Biomolecular NMR Assignments (2020) 14:265–268
https://doi.org/10.1007/s12104-020-09959-2

ARTICLE

^{1}H, ^{13}C, ^{15}N backbone and IVL methyl group resonance assignment of the fungal β-glucosidase from *Trichoderma reesei*

Eleni Makraki[1] · Marta G. Carneiro[2] · Alex Heyam[1] · A. B. Eiso[2] · Gregg Siegal[2,3] · Roderick E. Hubbard[1,4]

Received: 10 March 2020 / Accepted: 16 May 2020 / Published online: 19 June 2020

Abstract

β-glucosidases have received considerable attention due to their essential role in bioethanol production from lignocellulosic biomass. β-glucosidase can hydrolyse cellobiose in cellulose degradation and its low activity has been considered as one of the main limiting steps in the process. Large-scale conversions of cellulose therefore require high enzyme concentration which increases the cost. β-glucosidases with improved activity and thermostability are therefore of great commercial interest. The fungus *Trichoderma reseei* expresses thermostable cellulolytic enzymes which have been widely studied as attractive targets for industrial applications. Genetically modified β-glucosidases from *Trichoderma reseei* have been recently commercialised. We have developed an approach in which screening of low molecular weight molecules (fragments) identifies compounds that increase enzyme activity and are currently characterizing fragment-based activators of *Tr*Bgl2. A structural analysis of the 55 kDa apo form of *Tr*Bgl2 revealed a classical $(\alpha/\beta)_8$-TIM barrel fold. In the present study we present a partial assignment of backbone chemical shifts, along with those of the Ile (I)-Val (V)-Leu (L) methyl groups of *Tr*Bgl2. These data will be used to characterize the interaction of *Tr*Bgl2 with the small molecule activators.

Keywords *Tr*Bgl2 · β-glucosidase · Resonance assignment · Enzyme activators

Biological context

The depletion of fossil fuel in combination with the increasing demand for energy worldwide has instigated research on alternative and sustainable energy sources such as biofuels. Lignocellulosic (LC) biomass such as wood, agricultural residues and dedicated energy crops are abundant and available at low cost and have received considerable global attention as the most promising alternative, renewable source for biofuel production. LC biomass, the structural backbone of all plant cell walls, is composed mainly of cellulose, in combination with hemicellulose and lignin (Service 2007). A mixture of enzymes that together are known as cellulase, catalyze cellulose degradation and comprise three categories of enzymes; endoglucanases (EC 3.2.1.4), exoglucanases or cellobiohydrolases (EC 3.2.1.91) and β-glucosidases (EC 3.2.1.21) (Sticklen 2008) (Brethauer and Studer 2015). Endoglucanases cleave the internal β-1,4-glycosidic bonds of cellulose microfibrils releasing small fragments. Subsequently, exoglucanases or cellobiohydrolases (CBH) act on the reducing and non-reducing ends resulting in short chain cello-oligosaccharides such as cellobiose, which are hydrolysed into glucose by the action of β-glucosidases (Sticklen 2008). However, the activity of β-glucosidase is a rate limiting step which results in accumulation of cellobiose and subsequent inhibition of other cellulases (Lynd et al. 2002; Bommarius et al. 2008; Resa and Buckin 2011; Brethauer and Studer 2015). Currently, this is overcome by using high concentrations of β-glucosidases increasing the cost of large-scale conversions. Many efforts have therefore been directed towards the improvement of catalytic activity

✉ Marta G. Carneiro
mcarneiro@zobio.com

✉ Roderick E. Hubbard
roderick.hubbard@york.ac.uk

[1] YSBL, Department of Chemistry, University of York, Heslington, York, UK

[2] ZoBio BV, J.H. Oortweg 19, 2333 CH Leiden, The Netherlands

[3] Division of Medicinal Chemistry, Amsterdam Institute for Molecules, Medicines and Systems (AIMMS), Vrije Universiteit, De Boelelaan 1108, 1081 HZ Amsterdam, Netherlands

[4] Vernalis Research, Granta Park, Abington, Cambridge, UK

and thermostability of the enzyme using mainly traditional genetic approaches (Lee et al. 2012).

Trichoderma reesei produces large amounts of thermostable cellulolytic enzymes, which make them attractive targets for industrial applications (Gao et al. 2017). Genetically modified β-glucosidases from *Trichoderma reesei* have been commercialized and are included in cellulolytic enzyme cocktails that have been developed by companies. Recently, Jeng et al. (2011) solved the crystal structure of the 55 kDa β-glucosidase from *Trichoderma reesei* (*Tr*Bgl2) at 1.63 Å resolution, which was found to adopt a classical $(\alpha/\beta)_8$-TIM barrel fold and be in tight association with a TRIS molecule at the active site (Jeng et al. 2011).

In previous studies, we have identified fragment-based activators of enzyme activity (Darby et al. 2014). We are currently characterizing fragment-based activators of *Tr*Bgl2. The work reported here is a partial assignment of backbone chemical shifts, along with those of the Ile (I)-Val (V)-Leu (L) methyl groups of *Tr*Bgl2. These assignments will provide the basis for a study of the interaction of *Tr*Bgl2 with the small molecule activators.

Methods and experiments

Sample preparation

The codon-optimized DNA sequence for *Tr*Bgl2 was inserted into the pET-YSBLIC3C plasmid with an N-terminal His_6 cleavable tag (Fogg and Wilkinson 2008). The plasmid was transformed into the E. coli BL21 (DE3) bacterial strain. Expression was performed using the method from Gans et al. (2010) with some modifications. Labelled samples were grown in M9/D_2O minimal media containing $^{15}NH_4Cl$ (1.0 g/L) and $^{13}C_6$-glucose (1.5 g/L). The cells were incubated at 37 °C until reaching an OD_{600} of 0.4 when 2-Keto-3-(methyl-d3)-butyric acid- 1,2,3,4-13C4, 3-d sodium salt (SIGMA, Cat. 637858) and 2-Ketobutyric acid-13C4,3,3-d2 sodium salt hydrate (SIGMA, Cat. 607541) were added (Gans et al. 2010; Tugarinov and Kay 2003). The cells were further incubated at 37 °C until reaching an OD_{600} of 0.7–0.8. The temperature was then decreased to 16 °C and the protein expression was induced with 0.5 mM IPTG and cells were harvested after overnight incubation at 16 °C. The cells were resuspended in a buffer containing 20 mM TRIS–HCl pH 7.5, 400 mM NaCl, 10 mM Imidazole and then disrupted by using the Cell Disruption System (Benchtop, Constant Systems Ltd). After centrifugation, the soluble *Tr*Bgl2 was initially purified by Ni^{2+}-NTA chromatography as described previously (Jeng et al. 2011). The sample was then loaded onto a size exclusion column (16/600 Superdex 200, GE Healthcare), and eluted with a buffer containing 50 mM TRIS–HCl pH 8, 100 mM NaCl, 3 mM DTT. The purified protein was concentrated by ultrafiltration with 10 kDa cut-off filter and analyzed by SDS-PAGE. Samples were prepared with 10% D_2O (v/v) yielding concentration of 0.25 mM.

NMR experiments

All NMR data were collected at 298 K on a Bruker Avance Neo 700 MHz and a Bruker Avance III 850 MHz spectrometers equipped with triple-resonance cryogenic probes. Sequential backbone assignments were obtained using Transverse relaxation-optimized spectroscopy (TROSY) versions (Pervushin et al. 1997) of conventional three-dimensional experiments (HNCO, HN(CA)CO, HNCA, HN(CO)CA, HN(CO)CACB, HNCACB and HN(CA)CB) (Sattler et al. 1999). IVL methyl group assignments were obtained using a two-dimensional constant time (CT) ^1H-$^{13}C_{methyl}$ HMQC (Tugarinov and Kay 2003) and three-dimensional HMCMC, HMCM(C)CB, HMCM(CC)CA and HMCM(CCC)CO (Sinha et al. 2013). Additionally, a 3D 13Ch-13CH3 SOFAST-HMQC-NOESY-HMQC with a mixing time of 200 ms (Rossi et al. 2016), a 3D 15N-resolved NOESY with a mixing time of 400 ms, and a 3D 13C-resolved NOESY with a mixing time of 400 ms were used. All methyl assignment experiments were recorded with non-uniform sampling (NUS). NMR data were processed with Topspin 3.2 (Bruker), or with NMRPipe (Delaglio et al. 1995) and SMILE (Ying et al. 2017) and analyzed with Sparky (Lee et al. 2015).

Assignment and data deposition

Spectra of *Tr*Bgl2 were initially assessed in a phosphate-based buffer at pH 6.0 and a TRIS-based buffer at pH 8.0. Spectral quality was markedly improved in TRIS buffer at pH 8.0, with several additional peaks becoming visible (data not shown). This observation, together the fact that a TRIS

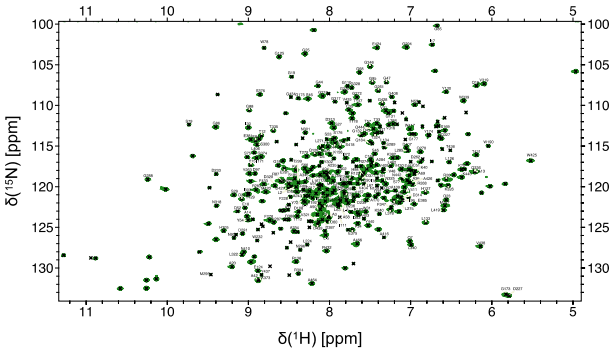

Fig. 1 The assigned 700 MHz [^{15}N,^1H] TROSY NMR spectrum of *Tr*Bgl2. Assigned backbone N–H cross peaks are labelled with the corresponding residue number

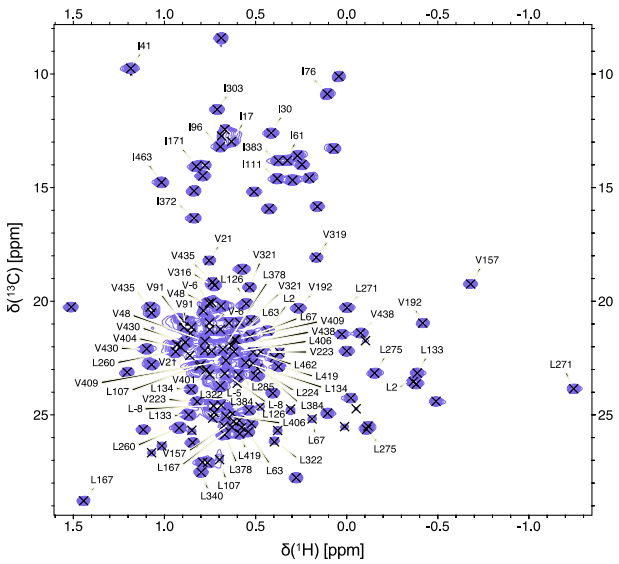

Fig. 2 The assigned 700 MHz [^{13}C,^{1}H] HSQC spectrum of an IVL methyl protonated sample of *Tr*Bgl2. Methyl side chain resonances are indicated by the one-letter amino acid code and the sequence number is shown in black

molecule was found at the active site of the crystallographic structure (Jeng et al. 2011), suggests that TRIS might stabilize apo *Tr*Bgl2.

The chemical shift assignment of *Tr*Bgl2 was performed in semi-automatic mode using the FLYA algorithm (Schmidt and Guntert 2012). The FLYA assignments were manually inspected and extended. Recombinant *Tr*Bgl2 consists of 488 amino acids (including an N-terminal His_6 cleavable tag) and has a molecular weight of 55 kDa. In total, we were able to assign 265 backbone amide resonances and 55% of the methyl group resonances of the IVL residues, from 459 non-proline residues (Figs. 1 and 2). Assignment completeness is limited possibly due to incomplete back-exchange of perdeuterated amide groups, with only ~ 75% of the expected resonances being observed in the [^{15}N,^{1}H] TROSY spectrum of perdeuterated *Tr*Bgl2. Nonetheless, the assignments here reported provide a useful NMR basis for studies of the interaction of *Tr*Bgl2 with small molecules. The assignment of chemical shifts of *Tr*Bgl2 has been deposited into BMRB (https://www.bmrb.wisc.edu/) with accession number 50158.

*Tr*Bgl2 secondary structure motifs were predicted by the TALOS-N software (Shen and Bax 2013) using the assigned backbone resonances as input data. The software indicated that the overall secondary structures are in agreement with the X-ray structure of *Tr*Bgl2 (PDB ID 3AHY) (Jeng et al. 2011) (Fig. 3).

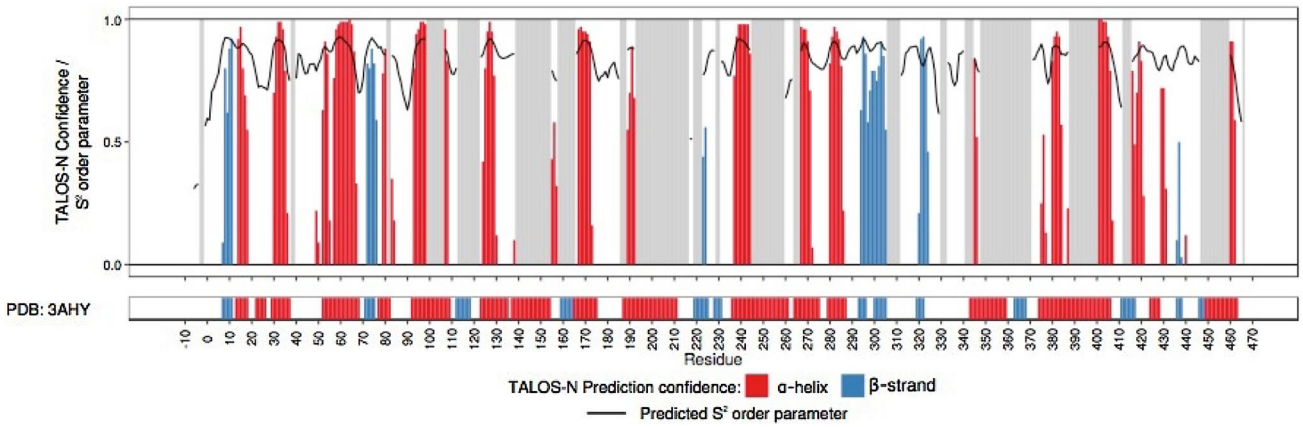

Fig. 3 Secondary structure prediction of *Tr*Bgl2 analyzed with TALOS-N using the assigned chemical shifts compared to the secondary structure of the X-ray structure of *Tr*Bgl2. Top: colored bars (red and blue bars indicate α-helix and β-strands respectively) show the secondary structure type predicted by TALOS-N. The bar height represents the prediction confidence. The black line shows the predicted S^2 order parameter, a measure of flexibility. Bottom: the secondary structure of *Tr*Bgl2 as determined by X-ray crystallography (PDB ID 3AHY) (Jeng et al. 2011)

Acknowledgements EM is very grateful to Michael Plevin (University of York) for his help and support in sample preparation. This work was supported by the European Union's Horizon2020 MSCA Programme under grant agreement 675899 (FRAGNET); research in the group of R.E.H. was additionally supported by research Grants from the BBSRC (BB/N008332/1) and institutional infrastructure support from funds provided by the Wellcome Trust and EPSRC.

Funding This work was supported by the European Union's Horizon2020 MSCA Programme under Grant agreement 675899 (FRAGNET); research in the group of R.E.H. was additionally supported by Research Grants from the BBSRC (BB/N008332/1) and institutional infrastructure support from funds provided by the Welcome Trust and EPSRC.

Data availability The data that support the findings of this work are available on BMRB with the following entry assigned accession number: 50158.

Compliance with ethical standards

Conflict of interest I confirm that all authors of the manuscript have no conflict of interest to declare.

Informed consent I confirm that all authors of the manuscript have given consent to participate to this work.

References

Bommarius AS et al (2008) Cellulase kinetics as a function of cellulose pretreatment. Metab Eng 10(6):370–381. https://doi.org/10.1016/j.ymben.2008.06.008

Brethauer S, Studer MH (2015) Biochemical conversion processes of lignocellulosic biomass to fuels and chemicals—a review. https://www.ingentaconnect.com/content/scs/chimia/2015/00000069/00000010/art00002

Darby JF et al (2014) Discovery of selective small-molecule activators of a bacterial glycoside hydrolase. Angew Chem Int Ed Engl 53(49):13419–13423. https://doi.org/10.1002/anie.201407081

Delaglio F et al (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR 6(3):277–293. https://doi.org/10.1007/bf00197809

Fogg MJ, Wilkinson AJ (2008) Higher-throughput approaches to crystallization and crystal structure determination. Biochem Soc Trans 36(4):771–775. https://doi.org/10.1042/bst0360771

Gans P et al (2010) Stereospecific isotopic labeling of methyl groups for NMR spectroscopic studies of high-molecular-weight proteins. Angew Chem Int Ed Engl 49(11):1958–1962. https://doi.org/10.1002/anie.200905660

Gao J et al (2017) Production of the versatile cellulase for cellulose bioconversion and cellulase inducer synthesis by genetic improvement of *Trichoderma reesei*. Biotechnol Biofuels 10:272. https://doi.org/10.1186/s13068-017-0963-1

Jeng WY et al (2011) Structural and functional analysis of three beta-glucosidases from bacterium *Clostridium cellulovorans*, fungus T*richoderma reesei* and termite *Neotermes koshunensis*. J Struct Biol 173(1):46–56. https://doi.org/10.1016/j.jsb.2010.07.008

Lee HL et al (2012) Mutations in the substrate entrance region of beta-glucosidase from *Trichoderma reesei* improve enzyme activity and thermostability. Protein Eng Des Sel 25(11):733–740. https://doi.org/10.1093/protein/gzs073

Lee W, Tonelli M, Markley JL (2015) NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. Bioinformatics 31(8):1325–1327. https://doi.org/10.1093/bioinformatics/btu830

Lynd LR et al (2002) Microbial cellulose utilization: fundamentals and biotechnology. Microbiol Mol Biol Rev 66(3):506–577. https://doi.org/10.1128/mmbr.66.3.506-577.2002

Pervushin K et al (1997) Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. Proc Natl Acad Sci USA 94(23):12366–12371. https://doi.org/10.1073/pnas.94.23.12366

Resa P, Buckin V (2011) Ultrasonic analysis of kinetic mechanism of hydrolysis of cellobiose by beta-glucosidase. Anal Biochem 415(1):1–11. https://doi.org/10.1016/j.ab.2011.03.003

Rossi P et al (2016) 15N and 13C-SOFAST-HMQC editing enhances 3D-NOESY sensitivity in highly deuterated, selectively [1H,13C]-labeled proteins. J Biomol NMR 66(4):259–271. https://doi.org/10.1007/s10858-016-0074-5

Sattler M, Schleucher J, Griesinger C (1999) NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. Progress Nucl Magnet Resonan Spectrosc 34:93–158. https://doi.org/10.1016/S0079-6565(98)00025-9

Schmidt E, Guntert P (2012) A new algorithm for reliable and general NMR resonance assignment. J Am Chem Soc 134(30):12817–12829. https://doi.org/10.1021/ja305091n

Service RF (2007) Cellulosic ethanol: biofuel researchers prepare to reap a new harvest. Science 315:1488–1491

Shen Y, Bax A (2013) Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. J Biomol NMR 56(3):227–241. https://doi.org/10.1007/s10858-013-9741-y

Sinha K, Jen-Jacobson L, Rule GS (2013) Divide and conquer is always best: sensitivity of methyl correlation experiments. J Biomol NMR 56(4):331–335. https://doi.org/10.1007/s10858-013-9751-9

Sticklen MB (2008) Plant genetic engineering for biofuel production: towards affordable cellulosic ethanol. Nat Rev Genet 9(6):433–443. https://doi.org/10.1038/nrg2336

Tugarinov V, Kay LE (2003) Ile, Leu, and Val methyl assignments of the 723-residue malate synthase G using a new labeling strategy and novel NMR methods. J Am Chem Soc 125(45):13868–13878. https://doi.org/10.1021/ja030345s

Ying J et al (2017) Sparse multidimensional iterative lineshape-enhanced (SMILE) reconstruction of both non-uniformly sampled and conventional NMR data. J Biomol NMR 68(2):101–118. https://doi.org/10.1007/s10858-016-0072-7