Primary and 3D structure of chitobiase from *Serratia marcescens*

Tews, I., Vincentelli, R., Perrakis, A., Dauter, Z., Wilson, K. S. and Vorgias, C. E.*

European Molecular Biology Laboratory (EMBL), c/o DESY, Notkestrasse 85, 22603 Hamburg, Germany. Tel: +49-40-89902118, Fax:+49-40-89902149
*Corresponding author

Summary

A 4.5 kb DNA fragment of the genomic DNA of the soil bacterium *Serratia marcescens* was cloned into pEMBL18 (1). The construct was used to obtain the DNA sequence of chitobiase gene and purify the protein which was consitutively expressed. A single open reading frame encoding a protein of 888 amino acids has been identified. It shows high similarity to other known bacterial chitobiases and hexaminidases (2). Expression of the chitobiase gene in *E. coli* yielded an enzymatically active protein with an estimated molecular weight of 97 kDa. The deduced amino acid protein is 27 amino acids longer at the N-terminus than that determined by amino acid sequencing of the purified protein. This suggests that a leader sequence of 27 amino acids is removed during transport of the enzyme across the cell membrane into the periplasm. The purification protocol for the recombinant enzyme published in (3) has been optimized. Alternative crystallisation conditions producing better diffracting crystals are also presented. Complete native data to 1.9 Å resolution were collected. Two heavy atom derivatives were used for phasing by multiple isomorphous replacement (MIR). The overall fold of the structure shows an eight stranded αβ barrel and three other domains. Refinement of the structure is underway.

Key words: Chitobiase, gene sequence, X-ray structure, αβ-barrel, *Serratia marcescens*.

Introduction

Chitin is the second most abundantly distributed polysaccharide throughout nature. This homopolymer of N-acetyl-D-glucosamine is not only the major constituent of the fungal cell wall and the arthropod exosceleton but also an important nutrient source of carbon and nitrogen in the marine environment (4).

Our interest is the elucidation of the mechanism of chitin degradation by chitinolytic enzymes. These enzymes are produced and secreted from chitinolytic bacteria and are Chitinases (EC. 3. 2. 1. 14) and Chitobiases (EC. 3. 2. 1.30). Chitinases have been classified into families 18 and 19 of glycosyl hydrolases (5). They hydrolyze chitin to oligosaccharides of which N, N'-diacetyl-glucosamine is the predominant product. N, N'-diacetyl-glucosamine is the substrate for Chitobiase (trivial name for N-acetyl-glucosaminidase) which is classified into family 20 of glycosyl hydrolases (5).

The application of crystallographic techniques to investigate the molecular structure of these enzymes will provide us with valuable information in order to elucidate the mechanism of chitin degradation. The X-ray structure of Chitinase A is presented in another paper in this volume. Crystallographic data and the overall folding of the preliminary model of Chitobiase are described here.

Cloning of the Chitobiase gene

The first identification of a genomic fragment containing Chitobiase activity and the construction of a Chitobiase producing plasmid is described in (1). The recombinant bacterial strain was kindly provided by A. Oppenheim (Hebrew University, Jerusalem, Israel) and Ilan Chet (Hebrew University, Rehovot, Israel). The gene was mapped by restriction analysis.

Nucleotide sequencing of Chitobiase gene

Two genomic clones of *Serratia marcescens* were initially used for sequencing. DNA fragments of 3.7 kb and 2.8 kb have been subcloned into pEMBL18 (1). Each of these fragments contains a part of the chitobiase gene. The clones were named pCB13 and pCB15, as shown in Figure I. Further fragmentation of the gene and subcloning into pUC vectors allow us fast access to several regions of the Chitobiase gene using universal primers. DNA sequencing was carried out using also synthetic primers. GC-rich regions were sequenced several times using

additional primers. The entire Chitobiase gene was double stranded sequenced using the Sequenase® kit from USB according the instructions of the manufacturers.

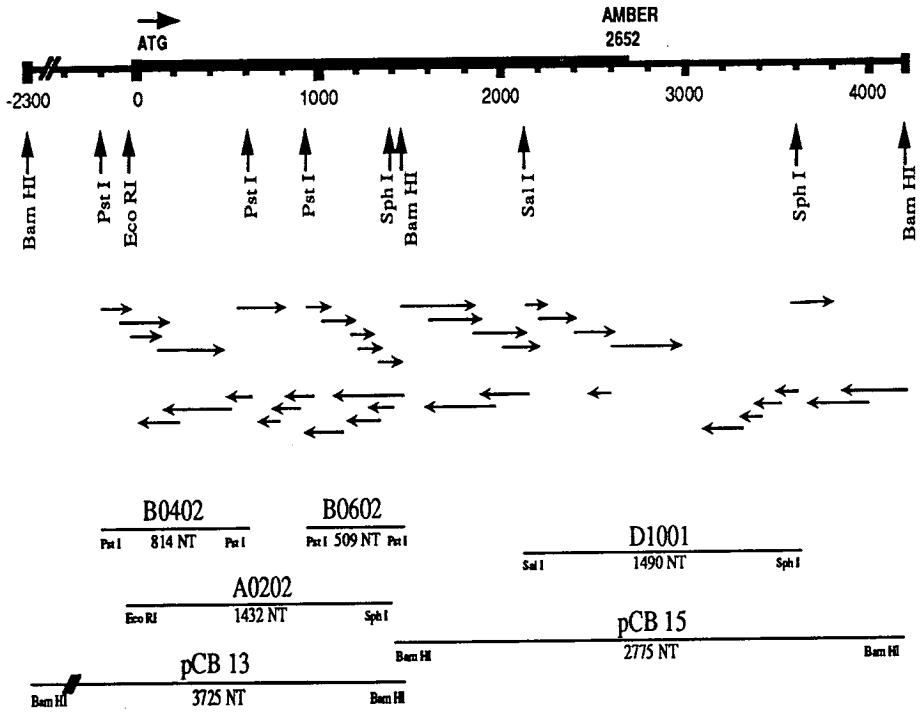

Figure 1. Restriction map, subcloning and sequencing strategy of the *Serratia marcescens* Chitobiase gene

The first twenty amino acids of the purified Chitobiase were determined by protein microsequencing carried out by Dr. R. Kellner (EMBL Heidelberg). Comparison of this sequence with the deduced amino acid sequence from the gene, led to the conclusion that the mature chitobiase was processed *in vivo* since the 27 amino acid signal peptide was cleaved. The cleavage occured between Ala27 and Asp28 during transport of the protein througth the cell membrane into the periplasm of *E. coli*. Figure 2 below shows the amino acid sequence of the leader peptide.

M N A F K **L S A L A R L T A T M G F L G G M** G S A M A | DQQLV

Figure 2. Amino acid sequence of the signal peptide of Chitobiase showing the positively charged region (underlined) and the hydrophobic region (bold). The vertical bar shows the cleavage site.

Comparison of Chitobiase to other carbohydrate hydrolases

Serratia marcescens Chitobiase shows a 54% sequence similarity to *Vibrio harveyi* Chitobiase (2) and 42 % to *Vibrio vulnificus* β-N-acetylhexosaminidase (6). The latter enzyme has hydrolytic activity for chitotriose, chitobiose, N-acetyl-glucosamine and N-acetyl-galactosamine. Sequence comparisons were carried out using the program ClustalW (7).

The amino acid sequence of these three prokaryotic enzymes align with five eukaryotic oligomeric hexosaminidases that hydrolyze G_{M2} gangliosides. The above mentioned enzymes are classified into family 20 of glycosyl hydrolases (5).

Expression and purification of chitobiase

The two step purification procedure of the recombinant protein from the bacterial periplasm (3) has been slightly modifed as summarized here: cells of the *E. coli* clone expressing Chitobiase were fractionated into periplasm and cytoplasm. The 55% to 85% saturated ammonium sulfate precipitate of the periplasmic protein fraction was applied on a S-Sepharose colum at pH 6.4. The chromatographic peak showing Chitobiase activity was analysed by SDS-PAGE. Finally, fractions containing Chitobiase were combined and applied on a Mono-S FPLC column at pH 6.2. The protein was concentrated to 30 mg/ml using the last column.

Crystallization of chitobiase

The first crystallization of Chitobiase into monoclinic crystals was achieved in PEG as previously described (8). In the meantime crystallisation experiments using various conditions were carried out. The best diffracting crystals have an orthorhombic form and were obtained from ammonium sulphate using hanging drop vapour diffusion.

The orthorhombic form can be grown to a size of up to 0.8 x 0.6 x 0.15 mm^3 (although slightly smaller crystals diffracted better) and used for data collection and structure analysis. Table I summarizes the two crystallization conditions and the physical parameters of both crystal forms.

Crystal form	monoclinic	orthorhombic
Space group	$P2_1$	$P2_12_12$
Unit cell size	62.2 * 133.2 * 55.1 Å, 101.0 °	110.7 * 99.9 * 87.7 Å
Precipitant	16.6 % PEG 4000	61.5 % Ammonium sulphate
Additives	7.5 % Isopropanol	1.5 % Isopropanol
Buffer	100 mM Cacodylate pH 5.6	100 mM Cacodylate pH 5.6
Vm	2.4 $Å^3/d$	2.4 $Å^3/d$
Solvent content	45 %	43 %
Resolution limit	2.7 Å	1.9 Å

Table I: Crystallization conditions and physical parameters of the two different crystal forms of Chitobiase.

Data collection and phase determination

X-ray diffraction data were collected to 1.9 Å resolution using synchrotron radiation at the EMBL beamline X31 (wavelength 1.0 Å) at HASYLAB, DESY, Hamburg. The data were 99.9 % complete to 1.9 Å with a R_{merge} of 6.2 %.

Resolution limits	10 Å to 1.9 Å
Number of unique reflections	82,900
Reflections missing	67
R_{merge}	6.2
R_{merge} in high resolution bin (1.88 Å to 1.92 Å)	23.8 %
Percentage reflections greater than 3σ	85.3 %
Percentage reflections greater than 3 σ (1.88 Å to 1.92 Å)	64.2 %
Reflections missing (1.88 Å to 1.92 Å)	5

Table II: Statistics of the native data collected of the orthorhombic form of Chitobiase crystals.

The 3D structure of Chitobiase was solved by multiple isomorphous replacement (MIR). One heavy atom derivative was prepared by soaking the crystals in mother liquor containing 5 mM K_2AuCl_4. Within thirty minutes the crystals turned red and after 20 hours dark purple when the data collection began (95 % completeness, 2 Å resolution, 6.2 % R_{merge}). Derivatized and native crystals were isomorphous (R_{iso} of 10.1 %, 8 Å to 2.5 Å). Three sites on Chitobiase bound Au according to the Patterson map. Single Isomorphous Replacement (SIR) with this derivative gave a mean figure of merit (FOM) of 0.43.

A second heavy atom derivative was necessary to solve the structure. Platinum was a potential candidate. However the Chitobiase crystals have different unit cell dimensions when platinum salts were added to the mother liquor. Finally Platinum(II)-(2,2'-, 6',2"-) tetrapyridine-chloride was found to keep the Chitobiase crystals isomorphous to native. This was obtained by soaking of the crystal for 50 hours in mother liquor to a concentration between 10 and 20 mM of this compound. Two datasets of the derivatized Chitobiase crystals were collected as summarized in Table III. Only the isomorphous signal was used in the low resolution dataset (20 Å to 4.8 Å). The other dataset, although collected to 2.2 Å, has useful phasing power only to 4 Å for the isomorphous and to 3.5 Å for the anomalous signal.

		eff. Res. (Å)	R_{merge} (%)	R_{iso} (%)	Phpower	$Phpower_{ano}$
1	K_2AuCl_4	2.5	6.2	10.1	2.4	2.2
2	< Platinum(II)-(2,2´-6´2"-	4.8	6.9	14.5	1.7	not used
	Tetrapyridine)-chloride >	4.0 (3.5)	6.8	17.7	1.5	1.5

Table III Phasing statistics of two different heavy atom derivatives of the Chitobiase crystals.

Heavy atom refinement, map calculations and density modification were carried out using the PHASES package by W. Furey (9). A cutoff of six sigma was applied to all data. A resolution cutoff was applied where the phasing power of the derivatives dropped to below 1.1. The phase set to 3 Å had an overall figure of merit of 0.74, at 2.5 Å the overall figure of merit was 0.69. Maps based on both phase sets were calculated and subjected to solvent flattening.

Model building and initial refinement of chitobiase

The program ´O´ was used for model building (10). The 3 Å map was skeletonized. It was possible to trace the complete chain in the 2.5 Å MIR map. The model presented here comprises 92 % of the side chains, all visible in the MIR density. The rest of the side chains were set to serine to allow the refinement programs to refine the Cβ position. The geometry of the protein backbone was regularized before the first refinement was started.

Refinement was carried out using ARP (11). ARP was run with restrained Hendrickson-Konnert refinement (12). In every cycle ARP analyses the $3F_{obs}$-$2F_{calc}$ density map to remove poorly defined atoms and inspects the F_{obs}-F_{calc} density

maps to add water molecules. The final model after 27 cycles contained 776 waters, the crystallographic R_{factor} dropped from 49.5 % to 21.8 %. Ten percent of the data were removed from the refinement set to monitor refinement using R_{free} (13). R_{free} converged from 49.7 % to 23.2 %. Insertion of the remaining amino acids and correction of the model are underway.

The fold of chitobiase

Chitobiase has an eight stranded αβ-barrel structure (= domain III, see figure 3) with three additional domains.

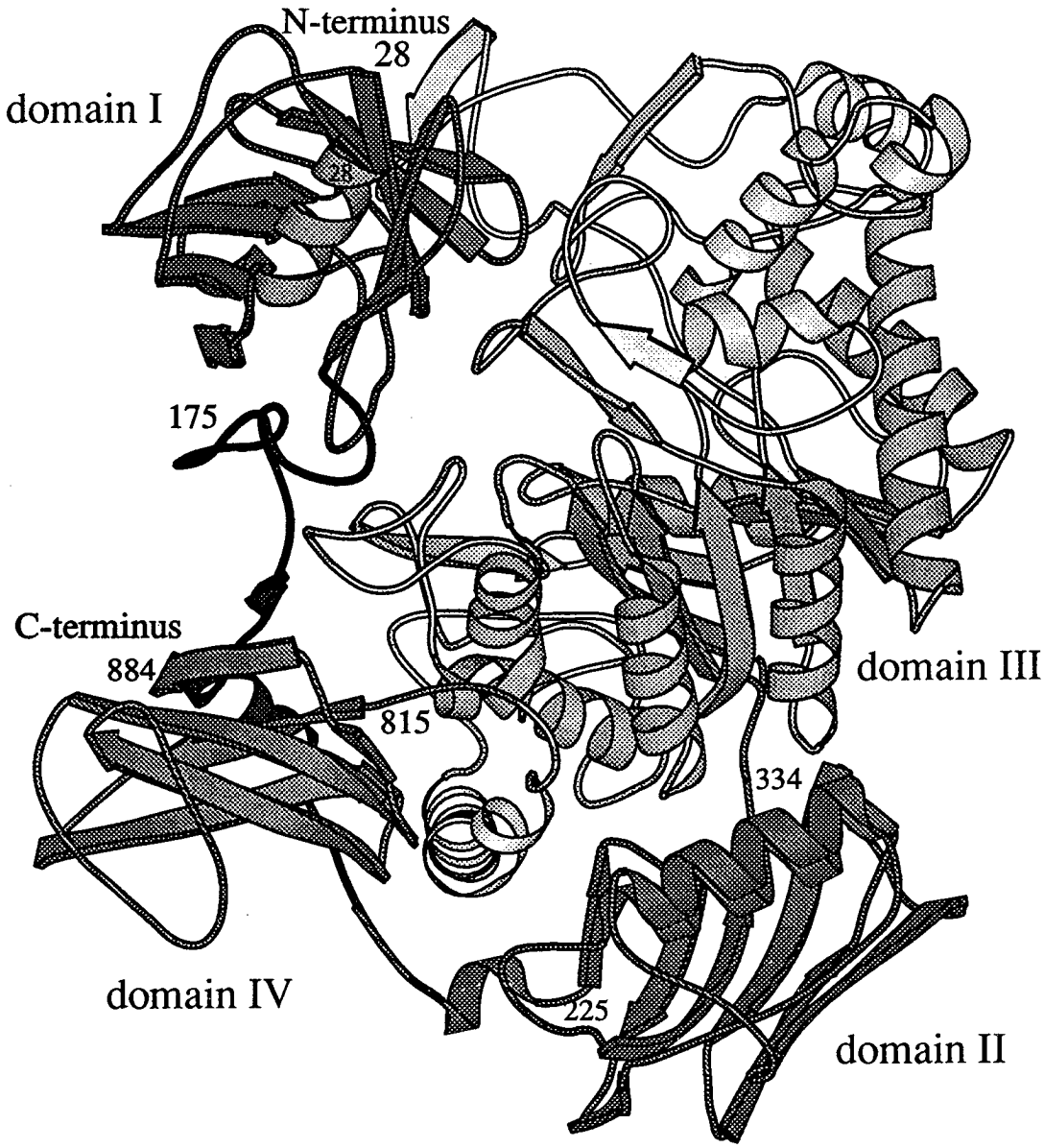

Figure 3. Schematic diagram illustrating the structure of Chitobiase.

Domain I comprises residues 28 to 175. This the N-terminal domain. Two β-pleated sheets wrap around a hydrophobic core. The motif starts with a three turn α-helix that points into solvent - an ideal position for processing as the leader peptide is chopped off here (upper left in figure 3, residue 28).

Domain I is connected to domain II by a fifty amino acid long linker (residues 175 to 225). It folds around the αβ-barrel (domain III) as shown in figure 3 (shaded dark). A short, two turn helix in close contact to domain IV can be seen.

Domain II (residues 225 to 334) shows two parallel helices (seen side-on in figure 3 one on top of the other). A five stranded antiparallel β-sheet faces the solvent. The β-strands are parallel to the helices.

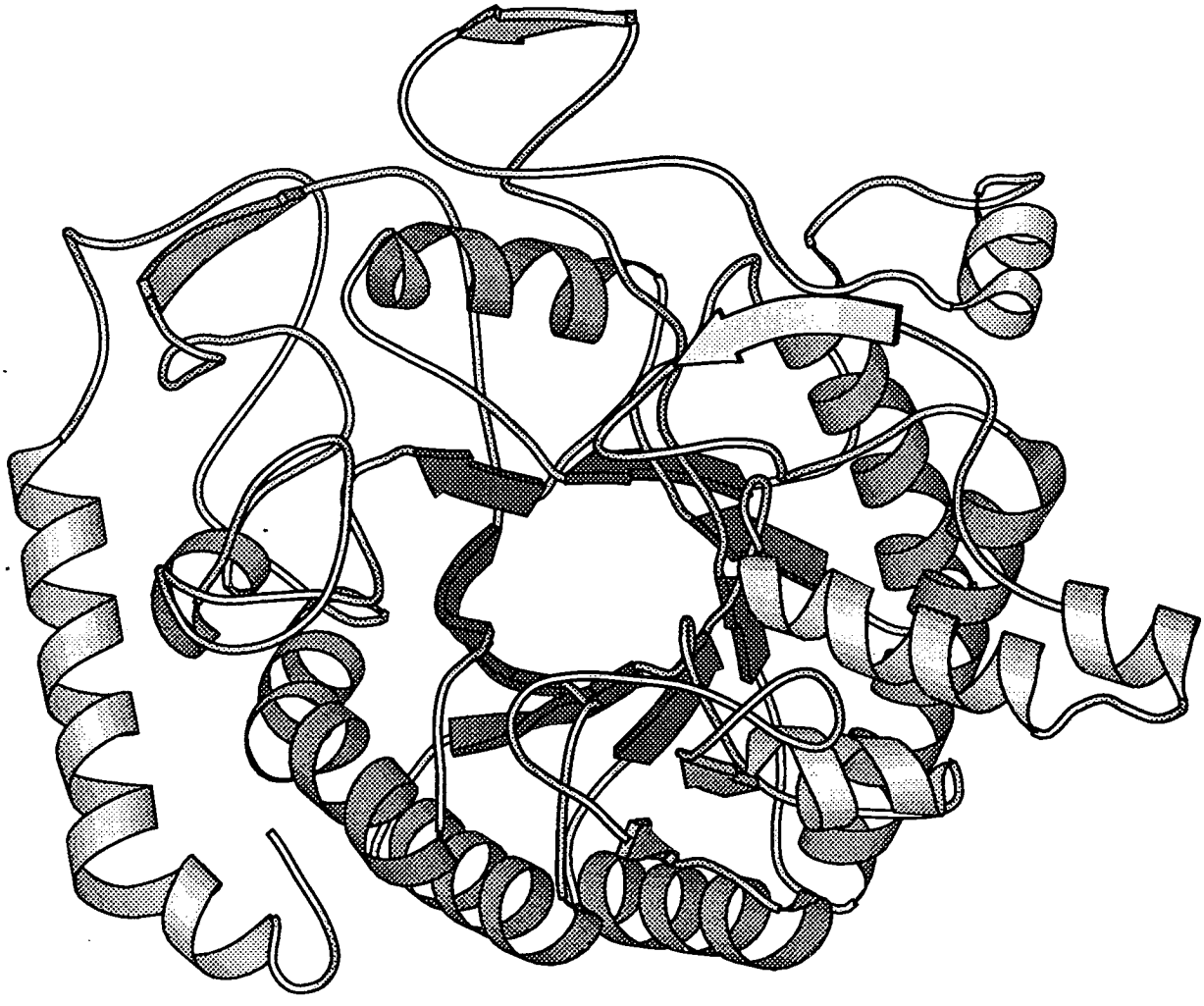

Figure 4. Schematic diagram illustrating the structure of domain III of Chitobiase.

Domain III forms an αβ-barrel motif, see figure 4. It comprises 465 amino acids (residues 340 to 815). Eight β-strands inside and seven helices on the outside were found. The eighth helix was replaced by a β-strand. Figure 3 shows the αβ-barrel side-on with the β-strands making up the core in a somewhat darker color to show the structure more clearly. The C-terminal end of the barrel faces towards domain I.

The active site -deduced from spacial organization and sequence comparisons- is likely to be at the C-terminus of the barrel. This is not surprising as all αβ-barrel enzymes determined so far had the active site at this end.

Most prominent insertions to the barrel motif are a loop towards domain I (close to the label '28' in figure 3, see also on the right hand side figure 4) and two helices pointing into solvent (lower middle in figure 4).

A 40 Å long helix expands around the barrel and finishes domain III (see upper side in figure 4, also seen viewed from the top between domains IV and II in figure 3). This helix has a kink after 4 turns where a glycine is found.

A small β-barrel at the C-terminal end forms domain IV (lower left corner in figure 3).

References

1 Kless, H., Sitrit, Y., Chet, I.and Oppenheim, A.B. (1989) Mol. Gen. Genet. **217**,471-473.

2 Somerville, C. C. and Colwell, R. R. (1993) Proc. Natl. Acad. Sci. USA **90**, 6751-6755.

3 Vorgias, C. E., Tews, I., Perrakis, A., Wilson, K. S., Oppenheim, A. B. (1993) *In* Chitin Enzymology. R.A.A. Muzzarelli (ed.) pp 417-422.

4 Yu, C., Lee, A. M., Bassler, B. L. and Roseman, S. (1991) J. Biol. Chem. **266**, 24260-24266.

5 Henrissat, B. (1991) Biochem. J. **280**, 309-316.

6 Soto-Gil, R., Zyskind, J. W. (1989) J. Biol. Chem. **264**, 14778-14783.

7 Higgins, G. D., Bleasby, J. A. and Fuchs, R. (1992) CABIOS **8**, 198-191.

8 Tews, I., Dauter, Z., Oppenheim, A. B., Vorgias, C.E. (1992) J. Mol. Biol. **228** 696-697.

9 Furey, W. N. (1990) American Crystallographic Association Series **2**, 18

10 Jones, T. A., Cowan, S. W., Zou, J.-Y., Kjeldgaard, M. (1991) Acta Cryst. **A47**, 110-119.

11 Lamzin, V. S., Wilson, K. S. (1993) Acta Cryst. **D49**, 129-147.

12 Konnert, J. H. and Hedrickson, W. A. (1980) Acta Cryst. **A36**, 344-350.

13 Bruenger, A. T, (1992) Nature **355**, 472-474.

14 Kraulis, P. (1991) J. Appl. Cryst. **24**, 946-950.