doi:10.1016/j.jmb.2008.06.034 J. Mol. Biol. (2008) 381, 913–927

JMB

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

# Structure of the Leech Protein Saratin and Characterization of Its Binding to Collagen

**Wolfram Gronwald[1,2], Jörg Bomke[1,3], Till Maurer[1], Barbara Domogalla[1], Fritz Huber[1,4], Frank Schumann[1], Werner Kremer[1], Florian Fink[2], Thomas Rysiok[5], Matthias Frech[3] and Hans Robert Kalbitzer[1]***

[1]*Institute of Biophysics and Physical Biochemistry, University of Regensburg, 93040 Regensburg, Germany*

[2]*Institute of Functional Genomics, University of Regensburg, 93040 Regensburg, Germany*

[3]*Molecular Interactions and Biophysics, Merck KGaA, Frankfurter Str. 250, D-642993 Darmstadt, Germany*

[4]*Lipofit GmbH, Regensburg, Germany*

[5]*Protein and Cell Sciences, Merck KGaA, Frankfurter Str. 250, D-642993 Darmstadt, Germany*

*Received 13 March 2008; received in revised form 9 June 2008; accepted 11 June 2008*
Available online 18 June 2008

***Edited by M. F. Summers***

The leech protein Saratin from *Hirudo medicinalis* prevents thrombocyte aggregation by interfering with the first binding step of the thrombocytes to collagen by binding to collagen. We solved the three-dimensional structure of the leech protein Saratin in solution and identified its collagen binding site by NMR titration experiments. The NMR structure of Saratin consists of one α-helix and a five-stranded β-sheet arranged in the topology ββαβββ. The C-terminal region, of about 20 amino acids in length, adopts no regular structure. NMR titration experiments with collagen peptides show that the collagen interaction of Saratin takes place in a kind of notch that is formed by the end of the α-helix and the β-sheet. NMR data-driven docking experiments to collagen model peptides were used to elucidate the putative binding mode of Saratin and collagen. Mainly, parts of the first and the end of the fifth β-strand, the loop connecting the α-helix and the third β-strand, and a short part of the loop connecting the fourth and fifth β-strand participate in binding.



*Keywords:* Saratin structure; collagen interaction; leech protein; NMR; hemostasis

## Introduction

Collagen plays an important structural role in the extracellular matrix of tissues, and under normal circumstances, it is not exposed to the flowing blood and its constituents. Upon injury of the vessel wall, collagen is exposed to the circulating cells and proteins, leading to activation and release of a variety of proaggregatory factors propagating aggregation and thrombosis.

In this process, thrombocytes recognize exposed collagen via the von Willebrand factor (vWF), which leads to their activation. The vWF hereby acts as a bridge between collagen and thrombocyte receptors.[1] Current findings show that depending on the collagen

*Corresponding author. E-mail address: hans-robert.kalbitzer@biologie.uni-regensburg.de.

Abbreviations used: vWF, von Willebrand factor; LAPP, leech antiplatelet protein; 3D, three-dimensional; NOE, nuclear Overhauser enhancement; hetNOE, heteronuclear NOE; NOESY, NOE spectroscopy; HSQC, heteronuclear single quantum coherence; RCSB, Research Collaboratory for Structural Bioinformatics; PDB, Protein Data Bank; SPR, surface plasmon resonance; TOCSY, total correlated spectroscopy.

type, either the A1 or the homologous A3 domain of vWF participates in this process. For both domains, X-ray structures have been solved, showing a typical α/β fold that contains a central β-sheet flanked by α-helices on each side.[2–4] The A1 domain of vWF appears to mediate binding to collagen type VI in the subendothelial matrix. At high shear rates, the A3 domain of vWF binds to collagen of types I and III in perivascular connective tissue,[4,5] leading to a conformational change of the vWF that enables its binding to the thrombocyte receptors GPIb/IX/V.[6] Together with the corresponding X-ray structure of the A3 domain and mutation studies, it could be shown that predominantly positively charged residues such as His1023 of the A3 domain participate in collagen binding and that the binding site is relatively flat and flexible.[4] However, presently, no structure of the A3 domain in complex with collagen is available. After initial binding, the thrombocytes are irreversibly attached to the collagen receptors GPIa/IIa, GPVI, GPIV, p65, and TIIICBP.[7] This, in turn, leads to thrombocyte activation, degranulation, and aggregation.[8] The activated thrombocytes release thromboxane A2, cytokines, mitogenic mediators, and vasoconstriction factors[9] that contribute, among others, to thrombus formation.[10] Therefore, inhibition of the first step of thrombocyte binding may lead to a more effective inhibition of the thrombocyte response in injured vessels.

In general, the hemostasis is indispensable for the life of vertebrates. In the event of external or internal injury, blood coagulation prevents bleeding to death. However, in some cases, for example, patients with a high disposition to blood coagulation, for a prevention of heart attacks, it is vitally important to suppress this mechanism. Different substances such as heparin, phenprocoumon, and acetylsalicylic acid have been in use for this purpose for many years. In 2001, a further coagulation inhibitor, the protein Saratin, was characterized.[11] Saratin, with a molecular mass of 12.06 kDa, was isolated from the saliva of the leech *Hirudo medicinalis.* It inhibits the first step of thrombocyte binding, that is, the interaction with the vWF by binding to collagen, presumably by binding to an overlapping epitope.[12] Therefore, Saratin affects the initial step of the blood coagulation, whereas the substances mentioned above affect coagulation later in the coagulation cascade. As a consequence, thrombocyte aggregation to vessel walls and thrombus formation is reduced.[11] Therefore, Saratin could be a powerful therapeutic component that locally prevents coagulation without interfering with the normal hemostatic functions. This view is supported by first studies that show the antithrombotic effect of Saratin on human atherosclerotic plaques.[13] As mentioned above, it is known that Saratin binds to collagen but the exact mode of interaction has still to be revealed. What is also known is that low concentrations of Saratin have no effect on vWF-independent collagen-induced platelet aggregation, whereas high Saratin concentrations inhibit both mechanisms.[11]

Additionally, identified molecules that interact with collagen include, as mentioned above, the A1 and A3 domain of the vWF, the homologous I domain of integrin α2β1,[14] and the leech antiplatelet protein (LAPP) from the leech *Haementeria officinalis*.[15] However, the A3 domain of the vWF and the I domain of integrin α2β1 interact in a different fashion with collagen. First, the collagen binding regions are located on different sites of the two molecules. Secondly, the binding site of the A3 domain is relatively flat[4] while integrin exhibits a distinct notch for collagen binding.[14] In addition, a metal ion participates vitally in collagen recognition of integrin, which is not required for the A3 domain. For LAPP, it is currently not known whether its collagen binding site is located in its flexible N-terminal domain or in the well-folded C-terminal part. Structural comparisons of LAPP with the collagen binding domains of the vWF and integrin did not reveal similar features from which the collagen binding site of LAPP could be identified. A recent study shows that Saratin interferes with the collagen interaction of the A3 domain of the vWF and of the I domain of integrin α2β1, indicating that overlapping binding epitopes on collagen may exist.[12] Structural information on the Saratin–collagen interaction is not available yet. As a prerequisite for the solution of these questions, the three-dimensional (3D) solution structure of Saratin has to be determined, a result we will present in this article. In addition, using the well-studied soluble triple-helical collagen model peptide (Gly–Lys–Hyp–(Gly–Pro–Hyp)$_{10}$–Gly–Lys)$_3$, which can replace collagen from natural sources in many respects,[16–18] and a data-driven docking algorithm,[19] we investigated the putative Saratin collagen interaction.

## Results

### NMR resonance assignment and secondary structure

The sequential assignment of $^{15}$N, $^{13}$C-enriched Saratin by multidimensional NMR spectroscopy has been described elsewhere.[20] These data (Biological Magnetic Resonance Bank entry 4973) were completed for the present study and deposited in the data base. More than 96% of the backbone carbonyl carbons, amide nitrogens, amide protons, α-carbons, α-protons, β-carbons, and β-protons were assigned as well as more than 95% of the side-chain protons and carbons. Additionally, 53% of the $^{13}$C shifts of the aromatic rings were assigned. The analysis of the $^{1}H^{\alpha}$, $^{13}C^{\alpha}$, $^{13}C'$, and $^{13}C^{\beta}$ chemical shifts employing the chemical shift index method[21] predicts six β-strands (B0: from Trp7 to Phe9, B1*: from Asn12 to Tyr15, B2: from Phe22 to Asp27, B3: from Tyr42 to Asp47, B4: from Glu52 to Tyr55, and B5: from Phe74 to Leu79), one α-helix (from Leu28 to Phe36), and, furthermore, an unstructured C-terminal part (approximately from