, CSIRO, Australia. Immobilized TGFα heptapeptides were synthesized on polyethylene pins using a kit supplied by Cambridge Research Biochemicals, Cheshire, U.K. Amino acid derivatives were from Millipore (U.K.) Ltd., Watford, and Novabiochem (U.K.) Ltd., Nottingham, U.K. 2,2′-Azinobis(3-ethylbenzothiazoline-6-sulfonic acid) (ABTS) was obtained from Sigma Chemical Co. The mouse monoclonal anti-EGF-R antibody (EGF-R1), sheep anti-mouse immunoglobulins conjugated to horseradish peroxidase, ^{125}I-mEGF, and [^{125}I]-5-iodo-2′-deoxyuridine (IUdR) were from Amersham International. Falcon 3912 Micro Test III (96-well) ELISA trays were supplied by Becton and Dickinson, Oxford, U.K. HN5 cells and primary human foreskin fibroblasts were cultured in Dulbecco's minimal essential medium (DMEM) supplemented with 10% fetal calf serum, 100 μg/mL streptomycin sulfate, and 100 μg/mL benzylpenicillin in an humidified atmosphere of 95% air and 5% CO_2 at 37 °C.

Synthesis of Immobilized Overlapping Heptapeptides of hTGFα. Overlapping heptapeptides which encompassed the primary sequence of hTGFα were synthesized on solid polyethylene pins in an 8 × 12 array, using a well-established procedure (Conlan et al., 1988), based upon the original method (Geysen et al., 1987). In all coupling reactions, the required L-amino acids were generally used as their activated pentafluorophenyl ester derivatives. Serine and threonine, however, were coupled as their hydroxybenzotriazole esters. The fluorenylmethoxycarbonyl (FMOC) group was used to protect N_α (Atherton & Sheppard, 1989; Carpino et al., 1982), and in general, side-chain functional groups were protected as their *tert*-butyl derivatives. In the case of arginine, the side-chain guanidino group was protected using the methoxytrimethylphenylsulfonyl group. All pins were prederivatized with FMOC-β-alanine via a precoupled, flexible linker consisting of hexamethylenediamine and acrylic acid. All syntheses were begun by deprotection of β-alanine, under mildly basic conditions, using 20% freshly distilled piperidine,

in DMF from which the contaminant dimethylamine had been removed by storage over 4-Å molecular sieves (Union Carbide). The lack of contaminants in the DMF solvent was confirmed colorimetrically immediately prior to synthesis (Stewart & Young, 1984). Following deprotection, and washing, the appropriate FMOC-amino acid ester was dissolved to a concentration of 30 mM in DMF which contained 30 mM 1-hydroxybenzotriazole as catalyst. This solution was then dispensed into the polyethylene microtitration trays in 100-μL aliquots. Amino acid coupling was carried out overnight at room temperature. Deprotection and coupling steps were repeated until peptides of the required length were synthesized, and finally the N-terminus of each peptide was acylated by reaction with DMF/Ac_2O/$(i\text{-}Pr)_2NEt$ (50:5:1). Side-chain deprotection, and subsequent neutralization, was accomplished by reaction with TFA/phenol/ethanedithiol (95:2.5:2.5) followed by washing with CH_2Cl_2/$(i\text{-}Pr)_2NEt$ (95:5). The derivatized pins were then washed with methanol for 18 h and dried in vacuo. Following this procedure, it was possible to synthesize several hundred peptides simultaneously for the analysis of EGF-R binding specificity. Positive (Pro-Leu-Ala-Gln) and negative (Gly-Leu-Ala-Gln) control peptides were also synthesized and assayed for appropriate reactivity with commercially supplied anti-PLAQ antibody. All heptapeptides were synthesized in duplicate. The positive reactivity of specific heptapeptides was confirmed by synthesis and reassay on a number of separate occasions.

Binding of Overlapping Heptapeptides from hTGFα to EGF-R. Heptapeptides covering the whole sequence of hTGFα were synthesized on polyethylene pins, as described, and tested for EGF-R binding in a modified ELISA (Figure 1). Plasma membrane vesicles were prepared from the squamous cell carcinoma cell line HN5 by hypotonic blebbing (Cohen et al., 1982). EGF-R was solubilized from the vesicles by incubation, for 1 h at room temperature, with 5% Triton X-100 in 20 mM HEPES (pH 7.4) containing 10% glycerol. After centrifugation at 10000*g* for 1 h, the supernatant containing the solubilized receptor was diluted 1:10 in PBS containing 1% BSA, 1% ovalbumin, and 0.1% Tween 20, giving a final protein concentration of 25 μg/mL. This receptor preparation was then applied to the immobilized peptides (175 μL/well) and incubated for 18 h at 4 °C. Binding of EGF-R to the heptapeptides was determined using a monoclonal anti-EGF-R antibody (EGF-R1, 0.5 μg/mL) which binds to the extracellular domain of EGF-R at a site distant from the ligand-binding domain (Waterfield et al., 1984). Complexes containing this monoclonal antibody were detected using sheep anti-mouse immunoglobulin conjugated to horseradish peroxidase and visualized using 2,2′-azinobis(3-ethylbenzothiazoline-6-sulfonic acid) (ABTS) as chromogen (Geysen et al., 1987). Following incubation, in the dark, for 30–45 min, the absorbance of each well was measured at 410 nm. These assays involving EGF-R were repeated at least five times. Solid-phase peptides were reused following the removal of bound receptor and antibodies by ultrasonication in hot buffered SDS and 2-mercaptoethanol (Geysen et al., 1987).

Molecular Modeling of hTGFα. Coordinates for six conformations consistent with experimental data, generated using distance geometry (Kaptein et al., 1988; Fesik & Zuiderweg, 1990) and constrained molecular dynamics (Brünger et al., 1986), were generously provided to us by Dr. T.-P. Kline and co-workers (Kline et al., 1990). These were determined to be of high quality using a variety of computational algorithms (Brown et al., 1989; Hempel & Brown, 1990) and are now available in the Brookhaven Protein Data Bank (Bernstein et al., 1977). The coordinates were converted into MacroModel format and visualized using the graphical modeling package FLIMMSY V1.1 on a Stardent 1000 workstation.[3] Structures were routinely rendered as solid CPK models using the default values of atomic van der Waals radii employed in the BATCHMIN computational software (Mohamadi et al., 1990). The conformational strain enthalpy of each hTGFα structure was computed using the AMBER potentials and parameters (Weiner et al., 1984, 1986) as implemented in BATCHMIN V3.0 (Mohamadi et al., 1990). Default values of all atomic partial charges were used in these calculations, and electrostatic, van der Waals, and hydrogen-bonding cutoffs were set to 12, 6, and 4 Å, respectively. The conformation possessing the lowest energy was routinely used in our analysis of the protein. No attempt was made to optimize any of the coordinate sets in these studies. All atoms, including hydrogens, in the protein segments corresponding to the heptapeptide sequences, bound by EGF-R, were color coded in order to highlight their location within the three-dimensional structure. In addition, the hydrophobicity of amino acids in these segments, together with Phe-15, was visualized on the basis of the usual classification of side-chain properties (Schulz & Schirmer, 1979).

RESULTS

Binding of EGF-R to Immobilized Overlapping Heptapeptides of hTGFα. Overlapping heptapeptides spanning the entire sequence of mature hTGFα were synthesized on polyethylene pins using a commercial kit according to established procedures (Geysen et al., 1987). Each peptide was connected to the solid phase through a flexible link consisting of a β-alanine residue coupled to the plastic via hexamethylenediamine and acrylic acid moieties. This technique permitted the simultaneous synthesis of many immobilized polypeptides, and for convenience these pins were mounted in a block corresponding to the 12 × 8 format of a microtiter immunoassay tray. The ability of specific immobilized heptapeptides to bind solubilized EGF-R was assayed using a modified ELISA (Figure 1). EGF-R, solubilized with detergent from plasma membrane vesicles, was incubated with the immobilized peptides for 18 h at 4 °C, and its binding was assayed using an anti-EGF-R monoclonal antibody (EGF-R1) which is known to interact solely with a polypeptide epitope in the extracellular domain of EGF-R distant from the ligand-binding region (Waterfield et al., 1984). The location of EGF-R1 was then determined with a sheep anti-mouse IgG antibody conjugated to horseradish peroxidase, the resultant complex being visualized in the usual manner. Control incubations performed in the absence of EGF-R showed that there was no direct recognition of the immoilized peptides by either the EGF-R1 antibody or the sheep anti-mouse IgG peroxidase conjugate. Well-defined binding of EGF-R was observed for heptapeptides cognate to residues 22–28, 28–34, 36–42, and 44–50 of hTGFα (Figure 1A). Five independent repeats of this experiment were carried out, comparable results being obtained in every case. As might be expected for overlapping peptides, some binding of EGF-R was observed on peptides adjacent to the major EGF-R-binding peptides. A notable exception was the heptapeptide cognate to residues 22–28 of hTGFα since no EGF-R binding was observed for the adjacent peptide possessing an identical primary sequence to residues

[3] N. G. J. Richards, M. H. Muir, and S. G. Chamberlin, unpublished results. The FLIMMSY graphical modeling package is freely available to academic users upon request (richards@qtp.ufl.edu).

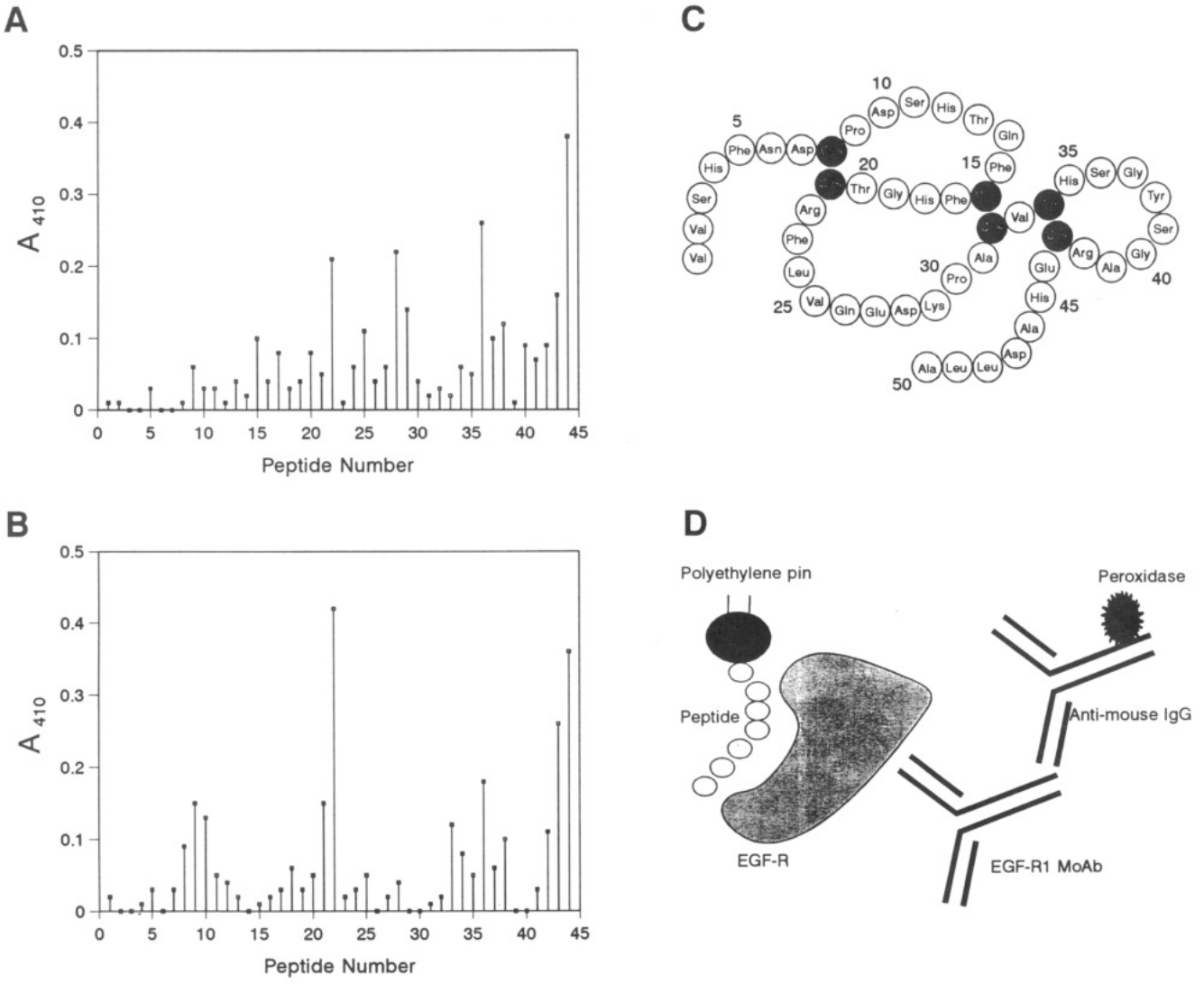

FIGURE 1: Binding of immobilized, overlapping heptapeptides, encompassing the complete primary sequence of hTGFα, to detergent-solubilized EGF-R. The mean absorbance (0.05) of the lowest 25% of the readings (background) has been subtracted from the measured values as recommended (Geysen et al., 1987). Heptapeptides are numbered according to the location of the N-terminal residue in hTGFα. For example, peptide 1 has the sequence cognate to residues 1–7 in hTGFα. The binding data are shown for the interaction of solubilized EGF-R with immobilized heptapeptides (A) in the absence and (B) in the presence of 0.8 nmol of mEGF. (C) Cartoon representation of the primary sequence of hTGFα, showing the location of the three conserved disulfide bonds and indicating the domain structure of the protein. (D) Schematic representation of the modified ELISA receptor-binding assay used to obtain these results.

21–27 of hTGFα. However, the latter peptide contained an N-terminal cysteine residue, and it is possible that binding of EGF-R may have been hindered by covalent cross-linking of immobilized peptides due to the formation of intermolecular disulfide bonds. On the other hand, the results of the assay were striking in that two of the immobilized sequences that appeared to be bound by EGF-R corresponded to segments 36–42 and 44–50 of hTGFα containing residues cognate to Gly-38, Tyr-39, Gly-40, Arg-42, and Leu-48. All five of these residues are not only conserved throughout the entire EGF/hTGFα family but have also been shown to be critical for biological activity in experiments employing site-directed mutagenesis.

To probe the specificity of binding of solubilized EGF-R to these putative ligands, we examined the ability of an EGF-R ligand to compete for binding to the receptor. In these studies, we used mEGF, prepared from murine salivary glands, since insufficient amounts of hTGFα were available for use an an authentic ligand. However, these control experiments using mEGF proved inconclusive, because the amount of heptapeptide immobilized on each individual pin was ca. 50 nmol, and excess, or even equivalent, amounts of the competing protein ligand still could not be attained. Thus we were only able to incubate solubilized EGF-R with immobilized heptapeptides in the presence of 0.8 nmol of mEGF. Binding of EGF-R to heptapeptide 28–34 was indeed blocked, but a reduction in EGF-R binding to the other three sequences of interest, as a consequence of the addition of mEGF, was not observed. A further surpising finding was that, in the presence of mEGF, the interaction of EGF-R to heptapeptide 22–28 was consistently enhanced (Figure 1B). One possible explanation for these observations was that EGF-R has binding sites for these four oligopeptides which are unrelated to those involved in binding the mitogenic protein ligands, TGFα and EGF, as has been observed in the case of laminin which binds to sites upon EGF-R distinct from those for TGFα or EGF (Panayotou et al., 1989). However, we believe this to be unlikely for several reasons. First, two of the four peptides which interacted strongly with EGF-R contained a number of conserved residues which mutagenesis experiments indicate are important to receptor binding. Second, in our original experimental design, we had anticipated that even an unfavorable molar ratio of

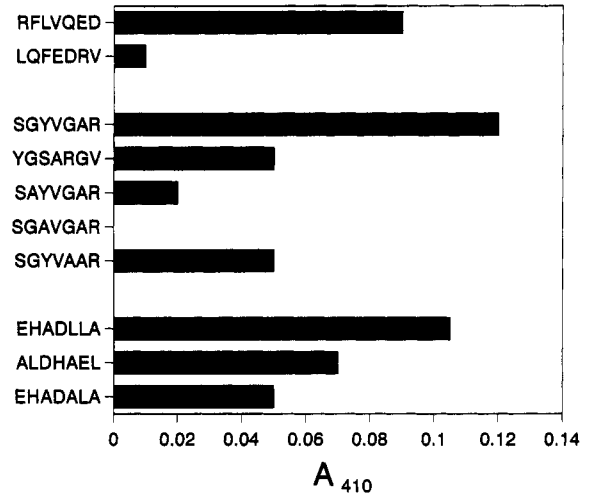

FIGURE 2: EGF-R binding to primary sequence variants of immobilized heptapeptides cognate to residues 22–28, 36–42, and 44–50 of hTGFα. Immobilized heptapeptides identical to the three protein segments, together with a series of analogues which either contained point substitutions or were randomized variants, were probed for their ability to bind EGF-R using the modified ELISA protocol. Binding data are shown after correction for background absorbance as described in Figure 1.

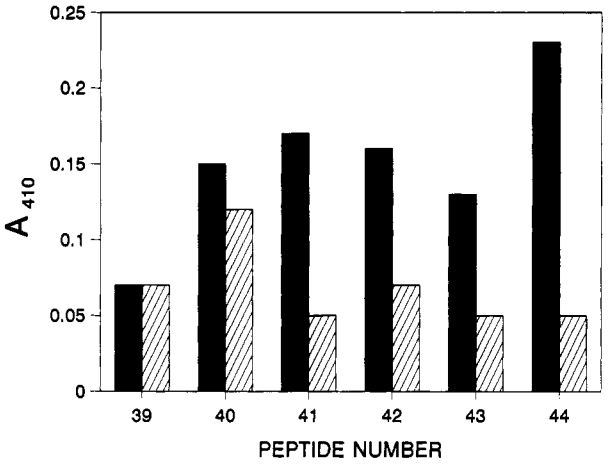

FIGURE 3: Effect of soluble hTGFα(44–50) peptide on binding of EGF-R to immobilized overlapping heptapeptides cognate to residues 39–50 of hTGFα. Immobilized hTGFα heptapeptides were probed for their ability to bind EGF-R in the absence (filled) or presence (hatched) of 150 nmol of soluble hTGFα heptapeptide cognate to residues 44–50 using the modified ELISA protocol. Binding data are shown after correction for background absorbance as described in Figure 1. Peptide numbers correspond to those shown in Figure 1.

competing ligand to immobilized peptide (approximately 1:60) would have been sufficient to block binding of EGF-R, as short solution-phase peptides, cognate to EGF or TGFα segments, had been reported to bind EGF-R with low affinity (Komoriya et al., 1984; Tam et al., 1991). However, the dissociation constant of EGF-R with immobilized peptides has not been determined. Third, although our choice of mEGF as the competing ligand was based purely on its availability, this did not seem unreasonable given that binding of EGF and TGFα to EGF-R appeared to be mutually competitive. However, more recent data seem to imply that the molecular nature of the interaction of EGF with EGF-R may not be identical to that of TGFα (Winkler et al., 1989; Lax et al., 1991; Korc et al., 1991), and we now recognize that such differences may become particularly apparent in our experimental situation where binding of small segments of a protein ligand are being studied. Hence it is possible that, due to its small size, a peptide fragment may still be able to occupy its site on a receptor in the presence of a heterologous competing ligand.

Effect of Sequence Randomization of Solid-Phase TGFα Heptapeptides upon EGF-R Binding. So as to seek alternative confirmatory evidence for the specificity of EGF-R binding, we synthesized randomized, or mutated, sequences for some of the heptapeptides for which the receptor had exhibited affinity. All peptides were synthesized under the same set of standard conditions such that the amount of peptide on each pin was equimolar. Randomization of the primary structure of peptides corresponding to residues 22–28, 36–42, and 44–50 to hTGFα reduced binding by 89%, 58% and 37%, respectively (Figure 2). Significantly, in the case of the heptapeptide corresponding to residues 36–42, which contained the conserved Gly-Tyr-X-Gly segment, point substitution of alanine for the three conserved amino acids resulted in a respective reduction of 83%, 100%, and 50% in receptor affinity. Substitution of the residue cognate to Leu-48 by alanine in the peptide of sequence corresponding to amino acids 44–50 of TGFα caused a 50% reduction in EGF-R binding. As the results from neither the randomization nor the point substitution of the heptapeptide cognate to residues 44–50 were entirely convincing, we further examined the specificity of the interaction of EGF-R with this immobilized fragment using a soluble heptapeptide cognate to residues 44–50 of hTGFα. When 150 nmol of soluble heptapeptide was added to EGF-R, binding of the receptor to the immobilized heptapeptide cognate to residues 44–50 of hTGFα was reduced by 82% (Figure 3). The specificity of the interaction also appeared evident because the soluble peptide blocked binding of EGF-R to peptides containing overlapping, related sequences but failed to block binding to those of unrelated primary sequence. When taken overall, we believe that these data support the hypothesis that sequence-specific recognition of the immobilized heptapeptides by EGF-R was being observed in this assay and, therefore, that residues cognate to amino acids which are both highly conserved within the EGF/TGFα family and important in mediating the binding of the intact protein to EGF-R were highly influential in determining the observed binding of EGF-R to specific heptapeptides.

Molecular Model of the Proposed Receptor-Binding Site of hTGFα. To interpret the results of the peptide-binding studies on a molecular level, we employed computer graphics to visualize the three-dimensional structure of hTGFα in aqueous solution, as determined using data for interproton distances determined by ^{1}H NMR techniques (Kline et al., 1990). Although the disulfide bonds impart intradomain rigidity, especially in the N-terminal domain, to the folded form of hTGFα, the six individual solution-phase structures built from NMR measurements adopted a variety of shapes, ranging from a compact, flexed form (Figure 4A) to an extended "rodlike" conformation (Figure 4D). This variation could be traced, in part, to modified main-chain torsion angles in the amino acid Val-33 (Table I) which separates the N- and C-domains of hTGFα.

The protein segments cognate to the immobilized heptapeptide sequences which were tightly bound by EGF-R were identified in each three-dimensional hTGFα structure. Significantly, in every conformation, all four segments were located on a *single* face of the molecule and formed an almost continuous surface which stretched from the β-sheet of the B-loop to the flexible C-terminal "tail". This is well illustrated by consideration of the most compact hTGFα structure, which

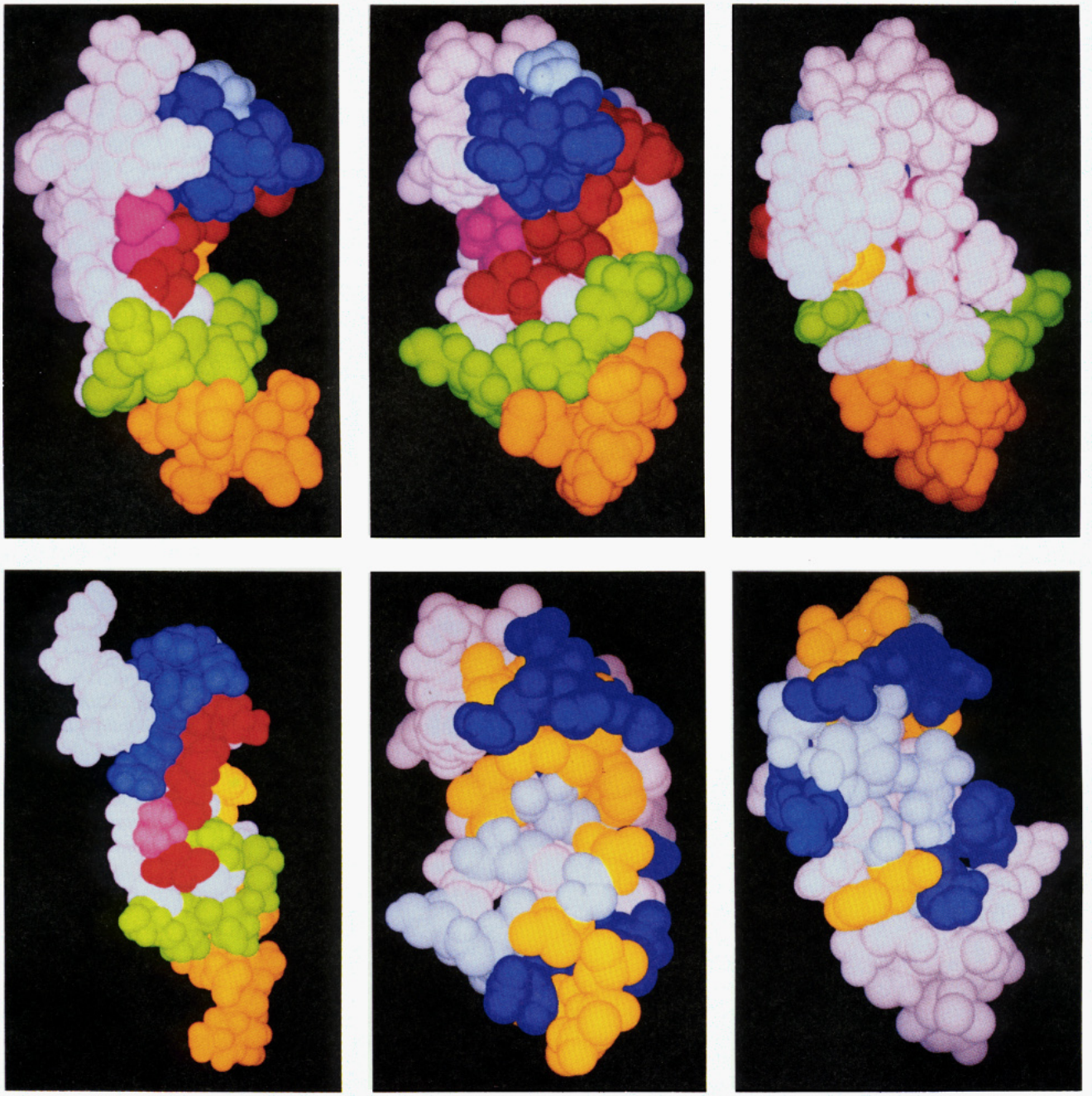

FIGURE 4: Color-coded representations of the three-dimensional solution structure of hTGFα, showing the location of the four heptapeptide sequences recognized by solubilized EGF-R and the nature of the proposed receptor-binding cavity of the protein ligand. In (A)–(D), all atoms in the protein, including hydrogens bound to carbon atoms, are shown in the following colors: blue, residues 22–27; red, 29–32 and 34; green, 36–42; orange, 44–50; aqua, Asp-28, which is contained in two of the bound heptapeptide sequences; purple, Val-33; yellow, Phe-15, which appears important for receptor binding and/or activation (Defeo-Jones et al., 1988). (A, top left) The most compact tertiary hTGFα structure, viewed from the side of the proposed receptor-binding cavity. (B, top center) The identical hTGFα structure, looking directly into the proposed receptor-binding cavity. (C, top right) Rear view of the hTGFα structure shown in (B). (D, bottom left) Location of the four heptapeptide sequences upon a representation of the most extended hTGFα conformation, shown in an orientation identical to that in (A), revealing the location of the receptor-binding sequences on a single fact of the protein. In (E) and (F), protein atoms, excluding hydrogens bound to carbon atoms, are shown in the following colors: dark blue, hydrophilic residues; yellow, hydrophobic residues; aqua, amphiphilic residues. (E, bottom center) Representation of the hydrophobic character of residues defining the potential EGF-R binding cavity in the identical structure and orientation to those shown in (B). Note the hydrophobic nature of the interior of the cavity and the hydrophilic residues defining the edges. (F, bottom right) Representation of the hydrophobic character of residues on the rear face of the protein in an orientation identical to that shown in (C).

was also of lowest potential energy (Figure 4A–C). We note that sequences corresponding to the EGF-R-binding peptides defined a distinct cavity whose chemical and structural features were consistent with its role as an EGF-R-binding site. For example, charged residues were predominantly located about its periphery, while the interior was comprised of either hydrophobic or amphiphilic residues (Figure 4E,F). Such a pattern is commonly found in protein, and carbohydrate, receptor-binding sites where initial receptor contact is mediated by charge–charge interactions on the surface, followed by burial of the hydrophobic core within the receptor (Lemieux, 1982). Furthermore, the aromatic residue Phe-15, which is highly conserved throughout the EGF/TGFα protein family, was found to lie in a central position in the proposed receptor-

Table I: Energies and Main-Chain Torsion Angles for Val-33 Determined from Six Solution-Phase hTGFα Structures Consistent with ^{1}H NMR Data

hTGFα structure no.	Val-33 main-chain torsion angles		AMBER energy (kJ/mol)
	φ	ψ	
1	−115.9	122.8	−1782.5
2	−132.3	95.6	−1823.4
3	−138.8	80.7	−1776.4
4	−120.5	87.1	−1952.9
5	−134.9	109.3	−1482.8
6	−135.6	127.7	−1827.4

binding cavity. This observation was also consistent with the results of mutagenesis experiments indicating the importance of Phe-15 in receptor binding (Defeo-Jones et al., 1988).

DISCUSSION

Identification of the regions of biologically active polypeptides which bind to specific cellular receptors often forms the basis of a molecular description of the ligand–receptor interaction and aids the subsequent development of compounds with potential therapeutic value. We have explored the usefulness of peptide mapping techniques (Geysen et al., 1987) as a rapid screen for putative receptor-binding sequences of a protein ligand using the well-characterized EGF receptor and overlapping heptapeptides encompassing the entire primary sequence of hTGFα. This system was chosen since a variety of independent experimental approaches had already been used to identify putative receptor-binding residues in hTGFα and a three-dimensional structure of hTGFα in aqueous solution was available (Kline et al., 1990). An important advantage of our procedure, as compared with conventional peptide synthesis, and which is shared by other recent methods for generating peptide epitopes (Fodor et al., 1991; Smith, 1991; Lam et al., 1991, Houghten et al., 1991), is that hundreds of peptides can be simultaneously synthesized, potentially permitting definition of receptor-binding sites to high resolution.

Four peptides which appeared to bind EGF-R in a *sequence-specific* manner were identified using the immobilized overlapping heptapeptides in hTGFα, in spite of potential entropy-based problems associated with constraining linear heptapeptides into a conformation identical to the cognate sequence in intact, folded hTGFα. A possible explanation for this apparent success might be associated with observations that the binding affinity of parts of a ligand can be increased by several orders of magnitude when presented as a copolymer (Spaltenstein & Whitesides, 1991). Unfortunately, relatively small amounts of mEGF failed to block binding of EGF-R to these heptapeptides. On the other hand, binding of EGF-R to the four solid-phase hTGFα heptapeptides was indicated as being sequence specific through the use of "mutated" and randomized peptide variants *and* to be dependent on the presence of residues which not only corresponded to highly conserved amino acids in the EGF/TGFα family but which also had been shown to be essential for receptor-binding activity by site-directed mutagenesis experiments. Several difficulties, however, are associated with simple interpretations of the results of the mEGF blocking experiments. First, enhancement of ligand-binding affinity of EGF-R may have resulted from its treatment with EGF. This phenomenon has been described for cytoskeletally associated EGF-R (Roy et al., 1991) and the recombinant extracellular domain of EGF-R (Hurwitz et al., 1991), and we have observed similar effects due to mEGF using our own detergent-solubilized EGF-R preparations (unpublished observations) (Richter, 1991). Alternatively, the failure of mEGF to block binding of EGF-R to the immobilized hTGFα heptapeptides may reflect differences in the binding of EGF and hTGFα to EGF-R (Winkler et al., 1989). This may also underly the observed enhancement of the interaction of EGF-R with the immobilized heptapeptide corresponding to residues 22–28 of hTGFα upon the addition of mEGF.[4]

Localization of the sequences corresponding to the four heptapeptides bound by EGF-R on solution-phase hTGFα structures (Figure 4) identified a cavity comprised of amino acids from both the N- and C-terminal domains, with the potential ability to interact with, and bind to, EGF-R. For example, not only was this structural feature constructed mainly from highly conserved residues in the B- and C-loops, which had been shown to be important from experiments using mutants of TGFα and EGF, but the outer residues possessed hydrophilic character while hydrophobic residues were located more centrally. Finally, Phe-15 and Leu-48, which also appear critical to biological activity, lay within this potential EGF-R binding cavity. Such a model of the receptor-binding region is therefore consistent both with recent studies using antibodies that implicate the B-loop of EGF in receptor binding (Katsuura & Tanaka, 1989) and with a "mitten" conformation recently proposed to be the active form of hTGFα (Kohda et al., 1989).

A feature of this model for the interaction of hTGFα with EGF-R is that binding involves regions in *both* the N- and C-terminal domains, which are linked by Val-33. Experimental evidence for the importance of Val-33 has been provided by the observation of a marked loss in the biological activity of hTGFα when this residue was replaced by its enantiomer (Tam et al., 1990). Given the constraints upon the secondary and tertiary structures imposed by the three disulfide bonds, this "mutation" might not have been expected to affect the gross overall folding of either of the two protein domains, although little structural information was presented. In accord with the concept that Val-33 might function as a molecular hinge, controlling the relative orientation of the N- and C-domains, flexibility about this residue was indicated by the fact that six solution structures for hTGFα are consistent with observed NOE data (Figure 4 and Table I) (Kline et al., 1990). It is therefore possible that variation in the main-chain torsion angles of Val-33 might alter the position of the C-domain relative to the β-sheet of the B-loop in the N-domain, facilitating positioning of the two hTGFα domains upon the EGF-R as part of the underlying mechanism of receptor dimerization and cell signaling (Canals, 1992; Schlessinger, 1988; Yarden & Schlessinger, 1987). Given the essential role of Leu-48, in the biological activity of hTGFα, and that the highly flexible C-terminal segment becomes structured upon interaction with lipid (Kohda & Inagaki, 1992), we postulate that this region of the C-domain is involved in the initial stages of receptor–ligand interaction. Subsequent modification of the main-chain torsion angles of Val-33 might then "clamp" the N-domain onto the receptor inducing a conformational change. Further evidence for conformational

[4] An alternate approach to settling the question of binding specificity might be to determine whether the free heptapeptides compete with ^{125}I-mEGF for binding to EGF-R. However, no such competition has been reported for any linear peptide comprising only seven residues in experiments with solubilized EGF-R. In unpublished experiments we have also prepared the immobilized complex of mEGF and EGF-R but have been unable to displace bound ^{125}I-mEGF even with intact protein ligands. Hence, if linear heptapeptides were used in a similar experiment, any absence of competition with mEGF would not unambiguously show that the observations reported here arose from nonspecific binding.

reorganization in hTGFα has also come from independent NMR investigations of its three-dimensional structure (Tappin et al., 1989; Harvey et al., 1991). Such a multidomain model explains the failure of short peptides to mimic adequately the biological activity of the intact hTGFα protein (Darlak et al., 1988; Tam et al., 1991). We also note that although the "hinge" residue is not conserved across the whole family of EGF-R ligands, all known TGFα sequences possess valine at position 33, while every known EGF protein has an asparagine residue at the homologous position (Asn-32). Such a structural difference might introduce variation into the conformational flexibility of the two families of proteins and hence account for the observation that EGF and TGFα appear to bind differently to EGF-R (Winkler et al., 1989). Other, more recent, work using site-directed mutants of the conserved residue Tyr-38 in hTGFα has provided further evidence for a dissimilar interaction between EGF-R and these two protein ligands (Engler et al., 1991).

Overall, this multidomain model provides a simple, molecular mechanism whereby binding of hTGFα to EGF-R can trigger a conformational change in the ligand which may, in turn, alter the conformation of the receptor as part of the processes involved in generation of any mitogenic signal, in accord with recent experimental evidence (Greenfield et al., 1989). We are currently pursuing site-directed mutagenesis of the hinge residue in intact hTGFα so as to test this hypothesis, and our results will be reported in due course.

ACKNOWLEDGMENT

We thank Dr. T.-P. Kline and co-workers (SKB) for the generous provision of coordinate data for hTGFα. We also acknowledge Dr. Dan Purich for critical comments concerning the manuscript.

REFERENCES

Atherton, E., & Sheppard, R. C. (1989) *Solid-Phase Peptide Synthesis; A Practical Approach*, IRL Press, Oxford.
Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T., & Tasumi, M. (1977) *J. Mol. Biol. 112*, 535–542.
Brown, F. K., Hempel, J. C., Dixon, J. S., Amato, S., Mueller, L., & Jeffs, P. W. (1989) *J. Am. Chem. Soc. 111*, 7328–7333.
Brünger, A. T., Clore, G. M., Gronenborn, A. M., & Karplus, M. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 3801–3805.
Campbell, I. D., Cooke, R. M., Baron, M., Harvey, T. S., & Tappin, M. J. (1989) *Prog. Growth Factor Res. 1*, 13–22.
Campbell, I. D., Baron, M., Cooke, R. M., Dudgeon, T. J., Fallon, A., Harvey, T. S., & Tappin, M. J. (1990) *Biochem. Pharmacol.* 40, 35–40.
Campion, S. R., Matsunami, R. K., Engler, D. A., & Niyogi, S. K. (1990) *Biochemistry 29*, 9988–9993.
Canals, F. (1992) *Biochemistry 31*, 4493–4501.
Carpenter, G., & Cohen, S. (1990) *J. Biol. Chem. 265*, 7709–7712.
Carpino, L. A., & Han, G. Y. (1972) *J. Org. Chem. 37*, 3404.
Cohen, S., Ushiro, H., Stoscheck, C., & Chinkers, M. (1982) *J. Biol. Chem. 257*, 1523–1531.
Conlan, J. W., Clarke, I. N., & Ward, M. E. (1988) *Mol. Microbiol. 2*, 673–679.
Cooke, R. M., Wikinson, A. J., Baron, M., Pastore, A., Tappin, M. J., Campbell, I. D., Gregory, H., & Sheard, B. (1987) *Nature 327*, 339–341.
Cooke, R. M., Tappin, M. J., Campbell, I. D., Kohda, D., Miyake, T., Fuwa, T., Miyazawa, T., & Inagaki, F. (1990) *Eur. J. Biochem. 193*, 807–817.
Cowley, G. P., Smith, J. A. & Gusterson, B. A. (1986) *Br. J. Cancer 53*, 223–229.
Darlak, K., Franklin, G., Woost, P., Soonenfeld, E., Twardzik, D., Spatola, A., & Schultz, G. (1988) *J. Cell. Biochem. 36*, 341–352.
Defeo-Jones, D., Tai, J. Y., Wegrzyn, R. J., Vuocolo, G. A., Baker, A. E., Payne, L. S., Garsky, V. M., Oliff, A., & Rieman, M. W. (1988) *Mol. Cell. Biol. 8*, 2999–3007.
Defeo-Jones, D., Tai, J. Y., Vuocolo, G. A., Wegrzyn, R. J., Schofield, T. L., Oliff, A., & Rieman, M. W. (1989) *Mol. Cell. Biol. 9*, 4083–4086.
Derynck, R. (1988) *Cell 54*, 593–595.
Dudgeon, T. J., Baron, M., Cooke, R. M., Fallon, A., & Campbell, I. D. (1990) *FEBS Lett. 261*, 392–396.
Engler, D. A., Matsunami, R. K., Campion, S. R., Stringer, C. D., Stevens, A., & Niyogi, S. K. (1988) *J. Biol. Chem. 263*, 12384–12390.
Engler, D. A., Montelione, G. T., & Niyogi, S. K. (1990) *FEBS Lett. 271*, 47–50.
Engler, D. A., Hauser, M. R., Cook, J. S., & Niyogi, S. K. (1991) *Mol. Cell. Biol. 11*, 2425–2431.
Eppstein, D. A., Marsh, Y. V., Schryver, B. B., & Bertics, P. J. (1989) *J. Cell. Physiol. 141*, 420–430.
Fesik, S. W., & Zuiderweg, E. R. P. (1990) *Q. Rev. Biophys. 23*, 97–131.
Fodor, S. P. A., Read, J. L., Pirrung, M. C., Stryer, L., Lu, A. T., & Solas, D. (1991) *Science 251*, 767–773.
Geysen, H. M., Rodda, S. J., Mason, T. J., Tribbick, G., & Schoofs, P. G. (1987) *J. Immunol. Methods 102*, 259–274.
Greenfield, C., Hiles, I., Waterfield, M. D., Federwisch, M., Wollmer, A., Blundell, T. L., & McDonald, N. (1988) *EMBO J. 8*, 4115–4123.
Gregory, H., Thomas, C. E., Young, J. A., Willshire, I. R., & Garner, A. (1988) *Regul. Pept. 22*, 217–226.
Han, K.-H., Ferretti, J. A., Niu, C.-H., Lokeshwar, V., Clarke, R., & Katz, D. (1988) *J. Mol. Recognit. 1*, 116–123.
Harvey, T. S., Wilkinson, A. J., Tappin, M. J., Cooke, R. M., & Campbell, I. D. (1991) *Eur. J. Biochem. 198*, 555–562.
Heath, W. F., & Merrifield, R. B. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 6367–6371.
Hempel, J. C., & Brown, F. K. (1989) *J. Am. Chem. Soc. 111*, 7323–7327.
Higashiyama, S., Abraham, J. A., Miller, J., Fiddes, J. C., & Klagsbrun, M. (1991) *Science 251*, 936–939.
Hoeprich, P. D., Langton, B. C., Zhang, J., & Tam, J. P. (1989) *J. Biol. Chem. 264*, 19086–19091.
Hommel, U., Dudgeon, T. J., Fallon, A., Edwards, R. M., & Campbell, I. D. (1991) *Biochemistry 30*, 8891–8898.
Houghten, R. A., Pinilla, C., Blondelle, S. E., Appel, J. R., Dooley, C. T., & Cuervo, J. H. (1991) *Nature 354*, 84–86.
Hurwitz, D. R., Emanuel, S. L., Nathan, M. H., Sarver, M., Ullrich, A., Felder, S., Lax, I., & Schlessinger, J. (1991) *J. Biol. Chem. 266*, 22035–22043.
Kaptein, R., Boelens, R., Scheek, R. M., & van Gunsteren, W. F. (1988) *Biochemistry 27*, 5389–5395.
Katsuura, M., & Tanaka, S. J. (1989) *J. Biochem. (Tokyo) 106*, 87–92.
Kline, T.-P., Brown, F. K., Brown, S. C., Jeffs, P. W., Kopple, K. D., & Mueller, L. (1990) *Biochemistry 29*, 7805–7813.
Kohda, D., & Inagaki, F. (1992) *Biochemistry 31*, 677–685.
Kohda, D., Shimada, I., Miyake, T., Fuwa, T., & Inagaki, F. (1989) *Biochemistry 28*, 953–958.
Komoriya, A., Hortsch, M., Meyers, C., Smith, M., Kanety, H., & Schlessinger, J. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 1351–1355.
Korc, M., Chandrasekar, B., & Shah, G. N. (1991) *Cancer Res. 51*, 6243–6249.
Lam, K. S., Salmon, S. E., Hersh, E. M., Hruby, V. J., Kazmierski, W. M., & Knapp, R. J. (1991) *Nature 354*, 82–84.
Lax, I., Fischer, R., Ng, C., Segre, J., Ullrich, A., Givol, D., & Schlessinger, J. (1991) *Cell Regul. 2*, 337–345.
Lazar, E., Watanabe, S., Dalton, S., & Sporn, M. B. (1988) *Mol. Cell. Biol. 8*, 1247–1252.

Lazar, E., Vicenzi, E., Van Oberghen-Schilling, E., Wolff, B., Dalton, S., Watanabe, S., & Sporn, M. B. (1989) *Mol. Cell. Biol. 9*, 860–864.
Lemieux, R. E. (1982) in *Frontiers of Chemistry* (Laidler, K. J., Ed.) Pergamon Press, Oxford.
Massagué, J. (1990) *J. Biol. Chem. 265*, 21393–21396.
Matsunami, R. K., Campion, S. R., Niyogi, S. K., & Stevens, A. (1990) *FEBS Lett. 264*, 105–108.
Matsunami, R. K., Yette, M. L., Stevens, A., & Niyogi, S. K. (1991) *J. Cell. Biochem. 46*, 242–249.
Mohamadi, F. M., Richards, N. G. J., Guida, W. C., Liskamp, R. M. J., Lipton, M. A., Caufield, C. E., Chang, C., Hendrickson, T., & Still, W. C. (1990) *J. Comput. Chem. 11*, 440–467.
Montelione, G. T., Wüthrich, K., Nice, E. C., Burgess, A. W., & Scheraga, H. A. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 5226–5230.
Nelson, J., Stewart, R., McGivern, M., Bailie, J. R., Walker, B., Murphy, R. F., & Wilson, D. J. (1991) *Carcinogenesis 12*, 1823–1829.
Nestor, J. J., Newman, S. R., DeLustro, B., Todaro, G. J., & Schreiber, A. B. (1985) *Biochem. Biophys. Res. Commun. 129*, 226–232.
Panayotou, G., End, P., Aumailley, M., Timpl, R., & Engel, J. (1989) *Cell 56*, 93–101.
Richter, A. (1991) Ph.D. Thesis, University of Southampton, U.K.
Roy, L. M., Gittinger, C. K., & Landreth, G. E. (1991) *J. Cell. Physiol. 146*, 63–72.
Schlessinger, J. (1988) *Trends Biochem. Sci. 13*, 443–447.
Schulz, G. E., & Schirmer, R. E. (1979) *Principles of Protein Structure*, Springer-Verlag, New York.
Shoyab, M., Plowman, G. D., McDonald, V. L., Bradley, J. G., & Todaro, G. J. (1989) *Science 243*, 1074–1076.
Smith, G. P. (1991) *Curr. Opin. Biotechnol. 2*, 668–673.
Spaltenstein, H., & Whitesides, G. M. (1991) *J. Am. Chem. Soc. 113*, 686–687.
Stewart, J. M., & Young, J. D. (1984) *Solid-Phase Peptide Synthesis*, Pierce Chemical Co., Rockford, IL.
Tam, J. P., Lin, Y.-Z., Wu, C.-R., Shen, Z.-Y., Galantino, M., Liu, W., & Ke, X.-H. (1990) in *Proceedings of the 11th American Peptide Symposium* (Rivier, J., & Marshall, G., Eds.) pp 75–77, ESCOM, Leiden.
Tam, J. P., Lin, Y.-Z., Liu, W., Wang, D.-X., Ke, X.-H., & Zhang, J.-W. (1991) *Int. J. Pept. Protein Res. 38*, 204–211.
Tappin, M. J., Cooke, R. M., Fitton, J. E., & Campbell, I. D. (1989) *Eur. J. Biochem. 179*, 629–637.
Ullrich, A., & Schlessinger, J. (1990) *Cell 61*, 203–212.
Waterfield, M. D., Mayes, E. L. V., Stroobant, P., Bennet, P. L. P., Young, S., Goodfellow, P. N., Banting, G. S., & Ozanne, B. (1984) *J. Cell. Biochem. 20*, 149–161.
Weiner, S. J., Kollman, P. A., Case, D. A., Singh, U. C., Ghio, C., Alagona, G., Profeta, S., Jr., & Weiner, P. K. (1984) *J. Am. Chem. Soc. 106*, 765–784.
Weiner, S. J., Kollman, P. A., Nguyen, D. T., & Case, D. A. (1986) *J. Comput. Chem. 7*, 230–236.
Winkler, M. E., O'Connor, L., Winget, M., & Fendly, B. (1989) *Biochemistry 28*, 6373–6378.
Yarden, Y., & Schlessinger, J. (1987) *Biochemistry 26*, 1434–1442.