Volume 301, number 3, 253–257 FEBS 10958 April 1992

cDNA-derived amino-acid sequence of lentil seedlings' amine oxidase

Antonello Rossi[a], Raffaele Petruzzelli[b] and Alessandro Finazzi Agrò[a]

Dipartimenti di [a]Medicina Sperimentale e [b]Biologia, Università di Roma 'Tor Vergata', Via O. Raimondo, 00173 Roma, Italy

Received 9 January 1992; revised version received 18 February 1992

Lentil seedlings' amine oxidase (LSAO) cDNAs were identified in a cDNA λgt 10 library by plaques hybridization. The nucleotide sequence of a 2111 bp clone was determined. It contains part of a signal peptide, the complete sequence coding for the mature protein and the 3'-untranslated region of the mRNA. The deduced protein sequence shows that the mature protein is composed of 569 amino acids with a molecular mass of 67 kDa, also taking into account the glucidic component. The LSAO cDNA was identified by sequencing the N-terminal part of the protein and several tryptic peptides. The protein sequence shows a characteristic hexapeptide present in amine oxidases containing 6-hydroxydopa as the organic cofactor. Three conserved histidines might be the ligands of copper bound to the enzyme.

6-Hydroxydopa; Amine oxidase; cDNA cloning; *Lens culinaris*

1. INTRODUCTION

Copper containing amine oxidases (AOs) (EC 1.4.3.6) catalyze the oxidative deamination of polyamines to the corresponding aldehydes with the production of ammonia and hydrogen peroxide. AOs are widespread among living organisms. Though the function of these enzymes still awaits to be elucidated, their action on polyamines grants an important physiological role. In fact polyamines are involved in a variety of processes in plants. Putrescine levels have been found to increase during ripening of fruits of a tomato cultivar with a prolonged storage life [1]. Accumulation of polyamines has been shown to induce chilling tolerance in maize [2], citrus [3] and zucchini squash [4]. A recent study shows that the most tolerant varieties of wheat to osmotic stress are those with the highest putrescine accumulation ability [5]. Putrescine accumulation is also involved in the response of plants to acid rain [6], low oxygen concentration [7] and to ozone treatment [8]. These examples show that accumulation of polyamines is a response to a great variety of stresses (see [9] for a review). Accumulation of polyamines is also observed during seedling development [10]. In lentils this accumulation is paralleled by an increase of amine oxidase activity [11] that reaches a peak when the seedlings are growing faster [12]. Despite the huge amount of studies conducted, little is known about the significance of these biochemical changes and the possible involvement of AOs.

Lentil seedlings' amine oxidase (LSAO) is a dimeric enzyme with a molecular weight of 140 kDa [13]. Each subunit contains a copper ion and an organic cofactor. The nature of the organic cofactor has been a matter of debate for long time. Recently it has been identified as 6-hydroxydopa (TOPA) in bovine serum amine oxidase (BSAO) [14]. Evidence has been presented for the presence of TOPA also in LSAO [15]. In this paper we describe the isolation and sequence analysis of the LSAO cDNA. An esapeptide presumably containing the TOPA cofactor was found as well as three sequences containing possible copper ligands.

Abbreviations: AO, amine oxidase; bp, base pair(s); BSAO, bovine serum amine oxidase; cDNA, DNA complementary to mRNA; Da, dalton; HPAO, *Hansenula polymorpha* amine oxidase; kDa, kilodalton; kb, kilobase (pair); LSAO, lentil seedlings' amine oxidase; nt, nucleotide(s); PTH, phenyl-thiohydantoin; SDS, sodium dodecyl sulphate; TOPA, 6-hydroxydopa.

Correspondence address: Prof. Alessandro Finazzi Agrò, Dipartimento di Medicina Sperimentale e Scienze Biochimiche, Università di Roma 'Tor Vergata', Via Orazio Raimondo, I-00173 Roma, Italy. Fax: (39) (6) 7234793.

2. MATERIALS AND METHODS

2.1. *Plant material*

Lentil (*Lens culinaris*) seeds were soaked for 12 h in autoclaved tap water at room temperature. The seeds were then grown in a greenhouse at 25°C in the dark for 5 days and watered daily with tap water.

2.2. *Enzyme isolation and amino-acid sequence analysis of LSAO*

Amine oxidase was purified from lentil seedlings as described by Flores et al. [13].

LSAO (5 mg) was solubilized in 0.1 M ammonium bicarbonate pH 7.8 in the presence of 2 M urea and digested with trypsin (enzyme/LSAO ratio 1:50 (w/w)) at 37°C, overnight. Tryptic peptides were purified on a reverse-phase column (Aquapore RP 300, 4.6 × 250 mm, Applied Biosystems) developed in 60 min with a linear gradient of 0% to 60% acetonitrile in 0.2% trifluoroacetic acid, generated in a Beckman model 340 instrument at a flow rate of 1.0 ml/min. The elution of peptides was monitored both at 220 and 280 nm with a Beckman model 165 spectrophotometer.

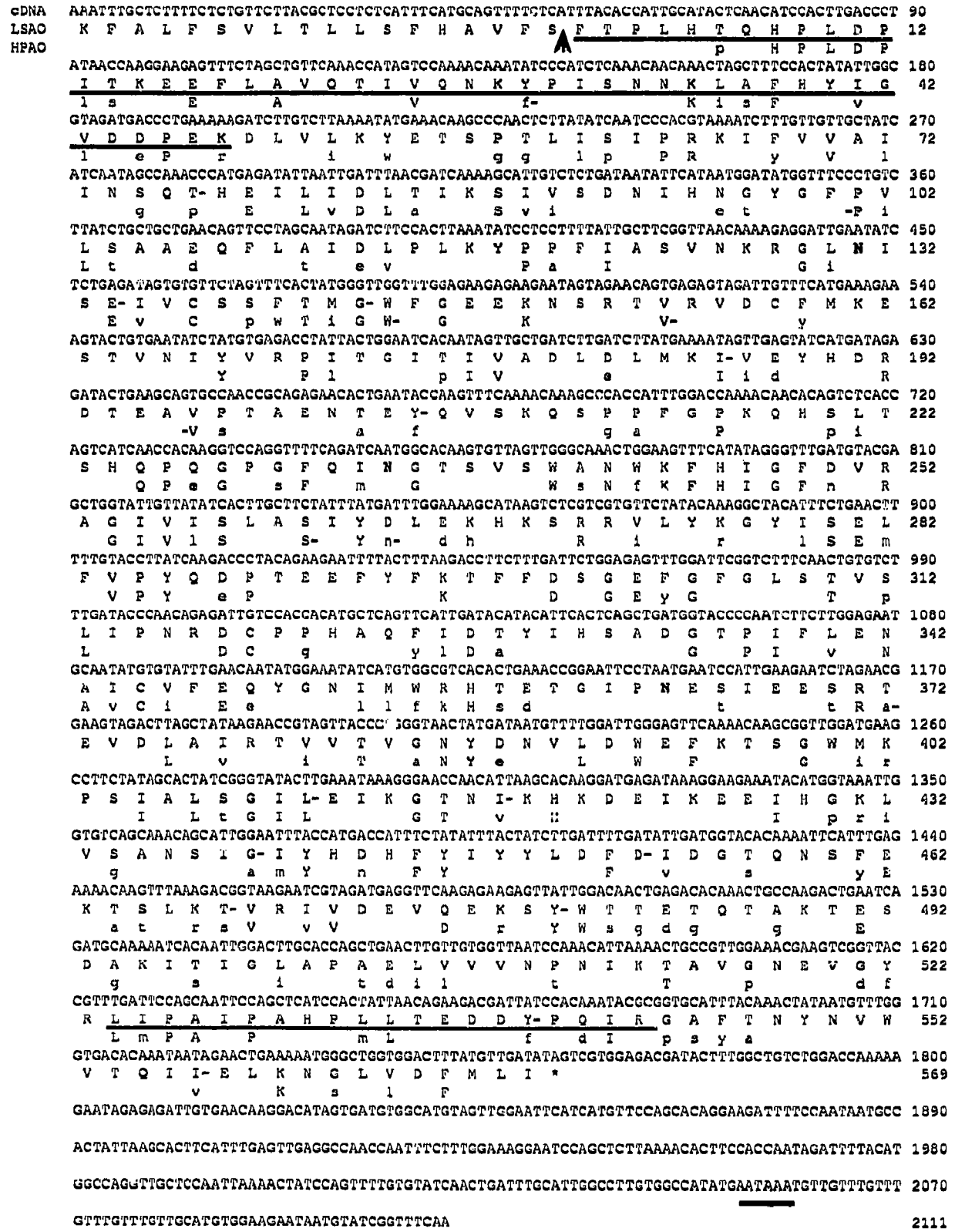

Fig. 1. Comparison of cDNA and the amino-acid sequence of LSAO with the sequence of yeast (*Hansenula polymorpha*) amine oxidase (HPAO) [24]. Numbering of the amino-acid sequence refers to the amino terminus of the mature protein. In the HPAO sequence only identical (upper case) or similar (lower case) residues to the LSAO sequence are shown. Hyphens indicate a gap introduced for maximizing matches. The features marked are: the cleavage site of a signal peptidase (vertical arrow), peptides directly sequenced by Edman degradation (underlined), possible glycosylation sites (asparagines in bold), the polyadenilation site (thick underlined).

Automated Edman degradation was carried out using an Applied Biosystems model 470A gas phase sequencer equipped with an Applied Biosystems model 120A PTH analyzer for the on-line detection of the PTH-amino acids.

2.3. *Construction and screening of the cDNA library*

Total RNA was extracted from 5-days-old lentil seedlings grown in the dark as described previously [16]. Polyadenylated RNAs were isolated by affinity chromatography on oligo(dT)-cellulose [17]. 10 g of seedlings yielded 2 mg of total RNA and 50 μg of poly(A)$^+$ RNA. The cDNA was synthesized using the Riboclone cDNA synthesis system (Promega) and tailored with *Eco*RI linkers using the Riboclone *Eco*RI linker ligation system (Promega). The yield of cDNA was 1 μg starting from 5 μg of poly(A)$^+$ RNA.

The tailored cDNA was then ligated into dephosphorylated, *Eco*RI cut, λgt 10 and the resulting recombinant phage DNAs were packaged using the Packagene system (Promega). The recombinant phages were plated on *E. coli* C600 Hfl for screening. 100 ng of cDNA yielded 500,000 plaques. The library was screened by plaque hybridization [18] using end-labelled oligonucleotides with [γ-^{32}P]ATP (Amersham).

Oligonucleotides sequences were 5'-TGNGGRTARTCRTCYTC-3' and 5'-ATNGGRTAYTTKATKTG-3' (N=A,T,C, or G; R=A or G; Y=C or T; K=G or T). They were purchased from GENESET.

2.4. *DNA sequencing*

cDNAs to be sequenced were subcloned into pBS (Stratagene) and both strands were sequenced using the Sequinase version 2.0 kit (USB) applied to plasmid DNA [19]. Nested deletions were made with Erase-a-base system (Promega). The primers used for sequencing were the M13-20 primer present in the Sequinase kit and the reverse sequencing primer 5'-CAGGAAACAGCTATGAC-3' (Promega).

Since there are two *Eco*RI sites and since the library was made with *Eco*RI linkers the cDNA was also sequenced directly in λgt 10 to confirm the sequences around the *Eco*RI sites. The possibility of having chimeric clones was excluded because the first *Eco*RI site is within the coding region of the protein, for which there is sequence information, and the second one does not contain the characteristic sequence of the linker.

2.5. *Sequence comparison and hydropathy analysis*

Conservative amino acid substitution was as reported in ref. [20].

Hydrophobicity values were calculated using a window size of 9 residues according to Kyte and Doolittle [21].

3. RESULTS

Total mRNA from 5-days-old lentil seedlings grown in the dark was taken to construct a cDNA library in the λgt 10 vector. This is the richest source of mRNA coding for LSAO, as previously shown [12].

Two degenerated oligonucleotides representing DNA sequences complementary to all possible coding sequences for the tryptic peptides GluAspAspTyrProGln and GlnAsnLysTyrProIle were used to screen the library by plaque hybridization. Approximately 50,000 recombinant phages were screened, from which two positives clones were found with inserts of 2000 and 2100 bp, respectively. They were subcloned in pBS and subjected to sequence analysis. The longest insert contained the full coding region for the mature protein plus a short hydrophobic peptide besides the 3'-untranslated region of the LSAO mRNA.

The identity of the LSAO cDNA was established by the correspondence with the N-terminal sequence of

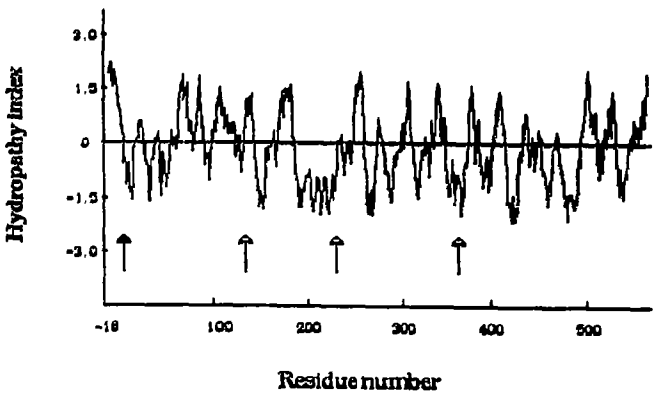

Fig. 2. Hydropathy plot of the LSAO amino-acid sequence derived from cDNA. The black-headed arrow indicates the N terminus of the mature protein. The white-headed arrows indicate three possible glycosylation sites. Numbering refers to the amino terminus of the mature enzyme. Hydrophobicity values were calculated using a window size of 9 residues according to Kyte and Doolittle [21]. Values indicating hydrophobic and hydrophilic regions are above and below zero, respectively.

LSAO protein and with several other tryptic peptides of the deduced amino acid sequence (Fig. 1). The open reading frame of 1707 nucleotides (Fig. 1) encodes a polypeptide of 569 amino acids with a molecular weight of 64,321 Da that, taking into account the presence of about 3% carbohydrates (G. D'Andrea, personal communication), gives a molecular mass of 67 kDa for the monomer, which is in agreement with the reported value of 70 kDa [13]. The amino-acid composition of the deduced protein sequence is similar to that determined from purified LSAO [13]. There are three putative glycosylation sites (Fig. 1) of which two only bear a carbohydrate chain (G. D'Andrea, personal communication). The sites at position 234 and 364 seem the most likely candidates since the other site appears to be located in a hydrophobic region (Fig. 2).

The deduced amino acid sequence of the largest insert does not contain methionines upstream the N-terminal residue (phenylalanine) of the mature protein, and therefore it is part of a supposed signal peptide which is incomplete in this clone. The cleavage site (Figs. 1 and 2) conforms to the '(−3,−1) rule' [22]. The deduced protein sequence shows a region of six highly conserved amino acids among copper containing AOs (Fig. 3) [23] where the tyrosine in the third position was found to be modified into TOPA in BSAO.

In an attempt to find out the ligands of the copper, homologies were searched for in already known sequences of copper-containing AOs. Fig. 1 shows three regions of significant similarity between LSAO and yeast AO [24]. They all contain a conserved histidine residue.

There is a 3'-untranslated sequence of 350 nucleotides besides the 1707 nucleotides of the open reading frame.

Source		Method	Ref.
Bovine serum	LNXDYV	Edman degradation	14
Porcine serum	LNXDYV	Edman degradation	23
Pig kidney	YNXDY	Edman degradation	23
Yeast	ANXEYC	Edman degradation	23
Yeast	ANYEYC	cDNA	24
Pea seedlings	GNXDNV	Edman degradation	23
Lentil seedlings	GNYDNV	cDNA	This work

Fig. 3. Comparison of the amino-acid sequence around TOPA in several amine oxidases obtained either from cDNA sequencing or direct sequence of the purified enzymes. X indicates a blank position obtained by Edman degradation.

It contains the eucaryotic polyadenylation signal AATAAA (Fig. 1).

4. DISCUSSION

A previous study showed that lentil seedlings grown in the dark for five days are the best source of LSAO cDNA [12]. This was the starting material for the construction of a cDNA library in λgt 10. Oligonucleotides constructed with reference to tryptic peptides were used to isolate LSAO cDNA clones by plaque hybridization. Sequencing the largest positive clone yielded an open reading frame of 1707 nucleotides and a 3'-untranslated region of 350 nucleotides. The latter contain a typical polyadenylation signal (Fig. 1).

The deduced protein sequence shows a putative signal peptide extending 18 residues upstream the N-terminal Phe and a peptidase cleavage site which was confirmed by sequencing the N-terminal part of the protein. There are no methionines present in this peptide, hence the sequence of the presumptive signal peptide appears to be incomplete.

The N-terminal sequence, together with that of 4 other tryptic peptides (Fig. 1) gave direct proof on the identity of this cDNA with LSAO. The predicted sequence of the mature protein shows that the monomer is composed of 569 amino acids with a calculated molecular weight of 67 kDa, in good agreement with a previous report [13]. It also shows three putative glucosylation sites. LSAO is in fact a glycoprotein containing two carbohydrate chains per monomer (G. D'Andrea, personal communication), therefore only two of these sites actually bind carbohydrates. Inspection of the hydropathy plot in Fig. 2 suggests that Asn-234 and Asn-364 may be *N*-glycosylated, since Asn-130 sits in a hydrophobic environment.

Copper-containing AOs have two different cofactors. One is cupric copper and the other has been recently identified as TOPA [14] which most probably results from a post-synthetic modification of a tyrosine. The amino-acid sequence of LSAO shown in Fig. 3 is very similar to the active-site peptide of bovine serum AO, where the presence of TOPA has been demonstrated at the position X [14]. This sequence is also present in other AOs where the presence of TOPA has been inferred from resonance Raman studies [23,25]. Therefore this sequence could represent a consensus sequence for TOPA quinone, which can be formed from the tyrosine located between an asparagine and an aspartic acid as proposed previously [23]. It must also be pointed out that recent EPR evidence has suggested the presence of TOPA quinone in LSAO [15]. From all this data it can be inferred that the tyrosine at position 387 of LSAO cDNA is modified to TOPA in the mature protein.

A sequence comparison between yeast AO [24] and LSAO was carried out in order to find out possible ligands for the copper. The similarities between these two proteins besides the region of the organic cofactor lie in three positions only, as shown in Fig. 1. All these three sequences contain a histidine residue. Previous spectroscopic studies have shown that the copper ion in BSAO should have three histidines coordinated [26]. Therefore the histidines at positions 8, 246 and 357 can be reasonably proposed as copper ligands.

Expression of active LSAO in vivo and site-directed mutagenesis will demonstrate if the histidines 8, 246 and 357 are actually the copper ligands. The importance of individual amino acids for the possible post-translational modification leading to the appearance of 6-hydroxydopa will be also assessed.

Acknowledgements: The authors gratefully acknowledge the help of Prof. D. Barra, University of Roma 'La Sapienza' for sequencing a few tryptic peptides. The authors wish to thank Mr. Maurizio Buganza for his help in subcloning and sequencing the LSAO clones. This research was supported by the National Research Council of Italy, Special Project RAISA, Sub-project N. 2, Paper N. XXX.

REFERENCES

[1] Saftner, R.A. and Baldi, B.G. (1990) Plant Physiol. 92, 547–550.
[2] Songstad, D.D., Duncan, D.R. and Widholm, J.M. (1990) J. Exp. Bot. 41, 289–294.
[3] Kushad, M.M. and Yelenosky, G. (1987) Plant Physiol. 84, 692–695.
[4] Kramer, G.F. and Wang, C.Y. (1990) J. Plant Physiol. 136, 115–119.
[5] Erdei, L., Trivedi, S., Takeda, K. and Matsumoto, H. (1990) J. Plant Physiol. 137, 165–168.
[6] Santerre, A., Markiewicz, M. and Villanueva, V.R. (1990) Phytochemistry 29, 1767–1769.
[7] Reggiani, R. and Bertani, A. (1989) J. Plant Physiol. 135, 375–377.
[8] Langebartels, C., Kerner, K., Leonardi, S., Schraudner, M., Trost, M., Heller, W. and Sandermann, H. Jr. (1991) Plant Physiol. 95, 882–889.
[9] Flores, H.E. (1990) in: Stress Responses in Plants: Adaptation and Acclimation Mechanisms (R.G. Alscher and J.R. Cummings, Eds.) Vol. 12, pp. 217–239, Wiley-Liss, New York.
[10] Slocum, R.D., Kaur-Sawhney, R. and Galston, A.W. (1984) Arch. Biochem. Biophys. 235, 283–303.
[11] Federico, R. and Angelini, R. (1988) Planta 173, 317–321.
[12] Maccarrone, M., Rossi, A., Avigliano, L. and Finazzi Agrò, A. (1991) Plant Sci. 79, 51–55.
[13] Floris, G., Giartosio, A. and Rinaldi, A. (1983) Phytochemistry 22, 1871–1874.

[14] Janes, S.M., Mu, D., Wemmer, D., Smith, A.J., Kaur, S., Maltby, D., Burlingame, A.L. and Klinman, J.P. (1990) Science 248, 981–987.
[15] Pedersen, J.Z., El-Sherbini, S., Finazzi Agrò, A. and Rotilio, G. (1992) Biochemistry 31, 8–12.
[16] Lottemann, J., Schell, J. and Willmitzer, L. (1987) Anal. Biochem. 163, 16–20.
[17] Aviv, H. and Leder, P. (1972) Proc. Natl. Acad. Sci. USA 69, 1408–1412.
[18] Wahl, G.M. and Berger, S.L. (1987) in: Methods in Enzymology (S.L. Berger and A.R. Kimmel, Eds.) Vol. 152, pp. 415–423, Academic Press, San Diego, CA.
[19] Chen, E.J. and Seeburg, P.H. (1985) DNA 4, 165–170.
[20] Dayhoff, M.O., Schwartz, R.M. and Orcutt, B.C. (1978) in: Atlas of Protein Sequence and Structure (M.O. Dayhoff, Ed.) Vol. 5, Suppl. 3, pp. 345–352, Natl. Biomed. Res. Found., Silver Spring, Washington D.C.
[21] Kyte, J. and Doolittle, R.F. (1982) J. Mol. Biol. 157, 105–132.
[22] von Heijne, G. (1986) Nucleic Acids Res. 14, 4683–4690.
[23] Janes, S.M. (1990) Ph.D. Thesis, University of California, Berkeley, USA.
[24] Bruinenberg, P.G., Evers, M., Waterham, H.R., Kuipers, J., Arnberg, A.C. and AB, G. (1989) Biochim. Biophys. Acta 1008, 157–167.
[25] Brown, D.E., McGuirl, M.A., Dooley, D.M., Janes, S.M., Mu, D. and Klinmann, S.P. (1991) J. Biol. Chem. 266,. 4049–4051.
[26] Scott, R.A. and Dooley, D.M. (1985) J. Am. Schem. Soc. 107, 4348–4350.

Note added in proof

The nucleotide sequence data reported will appear in the EMBL, GenBank and DDJB Nucleotide Sequence Databases under the accession number X64201.