Volume 230, number 1,2, 72–76 FEB 05692 March 1988

# Regional distant sequence homology between amylases, α-glucosidases and transglucanosylases

Birte Svensson

*Department of Chemistry, Carlsberg Laboratory, Gamle Carlsberg Vej 10, DK-2500 Copenhagen Valby, Denmark*

Received 25 January 1988

Amylases possess short, conserved regions near functional side chains. Sequence comparison extends this relationship to comprise a maltase and a cyclodextrin glucanotransferase. Similarity also exists with intestinal sucrase-isomaltase and fungal glucoamylase near identified essential carboxyl groups. Homology between COOH-terminal regions of glucoamylase and cyclodextrin glucanotranserase may indicate raw-starch binding areas. It is suggested that amylases, α-glucosidases, and transglucanosylases acting on 1,4- and 1,6-α-glucosidic linkages share key structural features in the active centres.

Catalytic carboxyl group; Substrate-binding residue; Segment alignment score; α/β-Barrel

## 1. INTRODUCTION

Functional residues can unveil similarities between proteins of no conspicuous structural resemblance, just as aligning distantly related sequences can focus on groups important for function and structure [1,2]. The essential Trp 120 of *Aspergillus niger* glucoamylase [3] thus disclosed a weak homology near Trp 83 in the active site of Taka-amylase A, referred to below as TAA [4,5]. α-Amylases align in homologous regions [5–8] constituting in total only 4–7% of the sequences and the majority of the side chains proposed to interact with amylose are at or near COOH-terminal ends of conserved strands of the TAA α/β-barrel [4,9]. Based on homology pointed out by functional residues it is suggested in the present communication that amylases, α-glucosidases, and transglucanosylases are related with respect to active site structures.

*Correspondence address:* B. Svensson, Department of Chemistry, Carlsberg Laboratory, Gamle Carlsberg Vej 10, DK-2500 Copenhagen Valby, Denmark

## 2. MATERIALS AND METHODS

The program RELATE and the Mútation Data Matrix were used to screen for homology by comparing either an entire protein (table 1) with selected TAA regions or all segments of 8 residues in two proteins. Scores were calculated with the program ALIGN using a matrix bias of 2 and a gap penalty of 8 ([10] and National Biochemical Research Foundation, Georgetown, Washington, DC).

## 3. RESULTS AND DISCUSSION

α-Amylases sequenced recently possess the 5 previously described conserved regions [5–8] (fig.1A–E). A sixth region (fig.1F) includes the essential Trp 120 in *A. niger* glucoamylase (gaAn) [3] and Taka-amylase A (TAA) Trp 83, located in binding subsite 3 ($B_3$) [4]. Subsite numbers [4] refer to the association of 7 consecutive glucosyl residues of amylose with $B_1$–$B_7$. Cleavage takes place between $B_4$ and $B_5$.

Affinity labelling of sucrase-isomaltase (S-I) by conduritol B-epoxide identified the catalytic -$COO^-$ groups Asp 505 and Asp 1394 [11–13]. In a fungal glucoamylase (gaAn) Asp 176, Glu 179 and Glu 180 were protected by a pseudo-



*Published by Elsevier Science Publishers B.V. (Biomedical Division)*
00145793/88/$3.50 

Table 1

Amylases, α-glucosidases and transglucanosylases compared

| Code | Enzyme | EC no. | Source | Reference |
|---|---|---|---|---|
| aBli | α-amylase | 3.2.1.1 | *B. licheniformis* | [a] |
| aBst | α-amylase | 3.2.1.1 | *B. stearothermophilus* | [b] |
| aBst′ | α-amylase | 3.2.1.1 | *B. stearothermophilus* | [c] |
| aBam | α-amylase | 3.2.1.1 | *B. amyloliquefaciens* | [d] |
| aBsu | α-amylase | 3.2.1.1 | *B. subtilis* | [e,f] |
| aShy | α-amylase | 3.2.1.1 | *Streptomyces hygroscopicus* | [g] |
| aB1 | α-amylase | 3.2.1.1 | barley isozyme 1 | [h] |
| aB2 | α-amylase | 3.2.1.1 | barley isozyme 2 | [i] |
| aDm | α-amylase | 3.2.1.1 | *Drosophila melanogaster* | [j] |
| PPA | α-amylase | 3.2.1.1 | pig pancreas (rat and mouse) | [r–n] |
| HAP/S | α-amylase | 3.2.1.1 | human pancreas/saliva | [o] |
| TAA | α-amylase | 3.2.1.1 | *Aspergillus oryzae* (Taka-amylase A) | [p] |
| M | maltase | 3.2.1.20 | *Saccharomyces cerevisiae* | [q] |
| I | isomaltase | 3.2.1.10 | rabbit intestine | [r] |
| S | sucrase | 3.2.1.48 | rabbit intestine | [r] |
| gaAn | glucoamylase | 3.2.1.3 | *Aspergillus niger* | [s,t] |
| gaRh | glucoamylase | 3.2.1.3 | *Rhizopus oryzae* | [u] |
| gaSd | glucoamylase | 3.2.1.3 | *Saccharomyces cerevisiae (diastaticus)* | [v] |
| CGT | cyclodextrin glucanotransferase | 2.4.1.19 | *B. macerans* | [x] |
| AM | amylomaltase | 2.4.1.25 | *Streptococcus pneumoniae* | [y] |
| BE | branching enzyme | 2.4.1.18 | *E. coli* | [z] |

[a] Yuuki, T. et al. (1985) J. Biochem. 98, 1147–1156. [b] Nakajima, R. et al. (1985) J. Bacteriol. 163, 401–406. [c] Ihara, H. et al. (1985) J. Biochem. 98, 95–103. [d] Takkinen, K. et al. (1983) J. Biol. Chem. 258, 1007–1013. [e] Yang, M. et al. (1983) Nucleic Acids Res. 11, 237–249. [f] Yamazaki, H. et al. (1983) J. Bacteriol. 156, 327–337. [g] Hoshiko, S. et al. (1987) J. Bacteriol. 169, 1029–1036. [h] Rogers, J.C. and Milliman, C. (1983) J. Biol. Chem. 258, 8169–8174. [i] Rogers, J.C. (1985) J. Biol. Chem. 260, 3731–3738. [j] Boer, P.H. and Hickey, D.A. (1986) Nucleic Acids Res. 14, 8399–8411. [j] Kluh, I. (1981) FEBS Lett. 136, 231–234. [l] Pasero, L. et al. (1986) Biochim. Biophys. Acta 869, 147–157. [m] MacDonald, R.J. et al. (1980) Nature 287, 117–122. [n] Hagenbüchle, O. et al. (1980) Cell 21, 179–187. [o] Nakamura, Y. et al. (1984) Gene 28, 263–270. [p] [5]. [q] Hong, S.H. and Marmur, J. (1986) Gene 41, 75–84. [r] [11]. [s] Svensson, B. et al. (1983) Carlsberg Res. Commun. 48, 529–544. [t] Boel, E. et al. (1984) EMBO J. 3, 1097–1102. [u] Ashikari, T. et al. (1986) Agric. Biol. Chem. 50, 957–964. [v] Yamashita, I. et al. (1985) J. Bacteriol. 161, 567–573. [x] Takano, T. et al. (1986) J. Bacteriol. 166, 1118–1122. [y] Lacks, S.A. et al. (1982) Cell 31, 327–336. [z] Baecker, P.A. et al. (1986) J. Biol. Chem. 261, 8738–8743

tetrasaccharide inhibitor, acarbose, against inactivating modification by carbodiimide [14]. These results guided the local alignment of gaAn and S-I (fig.1A). α-Amylases, α-glucosidases and two transglucanosylases show good homology in the same region and in spite of the modest scores obtained for glucoamylase it seems possible that all of the listed starch- and oligosaccharide-degrading enzymes are structurally related. In the 3-dimensional (3-D) structure of porcine pancreatic α-amylase (PPA) Asp 197 and Asp 300 have been identified as catalytic residues [9] in accordance with the present alignments (fig.1A,B). TAA Asp 206, Glu 230 and Asp 297 are all close to the scissile 1,4-α-glucosidic bond, however, Glu 230 and Asp 297 were suggested as performing the catalysis in that enzyme [4]. Although carbodiimides react with COOH groups [15], the alignments lead one to infer that similar stereochemical relationships of catalytic groups to substrate exist in the different enzymes. It is assumed, therefore, that gaAn Glu 180, like the homologous aspartates of S-I, acts as a base in the catalysis and mechanisms have been proposed [16,17] which are consistent with the retention and inversion of the product configuration as found for isomaltase [18] and glucoamylase [19].

Clear homology exists between α-amylases,

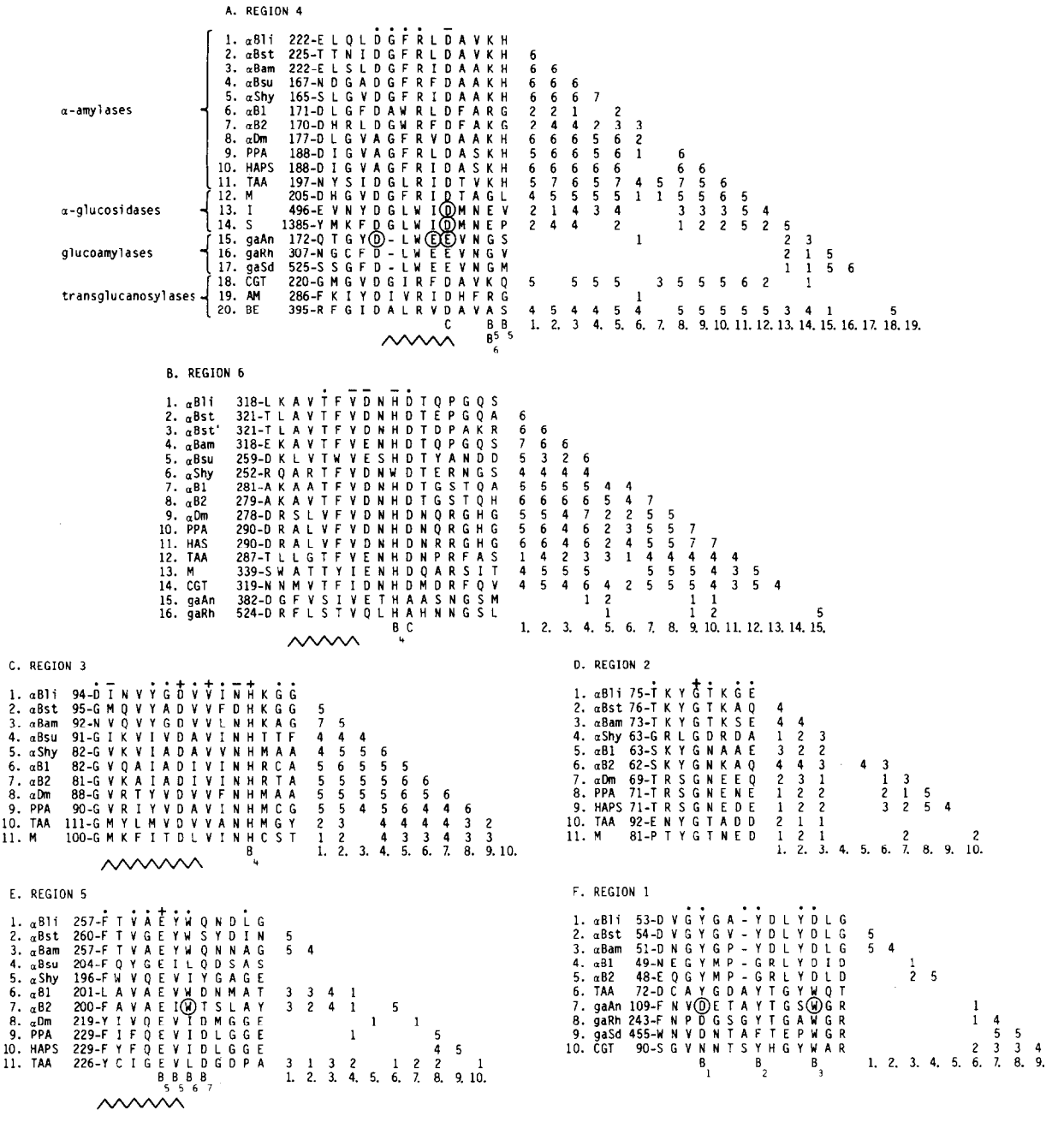


Fig.1. (A–F) Alignment of regions 1–6 from amylases, $\alpha$-glucosidases and transglucanosylases. Where possible the residue numbers refer to mature proteins. Encircled residues have been identified experimentally as interacting with inhibitors. $B_1$–$B_7$ signify residues proposed to participate in substrate binding in TAA subsites 1–7 [4]. C indicates proposed catalytic residues. $\beta$-Strands of the TAA $\alpha/\beta$-barrel [4] are indicated by zigzag lines. (+) Invariant positions; (−, ·) positions occupied by residues from one and two exchange groups [10], respectively. Scores ($s$) of pairwise alignments in SD units ($\sigma$) are grouped: $1 = 2.3\sigma < s \leq 2.8\sigma$; $2 = 2.8\sigma < s \leq 3.3\sigma$; $3 = 3.3\sigma < s \leq 3.8\sigma$; $4 = 3.8\sigma < s \leq 4.8\sigma$; $5 = 4.8\sigma < s \leq 6.8\sigma$; $6 = 6.8\sigma < s \leq 8.8\sigma$; $7 = 8.8\sigma < s \leq 11\sigma$. The corresponding probabilities are given in [10]. The calculations were produced with segments corresponding to TAA: A = 197–210; B = 287–302; C = 111–124; D = 92–99; E = 226–235; F = 72–85. A segment is omitted when all $s \leq 2.3\sigma$. PPA represents mammalian $\alpha$-amylases with the following exceptions: rat and mouse salivary enzymes have Thr 75, rat and mouse pancreas enzymes have Ala 199, and HAP has Arg 302. $\alpha$Bsu Ser 105 is reported in the reference cited in table 1, footnote e, and PPA Asp 77 and Ile 196 in those cited in footnote k.

maltase (M) and cyclodextrin glucanotransferase (CGT) in region 6 (fig.1B) comprising an indicated catalytic aspartic acid residue [4,9]. Glucoamylases lack this acid group and seem distantly related. Regions 3 (fig.1C) and 2 (fig.1D) are recognized in $\alpha$-amylases and maltase. Region 3 possesses an invariant histidyl residue that participates in binding of substrate [4,9]. The conserved PPA Asn 100 is liganded to $Ca^{2+}$ in the 3-D structure [9].

The first and fourth residues in TAA (Glu 230-Asp 233) and PPA (Glu 233-Asp 236) (fig.1E) are thought to hydrogen bond substrate in $B_5$ and $B_7$ and the adjacent bulky side chains to form van der Waals' contacts in $B_5$ and $B_6$ [4,9]. Trp 206 of barley $\alpha$-amylase 2 ($\alpha$B2) has been confirmed to participate in binding, since the inhibitor aplanin hinders inactivating chemical modification of this residue [20].

TAA Asp 72-Thr 85 forms a loop with aromatic residues in $B_1$–$B_3$ between a strand and the following helix of the $\alpha/\beta$-barrel [4] (fig.1F). The absence of homology for some $\alpha$-amylases might reflect that spatial constraints are smaller in region 1 than for the $\beta$-strand containing regions 3–6 (fig.1). In the crystal structure of PPA, Trp 58, Trp 59 and Tyr 62 do participate in substrate binding [9] and match binding residues identified in region 1 of TAA [4] and gaAn [3,14]. Despite this, common role alignment scores, however, only indicate local homology for PPA with other mammalian $\alpha$-amylases. No similarity was detected for $\alpha$-glucosidases which lack subsites equivalent to $B_1$–$B_3$. The probabilities [10] that the apparent homology of TAA with gaAn and CGT occurred by chance are about $10^{-2}$ ($2.3\sigma$) and $4 \times 10^{-4}$ ($3.3\sigma$), respectively. The values ranged from $10^{-4}$ to $10^{-6}$ for the comparisons of CGT and glucoamylases (fig.1F and [10]).

In gaAn both Trp 120 [3] and Asp 112 [14], corresponding to TAA Tyr 75 in $B_1$, are protected by acarbose. Trp 120 is in a subsite at a distance from the catalytic site [3,21] located in glucoamylase between subsites 1 and 2 [22]. The substrate analogue thus binds with a terminal glucosyl residue in subsite 1 and the reducing end towards Trp 120, equivalent to TAA Trp 83 in $B_3$ (fig.1F). This orientation of the $\alpha$-glucan chain is opposite to that indicated for amylose in TAA [4].

Sequence homology in addition exists in COOH-terminal regions of gaAn and CGT (fig.2), two enzymes that degrade raw starch [23,24]. gaAn without the COOH-terminal segment [24,25] or with oxidized tryptophans in that region [26] retains full activity on soluble starch, but is unable to bind onto raw starch [24,26]. The indicated homology possibly comprises raw-starch binding residues.

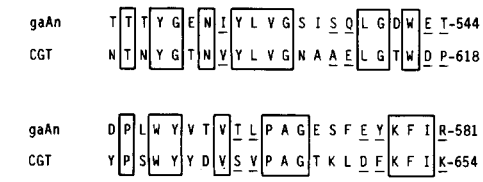


Fig.2. Sequence homology in COOH-terminal regions of glucoamylase G1 from *A. niger* (gaAn) and cyclodextrin glucanotransferase from *B. macerans* (CGT). Identical positions are boxed and conserved exchange groups [10] are underlined.

In summary, biochemical evidence [3,12,14,20] and local homology, characterized by alignment scores [10], provide support for the thesis that certain starch and saccharide hydrolases and transglucanosylases form a family of distantly related proteins. Carbohydrases and transglycosylases were previously claimed, on mechanistic grounds, to belong to the same class of enzymes [27,28]. Regions of CGT and M align colinearly with the $NH_2$-terminal domain of TAA and PPA, suggesting that these enzymes share the $\alpha/\beta$-barrel fold [4,9], while glucoamylases, S-I, AM, and BE appear less closely related and no homology is seen with $\beta$-amylases [29,30]. Structure predictions utilizing motifs [31], for example from the $\alpha/\beta$-barrel, can further examine the possible relationship. The presented thesis can be tested by analysis of relevant 3-D structures, sequences and essential residues and by mutagenesis. Work in progress on site-specific mutation of gaAn [32] has confirmed earlier identification of essential residues [3,14]. The investigations may lead to a generalized tertiary structural template for active centres of starch hydrolases and related enzymes and enable engineering, guided by sequence homology, of carbohydrase performance and specificity.

*Acknowledgements:* M. Ottesen, F.M. Poulsen and W.R. Taylor are thanked for stimulating discussions.

## REFERENCES


[1] Argos, P., Landy, A., Abremski, K., Egan, J.B., Haggard-Ljungquist, E., Hoess, R.H., Kahn, M.L., Kalionis, B., Narayana, S.V.L., Pierson, L.S. iii, Sternberg, N. and Leong, J.M. (1986) EMBO J. 5, 433–440.

[2] Argos, P. (1987) J. Mol. Biol. 193, 385–396.

[3] Clarke, A.J. and Svensson, B. (1984) Carlsberg Res. Commun. 49, 559–566.

[4] Matsuura, Y., Kusunoki, M., Harada, W. and Kakudo, M. (1984) J. Biochem. 95, 697–702.

[5] Toda, H., Kondo, K. and Narita, K. (1982) Proc. Jap. Acad. 58, Ser.B, 208–212.

[6] Friedberg, F. (1983) FEBS Lett. 152, 139–140.

[7] Rogers, J.C. (1985) Biochem. Biophys. Res. Commun. 128, 470–476.

[8] Nakajima, R., Imanaka, T. and Aiba, S. (1986) Appl. Microbiol. Biotechnol. 23, 355–360.

[9] Buisson, G., Duée, E., Haser, R. and Payan, F. (1987) EMBO J. 6, 3908–3916.

[10] Dayhoff, M.O., Barker, W.C. and Hunt, L.T. (1983) Methods Enzymol. 91, 524–545.

[11] Hunziker, W., Spiess, M., Semenza, G. and Lodish, H.F. (1986) Cell 46, 227–234.

[12] Quaroni, A. and Semenza, G. (1976) J. Biol. Chem. 251, 3250–3253.

[13] Ross, W.C.J. (1950) J. Chem. Soc., 2257–2272.

[14] Svensson, B., Clarke, A.J., Møller, H., Svendsen, I. and Håkansson, K. (1986) XIIIth International Carbohydrate Symposium, Ithaca, NY, abstr.C76.

[15] Carraway, K.L. and Koshland, D.E. jr (1972) Methods Enzymol. 25, 616–623.

[16] Robyt, J.F. (1984) in: Starch: Chemistry and Technology (Whistler, R. et al. eds) pp.87–123, Academic Press, New York.

[17] Semenza, G. (1987) in: Mammalian Ectoenzymes (Kenny, A.J. and Turner, A.J. eds) pp.265–287, Elsevier, Amsterdam, New York.

[18] Semenza, G., Curtius, H.-C., Raunhardt, O., Hore, P. and Müller, M. (1969) Carbohydr. Res. 10, 417–428.

[19] Ono, S., Hiromi, K. and Hamauau, Z. (1965) J. Biochem. 57, 34–38.

[20] Gibson, R.M. and Svensson, B. (1987) Carlsberg Res. Commun. 52, 373–379.

[21] Clarke, A.J. and Svensson, B. (1984) Carlsberg Res. Commun. 49, 111–122.

[22] Hiromi, K., Ohnishi, M. and Tanaka, A. (1983) Mol. Cell. Biochem. 51, 79–95.

[23] Kobayashi, S., Kainuma, K. and Suzuki, S. (1978) Carbohydr. Res. 61, 229–238.

[24] Svensson, B., Pedersen, T.G., Svendsen, I., Sakai, T. and Ottesen, M. (1982) Carlsberg Res. Commun. 47, 55–69.

[25] Svensson, B., Larsen, K. and Gunnarsson, A. (1986) Eur. J. Biochem. 154, 497–502.

[26] Svensson, B., Clarke, A.J. and Svendsen, I. (1986) Carlsberg Res. Commun. 51, 61–73.

[27] Kitahata, S., Brewer, C.F., Genghof, D.S., Sawai, T. and Hehre, E.J. (1981) J. Biol. Chem. 256, 6017–6026.

[28] Hehre, E.J., Genghof, D.S., Sternlicht, H. and Brewer, C.F. (1977) Biochemistry 16, 1780–1787.

[29] Kawazu, T., Nakanishi, Y., Uozumi, N., Sasaki, T., Yamagata, H., Tsugagoshi, N. and Udaka, S. (1987) J. Bacteriol. 169, 1564–1570.

[30] Kreis, M., Williamson, M., Buxton, B., Pywell, J., Hejgaard, J. and Svendsen, I. (1987) Eur. J. Biochem. 169, 517–525.

[31] Taylor, W.R. (1986) J. Mol. Biol. 188, 233–258.

[32] Sierks, M.R., Ford, C., Reilly, P.J. and Svensson, B. (1987) 9th Conference on Enzyme Engineering, Santa Barbara, CA, abstr. I-15.