Structure-based Minimization of Transforming Growth Factor-α (TGF-α) through NMR Analysis of the Receptor-bound Ligand

DESIGN, SOLUTION STRUCTURE, AND ACTIVITY OF TGF-α 8–50*

(Received for publication, April 24, 1998, and in revised form, July 22, 1998)

Campbell McInnes‡, Jianjun Wang§§, Ala-Eddin Al Moustafañ, Cedric Yansouni†, Maureen O’Connor-McCourt¶, and Brian D. Sykes¶

From the ‡Protein Engineering Network of Centres of Excellence, 713 Heritage Medical Research Centre, University of Alberta, Edmonton, Alberta T6G 2S2, Canada and the ¶Biotechnology Research Institute, Montreal, Quebec H4P 2R2, Canada

The investigation of a N-terminally truncated human transforming growth factor-α (TGF-α; residues 8–50) has been completed to determine the contribution of the N terminus to receptor binding and activation. The deletion protein was proposed and designed through study of NMR relaxation and nuclear Overhauser enhancement data obtained from the TGF-α-epidermal growth factor (EGF) receptor complex, which indicated that the residues N-terminal to the A loop remain flexible in receptor-bound TGF-α and thus suggested their lack of involvement in receptor binding (Hoyt, D. W., Harkins, R. N., Pagila, R. N., Debanne, M. T., O’Connor-McCourt, M., and Sykes, B. D. (1994) Biochemistry 33, 15283–15292; McInnes, C., Hoyt, D. W., Harkins, R. N., Pagila, R. N., Debanne, M. T., O’Connor-McCourt, M., and Sykes, B. D. (1996) J. Biol. Chem. 271, 32204–32211). TGF-α 8–50 was shown to have approximately 10-fold lower affinity for the receptor than the native molecule in an assay quantifying the ability to compete with EGF for binding and to have a similar reduction in activity as indicated by a cell proliferation assay. NMR solution structural calculations on this molecule demonstrate correct formation of the three disulfide bonds of TGF-α 8–50 and have established the presence of native secondary structure in the B and C loops of the protein. However, some perturbation of the global fold with respect to the orientation of the subdomains was observed. These results suggest that although the N-terminal residues do not contribute directly to binding, they make a significant contribution in defining the conformation of the growth factor, which is required for complete binding and activity and is therefore significant in terms of producing native folding of TGF-α. They also show that information obtained from the receptor-bound ligand can be used to guide the design and minimization of TGF-α analogues. The implications of the study of TGF-α 8–50 for the design and synthesis of reductants of this growth factor are therefore discussed.

The binding of human TGF-α* to the EGF receptor initiates a number of cell proliferation events including wound healing and embryogenesis (1, 2). Furthermore, this mitogenic protein of 50 residues is involved in the transformation of normal cells into malignant growths (3–5) and is also believed to promote angiogenesis (6). It is thus apparent that this polypeptide is a significant target from a pharmaceutical and drug design perspective, and to this end, considerable effort has been made to elucidate the essential structural features of this tricyclic growth factor required for binding and function (7, 8). Many attempts have also been made to synthesize reductant molecules that display a similar biological profile to the native TGF-α, although for the most part these efforts have met with limited success (9–11). NMR structures of the free ligand have been determined by several groups (12–15), and recently studies have been published from this laboratory on the use of NMR relaxation and NOE measurements in determining the essential components of the TGF-α-EGFR extracellular domain complex (16, 17). The data from the latter experiments suggest that A and C loops and the C-terminal tail of TGF-α contain residues (for the sequence of TGF-α and location of disulfides forming the three loops, see Fig. 1) that form the major binding interface with the receptor and that the N-terminal amino acids outside the A loop remain flexible in the receptor-bound species. Since the consensus from the relaxation and NOE data was that the N terminus of TGF-α does not play a role in the receptor-ligand interaction, the N-terminal deletion mutant, TGF-α 8–50 was synthesized and characterized in terms of NMR structure and biological activity. These experiments were performed to clarify the requirements of the growth factor structure necessary for receptor binding and activation and to determine the contribution of the N-terminal tail to the formation of the global fold of the native molecule. Thus the rationale is to assess any functional changes of the truncated protein with respect to native TGF-α in terms of structural variation and concurrently in doing this to ascertain if the N-terminal tail residues are required for establishing the native conformation of the protein. The truncated TGF-α was also studied to demonstrate that structural information obtained from NMR experiments on the receptor-bound ligand can be used toward the design of a “minimized” TGF-α where nonessential regions of the protein are deleted while retaining near native levels of receptor binding and activation.

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ Present address: Dept. of Medical Biochemistry School of Medicine, Southern Illinois University, Carbondale, IL 62901-4413.

§ To whom correspondence should be addressed. Tel.: 403-492-6540; Fax: 403-492-1473; E-mail: bds@polaris.biochem.ualberta.ca.

¶ The abbreviations used are: TGF-α, transforming growth factor-α; EGF, epidermal growth factor; EGFR, epidermal growth factor receptor; NOE, nuclear Overhauser enhancement; NOESY, nuclear Overhauser enhancement spectroscopy; TOCSY, total correlation spectroscopy; HPLC, high pressure liquid chromatography; r.m.s., root mean square.
Solution Structure of a Truncated TGF-α

FIG. 1. Schematic representation of the primary and loop structure of TGF-α. The seven N-terminal residues removed in the deletion protein are indicated by shading as are the six cysteine pairs that form the disulfide loops. The disulfide bonds are represented as thick lines, and the loops formed by these are labeled as A, B, and C.

EXPERIMENTAL PROCEDURES

Synthesis and Purification of TGF-α 8–50

TGF-α 8–50 was synthesized using standard solid phase synthesis techniques on a Automated Biosystems model 450A peptide synthesizer (Foster City, CA) using t-butyloxycarbonyl methodology. The N terminus was acetylated using acetic anhydride, and the peptide was liberated from the resin using a mixture of hydrogen fluoride, anisole and 1,2-ethanediol. The crude material was extracted with glacial acetic acid, lyophilized, and then dissolved in 10% acetic acid. The peptide was then purified using reversed phase HPLC (Beckmann System Gold, fullerton, CA) using a Synchropak RP-18 (250 × 21.2 mm inner diameter) column (Synchrom, Lafayette, IN) at a flow rate of 5 ml/min and subsequently oxidized in air using a solution containing 1.0 M urea, 0.1 M Tris, 1.5 mM oxidized and 0.75 mM reduced glutathione, pH 8.0 (10). A final purification step was used to remove oxidation byproducts, and the final purity was confirmed through reverse phase analytical HPLC. Electrospray mass spectrometry (VG Biotech, Cheshire, UK) verified the presence of a species of the desired molecular weight.

NMR Sample Preparation

To a lyophilized sample of TGF-α was added 460 μl of buffer containing 50 mM potassium phosphate, 10 mM potassium chloride, 1 mM ethylene diamine tetracetic acid, pH 7.0. The solution was adjusted to pH 6.5 by the addition of small aliquots of 0.5 N NaOD or 0.5 N HCl bringing the final volume to 500 μl.

NMR Spectroscopy

One- and two-dimensional 1H NMR spectra for TGF-α 8–50 were collected on a Varian Unity spectrometer operating at 599.9 MHz. TOCSY (18) and NOESY (19) experiments were acquired at 298 K and referenced relative to an internal sodium 2,2-dimethyl-2-silapentane-5-sulfonate (internal standard), and ethylene diamine tetraacetic acid, 0.5 mM sodium azide, 0.15 M NaCl. The solution was adjusted to pH 6.5 by the addition of small aliquots of 0.5 N NaOD or 0.5 N HCl bringing the final volume to 500 μl.

Structure Calculation of TGF-α 8–50

For TGF-α 8–50, all cross-peaks were peak-integrated using the program VNR (VNR 5.1A, Varian Associates, Palo Alto, CA) of which a subset were assigned based on the chemical shift assignments and possible NOE distances in the native protein (Protein Data Bank code 1yug; Ref 15.) calculated using the programs NMRPipe (24) and PIPP (25) and a cut-off distance of 12 Å. The suite of programs CAMRA (26) was implemented to convert the volume integral information to distance restraints. NOE cross-peak intensities were classified as strong (upper boundary, 2.3 Å), medium (3.0 Å), weak (4.0 Å), and very weak (5.0 Å), and the appropriate pseudoatom correction was added. The molecular dynamics/simulated annealing protocol in X-PLOR version 3.8 (27, 28) was utilized to calculate an initial set of structures derived only from the unambiguous distance restraints and native disulfides assumed in the starting structure. A program developed in our laboratory was then applied in conjunction with X-PLOR to modify NOE distances causing restraint violations and also to make use of those that have ambiguous assignments to generate a more refined structure. This program uses an approach similar to that of Nilges et al. (29) in that it allows the use of NOE distances that are ambiguous, because in most cases only one restraint will be possible based on distances from the ensemble of structures already calculated using unambiguous restraints. The program therefore removes restraints that are violated by greater than a user defined cut-off distance (0.5 Å) because the nonpossible assignments for the ambiguous restraints will give distance violations greater than the cut-off. The restraints that satisfy the distances in the initial structures will be retained. Because it is possible that a correct restraint can be violated in initial but not final structures, NOEs that are removed can be returned to the restraint file in later runs for restraint calculation. As the structures are refined in successive runs, these restraints should now be satisfied and are incorporated in the final data set.

This procedure as a whole allows the restraints that satisfy the distances in the successive structural ensembles to be retained and thus facilitates the incorporation of a greater number of restraints in the X-PLOR calculation. For the distance restraints that are consistently

| Residue | H2 | H3 | H4 | Other protons |
|---------|----|----|----|--------------|
| Cys8    | 8.50 | 4.40 | 3.06 | 2.97 | CH3CO2H 1.83 |
| Pro9    | 8.44 | 4.44 | 2.21 | 1.94 | H2O 1.72, 1.72, H3O 3.24 |
| Asp10   | 8.48 | 4.44 | 2.69 | 2.69 |
| Ser11   | 8.19 | 4.25 | 3.82 | 3.82 |
| His12   | 8.67 | 4.45 | 2.72 | 2.25 |
| Thr13   | 7.79 | 4.21 | 4.15 | H 1.19 |
| Gln14   | 8.40 | 4.21 | 1.99 | 1.81 | H2O 2.13, 2.13, H'2O 7.36, 6.75 |
| Phe15   | 7.98 | 4.20 | 3.06 | 2.85 | H3O 7.17, H'3O 7.18 |
| Cys16   | 8.21 | 4.33 | 2.46 | 2.04 |
| Phe17   | 8.33 | 4.22 | 2.69 | 2.69 |
| His18   | 6.87 | 4.45 | 2.72 | 2.25 |
| Gly19   | 7.35 | 3.90 | 3.90 |
| Thr20   | 8.41 | 4.51 | 3.95 | H 1.27 |
| Arg21   | 8.99 | 5.14 | 3.25 | 3.14 |
| Arg22   | 9.43 | 4.61 | 1.72 | 1.54 |
| Phe23   | 8.93 | 4.59 | 3.01 | 2.86 | H3O 6.88, H3O 7.16, H'7.24 |
| Leu24   | 8.28 | 4.43 | 1.77 | 1.46 | H'2O 0.76, 0.72 |
| Val25   | 8.15 | 3.68 | 2.08 | H 1.12, 1.00 |
| Glu26   | 8.74 | 4.10 | 2.08 | 2.08 |
| Gln27   | 7.70 | 4.21 | 1.73 | H'2O 2.19, 2.06 |
| Asp28   | 7.98 | 4.28 | 3.07 | 2.28 |
| Leu29   | 8.05 | 4.72 | 1.74 | 1.58 | H2O 1.24, 1.24, H' 2.95, 2.95 |
| Pro30   | 7.81 | 4.76 | 1.76 | 1.66 | H2O 1.45, 1.45, H' 3.54, 3.43 |
| Ala31   | 8.77 | 4.73 | 1.37 |
| Cys32   | 8.78 | 5.20 | 2.77 | 2.62 |
| Arg33   | 9.04 | 4.23 | 2.05 | H2O 0.95, 0.92 |
| Cys34   | 9.13 | 5.06 | 3.44 | 2.74 |
| His35   | 8.72 | 4.87 | 3.40 | 2.88 |
| Ser36   | 8.66 | 4.37 | 2.92 | 3.92 |
| Gly37   | 8.82 | 4.15 | 3.56 |
| Tyr38   | 8.02 | 5.28 | 2.92 | 2.77 | H3O 6.67, H3O 6.40 |
| Leu39   | 9.26 | 4.85 | 2.92 | 2.44 |
| Gly40   | 8.10 | 5.04 | 3.90 |
| Ala41   | 9.24 | 4.18 | 1.56 |
| Arg42   | 8.83 | 4.71 | 2.26 | 2.26 |
| Cys43   | 8.39 | 3.96 | 2.92 | 2.44 |
| Phe44   | 10.05 | 4.25 | 1.62 | 1.62 | H2O 1.83, 1.37 |
| His45   | 8.55 | 5.13 | 3.08 | 2.99 | H2O 8.10, H' 7.03 |
| Ala46   | 8.73 | 4.17 | 1.15 |
| Leu47   | 8.12 | 4.42 | 2.48 | 2.21 |
| Leu48   | 8.07 | 4.25 | 1.58 | 1.58 |
| His49   | 8.19 | 4.31 | 1.64 | 1.57 |
| Ala50   | 7.59 | 4.05 | 1.28 |
violated by less than the specified cut-off in the calculated structures, the program automatically modifies the upper bound of distance restraints by the average value of the distance violation for a user specified number of structures in which the violation occurs. The modified restraint file is then used in the next round of X-PLOR calculations in an iterative procedure until an structural ensemble is obtained of the desired quality. The reproducibility and accuracy of the ensemble generated for TGF-α 8–50 was examined using the program PROCHECK (30), which checks the distribution of backbone dihedral angles in the structures.

**Binding and Activity of TGF-α 8–50**

**Cell Culture**—The HBE4-EGF7 cell line was used and was established by transfection of primary human bronchial epithelial cells with a plasmid construct expressing the human papillomavirus type 16 E6 and E7 genes (31). The cell line was maintained in the Keratinocyte-SF (KSF) medium without supplements and incubated at 37 °C in an atmosphere containing 5% CO₂.

**Proliferation Assay**—Growth curves were established by measuring the incorporation of [³H]methylthymidine into the acid insoluble fraction of cellular extracts. Briefly, approximately 2 x 10⁴ cells were plated in 1 ml KSF into each well of 12-well tissue culture plates. After 1 day, the medium was changed either with fresh KSF (control) or medium containing TGF-α or TGF-α 8–50. Cells were allowed to grow for 3 days with [³H]methylthymidine (1 μCi/ml, Amersham Canada) being added during the last 20–24 h. After sequential and multiple washings in cold phosphate-buffered saline, 10% trichloroacetic acid, and 95% ethanol, the plates were allowed to dry, the cells were solubilized in 2% SDS, and the radioactivity was quantitated by liquid scintillation chromatography. The experimental points were determined in triplicate.

**EGF Competition Assay**—Cells (2 x 10⁴) were plated in 24-well plates in KSF and grown to confluence. 24 h before performing the experiment, the medium was changed with KSF. Cells were washed three times with binding buffer containing Dulbecco’s modified Eagle’s medium, 25 mM Heps, 5 mM MgCl₂, 100 μg/ml bacitracin, 0.1% bovine serum albumin, pH 7.4, and incubated in 0.5 ml of binding buffer containing various concentrations of TGF-α/TGF-α 8–50 with 0.05 mM ³¹²EFG at 0 C for 2 h. After incubation, the cells were washed twice with ice-cold binding buffer and three times with ice-cold phosphate-buffered saline. Cells were solubilized with 1 N NaOH and counted in a γ counter. The nonspecific binding was determined in the presence of 1 μM unlabeled TGF-α or TGF-α 8–50. Data were analyzed using the method of Scatchard.

**RESULTS**

**Analysis of Chemical Shift Variations between Truncated and Native TGF-α**—The deletion of the N-terminal tail of TGF-α raises the question of whether or not native disulfide bonds are present in the truncated form of the protein and if it is able to assume a similar global fold to the intact protein. For the purpose of determining whether the correct pairing of the disulfides in TGF-α 8–50 is present after synthesis and oxidation, the chemical shifts of the α-protons were compared with those of the native protein acquired under similar conditions, i.e. pH 6.5 and 30 °C. The δα values of the comparison (chemical shift differences) are illustrated in Fig. 2, and it can be seen that there are no large deviations in these shifts. This suggests that the three disulfide links between Cys⁸ and Cys¹, Cys¹⁶ and Cys²², and Cys³⁴ and Cys⁴⁴ are present as observed in the intact TGF-α. The presence of the native disulfide linkages was confirmed through observation of NOEs between cysteines 8 and 21 (Fig. 3) and 34 and 43 (not shown). Also from the absence of major deviations the assumption can be made that a structure cognate to the native molecule exists. However, despite the lack of large differences in the chemical shifts of the deletion mutant, significant changes of between 0.05 and 0.2 ppm were observed. These occur for the most part in the A and B loops and indicate that deletion of the residues N-terminal to the A loop effects changes in the secondary structure and perhaps in the overall fold of the molecule. In particular, the residues undergoing the largest variation are those of the A loop including Tyr¹³, Gln¹⁴, and Phe¹⁵, which have resonances differing from the native protein by 0.17, 0.13, and 0.13 ppm respectively. Of the B loop α-protons, those that exhibited the most significant changes were Gly¹⁹ and Thr²⁰ in addition to the α-resonances of all the residues between Val²³ and Lys²⁹. For the hinge region of TGF-α 8–50 (residues close to Val²³), which allows flexibility between the two subdomains (A and B loops comprise the N subdomain, whereas the C loop and tail form the C subdomain), δα values of between 0.04 and 0.09 ppm were observed for the residues between Cys²⁷ and Ser³⁶ and thus are indicative of a possible variation in the conformation of this region. An important observation of the shifts for the residues of the C loop is that with the exception of Gly⁴⁰ and Glu⁴⁴ there are no differences larger than 0.05 ppm, suggesting that the reverse double hairpin secondary structure is present in the truncated protein. The δα α-H value for Glu⁴⁴ of 0.08 ppm is of special interest because this residue forms interactions critical to the stabilization of the orientation of the N- and C-terminal subdomains.

**Receptor Binding and Activation by TGF-α 8–50**—Investigation of the biological potency of TGF-α 8–50 was performed to assess the contribution of the N-terminal tail to receptor bind-
Non-allowed regions, which are Phe15 and Cys16, are located in residues excluding glycine, proline, and end residues. Analysis of the average of the 35 structures fit to the experimental data, and in terms of the component and for the ensemble of 35 structures are shown in Table II and make the best use of ambiguous NOEs. The structure statistics calculated for the N-terminally deleted protein using NOE-competition assay are shown in Fig. 4. From the plot it is apparent that it requires approximately 10-fold more TGF-8–50 than EGF to displace iodinated EGF from HBE cells than it does for native TGF-α, thus indicating that TGF-α exhibits similar structural characteristics in comparison with the global fold of the native molecule. Our results indicate that although the binding and activation are diminished 10-fold, TGF-α 8–50 still retains a significant level of potency in terms of competition with EGF for the receptor and ability to effect cell proliferation. The plethora of studies on synthetic fragments of TGF-α have for the most part failed to exhibit biological activity close to the order of that observed for the native molecule. Our results differ slightly from those of Tam et al. (10), who previously synthesized TGF-α 8–50 and reported binding and mitogenic activity of 3%. Our data imply that TGF-α 8–50 is more potent; however, this variation may be the result of different methods used for assaying activity. In our case, the protein synthesized in this laboratory was fully characterized by NMR and thus the possibility of ambiguity in the structure and correct folding of the disulfides is not present, and therefore the biological data remain in the structural ensemble of the truncated molecule since the B loop β-sheets are present in TGF-α and native TGF-α as shown by backbone r.m.s. deviations from ideal geometry (Å) and 1.304 for the heavy atom r.m.s. deviations, which are very favorable in their comparability with the ensemble recently generated for TGF-α using similar conditions for data acquisition (15).

**DISCUSSION**

**Biological Potency of TGF-α 8–50**—TGF-α 8–50 was proposed and designed based on the results of the relaxation and NOE analysis data of the ligand-receptor complex, which suggested that the N terminus does not contribute to the complexion process and that the truncated protein should retain the ability to associate with the EGF receptor and initiate cell proliferation. **Table II**: Statistics for ensemble of NMR derived structures of TGF-α 8–50

| Statistics for ensemble of NMR derived structures of TGF-α 8–50 |
|---------------------------------|
| **XPLOR energies (kcal/mol)** |
| Total energy                   | 73.5 ± 0.326 |
| Bond energy                    | 3.12 ± 0.018 |
| Angle energy                   | 57.0 ± 0.167 |
| Improper energy                | 5.01 ± 0.029 |
| van der Waals energy           | 3.72 ± 0.057 |
| NOE energy                     | 4.64 ± 0.058 |

| r.m.s. deviations from ideal geometry (Å) |
|------------------------------------------|
| Bond                                     | 0.00224 ± 0.00001 |
| Angle                                    | 0.580 ± 0.00116 |
| Improper                                 | 0.309 ± 0.00109 |

| r.m.s. deviations from experimental restraints (Å) |
|--------------------------------------------------|
| Total NOE restraints (409)                        | 0.0147 ± 0.00010 |
| Intra-residue (154)                              | 0.0116 ± 0.00062 |
| Sequential (126)                                 | 0.00730 ± 0.00036 |
| Medium range (33)                                | 0.0181 ± 0.0043 |
| Long range (94)                                  | 0.0113 ± 0.0029 |

| r.m.s. deviations of cartesian coordinates ($) |
|------------------------------------------------|
| Residues 15–47 (backbone)                       | 0.844 ± 0.119 |
| Residues 15–47 (heavy)                          | 1.304 ± 0.144 |
| Residues 15–47 (all)                            | 1.477 ± 0.135 |
| Residues 19–24, 29–34, 38–46 (backbone)         | 0.654 ± 0.095 |
| Residues 19–24, 29–34, 38–46 (heavy)            | 1.083 ± 0.131 |
| Residues 19–24, 29–34, 38–46 (all)              | 1.304 ± 0.126 |

**Solution Structure of TGF-α 8–50**—The NMR-derived solution structure of TGF-α 8–50 exhibits similar structural characteristics in comparison with the global fold of the native protein as has been determined by several groups to date (12–15). The overall topology of the truncated protein superimposes favorably to that of the intact TGF-α as shown by backbone r.m.s. deviations of 2.78 Å for residues 15–47. The identity between the two structures is demonstrated in Fig. 6, which illustrates the Richardson diagrams of the native and N-terminally deleted TGF-α. The secondary structure elements present in TGF-α remain in the structural ensemble of the N-terminally truncated molecule since the B loop β-sheet comprised of residues 19–24 and 29–34 and the C-terminal

Figure 4: Plot of data from the competition assay of TGF-α and TGF-α 8–50. The graph illustrates the percentage of [125I]-EGF displaced from the EGF receptor on Human BE cells as a function of ligand concentration.
reverse double hairpin loop between residues 38 and 46 were observed. The partial helix present in some of the TGF-α conformations is no longer present in TGF-α 8–50 because the A loop shows considerably flexibility. Some NOEs diagnostic of an α-helix are observed for the truncated protein; however, these are insufficient to restrain the A loop into a helical conformation. The conformational differences of the A loop as observed in the structural ensembles are supported by the α-H resonances, which show considerable variation when compared with the intact protein, and indeed as previously mentioned, the most significant changes in the resonance frequency of the α-H of the N-terminally deleted molecule occur in this region.

Upon closer examination of the NMR structure of TGF-α 8–50, significant differences in the conformations of the two
The truncated TGF-α cluster of Phe17, His18, and Tyr38 is more closely packed in that stabilize the orientation of the two regions. The hydrophobic clusters in the chemical shift data, which indicate that although notable variations in the N-terminally deleted by the calculated structures are corroborated by the partial helix present in the intact protein is no longer present in TGF-α. Further evidence for this structural change is given by the absence/disruption of the third strand of the β-sheet formed by the interactions of Phe3 and Asn6 of the N terminus with Phe23 and Leu24 of the second strand. Disruption of these interactions does not destabilize the B loop β-sheet; however, it does allow the turn greater freedom to adopt a dissimilar conformation.

The most significant and consequential variation in the structure of the two proteins is in the relative orientation of the C-terminal subdomains comprised of residues 34–50. As can be observed from the superposition of TGF-α 8–50 to TGF-α in Fig. 6, this region in the deletion protein is angled almost perpendicular to the N-terminal subdomain with the first strand of the reverse double hairpin folded away from the hinge region from Cys29 to Ser36. The differences in the arrangement of the C-terminal subdomain for the two molecules as illustrated by the calculated structures are corroborated by the chemical shift data, which indicate that although notable variations in the α-proton resonances occur in the region between Cys29 and Ser36 (Fig. 2), less significant variation in the α-H frequencies of the C loop and tail secondary structure are observed. The conformational transition of the C-terminal subdomain that occurs in the N-terminally deleted TGF-α is apparently a propagated effect and is the result of disruption of the interactions between the A and C loops. As previously mentioned, the partial helix present in the intact protein is no longer present in TGF-α 8–50. The consequence of deletion of the N-terminal tail is variation in the conformation of the A loop, which in turn results in altered interactions between the A and C loops. The specific interactions that differ from the native protein in defining the fold of the A and C subdomains in TGF-α 8–50 appear mostly to be in the packing of the residues that stabilize the orientation of the two regions. The hydrophobic cluster of Phe17, His18, and Tyr38 is more closely packed in the truncated TGF-α in addition to a more defined interaction between Phe15 and Arg24. The conformation of the interdomain residues His18 and His35 also varies in the N-terminally deleted molecule, and the association of the side chains of Glu44 and Phe17 is different for TGF-α 8–50, an observation that is confirmed by the chemical shift difference for Glu44 α-H of 0.08 ppm. These changes indicate the transition in domain structure between the two molecules that result from novel interactions between the residues of the A and C loops.

It is interesting that the new conformation of TGF-α 8–50 and orientation of the subdomains in the truncated molecule is of comparable if not better definition than that seen in the structural ensemble generated for native TGF-α. This suggests that the N-terminal tail is not required to give a stable protein fold, although it is necessary to produce the native structure of TGF-α in terms of orientation of the subdomains.

**Implications of TGF-α 8–50 Solution Structure for Functional Differences**—With the solution structure of TGF-α 8–50 available, the biological rationale for decreased activity of the truncated protein as indicated by the reduction in receptor binding affinity and mitogenic activity can be envisaged. The previous studies from this laboratory implementing relaxation and NOE analysis techniques to elucidate the residues and regions of TGF-α involved in complexation with EGFR (16, 17) were used to design the N-terminally deleted protein. The results obtained implicated residues from the A and C loops and C tail as forming the predominant receptor binding interface. Fig. 7 illustrates the average ensemble structures of TGF-α and TGF-α 8–50 shown as molecular Connolly surface representations and color coded according to the percentage of NOEs absent in the receptor-bound wild type TGF-α, i.e. those present in the free ligand but not in the bound protein.

The structure utilized for the native molecule in Fig. 7 is the most recently published structure for TGF-α as determined by Moy et al. (15), which was unavailable for our previous studies (17). When this structure is used to display the changes in NOE intensity upon addition of the EGFR, an almost completely contiguous receptor binding face is observed. This visualization gives an even more plausible scenario for the interactions of TGF-α with its receptor than indicated in the previous publication from this laboratory (17) since Arg32, Leu49, and Leu49 are now present on the same face of the protein structure as the other residues that lose the highest percentage of NOEs upon receptor binding, i.e. those colored in red and green. As the regions of the truncated protein that exhibit the most significant structural changes are the A and C loops, it can easily be seen that perturbation of the conformation of these regions...
would lead to a diminution in the affinity to the receptor and concomitantly to a decrease in mitogenic activity. The loss of the partial helical structure of the A loop in the native molecule most probably results in a lower affinity as the important residues His$^{12}$, Thr$^{13}$, Phe$^{15}$, and Phe$^{17}$ are presented to the receptor in a different manner. The chemical shift differences in the α-protons, strongly suggest that in particular Thr$^{13}$, Gln$^{14}$, and Phe$^{15}$ are in quite different environments in the N-terminally deleted TGF-α. The resonance and structural perturbations for the truncated TGF-α are in good agreement with the NOE analysis data, which implicated His$^{12}$ and Phe$^{15}$ on the receptor binding interface, and with prior mutagenesis studies that delineated Phe$^{15}$ as a critical residue for association of the ligand with the receptor (7). It can be observed from the molecular surface view of TGF-α 8–50 in Fig. 7, that His$^{12}$ is indeed displayed differently to the receptor as it is now no longer contiguous with the rest of the residues that form the primary interactions with the EGFR. In addition, The A-C interactions of Phe$^{15}$ to Arg$^{42}$ and Phe$^{17}$ to Glu$^{44}$ in the N-terminally deleted molecule present a quite different surface of these residues to the receptor and thus corroborate the postulation that the new orientation of the subdomains is responsible for the decrease in binding and activity of TGF-α 8–50.

From Fig. 7, it is also apparent that the residues of the C loop and C-terminal tail implied to be on the EGFR binding face namely Val$^{39}$, Gly$^{40}$, Arg$^{42}$, Glu$^{44}$, His$^{45}$, Leu$^{48}$, and Leu$^{49}$ are displayed in a different orientation due to the variant C-terminal subdomain conformation in TGF-α 8–50. It is evident that these structural changes alone are significant enough to explain the decreased binding and activity of the truncated protein and despite the observation that the N-terminal tail does not contribute directly to interactions with the EGFR, it makes an important contribution to the protein fold in terms of orientation of the subdomains and hence is necessary for the presentation of the binding determinants to the receptor. The results obtained from the truncated TGF-α also further confirm the multidermin binding model since it is the relative positioning of the two subdomains that is important to give a fully active protein.

**Toward Minimization of TGF-α through Truncation and Re-optimization**—Through the knowledge of the structural requirements of TGF-α for receptor complexation as delineated by the relaxation rate and NOE analysis data, this study has demonstrated that the proposal of an active truncated protein is possible when enough is known about the receptor binding site on the ligand. The synthesis of a TGF-α fragment with a respectable level of activity serves as a platform for further minimization and reoptimization of this growth factor for the purposes of the development of agonists and antagonists as potential therapeutics. Work currently in progress in this laboratory has the aim of engineering the truncated TGF-α through phage display to increase the affinity to or beyond native levels (32). This can then be followed by further minimization of the ligand in a manner analogous to the studies of Li et al. (33), who have met with considerable success in the minimization of the polypeptide hormone atrial natriuretic peptide. In the instance of TGF-α, however, the process is greatly aided by the availability of NMR data for elucidating the interactions of the minimized proteins with the receptor and thus using a structure-based iterative procedure for the identification of lead compounds demonstrating potential as therapeutics for diseases involving the overexpression or underexpression of TGF-α.

**Acknowledgments**—We thank Paul Semchuk for peptide synthesis and purification and also mass spectrometry, Gerry McQuaid and Bruce Lix for maintenance of NMR equipment, and Leigh Willard and Robert Boyko for computer programming assistance. We also acknowledge Richard Harkins at Berlex Biosciences for early discussions proposing the study of TGF-α 8–50.

**REFERENCES**

1. Bennet, N. T., and Schultz, G. S. (1993) *Am. J. Surg.* 165, 728–737
2. Bennet, N. T., and Schultz, G. S. (1993) *Am. J. Surg.* 166, 75–81
3. Frignet, S. A., and Lemoine, N. R. (1992) *Prog. Growth Factor. Res.* 4, 1–24
4. Stemmler, B. M., Brandt, R., Ciardiello, F., and Normanno, N. (1998) *Crit. Rev. Oncol. Hematol.* 19, 183–232
5. Rusc, V., Mendelsohn, J., and Dmitrovsky, E. (1996) *Cytokine Growth Factor. Rev.* 7, 133–141
6. Okamura, K., Morimoto, A., Hamanaka, R., Ono, M., Kohn, K., Uchida, Y., and Kuwano, M. (1992) *Biochem. Biophys. Res. Commun.* 186, 1471–1479
7. Groenen, L. C., Nee, E. C., and Burgess, A. W. (1994) *Growth Factors* 11, 235–257
8. Campion, S. R., and Niyogi, S. K. (1994) *Proc. Natl. Acad. Sci. U. S. A.* 81, 6367–6371
9. Kohda, D., Shinada, I., Miyake, T., Fuwa, T., and Inagaki, F. (1989) *Biochemistry* 28, 853–858
10. Kline, T. P., Brown, F. K., Brown, S. C., Jeffs, P. W., Kopple, K. D., and Hoyt, D. W., Harkins, R. N., Debanne, M. T., O'Connor-McCourt, M., and Sykes, B. D. (1993) *Biochemistry* 32, 7805–7813
11. Harvey, T. S., Wilkinson, A. J., Tappin, M. J., Cooke, R. M., and Campbell, I. D. (1991) *Eur. J. Biochem.* 196, 555–562
12. Muy, F. J., Li, Y-C., Rauenbuehler, P., Winkler, M. E., Scheraga, H. A., and Montelione, G. T. (1993) *Biochemistry* 32, 7334–7353
13. Hoyt, D. W., Harkins, R. N., Debanne, M. T., O'Connor-McCourt, M., and Sykes, B. D. (1996) *J. Biol. Chem.* 271, 32024–32211
14. Bax, A., and Davis, D. G. (1985) *Magn. Reson. Med.* 35, 355–360
15. Jeener, J., Meier, B. H., Bachmann, F., and Ernst, R. (1979) *J. Chem. Phys.* 71, 4546–4553
16. Tappin, M. J., Cooke, R. M., Fenton, J. E., and Campbell, I. D. (1988) *Eur. J. Biochem.* 179, 629–637
17. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) *X-Plor*, in press
18. Bru¨ nger, A. T. (1993) *J. Mol. Biol.* 230, 408–422
19. Brown, S. C., Weber, P. L., and Mueller, R. (1988) *J. Magn. Reson. 77* 166–169
20. States, D. J., Haberkorn, R. A., and Rueben, D. R. (1982) *J. Magn. Reson.* 48, 286–292
21. Jeener, J., Meier, B. H., Bachmann, F., and Ernst, R. (1979) *J. Chem. Phys.* 71, 4546–4553
22. Piotto, M., Saudek, V., and Sklenar, V. (1992) *J. Biomol. NMR* 2, 661–665
23. Groenen, L. C., Nee, E. C., and Burgess, A. W. (1992) *J. Magn. Reson.* 48, 286–292
24. Jeener, J., Meier, B. H., Bachmann, F., and Ernst, R. (1979) *J. Chem. Phys.* 71, 4546–4553
25. Groenen, L. C., Nee, E. C., and Burgess, A. W. (1992) *J. Biomol. NMR* 2, 661–665
26. States, D. J., Haberkorn, R. A., and Rueben, D. R. (1982) *J. Magn. Reson.* 48, 286–292
27. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) *X-Plor*, in press
28. Bru¨ nger, A. T. (1993) *X-Plor*, version 3.1, Yale University Press, New Haven, CT
29. Nilges, M., Macias, M. J., O’Donoghue, S. I., and Ochshinat, H. (1997) *J. Mol. Biol.* 269, 408–422
30. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. (1993) *J. Appl. Crystallogr.* 26, 283–290
31. Villalot, J., Liu, C., Emond, J., and Taso, M. S. (1994) *Exp. Cell Res.* 212, 36–41
32. Tirn, X.-B., Dallaire, P., Hoyt, D. W., Sykes, B. D., O'Connor-McCourt, M., and Malcolm, B. A. (1997) *J. Biochem. (Tokyo)* 122, 686–690
33. Li, B., Tom, Y. K., Oare, D., Yen, R., Fairbrother, W. J., Wells, J. A., and Cunningham, B. C. (1995) *Science* 270, 1657–1660
34. Kuhl, P. J. (1991) *J. Appl. Crystallogr.* 24, 946–950
35. Merritt, E. A., and Murphy, M. E. P. (1994) *Acta Crystallogr.* Sec. D 50, 869–873
36. Nicholls, A., Sharp, K., and Honig, B. (1991) *Proteins* 11, 281–296