

A Physical Model for TACE/peptide interactions

Sen Liu¹⁺

¹ Institute of Molecular Biology, Medical Sciences College, China Three Gorges University, Yichang 443002, China

Abstract. Tumor necrosis factor-alpha converting enzyme (TACE), also known as a disintegrin and metalloproteinase 17 (ADAM17), is a membrane-bound enzyme that releases many soluble proteins from their membrane-bound forms, including cytokines (e.g. TNF α), cytokine receptors (e.g. TNF-R), and adhesion proteins (e.g. L-selectin). Therefore, TACE has been thought as a potential therapeutic target in a lot of diseases. Currently, however, there are no complex structures available for TACE and its substrate proteins or peptides, which hampers the understanding of the molecular mechanism of TACE/substrate interaction, and the design of TACE inhibitors. Here, we present a physical model of TACE and a peptide from pro-TNF α , one of its most efficient substrates. This model is shown to be reasonable when comparing to experimental evidences. This model could provide molecular details on residue-residue interactions between TACE and its substrates, and explain the substrate selection of TACE. This complex model could serve as a good starting point for design of TACE inhibitor in future.

Keywords: TACE, ADAM17, protein-peptide model, substrate specificity.

1. Introduction

It is estimated that up to 4% of all cell surface proteins are proteolytically cleaved and released into the extracellular space via a process called ectodomain shedding [1]. Both type I and type II transmembrane proteins can go through the ectodomain shedding process, and the cleavage site is generally located in close proximity to the outer surface of the cell membrane [2]. This ectodomain-shedding process is of great importance not only for the down-regulation of membrane-bound proteins, but also for initiating or inhibiting autocrine and paracrine signalling by modulating the amount of soluble, functional proteins.

ADAM enzymes (a disintegrin and metalloproteinase) are major ectodomain shedding proteinases. ADAMs are type I transmembrane proteins, consisting of an N-terminal signal sequence followed by a prodomain, a metalloproteinase (catalytic) domain, a disintegrin domain, an EGF-like (cysteine-rich) domain, a single transmembrane domain, and a cytoplasmic tail [1, 3]. TACE (ADAM17) is one of the most well studied ADAM enzymes. At present, the ectodomain shedding of at least 76 proteins have been found to be mediated by TACE, including cytokines and their receptors such as TNF α and TNF-R, some ErbB ligands and their receptors, interleukin-6 receptor, and cell adhesion molecules such as L-selectin [2]. However, the substrate selection and specificity of TACE has not been well understood, due to the fact that no complex structures available for TACE and its substrates at present.

Many of the above target proteins of TACE are disease-related. For example, the elevated levels of TNF α and IL-6 are correlated with inflammation, and inhibiting TACE to down-regulate their soluble forms could be effective in controlling autoimmune diseases. Recent studies also revealed that TACE could be involved in driving inflammation-associated cancer, which is caused by a chronic activation of the immune system [4, 5]. Therefore, developing effective TACE inhibitors is attracting a lot of attention, and could be largely helped by a detailed understanding of the molecular mechanism of TACE/substrate binding.

⁺ Corresponding author. Tel.: +86-717-6397979; fax: +86-717-6397179.
E-mail address: pkustar@gmail.com.

Trying to understand the substrate selection and specificity of TACE, a theoretical model was built to provide details of TACE/substrate interaction at the molecular level. The peptide sequence profile predicted from the model agrees pretty well with previous experimental results, and suggested that the substrate specificity of TACE is indeed sequence dependent. The model could help our understanding of the substrate selection and specificity of TACE, as well as the design of TACE inhibitors.

2. Methods and Results

2.1. TACE-peptide Docking

TNF α is a main cytokine in autoimmune diseases. It was first identified in a 17-kDa secreted form (s-TNF α), but later research showed that it also exists in a non-cleaved 26-kDa membrane-bound precursor form (pro-TNF α) [6]. Pro-TNF α is cleaved by TACE to secrete the active s-TNF α , which can bind to its receptors (TNF-R I or TNF-R II) and induce inflammatory responses or apoptosis. The cleavage happens between the residues Ala76 and Val77 of pro-TNF α , and the recognized sequence, PLAQA-VRSS, is one of the most efficient substrates of TACE [7]. Therefore, this peptide sequence was chosen for complex building.

TACE has several different domains, and the catalytic domain is the one that has the sheddase activity and cleaves TACE substrates. A lot of crystal structures have been obtained for the catalytic domain of TACE (TACEcat), and one of the structures (PDB ID: 1BKC) was chosen for complex building because it has a reasonably high resolution (2.0 Å). What's more, in the structure of 1BKC, the TACEcat is in the complex form with a peptide-like inhibitor, N-{d,l-[2-(hydroxyamino-carbonyl)methyl]-4-methylpentanoyl}l-3-(tert-butyl)glycyl-l-alanine. We hope that the existence of the bound peptide-like inhibitor could have pre-defined some necessary main-chain and side-chain conformation of TACEcat for substrate binding, since the substrates should bind to the same area of TACEcat.

With the substrate sept-peptide and the structure of TACEcat, the FlexPepDocking-AbInitio protocol [8] in the protein design software package Rosetta was used to build up the complex model. Firstly, a linear peptide structure was constructed for the sept-peptide PLAQAVRSS as described in [8]. Then, a starting "complex" structure was obtained by putting the linear peptide as close as possible to the catalytic site of the structure of TACEcat, while avoiding obvious collisions. After that, fragment libraries were generated for the peptide using the protocol, followed by pre-repacking the side chains.

One more thing to do before running the docking step of the FlexPepDocking-AbInitio protocol is adding constraints to the "core" area of the structure of TACEcat including the zinc atom and the three His residues coordinating the zinc atom. The reason to do that is, the interactions between the zinc atom and the three His residues should keep intact, which could be destroyed while the backbone and side chains are moving during the docking step. The constraints were defined by the atom distances and dihedral angles found in the crystal structure. A distance constraint was also defined between the zinc atom and the carbonyl oxygen of the residue Ala (the P1 site) right before the cleavage bond (between A-V). Briefly, the carbonyl oxygen of the cleavage Ala and the three NE2 atoms of the three His residues were placed at the four virtual positions of the zinc defined in Rosetta.

Nearly 60,000 docking models were generated on a computer cluster, and the 500 lowest-score models were clustered by a 2 Å radius (the sept-peptide only) after rescoring the models according to the protocol instruction. The biggest cluster has nearly 100 models, including the lowest-score model, and the other clusters are much smaller. Therefore, the lowest-score model (Fig. 1) in the biggest cluster was chosen as the representative complex model for the following substrate specificity scanning. In the complex model, the sept-peptide runs across the substrate-binding cleft, which is relatively flat on the left-hand side but becomes notched toward the right, similar to the conformation found in the TACEcat-inhibitor structure [9]. The four residues -L-Q-V-S- are interacting with TACEcat very well by hydrophobic or electro-static interactions, with the residue V (the P1' site) residing in the S1' pocket of TACEcat, which is mainly hydrophobic.

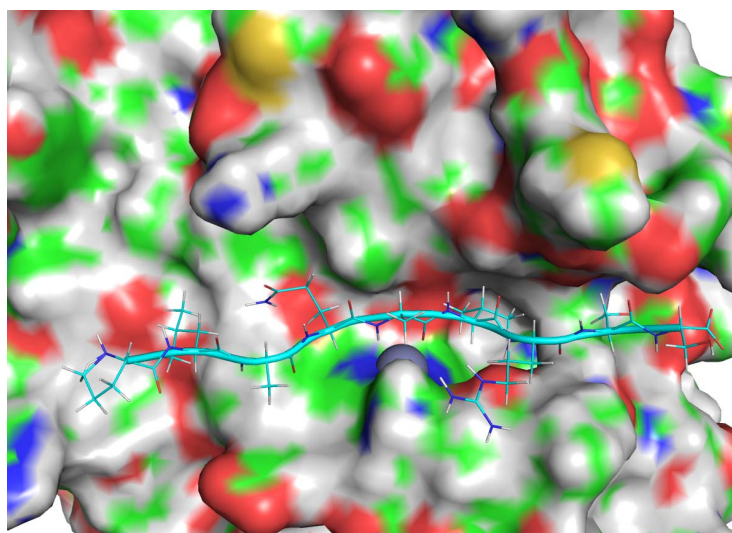


Fig. 1: The selected representative docking complex model of TACEcat and the sept-peptide PLAQAVRSS. TACEcat is shown in surface, and the zinc in grey sphere. The sept-peptide is shown in cartoon and lines, with the N-ter on the left-hand. The figure was prepared in Pymol.

2.2. Substrate Specificity Scanning

It's been a long arguing problem that if the sequence of the substrate cleavage site is important for the cleavage efficiency of TACE [2]. Experimental library scanning of the peptide sequence were used to map the substrate sequence specificity for TACE and did find some residue preferences for P4-P4' sites [10, 11]. However, lacking a TACE-peptide complex structure, it's still hard to explain the substrate specificity and selection of TACE. What's more, due to the size limit of the experimental library scanning, a lot of sequences could have been actually missed, which makes the scanning process not as thorough as expected. For example, in the ref [11], the Ala at the P1 site was basically kept unchanged.

Recently, a protocol called "backrub" was developed [12] and implemented in Rosetta. This protocol has been proved to be very useful in profiling sequence specificities by coupling to the derived sequence tolerance scanning protocol for both protein-peptide interactions [13] and protein-protein interactions in our recent study [paper under review]. Compared to the experimental library scanning, the computational sequence scanning has the potential, theoretically, to scan all possibilities at all sequence sites, which makes it a very good tool supplementary to experimental library scanning. Therefore, the backrub and sequence tolerance scanning protocols were used to predict the sequence preference of the substrate peptide of TACEcat using the above selected complex model. What's more, the backrub protocol can also optimize the backbone and side-chain conformations of the docking model.

The docking model from the above step was firstly used to generate 100 backrub models using the backrub protocol. The interface residues less than 6 Å away from each chain by C-alpha atom were defined as the pivot residues, and the same constraints on the core area were applied as well to keep it intact when the backbone made small backrub movements.

Then, the 100 low-score models from the backrub protocol were submitted for sequence tolerance scanning. The 20 amino acids except Cys were allowed to show up at all these 9 positions, so the total library size is about $2e17$. Genetic algorithm is applied in the sequence tolerance scanning protocol, and the population size was set to 4000, the generation 30. Finally, a position weight matrix (PWM) was generated from the scanning results (Fig. 2). As shown in Fig. 2, the prediction captured the critical sequence preference at these positions, except 256, 258, and 260, which are protruding out away from TACEcat as shown in Fig. 1 and could be important for main-chain conformation selection instead of affinity selection. The prediction result agrees with the previous experimental results pretty well. For example, in the ref [11], the P2 site prefers Ala, Ser, Arg and Glu, and these preferences are also captured in our scanning result as shown in Fig. 2 (the site 259). The P1' site was shown to prefer small hydrophobic residues such as Val, Leu, and Ile in the ref [11], which is also similar in our result (the site 261). Similar agreements were found for the

sites P2' and P3'. Therefore, the representative complex model we selected could be a pretty nice capture of the real complex structure.

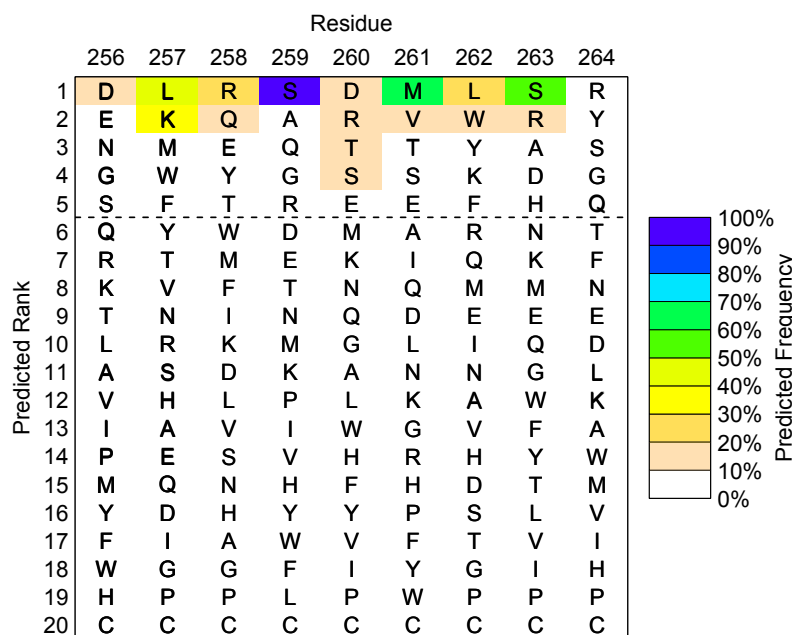


Fig. 2: The sequence tolerance scanning result for the sept-peptide (PLAQAVRSS) in the TACEcat-peptide model. The ID of the first residue of the sept-peptide is named as 256 in the complex model.

In this paper, we presented a computational physic model for the catalytic domain of TACE and one of its substrate peptide PLAQAVRSS. The nice agreement between our prediction results and previous experimental evidence probably means that our model is a reasonable capture of the real complex structure between them. We should mention that in the theoretical model, only the affinity, but not the enzymatic activity is considered. Therefore, the prediction result probably indicates that the substrate selection of TACE is indeed sequence dependent.

3. Acknowledgements

The helpful discussions from RosettaCommons are appreciated. This work was jointly supported by the National Natural Science Foundation of China (No. 21103098), the Key Discipline Construction Foundation of China Three Gorges University (No. 2011071001), and the Doctor Foundation of China Three Gorges University (No. 0620070083).

4. References

- [1] P. Saftig and K. Reiss. The "A Disintegrin And Metalloproteases" ADAM10 and ADAM17: novel drug targets with therapeutic potential? *Eur J Cell Biol* 2011, 90 (6-7): 527-535.
- [2] J. Scheller, A. Chalaris, C. Garbers, and S. Rose-John. ADAM17: a molecular switch to control inflammation and tissue regeneration. *Trends Immunol* 2011, 32 (8): 380-387.
- [3] M. Gooz. ADAM-17: the enzyme that does it all. *Crit Rev Biochem Mol Biol* 2010, 45 (2): 146-169.
- [4] A. Chalaris, N. Adam, C. Sina, P. Rosenstiel, J. Lehmann-Koch, P. Schirmacher, D. Hartmann, J. Cichy, O. Gavrilova, S. Schreiber, T. Jostock, V. Matthews, R. Hasler, C. Becker, M. F. Neurath, K. Reiss, P. Saftig, J. Scheller, and S. Rose-John. Critical role of the disintegrin metalloprotease ADAM17 for intestinal inflammation and regeneration in mice. *The Journal of experimental medicine* 2010, 207 (8): 1617-1624.
- [5] K. Brandl, L. Sun, C. Neppel, O. M. Siggs, S. M. Le Gall, W. Tomisato, X. Li, X. Du, D. N. Maennel, C. P. Blobel, and B. Beutler. MyD88 signaling in nonhematopoietic cells protects mice against induced colitis by regulating specific EGF receptor ligands. *Proceedings of the National Academy of Sciences of the United States of America* 2010, 107 (46): 19967-19972.
- [6] M. Kriegl, C. Perez, K. DeFay, I. Albert, and S. D. Lu. A novel form of TNF/cachectin is a cell surface cytotoxic

transmembrane protein: ramifications for the complex physiology of TNF. *Cell* 1988, 53 (1): 45-53.

- [7] M. S. Bahia and O. Silakari. Tumor necrosis factor alpha converting enzyme: an encouraging target for various inflammatory disorders. *Chem Biol Drug Des* 2010, 75 (5): 415-443.
- [8] B. Raveh, N. London, L. Zimmerman, and O. Schueler-Furman. Rosetta FlexPepDock ab-initio: simultaneous folding, docking and refinement of peptides onto their receptors. *PLoS One* 2011, 6 (4): e18934.
- [9] K. Maskos, C. Fernandez-Catalan, R. Huber, G. P. Bourenkov, H. Bartunik, G. A. Ellestad, P. Reddy, M. F. Wolfson, C. T. Rauch, B. J. Castner, R. Davis, H. R. Clarke, M. Petersen, J. N. Fitzner, D. P. Cerretti, C. J. March, R. J. Paxton, R. A. Black, and W. Bode. Crystal structure of the catalytic domain of human tumor necrosis factor-alpha-converting enzyme. *Proceedings of the National Academy of Sciences of the United States of America* 1998, 95 (7): 3408-3412.
- [10] M. H. Lambert, R. K. Blackburn, T. D. Seaton, D. B. Kassel, D. S. Kinder, M. A. Leesnitzer, D. M. Bickett, J. R. Warner, M. W. Andersen, J. G. Badiang, D. J. Cowan, M. D. Gaul, K. G. Petrov, M. H. Rabinowitz, R. W. Wiethe, J. D. Becherer, D. L. McDougald, D. L. Musso, R. C. Andrews, and M. L. Moss. Substrate specificity and novel selective inhibitors of TNF-alpha converting enzyme (TACE) from two-dimensional substrate mapping. *Comb Chem High Throughput Screen* 2005, 8 (4): 327-339.
- [11] C. I. Caescu, G. R. Jeschke, and B. E. Turk. Active-site determinants of substrate recognition by the metalloproteinases TACE and ADAM10. *Biochem J* 2009, 424 (1): 79-88.
- [12] C. A. Smith and T. Kortemme. Backrub-like backbone simulation recapitulates natural protein conformational variability and improves mutant side-chain prediction. *Journal of molecular biology* 2008, 380 (4): 742-756.
- [13] C. A. Smith and T. Kortemme. Structure-based prediction of the peptide sequence space recognized by natural and synthetic PDZ domains. *Journal of molecular biology* 2010, 402 (2): 460-474.