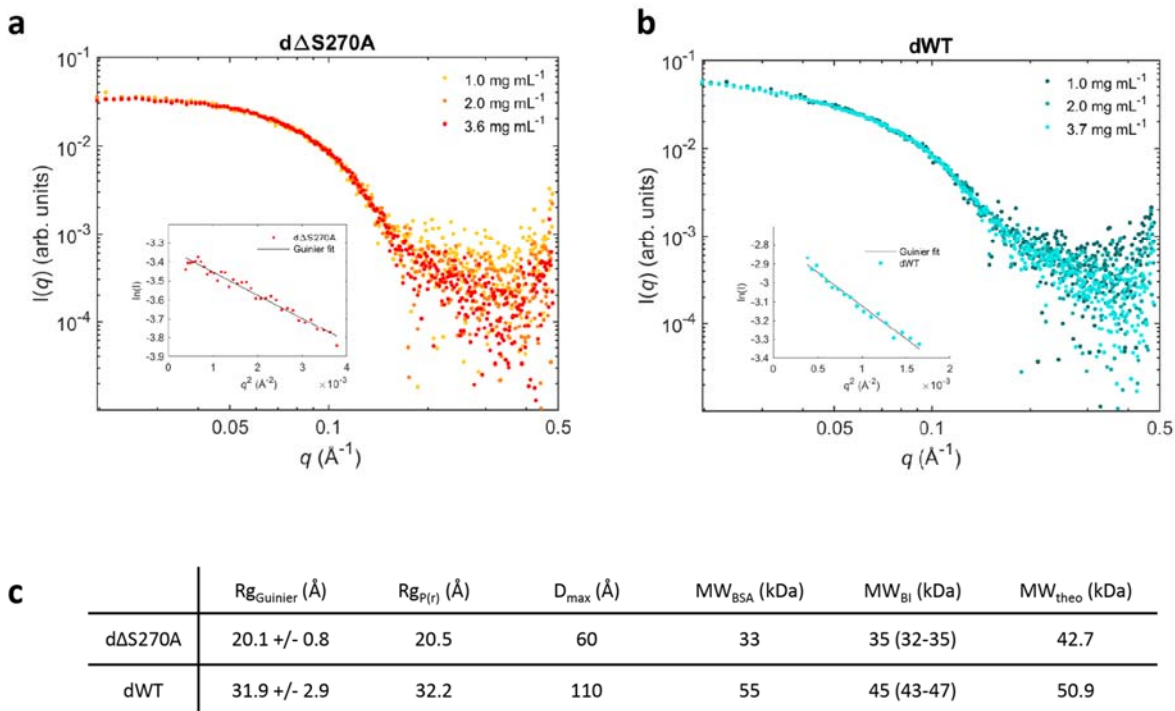


1 **Supplementary Information**

2 **The structural basis of fungal glucuronoyl esterase activity on natural substrates**

3 Ernst et al.

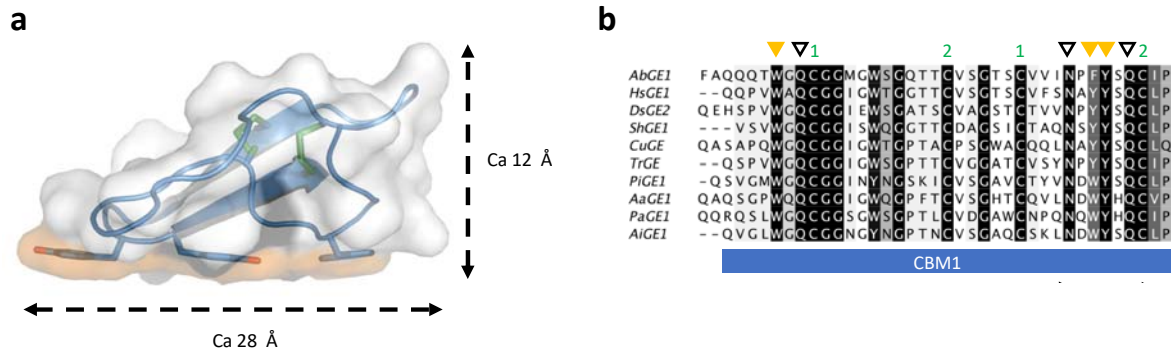


5

6 **Supplementary Figure 1 – SAXS data**

7 a) Experimental SAXS scattering profiles from dΔS270A measured at different concentrations. Insert:
 8 Guinier fit to the 3.6 mg mL⁻¹ data. b) Experimental SAXS scattering profiles from dWT. c) Summary of
 9 parameters from the primary SAXS data analyses. Two different approaches are included for MW
 10 estimation. MW_{BSA} is the classical approach, being concentration dependent and based on measurements
 11 of Bovine Serum Albumin (BSA) as reference. MW_{BI} is a recent addition to PRIMUS, being concentration
 12 independent and based on Bayesian Inference¹. Confidence intervals are given in parentheses. Source data
 13 for panel a and b are provided as a Source Data file.

14

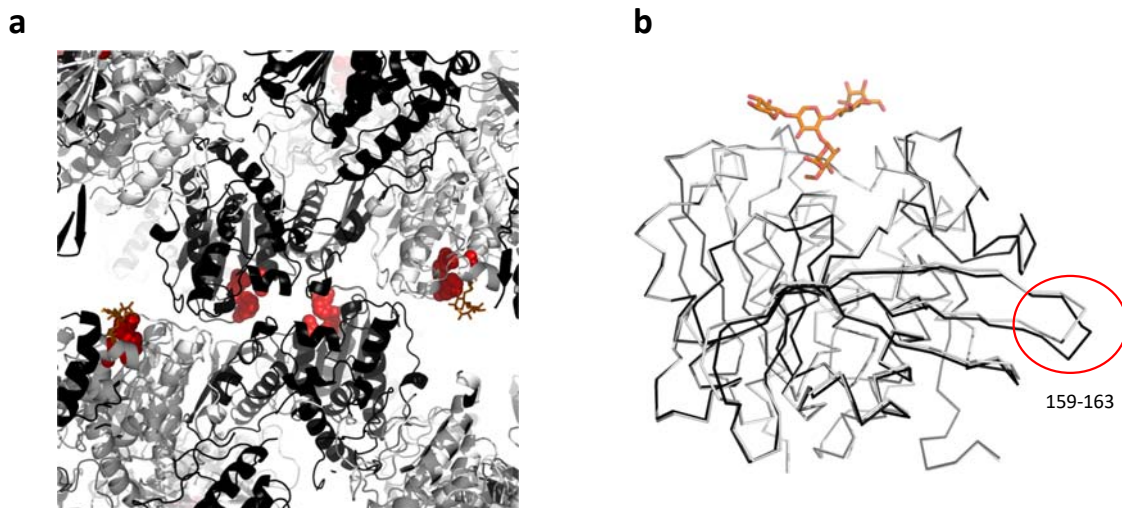


15
16

17 **Supplementary Figure 2 – *CuGE*-CBM1 homology model**

18 **a)** Homology model of *CuGE*-CBM1 (residues 3-38) made with SWISS-MODEL² using the structure of 1CBH³
 19 as template. The CBM1 domain is wedge-shaped with approximate dimensions 28 Å x 12 Å (thick end) x 20
 20 Å. The fold is stabilized by two disulfide bridges (green) and three conserved aromatic residues (orange)
 21 create a flat binding site expected to be implicated in binding of insoluble substrates. **b)** Multiple sequence
 22 alignment of CBM1 domains from functionally characterized fungal CE15 enzymes. Conserved residues
 23 forming the flat binding site are marked by orange triangles for aromatic residues and open triangles for
 24 polar residues. Cysteines involved in disulphide bridges are marked by green numbers.

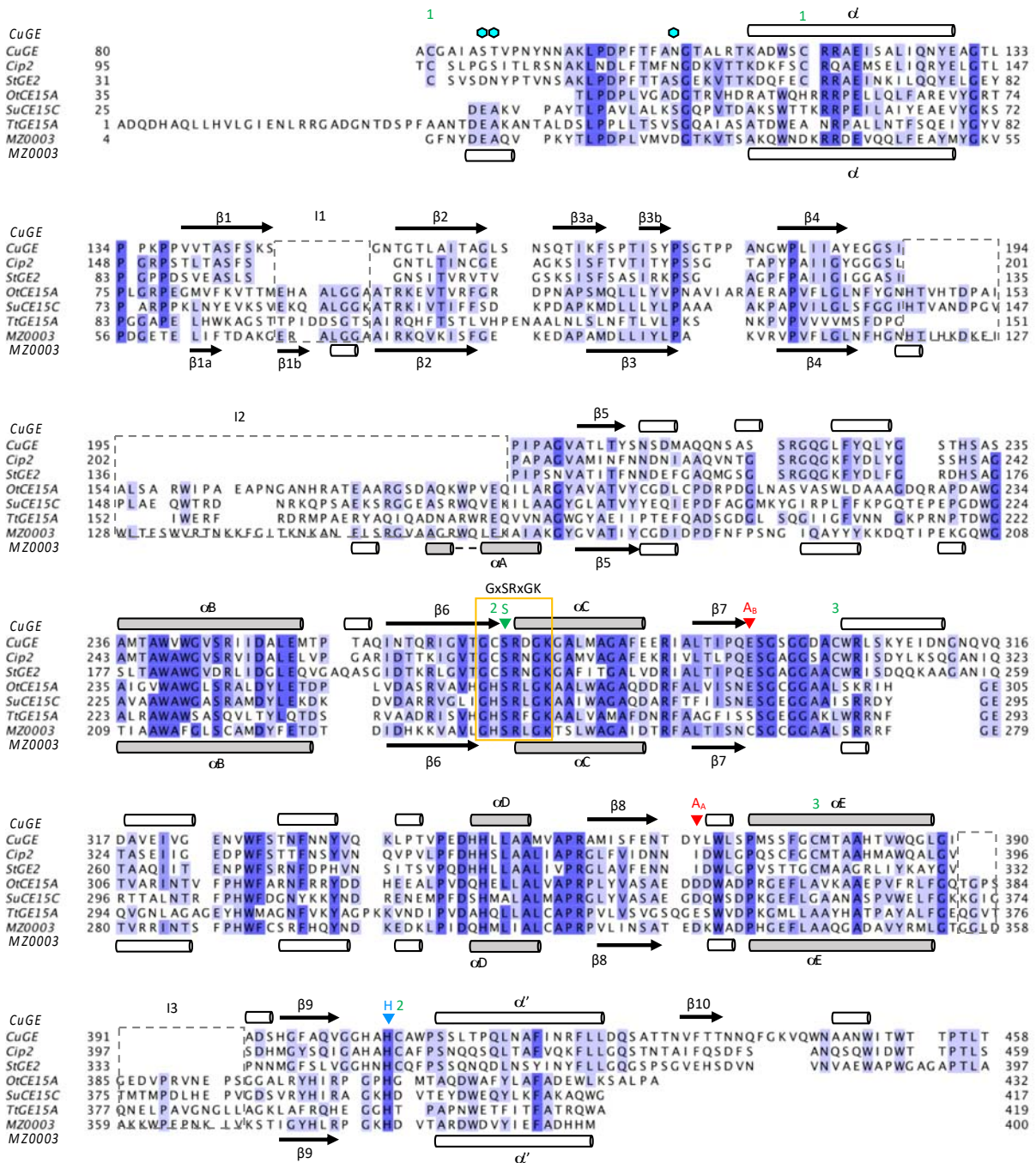
25



26

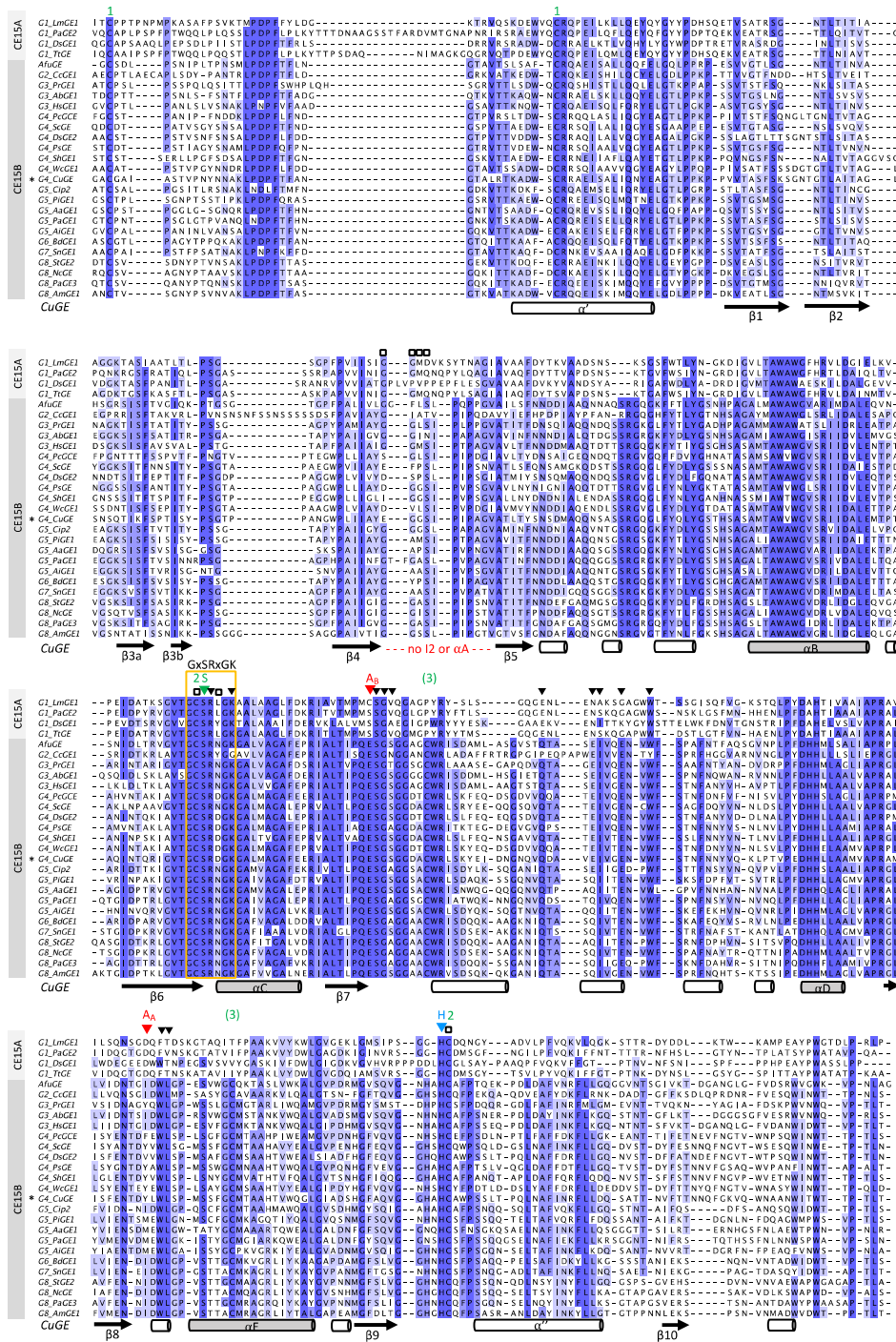
27 **Supplementary Figure 3 – CuGE crystal packing and two distinct active site environments**

28 **a)** Illustration of the crystal packing in the tetragonal space group $P4_12_12$. The two crystallographically
 29 independent molecules A (black) and B (white) differ in crystal packing environments and accessibility to
 30 the active site. To illustrate the location of the active site, the catalytic residues are shown as red spheres
 31 and the ligand from the $\Delta S270A:XU^{m4}XX^{OH}$ complex is shown in orange. Note that the active site in
 32 molecule A is partly occluded due to contacts in the crystal with a symmetry related molecule (A'), whereas
 33 the binding site in molecule B is solvent-exposed and easily accessible. **b)** Superposition of molecules A
 34 (black) and B (white) shown as C_α -traces. Molecule A and B superimpose with an RMSD of 0.32 Å for 359
 35 aligned C_α atoms (apo-dDS270A), reflecting only minor structural variations. The main difference concerns
 36 the β -turnb2-b3 (residues 159-163), which is affected by crystal packing of molecule B.



37

38 **Supplementary Figure 4 – Structure-based sequence alignment of all structurally characterized CE15 enzymes.**
 39 Secondary structure elements are shown for MZ0003 (6EHN)⁴ and *CuGE* (6RTV) as representatives for subgroups CE15-
 40 A and CE15-B, respectively. The major secondary structure elements are labelled according to the topology diagrams in
 41 Fig. 3 d-f in the main text. Residues comprising the catalytic triad are indicated with triangles: nucleophile (serine) in
 42 green, base (histidine) in blue and acid (glutamate/aspartate) in red. The two different structural locations of the acidic
 43 residue are labeled A_α or A_β, with reference to the configuration in α/β-hydrolase group A and B, respectively. The three
 44 major “inserts” in the bacterial structures are labelled I1, I2 & I3, green numbers indicated disulfide bridges and cyan
 45 hexagons symbolize glycosylation sites observed in the *CuGE* structures. The CE15-signature sequence “GxSRxGK” is
 46 highlighted with an orange box. PDB codes for all structures in the alignment: *CuGE* 6RTV; *Cip2* 3PIC; *StGE2* 4G4G;
 47 *OtCE15A* 6GS0; *SuCE15C* 6GRY; *TtGE15A* 6HSW, MZ0003 6EHN.



48
49
50
51
52
53
54
55
56
57
58
59
60

Supplementary Figure 5 Multiple sequence alignment (MSA) of functionally characterized fungal CE15-A and CE15-B enzymes. The alignment comprises sequences included in the nomenclature paper by Dilokpimol *et al.* on fungal CE15 GEs (subgroups G1-G8)⁵ and the *AfuGe* sequence from Mosbech *et al.*⁶ Catalytic residues are indicated with triangles as in Fig. 3 and Supplementary Fig. 4. Other residues within 5 Å from the XU^mX moiety in the CuGE:ligand complexes are marked with black triangles. Residues forming the putative monoligand docking site are marked with open squares. The two different structural locations of the acidic residue, that differentiate the subgroups CE15-A and CE15-B, are labeled A_A and A_B, respectively. Secondary structure elements from the CuGE structure (6RTV) are included for reference. The MSA suggests that the fungal CE15-A enzymes, apart from the distinct location of the acidic catalytic residue, display considerable structural similarity to the CE15-B counterparts. Two of the three disulfide bridges (green numbers) are conserved. Moreover, the loop between β4 and β5 is only a few residues longer in fungal CE15-A enzymes implying that these enzymes also lack I2 and the αA-helix and thus are predicted to have a relatively open binding site similar to fungal CE15-B enzymes.

```

1
Hyspu1_1114752 ITCPATPSRWRKADSVYKISITMPPDPFTYID-----GKTRVQSKAEWYACRQPEILNLFQEQYQYGYPDHSKETTITVTRSG-----
Eutla1|3629 -----MPPDFLPIARTLTDNGEGDGG---FANDVMTGKAGPRIQTPEEYWRCKRPEILLQLLEQYQYGYPDHSLETVTATRSG-----N
AE060465.1 IECAPIPSFPTFTWQELPQLQSSMPPDPFLPIAYTTPDNAA-----DVVVAGRGKGRVQTPEEYWRCKRPEILLQLLEQYQYGYPDHSEKVEATRSG-----
Spath2|2119719 IECAPIPSFPTFTWQELPQLQSSMPPDPFLPIAYTTPDNAA-----DVVVAGRGKGRVQTPEEYWRCKRPEILLQLLEQYQYGYPDHSEKVEATRSG-----
Chagl_1|11918 IQCPAPTSRWRFTWQQLPQLQSSLPDPFLPIQYTTTPSDAQ-----NIMAGKGGQGRVQTPEEYWRCKRPEILLQMLEQYQYGYPDHSEKVEATRSG-----
Thie2|2055580 IQCAALPSRFRFTWQQLPQLQSSLPDPFLPIQYTTTPSDAQ-----NIMAGKGGQGRVQTPEEYWRCKRPEILLQMLEQYQYGYPDHSEKVEATRSG-----
Pha1|17987 -----MPPDFLPIQYTTNNNAQSASAFQAQVMTGKGGQGRVQTPEEYWRCKRPEILLQMLEQYQYGYPDHSEKVEATRSG-----
Podan2|25 VQCAPLPSFPTFTWQQLPQLQSSLPDPFLPIKYTTTNDNAAGSS---TFARDVMTGNAPNIRSRAEWYQCRQPEILQFLQEQYQYGYPDHTQEKVEATRSG-----
Aurd3_1|1325771 IACPATPNSRFRKAGTFFIPIPTLPDPDPFTYIDG-----KTRVQSPPEWRDCKRPEILLRFLEQYQYGYPDHKLKETTATRSG-----
Gla1|1685 QQCPAFPERRRAPASLPSEATLPPDPKPYTTI-----DRVRSRADWDCKRAELKTLVQHLYGWNQDPTRRETVAASRDG-----
Copci1|4157 QQCPAFPERRRAPASLPSEATLPPDPKPYTTI-----DRVRSRADWDCKRAELKTLVQHLYGWNQDPTRRETVAASRDG-----
CBX90574.1 ITCPTTNPMPKKAASAFPSVKTMPPDPFFYIDG-----KTRVQSKDEWYQCRQPEILLKLLQEQYQYGYPDHSQETVSTRSG-----
Lepmu1|12448 ITCPTTNPMPKKAASAFPSVKTMPPDPFFYIDG-----KTRVQSKDEWYQCRQPEILLKLLQEQYQYGYPDHSQETVSTRSG-----
AB06204.1 VTCPTTNPMPKKAASAFPSVKTMPPDPFFYIDG-----KTRVQSQDEWLQCRQPEILLKLLQEQYQYGYPDHSQETVSTRSG-----
Hetan2|64706 SAGAPIRFRLLPQASDLPIIPLLPDPPTFRLLS-----PDRVRSRADW-ACRQKEIFITLVOQYLYGYPDHAETVSTRSG-----
Botbo1|53049 TNCPALPDAITFNGASLPVTLATLPDPDWTFFYD-----GSKMYSKQW-ACRKAELKTLVQHMVYQYQYGYPDHSEKVEATRSG-----
Dicsq1|58498 QGCAAPSAAQLPEPSDLPIISLTPDPPTFRLLS-----DRVRSRADWDCKRAELKTLVQHLYGWNQDPTRRETVAASRDG-----
Chagl_1|11918 IQCPAPLBARLEPSDLPVHITLPPDPPTFRLLS-----DRRVSHADW-PCRRAELKTLVQHLYGWNQDPTRRETVAASRDG-----
Gansp1|41841 QGCAAPLBARLEPSDLPVHITLPPDPPTFRLLS-----DRRVSHADW-PCRRAELKTLVQHLYGWNQDPTRRETVAASRDG-----
Trave1|34910 -----FPFIIDTLPPDPPTFRLLS-----NRVRSRADW-ECRREELMMLVQYLYGYPDHSRRETVAASRDG-----
Fomme1|132657 RSCSSLPNPLRFPDTLPIEALNPPKFFN-----NESVRSPEDW-ECRAELKTLVQHMVYQYGYPDHSEKVEATRSG-----

```

```

Hyspu1_1114752 TTLSLQVA-AGGKTTVTKMSINLPSGTS---TGGAPVVAIAIG---GIDNKAYLNQGLAVVTEDYSSVAADSNKTKGSFWDLYKQDQIGVLTAWAWGFHRLDGLI
Eutla1|3629 TCVSVEVS-AGGKTGRFTAQSLRAGAS---AESRVPVVINIG---GMQSQPYLQAGIUVGFDYTAAPDSNAKSGAFWALYNGRDIQVLTAWAWGFHRLDGLI
AE060465.1 NTNLIVYVT-AGGKQGSFRATISLPSGAS---ASNPAVVINIG---GMQNPYLSAGIAVAQFDYTTVSPDSNAKTGAFWSIYNGRDIQVLTAWAWGFHRLDGLI
Spath2|2119719 NTNLIVYVT-AGGKQGSFRATISLPSGAS---ASNPAVVINIG---GMQNPYLSAGIAVAQFDYTTVSPDSNAKTGAFWSIYNGRDIQVLTAWAWGFHRLDGLI
Chagl_1|11918 NTLSLTVTA-AGGKTGRFSAATVTLPSGAS---ASRPAVVINIG---GMQNGAYLSAGIAVAQFDYTSVAPDSNAKTGAFWSIYNGRDIQVLTAWAWGFHRLDGLI
Thie2|2055580 NTNLIAVIT-AGDKTGFKATLTLPSGAS---ASKPAVVINIG---GMQNPYLSAGIAVAQFDYTSVAPDSNAKTGAFWSIYNGRDIQVLTAWAWGFHRLDGLI
Pha1|17987 NTNLIAVIT-AGDKTGFKATLTLPSGAS---KDKPAVVINIG---GMQNPYLSAGIAVAQFDYTSVAPDSNAKTGAFWDDVKYGRDIQVLTAWAWGFHRLDGLI
Podan2|25 TLLQITVTGPQNKRGSFRATIQLPSGAS---SSRPAVVINIG---GMQNPYLSAGIAVAQFDYTTVAPDSNAKTGAFWSIYNGRDIQVLTAWAWGFHRLDGLI
Aurd3_1|1325771 NNLSSVSA-AGSKTGTFGAQLLPSGAS---SASRPAVVINIG---GIDNNAAYLNSGMVAUVSFDYTAAPDSNAKTGAFWALYNGRDIQVLTAWAWGFHRLDGLI
Gla1|1685 NSVTINIA-VGSRKSSDFKATINLPSGTT---SSAPAPVVAIAIG---GIDNNAAYLNSGMVAUVSFDYTTVAPDSNAKTGAFWALYNGRDIQVLTAWAWGFHRLDGLI
Copci1|4157 NTLLTINWS-AGGRNGEFTASLALPSGAS---PSNKVPPVVAIAIG---GVNNGAVTNARIAVATINVMNVAAPDSNAKTGTFWLSYNGRDIQVLTAWAWGFHRLDGLI
CBX90574.1 NTLLTITIA-AGGKTASIAATLTLPSG---SGPFPVVISIG---GMDVKSNTAGIAVAADFYTKVAADSNKSGSFWTLYNGRDIQVLTAWAWGFHRLDGLI
Lepmu1|12448 NTLLTITIA-AGGKTASIAATLTLPSG---SGPFPVVISIG---GMDVKSNTAGIAVAADFYTKVAADSNKSGSFWTLYNGRDIQVLTAWAWGFHRLDGLI
AB06204.1 NTLSTVTA-AGGKTASIAASLTLPSG---SGPFPVVISIG---GMDAKSNTAGIAVATFDYTRVAADSNKSGSFWTLYNGRDIQVLTAWAWGFHRLDGLI
Hetan2|64706 NALNTLVTS-VGRKASAFPATLTFNTVNTASQNPVPPVVINIPG---EIPTAPFLASGALVTIRVGDVAADSNKSGSFWTLYNGRDIQVLTAWAWGFHRLDGLI
Botbo1|53049 NTLSSVSA-AGGKTGTFASATLTLPTT---GQAPRPPVVAITP---SVDNNAFLSQGVALSFSVTAVAADSNKSGSFWTLYNGRDIQVLTAWAWGFHRLDGLI
Dicsq1|58498 INLTLSSVS-VDRKTFASFPANITLPSGA---SRANRVPVVAIATPLVPPVPPFLSGLSVAVAAFDVKYVANDSARIGAFWDLVYAWAWASKLILDAL
Gansp1|41841 SNLTLTVTS-VGGRTASFPANVTLPAS-RGGKVFPAVIAATPLVPPVPPFLSGLSVAVAAFDVKYVANDSARIGAFWDLVYAWAWASKLILDAL
Trave1|34910 DNLTINNF-VDRKTFASFPALITLFRPNA-THAGRVPVVAIIPG---SIDNCPFLDSGVALATFDVSTAVDSTARIGAFWTLYDDRIGLITAWAWAHRHRLDGLI
Fomme1|132657 SQITINIS-VGENSTFSGATLSMTSVN---ATPKAPVVAIIPG---SIDNCLATSVENSDAENIIVPGCAFWDLYSGRDIQVLTAWAWGFHRLDGLI

```

GxSRxGK

2S

```

Hyspu1_1114752 TDKVPEIDHTKVGVTGCSRLCKAALAAGLFDRITLVTMPMSSGVQGLGPPRYRHALS---GQDETLKNSKAGAWWTDVSIQFVCG---KSEQLPYDAHTIAAAIAP
Eutla1|3629 NLTVPFIDNPSLAGVTCGSRLLCKAALAAGLFDRILTLTTPMSSGVQGLGPPRYRHALS---GQDETLKNSKAGAWWTDVSIQFVCG---KSEQLPYDAHTIAAAIAP
AE060465.1 NLTVPFIDNPSLAGVTCGSRLLCKAALAAGLFDRILTLTTPMSSGVQGLGPPRYRHALS---GQDETLKNSKAGAWWTDVSIQFVCG---KSEQLPYDAHTIAAAIAP
Spath2|2119719 NLTVPFIDNPSLAGVTCGSRLLCKAALAAGLFDRILTLTTPMSSGVQGLGPPRYRHALS---GQDETLKNSKAGAWWTDVSIQFVCG---KSEQLPYDAHTIAAAIAP
Chagl_1|11918 NMTVLEIDAGRVGTGCSRLCKAALAAGLFDRITLVTMPMSSGVQGLGPPRYRHALS---GQDETLKNSKAGAWWTDVSIQFVCG---KSEQLPYDAHTIAAAIAP
Thie2|2055580 NMTVLEIDAGRVGTGCSRLCKAALAAGLFDRILTLTTPMSSGVQGLGPPRYRHALS---GQDETLKNSKAGAWWTDVSIQFVCG---KSEQLPYDAHTIAAAIAP
Pha1|17987 NMTVLEIDAGRVGTGCSRLCKAALAAGLFDRILTLTTPMSSGVQGLGPPRYRHALS---GQDETLKNSKAGAWWTDVSIQFVCG---KSEQLPYDAHTIAAAIAP
Podan2|25 QLTVPFIDNPSLAGVTCGSRLLCKAALAAGLFDRILTLTTPMSSGVQGLGPPRYRHALS---GQDETLKNSKAGAWWTDVSIQFVCG---KSEQLPYDAHTIAAAIAP
Aurd3_1|1325771 AMRVPEIDSRNGVTCGSRLLCKAALAAGLFDRITLVTMPMSSGVQGLGPPRYRHALS---GQDETLKNSKAGAWWTDVSIQFVCG---KSEQLPYDAHTIAAAIAP
Gla1|1685 ILTAPEIINSARVGTGCSRLCKAALAAGLFDRITLVTMPMSSGVQGLGPPRYRHALS---GQDETLKNSKAGAWWTDVSIQFVCG---KSEQLPYDAHTIAAAIAP
Copci1|4157 EAVVPEIDATKATVTCGSRLLCKAALAAGLFDRILTLTTPMSSGVQGLGPPRYRHALS---GQDETLKNSKAGAWWTDVSIQFVCG---KSEQLPYDAHTIAAAIAP
CBX90574.1 ELKVPFIDATKSGVTCGSRLLCKAALAAGLFDRILTLTTPMSSGVQGLGPPRYRHALS---GQDETLKNSKAGAWWTDVSIQFVCG---KSEQLPYDAHTIAAAIAP
Lepmu1|12448 ELKVPFIDATKSGVTCGSRLLCKAALAAGLFDRILTLTTPMSSGVQGLGPPRYRHALS---GQDETLKNSKAGAWWTDVSIQFVCG---KSEQLPYDAHTIAAAIAP
AB06204.1 ELKVPFIDATKSGVTCGSRLLCKAALAAGLFDRILTLTTPMSSGVQGLGPPRYRHALS---GQDETLKNSKAGAWWTDVSIQFVCG---KSEQLPYDAHTIAAAIAP
Hetan2|64706 ELKVPFIDATKSGVTCGSRLLCKAALAAGLFDRILTLTTPMSSGVQGLGPPRYRHALS---GQDETLKNSKAGAWWTDVSIQFVCG---KSEQLPYDAHTIAAAIAP
Botbo1|53049 ILRVPFIDATKSGVTCGSRLLCKAALAAGLFDRILTLTTPMSSGVQGLGPPRYRHALS---GQDETLKNSKAGAWWTDVSIQFVCG---KSEQLPYDAHTIAAAIAP
Dicsq1|58498 GEVVPFIDATKSGVTCGSRLLCKAALAAGLFDRILTLTTPMSSGVQGLGPPRYRHALS---GQDETLKNSKAGAWWTDVSIQFVCG---KSEQLPYDAHTIAAAIAP
Gansp1|41841 EQVVPFIDATKSGVTCGSRLLCKAALAAGLFDRILTLTTPMSSGVQGLGPPRYRHALS---GQDETLKNSKAGAWWTDVSIQFVCG---KSEQLPYDAHTIAAAIAP
Trave1|34910 EQVVPFIDATKSGVTCGSRLLCKAALAAGLFDRILTLTTPMSSGVQGLGPPRYRHALS---GQDETLKNSKAGAWWTDVSIQFVCG---KSEQLPYDAHTIAAAIAP
Fomme1|132657 EQVVPFIDATKSGVTCGSRLLCKAALAAGLFDRILTLTTPMSSGVQGLGPPRYRHALS---GQDETLKNSKAGAWWTDVSIQFVCG---KSEQLPYDAHTIAAAIAP

```

Aa

H2

```

Hyspu1_1114752 RALIDGCGGDFNNSKGTAVIVYPVPAKLVVYNLWLGVEDQIGMAIRSR---GGHCDNSGYTNVLPVFAKILKTPTRRNVDLTPW-TAMTTRYPWSDIIPKAA
Eutla1|3629 RAVIIDGCGTDFNNSKGTAVIVYPVPAKLVVYNLWLGVEDQIGMAIRSR---GGHCDNSGYTNVLPVFAKILKTPTRRNVDLTPW-TAMTTRYPWSDIIPKAA
AE060465.1 RAVIIDGCGTDFNNSKGTAVIVYPVPAKLVVYNLWLGVEDQIGMAIRSR---GGHCDNSGYTNVLPVFAKILKTPTRRNVDLTPW-TAMTTRYPWSDIIPKAA
Spath2|2119719 RAVIIDGCGTDFNNSKGTAVIVYPVPAKLVVYNLWLGVEDQIGMAIRSR---GGHCDNSGYTNVLPVFAKILKTPTRRNVDLTPW-TAMTTRYPWSDIIPKAA
Chagl_1|11918 RALVLDGCGTDFNNSKGTAVIVYPVPAKLVVYNLWLGVEDQIGMAIRSR---GGHCDNSGYTNVLPVFAKILKTPTRRNVDLTPW-TAMTTRYPWSDIIPKAA
Thie2|2055580 RALVLDGCGTDFNNSKGTAVIVYPVPAKLVVYNLWLGVEDQIGMAIRSR---GGHCDNSGYTNVLPVFAKILKTPTRRNVDLTPW-TAMTTRYPWSDIIPKAA
Pha1|17987 RALVLDGCGTDFNNSKGTAVIVYPVPAKLVVYNLWLGVEDQIGMAIRSR---GGHCDNSGYTNVLPVFAKILKTPTRRNVDLTPW-TAMTTRYPWSDIIPKAA
Podan2|25 RALVLDGCGTDFNNSKGTAVIVYPVPAKLVVYNLWLGVEDQIGMAIRSR---GGHCDNSGYTNVLPVFAKILKTPTRRNVDLTPW-TAMTTRYPWSDIIPKAA
Aurd3_1|1325771 RHLIIDGCGTDFNNSKGTAVIVYPVPAKLVVYNLWLGVEDQIGMAIRSR---GGHCDNSGYTNVLPVFAKILKTPTRRNVDLTPW-TAMTTRYPWSDIIPKAA
Gla1|1685 RALVLDGCGTDFNNSKGTAVIVYPVPAKLVVYNLWLGVEDQIGMAIRSR---GGHCDNSGYTNVLPVFAKILKTPTRRNVDLTPW-TAMTTRYPWSDIIPKAA
Copci1|4157 RAVIIDSQNSGQDFTDSKGTQITFPAKLVVYNLWLGVEDQIGMAIRSR---GGHCDNSGYTNVLPVFAKILKTPTRRNVDLTPW-TAMTTRYPWSDIIPKAA
CBX90574.1 RAVIIDSQNSGQDFTDSKGTQITFPAKLVVYNLWLGVEDQIGMAIRSR---GGHCDNSGYTNVLPVFAKILKTPTRRNVDLTPW-TAMTTRYPWSDIIPKAA
Lepmu1|12448 RAVIIDSQNSGQDFTDSKGTQITFPAKLVVYNLWLGVEDQIGMAIRSR---GGHCDNSGYTNVLPVFAKILKTPTRRNVDLTPW-TAMTTRYPWSDIIPKAA
AB06204.1 RAVIIDSQNSGQDFTDSKGTQITFPAKLVVYNLWLGVEDQIGMAIRSR---GGHCDNSGYTNVLPVFAKILKTPTRRNVDLTPW-TAMTTRYPWSDIIPKAA
Hetan2|64706 RAVIWDGEGQDWTNPEGSVYVFAASKLVFDWDLGIDGVKGVVHVRPPDDLHCLGSAYPAAQFVQKVFVFGTPTNVNFSNISP-PPHPEAYPWASLIPP-
Botbo1|53049 RAVIWDGEGQDWTNPEGSVYVFAASKLVFDWDLGIDGVKGVVHVRPPDDLHCLGSAYPAAQFVQKVFVFGTPTNVNFSNISP-PPHPEAYPWASLIPP-
Dicsq1|58498 RAVIWDGEGQDWTNPEGSVYVFAASKLVFDWDLGIDGVKGVVHVRPPDDLHCLGSAYPAAQFVQKVFVFGTPTNVNFSNISP-PPHPEAYPWASLIPP-
Gansp1|41841 RAVIWDGEGQDWTNPEGSVYVFAASKLVFDWDLGIDGVKGVVHVRPPDDLHCLGSAYPAAQFVQKVFVFGTPTNVNFSNISP-PPHPEAYPWASLIPP-
Trave1|34910 RAVIWDGEGQDWTNPEGSVYVFAASKLVFDWDLGIDGVKGVVHVRPPDDLHCLGSAYPAAQFVQKVFVFGTPTNVNFSNISP-PPHPEAYPWASLIPP-
Fomme1|132657 RAVIWDGEGQDWTNPEGSVYVFAASKLVFDWDLGIDGVKGVVHVRPPDDLHCLGSAYPAAQFVQKVFVFGTPTNVNFSNISP-PPHPEAYPWASLIPP-

```

61

62

63

64

65

66

67

68

Supplementary Figure 6 Multiple sequence alignment of the fungal CE15-A members, including non-characterized sequences. To further probe the variability among fungal CE15-A enzymes, all sequences assigned to the G1 subgroup in Dilokpimol *et al.*⁵ were included in the alignment. This small fungal subgroup appears highly conserved at sequence level, except for an insert present in the N-terminal region of some members only. Based on the structures of CuGE:ligand complexes it is predicted that the substrate binding site is formed by the catalytic triad (colored triangles) and residues at the positions marked by black triangles. The natural substrate(s) of fungal CE15-enzymes remains speculative.

74 **Supplementary Table 1 – Crystallization and soak conditions for the CuGE structures**

| | apo-dΔS270A | apo-Δ*dWT | dΔS270A:U^{m4}X | dΔS270A:U^{m4}XX^{-OH} | dΔS270A:XU^{m4}XX^{-OH} | Δ*dWT:U^{m4}XX^{-OH} |
|--|---|---|---|---|---|--|
| PDB accession code | 6RTV | 6RU2 | 6RU1 | 6RV7 | 6RV9 | 6RV8 |
| Vapor diffusion setup | Hanging drop | Sitting drop | Hanging drop | Sitting drop | Sitting drop | Sitting drop |
| Protein concentration (mg mL ⁻¹) | 10 | 10 | 5 | 10.9 | 10.9 | 10 |
| Reservoir condition | 0.1 M HEPES pH 7, 0.1 M KCl, 25 % w v ⁻¹ SOKALAN CP7 | 0.1 M ammonium acetate, 0.1 M bis-tris pH 5.5, 17 % w v ⁻¹ PEG10000 | 0.1 M HEPES pH 7, 0.1 M KCl, 25 % w v ⁻¹ SOKALAN CP7 | 0.1 M phosphate- citrate pH 4.2, 0.1 M KCl, 20 % w v ⁻¹ PEG8000 | 0.1 M phosphate- citrate pH 4.2, 0.2 M KCl, 20 % w v ⁻¹ PEG8000 | 0.1 M ammonium acetate, 0.1 M bis-tris pH 5.5, 17 % w v ⁻¹ PEG10000 |
| Drop volume (μL) | 2 | 0.4 | 3 | 0.4 | 0.4 | 0.4 |
| Ratio (protein:reservoir) | 1:1 | 3:1 | 2:1 | 3:1 | 3:1 | 3:1 |
| Temperature (°C) | 4 | 4 | 4 | 4 | 4 | 4 |
| Ligand | | | U ^{m4} X | U ^{m4} XX ^{-OH} | XU ^{m4} XX ^{-OH} | U ^{m4} XX ^{-OH} covalent intermediate |
| Product ID | | | O-AMX | O-UXXR | O-XUXXR | O-UXXR |
| Manufacturer | | | Megazyme | Megazyme | Megazyme | Megazyme |
| Soak condition | | | Ligand stock (25 mg mL ⁻¹) in reservoir solution was added 1:1 to the drop | Ligand stock (252 mM in 0.1 M Na ₃ -citrate pH 5.5, 0.2 M KCl, 20 % w v ⁻¹ PEG8000) was added 7:1 to the drop | Ligand stock (305 mM in 0.1 M Na ₃ -citrate pH 5.5, 0.2 M KCl, 20 % w v ⁻¹ PEG8000) was added 7:1 to the drop | Ligand stock (253 mM) in reservoir solution was added was added 1:1 to the drop |
| Soaking time (min) | | | 60 | 45 | 50 | 60 |
| Cryo condition | Reservoir condition supplemented with 26 % glycerol | Reservoir condition supplemented with 26 % glycerol | Ligand stock supplemented with 26 % glycerol | Ligand stock supplemented with 26 % glycerol | Ligand stock supplemented with 26 % glycerol | Ligand stock supplemented with 26 % glycerol |

75

76

77 **Supplementary Table 2 – Data collection and refinement statistics for the CuGE structures**

| | apo-dΔS270A | apo-Δ*dWT | dΔS270A:U^{m4}X | dΔS270A:U^{m4}XX^{-OH} | dΔS270A:XU^{m4}XX^{-OH} | Δ*dWT:U^{m4}XX^{-OH} |
|---|--|--|--|---|--|---|
| PDB accession code | 6RTV | 6RU2 | 6RU1 | 6RV7 | 6RV9 | 6RV8 |
| Data collection | | | | | | |
| Space group | <i>P</i> 4 ₁ 2 ₁ 2 | <i>P</i> 4 ₁ 2 ₁ 2 | <i>P</i> 4 ₁ 2 ₁ 2 | <i>P</i> 4 ₁ 2 ₁ 2 | <i>P</i> 4 ₁ 2 ₁ 2 | <i>P</i> 4 ₁ 2 ₁ 2 |
| Cell dimensions | | | | | | |
| <i>a</i> = <i>b</i> , <i>c</i> (Å) | 84.27, 261.13 | 84.45, 262.24 | 84.24, 260.82 | 84.78, 260.38 | 84.48, 261.73 | 84.75, 261.06 |
| Resolution (Å) ^a | 44.0 - 1.46 (1.51 - 1.46) | 44.2 - 1.96 (2.03 - 1.96) | 42.1 - 1.39 (1.44 - 1.39) | 49.3 - 1.73 (1.79 - 1.73) | 44.5 - 1.64 (1.70 - 1.64) | 44.1 - 1.85 (1.92 - 1.85) |
| <i>R</i> _{merge} ^b | 0.158 (0.827) | 0.252 (0.852) | 0.103 (0.844) | 0.167 (1.04) | 0.081 (0.867) | 0.106 (0.859) |
| $\langle I / \sigma I \rangle$ | 7.99 (1.90) | 6.64 (1.72) | 11.96 (1.78) | 10.39 (1.56) | 13.85 (1.68) | 28.48 (3.78) |
| Wilson B-value (Å ²) | 11.5 | 20.1 | 12.5 | 20.0 | 23.3 | 24.6 |
| Unique reflections | 163501 (16134) | 68186 (6594) | 188193 (18515) | 99921 (9775) | 116768 (11435) | 80254 (6618) |
| Completeness (%) | 100.0 (100.0) | 98.2 (97.1) | 99.5 (99.3) | 99.8 (100.0) | 99.5 (99.4) | 97.6 (82.0) |
| Redundancy | 9.7 (8.8) | 8.7 (8.5) | 7.4 (7.3) | 7.0 (7.1) | 8.1 (7.6) | 24.9 (19.3) |
| CC _{1/2} | 0.997 (0.555) | 0.989 (0.665) | 0.996 (0.693) | 0.997 (0.675) | 0.998 (0.683) | 0.999 (0.856) |
| Refinement | | | | | | |
| Resolution (Å) | 44.0 - 1.46 | 44.2 - 1.96 | 42.1 - 1.39 | 49.3 - 1.73 | 44.5 - 1.64 | 44.1 - 1.85 |
| No. reflections | 163458 (16133) | 67954 (6564) | 188018 (18507) | 99779 (9773) | 116315 (11427) | 80246 (6618) |
| <i>R</i> _{work} / <i>R</i> _{free} | 0.182/0.199 | 0.178/0.204 | 0.163/0.177 | 0.166/0.194 | 0.186/0.209 | 0.150/0.181 |
| No. atoms | | | | | | |
| Protein | 5936 | 5764 | 5954 | 5780 | 5738 | 5771 |
| Ligand/ion | 79 | 87 | 107 | 92 | 82 | 135 |
| Water | 757 | 542 | 734 | 518 | 527 | 579 |
| <i>B</i> -factors (Å ²) | | | | | | |
| Protein | 13.7 | 20.1 | 14.7 | 22.9 | 31.0 | 25.6 |
| Ligand/ion | 24.6 | 39.0 | 23.7 | 30.5 | 31.9 | 47.1 |
| Water | 24.7 | 26.7 | 25.6 | 29.3 | 33.2 | 34.2 |
| R.m.s. deviations ^c | | | | | | |
| Bond lengths (Å) | 0.01 | 0.01 | 0.01 | 0.03 | 0.03 | 0.03 |
| Bond angles (°) | 1.2 | 1.1 | 1.2 | 1.6 | 1.6 | 1.6 |
| Ramachandran (%) | | | | | | |
| Favored | 96 | 96 | 96 | 96 | 96 | 96 |
| Allowed | 4 | 4 | 4 | 4 | 4 | 4 |
| Outliers | 0 | 0 | 0 | 0 | 0 | 0 |
| Number of TLS groups | 10 | 17 | 11 | 8 | 10 | 13 |

78 One crystal was used for each structure. ^aValues in parentheses are for the highest-resolution shell. ^b $R_{merge} = \sum_{hkl} \sum_{i=1}^n |I_i(hkl) - \bar{I}(hkl)| / \sum_{hkl} \sum_{i=1}^n I_i(hkl)$. ^cR.m.s. deviations from the ideal
79 values reported by Engh & Huber⁷

80

81 **REFERENCES**

- 82 1. Hajizadeh, N. R., Franke, D., Jeffries, C. M. & Svergun, D. I. Consensus Bayesian assessment of
83 protein molecular mass from solution X-ray scattering data. *Sci. Rep.* **8**, 7204 (2018).
- 84 2. Waterhouse, A. *et al.* SWISS-MODEL: homology modelling of protein structures and complexes.
85 *Nucleic Acids Res* **46**, W296–W303 (2018).
- 86 3. Kraulis, J. *et al.* Determination of the three-dimensional solution structure of the C-terminal domain
87 of cellobiohydrolase I from *Trichoderma reesei*. A study using nuclear magnetic resonance and
88 hybrid distance geometry-dynamical simulated annealing. *Biochemistry* **28**, 7241–7257 (1989).
- 89 4. De Santi, C., Gani, O. A., Helland, R. & Williamson, A. Structural insight into a CE15 esterase from the
90 marine bacterial metagenome. *Sci. Rep.* **7**, 17278 (2017).
- 91 5. Dilokpimol, A. *et al.* Fungal glucuronoyl esterases: Genome mining based enzyme discovery and
92 biochemical characterization. *N. Biotechnol.* **40**, 282–287 (2018).
- 93 6. Mosbech, C., Holck, J., Meyer, A. & Agger, J. W. Enzyme kinetics of fungal glucuronoyl esterases on
94 natural lignin-carbohydrate complexes. *Appl. Microbiol. Biotechnol.* **103**, 4065–4075 (2019).
- 95 7. Engh, R. A. & Huber, R. Accurate bond and angle parameters for X-ray protein structure refinement.
96 *Acta Crystallogr. Sect. A* **47**, 392–400 (1991).

97