

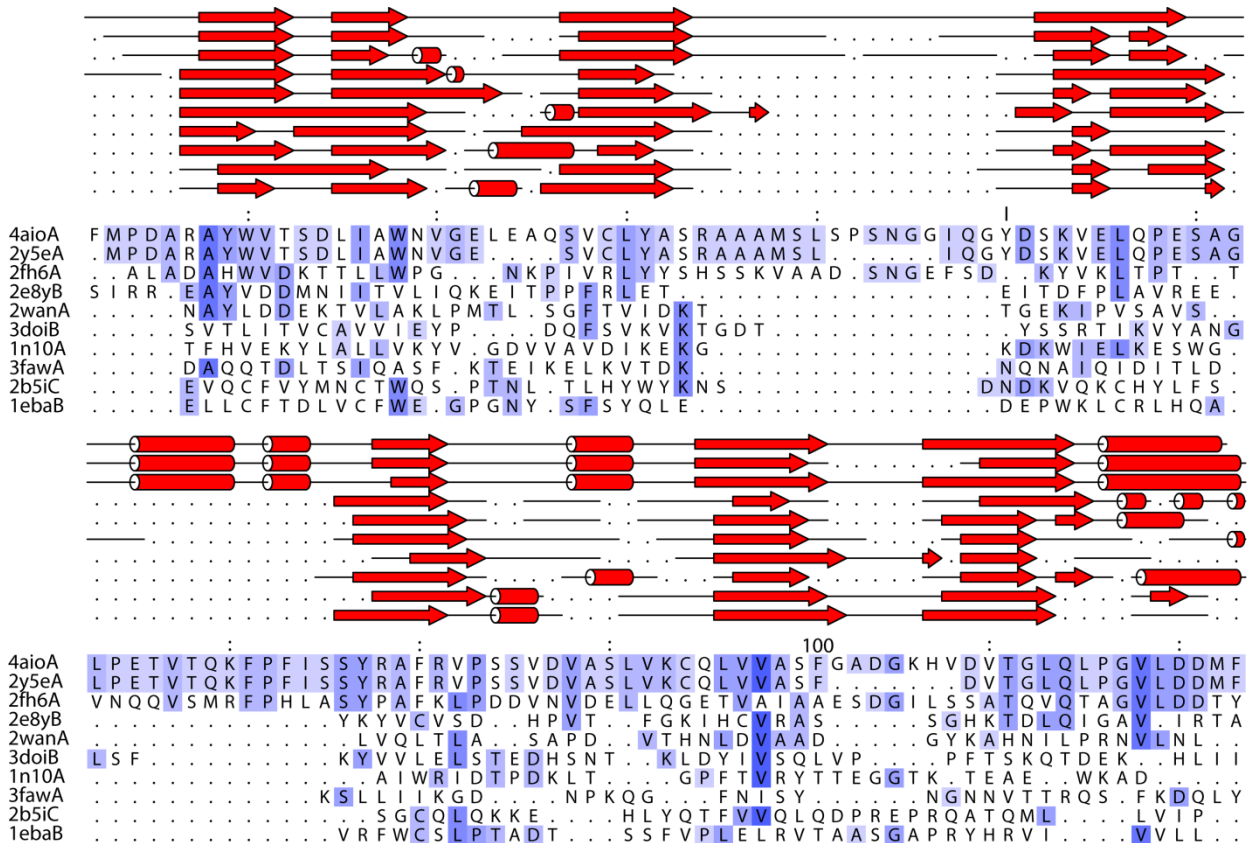
SUPPLEMENTARY MATERIAL

Native structure of the starch debranching enzyme barley limit dextrinase

Marie Sofie Møller^{ab}, Maher Abou Hachem^a, Birte Svensson^{a*} and Anette Henriksen^{b*}

^aEnzyme and Protein Chemistry, Department of Systems Biology, Technical University of Denmark, Søtofts Plads, Building 224, Kgs. Lyngby, DK-2800, Denmark, and ^bProtein Chemistry Group, Carlsberg Laboratory, Gamle Carlsberg Vej 10, Valby, DK-2500, Denmark

Correspondence email: bis@bio.dtu.dk; athx@novonordisk.com

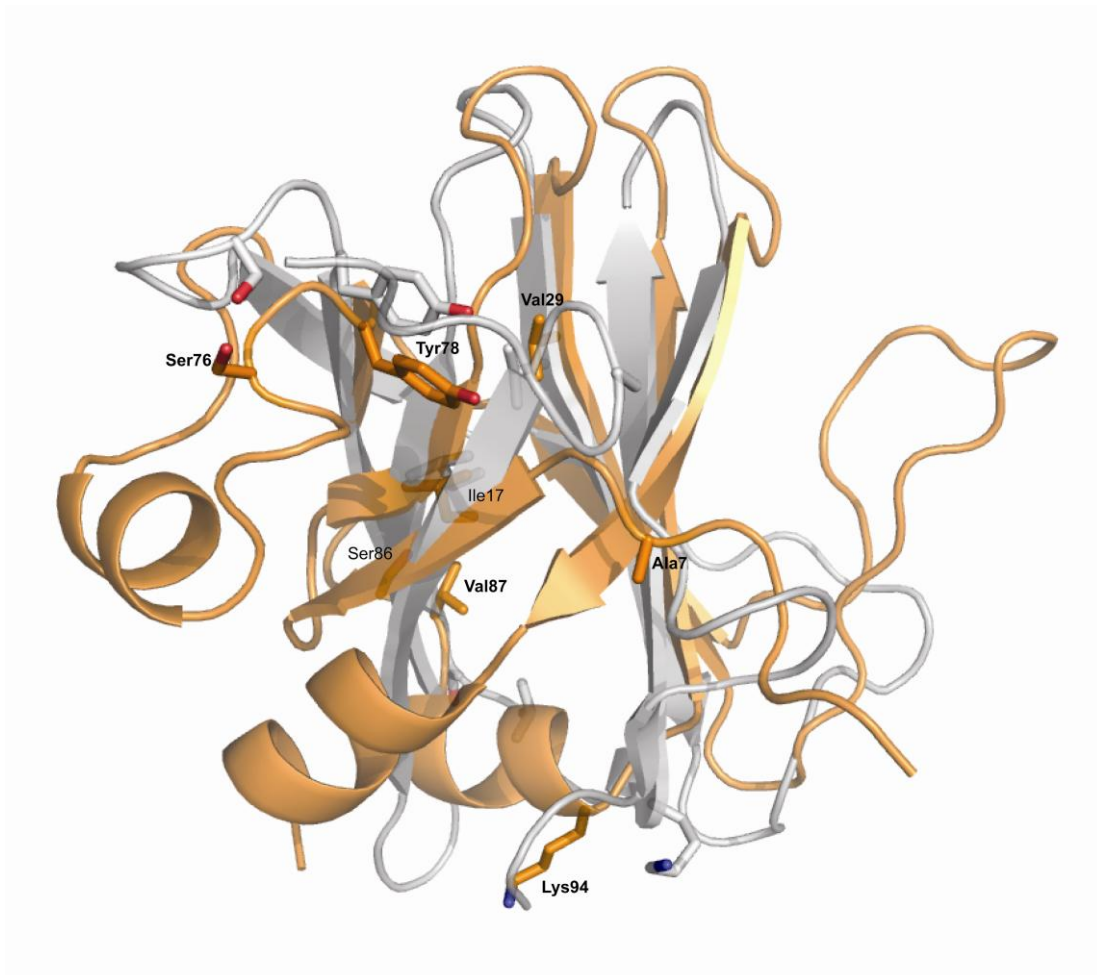


<i>PDB code</i>	2fh6	2e8y	2wan	3doi	1n10	1eba	3faw	2b5i	2vxq
2y5e	31	14	16	9	10	13	7	7	11
2fh6		16	15	2	13	7	14	12	10
2e8y			11	5	-	-	7	8	-
2wan				16	-	7	11	-	10
3doi					11	9	9	8	-
1n10						10	9	9	42
1eba							-	15	-
3faw								-	-
2b5i									5

Supplementary Fig. 1 Top: Structural alignment of domains with Z-score ≥ 5 in a DALI search (Holm & Rosenstrom, 2010) with the *HvLD* N-terminal domain (4aioA). The figure is generated using ALINE (Bond & Schuettelkopf, 2009) based on output from DALI. Conserved residues are highlighted in blue colours. The upper panel in the alignment show the secondary structure of the PDB entries in the alignment. Bottom: Pairwise sequence identity scores from a structural alignment between the protein domains identified in the DALI search.

LD: 6 D A R A Y W V T S ----- D L I A W N V G E L E A Q S V C L Y A S --- R A A A M S L S P S N G G I Q G Y D S K V E L Q P E S A G L P
 1111111111 111111111111 11111111 1111 1111111111111111
 2djm: 6 S A S V Q L D S Y N Y D G S T F S G K I Y V K **N I A Y S K** - K V T V V Y A D G S D N W N ----- N N G N I I A A S **F S G P I S G**

LD: E T V T Q K F P F I **S S Y** R A F R V P S **S V D** V A S L V **K C Q L** V V A S F G A A G G A V D V T G L Q L P G V L D D M 123
 11111111111111 1 1111111111 11111111 1
 2djm: ----- **S N Y E Y W T F S A S V K G I** --- **K E F Y I K Y E V S** ----- G K T Y **Y D N N N S A N Y Q V S** 105



Supplementary Fig. 2 Top: FATCAT alignment of the LD N-terminal domain and the N-terminal starch binding domain (CBM21) of glucoamylase from *Rhizopus oryzae* (2djm). The residues in red are identical residues between LD and the CBM21. The residues in bold are residues involved in ligand binding of the CMB21 (Tung *et al.*, 2008). Bottom: A superposition of the *HmLD* N-terminal domain (orange) and the N-terminal starch binding domain (grey) with conserved residues from the alignment shown as sticks.

Supplementary Table 1 Unique^a domains structurally similar to the LD N-terminal domain. The domains are identified by a DALI structural search^b (http://ekhidna.biocenter.helsinki.fi/dali_server/)

PDB entry	Organism	Protein activity	DALI Z-score	Rmsd	Alignment length	Sequence identity (%)	Reference	Possible function of aligned sequence
2Y5E <i>α-CD ligand</i>	<i>Hordeum vulgare</i>	Limit dextrinase	19.9	1.0	104	100	(Vester-Christensen <i>et al.</i> , 2010)	Whole domain - unknown function
2FH6 <i>Glucose ligand</i>	<i>Klebsiella aerogenes</i>	Pullulanase	16.7	1.8	112	31	(Mikami <i>et al.</i> , 2006)	Domain core - unknown function
2E8Y	<i>Bacillus subtilis</i> str. 168	Pullulanase	7.3	2.8	82	13	Malle <i>et al.</i> , unpublished	Domain core - unknown function
2WAN	<i>Bacillus acidopullulyticus</i>	Pullulanase	6.6	2.4	78	14	(Turkenburg <i>et al.</i> , 2009)	Domain core - unknown function.
3DOI	<i>Thermotoga maritime</i>	Esterase with preference for (C2–C10) esters	5.1	3.2	82	9	(Levisson <i>et al.</i> , 2009)	Domain core – Multimerisation of hexamer
1N10	<i>Phleum pratense</i>	Unknown, Grass pollen allergen Phl P1	5.1	2.6	73	10	Fedorov <i>et al.</i> , unpublished	Domain core - unknown function
1EBA	<i>Homo sapiens</i>	Peptide-EPO receptor	5.1	3.3	77	13	(Livnah <i>et al.</i> , 1998)	Peptide ligand binding
3FAW	Group B <i>Streptococcus agalactiae</i>	Pullulanase	5.0	2.7	72	7	(Gourlay <i>et al.</i> , 2009)	Domain core - unknown function
2B5I	<i>Homo sapiens</i>	Cytokine receptor gamma chain	5.0	3.0	76	7	(Wang <i>et al.</i> , 2005)	Domain core – NAG binding

^aEntries with more than 90% identical amino acid sequence are only included with one representative entry.

^bOnly entries with a Dali Z-score ≥ 5.0 are included.

References

Bond, C. S. & Schuettelkopf, A. W. (2009). *Acta Cryst. D* **65**, 510–512.

Gourlay, L. J., Santi, I., Pezzicoli, A., Grandi, G., Soriani, M. & Bolognesi, M. (2009). *J. Bacteriol.* **191**, 3544–3552.

Holm, L. & Rosenstrom, P. (2010). *Nucleic Acids Res.* **38**, W545–W549.

Levisson, M., Sun, L., Hendriks, S., Swinkels, P., Akveld, T., Bultema, J. B., Barendregt, A., van den Heuvel, R. H. H., Dijkstra, B. W., van der Oost, J. & Kengen, S. W. M. (2009). *J. Mol. Biol.* **385**, 949–962.

Livnah, O., Johnson, D. L., Stura, E. A., Farrell, F. X., Barbone, F. P., You, Y., Liu, K. D., Goldsmith, M. A., He, W., Krause, C. D., Pestka, S., Jolliffe, L. K. & Wilson, I. A. (1998). *Nat. Struct. Biol.* **5**, 993–1004.

Mikami, B., Iwamoto, H., Malle, D., Yoon, H. J., Demirkan-Sarikaya, E., Mezaki, Y. & Katsuya, Y. (2006). *J. Mol. Biol.* **359**, 690–707.

Padavattan, S., Flicker, S., Schirmer, T., Madritsch, C., Randow, S., Reese, G., Vieths, S., Lupinek, C., Ebner, C., Valenta, R. & Markovic-Housley, Z. (2009). *J. Immunol.* **182**, 2141–2151.

Tung, J., Chang, M. D., Chou, W., Liu, Y., Yeh, Y., Chang, F., Lin, S., Qui, Z. & Sun, Y. (2008). *Biochem. J.* **416**, 27–36.

Turkenburg, J. P., Brzozowski, A. M., Svendsen, A., Borchert, T. V., Davies, G. J. & Wilson, K. S. (2009). *Proteins Struct. Funct. Bioinf.* **76**, 516–519.

Vester-Christensen, M. B., Abou Hachem, M., Svensson, B. & Henriksen, A. (2010). *J. Mol. Biol.* **403**, 739–750.

Wang, X. Q., Rickert, M. & Garcia, K. C. (2005). *Science.* **310**, 1159–1163.