Eur Biophys J (1988) 15: 339–342

European Biophysics Journal

Structural and functional domains of cellobiohydrolase I from *Trichoderma reesei*

A small angle X-ray scattering study of the intact enzyme and its core

P. M. Abuja[1], M. Schmuck[1], I. Pilz[1]*, P. Tomme[2], M. Claeyssens[2], and H. Esterbauer[3]

[1] Institut für Physikalische Chemie, Karl-Franzens Universität Graz, Heinrichstrasse 28, A-8010 Graz, Austria
[2] Laboratorium voor Biochemie, Fakulteit Wetenschappen, Rijksuniversiteit Gent, K. L. Ledergancksstraat 35, B-9000 Gent, Belgium
[3] Institut für Biochemie, Karl-Franzens Universität Graz, Halbärthgasse 5, A-8010 Graz, Austria

Received June 10, 1987 / Accepted in revised form October 2, 1987

Abstract. Limited proteolysis (papain) of the cellobiohydrolase I (CBH I, 65 kDa) from *Trichoderma reesei* led to the seperation of two functional domains: a core protein (55 kDa) containing the active site, and a C-terminal glycopeptide (10 kDa) implicated in binding to the insoluble matrix (cellulose). The quaternary structures of the intact CBH I and its core in solution are now compared by small angle X-ray scattering (SAXS) measurements. The molecular parameters derived for the core (Rg = 2.09 nm, D_{max} = 6.5 nm) and for the intact enzyme (Rg = 4.27 nm, D_{max} = 18 nm) indicate very different shapes. The resulting models show a "tadpole"-like structure for the intact enzyme where the isotropic part coincides with the core protein and the flexible tail part should be identified with the C-terminal glycopeptide. Thus in this enzyme, functional differentiation is reflected in structural peculiarities.

Key words: Cellobiohydrolase I, core CBH I, small angle X-ray scattering, model for solution structure, functional-structural domains

1. Introduction

Cellobiohydrolase I (CBH I, 1.4-β-D-glucan cellobiohydrolase, E.C. 3.2.1.91) is the major component (up to 60%) of the cellulase complex secreted by *Trichoderma* strains. It has been isolated from several commercial and domestic enzyme preparations (e.g. by Gum and Brown 1976; Nummi et al. 1983; Esterbauer et al. 1983; Hayn and Esterbauer 1985). Purified CBH I shows a single band in SDS-PAGE, corresponding to a molecular weight of approx. 60–68 kDa (Nummi et al. 1983; Esterbauer et al. 1983; Bikhabai et al. 1985). Upon IEF-PAG however it appears as an iso-enzyme mixture (pI 3.5–3.9) (Bikhabai et al. 1985).

The purified CBH I adsorbs strongly onto insoluble cellulose (Reese 1982) and is able to hydrolyze microcrystalline (Avicel) and amorphous forms of this substrate (Nummi et al. 1983; Hayn and Esterbauer 1985; Fägerstam and Pettersson 1980) as well as soluble cellodextrins and their chromophoric glycosides (van Tilbeurgh et al. 1984a).

New structural insights have been obtained by complete sequencing of the protein (Fägerstam et al. 1984) and its gene (Shoemaker et al. 1983; Teeri et al. 1983). The carbohydrate content of the glycoprotein has been estimated to be 7%–11% approximately and most *O*-glycosylation is found at a Ser-Thr rich amino acid sequence near the C-terminus (B-block) (Fägerstam et al. 1984). The presence of two possible structural domains has been postulated (Bikhabai and Pettersson 1984).

Limited proteolysis studies (van Tilbeurgh et al. 1986) on the CBH I from *Tr. reesei* VTT-D-80133 indicated that the enzyme contains two functional domains: a C-terminal glycopeptide (10 kDa, block BA) acting as the binding site for insoluble cellulose, whereas the core protein (55 kDa) contains the active (hydrolytic) site. As deduced from SAXS, the enzyme (purified from *Tr. reesei* MCG 77) has a tadpole-like form with a large, isotropic and rigid head and a flexible and mobile tail (Schmuck et al. 1986).

In the present study these results are compared with SAXS measurements on CBH I from a different *Trichoderma* strain (VTT-D-80133) and with those of its core protein.

2. Materials and methods

2.1. Preparation of the CBH I and its core protein

CBH I was purified from the culture filtrate of the fungus *Trichoderma reesei* QM 9414 (VTT-D-80133), a

* To whom offprint requests should be sent

Abbreviations: SAXS, small angle X-ray scattering; SDS-PAGE, SDS-polyacrylamide gel electrophoresis; IEF-PAG, polyacrylamide gel isoelectric focusing; cellobiohydrolase (CBH, 1,4-β-glucan cellobio hydrolase (E.C. 3.2.1.91)); D_{max}, maximum diameter; Rg, radius of gyration.

kind gift from VTT, (Espoo, Finland) by affinity chromatography (van Tilbeurgh et al. 1984 b). The core protein was prepared as described (van Tilbeurgh et al. 1986). 25 µl of a 2.2 mg/ml papain (Boehringer Mannheim, FRG) solution in 0.2 *M* phospate buffer, 5 m*M* L-cysteine and 2 m*M* EDTA (pH 7.0) were added to a CBH I solution (10 mg in 1 ml 50 m*M* ammonium acetate, pH 5.0), and after overnight incubation (37 °C) the core protein was isolated by affinity chromatography. Small amounts of uncleaved (intact) CBH I could be removed by DEAE-Trisacryl (IBF-LKB, 2 × 10 cm) ion exchange chromatography (500 ml gradient of 5–250 m*M* ammonium acetate, pH = 5.6). The very acidic core protein (pI = 3.0) elutes at the end of the chromatographic run. The homogeneity was tested both with SDS-PAGE and IEF-PAG (van Tilbeurgh et al. 1986). Final preparations were extensively dialysed (FM-10, Amicon) against bidistilled water. The molar absorption coefficient is 73,000 M^{-1} cm^{-1} and is unchanged for the core protein (P. Tomme unpublished).

2.2. Small angle X-ray scattering

A Kratky camera (improved compact type) was used with a slit collimation system on a Philips PW1370 X-ray generator equipped with a copper tube that was operated at 50 kV and 30 mA.

A series of five concentrations from 10 to 50 mg/ml was measured for CBH I core and 4 to 23 mg/ml for the intact enzyme. Each scattering curve was recorded up to 10 times in the range h = 0.1 to 5.1 nm^{-1} ($h = 4\pi \sin \theta/\lambda$; 2θ: scattering angle, $\lambda = 0.154$ nm: wavelength of the CuK_α-Line). Measuring temperature was 4 °C. Data evaluation, desmearing and indirect Fourier transformation were done as described (Glatter and Kratky 1982; Pilz et al. 1977; Glatter 1977).

Preliminary scattering experiments on lyophilized CBH I preparations showed a considerable amount of deterioration, probably due to the formation of aggregates. Therefore, for this study we used only samples which were stored at 4 °C in bidistilled water.

3. Results and discussion

3.1. Scattering curves and molecular parameters

Figure 1 shows a comparison of the scattering curves ($I(h)$-functions) of intact CBH I and its core with the calculated error bars. The higher stability and concentration of the core solutions resulted in an enhanced accuracy of the scattering curve.

Figure 2 shows the pair distance distribution function for the two proteins. For homogenous particles

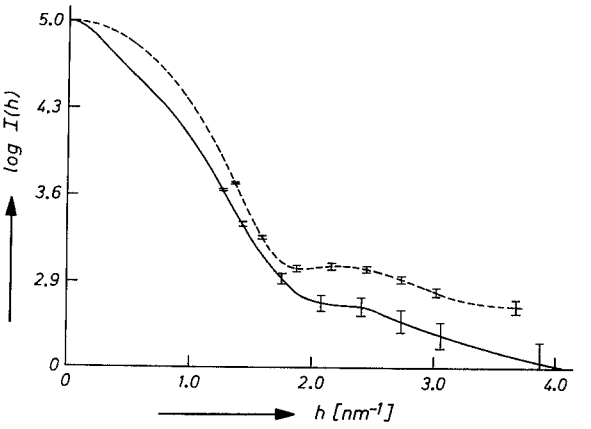

Fig. 1. Comparison of the experimental $I(h)$-functions of intact CBH I and core CBH I (with calculated error bars): $h = 4\pi \sin \theta/\lambda$, 2θ = scattering angle, $\lambda = 0.154$ nm (wavelength of the Cu K_α-line), I = scattering intensity (normalized to 10^5 at h = 0). Intact CBH I: (——); core CBH I: (– – –)

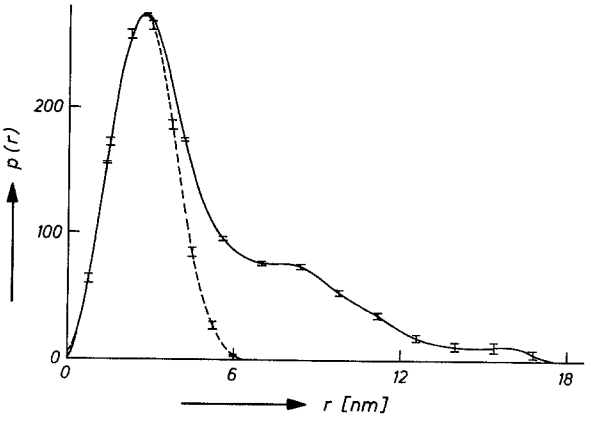

Fig. 2. Comparison of the $p(r)$-functions (electron pair distance distribution) of intact CBH I and core. $p(r)$ (in arbitrary units) is normalized to give the same maximum values for both proteins, r is the distance in real space. Intact CBH I: (——); core CBH I: (– – –)

the distance distribution function $p(r)$ (multiplied by 4π) represents the number of distances with length r found by the combination of any volume element with any other one within the particle. The $p(r)$-function becomes zero at the maximum distance of the particle.

The overall appearance of these functions clearly indicates that the core protein is a rather isotropic particle whereas the intact CBH I is composed of the core and a long tail, more or less like a tadpole. The maximum distance of the intact enzyme ($D_{max} = 18.0 \pm 0.5$ nm) exceeds that of the core ($D_{max} = 6.5 \pm 0.3$ nm) by 11.5 nm. The radius of gyration of the intact enzyme ($Rg = 4.27 \pm 0.2$ nm) is about twice as large as the value obtained for the core ($Rg = 2.09 \pm 0.2$ nm).

In addition, it should be noted that for intact CBH I preparations from two different strains, as described here (VTT-D-80133) for comparison and reported earlier (MCG 77) (Schmuck et al. 1986), exhibit the same molecular parameters.

3.2. Model calculations

Model calculations were performed with the MULTIBODY program (Glatter 1980). This permits the calculation of the $p(r)$ and $I(h)$-functions of models constructed from arbitrary spheres. These functions may then be compared to the experimental ones. Usually the fit of the $p(r)$-functions is regarded as a major criterion for the quality of the model, as the $I(h)$-function at larger angles is often influenced by minor variations in electron density which cannot be reconstructed owing to the resolutional limitations of the method.

A direct method to extract structural information from the experimental curves (e.g. Svergun et al. 1982) would have required additional assumptions about the symmetry and was therefore not used.

For CBH I (intact) we used the model already proposed by us (Schmuck et al. 1986) with some minor modifications.

In Figure 3 the experimental $p(r)$-function is compared with that of the model shown in the figure. Only a model with a tadpole-like shape gives satisfactory agreement. As the errors at higher r-values are larger than to be expected normally we have to assume that the tail of the molecule is not rigid but flexible in solution which somewhat limits the possible resolutions for the tail part of the model.

We must also point out that the mass distribution of the model should not be interpreted as an arrangement of the enzyme in subunits but as the result of the folding of a single polypeptide chain.

For CBH I core protein we obtained the best fit with an ellipsoid (axial ratio 1.5: 1.0: 1.0). Figure 4 gives a comparison of the experimental $p(r)$-function and that of the model shown in the figure. The core enzyme has the size and form of the "head" domain of intact CBH I as may be seen from Figs. 3 and 4. The most abundant electron pair distances at about 2.7 nm are the same for both core and intact enzyme, which is strong evidence that the core-enzyme and the head domain of the native enzyme have the same dimensions. This is corroborated by the similarity of the models for core enzyme and head domain of the intact enzyme. The parameters of the models are listed in Table 1.

Table 1. Molecular dimensions of CBH I and its core (data derived from the model)

Domain	Intact [nm]	Core [nm]
Head		
length	6.7	6.7
diameter	4.4	4.5
Tail		
length	12.9	–
max diameter	3.2	–

3.3. Primary structural data applied to the model

The complete amino acid sequence of CBH I has been determined (Fägerstam et al. 1984) and information on the repartition of disulfide bridges and glycosylation sites has been published (Fägerstam et al. 1984; Bikhabai and Pettersson 1984).

The papain cleavage point has been located at about residue 430 (van Tilbeurgh et al. 1986). From these results it can be deduced that the core protein

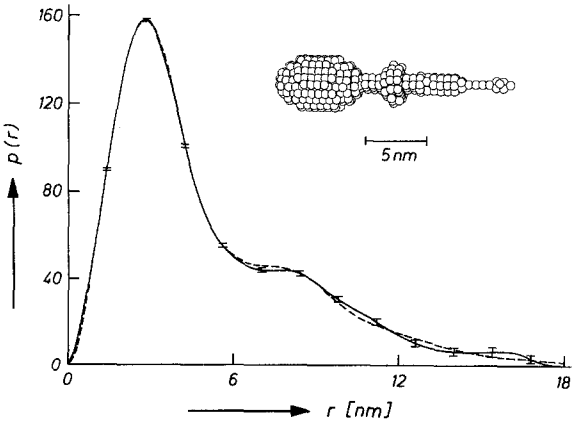

Fig. 3. Model structure and comparison of experimental and model $p(r)$-function of intact CBH I. The functions are normalized to an integral of 10^5. Experimental: (——); model: (– – –)

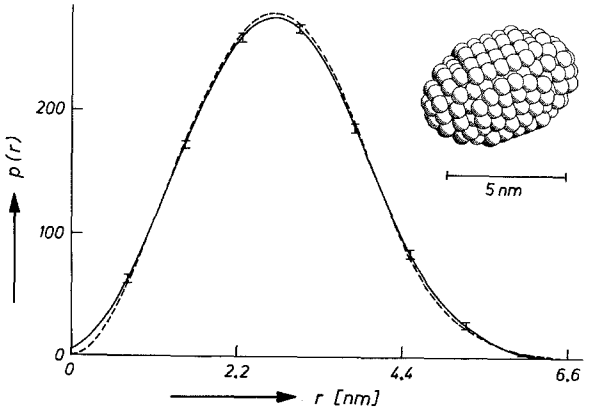

Fig. 4. Model structure and comparison of experimental and model $p(r)$-function of core CBH I. The functions are normalized to an integral of 10^5. Experimental: (——); model: (– – –)

(residues 1–429) contains three putative *N*-glycosylation sites and ten disulfide linkages. Two non-local bridges span 60 and 260 amino acid residues respectively and probably contribute to the compactness of the core protein as modelled here.

The tail part (residues 430–457) starts at a stretch of 14 amino acid residues particularly rich in proline (5 residues) and glycine (6 residues). Its repetitive GLY-ASP-PRO sequences undoubtedly represent a hinge region between the two domains (Rose et al. 1985). The following sequence (residues 445–465), rich in threonine (8 residues) and serine (4 residues), has to be identified with the heavily *O*-glycosylated block B, conserved in most other cellulases from *Trichoderma* (Teeri et al. 1987). It probably accounts for the collar-like structure in the tail part of the model. Block A, equally conserved, is represented by the last stretch of 30 residues and contains two disulfide linkages, tentatively attributed to (Bikhabai and Pettersson 1984): CYS 469-CYS 496 and CYS 480-CYS 486. It could induce some compactness or rigidity in this part of the tail domain.

Strong evidence has been obtained (van Tilbeurgh et al. 1986) for the implication of the C-terminal peptide (tail domain) in the interaction of the intact enzyme with microcrystalline cellulose. It was shown that the core protein contains the active (hydrolytic) site and accounts for the "swelling" or "structure disrupting" activity of the CBH I. The flexible, glycosylated tail domain could then serve as an "anchor", aligning the core to the scissable bonds on the insoluble substrate matrix.

Acknowledgements. The work of I. Pilz and P. M. Abuja was generously supported by the Jubiläumsfonds der Österreichischen Nationalbank. P. Tomme is an "IWONL"-bursar (Belgium). M. Claeyssens is indebted to the Belgian NPWO and NATO for grants.

References

Bhikhabai R, Pettersson G (1984) Isolation of cellulolytic enzymes from *Trichoderma reesei* QM 9414. Biochem J 222: 729–736

Bhikhabai R, Johansson G, Pettersson LG (1985) Cellobiohydrolase from *Trichoderma reesei*. Internal homology and prediction of secondary structure. Int J Appl Biochem 6: 336–345

Esterbauer H, Hayn M, Jungschaffer G, Taufratzhofer E, Schurz J (1983) Enzymatic conversion of lignocellulosic materials to sugars. J Wood Chem Technol 3: 261–287

Fägerstam LG, Pettersson LG (1980) The 1,4-β glucan cellobiohydrolase of *Trichoderma reesei* QM 9414. FEBS Lett 119: 97–100

Fägerstam LG, Petterson LG, Engström JA (1984) The primary structure of a 1,4-β glucan cellobiohydrolase from the fungus *Trichoderma reesei* QM 9414. FEBS Lett 167: 309–315

Glatter O (1977) A new method for the evaluation of small-angle scattering data. J Appl Crystallogr 10: 415–421

Glatter O (1980) Computation of distance distribution functions and scattering functions of models for small-angle scattering experiments. Acta Phys Austr 52: 234–256

Glatter O, Kratky O (eds) (1982) Small angle X-ray scattering. Academic Press, London New York

Gum EK, Brown RD Jr (1976) Structural characterization of a glycoprotein cellulase, 1,4-β glucan cellobiohydrolase from *Trichoderma viride*. Biochim Biophys Acta 446: 371–386

Hayn M, Esterbauer H (1985) Separation and partial characterization of *Trichoderma reesei* cellulase by fast chromatofocusing, J Chromatogr 329: 379–387

Nummi M, Niku-Paavola M-L, Lappalainen A, Enari T-M, Raunio V (1983) Cellobiohydrolase from *Trichoderma reesei*. Biochem J 215: 677–683

Pilz I, Glatter O, Kratky O (1977) Small-angle X-ray scattering. Methods Enzymol 61: 148–249

Reese ET (1982) Elution of cellulase from cellulose. Process Biochem 17: 2–6

Rose GD, Gierasch LM, Smith JA (1985) Turns in peptides and proteins. In: Advances in protein chemistry, vol 37. Academic Press, London New York, pp 1–97

Schmuck M, Pilz I, Hayn M, Esterbauer H (1986) Investigation of cellobiohydrolase from *Trichoderma reesei* by small-angle X-ray scattering. Biotechnol Lett 8 (6): 397–402

Shoemaker S, Schweickart V, Ladner M, Gelfand D, Kwok S, Myambo K, Innis M (1983) Molecular cloning of exo-cellobiohydrolase I derived from *Trichoderma reesei* strain L27. Bio/Technol 1: 691–695

Svergun DI, Feigin LA, Schedrin BM (1982) Small-angle scattering: direct structure analysis. Acta Crystallogr A38: 827–835

Teeri T, Salovouri I, Knowles J (1983) The molecular cloning of the major cellulase gene from *Trichoderma reesei*. Bio/Technol 1: 696–699

Teeri T, Lehtovaara P, Kaapinen S, Salovuori I, Knowles J (1987) Homologous domains in *Trichoderma reesei* cellulolytic enzymes: gene sequence and expression of cellobiohydrolase I. Gene 51: 43–52

Tilbeurgh H van, Claeyssens M, De Bruyne CK (1984a) The use of 4-methylumbelliferyl and other chromophoric glycosides in the study of cellulolytic enzymes. FEBS Lett. 149: 152–156

Tilbeurgh H van, Bikhabai R, Pettersson LG, Claeyssens M (1984b) Separation of endo- and exo-type cellulases using a new affinity chromatography method. FEBS Lett 169: 215–218

Tilbeurgh H van, Tomme P, Claeyssens M, Bikhabai R, Pettersson G (1986) Limited proteolysis of the cellobiohydrolase I from *Trichoderma reesei*: separation of functional domains FEBS Lett 204: 223–227