Crystallization and preliminary X-ray diffraction analysis of boar seminal plasma spermadhesin PSP-I/PSP-II, a heterodimer of two CUB domains

Antonio Romero[a], Paloma F. Varela[a], Libia Sanz[b], Edda Töpfer-Petersen[b], Juan J. Calvete[a,b,*]

[a]*Instituto de Química-Física 'Rocasolano' C.S.I.C., Madrid, Spain*
[b]*Institut für Reproduktionsmedizin, Tierärztliche Hochschule, Bünteweg 15, 30559 Hannover-Kirchrode, Germany*

Received 19 January 1996

Abstract Boar spermadhesin PSP-I/PSP-II (M_r 29000–30000), a non-covalent heterodimer of two CUB domains, was crystallized in two crystal forms. Complete diffraction data sets for hexagonal (space group $P6_{1,5}22$) and trigonal (space group $P3_{1,2}21$) crystals have been collected up to 2.9 and 2.5 Å resolution, respectively. Cell constants of the hexagonal and trigonal crystal forms are $a=b=87.2$ Å, $c=152.4$ Å, and $a=b=96.2$ Å, $c=70.8$ Å, respectively. The calculated packing parameters (V_m) are 2.8 and 3.2 Å^3/Da for the hexagonal and trigonal crystal forms, respectively, indicating that, in both cases, the asymmetric unit is constituted by one PSP-I/PSP-II heterodimer. This paper reports the first crystals of a protein built up by a CUB domain architecture.

Key words: Porcine fertilization; Boar spermadhesin PSP-I/PSP-II; CUB domain; Crystallization; X-ray diffraction analysis

1. Introduction

Porcine seminal plasma protein II (PSP-II) is a 14–16 kDa (116 amino acids) glycoprotein [1], which forms a non-covalent heterodimer with certain glycoforms of PSP-I (109 residues, 14–16 kDa) [2]. The PSP-I/PSP-II complex is a major component of boar seminal plasma [3]. Both subunits of the heterodimer belong to the spermadhesin protein family [4]. In the pig, this group of proteins include the polypeptides termed AQN-1, AQN-3, AWN, PSP-I, and PSP-II, which are major secretory products of the seminal vesicle epithelium [5] and coat the acrosomal cap of spermatozoa at ejaculation [6]. Members of the spermadhesin family display heparin-and/or carbohydrate-binding activity (reviewed in [4]). From their binding capabilities, spermadhesins are thought to play a pivotal role in aspects of porcine fertilization, i.e. sperm capacitation (heparin-binding activity) and sperm-egg interaction mediating the initial binding of spermatozoa to carbohydrate structures of glycoproteins of the extracellular matrix of the oocyte, the zona pellucida [4].

Proteins structurally related to boar spermadhesins have been characterized in bull (aSFP) [7] and horse (HSP-7) [8]. Porcine, bovine, and equine spermadhesins share 40–98% amino acid sequence identity, contain two conserved disulphide bridges between nearest-neighbor cysteine residues, but do not show discernible amino acid sequence similarity to any known protein structure [4].

A sequence pattern-search analysis has revealed that spermadhesins belong to a family of at least 16 proteins of diverse functionality, many of which are known to be involved in developmental processes such as embryogenesis or organogenesis (see Figs. 2 and 3 in [9]). The structure of these proteins is built by combination of several modules, all of them sharing the CUB domain, a 110 residue spanning extracellular module christened after the three proteins where it was first identified: *C*omplement subcomponents C1r/C1s, *U*egf, and *B*mp1 [9].

A consensus sequence property constructed for the CUB domain using a multiple sequence alignment of 31 copies revealed the presence of several rather conserved blocks interrupted by variable regions. Particularly, four cysteine residues are conserved, which in C1r/C1s and the spermadhesins have been shown to form the same arrangement of disulphide bridges between nearest neighbor cysteine residues (Cys^1-Cys^2; Cys^3-Cys^4) [4,10]. In addition to various conserved hydrophobic and aromatic residues, only a few other amino acid positions are nearly invariant [9]. The conserved hydrophobicity patterns cover the whole domain and is typical for an antiparallel β-sheet. Secondary structure predictions also indicated an all-β-structure. It has been proposed that the CUB module might adopt a nine-strand antiparallel β-barrel topology similar to that of an immunoglobulin variable domain [9]. The location of the two conserved disulphide bridges that would, respectively, connect the beginning of strands A and C and the N-terminal region of strand D with the end of strand E, support the proposed model [9]. Nevertheless, no structural data of proteins containing CUB domains have been reported.

Spermadhesins, 110–133 amino acid spanning polypeptides, form a subgroup of the CUB domain family as expected by their single domain architecture. To determine the three-dimensional conformation of the CUB domain, we have crystallized boar PSP-I/PSP-II spermadhesin complex. Here we report the X-ray diffraction analysis of two crystal forms of this heterodimer.

2. Materials and methods

Boar PSP-I/PSP-II heterodimer was isolated from the non-heparin-binding fraction of seminal plasma by a single size-exclusion chromatographic step [2]. The protein was dialyzed extensively against deionized (MilliQ) water and lyophilized until use.

Crystallization experiments were carried out at a constant room temperature of 22°C using the hanging-drop vapour diffusion method and 24-well tissue culture tray (Linbro). The initial PSP-I/PSP-II complex concentration was 15 mg/ml in 50 mM sodium phosphate, pH 7.0. Crystallization conditions were tested by mixing droplets of the protein solution (3–5 μl) with 2–5 μl of crystallization (reservoir) buffer (30% (w/v) polyethylene glycol (MW 2000), 0.1 M ammonium

*Corresponding author. Fax: (49) (511) 953 7172.
E-mail: JCalvete@Repro.TiHo-Hannover.De

Abbreviations: PSP, porcine seminal plasma protein; AWN, AQN, proteins of the spermadhesin family denominated after their first three N-terminal amino acids in the standard one-letter code; aSFP, bovine acidic seminal fluid protein; HSP, horse seminal plasma protein

SSDI S0014-5793(96)00133-0

acetate, pH 6.5). The mixture was allowed to equilibrate against 1 ml reservoir solution.

For N-terminal amino acid sequence of the crystallized protein, 3–5 crystals were washed five times with harvesting buffer (35% polyethylene glycol 2000 K, 100 mM ammonium acetate, pH 6.5), dissolved in distilled water, lyophilized, and analyzed with an ABI 473A automated sequencer following the manufacturer's instructions.

X-ray intensity data were measured on an imaging plate system (MAR Research, Hamburg) using CuK_α radiation from a Rigaku RU200 rotating anode generator operating at 5.4 kW. Images were processed and scaled using the programs MOSFLM, version 5.2 [11] and CCP4 package [12].

3. Results and discussion

Initial crystallization experiments on boar seminal plasma PSP-I/PSP-II complex were carried out using the screening protocol described by Jankarik et al. [13]. Small crystals grew in 2 weeks when polyethylene glycol (MW 4000) was used as precipitant agent. However, these crystals were not suitable for X-ray diffraction experiments. Attempts to improve the quality of the crystals were performed using different molecular weight polyethylene glycol, and varying the pH of the solutions, temperature and protein concentration. Hexagonal prism crystals with approximate dimensions $0.5\times0.3\times0.3$ mm^3 (Fig. 1A) were grown at a room temperature of 23°C in droplets containing a 1:1 (v/v) mixture of protein (15 mg/ml) to reservoir solution, which was 30% polyethylene glycol (MW 2000), 0.1 M ammonium acetate, pH 6.5. Under these conditions crystallization was perfectly reproducible with crystals of maximum size appearing after 1 month. These crystals belonged to the hexagonal space group $P6_{1,5}22$, had unit cell dimensions $a = b = 87.2$ Å, $c = 152.4$ Å, and diffracted to 2.9 Å resolution (Table 1).

Another crystalline form with a rombohedral habit, which diffracted beyond 2.0 Å resolution, was obtained using the same crystallization conditions as above, but starting with a mixture of 5 µl protein solution (15 mg/ml) and 2 µl of crystallization buffer. Crystals grew in approximately 4 months with a mean size of $0.8\times0.4\times0.4$ mm^3 (Fig. 1B). Their space group was determined to be trigonal $P3_{1,2}21$ and their unit cell dimensions were $a = b = 96.2$ Å, $c = 70.8$ Å (Table 1).

Data collection and statistics for both data sets are summarized in Table 1. The calculated packing parameter V_m [14] was 2.8 and 3.2 Å^3/Da for the hexagonal and trigonal crystal forms, respectively, which are in good agreement with that expected from their respective crystal geometry assuming one molecule of the heterodimer PSP-I/PSP-II (M_r 29000–30000) per asymmetric unit.

N-terminal amino acid sequence analysis of the crystallized protein showed two residues in a 1:1 equimolar ratio in each Edman degradation cycle: L/A, D/R, Y/I, H/N, G/A. This confirmed that the crystallized material was a heterodimer of PSP-I (N-terminal sequence LDYHA) and PSP-II (N-terminal sequence ARING) [1,2].

PSP-I/PSP-II is a heterodimer of two members of the spermadhesin protein family (reviewed in [15]) each of which is built up by a single CUB-domain architecture. Conformational analysis of spermadhesin PSP-I/PSP-II heterodimer by differential scanning calorimetry, FTIR spectroscopy and circular dichroism [16] indicated that the thermal unfolding of PSP-I/PSP-II was irreversible and followed a one-step transition with melting temperature (T_m) of 60.4°C. Deconvolution of circular dichroism spectra using a convex constraint analysis indicated that the relative percentage of secondary structural elements of PSP-I/PSP-II was 53% β-structure and 21% turns. No α-helical structure could be assigned. These data support the proposed all-β-structure single domain architecture for spermadhesins although does not rule out other all-β conformations, like the jelly roll topology which appears to be a common structural motif of a variety of carbohydrate-binding proteins including legume lectins ([17] and references therein), mammalian galectins [18,19], bacterial glucanases [20,21], human serum amyloid P component [22], and peptide : *N*-glycosidase F (PNGase F) from *Flavobacterium meningosepticum* [23].

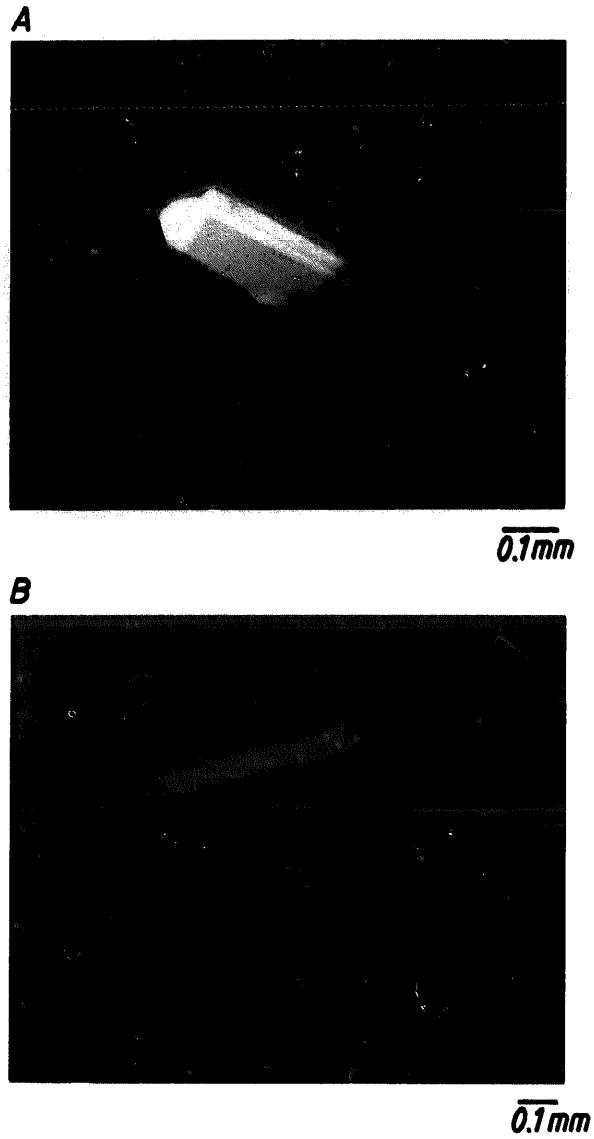

Fig. 1. (A) Hexagonal crystal of the PSP-I/PSP-II heterodimer complex grown in 30% polyethylene glycol 2000, 0.1 M ammonium acetate, pH 6.5, at an estimated final protein concentration of 15 mg/ml. Crystal size is $0.5\times0.2\times0.2$ mm^3. (B) Trigonal crystal obtained with the same crystallization buffer, but with a final protein concentration of about 37.5 mg/ml. The crystal size is $0.8\times0.4\times0.4$ mm^3.

Complete sets of data for two isomorphous heavy-atom derivatives of PSP-I/PSP-II hexagonal and trigonal crystals have been collected and their Patterson maps have been ana-

Table 1
Data collection and processing statistics

	Hexagonal crystals	Trigonal crystals
Cell constants	$a=b=87.2$ Å, $c=152.4$ Å	$a=b=96.2$ Å, $c=70.8$ Å
Space group	$P6_{1,5}22$	$P3_{1,2}21$
Crystal to detector distance	200 mm	130 mm
Maximum resolution	2.9 Å	2.5 Å
Rotation for each exposure	1°	1°
Time for each image	1500 s	900 s
Total rotation for a data set	40°	60°
Number of mesured reflections	34586	35213
Number of unique reflections	7744	10621
R_{merge}[a]	10.2%	5.7%
Completeness ($I>2\sigma$)	97.4%	90.4%
Last resolution shell	35.2%	48.3%

[a] $R_{merge}=\Sigma_j(|I_j-<I>|)/\Sigma<I>$, I_j is the observed scaled intensity of each unique reflection j and $<I>$ is the mean value of multiple observations.

lyzed (data not shown). Resolution of the three-dimensional structure of boar PSP-I/PSP-II will reveal the topology of the CUB domain and show whether the structural prediction holds true.

Spermadhesins are β-galactoside-binding lectins which act as sperm-associated binding molecules for zona pellucida glycoconjugates [24]. The carbohydrate-and zona pellucida glycoprotein-binding pocket of boar spermadhesins have been located around asparagine-50 [25,26]. This binding surface resides on the CC'C" β-strands of the current CUB-domain model. In addition, a composite heparin-binding domain has been mapped on the GFCC'C" face (in the immunoglobulin fold nomenclature; reviewed in ref [27]) of the proposed spermadhesin model, which is in an opposite location to the carbohydrate-binding region [28]. Structure determination of PSP-I/PSP-II is an important step towards elucidation of the structural requirements for ligand binding and, therefore, for understanding the molecular basis of porcine fertilization.

Acknowledgements: This work has been financed by grants Tö 114/3-1 from the Deutsche Forschungsgemeinschaft, Bonn, Germany, and PB92-0096 and PB93-0120 from the Dirección General de Investigación Científica y Técnica, Madrid, Spain. The authors wish to thank Dr. Guillermo Gímenez Gallego (Centro de Investigaciones Biológicas, CSIC, Madrid) for N-terminal sequence analyses, and Dr. Angel Vegas (Instituto de Química-Física, CSIC, Madrid) for helping with space group determination. X-ray diffraction data of hexagonal crystals with a 300 mm diameter Image Plate system were recorded in Prof. Robert Huber's department at the Max-Planck-Institut für Biochemie, Martinsried, Germany. Traveling was financed by Acciones Integradas Hispano-Alemana.

References

[1] Kwok, S.C.M., Yang, D., Dai, G., Soares, M.J., Chen, S. and McMurtry, J.P. (1993) DNA Cell Biol. 12, 605–610.
[2] Calvete, J.J., Mann, K., Schäfer, W., Raida, M., Sanz, L. and Töpfer-Petersen, E. (1995) FEBS Lett. 365, 179–182.
[3] Rutherfurd, K.J., Swiderek, K.M., Green, C.B., Chen, S., Shively, J.E. and Kwok, S.C.M. (1992) Arch. Biochem. Biophys. 295, 352–359.
[4] Calvete, J.J., Sanz, L., Dostálová, Z. and Töpfer-Petersen, E. (1995) Fertilität 11, 35–40.
[5] Dostàlovà, Z., Calvete, J.J., Sanz, L. and Töpfer-Petersen, E. (1994) Biochim. Biophys. Acta 1200, 48–54.
[6] Dostálová, Z., Calvete, J.J. and Töpfer-Petersen, E. (1995) Biol. Chem. Hoppe-Seyler 376, 237–242.
[7] Wempe, F., Einspanier, R. and Scheit, K.H. (1992) Biochem. Biophys. Res. Commun. 183, 232–237.
[8] Calvete, J.J., Nessau, S., Mann, K., Sanz, L., Sieme, H., Klug, E. und Töpfer-Petersen, E. (1994) Reprod. Dom. Anim. 29, 411–426.
[9] Bork, P. and Beckmann, G. (1993) J. Mol. Biol. 231, 539–545.
[10] Gagnon, J. and Arlaud, G.J. (1985) Biochem. J. 225, 135–142.
[11] Leslie, A.G.W. (1991) in: Crytallographic Computing V (Moras, D., Podjarny, A.D. and Thierry, J.P., eds.), Oxford University Press, Oxford, pp. 27–38.
[12] CCP4 (1994) Collaborative Computational Project, number 4. Acta Crystallogr. D50, 760–763.
[13] Jankarik, J. and Kim, S.M. (1991) J. Appl. Crystallogr. 24, 409–411.
[14] Matthews, B.W. (1968) J. Mol. Biol. 33, 491–497.
[15] Calvete, J.J., Sanz, L., Dostàlovà, Z. and Töpfer-Petersen, E. (1995) Fertilität 11, 35–40.
[16] Menéndez, M., Gasset, M., Laynez, J., López-Zúmel, C., Usobiaga, P., Töpfer-Petersen, E. and Calvete, J.J. (1995) Eur. J. Biochem. 234, 887–896.
[17] Loris, R., Van Overberge, D., Dao-Thi, M.-H., Poortmans, F., Maene, N. and Wyns, L. (1994) Proteins: Struct. Funct. Genet. 20, 330–346.
[18] Lobsanov, Y.D., Gitt, M.A., Leffler, H., Barondes, S.H. and Rini, J.M. (1993) J. Biol. Chem. 268, 27034–27038.
[19] Liao, D.-I., Kapadia, G., Ahmed, H., Vasta, G.R. and Herzberg, O. (1994) Proc. Natl. Acad. Sci. USA 91, 1428–1432.
[20] Keitel, T., Simon, O., Borriss, R. and Heinemann, U. (1993) Proc. Natl. Acad. Sci. USA 90, 5287–5291.
[21] Hahn, M., Pons, J., Planas, A., Querol, E. and Heinemann, U. (1995) FEBS Lett. 374, 221–224.
[22] Emsley, J., White, H.E., O'Hara, B.P., Oliva, G., Srinivasan, N., Tickle, I.J., Blundell, T.L., Pepys, M.B. and Wood, S.P. (1994) Nature 367, 338–345.
[23] Norris, G.E., Stillman, T.J., Anderson, B.F. and Baker, E.N. (1995) Structure 2, 1049–1059.
[24] Dostálová, Z., Calvete, J.J., Sanz, L. and Töpfer-Petersen, E. (1995) Eur. J. Biochem. 230, 329–336.
[25] Calvete, J.J., Solís, D., Sanz, L., Díaz-Mauriño, T., Schäfer, W., Mann, K. and Töpfer-Petersen, E. (1993) Eur. J. Biochem. 218, 719–725.
[26] Calvete, J.J., Sanz, L., Dostálová, Z. and Töpfer-Petersen, E. (1993) FEBS Lett. 334, 37–40.
[27] Bork, P., Holm, L. and Sander, C. (1994) J. Mol. Biol. 242, 309–320.
[28] Calvete, J.J., Dostálová, Z., Sanz, L., Adermann, K., Thole, H.H. and Töpfer-Petersen, E. (1996) FEBS Lett. 379, 207–211.