Peptide Design Aided by Neural Networks: Biological Activity of Artificial Signal Peptidase I Cleavage Sites†

Paul Wrede,*,‡ Olfert Landt,§ Sven Klages,|| Afshin Fatemi,⊥ Ulrich Hahn,⊥ and Gisbert Schneider‡,#

Freie Universität Berlin, Universitätsklinikum Benjamin Franklin, Institut für Medizinische/Technische Physik und Lasermedizin, AG Molekulare Bioinformatik, Krahmerstrasse 6-10, D-12207 Berlin, Germany, TIB MOLBIOL, Tempelhofer Weg 11-12, D-10829 Berlin, Germany, Max-Planck-Institut für Molekulare Genetik, Ihnestrasse 63, D-14195 Berlin, Germany, and Universität Leipzig, Fakultät für Biowissenschaften, Pharmazie und Psychologie, Institut für Biochemie, Talstrasse 33, D-04103 Leipzig, Germany

Received October 21, 1997; Revised Manuscript Received January 26, 1998

ABSTRACT: De novo designed signal peptidase I cleavage sites were tested for their biological activity in vivo in an *Escherichia coli* expression and secretion system. The artificial cleavage site sequences were generated by two different computer-based design techniques, a simple statistical method, and a neural network approach. In previous experiments, a neural network was used for feature extraction from a set of known signal peptidase I cleavage sites and served as the fitness function in an evolutionary design cycle leading to idealized cleavage site sequences. The cleavage sites proposed by the two algorithms were active in vivo as predicted. There seems to be an interdependence between several cleavage site features for the constitution of sequences recognized by signal peptidase. It is concluded that neural networks are useful tools for sequence-oriented peptide design.

The final step of eubacterial nonlipoprotein targeting and translocation is the interaction of signal peptidase I (SP I) with the protein precursor. In *Escherichia coli*, cleavage of the N-terminal signal peptide containing the secretory targeting information leads to the release of the mature protein from the plasma membrane and, as a consequence, to protein export (*1*−*3*). Cleavage of the signal peptide is strictly correlated with secretion of the protein (*4*). The signal peptide is tripartite in a targeting signal encompassing a basic N-terminal and an apolar central (core) region and the C-terminal cleavage site portion (*2*, *5*−*7*). Parts of the mature protein also seem to be required for proper precursor processing by SP I (*8*−*10*). Especially the positions −6, −3, −1, and +1 were shown to be relevant for the formation of a processing-competent structure of the cleavage site region, presumably a β-turn following an upstream more N-terminal α-helical segment (*11*−*16*). As a consequence, both processing and translocation efficiency are influenced by the properties of SP I target sites. It is, therefore, worthwhile trying to develop artificial cleavage sites for fusion protein design that are efficiently recognized and hydrolyzed by SP I. It should be emphasized that signal peptides including their cleavage site sequences do not share a significant degree of sequence identity, although they are recognized by a single protease, SP I (*17*, *18*). SP I is a special type of serine protease, and except for the natural substrates and designed substrate derivatives, no inhibitors for this enzyme are known. The reaction probably works via a catalytic serine-lysine dyad (*19*−*21*). For a deeper understanding of the SP I reaction mechanism and protein engineering applications, designed peptidic substrates will be very helpful (*22*).

Table 1: Designed SP I Cleavage Site Sequences

no	sequence identifier	amino acid sequence[a]	predicted activity[b]	measured activity[c]
1	OmpA/ RNase T1	LAGFATVAQA*AC		++[d]
2	PROSA1	VVIMSASAMA*AC	++	++
3	PROSA2	VVIMSASAMS*AC	++	+
4	PROSA3	VVVFSGSGFG*AC	++	+
5	PROSA4	LLIFSASAFA*AC	++	+
6	PROSA5	VVIMRASAMA*AC	−	−
7	SME1	FFFFGWYGWA*REAC	++	++
8	SME2	FFFFGWYGWA*AC	++	++

[a] Residues selected by the computer programs are underlined. Asterisks denote the predicted SP I cleavage sites. [b] Predictions were made on basis of the computer program PROSA (sequences 2−6), and neural networks trained on SP I cleavage site recognition (sequences 7 and 8). [c] Diameters of the halos around RNase T1-secreting *E. coli* colonies were compared to the reference (sequence 1). [d] ++, activity comparable to the reference (sequence 1); +, less active than the reference; −, inactive.

We have developed two computer-based techniques for sequence-oriented peptide design and applied these to the generation of artificial *E. coli* SP I target sites (*23*): the PROSA (protein sequence analysis) technique is a simple statistical procedure based on the frequencies of residue

† G.S. and U. H. were supported by the Fonds der Chemischen Industrie. This project was granted by the Deutsche Forschungsgemeinschaft (Sfb 312 "Gerichtete Membranprozesse") and Ha 1366/2-3 to U. Hahn.

* Corresponding author.
‡ Freie Universität Berlin.
§ TIB MOLBIOL.
|| Max-Planck-Institut für Molekulare Genetik.
⊥ Universität Leipzig.
Present address: F. Hoffman-La Roche Ltd., Pharmaceutical Division, Molecular Design & Bioinformatics, CH-4070 Basel, Switzerland.

S0006-2960(97)02603-2 CCC: $15.00
Published on Web 02/28/1998

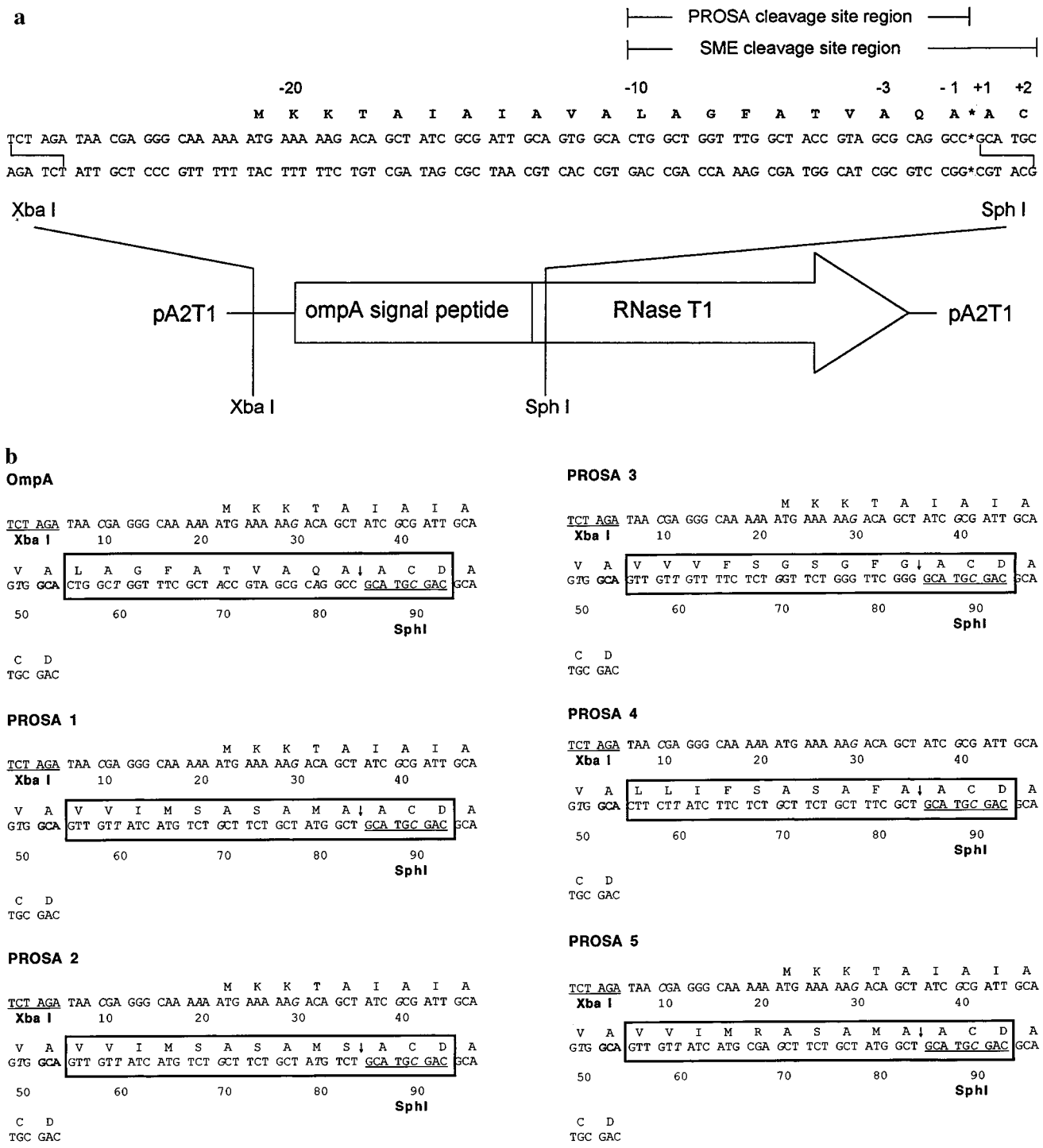

FIGURE 1: A scheme of the SP I cleavage site cassette in the plasmid expression vector pA2T1. The OmpA signal peptide sequence is given. PROSA-designed cleavage sites encompassed the relative positions −10 to −1, SME-designed cleavage sites the positions −10 to +2. The cassette between the *Xba*I and *Sph*I sites was used. *Xba*I,*Sph*I: restriction sites Figure 1b chemically synthesized nucleotide sequences to which the different PROSA sequences have been fused. The *Xba*I site fragment till *Sph*I for the wild-type OmpA cleavage site sequence is given. The boxed sequences are derived by the PROSA design algorithm (sequences PROSA-01 to PROSA-05) or wild-type sequence (OmpA), respectively. *Construction of cleavage site sequence cassettes*. A pair of constant N-terminal oligonucleotides between the *Xba*I site on the left and different pairs of primers for the altered signal peptide sequence overlapping nine bases coding for Ala-Val-Val at the beginning of the signal peptide were used to construct the cassette. The C-terminal ends of the variated part had *Sph*I compatible ends and an internal SphI site. This allowed to eliminate the amino acids Ala-Cys-Asp (underlined) by a *Sph*I digest to start with the native N-terminus of RNase T1 (Ala-Cys-Asp) or to leave an additional spacer formed by the amino acids Ala-Cys-Asp in the case of steric hindrance or impossibility to cut off the RNase from the signal peptide sequence because of the altered last amino acid at −1 of the signal peptide. The sequence PROSA-05 is inactive since it contains an arginine in position (−6) instead of a serine as in PROSA-01 (Materials and Methods).

classes in a set of aligned amino acid sequences (*5*). It generates "consensus" sequences that have in common a physicochemical property pattern which is thought to constitute an underlying functional or structural feature. The second design technique used here (SME, "simulated molecular evolution") employs artificial neural networks and

an evolutionary algorithm for peptide development (*24*). The neural networks were used to extract essential cleavage site features from a set of known cleavage and noncleavage site examples. After successful feature extraction these systems provided a mathematical model of SP I target sites which was used as a heuristic "fitness" function for a systematic search in sequence space, similar to SAR (structure−activity relationship) functions used in drug design approaches (*24*−*28*). The putative cleavage sites found were predicted to be excellent SP I substrates.

The export of a fusion protein by *E. coli* served as in vivo test system for both types of designed cleavage sites, as given by PROSA and the SME algorithms. The fusion protein was constructed from the genes coding the OmpA signal peptide with a natural N-terminal secretion signal, an SP I cleavage site cassette containing either designed or wild-type sequences, and the gene for ribonuclease T1 (RNase T1) of the mould fungus *Aspergillus oryzae* as reporter (*29*−*32*). Five PROSA-designed cleavage site sequences and two SME-designed sequences were tested for their biological activity in the fusion protein expression vector.

RESULTS AND DISCUSSION

In previous experiments, potential SP I cleavage site sequences were proposed employing the computer-based algorithms PROSA and SME (*5*, *24*) (Table 1). These stretches of amino acids consisted of 10 (PROSA-designed sequences) or 12 residues (SME-designed sequences), respectively. String matching in the PIR−International protein database (*33*) revealed that the generated peptides have no identical known counterpart. We tested the de novo designed cleavage site sequences in vivo by constructing an *E. coli* vector, pA2T1 (*31*), carrying the fusion gene of RNase T1 and the OmpA signal sequence together with a replaceable cleavage site sequence (Figure 1). Export of RNase T1 was easily detected and semiquantitatively measured using indicator plates (*34*). A color test allowed for the identification of RNase T1-secreting colonies which showed a purple halo against the blue medium. The diameter of a halo is directly related to the amount of RNase T1 secreted by *E. coli* colonies (*34*).

PROSA-Designed SP I Cleavage Sites. Five different PROSA sequences were tested for their substrate compatibility (Table 1). The designed active PROSA-sequences obey the basic features required for SP I substrate sequences (*6*, *15*, *35*). Positions −3 and −1 are occupied by small hydrophobic residues ("-3, −1 rule"), position −6 is a small residue, and a hydrophobic core region encompasses positions −10 to −7. PROSA1 revealed the highest biological activity of the PROSA-designed sequences (Table 1). The diameters of the red halo around colonies containing the natural-occurring OmpA cleavage site sequence, and the colonies expressing the sequence PROSA1 were identical. PROSA5 was inactive, as expected. PROSA5 had the identical amino acid sequences as PROSA1 besides a positively charged arginine residue in position −6, which was selected manually. Many comparative studies led to the conclusion that position −6 of SP I cleavage sites is preferred by serine, proline, or glycine while a positively charged amino acid has never been observed (*36*, *37*). Earlier studies suggested that this position belongs to the transition region from the hydrophobic core to the cleavage site region (*7*).

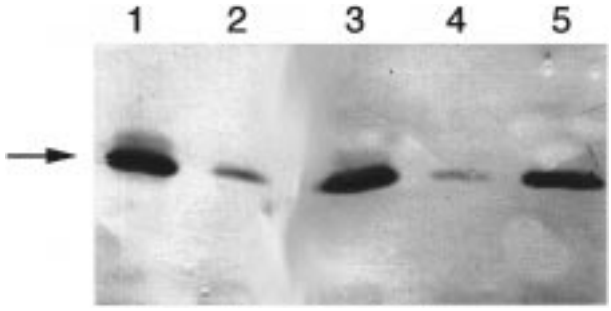

FIGURE 2: Immunoblot of pA2T1 constructions containing PROSA-designed cleavage sites expressed in *E. coli* cells. Exported proteins were isolated from the periplasm of transformed *E. coli* cells. The proteins were separated by a 14% SDS−PAGE and incubated with rabbit RNase T1 antiserum. Imaging was edited for clarity. Lane 1: PROSA-01 sequence (VVIMSASAMA*ACD). Lane 2: PROSA-03 sequence (VVVFSGSGFG*ACD). Lane 3: OmpA (control) (LAGFATVAQA*ACD). Lane 4: PROSA-02 sequence (VVIMSASAMS*ACD) Lane 5: Protein RNase T1.

The migration behavior of the processed fusion protein in a Western blot and subsequent immunodetection analysis was identical with that of the purified mature RNase T1 (Figure 2). N-terminal sequencing by Edman degradation proved that the N-termini of both processed PROSA1 and OmpA fusion proteins started with Ala-Cys-Asp-Tyr-Thr, which corresponds to the N-terminal sequence of mature RNase T1 (data not shown) (cf. Figure 1a).

A comparison of the two sequences PROSA1, PROSA2, with PROSA3 in their secretion efficiency according to the halo diameters give some hints to the role of position −1. The replacement of Ala (PROSA1) by Ser (PROSA2) led to a reduction of the halo size, indicating that the secretion efficiency was reduced. When both alanines in positions −1 and −3 were replaced by glycine (PROSA3), the halo size was somewhat smaller but large enough to be designated by a single "+". It is possible that *E. coli* SP I requires a defined structural element in the substrate for efficient hydrolysis which is favored by an alanine in −1 and to a lesser extent by the more flexible and less hydrophobic glycine or a more polar residue like serine.

SME-Designed SP I Cleavage Sites. Another approach toward a rational design of peptides is the application of evolutionary algorithms as a sequence-generating procedure that is coupled to a selection mechanism represented by a trained artificial neural network (*24*). The trained neural network serves as a fitness function structuring the sequence space into regions of higher and lower fitness values (*27*). The SME-design cycle led to a new potential SP I cleavage site which is different from the PROSA-designed sequences (Table 1, SME1 sequence). The SME1 sequence FFFFGWYGWA*RE was identified as an idealized SP I target site (*24*).

We tested the biological activity of the SME-designed sequence in the same system described for the PROSA-sequences. The SME1 cleavage site is recognized by SP I as well as the OmpA/RNase T1 reference (Table 1), i.e., the pA2T1-transformed *E. coli* cells with the OmpA or SME-designed cleavage site sequence exported the reporter protein RNase T1 equally well according to the halo size. An SDS−PAGE gel analysis of the exported proteins shows an

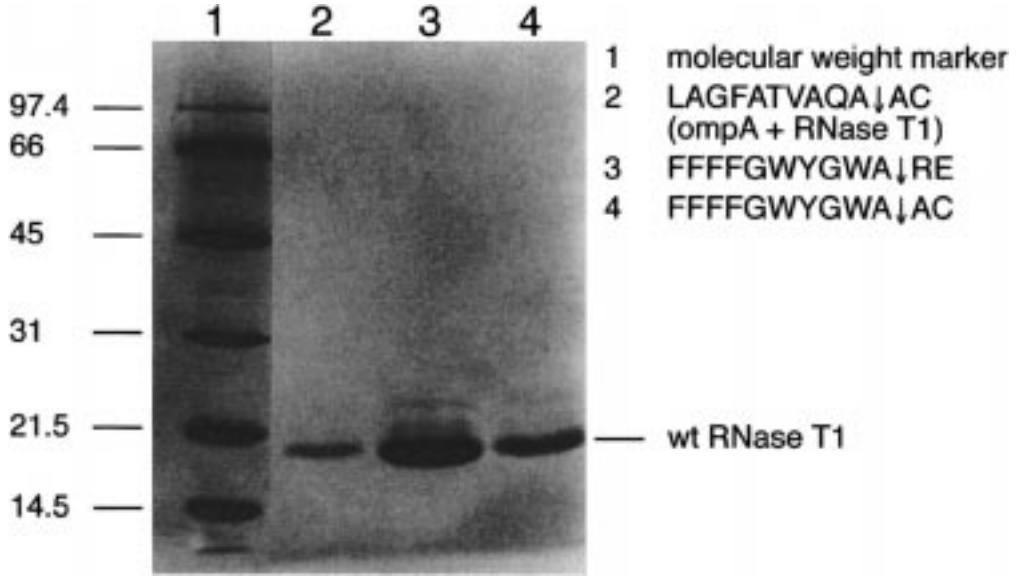

FIGURE 3: SDS−PAGE, Coomassie brillant blue stained, showing the migration of exported and SP I-processed fusion proteins (OmpA/cleavage site/RNase T1) from pA2T1 transformed *E. coli* cells. Lane 1: molecular weight marker. Lane 2: OmpA/RNase T1 fusion protein (LAGFATVAQA*AC). Lane 3: SME1 (FFFFGWYGWA*REAC). Lane 4: SME2 (FFFFGWYGWA*AC). Tick (right side) indicates MW_{app} = 17 500 position for the wt RNase T1 (*29*).

identical migration behavior of the processed RNase T1 molecules (Figure 3). From this result, we conclude that the SME1 cleavage site is recognized by SP I at the predicted position WA*RE. Strikingly, the proposed RNase T1 could not be purified by ionic exchange chromatography or other ion absorbing procedures as it would be necessary for a MALDI analysis. The presence of positively charged amino acids at the N-terminus of the mature protein should affect secretion adversely (*3*, *38*). Since the positively charged arginine in position +1 is a neighbor of the negatively charged glutamic acid residue in position +2, the adverse effect might be compensated. The SME1 sequence with the two oppositely charged amino acid residues did not seem to impair protein export. Omitting these two residues in sequence SME2 led to a comparable secretion efficiency (Table 1). According to an SDS−PAGE run (Figure 3, lane 4) the processed SME2-construct showed almost the same migration behavior as the processed SME1-construct (Figure 3, lane 3). To verify this conclusion a molecular mass determination of processed RNase T1 by MALDI (matrix-assisted laser desorption) mass spectroscopy (ms) was performed. The *m*/*z* ratio corresponds to a molecular mass of 11 085 Da (Figure 4). This finding proves that the designed SME 2 sequence is recognized and processed by SP I as predicted. This proves that the cleavage occurred at the expected position, namely at WA*AC.

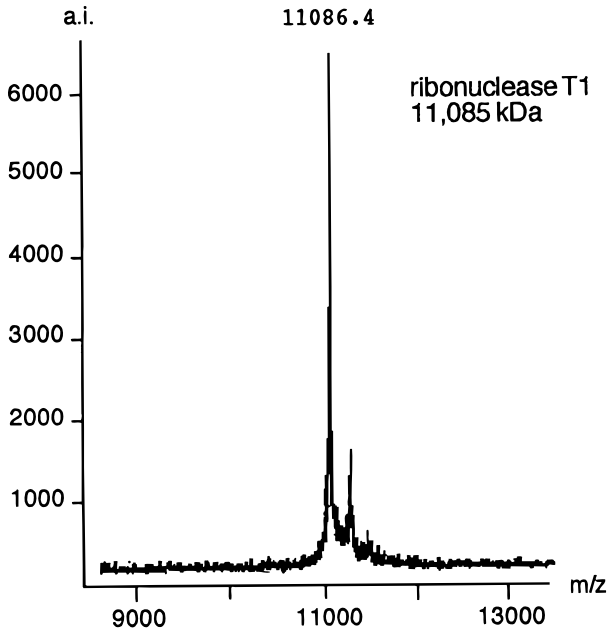

FIGURE 4: Molecular mass spectroscopic (MALDI) analysis of the exported and SP I-processed OmpA/SME 2 cleavage site/RNase T1 fusion protein. The calculated mass of RNase T1 is MW = 11085 Da, and the experimentally determined value is MW = 11086 Da.

The experiments presented here show that both the PROSA and the SME design procedure led to artificial SP I target sites that are active in vivo. A specific property of neural filter systems for pattern recognition is parallel data processing. Interactions between residues within the sequence window under investigation can easily be considered (*25*, *39*). During the training phase of neural network, development weight factors for each sequence position are optimized. Useful weight factors are related to the relevance of each sequence position with regard to the functional or structural feature analyzed, e.g., SP I substrate compatibility (*40*). The in vivo experimental results confirm that the trained artificial neural network applied for cleavage site design represented a useful mathematical model of SP I target sites (*41*), although its predictive value is limited compared to other approaches (*42*). However, our results confirm the assumption that essential SP I cleavage site features are locally encoded by a short stretch of residues. Therefore, the [−10,−1/+2] windows considered here for design were sufficient.

It should be taken into account that the SME design started from a random sequence, and after a small number of computer-based optimization cycles, a biologically active sequence was found in sequence space (*24*). During this

procedure, the positions −2 and −6 were comparably restricted in the variability of the tolerated residues. In position −6 of SME1, a glycine was selected, and position −2 is occupied by tryptophan. Although a proline residue is predominant in position −6 of natural SP I cleavage sites, glycine was selected by SME. In contrast to our findings, −6 glycine mutants of M13 procoat protein were inactive in vivo (*15*). This supports the assumption that the SME1 sequence was designed in parallel "as a block" considering possible interactions and dependencies between residues. The neural network model used in the design procedure clearly recognized crucial cleavage site features that are more complex than just residue frequencies at individual positions. The fact that in −6 of SME1 a glycine was automatically selected may be a consequence of the necessity to adapt the designed sequence to a more important cleavage site feature than the feature represented by the −6 position. In −1, alanine was selected, which is in perfect accordance with the PROSA-designed sequences and the above discussion of the importance of this residue position (see previous section). It is reasonable to assume that the bulky hydrophobic tryptophan residue in position −2, flanked by the alanine in −1 and the glycine in −3, represents an idealized predominant SP I cleavage site feature ("-3,−1 box"). The hydrophobic core region of the SME1 sequence is also pronounced by the four phenylalanines in a row (positions −10 to −7), which is not observed in natural signal peptides. Probably only a reasonable combination of several individual features constitutes an active SP I target site (*43*, *44*), and the SME1 sequence is an idealized version which was designed de novo.

MATERIALS AND METHODS

Construction of Signal Peptide Gene Cassettes. At the C-terminus of the newly designed signal peptides, we introduced a repeat of the terminal three amino acids, Ala-Cys-Asp, and therewith also a repetition of the SphI recognition site. This yielded either a signal peptide with the corresponding amino acid spacer in case of steric hindrance or on the other hand allowed to eliminate these additional three codons via a *Sph*I digestion.

Cassettes coding for the different signal peptide variants were composed of four oligonucleotides each. Two of them (upstream part) comprised fixed sequences representing the constant region, the other two (downstream region) contained varying sequences coding for the different variants (Figure 1). A pair of constant N-terminal oligonucleotides between the *Xba*I site on the left and different pairs of primers for the altered signal peptide sequence overlapping nine bases coding for Ala-Val-Val at the beginning of the signal peptide were used to construct the cassette. The C-terminal ends of the variated part had SphI compatible ends and an internal *Sph*I site. All oligodeoxy-ribonucleotides were synthesized on a Pharmacia Gene assembler using modified protocols. Products were purified by MonoQ FPLC anion exchanger chromatography. They were hybridized, ligated into the *Sph*I and *Xba*I sites of the RNase T1 expression/secretion vector pA2T1 to substitute the original cassette and used to transform *E. coli* DH5a as described (*30*). The correctness of each mutant was verified by DNA sequencing. Further standard cloning procedures were performed as described (*45*).

RNase T1 secreting clones were identified by the red halos on RNA/toluidine blue indicator plates (*34*).

Immunodetection. Western blots were incubated with antisera against RNase T1 from rabbit. Bound antibodies were detected with HRPO (horseradish peroxidase) conjugated mouse antibodies against rabbit IgGs. Immunodetection was performed as described in ref *45*.

ACKNOWLEDGMENT

Georg Büldt and Gerhard Müller are thanked for working facilities, many helpful discussions and encouragement. Stefan Müller is thanked for preparing the figures. Werner Schröder performed the N-terminal sequencing.

REFERENCES

1. Pugsley, A. P. (1989) *Protein targeting*, Academic Press, San Diego.
2. Wickner, W., Driessen, A. J., and Hartl, F.-U. (1991) *Annu. Rev. Biochem. 60*, 101−124.
3. von Heijne, G. (1994). in *Concepts in Protein Engineering and Design* (Wrede P., and Schneider, G., Eds.) Walter de Gruyter, Berlin, New York, 263−279.
4. Laforet, G. A., and Kendall, D. A. (1991) *J. Biol. Chem. 266*, 1326−1334.
5. Schneider, G., and Wrede, P. (1993) *Protein Seq. Data Anal. 5*, 227−236.
6. von Heijne, G. (1983) *Eur. J. Biochem. 133*, 17−21.
7. von Heijne, G. (1990) *J. Membr. Biol. 115*, 195−201.
8. Dev, I. K., Ray, P. H., and Novak, P. (1990) *J. Biol. Chem. 265*, 2009−20072.
9. Barkocy-Gallagher, G. A., and Bassford, P. J. Jr. (1992) *J. Biol. Chem. 267*, 1231−1238.
10. Nilsson, I., and von Heijne, G. (1992) *FEBS Lett. 299*, 243−246.
11. Kuhn, A., and Wickner, W. (1985) *J. Biol. Chem. 260*, 15914−15918.
12. Cornell, D. G., Dluhy, R. A., Briggs, M. S., McKnight, C. J., and Gierasch L. M. (1989) *Biochemistry 28*, 2789−2797.
13. Bird, P., Gething, M. J., and Sambrook, J. (1990) *J. Biol. Chem. 265*, 8420−8425.
14. Bruch, M. D., and Gierasch, L. M. (1990) *J. Biol. Chem. 265*, 3851−3858.
15. Shen, L. M., Lee, J.-I., Cheng, S., Jutte, H., Kuhn, A., and Dalbey, R. E. (1991) *Biochemistry 30*, 11775−11781.
16. Jones, J. D., and Gierasch, L. M. (1994) *Biophys. J. 67*, 1546−1561.
17. Watson, M. E. E. (1984) *Nucleic Acids Res. 12*, 4155−4174.
18. Randall, L. L., and Hardy, S. J. S. (1989) *Science 243*, 1156−1159.
19. Black, M. T., Munn, J. G. R., and Allsop, A. E. (1992) *Biochem. J. 282*, 539−543.
20. Dalbey, R. E., and von Heijne, G. (1992) *Trends Biochem. Sci. 17*, 474−478.
21. Paetzel, M., and Dalbey, R. E. (1997) *Trends Biochem. Sci. 22*, 28−31.
22. Nilsson, I., and von Heijne, G. (1991) *J. Biol. Chem. 266*, 3408−3410.
23. Schneider, G., Lohmann, R., and Wrede, P. (1994) in *Concepts in Protein Engineering and Design*, (Wrede P., and Schneider G., Eds.) Walter de Gruyter, Berlin, New York, 281−317.
24. Schneider, G., and Wrede, P. (1994) *Biophys. J. 66*, 335−344.
25. Schneider, G., Schuchhardt, J., and Wrede, P. (1994) *Comput. Appl. Biosci. 10*, 635−645.
26. Schneider, G, Schuchhardt, J., and Wrede, P. (1995) *Biophys. J. 68*, 434−447.
27. Schneider, G., Schuchhardt, J., and Wrede, P. (1995) *Biol. Cybern. 73*, 245−254.
28. Devillers, J. (1996) *Neural networks in QSAR and drug design*, Academic Press, London.

29. Quaas, R., McKeown, Y., Stanssens, P., Frank, R., Blöcker, H., and Hahn, U. (1988) *Eur. J. Biochem. 173*, 617−622.
30. Quaas, R., Grunert, H.-P., Kimura, M., and Hahn, U. (1988) *Nucleosides Nucleotides 7*, 619−623.
31. Grunert, H. P., Zouni, A., Beinecke, M., Quaas, R., Georgalis, Y., Saenger, W., and Hahn, U. (1991) *Eur. J. Biochem. 204*, 947−961.
32. Hahn, U. and Heinemann, U. (1994) in *Concepts in Protein Engineering and Design*, (Wrede P., and Schneider, G., Eds.) Walter de Gruyter, Berlin, New York, 109−168.
33. Barker, W. C., George, D. G., Mewes, H.-W., and Tsugita, A. (1992) *Nucleic Acids Res. 20*, 2023−2026.
34. Quaas, R., Landt, O., Grunert, H. P., Beincke, M., and Hahn, U. (1989) *Nucleic Acids Res. 17*, 3318.
35. Perlman, D., and Halvorson, H. A. (1983) *J. Mol. Biol. 167*, 391−409.
36. Goldstein, J., Lehnhardt, S., and Inouye, M. (1991) *J. Biol. Chem. 266*, 14413−14417.
37. Hoyt, D. W., and Gierasch, L. M. (1991) *J. Biol. Chem. 266*, 14406−14412.
38. Andersson, H., and von Heijne, G. (1991) *Proc. Natl. Acad. Sci. 88*, 9751−9754.
39. Rumelhart, D. E., and McClelland, J. L. (1986) in *Parallel Distributed Processing* (The PDP Research Group, Eds.) MIT Press, Cambridge, MA.
40. Schneider, G., Röhlk, S., and Wrede, P. (1993) *Biochem. Biophys. Res. Commun. 194*, 951−959.
41. Claros, M. G., Brunak, S., and v. Heijne, G. (1997) *Curr. Opin. Struct. Biol. 7*, 394−398.
42. Nielsen, H., Engelbrecht. J., Brunak, S., and von Heijne, G. (1997) *Protein Eng. 10*, 1−6.
43. Laforet, G. A., Kaiser, E. T., and Kendall, D. A. (1989) *J. Biol. Chem. 264*, 14478−14485.
44. Izard, J. W., Rusch, S. L., and Kendall, D. A. (1996) *J. Biol. Chem. 271*, 21579−21582.
45. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning*, Cold Spring Harbor Laboratory Press, Plainview, NY.

BI9726032