3D STRUCTURES OF EGF AND TGF-ALPHA DETERMINED BY NMR

DAISUKE KOHDA and FUYUHIKO INAGAKI

Department of Molecular Physiology, the Tokyo Metropolitan Institute of Medical Science, 18-22, Honkomagome 3-chome, Bunkyo-ku, Tokyo 113, Japan

Abstract - The 3D structures of epidermal growth factor (EGF from mouse and human) and type α transforming growth factor (TGF α from human) in solution were studied by two-dimensional proton NMR spectroscopy. The pH of the samples were pH 2-3 for EGF and pH 4.9 for TGF α . The backbone foldings of EGF and TGF α are similar to each other, which is consistent with the fact that they compete for binding to the same receptor. The dynamical properties, however, are different; TGF α is unusually flexible at acidic pHs. We found that a proline at residue 29 is responsible for this difference.

By means of 2D NMR pH titration experiment, the resonance assignment of mouse EGF at pH 6.8 was accomplished. On the basis of the assignment, the 3D structure of mouse EGF at the physiological pH was calculated. Then, the conformation of mouse EGF bound to micelles at pH 6.8 was determined. EGF interacts with micelles through the C-terminal tail region, which adopts an amphiphilic structure on binding to micelles. We propose a "mitten" model to describe the interaction mode of EGF/TGF α with the receptor.

Key words epidermal growth factor, transforming growth factor α , 3D structure determination, 2D NMR, micelle

Epidermal growth factor (EGF from mouse and human, 53 res.) and type α transforming growth factor (TGF α from human, 50 res.) are structurally related peptide hormones, and competitively bind to the same receptors on cell surface with similar affinity[1].

	1	10	20	30	40	50
mEGF :	NSYPGQPS	SYDGYCL	NGGVCMH	IESLDSY	TNCNVIG	YSGDRQ
hEGF :	NSDSECP	LSHDGYC	LHDGVCM	YIEALDK	YACNCV	VGYIGER
hTGF α :	VVSHFND	CPDSHTQ	CFH-GT	CRFLVQ	EDKPACV	CHSGYVG
						ARCEHADLLA

The residue numbers of TGF α in this paper refer to the EGF sequence unless specified otherwise. The determination of the 3D structures of these growth factors is a key step for understanding the structure-function relationship of EGF/TGF α . At the present time, no x-ray crystal structures are available. Nuclear magnetic resonance (NMR) spectroscopy is an alternative powerful method to elucidate the three-dimensional (3D) structures of proteins in solution[2]. This involves the sequence-specific resonance assignment of the protein and the calculation of the 3D structure that is consistent with the NMR information. The solution conditions for NMR studies can be varied over a wide range. We take this advantage for investigating the solution structures of EGF at various pHs and in the presence of micelles.

It should be noted here that NMR studies on EGF/TGF α have been also reported by I. D. Campbell's group, Univ. of Oxford, UK; G. T. Montelione's group, Univ. of Michigan, USA; L. Mueller's group, Smith Kline & French Laboratories, USA. Their results are consistent with ours.

EXPERIMENTAL

Materials Mouse EGF was purified from submaxillary glands of adult male mice by chromatography on Bio-Gel P10, DEAE-cellulose and reversed-phase HPLC[3]. Recombinant human EGF and human TGF α were produced in *E. coli* by Wakunaga Pharmaceuticals Company Hiroshima Japan. Perdeuterated

dodecylphosphocholine (DPC) was synthesized according to the method of Brown[4].

NMR Measurements The protein samples were dissolved at 3-6 mM in 99.8% $^2\text{H}_2\text{O}$ or in 90% $^1\text{H}_2\text{O}$ /10% $^2\text{H}_2\text{O}$. The sample pH was 2.0 or 6.8 for mouse EGF, 2.7 for human EGF and 4.9 for TGF α . ^1H NMR spectra at 400 MHz and at 500 MHz were recorded on JEOL JNM series spectrometers at a probe temperature of 28°C. DQF-COSY, HOHAHA, and NOESY spectra were recorded in the phase-sensitive mode.

Sequence-Specific Resonance Assignment Complete sequence-specific resonance assignments of mouse EGF(pH 2.0) human EGF(pH 2.8) and human TGF α (pH 4.9) were obtained [3,5,6] according to the conventional sequential assignment method developed by Wüthrich[2]. The assignments of mouse EGF (pH 6.8) were achieved by the combination of the sequential assignment method and 2D NMR pH titration method[7]. The assignment of mouse EGF (pH 6.8) in the presence of DPC micelles were obtained by the spectral comparison with the assignment in the absence of micelles, and confirmed by the sequential assignment[8].

Computing Procedure All calculations were achieved on a Silicon Graphics Iris 4D/70G workstation using the program XPLOR(ver1.5 and 2.1, Polygen, USA)[9]. The computing procedure is based on the variable target function protocol[10] and the simulated annealing protocol[11] as previously described[12]. Another protocol (YASAP) was also used[9]. Analyses of calculated coordinates were carried out using XPLOR, and structures were displayed and plotted using the program QUANTA(ver 3.1 Polygen, USA).

RESULTS AND DISCUSSION

3D structures of EGFs We determined the 3D structure of mouse EGF at acidic pH[13]. A set of structures was obtained using XPLOR from randomly generated initial structures. Converged ten structures were shown in Fig. 1. To access the convergence of the structures, the atomic root-mean-square difference (RMSD) is used. The average of the RMSD value about the mean structure is 0.9 Å, which is roughly comparable in quality to 3 Å resolution in the case of x-ray structures.

Fig.1 Superposition of 10 structures of mouse EGF (pH 2.0)

Fig.2 Backbone topology of mouse EGF

Fig. 2 shows a schematic drawing of the backbone topology of mouse EGF. EGF consists of two structural domains; the N-terminal domain and the C-terminal domain. The N-terminal domain(1-32) is characterized by an antiparallel β -sheet structure. The N-terminal segment, 1-4, is loosely attached to this β -sheet to form a triple-stranded sheet. Following the N-terminal segment, the backbone chain goes down as a right-handed helix-like structure at 5-15 and turns at 16-19 to lead into the β -sheet. The C-terminal domain(33-53) consists of a β -turn at 34-37, a left-handed five-residue loop at 39-43, a small stretch of antiparallel β -sheet at 37-38 + 44-45, and a helical part 46-50. Beyond residue 51 to the C-terminus the chain is less well defined.

The 3D structure of human EGF was also determined. As expected from 75% sequence homology of mouse and human EGFs, the 3D structures of the two EGFs were found to be almost identical to each other.

3D structure of TGF α TGF α has two extra residues at the N-terminus and lacks four residues at the C-terminus. The comparison of the 3D structures of the main cores shows the identical topology of the backbone foldings. A main difference in the sequence of TGF α when compared with that of EGF is the deletion at residue 17. This deletion is easily accommodated in the turn, 16-19. Another difference is the substitution of a tyrosine at residue 29 to a proline residue. Prolines are unsuitable for β -sheet structures. However, the β -sheet in the N-terminal domain of TGF α appears to be distorted slightly to accommodate the proline replacement.

Difference in Dynamical Properties In the previous section, we emphasized the structural similarity between EGF and TGF α , but large difference was observed in their dynamical properties. Amide protons involved in hydrogen bonds can be identified as slowly exchangeable amide protons. Lyophilized protein powder was dissolved in $^2\text{H}_2\text{O}$ at pH 3, and a NMR spectrum was recorded after 24h. More than ten amide protons were left in the case of mouse and human EGFs, suggesting a single stable conformation. In contrast, TGF α did not contain any slow amide protons. We inferred that the proline replacement made the TGF α molecule more flexible than EGF, so that a mutant human EGF was prepared in which the tyrosine at position 29 was replaced by a proline (in collaboration with Dr. S. Yokoyama, Univ. of Tokyo, and Dr. T. Miyazawa, Protein Engineering Institute, Japan). As expected, no slow amide protons were observed. Thus the proline residue at residue 29 is responsible for the unusual flexibility of TGF α and the mutant EGF[unpublished results].

2D NMR pH Titration Experiment Twelve 2D HOHAHA spectra of mouse EGF in $^2\text{H}_2\text{O}$ were recorded in the pH range, 1.5-9. This permits determination of pK_a values of 9 ionizable groups in the protein, and the sequence-specific resonance assignment at neutral pHs[7]. On the basis of the assignment, we calculated the 3D structure of mouse EGF at pH 6.8[unpublished results]. With the present semiquantitative interpretation of NOEs, we found the structures at the two pHs were indistinguishable. The 2D titration experiment, however, revealed that the deprotonation of the N-terminal α -amino group induced a widespread conformational change over the β -sheet of the N-terminal domain of mouse EGF[7].

3D Structure of EGF as Bound to Micelles The interaction of mouse EGF with micelles of a phospholipid analog, dodecylphosphocholine (DPC) was investigated at pH 6.8[8]. The use of perdeuterated detergents is a key to eliminate the huge proton resonances from the detergent micelles, and spin diffusion effects which obscure the quantitative analysis of NMR data[4]. The chemical shifts were compared with those in the absence of DPC. DPC induced large chemical shift changes of the resonances from the residues in the C-terminal tail, 46-53, while a little perturbation on the residues in the main core, 1-45. Starting from the 3D structure in the absence of DPC, micelle-bound structures were calculated. The C-terminal tail, 46-53, of mEGF was found to change its conformation to form an amphiphilic structure when bound to the micelles (Fig. 3).

Fig.3 The conformation of the C-terminal domain (33-53) of mouse EGF as bound to DPC micelles.

A Mitten Hypothesis We propose a "mitten" model as gross structures of EGF and TGF α (Fig. 4)[6,13]. EGF and TGF α bind to the receptor as a mitten would grasp an object. We infer that micelles mimic a putative hydrophobic pocket on the receptor rather than biomembranes. The conformation of the C-terminal tail region, 46-53, may be changed and adopts an amphiphilic structure when bound to the receptor.

Fig.4 A "mitten" model. This model describes the interaction mode of EGF/TGF α with the receptor.

ACKNOWLEDGMENT

We thank Wakunaga Pharmaceuticals Company, especially Drs. Tetsuo Miyake and Toru Fuwa, for their kind offer of human EGF and human TGF α . We also thank Drs. Kyoze Hayashi (Gifu Pharmaceutical University), Nobuhiro Go (Kyoto University), S. Yokoyama (The University of Tokyo), and T. Miyazawa (Protein Engineering Institute) for collaboration. We are grateful for financial support from the HFSP (Human Frontier Science Project) organization.

REFERENCES

1. G. Carpenter, and S. Cohen, *Annu. Rev. Biochem.* **48**, 193 (1979).
2. K. Wüthrich, K. in *NMR of Proteins and Nucleic Acids*, J. Wiley, New York, 1986.
3. D. Kohda, and F. Inagaki, *J. Biochem. (Tokyo)* **103**, 554 (1988).
4. L. R. Brown, *Biochim. Biophys. Acta*, **557**, 135 (1979).
5. R. M. Cooke, M. J. Tappin, I. D. Campbell, D. Kohda, T. Miyake, T. Fuwa, T. Miyazawa, and F. Inagaki, *Eur. J. Biochem.* **193**, 807 (1990).
6. D. Kohda, I. Shimada, T. Miyake, T. Fuwa, and F. Inagaki, *Biochemistry* **28**, 953 (1989).
7. D. Kohda, T. Sawada, and F. Inagaki, *Biochemistry* **30**, 4896 (1991).
8. D. Kohda and F. Inagaki, *Biochemistry* submitted.
9. A. T. Brünger, *XPLOR (version 1.5 & 2.1) manual* (1988,1990).
10. M. Nilges, A. M. Gronenborn, A. T. Brünger, and G. M. Clore, *Protein Engineering* **2**, 27 (1988).
11. M. Nilges, G. M. Clore, and A. M. Gronenborn, *FEBS Lett.* **229**, 317 (1988).
12. J.-M. Lancelin, D. Kohda, S. Tate, Y. Yanagawa, T. Abe, M. Satake, and F. Inagaki, *Biochemistry* **30** 6908 (1991).
13. D. Kohda, N. Go, K. Hayashi, and F. Inagaki, *J. Biochem. (Tokyo)* **103**, 741 (1988).