Boar spermadhesin PSP-II: location of posttranslational modifications, heterodimer formation with PSP-I glycoforms and effect of dimerization on the ligand-binding capabilities of the subunits

Juan José Calvete[a,b,*], Karlheinz Mann[c], Wolfram Schäfer[c], Manfred Raida[d], Libia Sanz[a], Edda Töpfer-Petersen[a]

[a]*Institut für Reproduktionsmedizin, Tierärztliche Hochschule, Hannover, Germany*
[b]*Instituto de Química-Física 'Rocasolano' CSIC, Madrid, Spain*
[c]*Max-Planck-Institut für Biochemie, Martinsried, Germany*
[d]*Niedersächsisches Institut für Peptid-Forschung GmbH, Hannover, Germany*

Received 7 April 1995

Abstract Spermadhesin PSP-II was isolated from the non-heparin-binding fraction of boar seminal plasma; its disulphide bridge pattern, and the location of a single N-glycosylation site were established. PSP-II forms a heterodimer with specific N-glycoforms of PSP-I. Although both subunits possess heparin-binding capability, the PSP-I/PSP-II complex does not. The heterodimer contains binding sites for zona pellucida glycoproteins and soybean trypsin inhibitor located in the PSP-II subunit. However, the PSP-I/PSP-II heterodimer binds only loosely to the sperm surface and is easily removed during in vitro capacitation, suggesting that the zona pellucida binding activity may not be relevant for gamete interaction. Our results show that dimerization of spermadhesins PSP-I and PSP-II markedly affects their binding capabilities.

Key words: Boar spermadhesin; Seminal plasma protein; Zona pellucida binding protein; Glycosylation; Soybean trypsin inhibitor binding protein

1. Introduction

The initial step of fertilization involves interaction of sperm surface lectins with oligosaccharides of the oocyte's zona pellucida [1]. Boar spermadhesins are a family of carbohydrate-binding proteins with molecular masses between 12–16 kDa. Spermadhesins AQN-1, PSP-I, PSP-II, AQN-3, and AWN (isoforms 1 and 2) are major secretory products of the seminal vesicle epithelium, accounting for over 80% of the total seminal fluid proteins [2,3]. AWN-1 is, in addition, synthesized by tubuli recti and rete testis epithelial cells [4]. Boar spermadhesins coat the acrosomal cap of spermatozoa [2], and are thought to play a role in porcine sperm capacitation and gamete recognition and binding [5].

The amino acid sequences of five boar spermadhesins have been determined: AQN-1 [6], PSP-I (AQN-2) [7,8], PSP-II [9], AQN-3 [10], and AWN (1 and 2) [11]. AWN-1 and AWN-2 differ only in that the latter contains an acetylated N-terminal residue. N-glycoforms of AQN-3 [7] and N- and O-glycosylated AWN species [12,13] have also been identified. Glycosylation not only contributes to the structural diversity of the protein family but also affects the ligand-binding capabilities of the glycosylated spermadhesins, i.e. it abolishes their zona pellucida-binding activity without impairing heparin binding [8,12,13].

AQN-1, AQN-2, AQN-3, and AWN are heparin-binding spermadhesins [14]. Here, we show that PSP-II is a major component of the non-heparin-binding fraction of boar seminal plasma. We have determined the location of disulphide bridges and its single N-glycosylation site. In addition, we show that PSP-II exist as a heterodimer with glycoforms of PSP-I, and report the effect of dimerization on the ligand-binding capabilities of the isolated subunits.

2. Materials and methods

Freshly ejaculated boar sperm was collected with an artificial vagina. Spermatozoa and seminal plasma were separated by centrifugation at $12,000 \times g$ for 15 min. Seminal plasma was fractionated by affinity chromatography on heparin-Sepharose as described [14]. The non-heparin-binding fraction was dialyzed against distilled water, lyophilized, and the proteins purified by reverse-phase HPLC on a Lichrocart (Merck) preparative (250 × 10 mm, 10 μm particle size) column eluted at 3 ml/min with a mixture of 0.1% trifluoroacetic acid in water (A) and acetonitrile (B), isocratically (25% B) for 5 min, followed by 25–45% B for 5 min, 45–60% B for 15 min, and 60–70% B for 5 min.

Determination of the apparent molecular masses of the proteins of the non-heparin-binding fraction, or of reverse-phase HPLC-purified protein, were estimated by gel-filtration chromatography on either a 2,000 × 5 cm Sephadex G-50 column or a Superose-12 (Pharmacia) column in 50 mM Tris-HCl, 150 mM NaCl, 1 mM EDTA, 0.025% sodium azide, pH 7.4. For calibration, a mixture of dextran blue 2000, bovine serum albumin (67 kDa), ovalbumin (43 kDa), ribonuclease A (14 kDa), and cyanocobalamine were used as molecular mass standards.

The ability of purified proteins to bind biotinylated ligands (soybean trypsin inhibitor, heparin, and zona pellucida glycoproteins) was determined by ELISA [8].

SDS-polyacrylamide gel electrophoresis was done according to Laemmli [15].

Amino acid analyses were done with a Pharmacia Alpha Plus analyzer after sample hydrolysis with 6 N HCl at 110°C for 24 h. For amino sugar and neutral sugar analyses, the samples were hydrolyzed with 4 N HCl for 4 h or 2 N HCl for 2 h, respectively, at 110°C. For sialic acid determination, the samples were hydrolyzed with 0.2 N trifluoroacetic acid for 1 h at 80°C. After drying the hydrolyzates in a SpeedVac, the monosaccharides were resolved on a CarboPac PA1 column (4 × 250 mm) eluting at 1 ml/min isocratically with either 16 mM NaOH (for amino and neutral sugars) or with 20 mM NaOH in 60 mM sodium acetate (for sialic acid), and analyzed using a Dionex DX-300 analyzer equipped with pulsed amperometric detector and the AI-450 chromatography software [14].

*Corresponding author: Institut für Reproduktionsmedizin, Tierärztliche Hochschule Hannover, Bünteweg 15, 30559 Hannover-Kirchrode, Germany. Fax: (49) (511) 953 8504.

SSDI 0014-5793(95)00452-1

Purified PSP-II (10 mg/ml in 70% (v/v) formic acid) was degraded with cyanogen bromide (100 mg/ml final concentration) overnight at room temperature in N_2-flushed tubes in the dark. Enzymatic digestions of purified proteins or fragments (2–5 mg/ml in 100 mM Tris-HCl, 150 mM NaCl, 1 M guanidine hydrochloride, pH 8.0) were performed with chymotrypsin, trypsin (Sigma) or endoproteinase Lys-C (Boehringer Mannheim) overnight at 37°C with an enzyme: substrate ratio of 1:100 (w/w). CNBr- and proteolytic fragments were isolated by reverse-phase HPLC on a Lichrospher RP-100 C18 column (250 × 4 mm, 5 μm particle size) (Merck) using the same chromatographic systems as above.

N-terminal sequence analyses were done with Applied Biosystems sequencers 473A and Procise following the manufacturer's instructions.

Deglycosylation of isolated proteins with PNGaseF was done as in [8].

The molecular masses of proteins were determined by matrix-assisted laser desorption ionization time-of-flight mass spectrometry using a fOCUS LD-TOF (GSG, Karlsruhe) mass spectrometer and α-cyano-4-hydroxycinamic acid as matrix. The molecular masses of peptides were measured by either electrospray ionization mass spectrometry using ac Sciex API-III LC/MS/MS triple quadropole instrument, or by fast atom bombardment liquid secondary-ion ionization using a MAT 900 mass spectrometer.

3. Results and discussion

3.1. Isolation and structural characterization of non-heparin-binding boar seminal plasma proteins

Boar seminal plasma was fractionated by affinity chromatography on a heparin-Sepharose column [14] and the proteins in the flow-through fraction were purified by reverse-phase HPLC (Fig. 1). SDS-polyacrylamide gel electrophoretic analysis of the chromatographic peaks showed that both of them contained heterogeneous material with apparent molecular masses of 16–18 kDa (Fig. 1, inset, lanes a and c). However, in each case, a single N-terminal amino acid sequence was obtained. Furthermore, both proteins are N-glycosylated (Table 1) and treatment of the isolated protein fractions with PNGaseF yielded a single electrophoretic band (Fig. 1, inset, lanes b and d), indicating that the proteins of each peak are glycoforms.

The N-terminal amino acid sequences of the isolated proteins were:

LDYHAXGGRLTDDYGTIFTYKGPKTE (peak 1) and
ARINGPDEXGRVIKDTSGSISNTDRQKNLXTWTILMK-PDQKVRMAIPY (peak 2).

In some preparations, peak 2 contained a mixture of two N-terminal sequences (ARINGPDE... and RINGPDE...), indicating that partial N-terminal degradation may occur.

The N-terminal sequences of the proteins in peaks 1 and 2 are identical with those of PSP-I and PSP-II, respectively [7–9], except for position 8 of PSP-II where glutamic acid was found

Table 1
Carbohydrate content of PSP-I and PSP-II isolated from the non-heparin-binding fraction of boar seminal plasma by reverse-phase HPLC as shown in Fig. 1

Monosaccharide	PSP-I	PSP-II
	(mol/mol protein)	
Fucose	0.6	0.3
Galactose	1.5	1.2
Manose	2.5	2.2
Glucosamine	3.7	3.9
Galactosamine	0.4	0.6
N-acetyl neuraminic acid	1.1	1.4

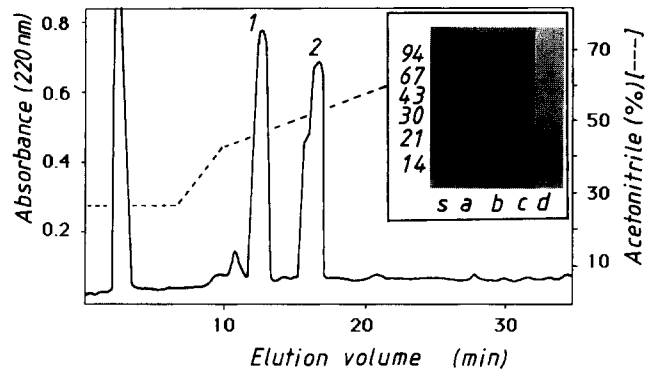

Fig. 1. Isolation by reverse-phase HPLC of the proteins of the non-heparin-binding fraction of boar seminal plasma. Inset: SDS-(15%) polyacrylamide gel electrophoresis of the intact proteins recovered in peaks 1 (lane a) and 2 (lane c). Lanes b and d show the same proteins as in lanes a and c, respectively, after deglycosylation with PNGaseF. Lane s, molecular mass standards, from top to bottom: phosphorylase b, bovine serum albumin, ovalbumin, carbonic anhydrase, soybean trypsin inhibitor, and lysozyme, whose molecular masses are indicated at the left.

in the protein ([7] and this work) but glutamine has been reported in a PSP-II cDNA [9].

The identity of the amino acid sequence of the protein of peak 1 with that of PSP-I was further verified by reverse-phase chymotryptic peptide mapping and characterization of the proteolytic fragments by amino acid and mass spectrometric analyses (not shown). In addition, the location of the two disulphide bridges (between cysteine residues 6–27 and 50–71) and the glycosylation site at asparagine 47 [8] were conserved. Furthermore, the molecular mass of PSP-I deglycosylated with PNGaseF, determined by mass spectrometry, was 11,982 ± 4 Da, which is close to the average molecular mass of PSP-I calculated from its amino acid sequence (11,979.7 Da) [7,8].

The amino acid sequence of PSP-II was established by combination of N-terminal sequence analysis and characterization of overlapping chymotryptic, tryptic, endoproteinase Lys-C, and cyanogen bromide-derived peptides. Except at position 8 (see above) our sequence is identical to a sequence obtained by cDNA technology (9) (Fig. 2). Analysis of proteolytic fragments of PSP-II showed that peak T-11 ($M + H^+$ = 1,579.8 Da) contained two tryptic peptides: ^{49}LNLAXGK^{55} and ^{73}LXAGAAIVF^{81} connected by a single disulphide bridge between cysteines 53 and 74. A second disulphide bond was located in a cyanogen bromide fragment corresponding to residues 1–36. This fragment ($M + H^+$ = 3,988.6 ± 3.8 Da) did not contain free cysteines but two ethylpyridylcysteine residues/mol peptide were quantitated by amino acid analysis following reduction and treatment with 4-vinylpyridine. Furthermore, the molecular mass of the derivatized peptide was 4,201 ± 3.1 Da. These data clearly showed that cysteines 9 and 30 form a disulphide bridge. The two disulphide bridges of PSP-II are conserved in all other members of the boar spermadhesin family (Fig. 2). In addition, the same cystine pattern has been reported in aSFP, a bovine seminal plasma polypeptide which is structurally related to boar spermadhesins [17]. This feature may be important for maintaining the correct three-dimensional structure of spermadhesins. Other conserved characteristics in the

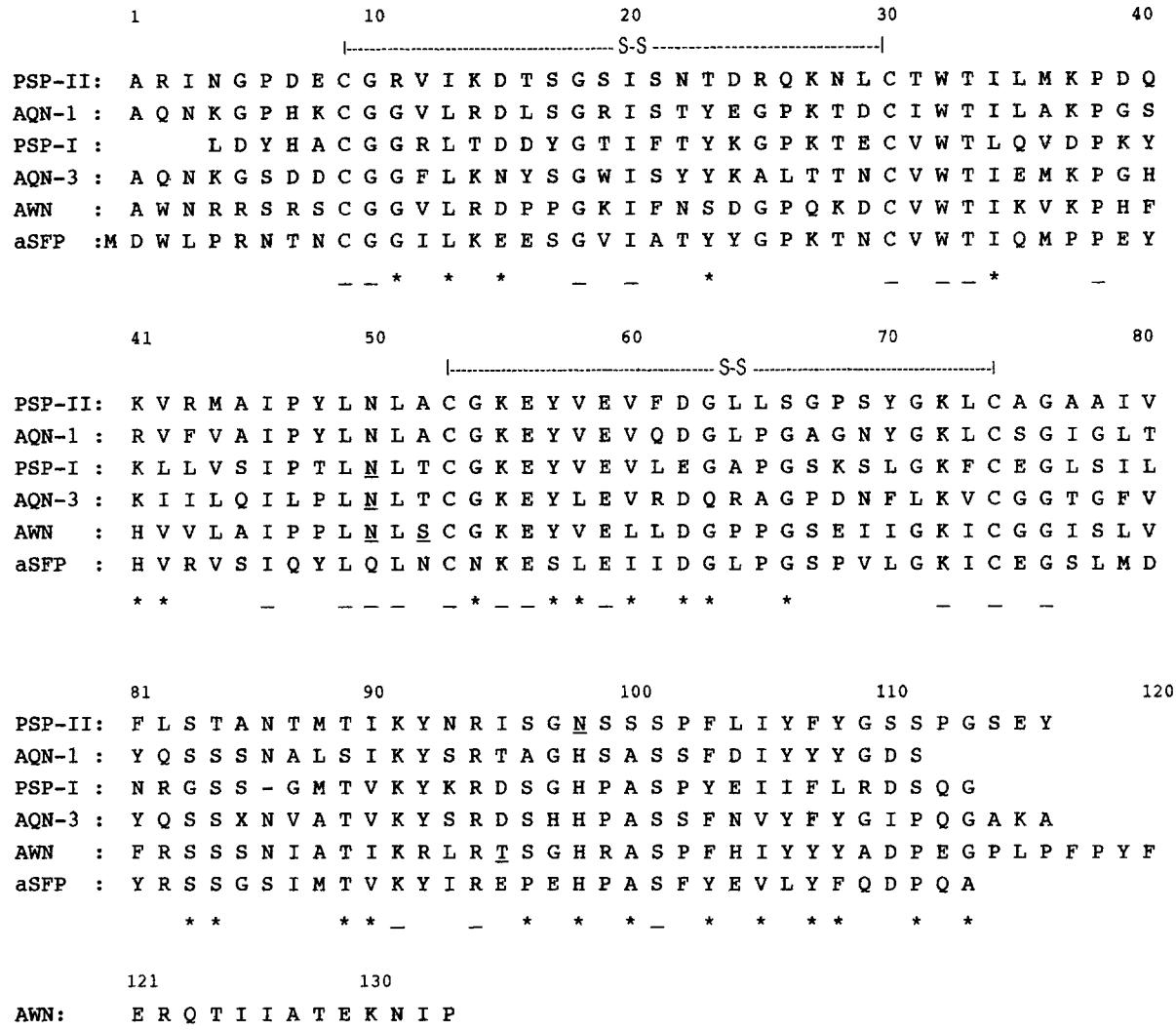

Fig. 2. Alignment of the amino acid sequences of spermadhesins. Positions with absolutely conserved residues (–) and conservative substitutions (⋆) in all polypeptides are shown $\underline{N}^{50}$, asparagine residues which are constitutively glycosylated in PSP-I (8) and in glycoforms of AQN-3 (8) and AWN (12,13); $\underline{N}^{98}$, glycosylated asparagine in PSP-II (this work); $\underline{S}^{52}$ and $\underline{T}^{95}$, serine and threonine residues which are O-glycosylated in glycoforms of AWN (13).

porcine and bovine proteins, which may define a consensus sequence pattern, are summarized in Fig. 2.

Amino acid and carbohydrate analyses of proteolytic fragments revealed that PSP-II possessed a single N-glycosylation site at asparagine 98. Mass spectrometric analysis of the PSP-II glycoform yielded a broad peak between 14,300 and 15,000 Da, whereas the average molecular mass measured for PNGaseF-deglycosylated PSP-II was 12,410 ± 8 Da. This mass is slightly lower than that calculated for the complete PSP-II sequence (116 residues, 12,643.5 Da) and may correspond to a truncated form of PSP-II (residues 2–115, calculated molecular mass: 12,409.4). C-terminal degradation of glycosylated forms of AWN-1 following treatment with PNGaseF has been reported previously [12].

3.2. *PSP-II forms a heterodimer with specific glycoforms of PSP-I*

To determine the aggregation state of the non-heparin-binding proteins of boar seminal plasma, the heparin-Sepharose flow-through fraction was chromatographed on either an FPLC Superose-12 or a Sephadex G-50 column. In both cases, a single symmetrical peak eluting with an approximate apparent molecular mass of 32 kDa was obtained, (Fig. 3). Reverse-phase HPLC analysis of fractions along this peak (labeled 'o' in Fig. 3) and quantitation of isolated PSP-I and PSP-II by amino acid analysis revealed that the glycoproteins were present in equimolar amounts. On the other hand, isolated PSP-I and PSP-II eluted from the size-exclusion columns as monomeric (16–18 kDa) proteins (Fig. 3, · · ·). Similar results have been

Table 2
Ligand binding capabilities of the PSP-I/PSP-II complex and its isolated subunits, as determined using a solid-phase assay and biotinylated ligands

Ligand	PSP-I/PSP-II	PSP-I	PSP-II
Heparin	–	+	+
Soybean trypsin inhibitor	+	–	+
Zona pellucida glycoproteins	+	–	+

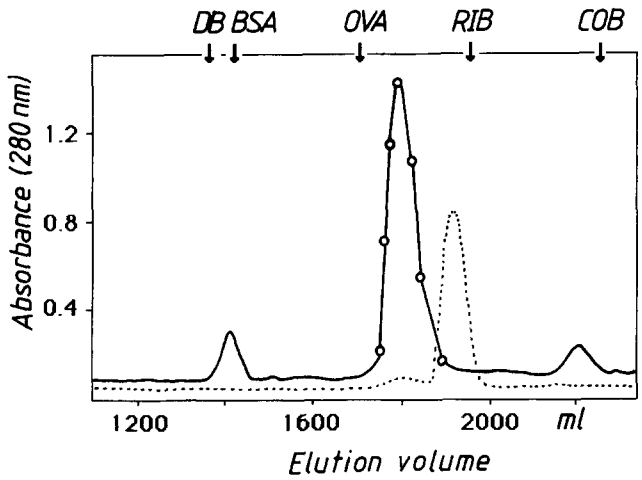

Fig. 3. Fractionation of the non-heparin-binding fraction of boar seminal plasma by gel filtration on Sephadex G-50. The elution position of molecular mass standards is indicated at the top: DB, dextran blue 2000 (< 2,000 kDa); BSA, bovine serum albumin (67 kDa); OVA, ovalbumin (43 kDa); RIB, ribonuclease A (14 kDa); COB, cyanocobalamine (1,350 Da). The symbol 'o' along the PSP-I/PSP-II peak (· · ·) indicates the elution position of samples which were analyzed by reverse-phase HPLC (as in Fig. 1) followed by quantitation of isolated PSP-I and PSP-II by amino acid analysis. The discontinous line shows the eluting position of isolated PSP-I and PSP-II.

reported using ion-exchange and gel-filtration chromatography [7]. This strongly suggests that both components of the non-heparin-binding fraction of boar seminal plasma exist as heterodimers.

PSP-I has also been isolated from the heparin-binding fraction of boar seminal plasma [8]. Although heparin-binding and non-heparin-binding PSP-I molecules have identical amino acid sequences and location of posttranslational modifications (see above), they are differently glycosylated. Thus, although both heparin-binding and non-heparin-binding PSP-I glycoforms, contain a single N-linked carbohydrate, notably the latter (Table 1) but not the former [8] contains galactosamine, galactose and sialic acid. Thus, it appears that the type of glycosylation may determine which PSP-I glycoforms are recruited by PSP-II to form heterodimers.

3.3. *Effect of dimerization on the ligand binding capabilities of isolated PSP-I and PSP-II*

The ability of the PSP-I/PSP-II heterodimer and its isolated subunits to bind known ligands of other boar spermadhesins, i.e. soybean trypsin inhibitor (STI), heparin, and zona pellucida glycoproteins (ZPG) [5], was assessed using a solid-phase (ELISA) assay and biotinylated ligands. It was found (Table 2) that, although the heterodimer did not bind heparin, both subunits possessed this binding activity. In addition, the PSP-I/PSP-II complex and isolated PSP-II, but not PSP-I, showed affinity for soybean trypsin inhibitor and zona pellucida glycoproteins. These results indicate that the binding site(s) for STI and ZPG on PSP-I/PSP-II are located on PSP-II, and that dimerization impairs the heparin-binding capability of both subunits. Whether dimerization sterically blocks the heparin binding site(s), or affects the conformation of the heparin-binding regions remains unclear. On the other hand, glycosylation at asparagine 47 is the cause of the inability of PSP-I to bind STI and ZPG [5,8].

The zona pellucida glycoprotein-binding activity of PSP-I/PSP-II would suggest a role during gamete interaction. However, PSP-I/PSP-II have been found to bind only loosely to the sperm surface, and is absent from in vitro capacitated spermatozoa [2]. Interestingly, heparin-binding PSP-I molecules are present on capacitated sperm, suggesting that either the type of glycosylation or complex formation with PSP-II also influences the sperm coating characteristics of PSP-I (and/or PSP-II) glycoforms. Thus, although the biological function(s) of boar seminal plasma PSP-I/PSP-II remain obscure, our results show that heterodimerization of specific glycoforms of PSP-I with PSP-II exerts a profound effect on the ligand-binding characteristics and compartmentalization of these boar spermadhesins.

Acknowledgements: This work has been financed by grants Tö 114/3-1 (E.T-P.) from the Deutsche Forschungsgemeinschaft (Bonn, Germany) and PB92-0096 (J.J.C. and L.S.) from the Dirección General de Investigación Científica y Técnica (Madrid, Spain).

References

[1] Yanagimachi, R. (1994) In: The Physiology of Reproduction, 2nd Edn. (Knobil, E. and Neill, J.D. Eds.) Raven Press, New York, pp. 189–317.
[2] Dostálová, Z., Calvete, J.J., Sanz, L. and Töpfer-Petersen, E. (1994) Biochim. Biophys. Acta 1200, 48–54.
[3] Kwok, S.C.M., Soares, M.J., McMurtry, J.P. and Yurewicz, E.C. (1993) Mol. Reprod. Dev. 35, 244–250.
[4] Sinowatz, F., Amselgruber W., Töpfer-Petersen, E., Calvete, J.J., Sanz, L. and Plendl, J. (1995) Cell Tissue Res. (in press).
[5] Calvete, J.J., Sanz, L. and Töpfer-Petersen, E. (1994) Assisted Reprod. Technol./Androl. 6, 316–330.
[6] Sanz, L., Calvete, J.J., Mann, K., Schäfer, W., Schmid, E.R. and Töpfer-Petersen, E. (1992) Eur. J. Biochem. 205, 645–652.
[7] 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.
[8] 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.
[9] Kwok, S.C.M., Yang, D., Dai, G., Soares, M.J., Chen, S. and McMurtry, J.P. (1993) DNA Cell Biol. 12, 605–610.
[10] Sanz, L., Calvete, J.J., Mann, K., Schäfer, W., Schmid, E.R. and Töpfer-Petersen, E. (1991) FEBS Lett. 291, 33–36.
[11] Sanz, L., Calvete, J.J., Mann, K., Schäfer, W., Schmid, E.R., Amselgruber, W., Sinowatz, F., Ehrhard, M. and Töpfer-Petersen, E. (1992) FEBS Lett. 300, 213–218.
[12] Calvete, J.J., Sanz, L., Dostálová, Z. and Töpfer-Petersen, E. (1993) FEBS Lett. 334, 37–40.
[13] Calvete, J.J., Solís, D., Sanz, L., Díaz-Mauriño, T. and Töpfer-Petersen, E. (1994) Biol. Chem. Hoppe-Seyler 375, 667–673.
[14] Sanz, L., Calvete, J.J., Mann, K., Gabius, H-J. and Töpfer-Petersen, E. (1993) Mol. Reprod. Dev. 35, 37–43.
[15] Laemmli, U.K. (1970) Nature 227, 680–685.
[16] Anumula, K.R. and Taylor, P.B. (1991) Eur. J. Biochem. 195, 269–280.
[17] Einspanier, R., Krause, I., Calvete, J.J., Töpfer-Petersen, E., Klostermeyer, H. and Karg, H. (1994) FEBS Lett. 344, 61–64.