Abstract: Branching enzyme (BE) catalyzes formation of the branch points in glycogen and amylopectin by cleavage of the α,1-4-linkage and its subsequent transfer to the α,1,6-position. A novel BE encoded by an uncharacterized ORF (TK1436) was identified in the hyperthermophilic archaeon, *Thermococcus kodakaraensis* KOD1. TK1436 encodes a conserved protein showing similarity to members of glycoside hydrolase family 57 (GH-57 family). TK1436 orthologs are distributed in archaea of Thermococcales, cyanobacteria, some actinobacteria and a few other bacterial species. When recombinant TK1436 protein was incubated with amylase used as the substrate, a product peak was detected by high-performance anion exchange chromatography, eluting slower than the substrate. Isoamylase treatment of the reaction mixture significantly increased the level of short-chain α-glucans, indicating that the reaction product contained many α,1,6-branching points. TK1436 protein showed an optimal pH of 7.0, an optimal temperature of 70°C, and thermostability up to 90°C as determined by the iodine-staining assay. These properties were the same when a protein devoid of the C-terminal HhH motifs (TK1436ΔH protein) was used. The average molecular weight of branched glucan after reaction with TK1436ΔH protein was over 100 times larger than that of the starting substrate. These results indicate that TK1436 encodes a structurally novel BE belonging to the GH-57 family.

Key words: branching enzyme, glycoside hydrolase family 57, hyperthermophile, archaea

Branching enzyme (BE) [1,4-α-d-glucan: 1,4-α-d-glucan-6-α-d-(1,4-α-d-glucano)-transferase] (EC 2.4.1.18) catalyzes the formation of branch points in glycogen and amylopectin by cleavage of α,1,4-glucosidic bonds and its subsequent transfer to α,1,6-positions. BEs identified to date all belong to family 13 of the glycoside hydrolases (GH-13 family). The GH-13 family is a large enzyme family that constitutes about 20 different reactions and product specificities. This family includes most amyloytic enzymes showing preferences for hydrolysis and transglycosylation of α,1,4- and α,1,6-glucosidic bonds.

The “α-amylase family” contains many proteins in the GH-13 family, such as α-amylase (EC 3.2.1.1), pullulanase (EC 3.2.1.41), isoamylase (EC 3.2.1.68), neopullulanase (EC 3.2.1.135), cyclodextrin glucanotransferase (EC 2.4.1.19) and BE. Proteins in the α-amylase family have the four conserved sequence motifs where catalytic residues and most other residues involved in substrate binding are located. BE also contains the four sequence motifs, and amino acid residues located in these regions (such as Asp405, Glu458 and Asp526 in the *Escherichia coli* BE numbering) have been shown, by site-specific mutagenesis, to be important for BE activity. The three-dimensional structure of BE supports this view.

In this paper, we report the identification of a novel BE in the hyperthermophilic archaeon, *Thermococcus kodakaraensis* KOD1. The new BE belongs to family 57 of glycose hydrolyases (GH-57 family) and is structurally distinct from the previously characterized BEs of the GH-13/α-amylase family.

Identification of TK1436 as a maltodextrin-induced ORF.

*Th. kodakaraensis* KOD1 is a hyperthermophilic archaeon belonging to the order Thermococcales. It is an anaerobic heterotroph that utilizes peptide-related substrates in the presence of elemental sulfur (S). It also utilizes starch without a requirement for S. Consideration of genome information and biochemical characteristics of the strain led to the proposal of a glycolytic pathway relevant to starch metabolism.

Using a DNA microarray that covers 96.5% of *T. kodakaraensis* ORFs, we performed a transcriptome analysis to exhaustively identify genes relating to α-glucan metabolism in the strain. A maltodextrin (Amycol No. 3-L) containing maltooligosaccharides with a degree of polymerization (DP) of 1–12, was used as the source of α-glucan. Among genes whose transcription was induced by the addition of maltodextrin, an ORF encoding an unknown protein (TK1436) was identified. The signal intensity of TK1436 was induced 3.2-fold under glycolytic conditions (containing maltodextrin) compared to the level seen under peptide-utilizing conditions.

Primary structure of the TK1436 protein.

TK1436 encodes a protein of 675 aa with a calculated molecular mass of 78,545 Da (Fig. 1). It is an uncharacterized conserved protein annotated as a “probable α-amylase” by the *T. kodakaraensis* genome project.
According to the NCBI Conserved Domain Database, TK1436 is assigned to the COG1543 group. The functions of the group members are unknown.

A long N-terminal region (1–534 aa) of the TK1436 protein shows similarity to proteins in the GH-57 family. Phylogenetic analysis of family members indicated that the GH-57 family is divided into seven subfamilies and three independent members. Four subfamilies include proteins of characterized specificities (α-amylase, 4-α-glucanotransferase, amylpullulanase and α-galactosidase). The TK1436 protein belongs to a subfamily of which no protein has been previously characterized. Multiple alignments of subfamily members showed the existence of six conserved sequence regions (Regions A to F). Regions A, B, C, E and F essentially correspond to the five conserved conserved sequence regions (Regions A to F). The TK1436 protein was incubated at 70°C with melibiose, pullulan, maltoheptaose or amylose as the substrate, and the reaction mixtures were analyzed by TLC. No mobility change was observed in reaction mixtures containing melibiose or pullulan, showing that TK1436 protein contained neither α-amylase nor pullulanase activity. Similarly, hydrolysis or transglycosylation of maltoheptaose could not be observed, suggesting that TK1436 protein contained neither α-amylase nor 4-α-glucanotransferase activities. However, in the case of amylose, small but significant amounts of short-chain glucans were detected.

The distinct specificities of TK1436 protein for two α-glucans (maltoheptaose and amylose) are seemingly conflicting. But BE is reported to produce short-chain α-1,4-glucan as a side product when amylose is used as the substrate. We next asked, therefore, if TK1436 protein showed BE specificity. The reaction mixture after incubation with TK1436 protein was further treated with isoamylase, catalyzing hydrolysis of α-1,6-linkages. As a result, the levels of maltotriosaccharides increased greatly, suggesting that the reaction product contained many α-1,6-branch points. Next, reaction mixtures were analyzed in detail by HPAEC (Fig. 2). The mixture after reaction with TK1436 protein contained, in addition to short-chain α-1,4-glucans, a peak corresponding to high molecular weight product, eluting at around 43 min (Fig. 2, chromatogram 2). Isoamylase treatment of the reaction mixture significantly increased the amounts of short-chain α-1,4-glucans with concomitant disappearance of the high molecular product (Fig. 2, chromatogram 3). The amylose used in this experiment (amylose AS-30) contains negligible levels of α-1,6-branch points, as isoamylase treatment of the substrate did not change the chromatographic profile.

**Determination of enzymatic activity of the TK1436 protein.**

In order to elucidate the function of TK1436 protein, recombinant protein was prepared using *E. coli* as host cells. In addition to the whole-length protein (TK1436 protein, 675 aa), recombinant protein devoid of both copies of HhH motifs (TK1436 ΔH protein, 562 aa) was prepared (Fig. 1).

Using the whole-length protein, we first examined whether TK1436 protein showed any enzyme activities known in the GH-57 family. Enzyme activities that were detected in the family are α-amylase, 4-α-glucanotransferase, amylpullulanase and α-galactosidase. TK1436 protein was incubated at 70°C with melibiose, pullulan, maltoheptaose or amylose as the substrate, and the reaction mixtures were analyzed by TLC. No mobility change was observed in reaction mixtures containing melibiose or pullulan, showing that TK1436 protein contained neither α-galactosidase nor pullulanase activity. Similarly, hydrolysis or transglycosylation of maltoheptaose could not be observed, suggesting that TK1436 protein contained neither α-amylase nor 4-α-glucanotransferase activities. However, in the case of amylose, small but significant amounts of short-chain glucans were detected.

The distinct specificities of TK1436 protein for two α-glucans (maltoheptaose and amylose) are seemingly conflicting. But BE is reported to produce short-chain α-1,4-glucan as a side product when amylose is used as the substrate. We next asked, therefore, if TK1436 protein showed BE specificity. The reaction mixture after incubation with TK1436 protein was further treated with isoamylase, catalyzing hydrolysis of α-1,6-linkages. As a result, the levels of maltotriosaccharides increased greatly, suggesting that the reaction product contained many α-1,6-branch points. Next, reaction mixtures were analyzed in detail by HPAEC (Fig. 2). The mixture after reaction with TK1436 protein contained, in addition to short-chain α-1,4-glucans, a peak corresponding to high molecular weight product, eluting at around 43 min (Fig. 2, chromatogram 2). Isoamylase treatment of the reaction mixture significantly increased the amounts of short-chain α-1,4-glucans with concomitant disappearance of the high molecular product (Fig. 2, chromatogram 3). The amylose used in this experiment (amylose AS-30) contains negligible levels of α-1,6-branch points, as isoamylase treatment of the substrate did not change the chromatographic profile.

**Fig. 1.** Overall structure of TK1436 and TK1436 ΔH proteins.


**Fig. 2.** HPAEC analysis.

1 (the amylose AS-30 substrate, 2) products after treatment of amylose AS-30 with TK1436 protein, 3) isoamylase-digested products of 2, (4) products after treatment of amylose AS-30 with TK1436 ΔH protein, (5) isoamylase-digested products of 4. Peaks eluting at 40–45 min in chromatograms are the substrate (41.4 min, shown as S) and branched products (43.3 min, shown as P). The numbers on peaks indicate chain lengths.
file (data not shown). In summary, these results clearly indicate that TK1436 protein contains BE activity that was not previously identified in the GH-57 family of enzymes.

The chain length distribution of the reaction products showed DP 5–30, with two local maxima at DP 6 and DP 11. The branch length distribution is much different from that obtained with *Aquilax aeolicus* BE, products of which showed a branch chain distribution of DP 5–40, with a sharp peak at DP 10. The reaction product of TK1436 protein showed basically the same properties as that of TK1436 protein (Fig. 2, chromatograms 4 and 5).

**Enzymatic characterization of recombinant TK1436 and TK1436 ΔH proteins.**

Optimal pH activity values, optimal temperature activity values, and thermostabilities of TK1436 and TK1436 ΔH proteins were determined by the iodine-staining assay. Both proteins showed the same optimal pH value of 7.0, functioned optimally at a temperature of 70°C, and had the same specific activities (TK1436 ΔH protein had a 95.6±4.6% activity of TK1436 protein). Both proteins were stable up to 90°C and activities decreased gradually during incubation at 100°C. The thermostability is even higher than that of *A. aeolicus* BE, which is the most thermostable BE previously reported, with stability at temperatures up to 85°C. In Fig. 3, the data on TK1436 ΔH protein are shown as representative.

**Detailed characterization of reaction product and enzyme.**

Molecular weights of reaction products as well as products after treatment with TK1436 ΔH protein were determined by HPLC-MALLS-RI analysis. Amylose AS-30 was used as substrate, and a time course of reaction products was analyzed (Fig. 4). A sharp peak at 0 h of incubation corresponded to the substrate (peak 2) with a Mw average of 4.5×10^7. As the reaction proceeded, the substrate peak gradually disappeared, and two other peaks were generated. A peak appearing faster than peak 2 (peak 1) corresponded to branched products with a Mw average of 3.0×10^8–3.5×10^8. The value was about a hundred times greater than that of the starting substrate. On the other hand, the other peak appearing slower than peak 2 (peak 3) corresponded to short-chain glucans with a Mw average of 4.8×10^7–2.0×10^7.

To confirm that the branched products did not contain linkages other than α-1,4- and α-1,6-glucosidic bonds, reaction products after isoamylase digestion were treated with β-amylase, and analyzed by HPAEC. No glucan except for maltose and glucose was detected (data not shown) indicating that branched products contained only α-1,4- and α-1,6-linkages similar to glucans made by conventional BEs of the GH-13 family.

**GH-57 family as the second α-amylase family.**

The present identification of BE specificity in the GH-57 family is very significant because this finding proves the existence of all four basic activities in the GH-57 family pertaining to synthesis and decomposition of α-glucans (i.e., hydrolysis of α-1,4- and α-1,6-glucosidic linkages, and transglycosylation to form α-1,4- and α-1,6-glucosidic linkages). It thus seems that the GH-57 family can be regarded as another “α-amylase family” with a structural background distinct from that of the GH-13 family, although future investigations to prove the common catalytic mechanism are necessary.

The results in this paper allowed us to identify GH-57-
type BEs in a number of microorganisms (Fig. 5). In spite of the present identification of BE specificity in an archaean protein, distribution of the GH$_{57}$ type BEs in the archaean domain is limited only in the order of Thermococcales, and a rather large number of orthologs are present in the bacterial domain. In particular, it is intriguing that all the completely sequenced cyanobacteria contain a GH-57-type BE gene in addition to the conventional GH$_{13}$ type BE gene. It was reported that a mutant of Synechocystis sp. lacking the GH-13-type BE gene still produced $\alpha$-glucan with an appreciable level of $\alpha$-1,6-branches, although the source of the branch formation was unclear.$^{30}$ Now we assume that GH-57-type BE of the Synechocystis sp. is actually functional $\textit{in vivo}$ and that this is responsible for the branched formation. The GH-57-type BE gene is also distributed in all five sequenced Mycobacterium genomes, and in both of the Thermus genomes available. On the other hand, while most species in Bacillus and Clostridium contain a GH-13-type BE gene, only a GH-57-type BE gene is present in Bacillus halodurans and Clostridium acetobutylicum.

**Structural perspective.**

In the structural genomics project of Thermus thermophilus HB8 (http://www.thermus.org/), the crystal structure of a GH-57-type BE ortholog, TT1467 protein, was determined as a protein of unknown function, and the data are filed in the PDB database (Accession number: 1UFA). At present, TT1467 is termed TTHA1902. As TK1436 shares 44% amino acid identities with TTHA1902, we performed comparative modeling to predict the 3D structure of TK1436 protein. Figure 6 shows a model structure around proposed catalytic residues, Glu183 (a nucleophile) and Asp354 (an acid/base catalyst). Carboxyl groups from the two catalytic residues are positioned face to face. The distance between these carboxyl groups is calculated as 6.3 Å. This length is close to that (6.7 Å) of \textit{T. litoralis} 4-$\alpha$-glucanotransferase, the 3D structure of which is the only reported in the GH-57 family.\textsuperscript{14} We found five residues (His10, His145, Trp270, Trp407 and Trp416) located within 5 Å from either one of the carbon atoms in the catalytic carboxyl groups are indicated. The dotted line indicates the shortest distance between carbon atoms in the two catalytic carboxyl residues.

Fig. 5. Phylogenetic tree of the TK1436 orthologs.

Fig. 6. A catalytic center model of TK1436 protein.

Glu183 (a nucleophile) and Asp354 (an acid/base catalyst) are the two catalytic residues conserved in the members of the GH-57 family. Residues (His10, His145, Trp270, Trp407 and Trp416) located within 5 Å from either one of the carbon atoms in the catalytic carboxyl groups are indicated. The dotted line indicates the shortest distance between carbon atoms in the two catalytic carboxyl residues.
Trp407 and Trp416, are located close to Glu183 (a nucleophile) and appear to cover the residue. It is, therefore, predicted that these residues function to prevent acceptor sugar (and water) from approaching from this direction. Unlike the situation with these four residues, Trp270 located in Region D is only conserved in the TK1436 subfamily, suggesting that the residue has an important function in defining BE specificity.

CONCLUSION

Overall, the present identification of GH-57-type BEs is noteworthy since it is the first BE found and characterized outside of the GH-13 family. The discovery of an overlooked BE will contribute significantly to our understanding of a diversified glycogen biosynthesis in living organisms as well as to industrial application in the field of starch processing.

REFERENCES


10) H. Atomi, T. Fukui, T. Kanai, M. Morikawa and T. Imanaka: Description of *Thermococcus kodakaraensis* sp. nov., a well studied hyperthermophilic archaeaon previously reported as *Pyrococcus* sp. KOD1. *Archaea*, 1, 263–267 (2004).


30) S.H. Yoo, M.H. Spalding and J.L. Jane: Characterization of
GH-57 family に属する新規 Branching Enzyme の同定と解析
金井 保¹, 村上 平¹, 高田洋樹¹, 栗木 隆¹, 今中忠行¹
¹ 京都大学大学院工学研究科
(615-8510 京都市京北区京大宫)
² 江崎グリコ株式会社生物化学研究所
(555-8502 大阪市西淀川区帯島 4-6-5)

超好熱原産菌 Thermococcus kodakaraensis KOD1 株の DNA Microarray 解析の結果、マルトデキストリン代謝条件において特異的に転写量増加を示す機能未定 ORF（TK1436）が発見された。TK1436 は Glycoside hydrolase, family 57 (GH-57 family) に分類されるタンパク（675aa, 78.5 kDa）をコードしていた。GH-57 family は一次構造より七つのサブファミリーに分類されるが、TK1436 は GH-57 family で既知の 4 種の酵素（α-Amylase, 4-α-Glucanotransferase, Amylopullulanase, α-Galactosidase）とは異なるサブファミリーに属していた。TK1436 タンパクの機能解析を行うために、組換え TK1436 タンパクの調製を行い、Melibiose, Pullulran, Maltoheptaose および Amylose を基質として反応を行った。その結果、Amylose を基質とした場合に、微量ながらもマルトオリゴ糖含量の上昇が観察された。Branching enzyme (BE) は、Amylose を基質とし
て Glucogen や Amylopectin の分枝構造を作る酵素であるが、反応副産物として少量のマルトオリゴ糖を生成する。これまでは BE としては、GH-13 family に属するタンパクのみが知られていたが、我々は TK1436 が新規 BE である可能性について検討を行った。先ほどの反応産物を、さらに α-1,6 結合の切断を行う Isoamylase で処理を行った結果、多量のマルトオリゴ糖が遊離した。次に平均分子量 3.0×10⁴ の Amylose を基質として、反応副産物の分子量を測定した。その結果、生成物の分子量は 3.0×10⁴～3.4×10⁵ であり、基質の 100 倍以上の高分子であることが判明した。

★★★★★

[質問]
秋田県立大 鈴木
1) GH57 ファミリーの酵素は、GH13 ファミリーのも