ELSEVIER

Computer-aided molecular modeling of cathepsin E, a possible endothelin-converting enzyme

Massoud Mahmoudian

Department of Pharmacology, Iran University of Medical Sciences, Tehran, Iran

A three-dimensional model of human cathepsin E, a possible endothelin-converting enzyme, is constructed using computer-aided molecular modeling techniques. The structure of porcine pepsin, another aspartic protease, was used as a template. The final structure, after all gaps and deletions were made, was optimized using the AMBER-4 package. A dipeptide (Trp-Val) representing the substrate was docked in the putative active site and the whole structure was optimized after several runs of minimization and dynamics calculations. The result of this modeling study showed that the structure of cathepsin E is similar to that of porcine pepsin and has three disulfide bonds that are conserved in both enzymes. There are two Asp-Thr-Gly sequences at the active site of enzyme. The active site cavity is large enough to accommodate its substrate. © *1996 by Elsevier Science Inc.*

INTRODUCTION

Human endothelin is a 21-residue polypeptide produced by endothelial cells and displaying vasoconstrictor activity and a number of other pharmacological actions. It is believed that endothelin is generated by a single proteolytic cleavage of the bond linking residues Trp-21 and Val-22 of its immediate biosynthetic precursor, big endothelin, by an endopeptidase termed endothelin-converting enzyme (ECE).[1] Several experimental studies suggest that cathepsin E might be a good candidate for ECE.[1-3] Cathepsin E is an intracellular aspartic protease[4] that exists both as a disulfide-linked monodimer and as a monomer. Both forms of the enzyme are catalytically active, but the dimer is more stable.[5]

Since the three-dimensional (3D) structure of this enzyme has not been elucidated, molecular modeling techniques have been used to generate a reasonable structure for this enzyme to give a better understanding of its structure–function relationships and assist in the design of inhibitors that are needed for clarification of function.[6] The inhibitors of ECE would have potential as antihypertensive drugs.

METHODS

The sequence of human procathepsin[7] was aligned with that of porcine pepsinogen[8] using the MULTALIN program.[9] The result is shown in Figure 1. The secondary structure assignments were based on the consensus assignments from a battery of secondary structure prediction methods.[10-12] The predicted secondary structure of human cathepsin E shows a remarkable similarity to that of the 3D structure of porcine pepsin and is supported by the results of alignments. Therefore, the 3D structure of porcine pepsin[13] was used as a template to generate the backbone of cathepsin E. There are three gaps in the alignment of these two peptides (Figure 1). Inspection of the structure of porcine pepsin showed that these gaps are located in surface loops. Therefore, the gaps were corrected by deleting or adding the required residues manually, using the ALCHEMY package.[14] The resulting structure was further refined using the AMBER-4 set of programs.[15] Finally, a constructed dipeptide (Trp-Val) was docked in the active site of the enzyme in the place of porcine pepsin inhibitor.[13] The whole structure was put in a droplet of water (2 678 molecules) and further refined by molecular dynamics and minimization calculations using the AMBER-4 set of programs. A dielectric constant multiplicator factor of unity was used for these calculations.

RESULTS AND DISCUSSION

The results of the alignment of porcine propepsin and human procathepsin are shown in Figure 1. As can be seen from Figure 1, these two proteins are similar and 54% of their residues are identical. Furthermore, the predicted secondary structure of cathepsin E, using a battery of secondary

Color Plates for this article are on page 225.

Address reprint requests to Massoud Mahmoudian, Department of Pharmacology, Iran University of Medical Sciences, P.O. Box 14155-6138, Tehran, Iran.

Received 14 May 1996; accepted 14 August 1996.

655 Avenue of the Americas, New York, NY 10010

0263-7855/96/$15.00
PII S0263-7855(96)00065-3

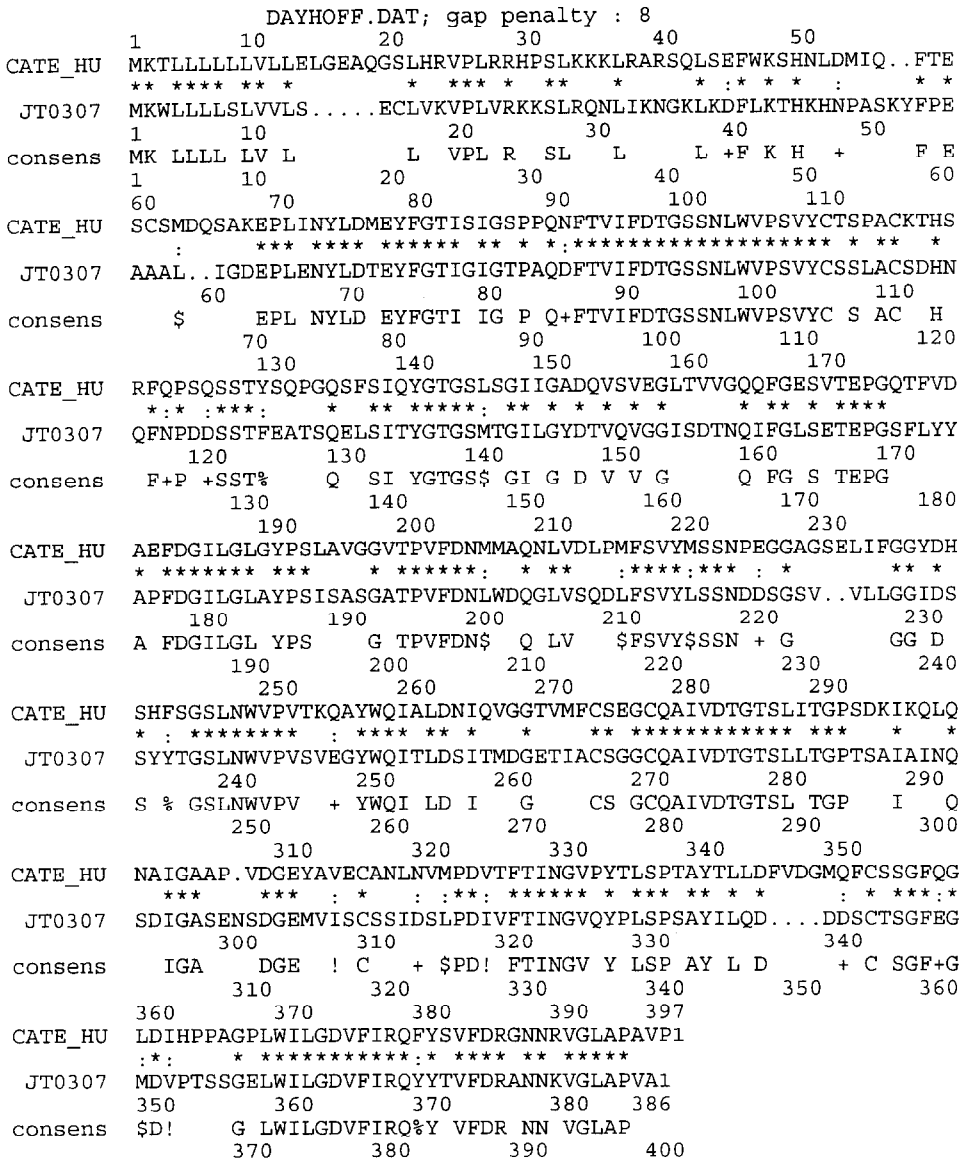

Figure. 1. Alignment of the sequences of human procathepsin (CATE-HU) with that of porcine pepsinogen (JY0307). An asterisk (*) indicates identical residues.

structure prediction methods,[10–12] is similar to that of porcine pepsin.[13] Therefore, it is reasonable to use the structure of porcine pepsin[13] as a template for constructing a model of human cathepsin E. The residues of porcine enzyme[13] were mutated to the corresponding residues of human cathepsin E and the gaps were corrected by manual addition and deletion of the required amino acids (Figure 1). The resulting structure was checked for incorrect binding of side chains, which were corrected manually. A total of 2 678 molecules of solvent water was added to the final structure. The structure of human cathepsin E, after several runs of energy minimization and dynamics calculations with the AMBER package,[14] is shown in Color Plate 1. The N-terminal residues of active porcine pepsin, whose structure[13] is used in this study, correspond to Ile-60 of propepsin. Thus our modeled structure of human cathepsin E has its N terminal at a serine residue, which corresponds to this isoleucine (Figure 1). It is believed that cathepsin E is synthesized as a preproenzyme[7] and during maturation it loses its first 56 residues.[4,5] Our model is slightly shorter than the mature enzyme and its N-terminus serine corresponds to Ser-9 of the mature form.

As can be seen from Color Plate 1, the secondary structure of human cathepsin E mainly consists of β sheets (32.5% of all residues). Eleven percent of residues are in the form of α helix and the rest are either turns or random coils. The secondary structure of the modeled cathepsin E (Color Plate 1 and Table 1) is similar to that of porcine pepsin. There are three disulfide bonds in the porcine pepsin: Cys-45–Cys-50, Cys-206–Cys-210, and Cys-249–Cys-282. All these residues and their corresponding disulfide bonds are preserved in the modeled structure of cathepsin E. There exists a large cavity, which could accommodate the substrate (Figure 1 and Color Plate 2).

Cathepsin E has a unique cysteine residue (Cys-4), which is located at the N terminus of the mature enzyme.[5] This

Table 1. Secondary structure of modeled human cathepsin E[a]

```
 10        20        30        40        50
 SAKEPLINYLDMEYFGTISIGSPPQNFTVIFDTGSSNLWVPSVYCTSPAC
       bbbtttbbbb   bttttb   bbbb     bb        ttt

 60        70        80        90       100
 KTHSRFQPSQSSTYSQPGQSFSIQYGTGSLSGIIGADQVSVEGLTVVGQQ
 tt     tt     b   bb  bbbttbbbbbbb bbbb  tt  bbbb

110       120       130       140       150
 FGESVTEPGQTFVDAEFDGILGLGYPSLAVGGVTPVFDNMMAQNLVDLPM
  bbbbb   aaaat    bbbb   ttt ttt   aaaaaaaat     b

160       170       180       180       200
 FSVYMSSNPEGGAGSELIFGGYDHSHFSGSLNWVPVTKQAYWQIALDNIQ
 bbbb  ttt      tt bb    ttb    bbbbbbbttttbbbbbbbb

210       220       230       240       250
 VGGTVMFCSEGCQAIVDTGTSLITGPSDKIKQLQNAIGAAPVDGEYAVEC
     bbb tt bbbb   tt  bbb aaaaaaaaaattb  bttb  b a

260       270       280       290       300
 ANLNVMPDVTFTINGVPYTLSPTAYTLLDFVDGMQFCSSGFQGLDIHPPA
 aaatt     bbbttbbb  taaaa               bbb     tt

310       320       330       340
 GPLWILGDVFIRQFYSVFDRGNNRVGLAPAVP
   bbb  aaaattbbbbbbbbtttbbbbbbb
```

[a]Amino acid residues are numbered as in the matured enzyme. a, α helix; b, β sheet; t, turn.

residue is responsible for dimerization of the mature form. Mutating this residue to alanine results in the production of only the monomeric form. The activity of the resultant enzyme will not be altered significantly compared with that of the native type. However, the stability of the mutated enzyme is markedly reduced.[5] Considering the fact that both subunits of the cathepsin E dimer are identical, only the monomeric structure was modeled. This model shows that the N terminus of cathepsin E is located at the surface of the molecule and is far away from its active site. This explains why mutation at the above-mentioned cysteine residue does not affect the catalytic activity of the enzyme.

Cathepsin E belongs to the family of asparate proteases and it is shown that the Asp-Thr-Gly sequence has been conserved in this enzyme.[16] The present modeled structure is in agreement with this experimental evidence and it is found that two Asp-Thr-Gly sequences are located around the cavity that could accommodate the substrate (Color Plate 2). It has been reported that cathepsin E is capable of hydrolyzing the peptide bond between Trp-21 and Val-22 of big endothelin 1 to produce active endothelin.[1] This dipeptide has been inserted into the porcine pepsin inhibitor to mark the active site of the modeled cathepsin E. However, to gain clues supporting the possibility of cathepsin E as ECE, further work is being carried out to model big endothelin and its complex with cathepsin E.

In conclusion, our modeling study shows that there is a high degree of similarity between cathepsin E and porcine pepsin, and the modeled structure is supported by experimental data.

ACKNOWLEDGMENTS

Thanks are due to Prof. P.A. Kollman (University of California, San Francisco, CA) for providing a copy of the AMBER-4 set of programs, and to Dr. A. Madadkar (Computer Department of National Research Center for Genetic Engineering and Biotechnology, Tehran, Iran) for his help during the course of this study. The coordinates of modeled human cathepsin E are available from the author on request (e-mail address: masmah73@vax.ipm.ac.ir).

REFERENCES

1 Lees, W.E., Kalinka, S., Meech, J., Capper, S.J., Cook, N.D., and Kay, J. Generation of human endothelin by cathepsin E. *FEBS Lett.* 1990, **273,** 99–100

2 Capper, S.J., Ella, S.J., and Kalinka, S. Development and application of three radioimmunoassays of endothelin of varying specificities. *J. Cardiovasc. Pharmacol.* 1991, **17**(Suppl.), S425–S426

3 Robinson, P.S., Lees, W.E., Kay, J., and Cook, N.D. Kinetic parameters for the generation of endothelins-1, -2, and -3 by human cathepsin E. *Biochem. J.* 1992, **284,** 407–409

4 Finley, E.M. and Kornfield, S. Subcellular localization and targeting of cathepsin E. *J. Biol. Chem.* 1994, **269,** 31259–31265

5 Fowler, S.D., Kay, J., Dunn, B.M., and Tatnell, P.J. Monomeric human cathepsin E. *FEBS Lett.* 1995, **366,** 72–74

6 Richards, A.D., Kay, J., Dunn, B.M., Besant, C.M., and Charlton, P.A. Inhibition of aspartic proteinases by synthetic peptides derived from the propart region of human protein. *Int. J. Biochem.* 1992, **24,** 297–301
7 Azuma, T., Pals, G., Mohandas, T.K., Couvreur, J.M., and Taggart, R.T. Human gastric cathepsin E. Predicted sequence localization to chromosome 1, and sequence homology with other aspartic proteinases. *J. Biol. Chem.* 1989, **264,** 16748–16753
8 Tsukagoshi, N., Ando, Y., Tomita, Y., Uchida, R., Takemura, T., Sasaki, T., Yamagata, H., Udaka, S., Ichihara, Y., and Takahashi, K. Nucleotide sequence and expression in *Escherichia coli* of cDNA of swine pepsinogen: Involvement of the amino-terminal portion of the activation peptide segment in restoration of the functional prorenim. *Gene,* 1988, **65,** 285–292
9 Carpet, F. Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res.* 1989, **16,** 10881–10889
10 Chou, P.Y. and Fasman, G.D. Prediction of the secondary structure of proteins from their amino-acid sequence. *Adv. Enzymol.* 1978, **47,** 45–148
11 Garnier, J., Osguthrope, D.J., and Robson, B. Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. *J. Mol. Biol.* 1978, **120,** 97–120
12 Kyte, J. and Doolittle, R.F. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* 1982, **157,** 105–132
13 Chen, L., Erickson, J.W., Rydel, T.J., Park, C.H., Neidhart, D.J., Luly, J., and Abad-Zapatero, C. Structure of a pepsin–renin inhibitor complex reveals a novel crystal packing induced by minor chemical alterations in the inhibitor. *Acta Crystallogr., Sect. B* 1992, **48,** 48–56
14 Alchemy, version II. Tripos Associte, Inc., 1699 S. Hanley Rd., Suite 303, St. Louis, Missouri 63144, 1900
15 Pearlman, D.A., Case, D.A., Caldwell, J.C., Seibel, G.L., Singh, U.C., Weiner, P., and Kollman, P.A. AMBER 4.0. University of California, San Francisco, 1991
16 Kageyama, T. Rabbit procathepsin E and cathepsin E. *Eur. J. Biochem.,* 1993, **216,** 717–728.