

**ISOLATION AND CHARACTERIZATION OF A TRYPSIN
INHIBITOR FROM THE SEEDS OF KOHLRABI
(BRASSICA NAPUS VAR. RAPIFERA)
BELONGING TO THE NAPIN FAMILY OF STORAGE PROTEINS**

by

IB SVENDSEN¹⁾, DIANA NICOLOVA²⁾, IVAN GOSHEV²⁾ and NICOLAY GENOV²⁾

¹⁾ Department of Chemistry, Carlsberg Laboratory
Gamle Carlsberg Vej 10, DK-2500 Copenhagen Valby

²⁾ Bulgarian Academy of Sciences, Institute of Organic Chemistry,
1113, Sofia, Bulgaria

Keywords: Circular dichroism, fluorimetry, amino acid sequence

A trypsin inhibitor with a K_m of 5×10^{-5} M has been isolated from kohlrabi (*Brassica napus* var. *rapifera*). Subtilisin DY is inhibited only weakly and chymotrypsin not at all. The inhibitor is closely related to napin as determined by amino acid sequence analysis which also showed the inhibitor to be polymorphous. The inhibitor has been further characterized by means of molecular weight determination using SDS gel-electrophoresis and by amino acid analysis, fluorimetry as well as circular dichroism. A simplified method for purification of napins is given.

1. INTRODUCTION

Over the past decade interest in the storage proteins of Brassicaceae has been increasing as regards their biological function, nutritional value and relevance to clinical research on allergens (16). Attention has also been directed to the possibilities of improving the amino acid composition of the seed proteins from a nutritional point of view by genetic engineering (6). The proteins have been analyzed especially in rape (*B. napus*) (14) which is the major oil-seed crop in many countries of the world. Its major basic low-molecular weight storage protein consists of two chains linked by disulfide bridges (14). Isolation and sequencing of napin cDNA clones revealed the protein to be synthesized as a precursor containing the polypeptide chains of both subunits (5). The primary structure has

also been determined directly on isolated napin (6). The aim of the present investigation was to isolate a trypsin inhibitor from the Brassicaceae, to characterize it physico-chemically and to determine its primary amino acid sequence. As the source was chosen seeds from kohlrabi (*Brassica napus* var. *rapifera*). The inhibitor is shown by amino acid sequencing to belong to the napin family of storage proteins and is the first reported example of a napin with protease inhibitory activity.

2. MATERIALS AND METHODS

Seeds of *Brassica napus rapifera* were purchased from the company "Sortovy semena" (Sofia, Bulgaria). Bovine trypsin (three times crystallized) and chymotrypsin were products of

Abbreviations: SDS = sodium dodecyl sulfate; TIBN = trypsin inhibitor from *Brassica napus* var. *rapifera*.

Fluka AG from Basel, Switzerland, clostripain from Sigma, USA, *S. aureus* V8 protease from ICN, USA and Endolys C protease from Boehringer, Germany.

Subtilisin DY was purified and isolated in a homogenous state as described by GENOV et al. (9). Sephadex G-25 and G-50 were obtained from Pharmacia (Uppsala, Sweden). Reagents used in sequencing were from Applied BioSystems.

2.1. Purification method

100 g of *Brassica napus* var. *rapifera* seeds were ground and extracted by stirring for 1 hour and 30 min. at room temperature with 1 liter of distilled water. The extract was centrifuged at 4,500 rpm for 15 min to obtain a clear supernatant. The pellet was reextracted with 500 ml of distilled water. The supernatants were combined and adjusted to pH 5.0 with 5 N HCl. After 12 hours at 4 °C the precipitated proteins were removed by centrifugation at 4,500 rpm and discarded. The supernatant was heated 5 min at 70 °C and cooled to room temperature. The mixture was centrifuged and the sediment containing thermolabile proteins discarded. The supernatant which had an inhibitory activity towards trypsin was lyophilized, the obtained powder dissolved in a minimal volume of 0.05 M ammonium acetate buffer pH 5.0 and gel filtrated through a column of Sephadex G-50 (1.5 × 80 cm) equilibrated and eluted with the same buffer at a flow rate 15 ml/h.

2.2. Determination of the inhibitory activity towards proteolytic enzymes

To a 2×10^{-6} M solution of the respective proteinase (trypsin, chymotrypsin or subtilisin DY) at pH 8.0 aliquots containing identical amounts of the inhibitor were added. The degree of inhibition was determined from the residual proteinase activity towards casein and expressed as percentage of inhibition relative to the activity in the control.

The inhibitory constant K_i of the highly purified trypsin inhibitor was determined as described in (7). Kinetic experiments with the homogeneous TIBN, trypsin and the substrate N^{α} -

benzoyl-L-arginine 4-nitroanilide were performed at 25 °C in 0.5 M Tris-HCl buffer, pH 8.2, containing 0.02 M $CaCl_2$. A recording spectrophotometer Shimadzu model 3000 was used. The concentration of the substrate varied between 1.5×10^{-4} and 5.2×10^{-4} mol l^{-1} and the enzyme concentration was 2.7×10^{-7} mol l^{-1} .

2.3. Polyacrylamide gel electrophoresis

Disc electrophoresis in 6 × 10 mm tubes was carried out according to LAEMMLI (12) using a 10% polyacrylamide gel. The electrophoresis was done at 3 mA per tube for 3 h and 30 min. For the molecular weight determination a kit of standard proteins Nr. MW-SDS-707 obtained from Sigma Chemical Co., St. Louis, MO, USA, was used.

2.4. Isoelectric focusing

Isoelectric focusing was performed by using an LKB-instrument model 2117 and power supply LKB 2103. A pH gradient from 3 to 10 was made using Servalyt. The voltage was kept at 1.1 kV.

2.5. Amino acid analyses

The amino acid composition of TIBN was calculated by using a BIOTRONIK automatic analyzer model 6001 after 24 h, 48 h and 72 h of hydrolysis in 6 M HCl in evacuated sealed tubes at 110 °C. Tryptophan content was estimated spectrophotometrically as described in (3).

2.6. Amino acid sequencing

Amino acid sequence analysis was performed using an Applied BioSystems gas phase sequencer model 470A equipped with an on-line HPLC model 120A. Reduction and vinylpyridination, CNBr cleavage, digestion with clostripain, *S. aureus* V8 protease, and EndoLysC protease was done as described previously (1). Separation of peptides was made either by gel chromatography on Biogel P6, Sephadex G-25 sf and G-50 sf in 30% acetic acid and by HPLC on a Vydac C_{18} column using a linear gradient of acetonitrile from 10% to 60% for elution.

2.7. Molar extinction coefficient determination

The molar extinction coefficient at 280 nm ($\epsilon_{M,280}$) of TIBN was determined as follows: the absorption of protein solution in 0.1 M phosphate buffer containing 0.15 M NaCl, pH 7.4, was measured at 280 nm. The protein concentration was determined by amino acid analysis assuming three moles of phenylalanine per mole TIBN. The obtained value was used for the calculation of $\epsilon_{M,280}$.

2.8. Reduction of TIBN

Reduction of TIBN was done according to the method of LÖNNERDAL and JANSON (14) or as described in (1).

2.9. Spectroscopic measurements

Absorption spectra were recorded with a Shimadzu model 3000 spectrophotometer.

Circular dichroism was measured with a Rousel Jouan Dichrographe III instrument. The spectra were recorded between 320 and 200 nm using cuvettes of 1 or 20 mm pathlength. The data were expressed in terms of mean residue ellipticity:

$$[\theta] = \frac{\nu}{10} \times \frac{M_r}{1 \cdot c} \text{ (degree} \cdot \text{cm}^2 \cdot \text{dmol}^{-1}\text{)}$$

where ν is the ellipticity measured in degrees, M_r is the molecular weight per mean residue, l is the optical path of the solution in cm and c is the protein concentration in g/ml. The α -helix content of the protein was calculated by the procedure of CHEN et al. (4) and the amount of the β -structure was determined according to (18).

Fluorescence measurements were performed with a Perkin-Elmer model LS 5 spectrofluorimeter equipped with a Data Station model 3600. The optical absorbance of the solutions was lower than 0.05 at the excitation wavelength to avoid inner filter effects. All measurements were performed at pH 7 and at 20 °C.

3. RESULTS

Crude extract prepared from the seeds of kohlrabi was tested for inhibitory activity toward trypsin, subtilisin DY and chymotrypsin. Equal amounts of the extract caused 100% inhibition of a 2×10^{-6} M solution of trypsin and 40% inhibition of the solution containing subtilisin with the same concentration. No inhibitory activity against chymotrypsin was detected. Heating of the extract at 70 °C for 25 min did not change its ability to inhibit trypsin.

Purification of the component possessing trypsin inhibitory activity was achieved with a procedure involving double precipitation of the concomitant proteins after adjustment of the extract to pH 5 and heating of the supernatant at 70 °C for 5 min followed by gel filtration through a column of Sephadex G-50. The elution was monitored by absorbance at 280 nm. In the elution profile shown in Figure 1 the first protein peak

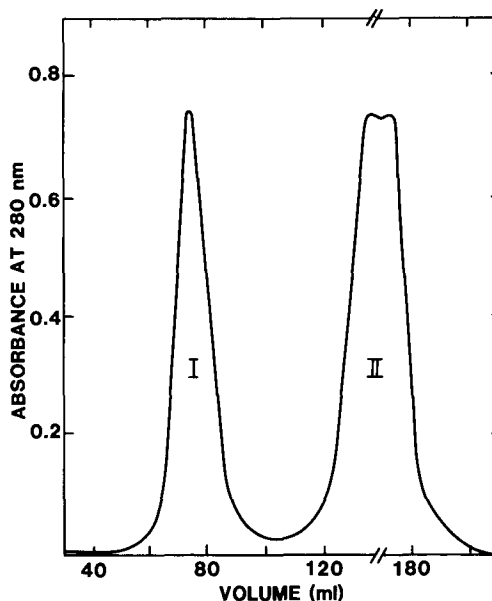


Figure 1. Final purification step of trypsin inhibitor from *Brassica napus* var. *rapifera* (TIBN). Gel filtration through a column of Sephadex G-50 (1.5 × 80 cm) equilibrated and eluted with a 0.05 M ammonium acetate buffer pH 5.0 at a flow rate of 15 ml/h.

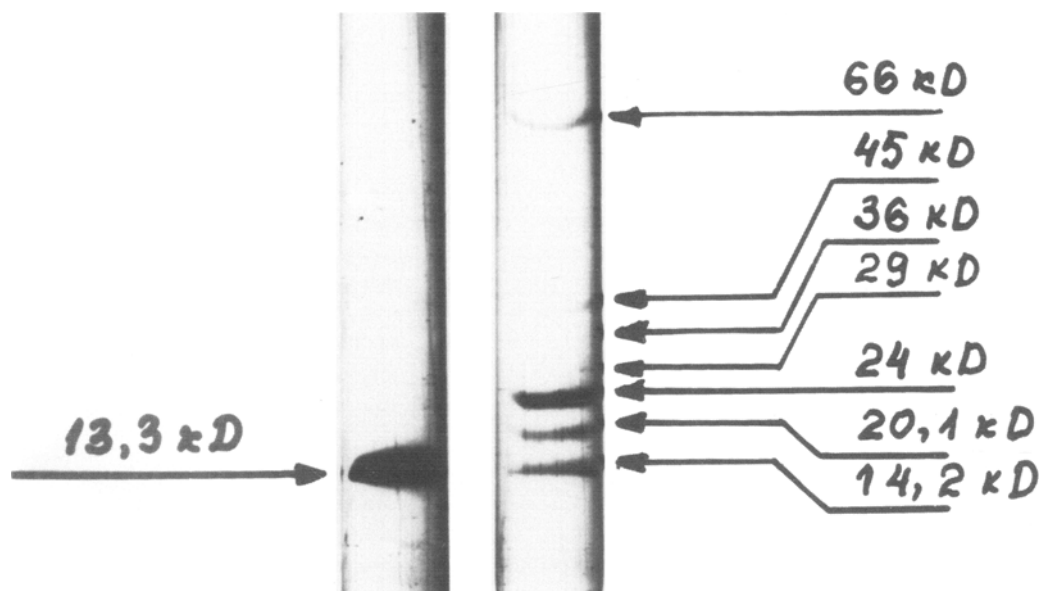


Figure 2. SDS-polyacrylamide gel electrophoresis of TIBN.

Table 1. Comparison of the amino acid composition of napins from the family Brassicaceae.

Amino acid	TIBN acid hydrolysis	TIBN sequence	<i>Sina 1</i> ^{a)}	pN 1 ^{b)}	pN 2 ^{b)}
Asx	3.8 (4)	2	2	4	3
Thr	4.5 (4-5)	4	4	5	5
Ser	5.5 (6)	4-5	7	3	4
Glx	25.4 (26)	30	28	31	29
Pro	12.6 (13)	15	16	10	11
Gly	9.3 (9)	8-9	8-9	6	8
Ala	7.5 (7-8)	8	7	6	7
Cys	6		8	7	8
Val	6.5 (6-7)	5-6	7	6	6
Met	3.2 (3)	3	2	3	3
Ile	4.4 (4)	3-4	4	4	4
Leu	7.3 (7)	8	9	7	8
Tyr	1.5 (2)	1	1	2	1
Phe	3.0 (3)	3	3	2	3
His	4.3 (4)	4	5	4	4
Lys	7.5 (7-8)	7-8	9	9	9
Arg	5.7 (6)	6-7	5-6	4	5
Trp ^{c)}	1	1	1	1	1
Total	117-122		127	113	119

^{a)} Data from (16); ^{b)} Data from (5); ^{c)} Determined spectrophotometrically according to (3). The variations in the amino acid compositions of the sequenced napins are due to gene polymorphisms.

contains TIBN. No inhibitory activity against trypsin, subtilisin DY or chymotrypsin was observed in the second peak. The inhibitor gave a single band on SDS-PAGE (Fig. 2) with a molecular weight of 13300 ± 1000 which is in good agreement with that found for other low molecular weight storage proteins from the family Brassicaceae (6, 14, 16). In the isoelectric focusing experiment the native protein appeared as a single peak with an isoelectric point of 10.1. However, focusing of the reduced TIBN has shown the sample to contain three components of pI 9.1, 8.4 and 10.1. It is suggested that the first two components represent the light and heavy chains of the inhibitor while the third one is unreduced TIBN. Results of gel electrophoresis and isoelectric focusing suggested the homogeneity of the purified protein.

Kinetic analysis of the interaction between TIBN and trypsin using N^α -benzoyl-L-arginine 4-nitroanilide as a substrate showed that the protein from the seeds of *Brassica napus* var. *rapifera* is a reversible competitive inhibitor of this proteinase. An inhibitory constant K_i of 5×10^{-5} M was calculated.

Table I shows the similarity between the amino acid compositions of TIBN and other storage proteins from the Brassica family. The inhibitor, like the other proteins included in the table is high in glutamic acid + glutamine (22% of the total content of amino acid residues) and proline (11%). The molecular weight of 13,500 calculated from the amino acid sequence of TIBN is in a very good agreement with that obtained from SDS-gel electrophoresis.

At pH 7.0 TIBN has an absorption spectrum with a minimum at 255 nm, maximum at 280 nm, shoulder at 289 nm and no visible absorption. The ultraviolet absorption spectrum is in accordance with the presence of aromatic residues in this protein. The molar extinction coefficient at 280 nm was calculated to be $7670 \text{ M}^{-1}\text{cm}^{-1}$. On the basis of the content of aromatic amino acids, the expected molar absorbance at this wavelength should be $6930 \text{ M}^{-1}\text{cm}^{-1}$. This value was calculated by using $\epsilon_{\text{M},280} = 5600 \text{ M}^{-1}\text{cm}^{-1}$ for tryptophan and $\epsilon_{\text{M},280} = 1330 \text{ M}^{-1}\text{cm}^{-1}$ for tyrosine (3).

Fluorescence emission spectra of TIBN are shown in Figure 3. The native protein contains

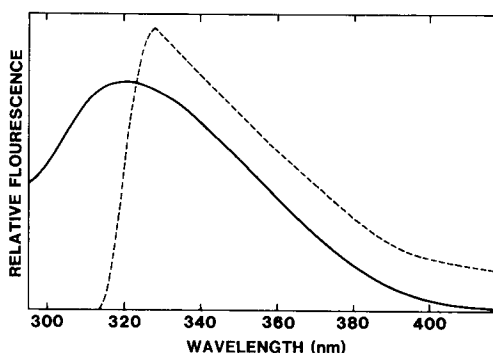


Figure 3. Fluorescence emission spectra of TIBN excited at 280 nm (—) and 295 nm (----).

two types of fluorophores: the side chains of one tryptophyl and one tyrosyl residue. Irradiation with light at 280 nm leads to a simultaneous excitation of both indole and phenolic groups. The emission spectrum upon excitation at 280 nm exhibits a maximum at 320 ± 1 nm and essentially retains the features of the tryptophan emission. The fluorescence maximum position of the tyrosyl residue is at 303-305 nm and does not depend on the microenvironment. However, the 280 nm-excited fluorescence spectrum of native TIBN shows a predominant contribution of the indole fluorophore. The possible explanation is an energy transfer at the singlet level from tyrosine to tryptophan which is a rather efficient process. This results in the quenching of a significant fraction of the tyrosine fluorescence. Some contribution of the disulfide bridges to the quenching process is also possible because it is known that these bridges are strong quenchers of the phenolic emission.

Native TIBN shows a peak with an emission maximum position at 329 ± 1 nm after excitation at 295 nm (Fig. 3). At this wavelength at least 93% of the incident light is absorbed by the indole group of the inhibitor. The wavelength of the emission maximum is typical of a tryptophyl residue in a hydrophobic environment. It can be concluded that the side chain of the single tryptophan of TIBN is deeply buried in the nonpolar interior of the molecule.

At pH 6.5 the circular dichroism spectrum of TIBN in the far ultraviolet spectral range was

amino acids in positions 49, 52, 53, 65, 70 and 74 in the correct chain. The C-terminal EndoLysC peptide appeared on HPLC as 3 peaks. Sequence analyses and amino acid analyses revealed the polymorphism in position 85 as well as the lack of an amino acid in this position.

4. DISCUSSION

The original intention of the present investigation was to isolate a trypsin inhibitor from kohlrabi. The amino acid composition (Table I) of the purified inhibitor with its high content of Glx and Pro pointed towards similarity to the napin storage proteins. This was verified when the amino acid sequence (Fig. 5) was completed. A search in the amino acid data base of NBRF-PIR disclosed that a protein belonging to the napin family of storage proteins had been isolated.

The amino acid sequences of the water-soluble storage protein napin from rape seeds (*Brassica napus* var. *oliefera*) has been obtained directly (6) and by deduction from nucleotide

sequences of genomic DNA (2, 10, 17) and cDNA (5). These sequence data shows polymorphism which also is apparent from the present study. Recently the primary structure of the major allergen of yellow mustard (*Sinapis alba* L.) another member of the Brassicaceae has been reported and shown to be a napin (16).

The napins seem to be widespread within the Brassicaceae, since they have been detected in radish (*Raphanus sativus*) (13) although not yet sequenced. A surprising phylogenetic relationship has been reported between the napins and a storage protein in the unrelated ostrich fern (*Matteuccia struthiopteris* L.) (19). No sequence data has yet been published for this storage protein.

A comparison of the amino acid sequence of TIBN to all known sequences among the members of the napin family is presented in Figure 6. Most of the sequences were deduced from nucleotide sequencing and thus also contain the posttranslationally processed peptide between the light and heavy chain. These linker peptides

Light chain:

	1	10	20	30
1a, b	P A G P F R I P K C R K E F Q Q A Q H L R A C Q Q W L H K Q A M Q S G G G P S			
2a				
2b			K	
3	S			
4a				S
4b	G			S
5			K	S
6				P S G

Heavy chain:

	1	10	20	30	40
1a	P Q G P Q Q R P P L L Q Q C C N E L H Q E E P L C V C P T L K G A A K A V K Q Q Q Q Q Q				
1b					V R
2a				S	I
2b	Q	H Y	Q	R	S
3				S	I
4a, b				S	V R
5				S	V R
6	Q	P P		S	R V R

	50	60	70	80	90
1a	G Q Q Q M • • • • • Q Q V P S R I Y Q T A T H L P K V C N I P Q V S V C P F K K T M P G P S				
1b	Q	G • • • • • M V			G
2a		G • • • • • K			Y
2b	E	G • • • • • K		Q	Y
3		G • • • • • K		Q	Y
4a, b	L E	G • • Q Q G P H V I		Q	
5	Q G	G • • Q Q L V I		Q	Y
6	Q G	M Q G Q Q M V I V	R R	I Q	G F Y

Figure 6. Comparison of the amino acid sequences of the light and heavy chains of the trypsin inhibitor of *Brassica napus* var. *rapifera* (TIBN) with those of different napins.

1a, 1b: TIBN. 2a, 2b: Napins N1 and N2 from rape seeds (5). 3: Napin A (nap A) from rape seeds (10). 4a, 4b: Allergen of yellow mustard, *Sma* I (16). 5: Napin 1 (BngNAP1) from rape seeds (2). 6: Napin (gNa) from rape seed (17). Amino acids at variance with TIBN in line 1a are given for the other sequences. • denotes deletions.

have been excluded from the comparison which thus is addressed only to the sequences of the mature proteins.

Polymorphisms occur in both chains, but two features are immediately apparent: Firstly, in most cases only two alternative amino acids are found in a particular position which may indicate the presence of alleles in the same gene. It was shown in 1935 by NAGAHURU U from Japan (20) that *Brassica napus* ($2n=38$) is an amphidiploid of *Brassica oleracea* ($2n=18$) and *Brassica campestris* ($2n=20$). It is tempting to speculate, that the polymorphisms primarily are the result of the ancient combination of the two parent genomes. Secondly, the section of the heavy chain from Val41 to Ile58 shows greater variation at the individual positions as well as variation in deletions ranging from 2 to 5 residues (fig. 6). These observations indicate the presence of several genes in each parental genome which may have arisen by gene duplication and subsequent mutational evolution. In agreement therewith napins have been shown to be encoded by a multigenic family ranging from 10 (17) to 16 (10) copies per genome. Although no three dimensional structure of the napins has yet been reported, it is evident from the C.D., fluorimetric and calculated data, that the molecule is well structured. The domain tolerating extensive deletions and mutations may signify a loop protruding (cf. 15) from a more compact structure.

Sequence comparisons among storage proteins from cereals (including a trypsin inhibitor from barley), castor bean, and *Brassica* (napin) have shown relationships indicating a common phylogenetic origin (11). No protease inhibitory function has hitherto been reported for other members of the napin family, probably because this was not suspected for a storage protein. The present investigation clearly demonstrates that the napin we have isolated from kohlrabi has an inhibitory effect towards trypsin, but none towards chymotrypsin.

The inhibitory site of TIBN has not been determined. A sequence comparison with barley trypsin inhibitor for which the inhibitory site has been determined reveals a faint similarity which may indicate that Arg42 of the heavy chain in TIBN is at the inhibitory site. This will also be in

agreement with the existence of the loop postulated above since it in several cases have been shown that the inhibitory site resides in a loop, see for example ref. (15).

ACKNOWLEDGEMENT

The skilled technical assistance of Mss. BODIL CORNELIUSSEN, LONE SØRENSEN and PIA BREDDAM is gratefully acknowledged.

REFERENCES

1. BALLANCE, G.M. & I. SVENDSEN: Purification and amino acid sequence determination of an endo-1,3- β -glucanase from barley. *Carlsberg Res. Commun.* 53, 411-420 (1988)
2. BASZCZYNSKI, C.L. & L. FALLIS: Isolation and nucleotide sequence of a genomic clone encoding a new *Brassica napus* gene. *Plant Mol. Biol.* 14, 633-635 (1990)
3. BEAVEN, G. & E. HOLIDAY: Ultraviolet absorption spectra of proteins and amino acids. *Adv. Protein Chem.* 7, 319-382 (1952)
4. CHEN, Y., J. YANG & H. MARTINEZ: Determination of the secondary structures of proteins by circular dichroism and optical rotatory dispersion. *Biochemistry* 11, p. 4120 (1972)
5. CROUCH, M., K. TENBARGE, A. SIMON & R. FERL: cDNA clones for *Brassica napus* seed storage proteins: Evidence from nucleotide sequence analysis that both subunits of napin are cleaved from a precursor polypeptide. *J. Mol. Appl. Gen.* 2, 273-283 (1983)
6. ERICSON, M., J. RODIN, M. LENMAN, K. GLIMELIUS, L.-G. JOSEFSSON & L. RASK: Structure of the rape-seeds 1.7 S storage protein, napin, and its precursor. *J. Biol. Chem.* 261, 14576-14581 (1986)
7. FRITZ, H., U. SEEMÜLLER & H. TSCHESCHE: General Review. In: *Methods of Enzymatic Analysis*. Bergmeyer, H. and Grasse, M., eds., Verlag Chemie GmbH, Weinheim, 3d edn., vol 5, *Enzymes 3: Peptidases, proteinases and their inhibitors*. pp. 74-98 (1984)
8. GARNIER, J., D.J. OSGUTHORPE & B. ROBSON: Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. *J. Mol. Biol.* 120, 97-120 (1978)
9. GENOV, N., M. SHOPOVA, R. BOTEVA, G. JORI & F. RICCHELLI: Chemical, photochemical and spectroscopic characterization of an alkaline proteinase from *Bacillus subtilis* variant DY. *Biochem. J.* 207, 193-200 (1982)

10. JOSEFSSON, L.-G., M. LENMAN, M. ERICSON & L. RASK: Structure of a gene encoding the 1.7 S storage protein, napin, from *Brassica napus*. J. Biol. Chem. 262, 12196-12201 (1987)
11. KREIS, M., B.G. FORDE, S. RAHMANN, B.J. MIFLIN & P.R. SHEWRY: Molecular evolution of the seed storage proteins of barley, rye and wheat. J. Mol. Biol. 183, 499-502 (1985)
12. LAEMMLI, U.K.: Cleavage of structural proteins during the assembly of the head of bacteriophage T₄. Nature 227, 680-685 (1970)
13. LAROCHE-RAYNAL, M. & M. DELSEMY: Identification and characterization of the mRNA for major storage proteins from radish. Eur. J. Biochem. 157, 321-327 (1986)
14. LÖNNERDAL, B. & J.-C. JANSSON: Studies on Brassica seed proteins. I. The low molecular weight proteins in rapeseeds. Isolation and characterization. Biochim. Biophys. Acta 278, 175-183 (1972)
15. MCPHALEN, C.A., I. SVENDSEN, I. JONASSEN & M.N.G. JAMES: Crystal and molecular structure of chymotrypsin inhibitor 2 from barley seeds in complex with subtilisin Novo. Proc. Natl. Acad. Sci. USA 82, 7242-7246 (1985)
16. MENENDEZ-ARIAS, L., I. MONEO, J. DOMINGUEZ & R. RODRIGUEZ: Primary structure of the major allergen of yellow mustard (*Sinapis alba* L.) seed, *Sin a* I. Eur. J. Biochem. 177, 159-166 (1988)
17. SCOFIELD, S. & M. CROUCH: Nucleotide sequence of a member of the napin storage protein family from *Brassica napus*. J. Biol. Chem. 262, 12202-12208 (1987)
18. SIEGEL, J.W. STEINMETZ & G. LONG: A computer-assisted model for estimating secondary structure from circular dichroic spectra: Comparison of animal lactate dehydrogenase. Analyt. Biochem. 104, 160-167 (1980)
19. TEMPLEMAN, T.S., D.B. STEIN & A.E. De MAGGIO: A fern spore storage protein is genetically similar to the 1.7 S seed storage protein of *Brassica napus*. Biochem. Genetics 177, 159-166 (1988)
20. U, N.: Genomic analysis in *Brassica* with special reference to the experimental formation of *B. napus* and peculiar mode of fertilization. Jap. J. Bot. 7, 389-452 (1935)