

A novel *Drosophila* serpin that inhibits serine proteases

Jin-Hua Han, Hong-Yan Zhang¹, Gi-Sik Min², Doris Kemler³, Carl Hashimoto*

Department of Cell Biology, Yale University School of Medicine, 333 Cedar Street, New Haven, CT 06520, USA

Received 25 January 2000

Edited by Horst Feldmann

Abstract Serpins define a large protein family in which most members function as serine protease inhibitors. Here we report the results of a search for serpins in *Drosophila melanogaster* that are potentially required for oogenesis or embryogenesis. We cloned and sequenced ovarian cDNAs that encode six distinct proteins having extensive sequence similarity to mammalian serpins, including residues important in the serpin inhibition mechanism. One of these new serpins in recombinant form inactivates, and complexes with, trypsin-like proteases *in vitro*. To our knowledge, these results represent the first evidence for a serpin in *Drosophila* that functions as a serine protease inhibitor.

© 2000 Federation of European Biochemical Societies.

Key words: Serpin; Serine protease inhibitor; *Drosophila melanogaster*

1. Introduction

Serpins constitute a large family of proteins found in animals, plants, and viruses (for review see [1]). Most serpins function as suicide substrate inhibitors of serine proteases. Such inhibitory serpins, typically about 400 amino acids in length, are cleaved by the target protease within a reactive center loop region of about 20 amino acids near the C-terminus. The amino acid N-terminal to the scissile bond (the P1 position) appears to be important for determining the specificity of serpins for particular proteases. Further N-terminal within the reactive center loop region are the residues of the hinge region which are highly conserved among inhibitory serpins. The hinge region appears to be mechanistically critical for the reactive center loop to undergo a dramatic conformational change upon serpin cleavage and for the formation of an essentially irreversible 1:1 complex of serpin and protease.

Serpins are involved in diverse biological processes. Many serpins found in human plasma regulate proteolytic reactions important in blood coagulation, fibrinolysis, the immune response, and inflammation. A serpin called maspin identified in mammalian breast tissue has been implicated in tumor suppression [2], whereas another serpin, PEDF, has been identified as a neurotrophic factor and as a regulator of angiogen-

esis [3,4]. PEDF is one example among several of a serpin that does not function as a serine protease inhibitor. Such non-inhibitory serpins may lack certain key residues, as in the reactive center loop region, necessary for inhibitory activity [1].

The genome of the fruit fly, *Drosophila melanogaster*, likely contains many serpin genes (Berkeley *Drosophila* Genome Project, unpublished). The first serpin from this organism to be described was the Acp76A protein, a component of the male accessory gland that is transferred to the female during mating [5]. More recent studies have demonstrated a role for another serpin in anti-fungal defense in *Drosophila* [6]. It has not yet been biochemically demonstrated that either of these serpins is active as a serine protease inhibitor.

We are interested in defining a developmental role for serpins using genetic analysis in *Drosophila*. A developmental process in *Drosophila* that may involve serpins is dorsoventral patterning of the embryo. Genetic studies have identified four serine proteases required to generate a ventral signal that initiates dorsoventral patterning, perhaps acting sequentially in a cascade like the mammalian blood clotting proteases [7]. While there is no genetic evidence that a serpin is involved in dorsoventral patterning, there is indirect biochemical evidence that one of the serine proteases in the dorsoventral cascade complexes with a serpin in the embryo [8].

Here we report the deduced sequences of six new *Drosophila* serpins identified in a search for serpins that potentially function in oogenesis or embryogenesis, including dorsoventral patterning. Sequence analysis suggests that the six serpins are likely to function as serine protease inhibitors. As a first step toward defining a developmental role for these serpins as serine protease inhibitors, we provide biochemical evidence that one of these serpins is a potent inhibitor of trypsin-like proteases *in vitro*.

2. Materials and methods

2.1. RT-PCR

We used standard methods for isolation and analysis of DNA and RNA, including RT-PCR [9]. Ovaries of Oregon R strain flies were manually dissected [10], and total RNA was extracted from ovaries with the Trizol reagent (Gibco-BRL, Gaithersburg, MD). Degenerate oligonucleotides were designed that corresponded to two amino acid sequences, NAVYFKG and DVNEEG, conserved between horseshoe crab and human serpins [11]. PCR reaction (40 cycles of 94°C 1 min, 50°C 1 min, and 72°C 1 min, followed by 72°C 7 min) using these oligonucleotides as primers yielded a 520 bp DNA fragment, which was cloned into the TA vector (Invitrogen, Carlsbad, CA). DNA sequencing of multiple clones of this fragment revealed that we had isolated partial cDNAs potentially encoding four different proteins, which by BLAST search of the GenBank database were found to be homologous in sequence to previously characterized serpins. 5' and 3' RACE reactions were performed to isolate full-length cDNAs, which were sequenced in both strands to derive the sequences of four different proteins (designated sp1–4). We subsequently found that sequen-

*Corresponding author. Fax: (1)-203-785 7226.
E-mail: carl.hashimoto@yale.edu

¹ Present address: Center for Genomics Research, Karolinska Institute, Stockholm, Sweden.

² Present address: Department of Biology, McGill University, Montreal, Que., Canada.

³ Present address: Department of MCD Biology, Yale University, New Haven, CT, USA.

Fig. 1. Primary structures of sp1-6. The complete amino acid sequences of sp1-6 and of the prototypal serpin human α_1 -proteinase inhibitor (α_1 -PI) [21] were aligned using MEGALIGN software (DNASTAR Inc., Madison, WI). Residues matching the consensus are boxed. Residues of the N-terminal signal sequences are in lowercase letters [22]. The underlined region is analyzed in Fig. 2. The accession numbers for sp1-6 are AJ251744-9, respectively.

reaction containing protease and sp6 as described above were withdrawn, mixed with SDS sample buffer, and immediately boiled for 5 min. Samples were analyzed by electrophoresis on SDS-PAGE gels, which were stained with Coomassie blue.

2.5. Determination of second-order rate constant

The second-order rate constant for interaction between protease and sp6 was determined by the discontinuous method under pseudo first-order conditions [14]. Protease inhibition was assayed as described above using 100 nM of purified sp6. The pseudo first-order rate constant (k_{obs}) was obtained from the slope of a plot of $\ln[E]$ versus time (t) according to the equation $\ln[E]_t = \ln[E]_0 - k_{\text{obs}}t$, in which $[E]_t$ and $[E]_0$ are protease activity at time t and time zero, respectively. The second-order rate constant (k) was determined with the equation $k = k_{\text{obs}}/[I]_0$, where $[I]_0$ is the initial sp6 concentration.

3. Results and discussion

3.1. Analysis of cDNAs encoding *Drosophila* serpins

At the time we began our studies, the male accessory gland protein Acp76a was the only serpin in *D. melanogaster* that had been identified. To identify additional serpins in this organism, we used a RT-PCR strategy (see Section 2). We used ovarian RNA as the template in the RT-PCR reaction to bias our search for serpins that could be candidate regulators of the serine proteases involved in dorsoventral patterning and that therefore would be maternally expressed like these proteases [7]. Using this strategy, we cloned cDNAs encoding four distinct polypeptide sequences similar to sequences of serpins from other organisms, as revealed by BLAST search of the GenBank database. In addition, we identified two *Drosophila* EST clones encoding two additional serpin-like sequences (see Section 2). The proteins potentially encoded by the different cDNAs were designated sp1–6. By in situ hybridization analysis, we mapped DNA encoding sp1, sp4, and sp5 to chromosomal locations 42E1–2, 42D1–5, and 88E3–5, respectively (data not shown). Although we did not map the chromosomal location of DNA encoding sp2, sp3, or sp6, ESTs matching the sequences for sp2 and sp6 exist within P1 clones that roughly map to 28E–29A and 55A–B, respectively (Berkeley *Drosophila* Genome Project, unpublished).

Full-length cDNAs were sequenced to derive the complete

	Hinge region	P1/P1'
sp1	EEGAEAA GATSVAVTNR/AGFST	
sp2	EEGAEAAA TALLFVRL/SVPMP	
sp3	EEGSEAAA TAVVFRYK/SIRSP	
sp4	EEGTEAAA TGMAVRRK/RAIMS	
sp5	EEGSTAAA T-VLFTYR/SARPV	
sp6	EEGTEAAA TGMIMMTR/MMTFP	
α_1 -PI	EKGTEAAG AMFLEAIPM/SIPPE	
AT	EEGSEAAA STAVVIAGR/SLNPN	
HNEI	EEGTEAAA ATAGIATFC/MLMPE	
SCCA-2	EEGTEAAA ATGVEVSLT/SAQIA	

Fig. 2. Predicted reactive center loop regions of sp1–6. C-terminal regions of sp1–6 (underlined in Fig. 1) were compared with the reactive center loop regions of human α_1 -PI, human antithrombin (AT) [23], human neutrophil elastase inhibitor (HNEI) [24], and mouse squamous cell carcinoma antigen-2 (SCCA-2) [25]. Boldface lettering highlights the hinge region important for the serpin inhibition mechanism as well as the residues (P1 and P1') flanking the scissile bond [26]. A fixed distance between hinge region and P1 residue and the common occurrence of serine as the P1' residue in inhibitory serpins were used as criteria to predict the reactive site of sp1–6 [1].

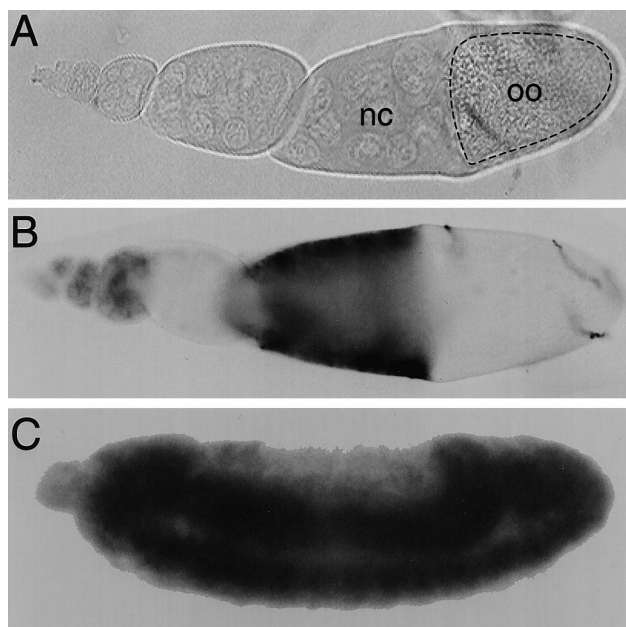


Fig. 3. Expression of sp2 RNA during oogenesis and embryogenesis. RNA in situ hybridization analysis of whole mount ovaries and embryos using RNA probes corresponding to the sense and antisense strands of sp2. A: During oogenesis, the oocyte (oo) develops within the egg chamber, which also contains 15 nurse cells (nc) and an epithelium of follicle cells surrounding the oocyte and nurse cells [17]. The dashed line roughly marks the oocyte border with nurse and follicle cells. The labeled egg chamber is at a stage just prior to the time when materials necessary for early embryogenesis are transferred from the nurse cells to the oocyte. Smaller egg chambers at earlier stages are visible to the left. These egg chambers were incubated with the sp2 sense probe to serve as a control for non-specific hybridization. B: Egg chambers at a similar stage as the labeled egg chamber in A show an intense hybridization signal in nurse cells when incubated with the sp2 antisense probe. Hybridization signal is also visible in younger egg chambers to the left. C: Example of a late-stage embryo incubated with the sp2 antisense probe showing intense hybridization signal in multiple tissues, including the nerve cord which runs along the curved ventral side of embryo toward the bottom of the figure.

primary structure of sp1–6, which range in length from 372 to 427 amino acids (Fig. 1). The level of sequence identity between any two proteins is 21–70%, with sp2 and sp3 most similar to each other and sp5 the most different from the others. Four of the proteins, excluding sp4 and sp6, are predicted to have a signal sequence at the N-terminus (Fig. 1).

BLAST search of the GenBank database revealed that all six proteins are new members of the serpin family, having sequence identity from 20 to 40% with serpins from mammals and other organisms (a comparison with α_1 -proteinase inhibitor is shown in Fig. 1). The closest homologue for sp4 appears to be human neuroserpin [15], whereas sp1, sp2, sp3, sp5 and sp6 appear to be most similar in sequence to members of the ov-serpin sub-family, which includes inhibitory serpins such as human plasminogen activator inhibitor-2 and non-inhibitory serpins such as chicken ovalbumin [16]. Despite this similarity, the five *Drosophila* serpins may belong in a distinct group, as they do not share all characteristics of ov-serpins. For example, in contrast to the case with ov-serpins, sp1, 2, 3, and 5 have a recognizable signal sequence at the N-terminus, as mentioned above (Fig. 1). In addition, sp6 is encoded in two exons, a different pattern than the seven-

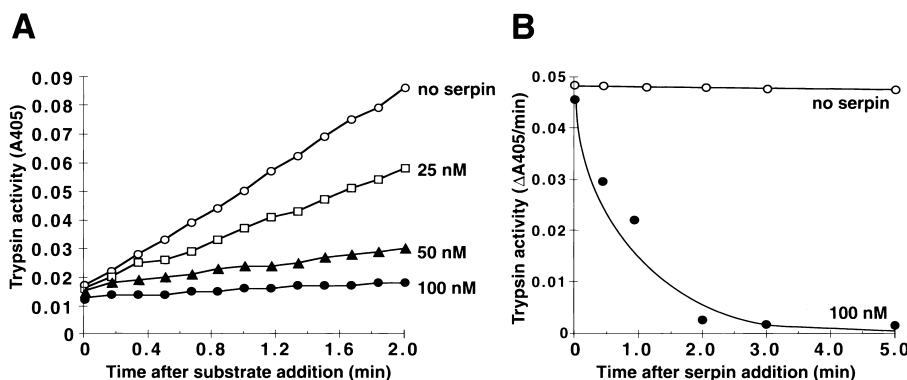


Fig. 4. Recombinant sp6 inhibits trypsin. A: Inhibition of trypsin by increasing amounts of sp6. The change in absorbance at 405 nm due to hydrolysis of a chromogenic substrate by trypsin was measured over a 2 min period after prior incubation of trypsin with sp6, as described in Section 2. Essentially complete inhibition is observed at 100 nM sp6. B: Time course of trypsin inhibition by sp6. Trypsin activity was measured as above after prior incubation of trypsin without (open circle) or with (closed circle) 100 nM sp6 for the time periods indicated. Fitted curves were generated using UltraFit software (Biosoft, Ferguson, MO).

exon structure of ov-serpin genes [1] (J.-H. Han, unpublished observation).

The C-terminal regions of sp1–6 can be aligned with the reactive center loop regions of serpins known to be active as serine protease inhibitors (Fig. 2). This alignment shows that sp1–6 all contain highly conserved residues of the hinge region important for the serpin inhibition mechanism, suggesting that they are likely to be active as serine protease inhibitors [1]. This alignment also predicts that the reactive site residue at the P1 position is a leucine for sp2 and either an arginine or a lysine for sp1 and sp3–6. Thus, as a serine protease inhibitor, sp2 should be active against a chymotrypsin-like protease, whereas the others would be active against trypsin-like enzymes.

3.2. Expression patterns of the *Drosophila* serpins

Although we cloned cDNAs for sp1–4 from ovarian RNA, we wanted to confirm that the RNAs encoding these serpins are expressed in the ovary. In addition, we wanted to know what cell type in the ovary expressed these serpins. By RNA in situ hybridization analysis, we found that RNAs encoding these four serpins are expressed predominantly in the nurse cells, which are principally responsible for providing the oocyte with materials necessary for early embryogenesis [17]. While expression of all four serpins was detected throughout the early stages of oogenesis, it appeared to peak during stages 9–10, just prior to the time when there is a major transfer of materials from the nurse cells to the oocyte (illustrated for sp2 in Fig. 3A,B). We detected comparably little expression of these serpins in the follicle cells that secrete eggshell

components around the oocyte. The expression of sp1–4 in ovarian nurse cells is compatible with a role for these serpins in oogenesis or embryogenesis.

Two types of observations suggest that sp1–4 function at other developmental stages and in other tissues. First, by RNA in situ hybridization, we detected sp1–4 transcripts throughout embryogenesis in a variety of different tissues (illustrated for sp2 in Fig. 3C). Second, the sequences of sp1–4 are represented in EST clones isolated from an adult head cDNA library (see Section 2).

The genes encoding sp5 and sp6 also appear to have complex expression patterns that include the ovary. While an EST clone for sp5 was isolated from an embryonic cDNA library (see Section 2), we also detected sp5 transcripts in the ovary by RT-PCR analysis (data not shown). In addition, the sequence of sp6 is found in EST clones from both larval/pupal and ovarian cDNA libraries (see Section 2).

3.3. Inhibition of serine proteases by recombinant serpin sp6

To test whether sp1–6 function as serine protease inhibitors in vitro, we wanted to make recombinant forms of these proteins. Of the serpins that were expressed with a His-tag in bacteria (see Section 2), only sp6 could be obtained in a preparation of sufficient purity and quantity necessary for biochemical studies (see Fig. 5, lane 9). To assay for inhibition, we incubated increasing amounts of recombinant sp6 with various serine proteases and measured the hydrolysis of synthetic chromogenic substrates by the proteases (see Section 2). Using this assay, we found that recombinant sp6 efficiently inhibited the activities of elastase, plasmin, thrombin, and trypsin (illustrated for trypsin in Fig. 4A,B).

To obtain a quantitative measure of inhibitory activity, we calculated the second-order rate constants for interaction between sp6 and proteases (see Table 1). By this criterion, recombinant sp6 inhibited trypsin about 30 times more than α_1 -proteinase inhibitor, and inhibited thrombin, a trypsin-like enzyme, as efficiently as antithrombin (in the absence of heparin). These results demonstrate that sp6 is a potent serine protease inhibitor. However, sp6 appeared to be weakest as an inhibitor against elastase. In addition, sp6 was inactive towards chymotrypsin and appeared to be cleaved by chymotrypsin (data not shown). This specificity suggests that the natural target of sp6 is a trypsin-like serine protease, consis-

Table 1
Second-order rate constants for interaction between protease and serpin ($M^{-1} \text{ min}^{-1}$)

	sp6 ^a	α_1 -PI ^b	Antithrombin ^{c,d}	Antithrombin+ heparin ^{c,d}
Trypsin	2.6×10^7	7.8×10^5	9.0×10^{6c}	6.6×10^{7c}
Thrombin	1.4×10^6	—	5.3×10^{5d}	2.2×10^{9d}
Elastase	5.3×10^4	3.9×10^9	—	—

^aRate constants were determined as described in Section 2.

^bData are from [18].

^cData are from [19].

^dData are from [20].

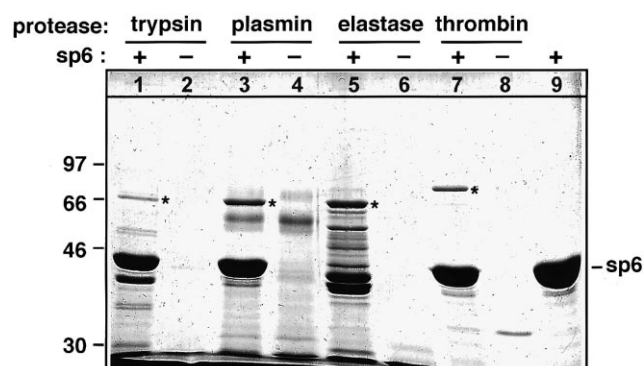


Fig. 5. Recombinant sp6 forms stable complexes with serine proteases. Complex formation between recombinant sp6 and various proteases was analyzed by SDS-PAGE as described in Section 2. Co-incubation of purified sp6 with trypsin, plasmin, elastase, or thrombin produced a band in the 70–80 kDa range (marked with an asterisk; lanes 1, 3, 5, and 7) matching the predicted size of a 1:1 complex of serpin and protease (molecular weight markers in kDa are shown on the left). This band was not seen in the sample of each protease (lanes 2, 4, 6, and 8) or sp6 (lane 9) alone. The presence of multiple bands between the putative sp6–elastase complex and free sp6 may be due to degradation of the complex, perhaps by a trace amount of free elastase. Trypsin, plasmin, and elastase migrated with the dye front and therefore are not visible, whereas thrombin is visible as a band of 37 kDa in lane 8. The bands visible in lane 4 likely represent contaminants in the plasmin sample.

tent with an arginine as its predicted reactive site P1 residue (see Fig. 2).

3.4. SDS-stable complexes of recombinant sp6 with serine proteases

A hallmark of the serpin inhibition mechanism is the formation of a stable complex of serpin and protease that cannot be dissociated by boiling in SDS [1]. To determine whether recombinant sp6 could form such a complex with its target proteases identified above, we incubated sp6 with protease and analyzed complex formation by SDS-PAGE (see Section 2). As shown in Fig. 5, incubation of sp6 with trypsin resulted in a polypeptide band of about 67 kDa (lane 1) which matches the predicted size of a complex composed of trypsin (25 kDa) and sp6 (42 kDa). By the same criterion, sp6 appeared to form a complex with plasmin (lane 3), elastase (lane 5), and thrombin (lane 7). Taken together with the protease inhibition data described above, the results of these experiments demonstrate that sp6 is a serine protease inhibitor.

Acknowledgements: We are grateful to Yue-Qing Tan for preparing the recombinant sp6 protein, to Bader Al-Anzi for help in making the

sp6 expression plasmid, and to Bill Konigsberg and Ellen LeMosy for helpful comments on the manuscript. This work was supported by a postdoctoral fellowship from the Swedish Natural Science Research Council to H.Z. and by Established Investigator Grant 9640187N from the American Heart Association to C.H.

References

- [1] Gettins, P.G.W., Patston, P.A. and Olson, S.T. (1996) Serpins: Structure, Function and Biology, R.G. Landes, Austin, TX.
- [2] Sager, R., Sheng, S., Pemberton, P. and Hendrix, M.J. (1997) *Adv. Exp. Med. Biol.* 425, 77–88.
- [3] Becerra, S.P. (1997) *Adv. Exp. Med. Biol.* 425, 223–237.
- [4] Dawson, D.W., Volpert, O.V., Gillis, P., Crawford, S.E., Xu, H., Benedict, W. and Bouck, N.P. (1999) *Science* 285, 245–248.
- [5] Coleman, S., Drahm, B., Petersen, G., Stolorov, J. and Kraus, K. (1995) *Insect Biochem. Mol. Biol.* 25, 203–207.
- [6] Levashina, E.A., Langley, E., Green, C., Gubb, D., Ashburner, M., Hoffmann, J.A. and Reichhart, J.-M. (1999) *Science* 285, 1917–1919.
- [7] LeMosy, E.K., Hong, C.C. and Hashimoto, C. (1999) *Trends Cell Biol.* 9, 102–107.
- [8] Misra, S., Hecht, P., Maeda, R. and Anderson, K.V. (1998) *Development* 125, 1261–1267.
- [9] Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A. and Struhl, K. (1998) *Current Protocols in Molecular Biology*, John Wiley and Sons, New York.
- [10] Spradling, A.C. and Mahowald, A.P. (1979) *Cell* 16, 589–598.
- [11] Miura, Y., Kawabata, S., Wakamiya, Y., Nakamura, T. and Iwanaga, S. (1995) *J. Biol. Chem.* 270, 558–565.
- [12] Suter, B. and Steward, R. (1991) *Cell* 67, 917–926.
- [13] Hoffmann, A. and Roeder, R.G. (1991) *Nucleic Acids Res.* 19, 6337–6338.
- [14] Han, J.-H., Van Deerlin, V.M.D. and Tollefsen, D.M. (1997) *J. Biol. Chem.* 272, 8243–8249.
- [15] Schimpf, S.P., Bleiker, A.J., Brecevic, L., Kozlov, S.V., Berger, P., Osterwalder, T., Krueger, S.R., Schinzel, A. and Sonderegger, P. (1997) *Genomics* 40, 55–62.
- [16] Remold-O'Donnell, E. (1993) *FEBS Lett.* 315, 105–108.
- [17] Spradling, A.C. (1993) in: *The Development of Drosophila melanogaster*, Vol. 1 (Bate, M. and Martinez Arias, A., Eds.), pp. 1–70, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- [18] Beatty, K., Bieth, J. and Travis, J. (1980) *J. Biol. Chem.* 255, 3931–3934.
- [19] Danielsson, A. and Bjork, I. (1982) *Biochem. J.* 207, 21–28.
- [20] Olson, S.T. and Bjork, I. (1991) *J. Biol. Chem.* 266, 6353–6364.
- [21] Long, G.L., Chandra, T., Woo, S.L., Davie, E.W. and Kurachi, K. (1984) *Biochemistry* 23, 4828–4837.
- [22] von Heijne, G. (1985) *J. Mol. Biol.* 184, 99–105.
- [23] Bock, S.C., Wion, K.L., Vehar, G.A. and Lawn, R.M. (1982) *Nucleic Acids Res.* 10, 8113–8125.
- [24] Remold-O'Donnell, E., Chin, J. and Alberts, M. (1992) *Proc. Natl. Acad. Sci. USA* 89, 5635–5639.
- [25] Bartuski, A.J., Kamachi, Y., Schick, C., Massa, H., Trask, B.J. and Silverman, G.A. (1998) *Genomics* 54, 297–306.
- [26] Schechter, I. and Berger, A. (1967) *Biochem. Biophys. Res. Commun.* 27, 157–162.