Rapid Communication

J. Biochem. **134**, 19–23 (2003)
DOI: 10.1093/jb/mvg108

Characterization and Preliminary Crystallographic Studies of EMS16, an Antagonist of Collagen Receptor (GPIa/IIa) from the Venom of *Echis multisquamatus*

Daiju Okuda[1], Katsunori Horii[2], Hiroshi Mizuno[2] and Takashi Morita[*,1]

[1]*Department of Biochemistry, Meiji Pharmaceutical University, 2-522-1, Noshio, Kiyose, Tokyo, 204-8588; and*
[2]*Department of Biochemistry, National Institute of Agrobiological Sciences, 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602*

Received March 13, 2003; accepted April 21, 2003

EMS16 is a member of the snake venom-derived C-type lectin family of proteins (CLPs) found in the venom of *Echis multisquamatus*. It binds to glycoprotein Ia/IIa (integrin α2β1), a major collagen receptor of platelets, acting as a potent antagonist of platelet aggregation and cell migration. Amino acid sequencing and cDNA cloning of EMS16 have revealed that it is composed of an A chain of 134 amino acid residues and a B chain of 128 residues. Crystals of EMS16 belong to space group $P2_12_12_1$, with unit-cell parameters a = 46.57, b = 59.93, and c = 115.74 Å, and diffract to a resolution of 1.9 Å. Phase determination is underway by means of molecular replacement with the structure of blood coagulation factor IX-binding protein (IX-bp) from habu snake venom (PDB code 1bj3) as the search model.

Key words: collagen receptor, C-type lectin-like protein, *Echis*, Glycoprotein Ia/IIa, snake venom.

Snake venom contains many protein components that affect haemostasis and the nervous system in various ways. A great many of these venom proteins are classified into superfamilies, such as metalloproteases, phospholipases and ion channel antagonists. The C-type lectin family proteins (CLPs) are some of the major components of snake venom. Even though these proteins have regions of amino acid sequence homology, the protein targets differ: some are directed at the haemostatic system, in which blood coagulation factors play an important role, while others are directed at plasma proteins and platelet glycoproteins. Recently, we reported the X-ray crystal structures of several CLPs, blood coagulation factors IX/X binding protein (IX/X-bp) (*1*), factor X binding protein (X-bp) (*2*), factor IX binding protein (IX-bp) (*3*), flavocetin-A (*4*), and bitiscetin (*5*). They all have a heterodimeric structure (αβ) or form heterodimeric tetramers $(\alpha\beta)_4$, and also have a unique 3D domain-swapping loop in their molecules.

The glycoprotein Ia/IIa complex (also commonly referred to as "integrin α2β1") is a major collagen receptor found on platelet membranes and is responsible for platelet activation with another collagen receptor, GPVI. Many researchers have studied the mechanisms of interaction between GPIa/IIa and GPVI during platelet aggregation, but because there are numerous complexes produced during signal transduction, as well as other factors, the true mechanism is not clearly understood. Currently, snake venom CLPs such as convulxin (*6*, *7*), a GPVI-agonist, and aggretin (rhodocytin) (*8*, *9*), a GPIa/IIa-agonist, are used as tools for studying the mechanism of collagen-induced platelet aggregation.

Recently, a new CLP, EMS16 from the venom of *Echis multisquamatus*, was reported by Marcinkiewicz *et al.* (*10*) as the first selective GPIa/IIa antagonist from snake venom. It binds to the integrin α2-I domain, which is the collagen-binding site of integrin α2, causing inhibition of collagen-induced platelet aggregation and cell migration.

The X-ray crystal structure of the integrin α2-I domain was previously reported by Emsley *et al*. (*11*). As the first step to clarify the binding mechanism in a complex of integrin α2-I domain and EMS16, the amino acid sequence of EMS16 was deduced by cDNA cloning, and its crystal structure analysis is underway.

MATERIALS AND METHODS

Materials—Pre-packed columns, Superdex 75 pg and S-Sepharose High Performance, were purchased from Amersham Pharmacia Biotech (UK), the COSMOSIL 5C18 AR-300 HPLC column was from Nacalai Tesque (Kyoto), endoprotease Asp-N was from Boehringer Mannheim, and N-glycosidase F was from Pierce. The DNA polymerase Advantage 2 polymerase mix and Pyrobest were purchased from CLONTECH (CA) and TaKaRa (Shiga), respectively. Chemicals of reagent grade were obtained from Amersham Bioscience (UK), Sigma (MO), or Wako Pure Chemicals (Osaka).

Protein Purification—The crude venom of *E. multisquamatus* was dissolved in 50 mM Tris-HCl buffer, pH 8.0, and insoluble materials were removed by centrifugation. The supernatant was fractionated by gel filtration with Superdex 75pg and collected fractions were assayed by determining the cross-reactivity with anti-IX/X-bp antibody. The cross-reacting fractions were pooled and purified by successive chromatographies on a Q-Sepharose HP column and then a reversed-phase HPLC column (COSMOSIL 5C18 AR-300). EMS16 was identified by N-

*To whom correspondence should be addressed. Tel/Fax: +81-424-95-8479, E-mail: tmorita@my-pharm.ac.jp

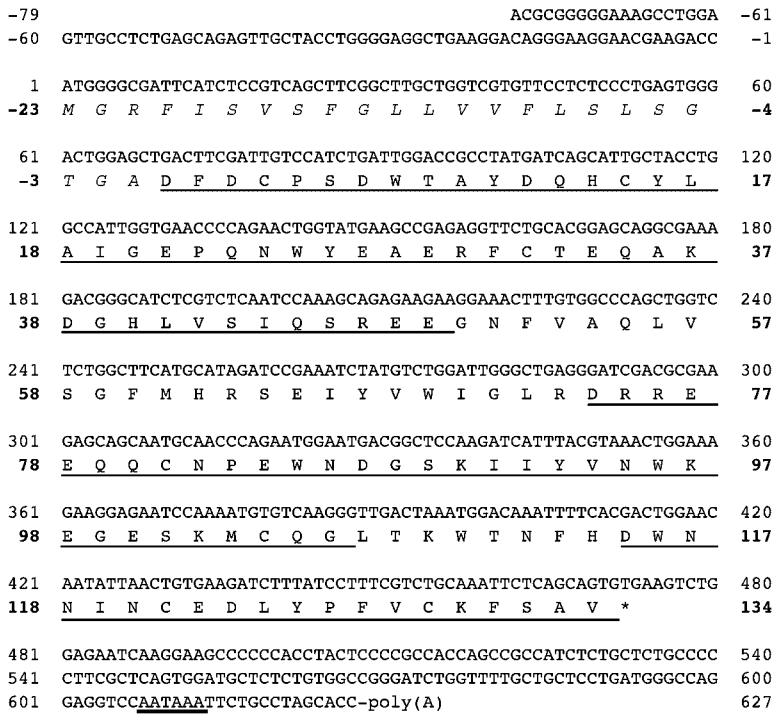

Fig. 1. **cDNA and deduced amino acid sequence of the EMS16 A chain.** The numbering of nucleotides and amino acids (bold) are indicated on both sides of the sequence. 1 indicates the position of the start codon and the N-terminus of the EMS16 A chain, respectively. The putative signal peptide, mature EMS16 A chain, and polyadenylation signal are indicated by italics, underlining and bold underlining, respectively. The cDNA sequence of the EMS16 A chain has been deposited in GenBank under accession number AB098253.

terminal amino acid sequence analysis of the intact and *S*-pyridylethylated proteins. Deglycosylation of EMS16 was carried out by treatment with *N*-glycosidase F at 37°C for 48 h.

Determination of the cDNA Sequence—cDNA cloning of EMS16 was carried out by the RT-PCR method using total RNA isolated from the venom gland of *E. multisquamatus*. The template for RACE-PCR was prepared by RT-PCR using MMLV reverse transcriptase. 5′- and 3′-RACE were carried out to determine the nucleotide sequences of both the 5′- and 3′- cDNA ends using Pyrobest (TaKaRa) with degenerate primers based on partial amino acid sequences of EMS16, and single primers identified by the nucleotide sequence. The nucleotide sequence was determined with a DSQ2000L DNA sequencer (Shimadzu, Kyoto)

Crystallization of EMS16—Initial crystallization experiments on EMS16 were performed using the microbatch technique (*12*) with Crystal Screen I (Hampton Research). Crystals were grown at 293 K under a thin layer of Al's oil (Hampton Research) on a 96-well plate (Nunc), 0.6 μl of the purified protein solution (5 mg/ml in 50 mM Tris

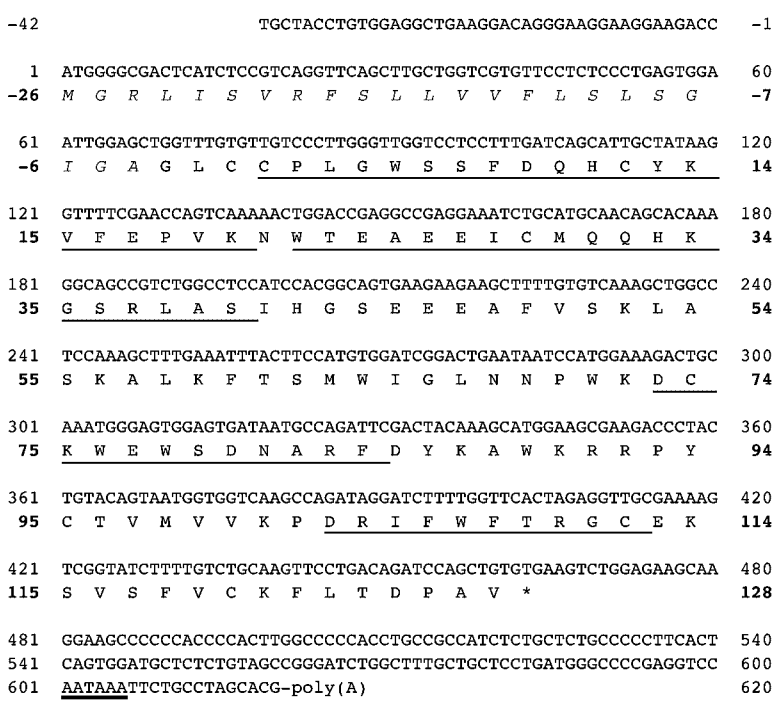

Fig. 2. **cDNA and deduced amino acid sequence of the EMS16 B chain.** The numbering of nucleotides and amino acids (bold) are indicated on both sides of the sequence. 1 indicates the position of the start codon and the N-terminus of the EMS16 B chain, respectively. The putative signal peptide, mature EMS16 B chain and polyadenylation signal are indicated by italics, underlining and bold underlining, respectively. The cDNA sequence of the EMS16 B chain has been deposited in GenBank under accession number AB098254.

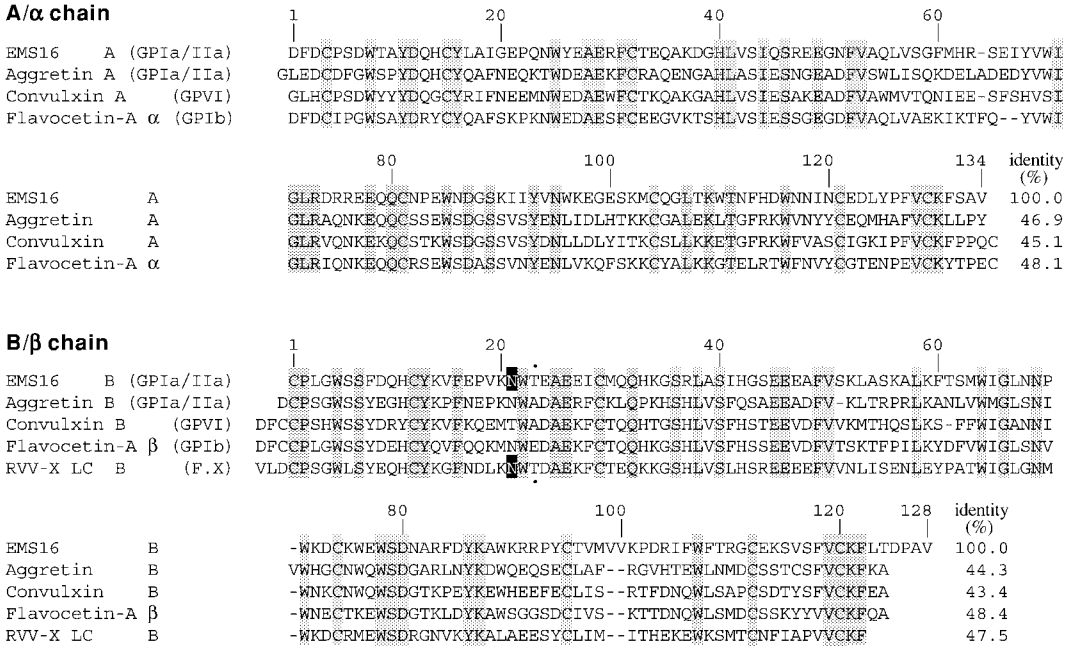

Fig. 3. **Alignment of the amino acid residues in the primary structures of snake venom C-type lectin proteins.** The numbering at the top refers to EMS16.The ligand of each CLP is shown in parentheses. The conserved amino acid residues are shaded and the positions of N-glycosylated Asn residues have a black background. The dots indicate the location of the consensus sequence of the N-glycosylation site in EMS16 and the RVV-X light chain B (*13*). The unconserved cysteine residues, including the C-terminus of the A chain and the N-terminal portion of the B chain of convulxin and flavocetin-A, and Cys115 in the B chain of aggretin, bridge the heterodimeric units (αβ) to give multimeric molecules $(\alpha\beta)_n$. RVV-X LC, Russell's viper venom coagulation factor X activator light chain.

HCl buffer at pH 7.5) being mixed with 1.2 μl of the precipitant solution. Initial investigation revealed protein microcrystals to be present in solutions containing the precipitant PEG8000 (Crystal Screen I: No. 42). The crystallization conditions were thereafter optimized by adjusting the volume ratio of protein solution to reservoir solution using the hanging-drop vapor-diffusion method. After mixing 1 μl of protein solution and 3 μl of precipitant solution (0.04 M potassium dihydrogen phosphate, 16% PEG8000, 20% glycerol suspended over 0.5 ml of reservoir solution), good diamond-shaped crystals appeared in about a week and grew to dimensions of 0.4 × 0.4 × 0.1 mm (Fig. 5).

RESULTS AND DISCUSSION

EMS16 was purified by three-step column chromatography (yield: 1.1 mg/g venom). The N-terminal sequences of the *S*-pyridylethylated (Pe-) subunits were NH_2-DFDCPSDWTAYDQHCYLAIGE for the Pe-A chain and NH_2-CPLGWSSFDQHCYKVFPVKXWTEAE for the Pe-B chain. These sequences are identical to those of EMS16, which were previously reported by Marcinkiewicz *et al*. (*10*). The *S*-pyridylethylated subunits of EMS16 were digested with endoprotease Asp-N and the peptide fragments were analyzed as to their amino acid sequences. The PTH-amino acid of residue 21 in the B chain could not be detected on sequencing of either the Pe-B chain or the Asp-N digested fragment.

The cDNAs encoding the EMS16 subunits were obtained by the RT-PCR method from the total RNA of the venom gland of *E. multisquamatus*. The nucleotide sequence of the EMS16 A chain comprised 706-bp, including a 79 bp 5′-untranslated region (UTR), a 479 bp open reading frame (ORF), and a 156 bp 3′-UTR (Fig. 1), while the B chain comprised 662-bp, including a 42 bp 5′-UTR, a 462 bp of ORF, and a 158 bp 3′-UTR (Fig. 2). Figure 3 shows the deduced amino acid sequence of mature EMS16 compared with those of other snake venom-derived CLPs that bind to platelet glycoproteins. The amino acid sequence of EMS16 was approximately 50% homologous with those of other snake CLPs, indicating that EMS16 belongs to the snake venom class of CLPs. The deduced amino acid sequence indicated that the 21st residue of the B chain is Asn. Alignment of the 21st to 23rd residues (Asn-Trp-Thr) suggested the possibility of N-glycosylation of the Asn at position 21 in the B chain,

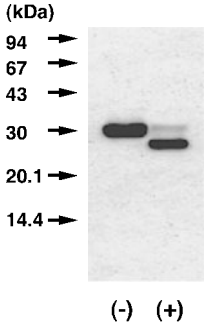

Fig. 4. **N-Glycosidase treatment of EMS16.** EMS16 (2 μg) was treated with *N*-glycosidase F at 37°C for 48 h. (–) without N-glycosidase F treatment; (+) after *N*-glycosidase F treatment. SDS-PAGE was performed under non-reducing conditions.

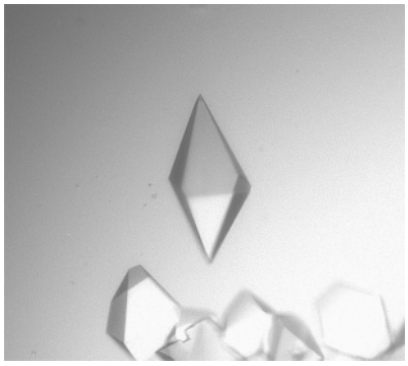

Fig. 5. **Crystals of EMS16.** Photograph of EMS16 crystals grown from PEG8000. The maximum dimension of the crystal was approximately 0.4 mm.

Table 1. **Data collection statistics.** Values in parentheses refer to the highest resolution shell.

Wavelength (Å)	1.5418
Space group	$P2_12_12_1$
Cell dimensions (Å)	a = 46.57 b = 59.93 c = 115.74
Resolution range (Å)	30–1.9 (1.97–1.9)
Measured reflections	91,401
Unique reflections	26,199
I/σ (I)	16.1 (5.6)
Completeness (%)	99.7 (99.7)
Mosaicity (°)	0.6
R_{merge}[a] (%)	3.1 (11.6)

[a]$R_{merge} = \Sigma_h \Sigma_i \mid I(h, i) - \langle I(h)\rangle \mid / \Sigma_h \Sigma_i I(h, i)$, where $I(h, i)$ is the intensity value of the ith measurement of h and $\langle I(h)\rangle$ is the corresponding mean value of $I(h)$ for all i measurements.

which also agrees with the result of amino acid sequencing of the B chain, in which the 21st residue could not be identified. As shown in Fig. 3, RVV-X (*Daboia russelli* viper venom coagulation factor-X activator) has an N-glycoside attached to Asn24 of light chain-B [previously called "light chain–1(LC1)" (*13*)]. The apparent mass of EMS16, as determined by SDS-PAGE, was decreased (31-kDa to 28-kDa) on treatment with N-glycosidase F (Fig. 4), resulting in an approximately 2.4 kDa difference between the 33,292 Da reported for EMS16 by Marcinkiewicz *et al*. (*10*) and the value of 30,942 Da calculated from the deduced amino acid sequence. This indicates that EMS16 very likely has an *N*-glycoside attached to the Asn at position 21 of the B chain.

Diffraction data for EMS16 were collected in-house using an R-AXIS IV^{++} imaging-plate detector mounted on a Cu$K\alpha$ rotating-anode generator equipped with an Osmic mirror system (Rigaku). A crystal was flash-cooled to 120 K in a stream of cold nitrogen-gas (Rigaku Cryogenic System) and then rotated through 90°, with 0.5° oscillation per frame. The raw data were processed and scaled using the program *CRYSTAL CLEAR* ver. 1.3 (Rigaku).

The crystals of EMS16 belong to orthorhombic space group $P2_12_12_1$, with unit-cell parameters a = 46.57, b = 59.93, and c = 115.74 Å, and diffract to a resolution of 1.9 Å. Analysis of the packing density showed that one molecule of EMS16 per asymmetric unit would yield a reasonable solvent content [V_M = 2.6 Å^{3}Da^{-1} and estimated solvent content V_{solv} = 53%; (*14*)]. The statistics for data collection are listed in Table 1.

Structural analysis of EMS16 was carried out by the molecular-replacement method using program *AMoRe* in the *CCP4* suite [Collaborative Computational Project, 1994; (*15*)]. A search model was constructed by homology modeling using *Insight II 2000* (Accelrys) based on the atomic coordinates of factor IX-binding protein (IX-bp) from the venom of the habu snake [PDB code 1bj3; (*3*)]. EMS16 and IX-bp show amino acid sequence homology of 54.7 and 43.9% for the A and B chains, respectively. After rotation and translation calculations, a clear peak was found with a correlation coefficient of 0.624 and an R factor of 39.0% (10–4 Å). Structural refinement is currently in progress using program *CNS* (*16*) and *XtalView* (*17*).

Platelet collagen receptors such as integrin α2β1 (GPIa/IIa) and GPVI play an important role in platelet activation. Many researchers are currently investigating the mechanism of platelet function, including signal transduction, whereby binding of collagen to integrin α2β1 and GPVI leads to platelet activation. The snake venom proteins that are specific for platelet glycoproteins are therefore useful tools for studying the process of platelet activation. For example, convulxin, which is a potent platelet agonist for GPVI, is a major tool being used in the study of the collagen-platelet interaction. However, there is no specific antagonist against collagen receptors, except for monoclonal antibodies to integrin α2β1 and GPVI. EMS16, a snake venom CLP, is the first specific antagonist of integrin α2β1, especially of the integrin α2-I domain, where collagen interacts to initiate platelet activation.

We have previously reported the crystal structures of several CLPs: coagulation factor IX/X-binding protein (bp) (*1*), factor X-bp (*2*), factor IX-bp (*3*), flavocetin-A (*4*), and bitiscetin (*5*). EMS16 shows significant sequence similarity to other CLPs, suggesting that the overall 3D structure of EMS16 should be similar, with unique 3D domain swapping. Determination of the structure of EMS16 and its complex with the integrin α2-I domain should provide valuable information regarding the collagen-integrin α2β1 interaction and permit the design of an anti-haemostasis drug.

This work was supported in part by a Grant-in-Aid for Scientific Research from the Ministry of Education, Science and Culture of Japan (T.M.). We would like to thank Dr. Yuri Utkin (Russian Academy of Science, Moscow, Russian Federation) for helping to catch the *Echis multisquamatus*, and Drs Michihisa Toriba and Atsushi Sakai (Japan Snake Institute, Gunma) for extracting the venom glands. We thank Tomoko Yokomizo and Yuko Tokunaga for their help with the peptide sequencing of EMS16.

REFERENCES

1. Mizuno, H., Fujimoto, Z., Koizumi, M., Kano, H., Atoda, H., and Morita, T. (1997) Structure of coagulation factors IX/X-binding protein, a heterodimer of C-type lectin domains. *Nat. Struct. Biol.* **4**, 438–441
2. Mizuno, H., Fujimoto, Z., Atoda, H., and Morita, T. (2001) Crystal structure of an anticoagulant protein in complex with the

Gla domain of factor X. *Proc. Natl Acad. Sci. USA* **98**, 7230–7234
3. Mizuno, H., Fujimoto, Z., Koizumi, M., Kano, H., Atoda, H., and Morita, T. (1999) Crystal structure of coagulation factor IX-binding protein from habu snake venom at 2.6 A: implication of central loop swapping based on deletion in the linker region. *J. Mol. Biol.* **289**, 103–112
4. Fukuda, K., Mizuno, H., Atoda, H., and Morita, T. (2000) Crystal structure of flavocetin-A, a platelet glycoprotein Ib-binding protein, reveals a novel cyclic tetramer of C-type lectin-like heterodimers. *Biochemistry* **39**, 1915–1923
5. Hirotsu, S., Mizuno, H., Fukuda, K., Qi, M.C., Matsui, T., Hamako, J., Morita, T., and Titani, K. (2001) Crystal structure of bitiscetin, a von Willebrand factor-dependent platelet aggregation inducer. *Biochemistry* **40**, 13592–13597
6. Berlanga, O., Tulasne, D., Bori, T., Snell, D.C., Miura, Y., Jung, S., Moroi, M., Frampton, J., and Watson, S.P. (2002) The Fc receptor gamma-chain is necessary and sufficient to initiate signalling through glycoprotein VI in transfected cells by the snake C- type lectin, convulxin. *Eur. J. Biochem.* **269**, 2951–2960
7. Suzuki-Inoue, K., Tulasne, D., Shen, Y., Bori-Sanz, T., Inoue, O., Jung, S.M., Moroi, M., Andrews, R.K., Berndt, M.C., and Watson, S.P. (2002) Association of Fyn and Lyn with the proline-rich domain of glycoprotein VI regulates intracellular signaling. *J. Biol. Chem.* **277**, 21561–21566
8. Suzuki-Inoue, K., Ozaki, Y., Kainoh, M., Shin, Y., Wu, Y., Yatomi, Y., Ohmori, T., Tanaka, T., Satoh, K., and Morita, T. (2001) Rhodocytin induces platelet aggregation by interacting with glycoprotein Ia/IIa (GPIa/IIa, Integrin alpha 2beta1). Involvement of GPIa/IIa-associated src and protein tyrosine phosphorylation. *J. Biol. Chem.* **276**, 1643–1652
9. Bergmeier, W., Bouvard, D., Eble, J.A., Mokhtari-Nejad, R., Schulte, V., Zirngibl, H., Brakebusch, C., Fassler, R., and Nieswandt, B. (2001) Rhodocytin (aggretin) activates platelets lacking alpha2beta1 integrin, glycoprotein VI, and the ligand-binding domain of glycoprotein Ibalpha. *J. Biol. Chem.* **276**, 25121–25126
10. Marcinkiewicz, C., Lobb, R.R., Marcinkiewicz, M.M., Daniel, J.L., Smith, J.B., Dangelmaier, C., Weinreb, P.H., Beacham, D.A., and Niewiarowski, S. (2000) Isolation and characterization of EMS16, a C-lectin type protein from *Echis multisquamatus* venom, a potent and selective inhibitor of the alpha2beta1 integrin. *Biochemistry* **39**, 9859–9867
11. Emsley, J., King, S.L., Bergelson, J.M., and Liddington, R.C. (1997) Crystal structure of the I domain from integrin alpha2beta1. *J. Biol. Chem.* **272**, 28512–28517
12. Chayen, N.E., Shaw-Stewart, P.D., Maeder, D.L., and Blow, D.M. (1990) An automated system for micro-batch protein crystallization and screening. *J. Appl. Crystallogr.* **23**, 297–302
13. Takeya, H., Nishida, S., Miyata, T., Kawada, S., Saisaka, Y., Morita, T., and Iwanaga, S. (1992) Coagulation factor X activating enzyme from Russell's viper venom (RVV-X). A novel metalloproteinase with disintegrin (platelet aggregation inhibitor)-like and C-type lectin-like domains. *J. Biol. Chem.* **267**, 14109–14117
14. Matthews, B.W. (1968) Solvent content of protein crystals. *J. Mol. Biol.* **33**, 491–497
15. Navaza, J. (1994) AMoRe: an automated package for molecular replacement. *Acta Cryst.* **D50**, 157–163
16. Brünger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., Read, R.J., Rice, L.M., Simonson, T., and Warren, G.L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Cryst.* **D54**, 905–921
17. MacRee, D.E. (1999) XtalView/Xfit – A Versatile Program for Manipulating Atomic Coordinates and Electron Density. *J. Struct. Biol.* **125**, 156–165