Model structure for the human blood coagulation agent β-factor XIIa

M. J. Ramos

Oxford Centre for Molecular Sciences and Physical Chemistry Laboratory, Oxford University, Oxford, UK

An improvement to the human blood coagulation agent β-factor XIIa three-dimensional model is proposed. The sequence alignment as well as the modeling procedures are presented and the minimized energy of the new model is reported before and after solvation of the active center.

Keywords: β-factor XIIa, sequence homology, sequence alignment, modeling

INTRODUCTION

The blood coagulation process is a complex cascade reaction system and in each of its stages a precursor protein converts the precursor of another protease into its enzimatically active form. The active sites of the proteases that arise from the carboxyl ends of the zymogens present a great degree of homology to the pancreatic serine proteases, despite the fact that the coagulation proteases are much larger than the pancreatic proteases.[1,2] Yet, the functional differences of both sets of enzymes must be translated somehow into differences in the respective active sites or putative extended substrate binding sites that surround the active site.[3–5]

One of the blood coagulation agents is the Hageman factor, or factor XII. Human factor XII is a glycoprotein (M_r = 80 000) that has been previously purified and characterised,[6,7] and it is the circulating precursor of the enzyme that proteolytically activates factor XI. When in the presence of kallikrein and an anionic surface, factor XII is transformed into α-factor XII*a* and, upon further proteolysis, it yields β-factor XII*a*. Both the α-factor XII*a* and β-factor XII*a* amino acid sequences have been determined;[8,4] these sequences, as well as the *c*DNA sequence,[1,9,10] have been confirmed by the determination of the organization of the human factor XII gene.[11] Alpha-factor XII*a* consists of two polypeptide chains (M_r = 52 000 and M_r = 28 000) held together by a disulphide bond. Similarly, β-factor XII*a* is also composed of two polypeptide chains: a heavy chain of 243 amino acid residues (the *H*-chain, M_r = 28 000) and a light chain of 9 amino acid residues (the *L*-chain, M_r = 2 000). A disulphide bond holds both chains together. Beta-factor XII*a* has six internal disulphide bonds in addition to the one between the chains.

The quest for a good inhibitor of the human blood coagulation agent β-factor XII*a* requires a structure of the protein. Despite the available sequence information, little is known about the tertiary structure of α- and β-factor XII*a*.[12] No X-ray diffraction studies have been successfully performed on blood coagulation proteins due to the lack of crystallization forms or of quantities needed for such purposes. Therefore, for the time being, computer models of those enzymes remain the only means of their visualization. There is a computer model[1] of β-factor XII*a*, but no available coordinates. This model was rebuilt from scratch, benefiting from the experience mentioned above. Although in general the new model closely resembles one described by Cool et al.,[1] it is believed that part of the sequence alignment can be significantly improved.

SEQUENCE ALIGNMENT AND MODELING PROCEDURES

The three-dimensional model for β-factor XII*a* was built based on sequence homology with the well-known pancreatic serine proteases, bovine trypsin,[13] chymotrypsin[14] and elastase.[15]

The sequence alignment used is identical to that of Cool et al.,[1] with one difference: Residues 34–41 have been assigned differently. (The numbering system is based on the one from chymotrypsinogen.[16]) The complete resulting amino acid sequence alignment of β-factor XII*a* and bovine trypsin is displayed in Figure 1.

As far as the building of the three-dimensional model is concerned, mutations, as well as insertions and deletions, have been carried out based on the known X-ray crystal structure of bovine pancreatic trypsin,[14] using the program QUANTA.[17] The molecular mechanics computer package AMBER[18] (united atom option) was used to minimize the structure, not only in vacuum but also solvating the active center.

RESULTS AND DISCUSSION

Janin and Chothia calculated[19] the hydrophobic contribution to the free energy of dissociation in the crystalline complexes of trypsin with pancreatic trypsin inhibitor and with soybean

Address reprint requests to Dr. Ramos at the Oxford Centre for Molecular Sciences and Physical Chemistry Laboratory, Oxford University, South Parks Road, Oxford OX1 3QZ, UK.
Received 3 September 1990; accepted 9 October 1990

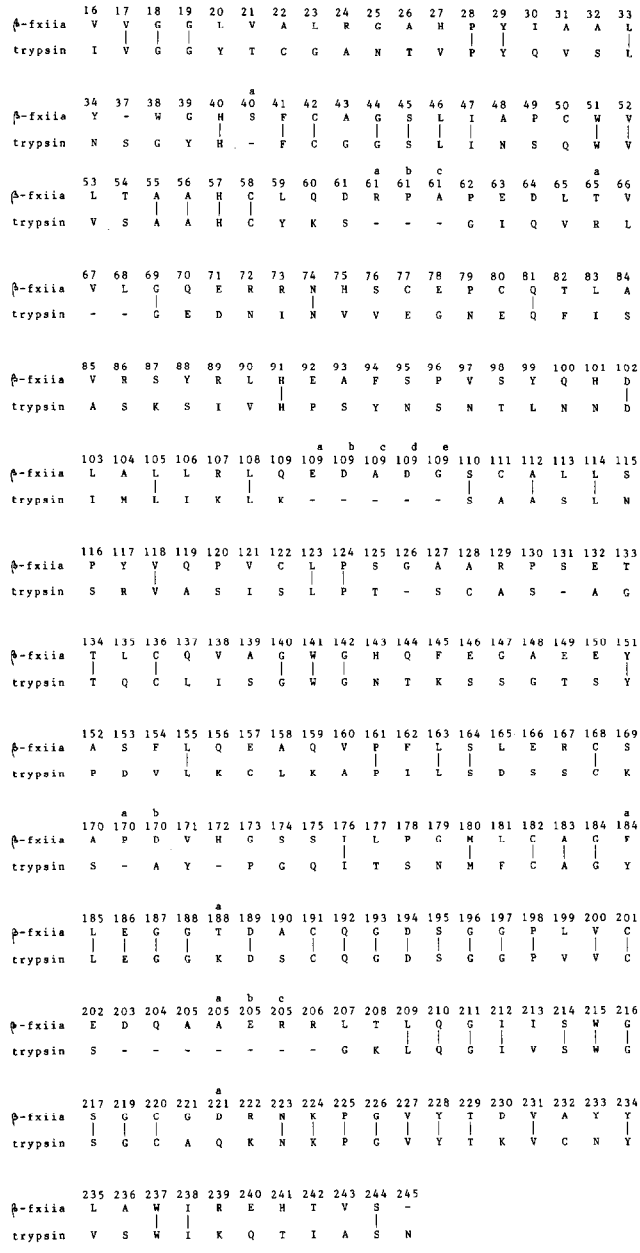

Figure 1. Amino acid sequence alignment of the H-*chain of human β-factor XII*a[4] *and bovine pancreatic trypsin.*[13] *The numbering system is based on the one from chymotrypsinogen.*[16] *Deletions and insertions are indicated by dashed lines, and vertical lines assign the corresponding positions of both enzymes. The one-letter amino acid code is used*

trypsin inhibitor. They also showed that hydrophobicity is the major factor in stabilizing the complexes. The residues that constitute the protein surface area buried in the complexes are, thus, of the utmost importance. The residue HIS 40 is one of the polar groups accessible to solvent in free trypsin and buried in the trypsin-inhibitor complexes. It is, consequently, one of the residues that constitutes the interface and contributes to the stabilization of those complexes. Its position is, therefore, relevant in this family of proteins.

The modified sequence alignment proposed here presents

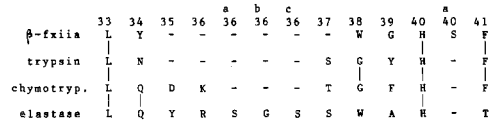

Figure 2. Excerpt of amino acid sequence alignment of the H-*chain of human β-factor XII*a,[4] *bovine trypsin,*[13] *chymotrypsin*[14] *and elastase.*[15] *The numbering system is based on the one from chymotrypsinogen.*[16] *Deletions and insertions are indicated by dashed lines. The one-letter amino acid code is used*

one more deletion and one more insertion than Cool's alignment.[1] These additions allow HIS 40 to occupy the same position in β-factor XII*a* as in trypsin. In fact, if one compares this alignment also with chymotrypsin and elastase sequences, it is possible to see that HIS 40 can be kept in the same place in all of these enzymes. Figure 2 presents an excerpt of these protein sequences around that region to highlight this. The modification in the sequence alignment falls into one of the variable sequence regions proposed by Furie et al.[3]

Although in Cool's model, HIS 40 is only one residue away from its expected position, this represents a significant difference that becomes apparent once the three-dimensional model is built and analyzed. Furthermore, in the trypsin-BPTI (bovine pancreatic trypsin inhibitor) complex, one of the H-bonds that plays an important role in the binding of the inhibitor is that formed between TYR 39 (in trypsin) and ILE 19 (in BPTI). With the new alignment proposed here, position 39 is occupied by GLY. (It is occupied by HIS in Cool's model.) This arrangement cannot lead to the formation of an H-bond with ILE 19 in BPTI or, indeed, any other residue in BPTI, and could be the reason why BPTI does not bind to β-factor XII*a*. Color Plate 1 displays both enzymes superimposed, the new β-factor XII*a* model, in purple and trypsin, in green; for clarity only the C_α backbone is shown.

The minimization of the three-dimensional model of β-factor XII*a* as described by Cool et al., and the minimization of the same model with the modification in the sequence alignment introduced here, converge virtually to the same value, −1825 kcal/mol. Given the inevitable crudity of molecular mechanics potentials it is not possible on energetic grounds to distinguish between the structural variants. A 75 Å cube of "Monte-Carlo water" has been centered in the active site of the new model and minimization of that system converges to an energy of −6112 kcal/mol.

It is gratifying to observe that, as expected, the minimizations lead to the pairing of twelve of the cysteines in disulphide bridging: residues 42–58, 50–111, 77–80, 136–201, 168–182 and 191–220. Color Plate 2 displays the minimized β-factor XII*a* model complete with side chains and the six disulphide bridges (in yellow) known to be present in β-factor XII*a*.[4] I am happy to provide the coordinates of the proposed model structure.

ACKNOWLEDGEMENTS

I would like to thank the Commission of the European Communities for a research grant, and also Richard Gilbert for helpful discussions.

REFERENCES

1 Cool, D.E., Edgell, C.-J.S., Louie, G.V., Zoller, M.J., Brayer, G.D. and MacGillivray, R.T.A. *J. Biol. Chem.* 1985, **260,** 13666

2 Jackson, C.M. and Nemerson, Y. *Ann. Rev. Biochem.* 1980, **49,** 765

3 Furie, B., Bing, D.H., Feldmann, R.J., Robison, D.J., Burnier, J.P. and Furie, B.C. *J. Biol. Chem.* 1982, **257,** 3875

4 Fujikawa, K. and McMullen, B.A. *J. Biol. Chem.* 1983, **258,** 10924

5 Furie, B. and Furie, B.C. *Cell* 1988, **53,** 505

6 Revak, S.D., Cochrane, C.G., Johnson, A.R. and Hugli, T.E. *J. Clin. Invest.* 1974, **54,** 619

7 Fujikawa, K. and Davie, E.W. *Methods Enzymol.* 1981, **80,** 198

8 McMullen, B.A. and Fujikawa, K. *J. Biol. Chem.* 1985, **260,** 5328

9 Tripodi, M., Citarella, F., Guida, S., Galeffi, P., Fantoni, A. and Cortese, R. *Nucleic Acids Res.* 1986, **14,** 3146

10 Que, B.G. and Davie, E.W. *Biochemistry* 1986, **25,** 1525

11 Cool, D.E. and MacGillivray, R.T.A. *J. Biol. Chem.* 1987, **262,** 13662

12 Baron, M., Norman, D., Willis, A. and Campbell, I.D. *Nature* 1990, **345,** 642

13 Marquart, M., Walter, J., Deisenhofer, J., Bode, W. and Haber, R. *Acta Crystallogr., Sect. B* 1983, **39,** 480

14 Cohen, G.H., Silverton, E.W. and Davies, D.R. *J. Mol. Biol.* 1981, **148,** 449

15 Sawyer, L., Shotton, D.M., Campbell, J.W., Wendell, P.L., Muirhead, H., Watson, H.C., Diamond, R. and Ladner, R.C. *J. Mol. Biol.* 1978, **118,** 137

16 Hartley, B.S. and Kauffman, D.L. *Biochem. J.* 1966, **101,** 229

17 QUANTA (version 2.1). Polygen Corporation, 200 Fifth Ave., Waltham, MA 02254, USA

18 Singh, U.C., Weiner, P.K., Caldwell, J.W. and Kollman, P.A. AMBER (version 3.1). Department of Pharmaceutical Chemistry, University of California, San Francisco, USA

19 Janin, J. and Chothia, C. *J. Mol. Biol.* 1976, **100,** 197