Biochem. J. (1989) 264, 309–311 (Printed in Great Britain) 

BIOCHEMICAL JOURNAL LETTERS

## Sequence homology between putative raw-starch binding domains from different starch-degrading enzymes

Primary structures of a wide variety of starch-degrading enzymes have been reported over the past few years. In general, they exhibit little or no homology but active site similarities have been indicated in the cases of α-amylases, cyclodextrin glucanotransferases (CGTases), debranching enzymes, exo-α-amylases, glucoamylases and α-glucosidases [1–6]. A branching enzyme and an amylomaltase have very distantly related sequences [1]. Some of the enzymes may have only discrete areas in common [1] whereas others are predicted to contain an α/β-barrel catalytic domain [1,2] similar to the three-dimensional structure of α-amylases [7–9]. The present communication addresses another common feature consisting of a terminal sequence motif (Fig. 1). It was earlier observed in part at the *C*-terminus of *Aspergillus niger* glucoamylase and *Bacillus macerans* CGTase [1] and has now been recognized in extended form in these enzymes, as well as four additional CGTases [10–13], two highly homologous α-amylases [14,15], two exo-α-amylases [5,6], and a β-amylase [16]. This motif is the first example of a clear homology between β-amylase and other starch-degrading enzymes [1,17]. Moreover, both endo- and exo-acting enzymes as well as enzymes that either retain or invert the configuration of the product are represented. The *N*-terminus of *Rhizopus oryzae* glucoamylase [18], which presents less similarity, is included (Fig. 1). For both of the glucoamylases the motif region has been demonstrated to be required for adsorption onto and degradation of granular starch [19–22].

The *Streptomyces* α-amylases readily align with both the catalytic and *C*-terminal (γ-crystallin-type) domains described in two tertiary structures of α-amylase [7,8,14,15,23]. The *Streptomyces* enzymes are, however, further extended at the *C*-terminal end by the present motif that is linked to the predicted γ-crystallin-type domain by a short sequence rich in glycine. This proposed domain organization is consistent with hydrophobic cluster analysis [23]. Despite exceptionally long *C*-terminal segments, two *B. subtilis* α-amylases [14,23–25] do not contain the motif. Searching the PIR data base of protein sequences (National Biomedical Research Foun-

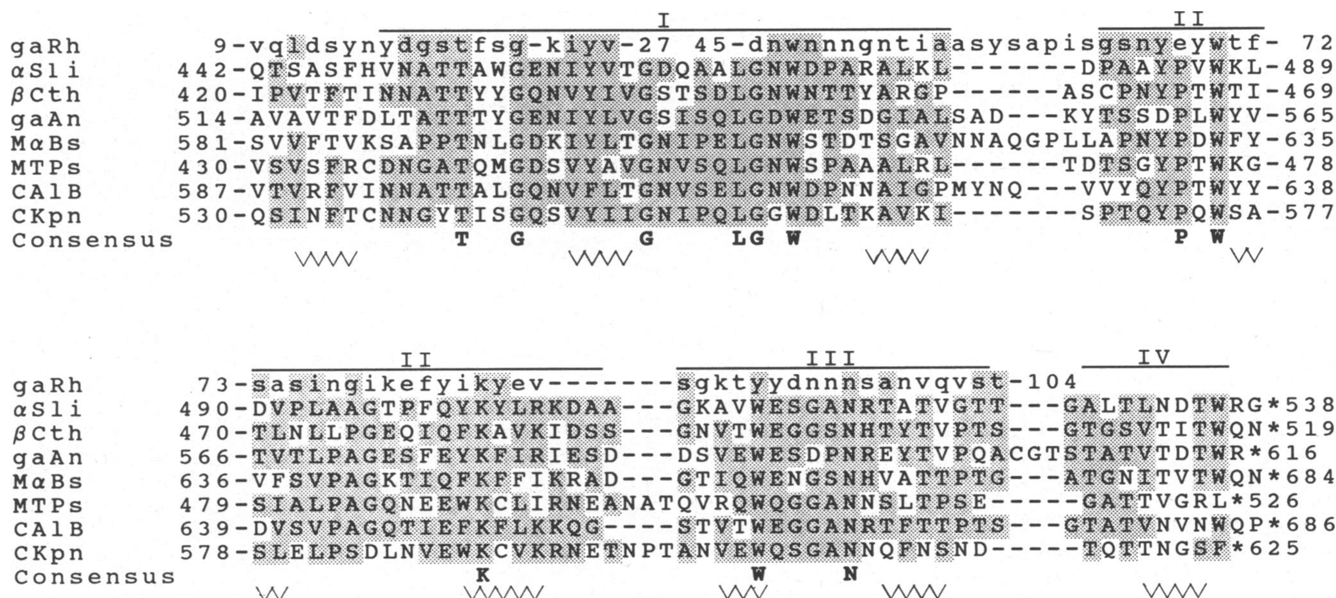


**Fig. 1. Comparison of *C*-terminal sequences of amylases and cyclodextrin glucanotransferases**

Residues from dominant exchange groups [39] are shaded. Gaps are indicated by dashes. ⋁⋁ indicates predicted β-sheet structure. Possible errors in predicted positions are ±2. Asterisks signify the *C*-terminal residues of the protein. αSli, α-amylase from *Streptomyces limosus* [14]; βCth, β-amylase from *Clostridium thermosulfurogenes* [16]; gaAn, glucoamylase from *Aspergillus niger* [20]; MαBs, maltogenic α-amylase from *Bacillus stearothermophilus* [5]; MTPs, maltotetraose-forming amylase from *Pseudomonas stutzeri* [6]; CA1B, CGTase from alkalophilic *Bacillus* sp. strain 1011 [12]; CKpn, CGTase from *Klebsiella pneumoniae* [13]. The *N*-terminal starch-binding region from *Rhizopus oryzae* glucoamylase [18] (gaRh) is shown in lower case.

dation, Washington, DC, U.S.A.) with the more distantly related *R. oryzae* sequence (residues 9–104) interestingly indicated similarity to the 75 *C*-terminal residues in barley α-amylase [26]. This result raises the question whether the cereal α-amylases, predicted to contain unusually short *C*-terminal regions of about 75 residues in length [9,23] actually conform to the γ-crystallin-type fold seen in two α-amylases [7,8] or whether they have a domain of the present type directly attached to the α/β-barrel.

The compared *C*-terminal parts consist of 96 to 103 amino acid residues containing four areas of sequence similarity (I–IV in Fig. 1) separated by short segments varying in length. Pairwise comparison of these sequences revealed 21–44% invariant residues. Well-established methods of structure prediction [27,28] indicate the presence of eight β-sheet elements which agrees with β-sheet structure earlier predicted for the *N*-terminal sequence of *Rhizopus* glucoamylase [29]. The following residues occur in all eight sequences: Thr-453, Gly-456, Trp-471, Trp-487, Lys-502 and Asn-519 (*S. limosus* α-amylase numbering). In addition, Glu-463, Leu-468, Gly-469, Pro-485 and Trp-514 are invariant among the *C*-terminal structures (Fig. 1). The conserved tryptophans are likely to participate in binding of glucosyl residues through hydrophobic or hydrogen bonding forces, as discussed recently for protein–carbohydrate complexes [30,31]. The glycines and proline may be critical for the conformation of the polypeptide chain, while threonine, lysine, asparagine and leucine could play a similar role and/or interact with starch.

We hypothesize that this conserved region constitutes a self-contained domain which is involved in binding granular starch since that function had already been postulated for the two glucoamylase sequences [19–22] in the comparison. A form of *A. niger* glucoamylase lacking the motif region is unable to adsorb onto and degrade raw starch. Moreover, a fusion protein of *E. coli* β-galactosidase and residues 484–616 from this glucoamylase absorbed onto starch [32]. Finally, the ability to bind starch was selectively destroyed without loss of catalytic activity by *N*-bromosucciminide oxidation of Trp-590 and Trp-615 in glucoamylase from *A. niger* (Trp-514 and Trp-536 in *S. limosus* α-amylase) in the presence of the inhibitor acarbose. Trp-543 and Trp-563 (Trp-471 and Trp-487 in *S. limosus* α-amylase), however, were not modified [33]. Since the *Rhizopus* glucoamylase lacks a part equivalent to region IV of the motif, the conserved aromatic group in region III (Trp-514 in *S. limosus* α-amylase) is speculated to be important for the interaction with granular starch. In addition to the glucoamylases, several CGTases [34,35] and the β-amylase from *Clostridium thermosulfurogenes* [36] are reported to have high affinity for raw starch.

Given the differences in function, sequence, and size represented by the investigated starch-degrading enzymes, it is remarkable how similar the terminal regions are with respect to size and sequence. They may constitute modern derivatives of a domain that could either have been attached to or removed from proteins during the course of evolution. Since most α-amylases do not have this domain, but can still attack granular starch [37], it is not universally required for degradation of raw starch. Preliminary X-ray diffraction analysis of *B. stearothermophilus* CGTase has been reported [38]. Therefore, a solved three-dimensional structure of the domain should provide further insight into the significance of the consensus residues indicated in the sequence comparison (Fig. 1).

Birte SVENSSON,* Hans JESPERSEN,*
Michael R. SIERKS* and E. Ann MacGREGOR†

*Department of Chemistry, Carlsberg Laboratory, Gamle Carlsberg Vej 10, DK-2500 Copenhagen Valby, Denmark, and †Department of Chemistry, University of Manitoba, Winnipeg, Manitoba, Canada R3T 2N2

1. Svensson, B. (1988) FEBS Lett. **230**, 72–76
2. MacGregor, E. A. & Svensson, B. (1989) Biochem. J. **259**, 145–152
3. Amemura, A., Chakraborty, R., Fujita, M., Noumi, I. & Futai, M. (1988) J. Biol. Chem. **263**, 9271–9275
4. Katsuragi, N., Takiozawa, N. & Murooka, Y. (1987) J. Bacteriol **169**, 2301–2306
5. Diderichsen, B. & Christiansen, L. (1988) FEMS Microbiol. Lett. **56**, 53–60
6. Fujita, M., Torigoe, K., Nakada, T., Tsusaki, K., Kubota, M., Sakai, S. & Tsujisaka, Y. (1989) J. Bacteriol. **171**, 1333–1339
7. Buisson, G., Duée, E., Haser, R. & Payan, F. (1987) EMBO J. **6**, 3908–3916
8. Matsuura, Y., Kusunoki, M., Harada, W. & Kakudo, M. (1984) J. Biochem. (Tokyo) **95**, 697–702
9. MacGregor, E. A. (1988). J. Protein Chem. **7**, 399–415
10. Takano, T., Fukuda, M., Monma, M., Kobayashi, S., Kainuma, K. & Yamane, K. (1986) J. Bacteriol. **166**, 1118–1122
11. Sakai, S., Kubota, M., Yamamoto, K., Nakada, T., Torigoe, K., Ando, O. & Sugimoto, T. (1987) J. Jpn. Soc. Starch Sci. **34**, 140–147
12. Kimura, K., Katoka, S., Ishii, Y., Takano, T. & Yamane, K. (1987) J. Bacteriol. **169**, 4399–4402
13. Binder, F., Huber, O. & Böck, A. (1986) Gene **47**, 269–277
14. Long, C. M., Virolle, M.-J., Chang, S.-Y., Chang, S. & Bibb, M. J. (1987) J. Bacteriol. **169**, 5745–4754
15. Virolle, M.-J., Long, C. M., Chang, S. & Bibb, M. J. (1988) Gene **74**, 321–334
16. Kitamoto, N., Yamagata, H., Kato, T., Tsukagoshi, N. & Udaka, S. (1988) J. Bacteriol. **170**, 5848–5854
17. Friedberg, F. & Rhodes, C. (1988) Protein Seq. Data Anal. **1**, 499–501
18. Ashikari, T., Nakamura, N., Tanaka, Y., Kiuchi, N., Shibano, Y., Tanaka, T., Amachi, T. & Yoshizumi, H. (1986) Agric. Biol. Chem. **50**, 957–964
19. Svensson, B., Pedersen, T. G., Svendsen, I., Sakai, T. & Ottesen, M. (1982) Carlsberg Res. Commun. **47**, 55–69
20. Svensson, B., Larsen, K., Svendsen, I. & Boel, E. (1983) Carlsberg Res. Commun. **48**, 529–544
21. Takahashi, T., Kato, K., Ikegami, Y. & Irie, M. (1985) J. Biochem. (Tokyo) **98**, 663–671
22. Svensson, B., Larsen, K. & Gunnarsson A. (1986) Eur. J. Biochem. **154**, 497–502
23. Raimbaud, E., Buléon, A., Pérez, S. & Henrissat, B. (1989) Int. J. Biol. Macromol., in the press
24. Yang, M., Galizzi, A. & Henner, D. (1983) Nucleic Acids Res. **11**, 237–249
25. Yamasaki, H., Ohmura, K., Nakayama, A., Takeichi, Y., Otozai, K., Yamasaki, K., Tamura, G. & Yamane, K. (1983) J. Bacteriol. **156**, 327–337
26. Rogers, J. C. & Milliman, C. (1983) J. Biol. Chem. **258**, 8169–8174
27. Garnier, J., Osguthorpe, D. J. & Robson, B. (1978) J. Mol. Biol. **120**, 97–120

28. Cid, H., Bunster, M., Arriagada, E. & Campos, M. (1982) FEBS Lett. **150**, 247–254
29. Tanaka, Y., Ashikari, T., Nakamura, N., Kiuchi, N., Shibano, Y., Amachi, T. & Yoshizumi, H. (1986) Agric. Biol. Chem. **50**, 965–969
30. Quiocho, F. A. (1986) Annu. Rev. Biochem. **55**, 287–315
31. Johnson, L. N., Cheetham, J., McLaughlin, P. J., Acharya, K. R., Barford, D. & Phillips, D. C. (1988) Curr. Top. Microbiol. Immunol. **139**, 81–134
32. Chen. L., Ford, C., Nikolov, Z., Reilly, P. & Svensson, B. (1988) Proc. Int. Congr. 16th (Toronto), abstr. 35.41.12
33. Svensson, B., Clarke, A. J. & Svendsen, I. (1986) Carlsberg Res. Commun. **51**, 61–73
34. Stavn, A. & Granum, P. E. (1979) Carbohydr. Res. **75**, 243–250
35. Kato, Y., Mikuni, K., Hara, K., Hashimoto, H., Nakajima, T., Kobayashi, S. & Kainuma, K. (1988) J. Ferment. Technol. **66**, 159–166
36. Saha, B. C., Shen, G.-J. & Zeikus, J. G. (1987) Enzyme Microb. Technol. **9**, 598–601
37. Sandstedt, R. M. & Ueda, S. (1969) J. Jpn. Soc. Starch Sci. **17**, 215–228
38. Kubota, M., Mikami, B., Tsujisaka, Y. & Morita, Y. (1988) J. Biochem. (Tokyo) **104**, 12–14
39. Dayoff, M. O., Barker, W. C. & Hunt, L. C. (1983) Methods Enzymol. **91**, 524–545

Received 24 July 1989