

## A molecular model of a point mutation (Val297Met) in the serine protease domain of protein C

Kyung Soon Song<sup>1,2</sup>, Young Sook Park<sup>1</sup>, Jong Rak Choi<sup>1</sup>, Hyun Kyung Kim<sup>1</sup> and Quehn Park<sup>1</sup>

<sup>1</sup> Department of Clinical Pathology, College of Medicine, Yonsei University, Seoul 120-752, Korea

<sup>2</sup> Corresponding author: Tel, 82-2-361-6470; Fax, 82-2-313-0956; E-mail, kssong@yumc.yonsei.ac.kr

Accepted 11 March 1999

Abbreviations: PCR, polymerase chain reaction; PC, protein C

### Abstract

**A heterozygous GTG to ATG (Val297Met) mutation was detected in a patient with inherited protein C deficiency and deep vein thrombosis. Cosegregation of the mutation with protein C deficiency was observed through a family pedigree study. Molecular models of the serine protease domains of wild type and mutant protein C were constructed by standard comparative method. Val 297 was found to be located in the hydrophobic core of the protein. Although the substitution of Met for Val does not greatly alter the hydrophobicity of the protein, it introduces a bulkier side chain, which yields steric hindrance between this residue and adjacent residues, such as Met364, Tyr393, Ile321, Ile323, and Val378. It seems that the Met can not fit into the tight packing into which it is trapped, thereby probably inducing misfolding and/or greater instability of the protein. Such misfolding and/or instability thereby eventually disturbs the catalytic triad, in consistent with the observed type I deficiency state.**

**Keywords:** protein C deficiency, deep vein thrombosis, polymerase chain reaction (PCR), direct sequencing, point mutation, protein modeling

### Introduction

Protein C (PC) is a vitamin K-dependent zymogen of a serine protease that plays a role in the down-regulation in the blood coagulation cascade (Kisiel *et al.*, 1976). Mature PC is a two chain glycoprotein with a light chain of 155 amino acids and a heavy chain of 262 amino acids connected by a disulfide bond (Bechmann *et al.*, 1985). The light chain consists of an amino-terminal domain

containing nine  $\gamma$ -carboxyglutamic acid (Gla) residues and two epidermal growth factor-like domains, and the heavy chain is a serine protease domain (Reitsma, 1997).

Human PC is encoded by a gene on chromosome 2q13-14 and spans about 11 kb (Foster and Davie, 1984). The coding sequence for mature mRNA is contained in 9 exons that separated by 8 introns (Plutzky *et al.*, 1986).

Individuals with hereditary PC deficiency tend to have an increased risk of thromboembolism (Griffin *et al.*, 1981) and a number of patients have been reported with severe homozygous or compound heterozygous PC deficiency as well as mild heterozygous PC deficiency (Reitsma *et al.*, 1991; Grandrille and Aiach, 1995; Alessi *et al.*, 1996). Two major subtypes of heterozygous PC deficiency have been delineated using immunological and functional assays. The classical or type 1 deficiency state is the most common form and is characterized by a reduction in both the immunologic and biologic activity of PC in the blood, while in families with a type II deficiency state, affected individuals have normal PC levels on immunologic examination yet possess lowered functional levels of PC (Reitsma *et al.*, 1995).

More than 160 different mutations have been described for the PC gene and are compiled on a regular basis in a database (Reitsma *et al.*, 1995).

In the present study, we characterized the gene defect of a family with type I PC deficiency. The recently reported X-ray crystal structure of serine protease domain of the activated PC (Mather and Oganessyan, 1996) was selected as the template for a molecular modeling to observe the change of the tertiary structure of the serine protease domain of PC due to the gene defect.

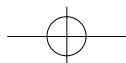
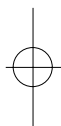
### Materials and Methods

#### Patients

The proband (I-5 in Figure 1) is 46 years old female and was suffered from the first episode of deep vein thrombosis, which was discovered by painful swelling of her right leg and venography. The patient, on referral, was on oral anticoagulant therapy and exhibited ecchymotic areas on both buttocks. Other inhibitory proteins associated with inherited thrombophilia including antithrombin III and protein S were normal in plasma activity (92% and 71%, respectively). In other members of her family, no thrombotic events had been observed.

#### Preparation and analysis of genomic DNA

Blood samples were drawn from the patient, her family members and normal individuals. Nine parts of blood



were mixed with one part of 3.8% trisodium citrate (v/v), and the plasma was separated by centrifugation at 2,500 g for 20 min at 4°C, then stored at -80°C until use. Genomic DNA was isolated from peripheral blood leukocytes as described previously (Miyata *et al.*, 1995).

**Protein C assays**

PC amidolytic activity was measured using Stachrom kit (Stago, Asnieres, France) by STA coagulyzer (Stago). Enzyme linked immunosorbent assay was performed for PC antigen using Asserachrom kit (Stago).

**Identification of mutation**

The exons and flanking intron regions of the PC gene were amplified by polymerase chain reaction (PCR) as described previously (Miyata *et al.*, 1994). The nucleotides of PC gene are numbered according to Foster *et al* (Foster *et al.*, 1986). Direct sequencing was performed with the USB PCR product sequencing kit using [ $\alpha$ -<sup>35</sup>S] dATP (1,000 Ci/mmol, Amersham, Buckinghamshire, UK). The primers used for direct sequencing were the same as those used for PCR.

**Molecular modeling**

The X-ray crystal structure of serine protease domain of the activated protein C has been reported recently (Mather and Oganessyan, 1996) and the coordinates of which are stored in PDB database (PDB ID code, 1AUT). The chimeric reference protein was not constructed for this

study and a known crystal structure of serine protease domain, instead, was selected as the template in a mole-cular modeling. All energy calculations were performed on an Alpha workstation using XPLOR 3.8 software through conjugated gradient energy minimization in order to obtain the best structural model. The constructed model showed the structurally conserved region among serine protease domain and the well-conserved catalytic triad. To obtain the mutant model of the serine protease domain of the activated PC, a native valine residue was replaced with a mutant methionine residue and then subjected it to energy minimization.

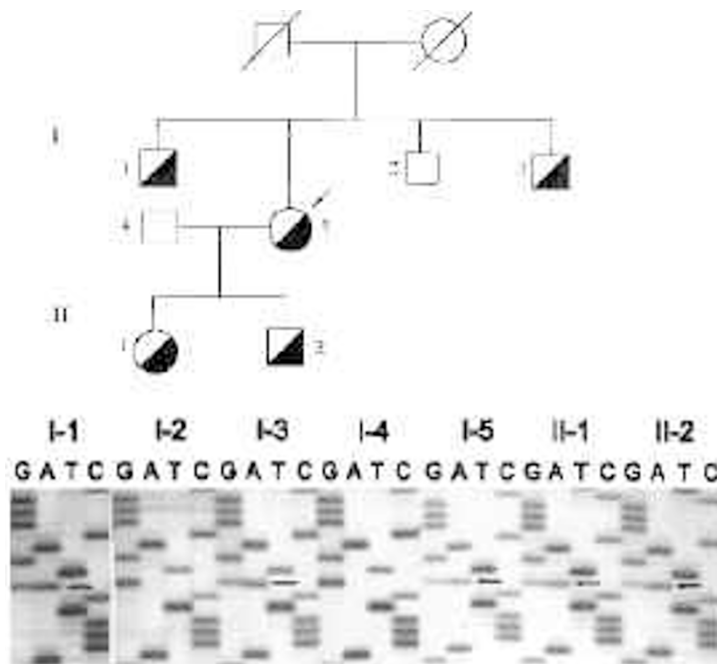
**Results**

**Protein C levels in the family members**

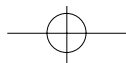
Plasma levels of both PC antigen and PC activity of the propositus (I-5), her younger brother (I-3), her daughter (II-1) and son (II-2) were low, which are consistent with type 1 protein C deficiency (Table 1). PC levels of other family members were in the normal ranges (70-130%).

**Genetic study**

The molecular basis of the protein C deficiency was determined by genetic analysis and a heterozygous missense mutation (G8604A) in exon 9 was detected in the propositus and several family members as shown in



**Figure 1.** Missense mutations in the exon 9 of the protein C gene in the propositus (arrow) and in her family pedigree. Half solid symbols represent affected male and female subjects with mutated protein C gene defect and heterozygosity for protein C deficiencies. Sequence analysis (below) represent the G to A substitution at nucleotide 8604 (arrow) for the individual affected subjects, which causes the amino acid substitution (Val297Met).



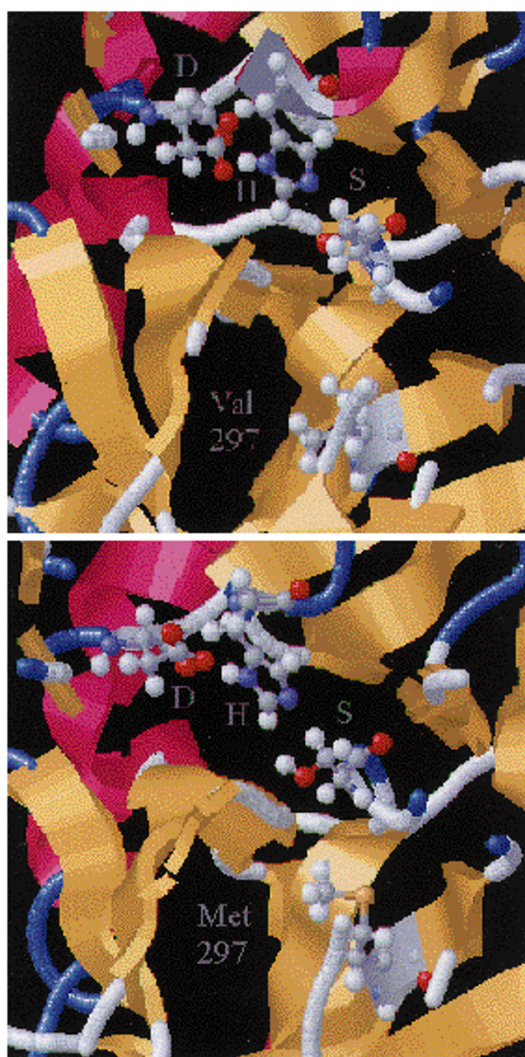
**Table 1.** Protein C levels on the family members. Concordant reductions in PC activity and PC antigen in affected family members (I-3, I-5, II-1, II-2) are consistent with type I PC deficiency and 4 of 5 affected members with the same mutation (except I-1) were cosegregated with PC deficiency.

Pedigree	I-1	I-2	I-3	I-4	I-5	II-1	II-2
Age(yr)	52	43	40	48	46	24	21
Activity(%)	78	102	10	95	3	52	58
Antigen(%)	73	100	5	110	1	48	43

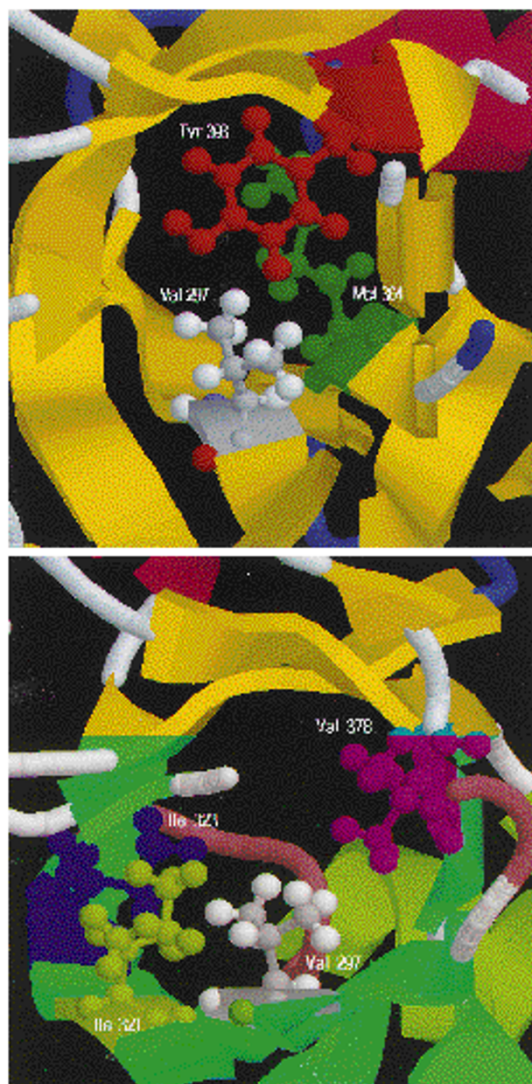
Figure 1, where GTG coding for Val 297 in the serine protease domain of PC was changed to ATG coding for Met. This mutation was cosegregated with heterozygous type I protein C deficiency in 3 family members (I-3, II-1, II-2).

**Molecular modeling**

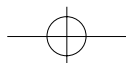
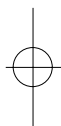
The constructed model showed the structurally conserved region among serine protease domain and the well-conserved catalytic triad (Figure 2A). The Val297 residue



**Figure 2.** Molecular model of serine protease domain of activated protein C. The  $\alpha$ -carbon- $\alpha$ -carbon 'bonds' are shown as white;  $\alpha$  helices are shown as red spirals;  $\beta$ -sheets are shown as yellow arrows. H, D, and S represent the residues of the catalytic triad; H: His211, D: Asp257, S: Ser360. Panel A: The wild type residue, Val297 is located in the hydrophobic core of the protein and the catalytic triad is highly conserved. Panel B: The mutated residue, Met297 appears to change the core conformation and, therefore, the catalytic triad is disturbed.



**Figure 3.** Detailed representations of interactions between Val297 and neighboring amino acids. Panel A: Met364 (green) and Tyr393 (red), against Val297. Panel B: Ile321 (purple), Ile323 (green) and Val378 (cobalt blue), against Val297.



where the mutation occurs is located deeply in the hydrophobic core of the protein and surrounded by a cluster of hydrophobic amino acids, such as Met364, Tyr393, Ile321, Ile323, and Val378 (Figure 3). The substitution of Met for Val introduces a bulkier side chain, which yields a steric disturbance between these residues. Such alleviated stresses appear to disrupt the tight packing required to form the protein core and interfere with the folding and stability of the molecule, together with a catalytic site eventually disturbed, as shown in Figure 2B. Possible explanation for a type I phenotype due to Val297Met substitution could be responsible for a conformational change of the protein, which leads to the protein being unable to fold stably.

## Discussion

Genomic analysis of a patient with type 1 protein C deficiency and thrombotic complications allowed us to detect a missense mutation (G8604A) in the propositus and in her family members. This mutation cosegregated with the protein C deficiency in 4 of 5 family members. Thus, this mutation identified in this family is likely to have a detrimental effect. However, other family members had not developed thrombosis by the time of testing. Furthermore, one (I-1) of family member with same mutation in this study had normal levels of PC antigen and activity. Therefore, the presence of the same mutation in a family showing a clinically recessive pattern of inheritance indicates that unidentified factors, apart from the type of protein C gene mutation, are likely to modulate the phenotypic expression of heterozygous protein C deficiency.

In the 1995 update of the PC database of mutations, about 30% of different mutations have occurred at CpG dinucleotides and are C to T or G to A transitions, comparable with a model of methylation-mediated deamination (Reitsma, 1996). As one of G to A transition, a novel heterozygous GTG to ATG (Val297Met) substitution was detected in an individual with probable inherited protein C deficiency and both venous and arterial thrombotic disease (Millar *et al.*, 1993) and several unrelated probands have been reported from the different countries (Reitsma *et al.*, 1995).

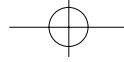
Several research groups have constructed molecular models of the serine protease domain of protein C (Wacey *et al.*, 1993; Krawczak *et al.*, 1996; Zheng *et al.*, 1996). These computer models have been used to study missense mutations causing protein C deficiency and amino acid substitutions causing type 1 deficiency were predicted to yield energetically unfavorable proteins by the adverse interactions with neighboring amino acids.

To investigate the structural defects of the serine protease domain caused by Val297Met substitution, a molecular model of the tertiary structure of the serine protease domain of the activated protein C was constructed

on the basis of the known X-ray crystal structure of the serine protease domain (Figure 2). The structural model indicates that Val297 is located deeply in the hydrophobic core of the protein. Val297 has a relatively small non-polar side chain and a cluster of hydrophobic amino acids surrounds this residue (Figure 3). Although the substitution of Met for Val does not greatly alter the hydrophobicity of the protein, it introduces a bulkier side chain that appears to generate steric hindrance between this mutated residue and surrounding amino acids. Thus, it seems that the Met can not fit into the tight packing into which it is trapped, thereby inducing misfolding and greater instability of the protein, together with a catalytic site eventually disturbed. The model of the mutated serine protease domain of the protein C shows that the  $\alpha$ -helical backbone possessing His211 in the catalytic site is not highly conserved and the catalytic triad, therefore, could not orient precisely, as shown in Figure 2B. The loss of both activity and antigen levels in Val297Met mutant would be associated with disruption of correct protein folding and structural instability of the protein, a condition associated with rapid intracellular degradation of the protein. Our findings seem to be consistent with the observation of the same genetic defect, which has been reported previously (Millar *et al.*, 1993). These molecular modeling experiments may help understanding of pathogenesis of hereditary protein C deficiency.

## References

- Alessi, M. C., Aillaud, M. F., Paut, O., Roquelaure, B., Alhencgas, M., Pellissier, M. C., Ghanen, N. and Juhan-Vague, I. (1996) Purpura fulminans in a patient homozygous for a mutation in the protein C gene-Prenatal diagnosis in a subsequent pregnancy. *Thromb. Haemost.* 75: 525-526
- Beckmann, R. J., Schmidt, R. J., Santerre, R. F., Plutzky, J., Crabtree, G. R. and Long, G. L. (1985) The structure and evolution of a 461 amino acid human protein C precursor and its messenger RNA, based upon the sequence of cloned human liver cDNAs. *Nucleic acids Res.* 13: 5233-5247
- Foster, D. and Davie, E. W. (1985) Characterization of a cDNA coding for human protein C. *Proc. Natl. Acad. Sci. USA* 81: 4766-4770
- Foster, D. C., Yoshitake, S. and Davie, E. W. (1986) The nucleotide sequence of the gene for human protein C. *Proc. Natl. Acad. Sci. USA* 82: 4673-4677
- Grandrille, S. and Aiach, M. (1995) French INSERM Network on Molecular Abnormalities Responsible for Protein C and Protein S Deficiencies. *Blood* 86: 2598-2605
- Griffin, J. H., Evatt, B., Zimmerman, T. S., Kleiss, A. J. and Wideman, C. (1981) Deficiency of protein C in congenital thrombotic disease. *J. Clin. Invest.* 68: 1370-1373
- Kisiel, W., Ericsson, L. H. and Davie, E. W. (1976) Proteolytic activation of protein C from bovine plasma. *Biochemistry* 15: 4893-4900
- Krawczak, M., Wacey, A. and Cooper, D. N. (1996) Molecular reconstruction and homology modelling of the catalytic domain of the common ancestor of the haemostatic vitamin K dependent serine proteinases. *Hum. Genet.* 98: 351-370
- Mather, T., Oganessyan, V., Hof, P., Huber, R., Foundling, S., Esmon, C. and Bode, W. (1996) The 2.8 Å crystal structure of Gla-domainless activated protein C. *EMBO. J.* 15: 6822-6831
- Millar, D. S., Wacey, A. I., Voke, J., Kakkar, V. V. and Cooper, D. N. (1993) A novel point



mutation (Val 297-> Met) in the serine proteinase domain of protein C in a patient with both venous and arterial thromboembolic disease. *Blood Coagul. Fibrinol.* 4: 631-633

Miyata, T., Zheng, Y. Z., Sakata, T., Tsushima, N. and Kato, H. (1994) Three missense mutations in the protein C heavy chain causing type I and type II protein C deficiency. *Thromb. Haemost.* 71: 32-37

Miyata, T., Zheng, Y. Z., Sakata, T. and Kato, H. (1995) Protein C Osaka 10 with aberrant propeptide processing: Loss of anticoagulant activity due to an amino acid substitution in the protein C precursor. *Thromb. Haemost.* 74: 1003-1008

Plutzky, J., Hoskins, J. A., Long, G. L. and Crabtree, G. R. (1986) Evolution and organization of the human protein C gene. *Proc. Natl. Acad. Sci. USA* 83: 546-550

Reitsma, P. H., Poort, S. R., Allaart, C. F., Briet, E. and Bertina, R. M. (1991) The spectrum of genetic defects in a panel of 40 Dutch families with symptomatic protein C deficiency type I: Heterogeneity and founder effects. *Blood* 78: 890-894

Reitsma, P. H., Bernardi, F., Doig, R. G., Gandrille, S., Greengard, J. S., Ireland, H., Krawczak, M., Lind, B., Long, G. L., Poort, S. R., Saito, H., Sala, N., Witt, I. and Cooper, D. N. (1995) Protein C deficiency: A data-base of mutations, 1995 update. *Thromb. Haemost.* 73: 876-889

Reitsma, P. H. (1996) Protein C deficiency: summary of the 1995 data-base update.

*Nucleic Acids Res.* 24: 157-159

Reitsma, P. H. (1997) Protein C deficiency: from gene defects to disease. *Thromb. Haemost.* 78: 344-350

Wacey, A. I., Pemberton, S., Cooper, D. N., Kakkar, V. V. and Tuddenham, E. G. D. (1993) A molecular model of the serine protease domain of the activated protein C: Application to the study of missense mutations causing protein C deficiency. *Br. J. Haematol.* 84: 290-300

Zheng, Y.-Z., Sakata, T., Matsusue, T., Umeyama, H., Kato, H. and Miyata, T. (1996) Six missense mutations associated with type I and type II protein C deficiency and implications obtained from molecular modelling. *Blood Coagulation and Fibrinolysis* 5: 687-696

