

Evolution and Adaptation of Enzymes' Active Sites: Enzymes Use Functionally Similar Active Site for Same Type of Reactions with Structurally Similar Substrates Irrespective of the Organism

Peramachi Palanivelu *

Senior Professor and Head (Retd.), Department of Molecular Microbiology, School of Biotechnology, Madurai Kamaraj University, Madurai – 625 021, India.

World Journal of Advanced Research and Reviews, 2025, 28(03), 1810-1835

Publication history: Received on 15 November 2025; revised on 21 December 2025; accepted on 24 December 2025

Article DOI: <https://doi.org/10.30574/wjarr.2025.28.3.4234>

Abstract

Enzymes have played a crucial role in the evolution of life on the planet Earth, as they could accelerate thermodynamically favourable reactions approximately a million times faster. The evolution of enzymes with their unique 3D structures and specific active sites for each reaction is still a mystery. Interestingly, now it is found that two different transferases from two different metabolic pathways use the same active site amino acids, suggesting that active sites of enzymes are designed based on the specific molecular interactions between the active site amino acids and their substrate molecules. The present data further reveal that not only the active site is designed for a particular biological reaction during evolution, but also adapted and conserved in all life forms, from viruses to humans. This is based on the fact that just two-point mutations, viz. S³⁵⁶ and Q³⁶³, instead of the conserved Phe³⁵⁶ and a positively charged residue K³⁶³, respectively, rendered the enzyme, fructosyl transferase from *Geobacillus stearothermophilus* inactive. The importance of these two amino acids in fructosyl transferases was further confirmed from analysis of hundreds of fructosyl transferase sequences from levan- and inulosucrases, and other glycoside hydrolase GH68 family of enzymes from a large number of organisms. The fructosyl transferases are bifunctional enzymes and function both as a hydrolase and as a transferase. The hydrolase uses a catalytic triad in its active site, viz. 2 Asp residues where one acts as the nucleophile and the other acts as the transition-state stabilizer and a Glu residue that acts as a general acid-base catalyst. Similarly, the transferase also uses a catalytic triad, viz. a positively charged amino acid residue, K/R/H, as the proton abstractor and an R at -3 from the proton abstractor as the donor selection amino acid and an Y or a branched chain amino acid with a G as the template-binding pair. It is interesting to note that the same catalytic triad is observed in the active sites of nucleotidyl transferases like DNA/RNA polymerases [1-3]. For example, the nucleotidyl transferase (*E. coli* DNA polymerase I) and the fructosyl transferase (*Bacillus subtilis* levanasucrase) possess more or less the same active site structure, (-R⁻⁴RSAK⁷⁵⁸A⁷⁵⁹N⁷⁶⁰F⁷⁶¹G⁷⁶²I⁷⁶³Y⁷⁶⁴GM⁻ and -D³⁵⁸S³⁵⁹R⁻³G³⁶⁰S³⁶¹K³⁶²M³⁶³T³⁶⁴I³⁶⁵G³⁶⁶T³⁶⁷D³⁶⁸I³⁶⁹Y³⁷⁰M³⁷¹L³⁷²GY⁻, respectively) where K⁷⁵⁸ in the *E. coli* enzyme is shown to be the proton abstractor, the R at -4 as the nucleotide selection amino acid and the -YG- as the template-binding pair. Furthermore, extensive analysis of different DNA/RNA polymerases and fructosyl transferases from a large number of organisms, from viruses to humans, revealed the same active site structure [1-3]. These data confirm that during evolution, not only similar active site is designed and used for the same types of reactions, but also conserved and adapted to all forms of life, from viruses to humans.

Keywords: Evolution of Life; Evolution of Enzyme Active Sites; Fructans; Fructosyl Transferases; Hydrolase Active Site; Fructosyl Transferase Active Site; Nucleotidyl Transferases; Nucleotidyl Transferase Active Site; Mechanisms of the Fructosyl and Nucleotidyl Transferases

* Corresponding author: Peramachi Palanivelu

1. Introduction

Understanding the origin of life is still a complex and a completely unsolved problem in biology. However, laboratory experiments suggest that life might have evolved in the early Earth by a series of progressive and spontaneous chemical reactions under the prevailing environmental conditions of the prebiotic era. The first phase of evolution, viz. the chemical evolution, actually set the very early stage for the development of life by the formation of simple organic molecules such as amino acids, simple sugars, fatty acids, nucleic acid bases, etc. from inorganic molecules. Such spontaneous formation of organic molecules from inorganic molecules is proved by groundbreaking experiments conducted by Urey [4] and Miller [5] in the 1950s where they simulated the conditions of early Earth's atmosphere and studied the formation of organic molecules. Miller and Urey [6] demonstrated that organic molecules such as amino acids could be spontaneously formed from inorganic compounds when an electrical arc (simulating lightning) was passed through a mixture of gases. Their seminal experiments suggested that electrical discharges might have played a significant role in the formation of organic compounds from a mixture of gases present in the primitive atmosphere.

Their findings of spontaneous formation of life's molecules in the primitive atmosphere were further confirmed by Criado-Reyes *et al.* [7]. They recently repeated the familiar Urey-Miller experiment, conducted about 70 years ago, but in a borosilicate vessel. Interestingly, they found a large number of additional compounds such as dipeptides, multi-carbon dicarboxylic acids, polycyclic aromatic hydrocarbon compounds (PAHs) and a complete panel of biological nucleobases in their reaction mixture. Furthermore, the natural formation of small peptides and oligonucleotides in the presence of inorganic catalysts like hydroxyapatite, clay, zeolites, etc. further advanced our knowledge on the evolution of life [8]. These experiments strongly support a widely accepted theory of "spontaneous generation," suggesting that chemical reactions could just start spontaneously when the reacting substances are kept in close proximity to each other. This theory of 'spontaneous formation' of life's molecules is further supported by a recent work by Mohajer *et al.* [9]. They have shown the spontaneous formation of urea, the simplest organic molecule and a key compound that connects chemistry and life and is also considered as a key molecule in the search for the origin of life. Its formation from ammonia and carbon dioxide requires either high pressures and temperatures, or, under milder conditions with catalysts or additional reagents. In contrast, Mohajer *et al.* [9] have demonstrated that the spontaneous formation of urea under ambient conditions (i.e., non-energetic conditions and without additional catalysts) from ammonia and carbon dioxide in the surface layer of aqueous droplets. Their experiment further corroborates that the evolutionary processes are spontaneous.

Further, additional advanced experiments on evolution have demonstrated that replicating autocatalytic RNA molecules could undergo spontaneous changes and that the variants of RNA molecules with the greatest autocatalytic activities prevailed in the prebiotic life [10]. Their results support the basis of the "RNA world" hypothesis. In support of this hypothesis, a recent study has shown that RNAs are not only autocatalytic, but also could help amino acids to assemble into peptides without a pre-existing protein machinery during evolution [10]. However, the "Metabolism First" camp argues that self-sustaining chemical networks should have evolved first for the self-replicating genetic molecules like RNA to use them. The 'Metabolism First' model suggests that life emerged as a result of simple chemical reactions, governed by the laws of chemistry and thermodynamics, rather than through improbable chance events.

The next crucial phase in the origin of life is the formation of 'protocells' that are enveloped with primitive type membrane structures, where the organic compounds that are formed are compartmentalized and the compartmentalized organic compounds with self-organization potential led to the formation of lower molecular weight compounds like simple oligopeptides, oligonucleotides, etc. Thus, they bridge chemistry and biology by establishing primitive metabolic activities inside them. The protocells later evolved into 'primitive cell' types. The primitive cells may be considered as the first living cells capable of basic metabolism and reproductive capabilities. Interestingly, small peptides, oligonucleotides that are assembled on inorganic matrixes like clay, silica, zeolites, etc., and other small molecules, nutrients, ions, etc. could also have diffused freely into the primitive cells expanding their metabolic capabilities [11]. It is also likely that at different prevailing environmental conditions on the planet Earth different types of protocells and primitive cells would have formed, giving rise to diversities. In the next phase, the glycan-based polymers would have formed a protective layer on the fragile cell membranes, leading to the formation of complete primitive cell types. The two crucial components of basic life processes, viz. enzymes (which are responsible for all the cellular reactions) and genetic materials (which carry the genetic information to the next generation) might have evolved much later.

Primitive types of catalytic activities in the primitive cells would have started with metal ions and minerals like clay, zeolites, etc. As each enzyme's active site is specific for its substrate, it would have formed after the binding of active site amino acids on to their respective existing substrates, followed by filling of the gaps between the active site amino acids with other amino acids and likely not *vice versa*. This is evidenced by the distance conservation between the active

site amino acids in enzymes. Then the enzymes could have formed their respective RNA molecules where three nucleobases (as a codon) recognize each of the amino acid and such codons are connected in the same order into a short RNA molecule representing that particular enzyme, which would later code for that enzyme. An intuitive addition of a Met codon at the start and a stop codon at the end make the translation events unidirectional. And, many such protein coding RNA molecules could have further assembled into a larger RNA molecule forming an RNA genome. The RNA genome, thus, carries all the molecular instructions of the enzymes in a cell. And, when the deoxyribonucleosides were available, the molecular instructions on the RNA molecule could have transferred to the more stable DNA molecule by complementary base-pairing. In the reverse direction, such DNA can give rise to RNA molecules (transcription) and the RNA molecules can give rise to enzymes (translation) which function in the cells as originally designed by nature. Therefore, it is tempting to speculate that the life's molecules could have evolved spontaneously, self-assembled and organized into primitive cell types, initiating all cellular reactions. The mutation and selection, addition(s) and deletion(s), are natural and progressive steps in the evolutionary process. As discussed elsewhere, enzymes have evolved over billions of years through natural selection, which act primarily at the level of biological function, and diverse types of enzymes are constantly evolved from mutation and selection, divergence of existing genes and recombination of domains. Further, adaptation to different new environments could also play a key role in further evolution of new enzymes with new biological functions. Besides, changes in the available chemicals, ions, minerals and environmental conditions could also favour further evolution of enzymes with new functions that are beneficial to the organisms.

In this communication, evidences are presented to show how two different unconnected enzymatic reactions, one in the carbohydrate metabolism and the other in the nucleic acid metabolism, use the same active site amino acids for polymerization reactions, which strongly supports the chemical basis of evolution enzymes and their active sites. Both the enzymes are classified under the Main class, 'Transferases', but one is a fructofuranosyl transferase and the other is a nucleotidyl transferase. The nucleotidyl transferases play the key role in DNA/RNA polymerases to polymerize the nucleobases into a DNA or an RNA. It is interesting to note that the requirement for such specialized nucleotidyl transferases to accurately replicate the genetic material is universal in all kingdoms of life and hence, they are also known as crucial 'life catalysts'. The properties, multiple sequence alignment (MSA) analysis and the mechanism of actions of DNA and RNA polymerases from viruses, plants to humans have been extensively analyzed and reported by the author [1-3]. Therefore, in this communication evidences are provided to support how these two different groups of enzymes use and have adapted the same active site amino acid structure to perform similar functions in all organisms.

1.1. Fructans and their Applications

Inulins, levans and fructooligosaccharides are collectively known as fructans. They mainly differ in their linkages and degree of polymerization (DP). The inulins and levans exhibit chain lengths of >20 DPs and the fructooligosaccharides are of only <20 DPs. Demand and interest for fructans have increased over the years, due to their versatile applications and with increased interest in healthy lifestyles as they exhibit many potential health benefits. In recent times, the global market for fructans has steadily increased, with an annual growth rate of >5%, and is expected to reach approximately 2.7 billion US dollars by 2026 [12]. The fructans are produced by a large number of organisms like bacteria, yeasts [13], higher fungi [14] and plants [15]. Bacterial levans (produced as exopolysaccharides, EPS) are much larger than those produced by plants and fungi with multiple branches and molecular weights ranging from 2 to 100 million Daltons [16]. Inulin-type fructans are mostly found in plants. The dicots accumulate them as long-term reserve carbohydrates, in underground storage organs such as roots and tubers and are mainly extracted from the plant sources like chicory, dahlia, yacon and Jerusalem artichoke as they accumulate up to 80% polysaccharides in their tubers [17-19]. In grasses, graminan (branched type fructans), levan, and neokestose-derived fructans mainly act as short-term storage compounds in stems, tiller bases, leaf sheaths, elongating leaf bases, and to a lesser extent in leaf blades and roots [17]. They find versatile applications in various industries and also in medicine. In industries, they are used as food thickeners, in cosmetics, as prebiotics, dietary fibres, mineral absorbing agents, anticancer agents, diabetic control agents, etc. Fructans exhibit many potential health benefits. As these fructans are non-digestible in our digestive system, they are considered as health-promoting 'prebiotics', i.e., they act as a food source for the beneficial gut bacteria. Therefore, they are gaining considerable attention the health industries as effective prebiotics to support the growth of beneficial gut bacteria, which play important roles in promoting digestive health, boosting the immune system, reducing the risk of chronic diseases, etc. [20-22]. Many bacteria, fungi and ~ 40,000 plant species use fructans as the primary carbohydrate reserves. Furthermore, it is found that in plants these fructans function as stabilizers of cell membranes and thus, confer tolerance to drought and frost. In plants, the fructan molecules are bound to the polar head groups of the lipid bilayer of the membrane to block water leakage during abiotic stress such as frost and drought [15, 23]. In addition to the above roles in industrial applications and as prebiotics, levan exhibits a wide range of specialized properties like film-forming ability, biodegradability, non-toxicity, self-aggregation, encapsulation, controlled release capacity, water retention, immunomodulatory activity, antimicrobial and anticancer activities, high biocompatibility

etc. With the strongest bio-adhesive properties and film-forming ability, levan could be used for healing wounds, burned tissues and for the development of bioresorbable electronic implants [24]. Furthermore, high DP fructans from bacteria are ideal substrates for the production of high fructose syrups (HFS), with very low glucose content. As levan is more soluble than inulin, it produces viscous solutions in water. In the industry, this property makes levan especially attractive as an emulsifier or encapsulating agent in a wide range of products, including biodegradable plastics, cosmetics, glues, textile coatings and detergents. These exceptional properties position levan as an attractive candidate for nature-based materials in food production, modern cosmetology, medicine, and pharmaceutical industries. Thus, with excellent medicinal properties and ease of production, microbial levans appear as a valuable and versatile biopolymer of the future.

1.2. Fructofuranosyl Transferases

The commonly used sugar, sucrose [α -D-glucopyranosyl-(1 \rightarrow 2)- β -D-fructofuranoside] is a non-reducing disaccharide, where its two monosaccharide units, viz. glucose and fructose are linked by a glycosidic bond. Unlike most of the other disaccharides, as both the reducing ends of the monosaccharide units are involved in the glycosidic bond formation, it becomes non-reducing. Furthermore, in sucrose, the glucose is in its pyranose form and the fructose is in its furanose form. It is metabolized mainly by at least four types of enzymes, i.e., i) invertases (EC 3.2.1.26), which hydrolyze the sucrose to produce 1 mole each of glucose and fructose, ii) inulosucrases (E.C 2.4.1.9), which hydrolyze the sucrose to produce inulin (with a molecular mass of about 10 Da) and glucose units, iii) levansucrases (E.C 2.4.1.10), which hydrolyze the sucrose to produce levan (with a molecular mass up to 1×10^6 Da) and glucose units and iv) dextranases (EC 2.4.1.5), which hydrolyze sucrose and produce dextran (with a molecular mass of about 1kDa to 2MDa) and fructose units. The chains of levan, inulin and dextran, like other biopolymers like starch, cellulose, and chitin grow stepwise by the repeated transfer of a hexosyl group (donor) to a growing acceptor molecule. These fructan-forming enzymes are commonly reported from bacteria, fungi and plants. Bacteria use the fructans and dextran as reserve EPS and utilize them by producing corresponding degradative enzymes like levanase, inulinase and dextranase when needed [25]. Among them, fructans are acid-labile, water-soluble, and polydisperse polymers of fructose built upon a sucrose starter unit. Between the two commonly occurring fructans, viz. inulin and levan, the inulosucrases polymerize fructosyl units into inulin which is mainly β -(2 \rightarrow 1) linked, whereas the levansucrases, form levan from fructosyl units which is mainly β -(2 \rightarrow 6) linked. Both belong to the Glycoside Hydrolase family GH68 (<http://afmb.cnrs-mrs.fr/CAZY>) and specifically belong to the hexosyltransferases [25]. Levansucrases, in general, display a five-bladed β -propeller architecture, where the catalytic residues which are responsible for sucrose hydrolysis are located in a negatively charged cavity [20, 26]. The microbial fructan synthesis is more efficient than the plant systems because it is synthesized from sucrose by a single enzyme reaction and exhibits homogeneous structures with occasional branches. In contrast, plant fructans are shorter and exhibit diverse structures, which result from a combination of catalytic actions of various enzymes, including sucrose:sucrose 1-fructosyltransferase (1-SST), fructan:fructan 1-fructosyltransferase (1-FFT), fructan:fructan 6G-fructosyltransferase (6G-FFT), and sucrose:fructan 6-fructosyltransferase (6-SFT). Fructosyltransferases (Fts) are also reported from fungi such as *Aspergillus*, *Penicillium*, *Aureobasidium*, *Kluyveromyces*, etc. Fungal Fts act on sucrose by cleaving the β -(2-1) linkage, releasing glucose, and then transferring the fructosyl units to an acceptor molecule. The fungal Ft from *Aspergillus* possesses both hydrolytic and transfructosylating activities as in other organisms. Interestingly, at sucrose concentrations >100 mM, the Fts exclusively exhibit a transfructosylation activity [27].

2. Materials and Methods

The protein sequence data of levansucrases, inulosucrases and other GH68 family of enzymes and DNA and RNA polymerases were obtained from PUBMED and SWISS-PROT databases. The advanced version of Clustal Omega was used for protein sequence analysis. Along with the conserved motifs identified by the bioinformatics analysis, and with the data already available from biochemical, site-directed mutagenesis (SDM) experiments and X-ray crystallographic analysis on these enzymes are used to confirm the possible amino acids at the active sites of these two different transferases, viz. fructofuranosyl and nucleotidyl transferases.

3. Results and Discussion

3.1. MSA analysis of levansucrases and other GH68 family glycoside hydrolases fructosyl transferases

Figure 1 shows the MSA of levansucrases and GH68 family of fructosyl transferases from different bacteria, (only the required regions for the discussions are shown). The sources of active and inactive fructosyl transferase enzymes are highlighted in yellow. The green arrow mark tentatively separates the hydrolytic and transferase domains. As the fructosyl transferases are bifunctional enzymes, they exhibit both hydrolase and transferase activities and each activity

is composed of a catalytic triad of three invariant or highly conserved amino acids. Furthermore, they also possess two highly conserved sucrose-binding sites. The N-terminal region is not conserved among them except for a few amino acids. However, an octapeptide -DVWDS/TWPL- is completely conserved in these enzymes where the second D is shown to be the catalytic nucleophile [18]. The second conserved peptide -EWSGSA- is found around 160 amino acids from the N-terminal. The first S is known to play an important role in binding to the Glc unit of sucrose. By SDM, the levansucrase of *Bacillus subtilis* SacB (-EWS¹⁶⁴), where the conserved S¹⁶⁴→A, showed that this is an important residue in maintaining the nucleophile position in the active site, suggesting that the motif is possibly involved in sucrose-binding [28]. The third highly conserved region in these enzymes is a large peptide -DNHTL/M/FRDPHYVED- located around 250 amino acids from the N-terminal. The -RDP- triad is conserved in all and is shown to bind the Fru unit of sucrose. The D in this triad is also shown to play a role in stabilizing the reaction intermediate during catalysis [18]. The acid-base catalytic E is found around 340 amino acids from the N-terminal and located in the highly conserved tetrapeptide -EIER- where the catalytic E is shown to act as the proton donor in catalysis [18]. The transferase module starts immediately after the hydrolytic module (indicated by a green arrow mark). The active site amino acids in the transferase module are highlighted in green. The catalytic site of transferases is also composed of a catalytic triad with an invariant basic amino acid K (proton abstractor), a template-binding pair -LG- (usually exists as a conserved triad -LGY/F) and an invariant amino acid R at -3, which is proposed in selecting incoming donor fructose units for polymerization. It is interesting to note that in the inactive levansucrase from *G. stearothermophilus*, the proton abstractor K is replaced by a Q (marked in red). Thus, this module is conserved in both fructosyl transferases and in nucleotidyl transferases. In the conserved pentapeptide -KWYLF- the invariant Y is also implicated in the interaction with the fructose donor and the F is replaced by an S in the inactive form of the levansucrase from *G. stearothermophilus* (marked in red).

CLUSTAL O (1.2.4) MSA of levansucrases and other GH68 family glycoside hydrolase fructosyl transferases from bacteria

tr A0A163L384 A0A163L384_9BACI	DMINQH--GDDKFTVPSFDASTIKNVESATKTDENGNKIKM	DVWDTWPL	QNADGTVAEYK	108
tr A0AAX2EH38 A0AAX2EH38_9BACI	EMIGQH--GDERYTVPSPDASSIKNIKSATKIDENGNEIKM	DVWDTWPL	QNPDGTVAODYN	113
tr A0A268AGA6 A0A268AGA6_9BACI	EMIGQH--GDERYTVPSPDASSIKNIKSATKIDENGNEIKM	DVWDTWPL	QNPDGTVAODYN	113
tr A0A084GQG0 A0A084GQG0_METIDI	DMINQH--GDARYTVPKFDAKSIENIPIASAKKTTESGETMDI	DVWDTWPL	QNADGTVAEYK	111
tr A0A612MEF7 A0A612MEF7_9BACI	DMINQH--GDARYTVPKFDESTIKNIPSAKKVTESGETIDM	DVWDTWPL	QNADGTVAEYK	111
tr A0A1H0C3I3 A0A1H0C3I3_9BACI	DMISQH--GDANFEVPKFDASSIKNIPSAKKYDKDGNLIMK	DVWDTWPL	QNADGTVANYK	110
tr A0A1N6Y262 A0A1N6Y262_9BACI	NMISQH--GDPKFTVPFAFDESTIRNIPAAKTTDENGNTIMM	DVWDTWPL	QNTDGTVAEYN	111
tr A0AA96N005 A0AA96N005_9BACI	DMIKQH--GDAKYTVPQFDASTIKNIPAAQKVDEENGNLMDI	DVWDTWPL	QNADGTVAEYN	111
tr A0A926NQY10 A0A926NQY10_9BACI	EMALQH--GNEKFEVPKFDDETIKNIPIASAEGL-----	DVWDTWPL	QNADGTVAEYK	101
tr A0A1G8FSH7 A0A1G8FSH7_9BACI	NMIDQH--GDPDFTVPSFDASTIENIPIASAKKINEEGSEIDM	DVWDTWPL	QNADGTVAEYN	111
tr A0A428MZE4 A0A428MZE4_9BACI	DMINQH--GDSDFTVPEFDASTIENIPIASATKTGENGNEIVM	DVWDTWPL	QNADGTVAEYN	111
tr A0A917B9N0 A0A917B9N0_HALAA	DMINQH--GDSDFTVPEFDASTIDNIPSATKTDENGNEIVM	DVWDTWPL	QNADGTVAEYN	111
tr A0A1R1RM81 A0A1R1RM81_9BACI	NIIKQH--DDPRFEVPKQFDASTIKNIPSAKKYDENGKLMID	DVWDTWPL	QNADGTVAEYK	111
tr A0A1H3V2M3 A0A1H3V2M3_9BACI	NMIRQH--GDPRYTVPPEFDASTIVNIPSQAKYDENGNLMDI	DVWDTWPL	QTDGDTVAKYK	111
tr A0A926S2X1 A0A926S2X1_9BACI	KIIGQH--GDSRYTVPQFDESTIKNIPSAKGYDEHGNVIDI	DVWDTWPL	QNADGTVANYN	111
tr A0A2N025D7 A0A2N025D7_9BACI	KMVDQQ--SDPRYTVPPTFDKTAIKNIPSAKKTDEGNVIDI	DVWDTWPL	QNADGTVAEYN	111
tr A0AA91TQD8 A0AA91TQD8_NIACI	KMTEQQ--NDSKFKVPFAFDESTIKNIPSAKGYDEKGNLIDI	DVWDTWPL	QNADGTVAEYN	111
tr A0A7X2V5Z1 A0A7X2V5Z1_9BACI	QIPGQIAEDKERTVPAFDASKLKNIESATGFDKDGKNLIDI	DVWDTWPL	MNADGTVADYK	110
tr A0A0M3RAN0 A0A0M3RAN0_9BACI	KLPKQQ--DNARFEVPKFDADSIRNIPTAQYDKGKNGNIDI	DVWDSWPL	QNADGTVAEYN	110
tr A0A3D8X658 A0A3D8X658_PRIMG	KIPQQQ--NSDQFKVPFAFDESTIKNIASAKGKDASGNTIDI	DVWDSWPL	QNADGTVATYH	110
tr A0A806U3J9 A0A806U3J9_PRIMG	KIPQQQ--KSEQFKVPFAFDKATIKNIASAKGYDKSGNLIDI	DVWDSWPL	QNADGTIANYH	110
tr A0AAP8K2B5 A0AAP8K2B5_PRIAR	KIPQQQ--KSEQFKVPFAFDKATIKNIASAKGYDKSGNLIDI	DVWDSWPL	QNADGTVANYH	110
tr A0AAP3FQJ4 A0AAP3FQJ4_BACVA	KIPEQQ--KNEKYQVPEFDPSTIKNISAKG-----	DVWDSWPL	QNADGTVANYH	101
sp P94468 LSC1_GEOSE	QIPEQQ--KNEKYQVPEFDSSTIKNISAKG-----	DVWDSWPL	QNADGTVANYH	101
sp P05655 LSC_BACSU	QIPEQQ--KNEKYQVPEFDSSTIKNISAKG-----	DVWDSWPL	QNADGTVANYH	101
tr B0N5W2 B0N5W2_9FIRM	QIPEQQ--KNEKYQVPEFDSSTIKNISAKG-----	DVWDSWPL	QNADGTVANYH	101
tr A0A9Q4DMY4 A0A9Q4DMY4_BACSC	QIPEQQ--KNEKYQVPEFDSSTIKNISAKG-----	DVWDSWPL	QNADGTVANYH	101
tr A0A5M8RY49 A0A5M8RY49_9BACI	KIPEQQ--KSEQFQVPFDQPAKTIQNIPSAKGYNSSGKLIDI	DVWDSWPL	QNADGTVATYH	108
tr A0A7Z1B6S1 A0A7Z1B6S1_9BACI	KIPEQQ--KSEQFQVPFDQPAKTIKNIPSAKGYKNGLIDI	DVWDSWPL	QNADGTVATYH	108
tr B6DVC6 B6DVC6_BACLI	KIPEQQ--KSEQFQVPFDQPAKTIKNIPSAKGYKNGLIDI	DVWDSWPL	QNADGTVATYH	108
tr A0AA90UC76 A0AA90UC76_9BACI	KIPEQQ--KSEQFQVPFDQPAKTIKNIPSAKGYKNGLIDI	DVWDSWPL	QNADGTVATYH	108
tr A0A089MCV7 A0A089MCV7_9BACI	KIPEQQ--KSEQFQVPFDQPAKTIKNIPSAKGYKNGLIDI	DVWDSWPL	QNADGTVATYH	108
tr A0A919RE8 A0A919RE8_9CLOT	KIPEQQ--KSGQFQVPFAFNAAISIQNISAKGSDQWGNAAIDI	DVWDSWPL	QNADGTVAEYN	110
tr A0A1ROXBD9 A0A1ROXBD9_9BACI	KLPEQQ--KSQSFQVPFDASTIENIPIAAKTDQWGNQIDI	DVWDTWPL	QNADGTVANYN	109
tr A0A2Z2K7W7 A0A2Z2K7W7_9BACI	KLPKQQ--NSQSFQVPFDASTIENIPIAAKTDQWGNQIDI	DVWDTWPL	QNADGTVASYK	112
tr E6U567 E6U567_ETHERY	KLPEQQ--KSEQFQVPFDASTIENIPIAAKTDQWGNQIDI	DVWDTWPL	QNADGTVANYN	110
tr A0A917S1W0 A0A917S1W0_9BACI	QIPGQQ--SSAQFEVPEFDASTLKNLHQ--VQGPDEGNPVDI	DVWDSWPL	QSPDGTVATYN	106
tr A0A1S8MP69 A0A1S8MP69_CLOPA	KIPGQQ--NSAQFKVPQFDSTIKNIPAAKKG-----	DVWDSWPL	QNADGTAANYH	101
tr A0A0H3J026 A0A0H3J026_CLOPA	KIPEQE--NSAQFKVPQFDSTLKNLILATGKDEAGNLLIDI	DVWDSWPL	QNADGTVANYN	109
tr R4K9X0 R4K9X0_CLOPA	KIPAQQ--NNAQFKVPQFDASTLKNIPSAKGYDEAGNLLIDI	DVWDSWPL	QNGDGTVSNYH	109
tr R4K861 R4K861_CLOPA	KISEQQ--NNAQFKVPQFDASTLKNIPSAKGYDEFGNLLIDI	DVWDSWPL	QNADGTVANYH	109
	KIPGQQ--NNAQFKVPQFDASTIKNISAKDYDKYGNLIDI	DVWDSWPL	QNADGTVANYH	109

tr A0A163L384 A0A163L384_9BACI	EWSGSAI	LFTDDGIRLFYTNRGDFDESK-E-LFGQTLTTAQVNVS	EWGKGDKDLQVDGIEDH	227
tr A0AAX2EH38 A0AAX2EH38_9BACI	EWGSAAFI	DGEIRLFYTNRGDFDESQG-IFGQTLTTAQVNVS	EWGKGDKDLQVDGIEDH	230
tr A0A268AGA6 A0A268AGA6_9BACI	EWGSAAFI	DGEIRLFYTNRGDFDESQG-IFGQTLTTAQVNVS	EWGKGDKDLQVDGIEDH	230
tr A0A084GQG0 A0A084GQG0_METID	EWGSAFTEDGE	VRLFYTNRNEAFNADAG-YYGQTLTTAQVNVS	EWGKGDKDLQVDGIEDH	230
tr A0A612MEF7 A0A612MEF7_9BACI	EWGSAFTEDGE	VRLFYTNRNEAFNADAG-YYGQTLTTAQVNVS	EWGKGDKDLQVDGIEDH	230
tr A0A1H0C3I3 A0A1H0C3I3_9BACI	EWGSAFTSDGR	IIRLFYTNRREPFTPESK-HYGQTLTTAQVNVS	EWGKGDKDLQVDGIEDH	229
tr A0A1N6Y262 A0A1N6Y262_9BACI	EWGSAFTSDGR	IIRLFYTNRREPFTPESK-HYGQTLTTAQVNVS	EWGKGDKDLQVDGIEDH	229
tr A0AA96N005 A0AA96N005_9BACI	EWGSAFTSDGR	IIRLFYTNRREPFTPESK-HYGQTLTTAQVNVS	EWGKGDKDLQVDGIEDH	230
tr A0A926NJ10 A0A926NJ10_9BACI	EWGSAFTSDGR	IIRLFYTNRREPFTPESK-HYGQTLTTAQVNVS	EWGKGDKDLQVDGIEDH	230
tr A0A1G8FSH7 A0A1G8FSH7_9BACI	EWGSAFTSDGR	IIRLFYTNRREPFTPESK-HYGQTLTTAQVNVS	EWGKGDKDLQVDGIEDH	230
tr A0A428MZE4 A0A428MZE4_9BACI	EWGSAFTSDGR	IIRLFYTNRREPFTPESK-HYGQTLTTAQVNVS	EWGKGDKDLQVDGIEDH	230
tr A0A917B9N0 A0A917B9N0_HALAA	EWGSAFTSDGR	IIRLFYTNRREPFTPESK-HYGQTLTTAQVNVS	EWGKGDKDLQVDGIEDH	230
tr A0A1R1RM81 A0A1R1RM81_9BACI	EWGSAFTSDGR	IIRLFYTNRREPFTPESK-HYGQTLTTAQVNVS	EWGKGDKDLQVDGIEDH	230
tr A0A1H3V2M3 A0A1H3V2M3_9BACI	EWGSAFTSDGR	IIRLFYTNRREPFTPESK-HYGQTLTTAQVNVS	EWGKGDKDLQVDGIEDH	229
tr A0A26S2X1 A0A26S2X1_9BACI	EWGSAFTSDGR	IIRLFYTNRREPFTPESK-HYGQTLTTAQVNVS	EWGKGDKDLQVDGIEDH	230
tr A0A2N025D7 A0A2N025D7_9BACI	EWGSAFTSDGR	IIRLFYTNRREPFTPESK-HYGQTLTTAQVNVS	EWGKGDKDLQVDGIEDH	230
tr A0AA91TQD8 A0AA91TQD8_NIACI	EWGSAFTSDGR	IIRLFYTNRREPFTPESK-HYGQTLTTAQVNVS	EWGKGDKDLQVDGIEDH	230
tr A0A7X2V5Z1 A0A7X2V5Z1_9BACI	EWGSAFTSDGR	IIRLFYTNRREPFTPESK-HYGQTLTTAQVNVS	EWGKGDKDLQVDGIEDH	223
tr A0A0M3RANO A0A0M3RANO_9BACI	EWGSAFTSDGR	IIRLFYTNRREPFTPESK-HYGQTLTTAQVNVS	EWGKGDKDLQVDGIEDH	230
tr A0A3D8X658 A0A3D8X658_PRIMG	EWGSAFTSDGR	IIRLFYTNRREPFTPESK-HYGQTLTTAQVNVS	EWGKGDKDLQVDGIEDH	224
tr A0A806U3J9 A0A806U3J9_PRIMG	EWGSAFTSDGR	IIRLFYTNRREPFTPESK-HYGQTLTTAQVNVS	EWGKGDKDLQVDGIEDH	224
tr A0AAP8K2B5 A0AAP8K2B5_PRIAR	EWGSAFTSDGR	IIRLFYTNRREPFTPESK-HYGQTLTTAQVNVS	EWGKGDKDLQVDGIEDH	214
tr A0AAP3FQJ4 A0AAP3FQJ4_BACVA	EWGSAFTSDGR	IIRLFYTNRREPFTPESK-HYGQTLTTAQVNVS	EWGKGDKDLQVDGIEDH	214
sp P94468 LSCI_GEOSE	EWGSAFTSDGR	IIRLFYTNRREPFTPESK-HYGQTLTTAQVNVS	EWGKGDKDLQVDGIEDH	214
sp P05655 LSC_BACSU	EWGSAFTSDGR	IIRLFYTNRREPFTPESK-HYGQTLTTAQVNVS	EWGKGDKDLQVDGIEDH	214
tr B0N5W2 B0N5W2_9FIRM	EWGSAFTSDGR	IIRLFYTNRREPFTPESK-HYGQTLTTAQVNVS	EWGKGDKDLQVDGIEDH	214
tr A0A9Q4DMY4 A0A9Q4DMY4_BACSC	EWGSAFTSDGR	IIRLFYTNRREPFTPESK-HYGQTLTTAQVNVS	EWGKGDKDLQVDGIEDH	214
tr A0A5M8RY49 A0A5M8RY49_9BACI	EWGSAFTSDGR	IIRLFYTNRREPFTPESK-HYGQTLTTAQVNVS	EWGKGDKDLQVDGIEDH	214
tr A0A7Z1B6S1 A0A7Z1B6S1_9BACI	EWGSAFTSDGR	IIRLFYTNRREPFTPESK-HYGQTLTTAQVNVS	EWGKGDKDLQVDGIEDH	222
tr B6DVC6 B6DVC6_BACLI	EWGSAFTSDGR	IIRLFYTNRREPFTPESK-HYGQTLTTAQVNVS	EWGKGDKDLQVDGIEDH	222
tr A0AA90J7C6 A0AA90J7C6_9BACI	EWGSAFTSDGR	IIRLFYTNRREPFTPESK-HYGQTLTTAQVNVS	EWGKGDKDLQVDGIEDH	222
tr A0A089MCV7 A0A089MCV7_9BACI	EWGSAFTSDGR	IIRLFYTNRREPFTPESK-HYGQTLTTAQVNVS	EWGKGDKDLQVDGIEDH	229
tr A0A919RYE8 A0A919RYE8_9CLOT	EWGSAFTSDGR	IIRLFYTNRREPFTPESK-HYGQTLTTAQVNVS	EWGKGDKDLQVDGIEDH	228
tr A0A1ROXB9D9 A0A1ROXB9D9_9BACL	EWGSAFTSDGR	IIRLFYTNRREPFTPESK-HYGQTLTTAQVNVS	EWGKGDKDLQVDGIEDH	231
tr A0A2Z2K7W7 A0A2Z2K7W7_9BACL	EWGSAFTSDGR	IIRLFYTNRREPFTPESK-HYGQTLTTAQVNVS	EWGKGDKDLQVDGIEDH	229
tr E6U567 E6U567_ETHHY	EWGSAFTSDGR	IIRLFYTNRREPFTPESK-HYGQTLTTAQVNVS	EWGKGDKDLQVDGIEDH	224
tr A0A917S1W0 A0A917S1W0_9BACI	EWGSAFTSDGR	IIRLFYTNRREPFTPESK-HYGQTLTTAQVNVS	EWGKGDKDLQVDGIEDH	221
tr A0A1S8MP69 A0A1S8MP69_CLOSA	EWGSAFTSDGR	IIRLFYTNRREPFTPESK-HYGQTLTTAQVNVS	EWGKGDKDLQVDGIEDH	228
tr A0A0H3J026 A0A0H3J026_CLOPA	EWGSAFTSDGR	IIRLFYTNRREPFTPESK-HYGQTLTTAQVNVS	EWGKGDKDLQVDGIEDH	228
tr R4K9X0 R4K9X0_CLOPA	EWGSAFTSDGR	IIRLFYTNRREPFTPESK-HYGQTLTTAQVNVS	EWGKGDKDLQVDGIEDH	229
tr R4K861 R4K861_CLOPA	EWGSAFTSDGR	IIRLFYTNRREPFTPESK-HYGQTLTTAQVNVS	EWGKGDKDLQVDGIEDH	229

tr A0A163L384 A0A163L384_9BACI	KSI	EGGDGKTYENVE-KAFEDR---NFN	NHHTFRDPHYIED	KGRKYLVFEANTGTEYG	282
tr A0AAX2EH38 A0AAX2EH38_9BACI	KSI	YDGGNGKTYENVN-KAFEDR---NFA	NHHTLRLDPHYIED	KGRKYLVFEANTGTEYG	285
tr A0A268AGA6 A0A268AGA6_9BACI	KSI	YDGGNGKTYENVN-KAFEDR---NFA	NHHTLRLDPHYIED	KGRKYLVFEANTGTEYG	285
tr A0A084GQG0 A0A084GQG0_METID	KSI	YDGGKDSKYYQTVD-KFAKDG---APN	NHHTFRDPHYIED	KGRKYLVFEANTGTEYG	285
tr A0A612MEF7 A0A612MEF7_9BACI	KSI	YDGGKDSKYYQTVD-KFAKDG---APN	NHHTFRDPHYIED	KGRKYLVFEANTGTEYG	285
tr A0A1H0C3I3 A0A1H0C3I3_9BACI	KSI	YDGGKDSKYYQTVD-KFAKDG---APN	NHHTFRDPHYIED	KGRKYLVFEANTGTEYG	285
tr A0A1N6Y262 A0A1N6Y262_9BACI	KSI	YDGGKDSKYYQTVD-KFAKDG---APN	NHHTFRDPHYIED	KGRKYLVFEANTGTEYG	285
tr A0AA96N005 A0AA96N005_9BACI	KSI	YDGGKDSKYYQTVD-KFAKDG---APN	NHHTFRDPHYIED	KGRKYLVFEANTGTEYG	285
tr A0A926NJ10 A0A926NJ10_9BACI	KSI	YDGGKDSKYYQTVD-KFAKDG---APN	NHHTFRDPHYIED	KGRKYLVFEANTGTEYG	285
tr A0A1G8FSH7 A0A1G8FSH7_9BACI	KSI	YDGGKDSKYYQTVD-KFAKDG---APN	NHHTFRDPHYIED	KGRKYLVFEANTGTEYG	285
tr A0A428MZE4 A0A428MZE4_9BACI	KSI	YDGGKDSKYYQTVD-KFAKDG---APN	NHHTFRDPHYIED	KGRKYLVFEANTGTEYG	285
tr A0A917B9N0 A0A917B9N0_HALAA	KSI	YDGGKDSKYYQTVD-KFAKDG---APN	NHHTFRDPHYIED	KGRKYLVFEANTGTEYG	285
tr A0A1R1RM81 A0A1R1RM81_9BACI	KSI	YDGGKDSKYYQTVD-KFAKDG---APN	NHHTFRDPHYIED	KGRKYLVFEANTGTEYG	285
tr A0A1H3V2M3 A0A1H3V2M3_9BACI	KSI	YDGGKDSKYYQTVD-KFAKDG---APN	NHHTFRDPHYIED	KGRKYLVFEANTGTEYG	285
tr A0A26S2X1 A0A26S2X1_9BACI	KSI	YDGGKDSKYYQTVD-KFAKDG---APN	NHHTFRDPHYIED	KGRKYLVFEANTGTEYG	285
tr A0A2N025D7 A0A2N025D7_9BACI	KSI	YDGGKDSKYYQTVD-KFAKDG---APN	NHHTFRDPHYIED	KGRKYLVFEANTGTEYG	285
tr A0AA91TQD8 A0AA91TQD8_NIACI	KSI	YDGGKDSKYYQTVD-KFAKDG---APN	NHHTFRDPHYIED	KGRKYLVFEANTGTEYG	285
tr A0A7X2V5Z1 A0A7X2V5Z1_9BACI	KSI	YDGGKDSKYYQTVD-KFAKDG---APN	NHHTFRDPHYIED	KGRKYLVFEANTGTEYG	285
tr A0A0M3RANO A0A0M3RANO_9BACI	KSI	YDGGKDSKYYQTVD-KFAKDG---APN	NHHTFRDPHYIED	KGRKYLVFEANTGTEYG	285
tr A0A90J7C6 A0A90J7C6_9BACI	KSI	YDGGKDSKYYQTVD-KFAKDG---APN	NHHTFRDPHYIED	KGRKYLVFEANTGTEYG	284
tr A0A3D8X658 A0A3D8X658_PRIMG	KSI	YDGGKDSKYYQTVD-QAFSEG---NF	NHHTFRDPHYIED	KGRKYLVFEANTGTEYG	284
tr A0A806U3J9 A0A806U3J9_PRIMG	KSI	YDGGKDSKYYQTVD-QAFSEG---NF	NHHTFRDPHYIED	KGRKYLVFEANTGTEYG	284
tr A0AAP8K2B5 A0AAP8K2B5_PRIAR	KSI	YDGGKDSKYYQTVD-QAFSEG---NF	NHHTFRDPHYIED	KGRKYLVFEANTGTEYG	284
tr A0AAP3FQJ4 A0AAP3FQJ4_BACVA	KSI	YDGGKDSKYYQTVD-QAFSEG---NF	NHHTFRDPHYIED	KGRKYLVFEANTGTEYG	284
sp P94468 LSCI_GEOSE	KSI	YDGGKDSKYYQTVD-QAFSEG---NF	NHHTFRDPHYIED	KGRKYLVFEANTGTEYG	284
sp P05655 LSC_BACSU	KSI	YDGGKDSKYYQTVD-QAFSEG---NF	NHHTFRDPHYIED	KGRKYLVFEANTGTEYG	284
tr B0N5W2 B0N5W2_9FIRM	KSI	YDGGKDSKYYQTVD-QAFSEG---NF	NHHTFRDPHYIED	KGRKYLVFEANTGTEYG	284
tr A0A9Q4DMY4 A0A9Q4DMY4_BACSC	KSI	YDGGKDSKYYQTVD-QAFSEG---NF	NHHTFRDPHYIED	KGRKYLVFEANTGTEYG	284
tr A0A5M8RY49 A0A5M8RY49_9BACI	KSI	YDGGKDSKYYQTVD-QAFSEG---NF	NHHTFRDPHYIED	KGRKYLVFEANTGTEYG	284
tr A0A7Z1B6S1 A0A7Z1B6S1_9BACI	KSI	YDGGKDSKYYQTVD-QAFSEG---NF	NHHTFRDPHYIED	KGRKYLVFEANTGTEYG	284
tr B6DVC6 B6DVC6_BACLI	KSI	YDGGKDSKYYQTVD-QAFSEG---NF	NHHTFRDPHYIED	KGRKYLVFEANTGTEYG	284
tr A0A90J7C6 A0A90J7C6_9BACI	KSI	YDGGKDSKYYQTVD-QAFSEG---NF	NHHTFRDPHYIED	KGRKYLVFEANTGTEYG	284
tr A0A089MCV7 A0A089MCV7_9BACI	KSI	YDGGKDSKYYQTVD-QAFSEG---NF	NHHTFRDPHYIED	KGRKYLVFEANTGTEYG	284
tr A0A919RYE8 A0A919RYE8_9CLOT	KSI	YDGGKDSKYYQTVD-QAFSEG---NF	NHHTFRDPHYIED	KGRKYLVFEANTGTEYG	284
tr A0A1ROXB9D9 A0A1ROXB9D9_9BACL	KSI	YDGGKDSKYYQTVD-QAFSEG---NF	NHHTFRDPHYIED	KGRKYLVFEANTGTEYG	284
tr A0A2Z2K7W7 A0A2Z2K7W7_9BACL	KSI	YDGGKDSKYYQTVD-QAFSEG---NF	NHHTFRDPHYIED	KGRKYLVFEANTGTEYG	284
tr E6U567 E6U567_ETHHY	KSI	YDGGKDSKYYQTVD-QAFSEG---NF	NHHTFRDPHYIED	KGRKYLVFEANTGTEYG	284
tr A0A917S1W0 A0A917S1W0_9BACI	KSI	YDGGKDSKYYQTVD-QAFSEG---NF	NHHTFRDPHYIED	KGRKYLVFEANTGTEYG	284
tr A0A1S8MP69 A0A1S8MP69_CLOSA	KSI	YDGGKDSKYYQTVD-QAFSEG---NF	NHHTFRDPHYIED	KGRKYLVFEANTGTEYG	284
tr A0A0H3J026 A0A0H3J026_CLOPA	KSI	YDGGKDSKYYQTVD-QAFSEG---NF	NHHTFRDPHYIED	KGRKYLVFEANTGTEYG	284
tr R4K9X0 R4K9X0_CLOPA	KSI	YDGGKDSKYYQTVD-QAFSEG---NF	NHHTFRDPHYIED	KGRKYLVFEANTGTEYG	284
tr R4K861 R4K861_CLOPA	KSI	YDGGKDSKYYQTVD-QAFSEG---NF	NHHTFRDPHYIED	KGRKYLVFEANTGTEYG	284

		Hydrolytic module ← → Transferase module	
r AOA163L384 AOA163L384_9BACI	L1VSNVTIDEIERPN	I1VKGKTSVVKDSILE QQQ1TYDQG-----	402
r AOAAX2EH38 AOAAX2EH38_9BACI	L1VSNVTIDEIERPN	I1VKGKTSVVKDSILE QQQ1TYEE-----	405
r AOA268AGA6 AOA268AGA6_9BACI	L1VSNVTIDEIERPN	I1VKGKTSVVKDSILE QQQ1TYDE-----	408
r AOA084GQG0 AOA084GQG0_METID	I1VKGKTSVVKDSILE QQQ1TDD-----	408	
r AOA612MEF7 AOA612MEF7_9BACI	I1VKGKTSVVKDSILE QQQ1TQD-----	405	
r AOA1N6Y262 AOA1N6Y262_9BACI	I1VKGKTSVVKDSILE QQQ1TQD-----	405	
r AOA96N005 AOA96N005_9BACI	I1VKGKTSVVKDSILE QQQ1TQD-----	405	
r AOA526NJI0 AOA526NJI0_9BACI	I1VKGKTSVVKDSILE QQQ1TQD-----	394	
r AOA1G8FSH7 AOA1G8FSH7_9BACI	I1VKGKTSVVKDSILE QQQ1TQD-----	407	
r AOA428MZE4 AOA428MZE4_9BACI	I1VKGKTSVVKDSILE QQQ1TQD-----	407	
r AOA917B9N0 AOA917B9N0_HALAA	I1VKGKTSVVKDSILE QQQ1TQD-----	407	
r AOA1R1RM81 AOA1R1RM81_9BACI	I1VKGKTSVVKDSILE QQQ1TQD-----	404	
r AOA1H3V2M3 AOA1H3V2M3_9BACI	I1VKGKTSVVKDSILE QQQ1TQD-----	404	
r AOA926520 AOA926520_9BACI	I1VKGKTSVVKDSILE QQQ1TQD-----	404	
r AOA2N025D7 AOA2N025D7_9BACI	I1VKGKTSVVKDSILE QQQ1TQD-----	405	
r AOA917QD8 AOA917QD8_NIACI	I1VKGKTSVVKDSILE QQQ1TQD-----	405	
r AOA1K2Y751 AOA1K2Y751_9BACI	I1VKGKTSVVKDSILE QQQ1TQD-----	399	
r AOA03RANO AOA03RANO_9BACI	I1VKGKTSVVKDSILE QQQ1TQD-----	407	
r AOA3D8X658 AOA3D8X658_PRIMG	I1VKGKTSVVKDSILE QQQ1TQD-----	400	
r AOA806U3J9 AOA806U3J9_PRIMG	I1VKGKTSVVKDSILE QQQ1TQD-----	400	
r AOAAP8K2B5 AOAAP8K2B5_PRIAR	I1VKGKTSVVKDSILE QQQ1TQD-----	399	
r AOAAP3FQJ4 AOAAP3FQJ4_BACVA	I1VKGKTSVVKDSILE QQQ1TQD-----	399	
p P94468 LSCI_GEOSE	I1VKGKTSVVKDSILE QQQ1TQD-----	390	
p P05655 LSC_BACSU	I1VKGKTSVVKDSILE QQQ1TQD-----	390SDM	
r B0N5W2 B0N5W2_9FIRM	I1VKGKTSVVKDSILE QQQ1TQD-----	390	
r AOA9Q4DMY4 AOA9Q4DMY4_BACSC	I1VKGKTSVVKDSILE QQQ1TQD-----	390	
r AOA5M8RY49 AOA5M8RY49_9BACI	I1VKGKTSVVKDSILE QQQ1TQD-----	390	
r AOA7Z1B6S1 AOA7Z1B6S1_9BACI	I1VKGKTSVVKDSILE QQQ1TQD-----	390	
r B6DVC6 B6DVC6_BACLI	I1VKGKTSVVKDSILE QQQ1TQD-----	390	
r AOAA90JC76 AOAA90JC76_9BACI	I1VKGKTSVVKDSILE QQQ1TQD-----	390	
r AOA089MCV7 AOA089MCV7_9BACI	I1VKGKTSVVKDSILE QQQ1TQD-----	390	
r AOA919RYE8 AOA919RYE8_9CLOT	I1VKGKTSVVKDSILE QQQ1TQD-----	390	
r AOA1R0XBD9 AOA1R0XBD9_9BACL	I1VKGKTSVVKDSILE QQQ1TQD-----	390	
r AOA222K7W7 AOA222K7W7_9BACI	I1VKGKTSVVKDSILE QQQ1TQD-----	390	
r E6U567 E6U567_ETHER	I1VKGKTSVVKDSILE QQQ1TQD-----	390	
r AOA917S1W0 AOA917S1W0_9BACI	I1VKGKTSVVKDSILE QQQ1TQD-----	390	
r AOA1S8MP69 AOA1S8MP69_CLOSA	I1VKGKTSVVKDSILE QQQ1TQD-----	390	
r AOA0H3J026 AOA0H3J026_CLOPA	I1VKGKTSVVKDSILE QQQ1TQD-----	390	
r R4K9X0 R4K9X0_CLOPA	I1VKGKTSVVKDSILE QQQ1TQD-----	390	
r R4K861 R4K861_CLOPA	I1VKGKTSVVKDSILE QQQ1TQD-----	390	

End of levensucrases and family GH68 fructosyl transferases		
r AOA163L384 AOA163L384_9BACI	I1VKGKTSVVKDSILE QQQ1TYDQG-----	486
r AOAAX2EH38 AOAAX2EH38_9BACI	I1VKGKTSVVKDSILE QQQ1TYEE-----	489
r AOA268AGA6 AOA268AGA6_9BACI	I1VKGKTSVVKDSILE QQQ1TYDE-----	488
r AOA084GQG0 AOA084GQG0_METID	I1VKGKTSVVKDSILE QQQ1TDD-----	488
r AOA612MEF7 AOA612MEF7_9BACI	I1VKGKTSVVKDSILE QQQ1TQD-----	486
r AOA1H0C3I3 AOA1H0C3I3_9BACI	I1VKGKTSVVKDSILE QQQ1TQD-----	485
r AOA1N6Y262 AOA1N6Y262_9BACI	I1VKGKTSVVKDSILE QQQ1TQD-----	487
r AOA96N005 AOA96N005_9BACI	I1VKGKTSVVKDSILE QQQ1TQD-----	488
r AOA926NJI0 AOA926NJI0_9BACI	I1VKGKTSVVKDSILE QQQ1TQD-----	489
r AOA1G8FSH7 AOA1G8FSH7_9BACI	I1VKGKTSVVKDSILE QQQ1TQD-----	490
r AOA428MZE4 AOA428MZE4_9BACI	I1VKGKTSVVKDSILE QQQ1TQD-----	490
r AOA917B9N0 AOA917B9N0_HALAA	I1VKGKTSVVKDSILE QQQ1TQD-----	487
r AOA1R1RM81 AOA1R1RM81_9BACI	I1VKGKTSVVKDSILE QQQ1TQD-----	487
r AOA1H3V2M3 AOA1H3V2M3_9BACI	I1VKGKTSVVKDSILE QQQ1TQD-----	487
r AOA926520 AOA926520_9BACI	I1VKGKTSVVKDSILE QQQ1TQD-----	486
r AOA2N025D7 AOA2N025D7_9BACI	I1VKGKTSVVKDSILE QQQ1TQD-----	498
r AOA917QD8 AOA917QD8_NIACI	I1VKGKTSVVKDSILE QQQ1TQD-----	487
r AOA7Z1B6S1 AOA7Z1B6S1_9BACI	I1VKGKTSVVKDSILE QQQ1TQD-----	493
r AOA03RANO AOA03RANO_9BACI	I1VKGKTSVVKDSILE QQQ1TQD-----	494
r AOA3D8X658 AOA3D8X658_PRIMG	I1VKGKTSVVKDSILE QQQ1TQD-----	484
r AOA806U3J9 AOA806U3J9_PRIMG	I1VKGKTSVVKDSILE QQQ1TQD-----	483
r AOAAP8K2B5 AOAAP8K2B5_PRIAR	I1VKGKTSVVKDSILE QQQ1TQD-----	483
r AOAAP3FQJ4 AOAAP3FQJ4_BACVA	I1VKGKTSVVKDSILE QQQ1TQD-----	473
p P94468 LSCI_GEOSE	I1VKGKTSVVKDSILE QQQ1TVNQ-----	473
p P05655 LSC_BACSU	I1VKGKTSVVKDSILE QQQ1TVNK-----	473
r B0N5W2 B0N5W2_9FIRM	I1VKGKTSVVKDSILE QQQ1TVNK-----	473
r AOA9Q4DMY4 AOA9Q4DMY4_BACSC	I1VKGKTSVVKDSILE QQQ1TINK-----	473
r AOA5M8RY49 AOA5M8RY49_9BACI	I1VKGKTSVVKDSILE QQQ1TVN-----	482
r AOA7Z1B6S1 AOA7Z1B6S1_9BACI	I1VKGKTSVVKDSILE QQQ1TVN-----	481
r B6DVC6 B6DVC6_BACLI	I1VKGKTSVVKDSILE QQQ1TVN-----	482
r AOAA90JC76 AOAA90JC76_9BACI	I1VKGKTSVVKDSILE QQQ1TVN-----	482
r AOA089MCV7 AOA089MCV7_9BACI	I1VKGKTSVVKDSILE QQQ1TVN-----	485
r AOA919RYE8 AOA919RYE8_9CLOT	I1VKGKTSVVKDSILE QQQ1TIDK-----	486
r AOA1R0XBD9 AOA1R0XBD9_9BACL	I1VKGKTSVVKDSILE QQQ1TLKGNKPY-----	494
r AOA222K7W7 AOA222K7W7_9BACI	I1VKGKTSVVKDSILE QQQ1TIEQEKNNKK-----	496
r E6U567 E6U567_ETHER	I1VKGKTSVVKDSILE QQQ1TIDGDQHRR-----	485
r AOA917S1W0 AOA917S1W0_9BACI	I1VKGKTSVVKDSILE QQQ1TADK-----	481
r AOA1S8MP69 AOA1S8MP69_CLOSA	I1VKGKTSVVKDSILE QQQ1TADK-----	484
r AOA0H3J026 AOA0H3J026_CLOPA	I1VKGKTSVVKDSILE QQQ1TADK-----	490
r R4K9X0 R4K9X0_CLOPA	I1VKGKTSVVKDSILE QQQ1TADK-----	495
r R4K861 R4K861_CLOPA	I1VKGKTSVVKDSILE QQQ1TTNK-----	488

A0A163L384_9BACI	Levansucrase, <i>Rossellomoreea marisflavi</i>
A0AAX2EH38_9BACI	Levansucrase, <i>Terribacillus saccharophilus</i>
A0A26AGA6_9BACI	GH 68 family protein, <i>Terribacillus saccharophilus</i>
A0A084GQG0_METID	Levansucrase, <i>Metabacillus indicus</i>
A0A6I2MEF7_9BACI	GH 68 family protein, <i>Metabacillus idriensis</i>
A0A1H0C3I3_9BACI	Levansucrase, <i>Fictibacillus solisalsi</i>
A0A1N6Y262_9BACI	Levansucrase, <i>Domibacillus encleensis</i>
A0AA96N005_9BACI	GH 68 family protein, <i>Domibacillus</i> sp.
A0A926NJI0_9BACI	GH 68 family protein, <i>Metabacillus arenae</i>
A0AA90JC76_9BACI	GH 68 family protein, <i>Bacillus haynesii</i>
A0A926S2X1_9BACI	GH 68 family protein, <i>Metabacillus arenae</i>
A0A1G8FSH7_9BACI	Levansucrase, <i>Alteribacillus bidgolensis</i>
A0A428MZE4_9BACI	GH 68 family protein, <i>Salibacterium salarium</i>
A0A917B9N0_HALAA	Levansucrase, <i>Halobacillus andaensis</i>
A0A1R1RM81_9BACI	GH 68 family protein, <i>Bacillus svezeyi</i>
A0A1H3V2M3_9BACI	Levansucrase, <i>Evansella caseinilytica</i>
A0A2N0Z5D7_9BACI	GH 68 family protein, <i>Niallia nealsonii</i>
A0AA91TQD8_NIACI	GH 68 family protein, <i>Niallia circulans</i>
A0A7X2V5Z1_9BACI	GH 68 family protein, <i>Metabacillus mangrovi</i>
A0A0M3RAN0_9BACI	Levansucrase, <i>Bacillus gobiensis</i>
A0A3D8X658_PRIMG	GH 68 family protein, <i>Priestia megaterium</i>
A0A806U3J9_PRIMG	Levansucrase, <i>Priestia megaterium</i>
A0AAP8K2B5_PRIAR	GH 68 family protein, <i>Prieszia aryabhatterai</i>
A0AAP3FQJ4_BACVA	Levansucrase, <i>Bacillus vallismortis</i>
P94468 LSCI_GEOSE	Inactive levansucrase, <i>Geobacillus stearothermophilus</i> , <i>sacB</i>
P05655 LSC_BACSU	Levansucrase, <i>Bacillus subtilis</i> , <i>sacB</i>
B0N5W2_9FIRM	Levansucrase, <i>Thomasoniella ramosa</i>
A0A9Q4DMY4_BACSC	Levansucrase, <i>Bacillus spizizenii</i>
A0A5M8RY49_9BACI	GH 68 family protein, <i>Bacillus svezeyi</i>
A0A7Z1B6S1_9BACI	GH 68 family protein, <i>Bacillus paralicheniformis</i>
B6DVC6_BACLI	Levansucrase, <i>Bacillus licheniformis</i>
A0AA96N005_9BACI	GH 68 family protein, <i>Domibacillus</i> sp
A0A089MCV7_9BACL	Levansucrase, <i>Paenibacillus graminis</i>
A0A919RYE8_9CLOT	GH 68 family protein, <i>Clostridium polyendosporum</i>
A0A1R0XBD9_9BACL	GH 68 family protein, <i>Paenibacillus odorifer</i>
A0A2Z2K7W7_9BACL	GH 68 family protein, <i>Paenibacillus donghaensis</i>
E6U567_ETHHY	GH 68 family protein, <i>Ethanoligenens harbinense</i>
A0A917S1W0_9BACL	GH 68 family protein, <i>Sporolactobacillus putidus</i>
A0A1S8MP69_CLOSA	Levansucrase, <i>Clostridium saccharobutylicum</i>
A0A0H3J026_CLOPA	Levansucrase, <i>Clostridium pasteurianum</i>
R4K9X0_CLOPA	Levansucrase, <i>Clostridium pasteurianum</i>
R4K861_CLOPA	Levansucrase, <i>Clostridium pasteurianum</i>

Figure 1 MSA of levansucrases and GH68 family glycoside hydrolase fructosyl transferases

3.2. MSA analysis of inulosucrases

Figure 2 shows the MSA of inulosucrases from different bacteria, (only the required regions for the discussions are shown). These are also bifunctional enzymes like the levansucrases and exhibit two different activities, viz. hydrolytic and transferase activities. The active site amino acids of hydrolytic and transferase domains are highlighted in yellow and green, respectively. Like the levansucrases, the inulosucrases are also composed of the same catalytic triad. The N-terminal is not conserved among them until it reaches the first catalytic site amino acid, D (nucleophile). The catalytic triad is found in the completely conserved motifs, the nucleophilic amino acid D is found in -WDSW-. The proposed Glc-binding motif is the conserved, -WSGSA- motif. In the other two highly conserved motifs, one harbours the acid-base catalytic amino acid E, -EI/VER- and the other is the proposed Fru-binding motif, -M/LRDP/AH-, where the -RD- pair is found invariant in all and shown to involve in catalysis by SDM. As found in the levansucrases, the catalytic triad in the second transferase domain of inulosucrases is also highly conserved, viz. the basic amino acid, K/R for proton abstraction, and a completely conserved fructose donor selection amino acid, R at -3, and a template-binding pair, -MI//VLG- at -22 from the proton abstractor and usually exists as a conserved triad. Interestingly, the active site amino acids in the transferase domain are the same as the nucleotidyl transferase domain found in the DNA/ RNA polymerases [1-3]. Most of the inulosucrases' active site amino acids are also confirmed by SDM experiments and X-ray crystallographic data. For example, in *Lactobacillus reuteri* the catalytic triad is composed of D²⁷², E⁵²³ and D⁴²⁴ (highlighted) and also confirmed by SDM analysis. The other three completely conserved residues, viz. W²⁷¹, W³⁴⁰ and R⁴²³ (highlighted) are located at the -1 subsite. Inulosucrase mutants W³⁴⁰→N and R⁴²³→H were virtually inactive, confirming the essential role played by these residues in the inulosucrase active site [29, 30].

SYW17616.1	TFDQDWQNNSYAYPEIDSSKVENLSAADMDSTPNVTGEVDTQAVQLDW	WDSW	PLQNAQ	270	
EOD02545.1	AKKIVNRDSSSQIPLFKGEKIVNMPGLEVKD	AE	---TGEEATMDI	WDSW PVQDPE 328	
tr Q7X481 Q7X481_LEUCI	IKLAQKDPKTTIPVFNASQIINNL PAS-IFKD	--AQ	---TNKVEKMDW	WDSW ALQDSK 363	
XAO78497.1	ADKLAEVDPQYAPIPYFKADQIQNLPAA-TAKD	--GQ	---TGKVANMDI	WDSW PVQDPT 329	
VTY24076.1	VASFLKQDSKLA VPYFKADTIINMPAF	-NTVD	--AQ	---TMEKEEIDV	WDSW PVQDAE 288
VUW83976.1	VASFLKQDSKLA VPYFKADTIINMPAF	-NTVD	--AQ	---TMEKEEIDV	WDSW PVQDAE 288
VTY27523.1	VASFLKQDSKLA VPYFKADTIINMPAF	-NTVD	--AQ	---TMKKEEIDV	WDSW PVQDAK 295
VTY49811.1	AQFLIAQDERYAI PYFNAAKAIKNMKA	-TTRD	--AQ	---TGQIADLDV	WDSW PVQDAK 254
CAH1856476.1	ADSLVDRDQRYAVPVFDAGKIELN PAA-YSRD	--AQ	---TGQYEHLDI	WDSW PVQNPQ 251	
pdb 8XVQ B	ADTLIKQDGRTV PFFFKASEIKNM PAA-TTKD	--AQ	---TNTIEPLDV	WDSW PVQDVR 280	
WVT67746.1	ADTLIKQDGRTV PFFFKASEIKNM PAA-TTKD	--AQ	---TNTIEPLDV	WDSW PVQDVR 280	
tr Q8GP32 Q8GP32_LIMRT	ADTLIKQDGRTV PFFFKASEIKNM PAA-TTKD	--AQ	---TNTIEPLDV	WDSW PVQDVR 280	
GFI19642.1	AKTLIEQDARYAI PFFNASKIKNM PAA-KTLD	--AQ	---SGKVEDLEI	WDSW PVQDAK 280	
ACZ67286.1	AKTLIEQDARYAVPFFNASKIKNM PAA-KTLD	--AQ	---TGKVEDLEI	WDSW PVQDAK 274	
	:	*	... : *: : *: .	
				... : *: .	

SYW17616.1	DVDYTYGVNQDTASAWVGANPLVQDSNLSFHENQPVEAEQ	WSGS	VLN SDN SIQLYYT SV	390
EOD02545.1	-----NDPSTQG	WSGS	AMLNADGSIQLFYTDV	407
tr Q7X481 Q7X481_LEUCI	-----AKPESQEW	WSGS	AVNSDDSIQLFYTRV	443
XAO78497.1	DKKT-----	GLEIFGDQEW	WSGSAYPLDDGSIQLFYTHS	416
VTY24076.1	-----	ALEDDQ	QWSGSATVNNSDGSIQLFYTMN	367
VUW83976.1	-----	ALEDDQ	QWSGSATVNNSDGSIQLFYTMN	367
VTY27523.1	-----	ALEDDQ	QWSGSATVNNSDGSIQLYYTKN	374
VTY49811.1	-----	ETPLTQEW	WSGSATVNEDGSIQLFYTKV	333
CAH1856476.1	-----	ASPLSQE	WSGS AIVNPDDSIQMFYTV	330
pdb 8XVQ B	-----	STAVSQE	WSGS AIVLN SDNSIQLFYTRV	359
WVT67746.1	-----	STAVSQE	WSGS AIVLN SDNSIQLFYTRV	359
tr Q8GP32 Q8GP32_LIMRT	-----	STAVSQE	WSGS AIVLN SDNSIQLFYTRV	359
GFI19642.1	-----	T-PVIQ	QWSGSATLNKDGSIQLYYTKV	358
ACZ67286.1	-----	T-PVIQ	QWSGSATLNKDGSIQLYYTKV	352
	*****	*	... : *: **	

SYW17616.1	NKQIFDYYESQ-EEKNGE--POVLNG	MRDPH	VIVV-NGQRYLAFESTTLVDNG--QGIVED	489
EOD02545.1	FSQWQNGTN-----SSVNNFQ	LRDAH	IFKDSDGTYYIAFETATGDLGDEAESS-AH	502
tr Q7X481 Q7X481_LEUCI	KDQLKP-----QADMFTL	RDPH	LIELDDGERYLTFEANTGIYDE--ASD-QQ	530
XAO78497.1	FDQWAHNVTMF DGHKEDFGGADNF	MRDPH	IVKDSQGNRYLVFEASTGDDY--QSE-DQ	530
VTY24076.1	YPKFMDTFNDDDHNDGIPDRADNYC	LRDPH	IIED-KGSRYLIFESNTGDENY--QGE-KQ	470
VUW83976.1	YPKFMDTFNDDDHNDGIPDRADNYC	LRDPH	IIED-KGSRYLIFESNTGDENY--QGE-KQ	470
VTY27523.1	YPKFMDTFNDDDHNDGIPDRADNYC	LRDPH	IIED-KGSRYLIFESNTGDENY--QGE-KQ	477
VTY49811.1	YQQWRSTFT-----	GADNIA	MRDPH VIEDENGDRYLVFEASTGTENY--QGE-DQ	429
CAH1856476.1	YPOQWRATNQ-----	GADNIA	LRD SHVVDDADGSRYLIFE GATGSQNY--QGE-HQ	424
pdb 8XVQ B	YDQWKATNK-----	GADNIA	MRDAH VIEDDNGDRYLVFEASTGLENY--QGE-DQ	453
WVT67746.1	YDQWKATNK-----	GADNIA	MRDAH VIEDDNGDRYLVFEASTGLENY--QGE-DQ	453
tr Q8GP32 Q8GP32_LIMRT	YDQWKATNK-----	GADNIA	MRDAH VIEDDNGDRYLVFEASTGLENY--QGE-DQ	453
GFI19642.1	YDQWKETNK-----	GADNIA	MRDAH VIDDDNGRNRYLVFEASTGTENY--QGD-DQ	454
ACZ67286.1	YQWKKTNK-----	GADNIA	MRDAH VIDDDKGDRYLVFEASTGTENY--QGA-DQ	448
	***	..	.* : ** *	.

	Hydrolase		Transferase			
SYW17616.1	E----NNNVT-----	VVLPPITATMAST	EIEP	PSI VPMNGRFILEVAT	YV	YSSIDD 582
EOD02545.1	MSDPKNPQVATDENG	KLVLYRPLVKT	YR	PSI IKLNGKYYLFTDGR	GR	LTATDTD 606
tr Q7X481 Q7X481_LEUCI	G-DYFNPILD-----	RLYKPLITAVGVT	DEIER	PSI PDI VPFNGRYYLFTD	SRFNE	SAADN 626
XAO78497.1	G-TENNPIVTA-----	EYDPIISAVMVS	DEIER	PSI PDI VKIGDTYYLFAASRLN	EGT	ND 642
VTY24076.1	D-NEKNP SVA-----	ELYTPLVTS	HMV	PSI PDI VKMGNRYYLFAASRIS	STD	DAE 566
VUW83976.1	D-NEKNP SVA-----	ELYTPLVTS	HMV	PSI PDI VKMGGRYYLFTASRIN	STD	DAE 566
VTY27523.1	D-NEKNP SVA-----	ELYTPLVTS	HMV	PSI PDI VKLGDGRYYLFTASRLN	EGS	NND 573
VTY49811.1	G-DKKTPEVD-----	QFYTPLL	SSTMV	PSI PDI VKIGGRYYLFAADTRL	N	GSNDY 525
CAH1856476.1	D-QGKTP TVA-----	QLYTPLI	TANMVS	PSI PDI VKLGNRYYLFAATRLN	EGS	GSNDY 520
pdb 8XVQ B	K-DEKNP KVA-----	ELYSPLI	SAPMVS	PSI PDI VKLGNRYYLFAATRLN	EGS	GSNDY 549
WVT67746.1	K-DEKNP KVA-----	ELYSPLI	SAPMVS	PSI PDI VKLGNRYYLFAATRLN	EGS	GSNDY 549
tr Q8GP32 Q8GP32_LIMRT	K-DEKNP KVA-----	ELYSPLI	SAPMVS	PSI PDI VKLGNKYYLFAATRLN	RGS	GSND 549
GFI19642.1	D-DVKNP SVA-----	KVYTP	PLISAPMVS	PSI PDI VRLGNKYYLFAAT	TRLN	EGS 550
ACZ67286.1	N-DTKNP GVE-----	KVYTP	PLISAPMVS	PSI PDI VRLGNKYYLFAAT	TRLN	EGS 544
	.	:	.*:..	**:*	..:..:***.	**..:..

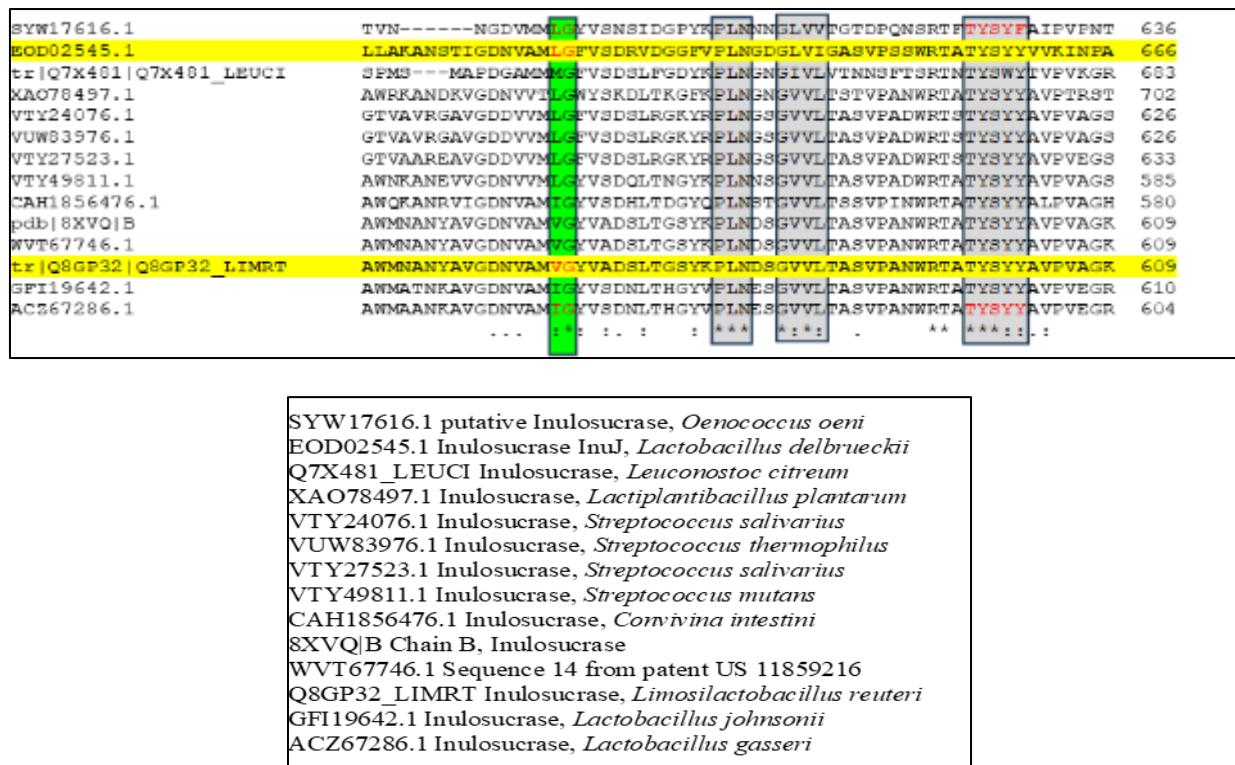


Figure 2 MSA of inulosucrases from different bacteria

3.3. 'Mix and Match' MSA analysis of levansucrases, inulosucrases and other Family GH68 fructosyl transferases

Figure 3 shows the 'Mix and Match' MSA of all the enzymes belonging to fructosyl transferases, viz. levansucrases, inulosucrases and other GH68 family glycoside hydrolases, (only the required regions for the discussions are shown). The active site amino acids of the hydrolytic and transferase domains from all three groups are highlighted in yellow and green, respectively. The active and inactive transferase enzyme from *B. subtilis* and *G. stereothermophilus* is highlighted in green and red, respectively. Strikingly, all three groups revealed high sequence similarities mostly at their catalytic sites only. Furthermore, it is interesting to note that the enzyme from *S. mutans*, the organism which causes dental caries in humans, synthesizes a β -(1 \rightarrow 2)-linked fructan (inulin type), whereas the *B. subtilis* (a nonpathogenic organism) produces a β -(2 \rightarrow 6)-linked polymer (levan type). As the hydrolytic site amino acids are the same, the structural difference of the fructans may be correlated with the sequence difference(s) around the transferase catalytic regions. For example, the comparison of the active site regions in levan and inulin synthetic enzymes show distinct differences in their distance motifs between the proton donor and the proposed template-binding pair. In fact, a longer distance motif between the proton abstractor and the template-binding pair is observed in inulosucrases of *S. mutans*, (e.g.), *B. subtilis* use, -SR³GSK¹MTIDGITSDIYML¹⁵GY- and *S. mutans* use, -SR³LNH¹GSNNDAWNKANEVVGDNVVML²¹GY-). It is interesting to note that in *S. mutans*, the EPS of dental plaque contains approximately 30% inulin, which protects the cells from antimicrobial materials present in saliva by forming biofilm [31].

The inulosucrase sequences start from the light blue highlighted sequence. Interestingly, they are larger proteins, and are almost double the size of the other two groups of transferases. There are big gaps in the N-terminal region and more or less no conserved amino acids or regions (data not shown) are observed until it reaches the first highly conserved tetrapeptide, -WDS/T/W/M-. This tetrapeptide harbours the invariant D, which acts as the catalytic nucleophile in the hydrolytic reaction. Again, there are gaps in the alignment until it reaches the second highly conserved tetrapeptide, -Q/D/EWSG/A- and is suggested to interact with the Glu unit of the sucrose molecule. This is followed by the next conserved tetrapeptide is -MF/LRDP/A/S- which is highly conserved in all three groups suggesting its importance. This tetrapeptide is suggested to interact with the Fru unit of the sucrose molecule. Substitution of Asp³⁰⁹ by Asn in the -FRD³⁰⁹P- motif of *A. diazotrophicus* levansucrase reduced the k_{cat} of the mutant enzyme 75-fold, indicating that Asp³⁰⁹ plays a major role in catalysis affecting the overall efficiency of sucrose hydrolysis. It is interesting to note that this tetrapeptide motif is conserved not only in the fructosyltransferases, but also in invertases, levanases, inulinases and sucrose-6-phosphate hydrolases, suggesting a common functional role for this motif in these enzymes [32]. Yanase *et al.* [33] found that the substitution of Asp¹⁹⁴, located in the -FRD¹⁹⁴P- motif of *Z. mobilis* levansucrase abolished sucrose

hydrolysis, suggesting its importance in catalysis. Furthermore, they found that the His²⁹⁶, in the transferase domain is crucial for catalysis of the transfructosylation reaction -FT-³ISH²⁹⁶HSTYADGLSGPDGVY¹⁶GF-. Not much conservation is observed until it reaches the next highly conserved tetrapeptide - QT/EI/VER-. The -ER- pair is completely conserved in all and the invariant -E- is known to act as the acid/base catalyst in the hydrolytic reaction.

Immediately after the hydrolytic module, the transferase active site domain is placed (highlighted in green). It is interesting to note that the transferase module is very similar to the nucleotidyl transferases module found in the DNA/RNA polymerases and composed of a similar catalytic triad, viz. a proton abstractor (usually a basic amino acid in all DNA/RNA polymerases and is conserved in the fructosyl transferases also) and a donor selection amino acid R at -3 to -4 from the proton abstractor and a template-binding pair -L/I/GY or -YG- (a highly conserved -YG- pair is shown as the template-binding pair in most of the DNA/RNA polymerases (Table 2). Interestingly, in all these group of fructosyl transferases a highly conserved template-binding triad, rather than a pair, is observed. The levansucrase from *G. stearothermophilus*, is reported to be inactive. The main difference between the active (from *B. subtilis*) and the inactive (from *G. stearothermophilus*) levansucrases is the proton abstractor, where the K is replaced by Q, respectively (marked in red). As the Q does not act as a proton abstractor in transferases, the enzyme could be inactive which is further confirmed by MSA analysis and SDM data. Interestingly, the same highly conserved amino acids at the catalytic sites of the two different transferases, viz. the fructosyl and nucleotidyl transferases suggest common evolutionary principles in designing active sites for same type of reactions in living cells.

CLUSTAL O (1.2.4) 'Mix and Match' MSA of levan- and inulosucrases and other GH68 family glycoside hydrolase fructosyl transferases.

sp Q43998 LSC_GLDI	PNI PADF-----PVINPDWVW WDTW ELIDIK-HADQ----FSYNGWEVIECILT	159
tr B2TF781B2F78_BEIIS	PD1 PADF-----POTNPDVWWDIWELADW-HONO----LSFOGWEVIFSLT	124
tr D0J9C21D0J9C2_HALJD	PIIYPPR-----EDAAPEIING WDTW PLER-DGSI----ATVGGWVIFSLT	71
tr AOA0M3KKU6 AOA0M3KKU6_ERMAE	PVIDIAF-----PVMSEEVIT WDTW PLRDF-DGEI----ISVNGWCIIFTLT	70
tr B2VCC3 B2VCC3_ERWT9	FLVDVAF-----PVMSEEVIT WDTW PLRDF-DGDI----SVSNNGCVIFLT	70
SYW17616.1	ENLSRADMVDSTPNVTTGEVDTQAVOLDW WDTW PLQNAOTGAAOSITNGKRYQIVIAMA	291
tr AOA163L304 AOA163L304_9BACI	KNIVESAT-K-----TDEMNKNIKMDW WDTW PLQNA-DGTV----AEYKGYHIVFGLA	117
tr AOAAX2MH38 AOAAX2MH38_9BACI	KNIKSAT-K-----IDENGNEIKMDW WDTW PLQNP-DGTV----ADYNGYQ1VFGLA	122
tr AOA268AGA6 AOA268AGA6_9BACI	KNIKSAT-K-----IDENGNEVKMDW WDTW PLQNP-DGTV----ADYNGYQ1VFGLA	122
tr AOA084Q000 AOA084Q000_METID	ENIPSAK-K-----TTESETMDLWD WDTW PLQNA-DGTV----AEYKGYHIVFGLA	120
tr AOA612MEF7 AOA612MEF7_9BACI	ENIPSAK-K-----VTESETIDMDW WDTW PLQNA-DGTV----AEYKGYHIVFGLA	120
tr AOA1H0C313 AOA1H0C313_9BACI	ENIPSAK-K-----YDKDGNL1KMDW WDTW PLQNA-DGTV----ANVKGYNLVFGLA	119
tr AOA1N6Y262 AOA1N6Y262_9BACI	RNI PAAT-K-----TDEMNNTIMMDW WDTW PLQNT-DGTV----AEYNGYHIVFGLA	120
tr AOA96N005 AOA96N005_9BACI	KNIPAAQ-K-----VDEMNQNLMDW WDTW PLQNA-DGTV----AEYNGYHIVFGLA	120
tr AOA926NJI0 AOA926NJI0_9BACI	KNIPSAK-G-----LDW WDTW PLQNA-DGTV----AEYKGYQ1VFGLA	110
tr AOA1G0F5M AOA1G0F5M_9BACI	ENIPSAK-K-----INEEGSEIDMDW WDTW PLQNA-DGTV----AEYNGYH1LFGLA	120
tr AOA428M2E4 AOA428M2E4_9BACI	ENIPSAK-K-----TGENQNEIVMDW WDTW PLQNA-DGTV----AEYNGYH1LFGLA	120
tr AOA917B9M0 AOA917B9M0_HALLA	DNIPSAK-K-----TDEMNNEIVMDW WDTW PLQNA-DGTV----AEYNGYH1LFGLA	120
tr AOA1R1RM01 AOA1R1RM01_9BACI	KNIPSAK-K-----YDENGKMLIDW WDTW PLQNA-DGTV----AEYKGYHIVFGLA	120
tr AOA1M3V2M3 AOA1M3V2M3_9BACI	VNT P5AQ-K-----YDENGNMLIDW WDTW PLQNT-DGTV----AEYNGFMTVFGLA	120
tr AOA926S2K1 AOA926S2K1_9BACI	KNIPSAK-G-----YDEMGNV1DMDW WDTW PLQNA-DGTV----ANVNGYQ1VFGLA	120
tr AOA2N025D7 AOA2N025D7_9BACI	KNIPSAK-K-----TDEMGNV1DMDW WDTW PLQNA-DGTV----AEYNGYQ1LFGLA	120
tr AOA91TQD8 AOA91TQD8_NIACI	KNIPSAK-G-----YDEKGNL1DMDW WDTW PLQNA-DGTV----AEYNGYQ1VFGLA	120
tr AOA7x2v5z1 AOA7x2v5z1_9BACI	KNIPSAK-G-----FDENGKNL1DMDW WDTW PLQNA-DGTV----ADYKGYQ1LFGLA	119
tr AOA0M3RANO AOA0M3RANO_9BACI	RNI PTAQ-C-----YDEKGNTFTLWD WDTW PLQNA-DGTV----AEYNGYH1LFTLA	119
tr AOA806U3J9 AOA806U3J9_PRIM0	KNI ASAK-G-----YDKSNSL1DMDW WDTW PLQNA-DGTV----ANVHGYQ1VFGALA	119
tr AOAAPP2K2 AOAAPP2K2_PRIM0	KNI ASAK-G-----YDKSGNL1DMDW WDTW PLQNA-DGTV----ANVHGYQ1VFGALA	119
tr AOA3D0X650 AOA3D0X650_PRIM0	KNI ASAK-G-----RDAGNT1DMDW WDTW PLQNA-DGTV----ATVHGYQ1VFGALA	119
tr DSDC071DSDC07_PRIM3	KNI ASAK-G-----ENASONT1DMDW WDTW PLQNA-DGTV----ATVHGYQ1VFGALA	119
tr AOAAPP3FOJ4 AOAAPP3FOJ4_BACVA	KNI SAK-G-----LDW WDTW PLQNA-DGTV----ANVHGYHIVFALA	110
sp P94460 LGCT_GEOGE	KNI SAK-G-----LDW WDTW PLQNA-DGTV----ANYHGYHIVFALA	110
sp P056551 LGCT_GEOGE	KNI SAK-G-----LDW WDTW PLQNA-DGTV----ANYHGYHIVFALA	110
tr B0G5W21H0N52_9FIRM	KNI SAK-G-----LDW WDTW PLQNA-DGTV----ANYHGYHIVFALA	110
tr AOA9Q4DMY4 AOA9Q4DMY4_BACSC	KNI SAK-G-----LDW WDTW PLQNA-DGTV----ANYHGYHIVFALA	110
tr AOA5M0R49 AOA5M0R49_9BACI	QNT P5AK-G-----YNS5SGKEL1DMDW WDTW PLQNA-DGTV----ATVNGYH1LFGLA	117
tr AOA7Z1H6S1 AOA7Z1H6S1_9BACI	KNIPMAK-G-----YNKNGEL1DMDW WDTW PLQNA-DGTV----ATVHGYNLVFGLA	117
tr BEDVC6 BEDVC6_BACLI	KNIPSAK-G-----YNKNGEL1DMDW WDTW PLQNA-DGTV----ATVHGYNLVFGLA	117
tr AOA2A90JCT6 AOA2A90JCT6_9BACI	KNIPSAK-G-----YNKNGEL1DMDW WDTW PLQNA-DGTV----ATVHGYNLVFGLA	117
tr AOA089MCV7 AOA089MCV7_9BACI	QNT SAK-G-----SDOWNSNAIDMDW WDTW PLQNA-DGTV----AEYNGYHIVFALA	119
tr AOA919YE0 AOA919YE0_9CLOT	KNIPSAI-K-----YDOWNSNQIDMDW WDTW PLQNA-DGTV----ANVNGYHIVFGLA	110
tr AOA1R0XBDS AOA1R0XBDS_9BACI	ENIPAAI-K-----EDENONKE1MDW WDTW PLQNA-DGTV----ASVKGYQ1VFGLA	121
tr AOA2Z2K7W7 AOA2Z2K7W7_9BACI	KNIPSAI-K-----VDESGNPIEMDW WDTW PLQNA-DGTV----ANVNGYHIVFGLA	119
tr EUS671E6E67_ERHMY	KNLH-VO-G-----PDEMGNPVFLDW WDTW PLQNA-DGTV----ATVNGYHIVFALA	115
tr AOA917S1W0 AOA917S1W0_9BACI	KNIPAAK-G-----LDW WDTW PLQNA-DGTA----ANVHGYNLVFGALA	110
tr AOA1R8MP69 AOA1R8MP69_CLOPA	KNILTAK-G-----YDPMAGNL1DMDW WDTW PLQNA-DGTV----ANVNGYNLVFGALA	118
tr AOA0H3J026 AOA0H3J026_CLOPA	KNIPSAK-G-----YDEAQNLI1DMDW WDTW PLQNA-DGTV----SNVHGYGYNIVFALA	110
tr RAK9X01R4K861_CLOPA	KNIPSAK-G-----YDEFGNL1DMDW WDTW PLQNA-DGTV----ANVHGYNLVFGALA	110
tr RAK861 R4K861_CLOPA	KNI SAK-G-----YDKYGNL1DMDW WDTW PLQNA-DGTV----ANVHGYNLVFGALA	118
ED002545.1	VNM PGLKVK-----DAETGCAEATMIDW WDTW PLQNA-DGTV----ENWNGYQ1LVVAMI	345
XAO78497.1	QNL PAAT-AK-----DGQTCVAKMDW WDTW PLQNTGEI----ANWNCGQLVIAAM	346
V7Y24076.1	INM PAFN-TV-----DAOTMEKEE1DMDW WDTW PLQDAE SOLV----SNWNGYQ1LVIAAM	305
VUW83976.1	INM PAFN-TV-----DAOTMEKEE1DMDW WDTW PLQDAE SOLV----SNWNGYQ1LVIAAM	305
V7Y27523.1	INM PAFN-TV-----DAQTMKKEE1DMDW WDTW PLQDAE SOLV----SNWNGYQ1LVIAAM	312
sp Q70XJ9 LSC_PRUSA	KNM PAFN-TV-----DAQTCMMAELDW WDTW PLQDAE SOLV----SNWNGYQ1LVIAAM	336
V7Y49811.1	KNM PAFN-TV-----DAOTGQIAADLW WDTW PLQDAE SOLV----INWNGYQ1LVVAMM	271
CAM1056476.1	ENLPAAV-SR-----DAQTCMKEE1DMDW WDTW PLQDAE SOLV----VNVHGYQ1LVVAMM	260
CAM1056006.1	QNP PAFN-TK-----DAQTNWVHMLDW WDTW PLQDAE SOLV----SNWNGYQ1LVVAMM	254
pdb18XVO1B	KNM PAFN-TK-----DAOTNTIEFLDW WDTW PLQDAE SOLV----ANWNGYQ1LVIAAM	297
WT67746.1	KNM PAFN-TK-----DAOTNTIEFLDW WDTW PLQDAE SOLV----ANWNGYQ1LVIAAM	297
tr Q8GP32 Q8GP32_LIMRT	KNM PAFN-TK-----DAQTNWVHMLDW WDTW PLQDAE SOLV----ANWNGYQ1LVIAAM	297
SP119642.1	KNM PAFN-TL-----DAQSGKVKEDLW WDTW PLQDAE SOLV----SNWNGYQ1LVIGMM	297
AC267286.1	KNM PAFN-TL-----DAQTCMKEE1DMDW WDTW PLQDAE SOLV----SNWNGYQ1LVIGMM	291

//End of levan-, inulosucrases and family GH68 fructosyl transferases	
sp Q43998 LSC_GLUDI	----- 524
tz B2IF78 B2IF78_BEII9	----- 524
tz D8J9C2 D8J9C2_HALJB	----- 428
tz A0A0M3KKU6 A0A0M3KKU6_ERWAE	----- 415
tz B2VCC3 B2VCC3_ERWT9	----- 415
SYW17616.1	KR8NGMYYTYAQVLDGNKTYWIDQRALSTKKK 1164
tz A0A163L384 A0A163L384_9BACI	----- 486
tz A0AACX2EH38 A0AACX2EH38_9BACI	----- 489
tz A0A268AGA6 A0A268AGA6_9BACI	----- 488
tz A0A084GQG0 A0A084GQG0_METID	----- 488
tz A0A6I2MEF7 A0A6I2MEF7_9BACI	----- 486
tz A0A1H0C313 A0A1H0C313_9BACI	----- 485
tz A0A1N6Y262 A0A1N6Y262_9BACI	----- 487
tz A0A96N005 A0A96N005_9BACI	----- 488
tz A0A926NJI0 A0A926NJI0_9BACI	----- 489
tz A0A1G8FSH7 A0A1G8FSH7_9BACI	----- 490
tz A0A428MZE4 A0A428MZE4_9BACI	----- 490
tz A0A917B9N0 A0A917B9N0_HALAA	----- 490
tz A0A1R1RM81 A0A1R1RM81_9BACI	----- 487
tz A0A1H3V2M3 A0A1H3V2M3_9BACI	----- 487
tz A0A926S2X1 A0A926S2X1_9BACI	----- 486
tz A0A2N0Z5D7 A0A2N0Z5D7_9BACI	----- 498
tz A0A91TQD8 A0A91TQD8_NIACI	----- 487
tz A0A7X2V5Z1 A0A7X2V5Z1_9BACI	----- 493
tz A0A0M3RAN0 A0A0M3RAN0_9BACI	----- 494
tz A0A806U3J9 A0A806U3J9_PRIMG	----- 483
tz A0AAP8K2B5 A0AAP8K2B5_PRIAR	----- 483
tz A0A3D8X658 A0A3D8X658_PRIMG	----- 484
tz D5DC07 D5DC07_PRIM3	----- 484
tz A0AAP3FQJ4 A0AAP3FQJ4_BACVA	----- 473
sp P944681 LSC_GEOSE	----- 473
sp P056551 LSC_BACSU	----- 473
tz B0NSW2 B0N5W2_9FIRM	----- 473
tz A0A8Q4DMY4 A0A8Q4DMY4_BACSC	----- 473
tz A0A5MSRY49 A0A5MSRY49_9BACI	----- 482
tz A0A7Z1B6S1 A0A7Z1B6S1_9BACI	----- 481
tz B6DVC6 B6DVC6_BACLI	----- 482
tz A0A90J0C76 A0A90J0C76_9BACI	----- 482
tz A0A089MCV7 A0A089MCV7_9BACL	----- 485
tz A0A919YE8 A0A919YE8_SCLOT	----- 486
tz A0A1ROXBD9 A0A1ROXBD9_9BACL	----- 494
tz A0A2Z2K7W7 A0A2Z2K7W7_9BACL	----- 496
tz E6U5671 E6U5671_ETHHY	----- 485
tz A0A917S1W0 A0A917S1W0_9BACL	----- 481
tz A0A1SSMP69 A0A1SSMP69_CLOSA	----- 484
tz A0A0H3J026 A0A0H3J026_CLOFA	----- 490
tz R4K9X0 R4K9X0_CLOFA	----- 495
tz R4K861 R4K861_CLOFA	----- 488
E0D02945.1	----- 1028
XAO78497.1	----- 937
VTY24076.1	----- 935
VWU83976.1	----- 935
VTY27523.1	----- 939
sp Q70XJ9 LSC_FRUSA	----- 779
VTY49881.1	----- 795
CAH188676.1	----- 663
CAH188606.1	----- 734
SDB187Q1	----- 736
WVT67746.1	----- 796
tz Q8GP32 Q8GP32_LIMRT	----- 796
GTI19642.1	----- 802
ACZ67286.1	----- 761

NB: The SDM and X-ray crystallographic data are highlighted in dark blue and light blue in bold, respectively.

B2IF78_BEII9 Levansucrase, <i>Beijerinckia indica</i>
D8J9C2_HALJB Levansucrase, <i>Halalkalicoccus jeotgali</i>
A0A0M3KKU6_ERWAE levansucrase, <i>Erwinia amylovora</i>
B2VCC3_ERWT9 levansucrase, <i>Erwinia tasmaniensis</i>
SYW17616.1 putative Inulosucrase, <i>Oenococcus oeni</i>
A0A163L384_9BACI Levansucrase, <i>Rosselloomorea marisflavi</i>
A0AACX2EH38_9BACI Levansucrase, <i>Terribacillus saccharophilus</i>
A0A268AGA6_9BACI GH68 family protein, <i>Terribacillus saccharophilus</i>
A0A084GQG0_METID Levansucrase, <i>Metabacillus indicus</i>
A0A6I2MEF7_9BACI GH68 family protein, <i>Metabacillus idriensis</i>
A0A1H0C313_9BACI Levansucrase, <i>Fictibacillus solisalsi</i>
A0A1N6Y262_9BACI Levansucrase, <i>Domibacillus encleensis</i>
A0AA96N005_9BACI GH68 family protein, <i>Domibacillus</i> sp.
A0A926NJI0_9BACI GH68 family protein, <i>Metabacillus arenae</i>
A0A1G8FSH7_9BACI Levansucrase, <i>Alteribacillus bidgolensis</i>
A0A428MZE4_9BACI GH68 family protein, <i>Salibacterium salarium</i>
A0A917B9N0_HALAA Levansucrase, <i>Halobacillus andensis</i>
A0A1R1RM81_9BACI GH68 family protein, <i>Bacillus svezeyi</i>
A0A1H3V2M3_9BACI Levansucrase, <i>Evansella caseinilytica</i>
A0A926S2X1_9BACI GH68 family protein, <i>Metabacillus arenae</i>
A0A2N0Z5D7_9BACI GH68 family protein, <i>Niella nealsonii</i>
A0AA91TQD8_NIACI GH68 family protein, <i>Niella circularis</i>
A0A7X2V5Z1_9BACI GH68 family protein, <i>Metabacillus mangrovi</i>
A0A0M3RAN0_9BACI Levansucrase, <i>Bacillus gobiensis</i>
A0A806U3J9_PRIMG Levansucrase, <i>Priestia megaterium</i>
A0AAP8K2B5_PRIAR GH68 family protein, <i>Priestia aryabhattai</i>
A0A3D8X658_PRIMG GH68 family protein, <i>Priestia (Bacillus) megaterium</i>
D5DC07_PRIM3 Levansucrase, <i>Priestia (Bacillus) megaterium</i>
A0AAP3FQJ4_BACVA Levansucrase, <i>Bacillus vallismortis</i>
P944681_LSC_GEOSE inactive levansucrase, <i>Geobacillus stearothermophilus</i> , <i>sacB</i>
P056551_LSC_BACSU Levansucrase, <i>Bacillus subtilis</i> , <i>sacB</i>

B0N5W2_9FIRM Levansucrase, <i>Thomasmavelia ramosa</i>
A0A9Q4DMY4_BACSC Levansucrase, <i>Bacillus spizizenii</i>
A0A5M8RY49_9BACI GH68 family protein, <i>Bacillus svezeyi</i>
B6DVC6_BACLI Levansucrase, <i>Bacillus licheniformis</i>
A0AA90JC76_9BACI GH68 family protein, <i>Bacillus haynesii</i>
A0A089MCV7_9BACI Levansucrase, <i>Paenibacillus graminis</i>
A0A919RYE8_9CLOT GH68 family protein, <i>Clostridium polyendosporian</i>
A0A1R0XBD9_9BACI GH68 family protein, <i>Paenibacillus odorifer</i>
A0A2Z2K7W7_9BACI GH68 family protein, <i>Paenibacillus donghaensis</i>
E6U567_ETHHY GH68 family protein, <i>Ethanoligenes harbinense</i>
A0A917S1W0_9BACI GH68 family protein, <i>Sporolactobacillus putidus</i>
A0A1S8MP69_CLOSA Levansucrase, <i>Clostridium saccharobyticum</i>
A0A0H3J026_CLOPA Levansucrase, <i>Clostridium pasteurianum</i>
A0A7Z1B6S1_9BACI GH68 family protein, <i>Bacillus paralicheniformis</i>
R4K9X0_CLOPA Levansucrase/Invertase, <i>Clostridium pasteurianum</i>
R4K861_CLOPA Levansucrase/Invertase, <i>Clostridium pasteurianum</i>
EOD02545.1 Inulosucrase InuJ, <i>Lactobacillus delbrueckii</i> subsp. <i>jakobsonii</i>
XAO78497.1 Inulosucrase, <i>Lactiplantibacillus plantarum</i>
VTY24076.1 Inulosucrase, <i>Streptococcus salivarius</i>
VUW83976.1 Inulosucrase, <i>Streptococcus thermophilus</i>
VTY27523.1 Inulosucrase, <i>Streptococcus salivarius</i>
Q70XJ9L SC_FRUSA Levansucrase, <i>Fructilactobacillus sanfranciscensis</i>
VTY49811.1 Inulosucrase, <i>Streptococcus mutans</i>
CAH1856476.1 Inulosucrase, <i>Comvivina intestini</i>
CAH1856006.1 Inulosucrase, <i>Comvivina praedatoris</i>
pdb8XVQJB Chain B, Inulosucrase
WVTI67746.1 Sequence 14 from patent US 11859216
Q8GP32_LIMRT Inulosucrase, <i>Limosilactobacillus reuteri</i>
GFI19642.1 Inulosucrase, <i>Lactobacillus johnsonii</i>
ACZ67286.1 Inulosucrase, <i>Lactobacillus gasseri</i>

Figure 3 'Mix and Match' MSA of all three groups of fructosyl transferases: levansucrases, inulosucrases and GH68 family glycoside hydrolases from Gram-positive and Gram-negative bacteria

4. Analyses of active site amino acids in the hydrolytic and transferase domains

The amino acids in the hydrolytic and transferase active sites are confirmed by SDM experiments, X-ray crystallographic and MSA analyses. All three groups, viz. the levansucrases, inulosucrases and other GH68 family of glycoside hydrolases invariably use three highly conserved amino acids in their active sites, as catalytic triad, for hydrolysis of sucrose molecule and another catalytic triad for subsequent fructosyl transferase activity.

4.1. Catalytic site amino acids in the hydrolase domain of levansucrases and inulosucrases

As discussed elsewhere, the hydrolytic domain's catalytic triad consists of two invariant Ds and an E in all three groups of enzymes. *B. subtilis* levansucrase (SacB) uses D⁸⁶, D²⁴⁷ and E³⁴² as the catalytic triad and they are shown to act as the nucleophile, transition-state stabilizer and the general acid-base catalyst (proton donor), respectively. On modification of the above amino acids to Ala by SDM, the enzyme lost its activity and thus, confirming their direct involvement in catalysis [18].

In the same way, the levansucrases from *L. reuteri* uses D²⁴⁹, D⁴⁰⁴ and E⁵⁰³ as the catalytic triad in its hydrolytic domain. On modification of the above amino acids to their amide forms, viz. D²⁴⁹→N, D⁴⁰⁴→N and E⁵⁰³→Q by SDM, about 10,000-fold reduction in sucrose hydrolytic activity was observed and thus, confirming their role on hydrolysis of sucrose molecule [34].

In *B. megaterium* levansucrase, the residues Asp⁹⁵, Asp²⁵⁷ and Glu³⁵² are proposed to form the catalytic triad. After the substitution of Ala to the proposed catalytic residues by SDM (Asp⁹⁵→A, Asp²⁵⁷→A, Glu³⁵²→A), it was found that the enzyme had lost its activity [35] and thus, confirming their direct involvement in the sucrose hydrolysis. Interestingly, the variants R²⁵⁶→A and E³⁵⁰→A were also nearly inactive.

Polsinelli *et al.* [36] have determined the 3D structure of a levansucrase from *Erwinia tasmaniensis* and found the catalytic triad involved in catalysis are, Asp⁴⁶, Asp²⁰³, and Glu²⁸⁷.

Not only the levansucrases, but also the inulosucrases also use the same active site amino acids in their hydrolytic domain. Ozimek *et al.* [34] proposed that the inulosucrase of *L. reuteri* might use D²⁷², D⁴²⁴ and E⁵²³ in the catalytic site of its hydrolytic module. This was further confirmed by SDM experiments. For example, the substitution of Ala to the proposed catalytic residues (Asp²⁷²→A, Asp⁴²⁴→A, Glu⁵²³→A), led to the loss of enzyme activity.

Pijning *et al.* [37] found that the three putative catalytic residues (D²⁷², D⁴²⁵, and E⁵²⁴) in an inulosucrase from *Lactobacillus johnsonii*. These three amino acids superposed almost perfectly with those of *B. subtilis* SacB and thus, confirming both the enzymes use the same amino acids at the active sites of their hydrolytic domains. Furthermore, they found that D²⁷²→N, a nucleophilic mutant of the enzyme, exhibited only a very low residual activity and the mutant/native enzyme activity ratio was about 8×10⁻⁶.

The crystal structure of an inulosucrase from a halophilic archaeal strain, *Halalkalicoccus jeotgali*, also exhibited the same pattern of amino acids at their hydrolytic domain. The catalytic triad is found to be composed of residues D⁴⁷, D¹⁹⁷, and E²⁶⁸ [38]. Furthermore, the 3D structure of the halophilic enzyme resembled closely to other fructansucrase structures.

The crystal structure of a fructosyl transferase from the fungus, *Aspergillus japonicus*, has shown that the amino acids in the catalytic triad are, Asp⁶⁰, Asp¹⁹¹, and Glu²⁹² and these three act as nucleophile, transition-state stabilizer, and general acid/base catalyst, respectively in the fungal enzyme and further suggested that they also governed the binding of the terminal fructose at the -1 subsite for the catalytic reaction. The active site amino acids in the fungal enzyme were further confirmed by SDM analysis. It was found that the mutants D⁶⁰→A, D¹⁹¹→A, and E²⁹²→A completely lost their activity [27].

MSA analysis has also confirmed that these three completely conserved amino acids constitute the catalytic triad in the hydrolytic site of all these fructosyl transferases (Fig. 3).

4.2. Catalytic site amino acids in the transferase domain of levansucrases and inulosucrase

The catalytic site amino acids in the transferase active site are arrived at from SDM experiments, X-ray crystallographic data and MSA analysis. As already discussed elsewhere, the fructosyl transferase from *G. stearothermophilus* was found to be inactive due to replacement of a K³⁶³ to Q³⁶³ in the proposed transferase domain. The importance of this highly conserved basic amino acid in this domain was further confirmed from analysis of hundreds of fructosyl transferase sequences from levansucrases, inulosucrases and other GH68 family glycoside hydrolase enzymes from a large number of organisms (Fig. 3). It is interesting to note that it has been shown a positively charged amino acid, viz. K/R/H is crucial for the functioning of nucleotidyl transferase catalytic reactions in DNA/RNA polymerases [1-3] where it initiates the catalytic reaction by abstraction of a proton from the substrate. Furthermore, in both the types of transferases, an invariant R at -3 or -4 from the proton abstractor (Tables 1 and 2) acts in the donor selection. The involvement of a basic amino acid as the proton abstractor and an R at -3 or -4 from the proton abstractor as nucleotide/fructose donor selection amino acid are further confirmed by SDM experiments, MSA analysis, and also from X-ray crystallographic data. The third template-binding pair (L/I/V/YG) in fructosyl transferases is suggested based on the sequence similarity from a large number of nucleotidyl transferases (DNA/RNA polymerases) (Tables 1 and 2 [1-3]).

The above conclusions are further substantiated by SDM experiments on large number of fructosyl and nucleotidyl transferases. Meng and Fütterer [18] found that SDM studies on the levansucrase of *B. subtilis* SacB revealed that modifications of the residues N²⁴², K³⁶³, and Y²³⁷ (numbering of *B. subtilis* levansucrase), which are located on the surface of the catalytic cavity affect the enzyme catalytic efficiency, transfructosylation, and hydrolysis ratio by interrupting the polymerization process. Therefore, they suggested that these residues could control the chain lengths of levans. However, they concluded that the interactions mediated by these residues had not yet been identified, and the existence of additional external acceptor-binding subsites, acting as structural determinants in the elongation of levans had not yet been clarified. Furthermore, in chain A, the interactions might resemble an acceptor anchoring state, with R³⁶⁰ and K³⁶³ maintaining direct contact with fructosyl-2 of the ligand, while Y⁴¹¹ might coordinate a water-mediated contact. Therefore, it is clear that these two residues, R³⁶⁰ and K³⁶³ play an important role in the transferase active site, which is also further supported by the present MSA analysis.

Further SDM experiments have conclusively proved that these two amino acids in the transferase catalytic site play a crucial role in levan synthesis. For example, the *B. megaterium* levansucrase SacB variants, viz. Y²⁴⁷→A, Y²⁴⁷→W, N²⁵²→A, D²⁵⁷→A, and K³⁷³→A revealed novel surface motifs remote from the sucrose-binding site with distinct influence on the polysaccharide product spectrum. Consistent with the eliminated polysaccharide synthesis, the K³⁷³→A mutant showed an increased hydrolytic activity of almost 33%. Furthermore, the five structures of the SacB variants (Y²⁴⁷→A, Y²⁴⁷→W,

$N^{252}\rightarrow A$, $D^{257}\rightarrow A$, and $K^{373}\rightarrow A$) obtained at resolutions between 2.0 and 1.75 Å supported a surface-modulated transfructosylation mechanism [39]. It is interesting to note that the D^{257} is in the highly conserved -LRD²⁵⁷P- motif of Fru-binding site and the others are in its vicinity and the K^{373} is implicated as the proton abstractor in the transferase active site. Homann *et al.* [35] also found that the R^{370} at -3 from the proton abstractor is crucial in the levansucrase of *B. megaterium* for the transferase activity. Furthermore, they found that the residues R^{370} and N^{252} seem to be crucial for polyfructan synthesis and are conserved in all levansucrases from Gram-positive bacteria.

Yanase *et al.* [33] found that the H^{296} , the proposed proton abstractor in the transferase active site is crucial for catalysis of the transfructosylation reaction in the levansucrase of *Z. mobilis* (Table 2).

Del Rio *et al.* [26] studied the transferase active site of an inulosucrase from *L. citreum* by SDM experiments. They found that R^{618} is an invariant amino acid in all inulosucrases and is placed at -3 from the proton abstractor K^{621} and hence, suggested it could play an important role in fructose donor selection. The transfructosylation activity and product selectivity factor found in the $R^{618}\rightarrow K$ mutant was only 50% with an order of magnitude decrease, respectively.

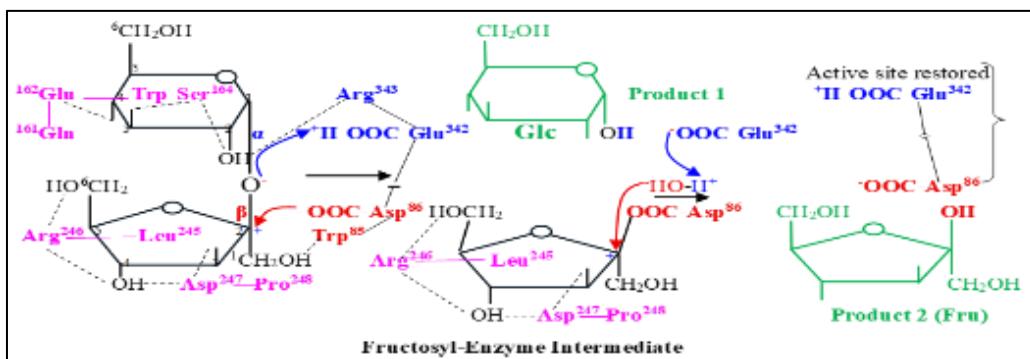
The above SDM data and the MSA analysis confirm that the highly conserved K and the R (at -3) play important roles in the transferase reactions in levan- and inulosucrases, where they act as proton abstractor and donor selection, respectively.

5. Mechanism of Hydrolytic and Transferase Reactions in Fructosyl Transferases

5.1. Mechanism of action of the hydrolytic reaction

It is interesting to note that both the hydrolytic and transferase modules in all three groups of fructosyl transferases invariably use catalytic triads for hydrolysis and transferase reactions. Furthermore, the sucrose-binding sites and the hydrolytic and transferase reactions are virtually identical in all three groups of enzymes, confirming that they use a fully conserved structural framework for the sucrose-binding, cleavage and transferase reactions. In bacteria, the fructan synthesis is somewhat simpler as compared to plant systems, because only one enzyme with the bifunctional property (exhibiting both hydrolytic and transferase activities) is involved. By SDM and mutational analyses, Chembert and Petit-Glatron [40] found that the hydrolytic and transferase activities could be separated in levansucrases, suggesting that they are two independent activities.

The hydrolytic domain is composed of a catalytic triad: a nucleophile (Asp), a general acid/base catalyst (Glu), and a transition-state stabilizer (Asp). The hydrolytic reaction follows a double-displacement mechanism involving two steps, (i.e.) i) protonation of the glycosidic oxygen and breaking the glycosidic linkage, and ii) a nucleophilic attack on the anomeric carbon of the donor resulting in the formation of the fructosyl-enzyme intermediate. In the latter step, a water, or a sucrose molecule or a fructan serves as an acceptor resulting in the hydrolysis or transfer of the fructose to the acceptor molecule, where the acceptor is elongated by one fructosyl unit, respectively. The reaction occurs with an overall retention of the anomeric configuration of the fructosyl residue.



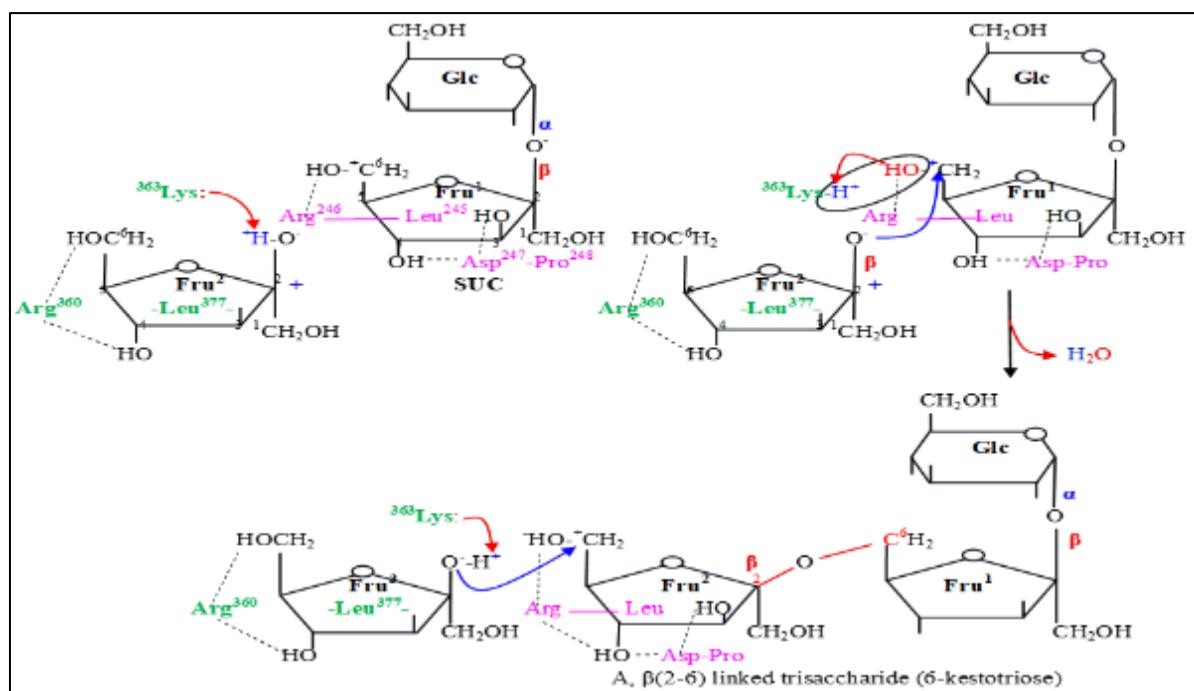
The amino acids in the catalytic triad involved in the hydrolysis of the sucrose molecule are; Asp^{96} , Glu^{342} and Asp^{247} (they act as nucleophile, proton donor and transition state stabilizer, respectively) [18]. The nucleophilic amino acid pair, 'Trp-Asp' and the catalytic acid-base amino acid pair 'Glu-Arg', are conserved in all levansucrases. The highly conserved -E¹⁶²WSGS- and -L²⁴⁵RDPH- motifs are involved in sucrose-binding and shown in magenta.

Figure 4 A schematic diagramme showing the hydrolytic active site amino acids and its proposed mechanism of action (numberings from the *B. subtilis* levansucrase, SacB)

During the first step of the protonation reaction, the acid/base catalyst, Glu, donates a H^+ to the glycosidic bond resulting in its cleavage, followed by a nucleophilic-electrophilic attack on the anomeric carbon of the fructose unit resulting in the formation of a fructosyl-enzyme intermediate, where the C-2 of fructose forms a covalent bond with the nucleophilic Asp residue [14, 13, 18, 25]. The fructosyl-enzyme intermediate formation is supported by kinetic analysis under steady-state conditions and was also successfully isolated and studied by Chambert and Treboul [41], where they found that the β -carboxyl group of the Asp formed a covalent bond with the fructose (Fig. 4). The second Asp acts as a transition-state stabilizer in both steps, and it is based on the observation that it forms strong hydrogen bonds with the C3 and C4 hydroxyls of the fructosyl unit [34] (Fig. 4).

5.2. Mechanism of action of the fructosyl transferase reaction

A second, separate module of the same enzyme accomplishes the fructosyl transferase reaction [40]. The synthesis of fructans starts with a transfructosylation reaction in which the hydrolyzed sucrose molecules in the first step provide the fructosyl donors in which a molecule of sucrose acts as the initial acceptor of the fructosyl units. The transfer of a fructose unit to the acceptor sucrose molecule can occur via O1 to form 1-kestotriose, (the basic unit for the formation of β -(2 \rightarrow 1)-linked inulin) or via O6 to form 6-kestotriose, (the basic unit for the formation of β -(2 \rightarrow 6)-linked levan (Fig. 5). The transferase module is also composed of a catalytic triad where a basic amino acid acts as a proton abstractor, a hydrophobic/aromatic amino acid (found downstream to the catalytic proton abstractor) forms the template-binding pair and again a basic amino acid (found upstream to the catalytic proton abstractor) involves in the donor sugar selection. The proposed fructosyl transferase reaction is very similar to the well-established nucleotidyl transferase reaction [1] as discussed elsewhere. In *B. subtilis* SacB enzyme, the catalytic K abstracts a proton from the donor fructose in the first step, leading to a nucleophilic-electrophilic attack between the incoming fructose unit and the acceptor sucrose molecule and thus, establishing a glycosidic bond between the donor and the acceptor in the second step. By the same mechanism, the fructosyl units are added sequentially to the growing chain forming a levan polymer. The initial linkage on the kestotriose, either β -(2 \rightarrow 6) or β -(2 \rightarrow 1) decides the formation of a levan or an inulin polymer, essentially using the same mechanism, but accomplished by two different enzymes, viz. the levansucrases or the inulosucrases, respectively (Fig. 5). For sucrose at concentrations of >100 mM, fructosyl transferases exhibit almost an exclusive transfructosylation activity [27].



The hydrolytic domain's Fru-binding active site amino acids are in magenta and the transferase domain's active site amino acids are in green. The active site amino acids in the transferase active site domain are: Arg³⁶⁰, Lys³⁶³ and -Leu³⁷⁷GY- (for incoming fructose- selection, proton abstraction and template-binding, respectively).

Figure 5 A schematic diagramme showing the transferase active site and its proposed mechanism action (numbering from the *B. subtilis* levansucrase, SacB)

Like the nucleotidyl transferases (RNA/DNA polymerases), the fructosyl transferases are also multifunctional enzymes and possess two active site domains, one for sucrose hydrolysis (Fig. 1) and the other one for fructosyl transfer, i.e., in the first step, the sucrose molecules are hydrolyzed and in the second step a fructose unit is added to an acceptor sucrose molecule to form a trisaccharide, either 6- or 1-kestotriose (depending on the enzyme type) which is followed by sequential additions of fructose units resulting in levan or inulin polymer, respectively (Fig. 2) [37].

6. Structure and Mechanism of Action of Nucleotidyl Transferases

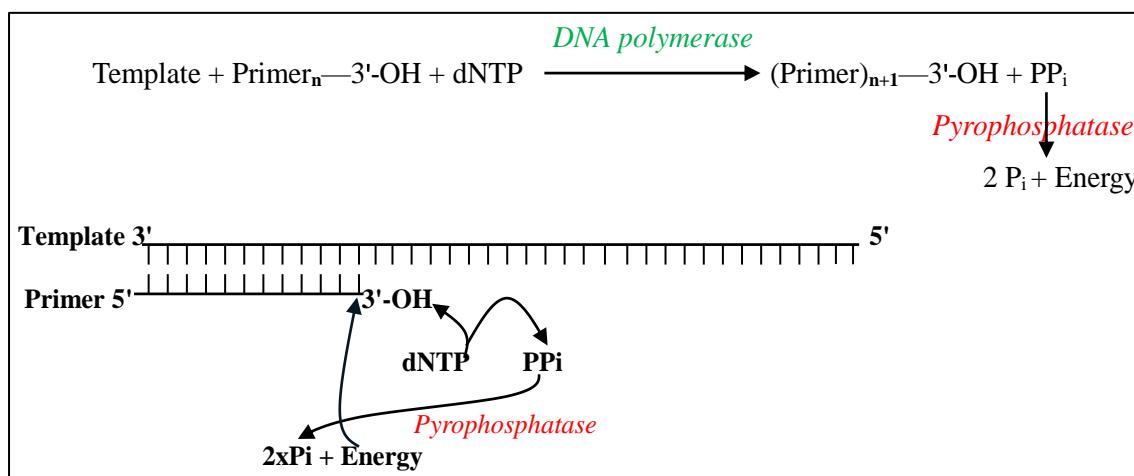
The nucleotidyl transferases also belong to the Main class, 'Transferases', like the fructosyl transferases, but these enzymes involve in the sequential addition of nucleoside triphosphates (NTPs) or deoxynucleoside triphosphates (dNTPs) to synthesize an RNA or a DNA molecule, respectively. DNA polymerases make sequential addition of dNTPs to the 3-OH of a primer on a template, whereas the RNA polymerases make sequential addition of NTPs on a template. Thus, both the polymerases belong to the subclass, nucleotidyl transferases. The properties and mechanisms of action of these crucial catalysts of life are already analyzed in detail and reported by this author [1-3].

DNA polymerases catalyze the polymerization of a new strand of DNA on a given template DNA, using a primer molecule. Except for viruses, there are always more than one enzyme in a living cell. For example, at least three different polymerases have been identified even in the simplest prokaryotes. There are at least five different DNA polymerases reported from *E. coli* and they are named DNA polymerases I, II, III, IV and V. The DNA polymerases, which undertake replication, are known as replicases. The *E. coli* DNA polymerase I is used as an example to explain the mechanism of action here. It was the first enzyme to be discovered and is encoded by *polA* gene. It is a single-subunit enzyme with a molecular mass of 10^3 kDa (928 amino acids). The main functions of this polymerase are: i) removal of the RNA primer after genome replication and subsequent filling of the gap arising due to such removal, ii) DNA repair and iii) removal of TT dimers in UV irradiated cells. Thus, it is a multifunctional enzyme and exhibits three different activities, viz. 5'→3' DNA polymerase activity, 3'→5' exonuclease activity (proofreading function) and 5'→3' exonuclease activity (DNA repair function). There are ~ 400 copies of the enzyme per cell.

6.1. Mechanism of Nucleotidyl Transferases

6.1.1. Dynamics of DNA polymerization

DNA polymerases synthesize a new strand of DNA on a given template using a primer and the four dNTPs (Fig. 6). They add nucleotides one at a time to the 3'-OH end of the primer as shown and thus, the new strand grows from 5'→3'. The overall reaction catalyzed by DNA polymerases can be written as,

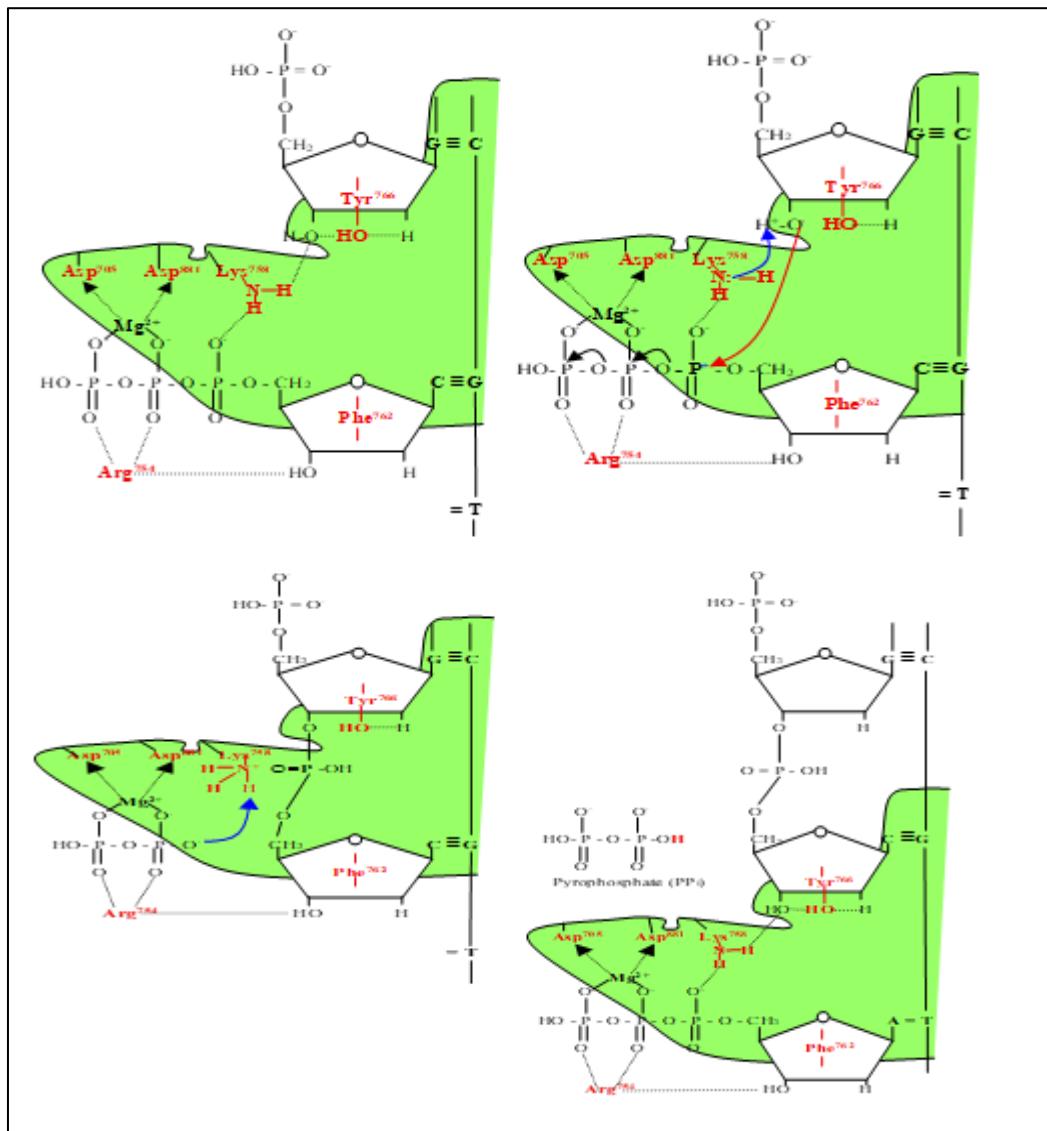


NB: PPi, Inorganic pyrophosphate; Pi, Inorganic phosphate

The high-energy inorganic pyrophosphate is hydrolyzed immediately by the enzyme pyrophosphatase into two molecules of inorganic phosphate resulting in release of energy, which is utilized for the polymerization reaction. Such coupling of reactions drives the polymerization process only in the forward direction, i.e., the polymerized nucleotides cannot be depolymerized.

Figure 6 A schematic diagramme showing the overall reaction of DNA polymerases

As discussed elsewhere, the DNA and RNA synthesis in cells are accomplished by the nucleotidyl transferases where they use either dNTPs or NTPs, respectively. A great deal of information is available on these enzymes from a large number of sources. X-ray crystallographic structures of a large number of polymerases from both prokaryotes and eukaryotes are available. Although they exhibit significant differences in their primary structures (probably due to divergent evolution), their crystal structures are remarkably the same in overall shape. For example, in general, all DNA polymerases, irrespective of their source, have a common protein fold that resembles the shape of a half-opened "right hand" and the polymerase domain exhibits three distinct subdomains, "thumb", "palm" and "fingers" with analogous functions in all organisms [42]. The mechanism of action of *E. coli* DNA polymerase I is shown in Fig. 7 [1]



Watson-Crick base pairing of incoming dNTP to the template at the active site; Electronic transitions at the active site for proton abstraction and an electrophilic-nucleophilic attack; Proton abstraction by the active site amino acid Lys followed by an electrophilic-nucleophilic attack resulting in the formation of 3'→5' phosphodiester bond with the concomitant release of pyrophosphate; Transfer of the proton abstracted by the Lys to the pyrophosphate resulting in the formation of inorganic pyrophosphate. The high-energy bond in the inorganic pyrophosphate is hydrolyzed immediately by the enzyme pyrophosphatase (usually associated with enzyme) and the energy released is utilized by the DNA polymerase to move (translocate) to the next nucleotide, already positioned on the template. As the PPi is removed constantly from the reaction, the polymerization reaction is always unidirectional.

Figure 7 A schematic diagramme showing the steps involved in the polymerization of nucleotidyl transferases (numberings from the *E. coli* DNA polymerase I)

7. Comparative analysis of the active site amino acid domains from the two different transferases

The substrates for the two different transferases, viz. the ribose and fructose have an important similarity in structure, in that, both have a five-membered furanose ring system, suggesting the same active site amino acids could be used for polymerization reactions. This is exemplified in the following two tables. Table 1 shows the highly conserved catalytic amino acids from a large number of DNA/RNA polymerases like, single subunit (SSU) and multisubunit (MSU) DNA polymerases, RNA polymerases, RNA transcriptases from viruses to humans. Irrespective of the source and type of the polymerase, these enzymes use a catalytic triad which consists of a proton abstractor (a basic amino acid, K/R), a template-binding pair located downstream to the proton abstractor (usually an -YG- pair) and a nucleotide selection amino acid (usually a basic amino acid, R, or in some cases a Q or N is also found) at about -4 from the proton abstractor.

Table 1 Conservation of catalytic domain(s) in viral, prokaryotic and eukaryotic RNA/DNA polymerases (Proposed and Confirmed)

Polymerase Type	Polymerase Catalytic Region
1. BACTERIOPHAGES	
SSU RNA Polymerases	
SSU DNA-dep RNAPs	
T7 Viral SSU RNA pol	-620WLA ^Y ⁸ GVTR ^R ⁴ SVT ^K ⁶³¹ RSVMTLA ^Y ⁹ GS-
SP6 Viral SSU RNA pol	-612WDSI ^Y ⁸ GTR ^R ⁴ SLT ^K ¹ KPVMTLP ^Y ⁹ GS-
T3 Viral SSU RNA pol	-WLA ^Y ⁸ GVTR ^R ⁴ SVT ^K ¹ RSVMTLA ^Y ⁹ GS ⁶⁴² -
BPK11 Viral SSU RNA pol	-WLQ ^Y ⁸ GVTR ^R ⁴ KVT ^K ¹ RSVMTLA ^Y ⁹ GS ⁶³⁴ -
2. HUMAN VIRUSES	
(+) Strand RNA Viruses (SARS-CoV)	
RNA-dep RNAPs (NSP12)	
SARS-CoV-1	-STMTN ^R ⁵ ^Q ⁴ YHQ ^K ¹ LKSIAATRGATVIGTSKF ^Y ²² ^G ⁵⁹⁷ -
SARS-CoV-2	-STMTN ^R ⁵ ^Q ⁴ YHQ ^K ¹ LKSIAATRGATVIGTSKF ^Y ²² ^G ⁵⁹⁷ -
MERS-CoV	-STMTN ^R ⁵ ^Q ⁴ YHQ ^K ¹ M LKSMMAATRGATCVIGTTKF ^Y ²² ^G ⁵⁹⁸ -
(-) Strand, Segmented RNA Viruses	
PB1 Catalytic Subunits of the RNA-dep RNAPs	
Human influenza A virus	-477MSKK ^K ³ SYIN ^R ¹ TGTFEFTSFFYR ^Y ¹⁴ ^G ^{FV} -
Human influenza B virus	-860RTKR ^N ⁴ RSIL ^N ¹ DQRNMILEEQCY ^Y ¹⁴ ^A ^{KC} -
Human influenza C virus	-464IRR ^N ⁴ AVCK ^K ¹ IGINMSLEKSY ^Y ¹³ ^{GSL} -
(-) Strand, Non-segmented RNA Virus	
RNA-dep RNAPs	
Human Respiratory Syncytial Virus	-291CL ^N ⁴ TLN ^K ¹ S LGLRCGFNNVLTQLFLY ¹⁸ ^{GD} - (catalytic core 1) -640QV ^Q ⁵ ILAEC ^K ¹ MIAENILQFFPESLTRY ^Y ¹⁷ ^{GD} - (catalytic core 2)
3. PROKARYOTIC AND EUKARYOTIC MSU DNA-dependent RNAPs (mRNAs)	
a) MSU RNAP Family (Rpb2: Initiation subunits)	
<i>E. coli</i> MSU RNAP β subunit	-539TR ^R ⁴ AGFEVR ^R ¹ DVHPTHY ^Y ⁸ GRV ⁵⁵⁸ -
<i>S. cerevisiae</i> MSU RNAP II Rpb2 subunit	-851F ^R ⁵ SLF ^R ¹ S YMDQEKKY ¹⁰ ^{GMS} ⁸⁸⁹ -
<i>S. pombe</i> MSU RNAP II Rpb2 subunit	-840F ^R ⁵ SIFYR ^R ¹ T YTDQEKKY ¹⁰ ^{GMS} ⁸⁵⁸ -
<i>A. thaliana</i> MSU RNAP II Rpb2 subunit	-819FRSYR ^R ⁵ DEEK ^K ¹ MGTLVKEDF ¹⁰ ^{GRP} ⁸⁴⁰ -
Human MSU RNAP II Rpb2 subunit	-906F ^R ⁵ SVFYR ^R ¹ S YKEQESK ^K ¹⁰ ^{GFDQ} ⁸²⁵ -
b) MSU RNAP Family (Rpb1: Elongation subunits)	
<i>E. coli</i> MSU RNAP β' subunit	-833NS ^V ⁴ DAVKV ^R ¹ S VVSC ⁶ ^D TD ⁷ FGVC ¹³ ^{AH} C ¹⁶ ^Y ¹⁷ ^G RD ^L ⁸⁰¹ -
<i>S. cerevisiae</i> MSU RNAPII Rpb1 subunit	-55DPR ^R ⁶ LGSI ^D R ⁹ N ¹ LKC ^C ⁵ QTC ⁸ QEGMNEC ¹⁵ PGH ^F ¹⁸ ^G HI ⁸⁴ -
<i>P. pastoris</i> MSU RNAPII Rpb1 subunit	-55DPR ^R ⁶ LGSI ^D R ⁹ N ¹ F ^K ^C ⁵ QTC ⁸ QEGMNEC ¹⁵ PGH ^F ¹⁸ ^G HM ⁸⁴ -
<i>A. thaliana</i> MSU RNAPII Rpb1 subunit	-54DTR ^R ⁷ LGTIDR ^K ¹ V ^C ⁴ ETC ⁷ MANMAEC ¹⁴ PGH ^F ¹⁸ ^G YL ⁸³ -
Human MSU RNAPII Rpb1 subunit	-59DPR ^R ⁶ QGVIER ^R ¹ T GRC ^C ⁵ QTC ⁸ AGNMTEC ¹⁵ PGH ^F ¹⁸ ^G HI ⁸⁸ -
4. ORGANELLAR SSU RNAPs (NEPs)	
Mitochondrial RNA pol (<i>S. cerevisiae</i>)	-TR ^R ⁴ KVV ^K ¹ Q TVMTNV ^Y ⁹ ^G V ¹⁰²⁴ -
Mitochondrial RNA pol (<i>A. thaliana</i>)	-QVDR ^R ⁴ KLV ^K ¹ Q TVMTSV ^Y ⁹ ^G V ⁷⁶² -
Mitochondrial RNA pol (<i>Nicotiana sylvestris</i>)	-QVDR ^R ⁴ KLV ^K ¹ Q TVMTSV ^Y ⁹ ^G V ⁷⁸⁸ -
Mitochondrial RNA pol (<i>H. sapiens</i>)	-TR ^R ⁴ KVV ^K ¹ Q TVMTVV ^Y ⁹ ^G V ¹⁰⁰¹ -
Chloroplast RNA pol (<i>Oryza sativa</i> , Japonica)	-QVD ^R ⁴ KLV ^K ¹ Q TVMTSV ^Y ⁹ ^G V ⁷³⁹ -
Chloroplast RNA pol (<i>A. thaliana</i>)	-QVD ^R ⁴ KLV ^K ¹ Q TVMTSV ^Y ⁹ ^G V ⁷⁷⁹ -

1. HUMAN and ANIMAL DNA VIRUSES	
DNA-dependent RNAPs	
Smallpox RNA pol	-SR ⁴ YPD R ¹ D SMVCHRILTY ¹² G K ³⁸⁴ -
Monkeypox virus RNA pol	-SR ⁴ YPD R ¹ D SMVCHRILTY ¹² G K ³⁸⁴ -
Cowpox virus RNA pol	-SR ⁴ YPD R ¹ D SMVCHRILTY ¹² G K ³⁸⁴ -
6. VIRAL DNA-dependent DNA Polymerases	
DNA pol (T7 phage)	-509NQIAAEPLT R ¹ DNAK ²² TFIY GFLY ¹ GAG
Human influenza A virus	-256ETLAR S ⁴ ICE K ¹ L EQSGLP V ⁹ G GN- #gRNA
Human influenza B virus	-256ENLAK N ⁴ ICE N ¹ L EQSGLP V ⁹ G GN- #gRNA
Human influenza C virus	-258ETVAQ K ⁴ ICE K ¹ L KESGLP V ⁹ G GN- #gRNA
Smallpox DNA pol	TEKAIYDSM Q ⁴ TYT K ⁶⁰ IVANSV Y ⁹ GLM-
Vaccinia DNA pol	-648TEKAIYDSM Q ⁴ TYT K ⁶⁶ IVANSV Y ⁹ GLM-
Mpox DNA pol	-648TEKAIYDSM Q ⁴ TYT K ⁶⁶ IVANSV Y ⁹ GLM-
7. PROKARYOTIC DNA-dependent DNA Polymerases	
DNA pol I (<i>E. coli</i>)	-745PLETVTSEQ R ⁴ RSAK ⁷⁵⁸ A IN F ⁷⁶² GLI V ⁸ GMS-
Taq DNA pol I (<i>T. aquaticus</i>)	-650PREAVDPLM R ⁴ RAA K ⁶⁶³ T INF GVL Y ⁹ GM-
Pfu DNA pol (<i>P. furiosus</i>)	-478ILLDY R ⁴ QKA K ⁴⁸⁹ L LANSF Y ⁹ G YYGYAK-
DNA pol II (<i>E. coli</i>)@	-480AKRQGNKPLS S ⁴ QALK ⁴⁹³ I IMNAF Y ⁹ G VL-
DNA pol III MSU (Replicase, <i>E. coli</i>)	-661ISYPDVQWQ H ⁴ ESL K ⁶⁷⁴ P VLEPT Y ⁹ GI-
8. EUKARYOTIC DNA-dependent DNA Polymerases	
Yeast α DNA pol (<i>S. cerevisiae</i>)	-931HKRVQCDIR Q ⁴ QAL K ⁹⁴⁴ L TANS M Y ⁸ G ⁹⁵² CL-
Yeast α DNA pol (<i>A. thaliana</i>)	-980LKYWELDIR Q ⁴ QAL K ⁹⁹³ L TANS M Y ⁸ G CL-
Animal α DNA pol (<i>H. sapiens</i>)	-937DLILQYDIR Q ⁴ KAL K ⁹⁵⁰ L TANS M Y ⁸ G CL-
Yeast ϵ DNA pol (<i>S. cerevisiae</i>)	-787MIVLYDSL Q ⁴ LAH K ⁸⁰⁹ V ILNSF Y ⁸ GYV-
Plant ϵ DNA pol (<i>A. thaliana</i>)	-770MVVYDSL Q ⁴ LAH K ⁷⁸² C ILNSF Y ⁸ GYV-
Animal ϵ DNA pol (<i>H. sapiens</i>)	-812MEVLYDSL Q ⁴ LAH K ⁸²⁴ C ILNSF Y ⁸ GYV-
Yeast δ DNA pol (<i>S. cerevisiae</i>)	-688FK RDV LNG R ⁴ AL K ⁷⁰¹ SANSV Y ⁸ GFT-
Plant δ DNA pol (<i>A. thaliana</i>)	-679LE KAV LDG R ⁴ AL K ⁶⁹² SANSV Y ⁸ GFT-
Animal δ DNA pol (<i>H. sapiens</i>)	-681LRR QV LDG R ⁴ AL K ⁶⁹⁴ V SANSV Y ⁸ GFT-
9. ORGANELLAR DNA-dependent DNA Polymerases	
<i>S. cerevisiae</i> Mitochondrial DNA pol y	-741LGCS R ⁴ NEA K ⁷⁴⁹ I FNYGRI Y ⁹ G AG-
<i>A. thaliana</i> Mitochondrial DNA pol y (pol 1)	-872E R ⁴ RKA K ⁸⁷⁷ M LNF SIA Y ⁹ G K-
<i>A. thaliana</i> Mitochondrial DNA pol y (pol 2)	-874E R ⁴ RKA K ⁸⁷⁹ M LNF SIA Y ⁹ G K-
Human Mitochondrial DNA pol y	-917TTVGIS R ⁴ EHA K ⁹⁴⁷ I FNYGRI Y ⁹ G AG-
<i>A. thaliana</i> Chloroplast DNA pol IA	-873E R ⁴ RKA K ⁸⁷⁸ M LNF SIA Y ⁹ G K-
<i>A. thaliana</i> Chloroplast DNA pol IB	-857E R ⁴ RKA K ⁸⁶² M LNF SIA Y ⁹ G K-

Adapted from Palanivelu [1-3, 43, 44]

NEPs, Nuclear Encoded Pols; #gRNA. Genomic RNA (replicase type); Repl, Replicase. Pfu, *Pyrococcus furiosus*.

The confirmed active site amino acids (by SDM, X-ray crystallography and active site-directed inhibitors) are highlighted in dark blue. In *E. coli* pol I, the replacement of K⁷⁵⁸ with Ala caused a 1,000-fold reduction in k_{cat} . The Ala substitutions result in moderate to severe effects on the polymerase activity of the individual mutant enzymes. Severe loss of activity is associated with R⁷⁵⁴→A, K⁷⁵⁸→A, F⁷⁶²→A, and Y⁷⁶⁶→A (highlighted in dark blue) in the *E. coli* DNA pol I. Tyr⁷⁶⁶ and Phe⁷⁷¹ are either involved in template-primer binding or are in the vicinity of the DNA-binding track. Residues Arg⁷⁵⁴, Lys⁷⁵⁸, Phe⁷⁶², and Tyr⁷⁶⁶ appear to be required for the binding of Mg.dTTP, while only Arg⁷⁵⁴ and Lys⁷⁵⁸ are utilized in the polymerization of Mn.dTTP [45 and references therein].

In *Taq* pol I, the invariants R⁶⁵⁹ and K⁶⁶³ in the active site were found to be immutable [45 references therein].

Table 2 shows the highly conserved catalytic amino acids in the hydrolytic and transferase active site domains of fructosyl transferases, viz. levansucrases and inulosucrases. It is interesting to note, that the fructosyl transferases also follow the same transferase active site amino acid pattern like the nucleotidyl transferases and their active site is also composed of a similar catalytic triad, where a basic amino acid (K/R/H) acts as the proton abstractor, a -L/I/V/YG- could act as the template-binding pair which is placed downstream to the proton abstractor and a donor (Fru) selection amino acid usually an R which is placed at -3 from the proton abstractor. It is interesting to note that the *G. stearothermophilus* enzyme is reported to be inactive where the regular basic amino acid K is replaced by a Q.

Table 2 Proposed amino acids in the active site of hydrolytic and transferase domains of levan- and inulosucrases from different organisms.

Invertase module's active site		Transferase module's active site			Fru selection/Proton abstractor/Template-binding pair
Fru- transferase type		Nucleophile	GBM	Proton Donor	FBM
1. Levansucrases					
Bacteria					
<i>Bacillus subtilis</i> (P05655)	-WDSW-	-QWSG-	-EIER-	-LRDP-	-S ¹ R ² GSK ³ M ⁴ TIDGITSNDIYM ⁵ L ⁶ G ⁷ YV-
<i>Geobacillus stearothermophilus</i> (P94468)	-WDSW-	-QWSG-	-EIER-	-LRDP-	-S ¹ R ² GSK ³ M ⁴ TIDGITSNDIYM ⁵ L ⁶ G ⁷ YV-
<i>Sphaerotilus salinarum</i> (A0A428MZE4)	-WDTW-	-QWSG-	-EIER-	-FRDP-	-A ¹ R ² GSK ³ M ⁴ TIDGIDDEDIYM ⁵ L ⁶ G ⁷ YV-
<i>Bacillus megaterium</i> (A0A3DBX658)	-WDSW-	-QWSG-	-EIER-	-LRDP-	-S ¹ R ² GSK ³ M ⁴ TIDGIGQDDVYM ⁵ L ⁶ G ⁷ YV-
<i>Paenibacillus graminis</i> (A0A089MCV7)	-WDSW-	-QWSG-	-EIER-	-LRDP-	-S ¹ R ² GSK ³ M ⁴ TIDGIGANDVYM ⁵ L ⁶ G ⁷ YV-
<i>Clostridium saccharobutylicum</i> (A0A1S8MP69)	-WDSW-	-QWSG-	-EIER-	-LRDP-	-S ¹ R ² GSK ³ M ⁴ TIDGISDKDIYM ⁵ L ⁶ G ⁷ YV-
<i>Clostridium pasteurianum</i> (R4K861)	-WDSW-	-QWSG-	-EIER-	-LRDP-	-S ¹ R ² GSK ³ M ⁴ TIDGIGNKDVYM ⁵ L ⁶ GFT-
<i>Gluconacetobacter diazotrophicus</i>	-WDTW-	-EWSG-	-QTER-	-FRDP-	-F ¹ T ² I ³ H ⁴ R ⁵ TTFAAGVDGPDGV ⁶ Y ⁷ G ⁸ FV-
<i>Beijerinckia indica</i>	-WDTW-	-EWSG-	-QTER-	-FRDP-	-F ¹ T ² I ³ H ⁴ R ⁵ TTYAAGVDGPDGV ⁶ Y ⁷ G ⁸ FV-
<i>Erwinia amylovora</i>	-WDTM-	-EWAG-	-QTER-	-FRDP-	-F ¹ T ² I ³ H ⁴ K ⁵ YTFAADNL ⁶ TGPDGV ⁷ Y ⁸ G ⁹ FV-
<i>Erwinia tasmaniensis</i>	-WDTM-	-EWAG-	-QTER-	-FRDP-	-F ¹ T ² I ³ H ⁴ K ⁵ YTFAADNL ⁶ TGPDGV ⁷ Y ⁸ G ⁹ FV-
<i>Zymomonas mobilis</i>	-WDTW-	-EWAG-	-QTER-	-FRDP-	-F ¹ T ² I ³ H ⁴ H ⁵ STYADGLSGPDGV ⁶ Y ⁷ G ⁸ FV-
2. Inulosucrases					
a) Eubacterial Inulosucrases					
<i>Lactobacillus delbrueckii</i> (EOD02545.1)	-WDSW-	-QWSG-	-EVER-	-LRDA-	-G ¹ R ² L ³ T ⁴ A ⁵ TDTDLLAKANSTIGDNVAM ⁶ L ⁷ G ⁸ FV-
<i>Streptococcus salivarius</i> (VTY24076.1)	-WDSW-	-QWSG-	-EVER-	-LRDP-	-S ¹ R ² I ³ K ⁴ S ⁵ TDAEGTVAVRGAVGDDVVM ⁶ L ⁷ G ⁸ FV-
<i>Streptococcus mutans</i> (VTY49811.1)	-WDSW-	-QWSG-	-ELER-	-MRDP-	-S ¹ R ² L ³ N ⁴ H ⁵ G ⁶ SNNDAWN ⁷ KANEVV ⁸ GDNVVM ⁹ L ¹⁰ G ¹¹ YV-
<i>Convivina intestine</i> (CAH1856476.1)	-WDSW-	-QWSG-	-EIER-	-LRDS-	-T ¹ R ² L ³ N ⁴ R ⁵ G ⁶ SNDYAWQKANR ⁷ VIGDNVAM ⁸ L ⁹ G ¹⁰ YV-
<i>Limosilactobacillus reuteri</i> (Q8GP32)	-WDSW-	-QWSG-	-EIER-	-MRDA-	-T ¹ R ² L ³ N ⁴ R ⁵ G ⁶ SNDDAWMN ⁷ N ⁸ YAV ⁹ GDNVAM ¹⁰ V ¹¹ G ¹² YV-
<i>Streptococcus thermophilus</i> (ACZ67286.1)	-WDSW-	-QWSG-	-EIER-	-MRDA-	-T ¹ R ² L ³ N ⁴ R ⁵ G ⁶ SNDDAWMA ⁷ ANKAV ⁸ GDNVAM ⁹ V ¹⁰ G ¹¹ YV-
b) HaloArchaea Inulosucrases					
<i>Halofera gibbonsii</i>	-WDTW-	-QWAGS-	--ELER-	-FRDP-	-F ¹ V ² S ³ H ⁴ I ⁵ VHTFAPGLTGYDAL ⁶ Y ⁷ G ⁸ FV-
<i>Natronococcus amylolyticus</i>	-WDTW-	-QWAGS-	--ELER-	-FRDP-	-F ¹ L ² S ³ H ⁴ I ⁵ EHTFAEGL ⁶ GYDAL ⁷ Y ⁸ G ⁹ FV-
<i>Halalalkalicoccus jeotgali</i>	-WDTW-	-QWAGS-	--ELER-	-FRDP-	-F ¹ V ² S ³ H ⁴ I ⁵ DHTFAPGL ⁶ GYDAL ⁷ Y ⁸ G ⁹ FV-
<i>Halomicromium</i> sp.	-WDTW-	-TWAGS-	--ELER-	-FRDP-	-F ¹ V ² S ³ H ⁴ I ⁵ DHTFAPGL ⁶ GYDAL ⁷ Y ⁸ G ⁹ FV-
<i>Haloarcula</i> sp	-WDTW-	-QWAGS-	--ELER-	-FRDP-	-F ¹ V ² S ³ H ⁴ I ⁵ LHTFAPGL ⁶ GYDAL ⁷ Y ⁸ G ⁹ FV-
<i>Halorubrum saccharovorum</i>	-WDTW-	-QWAGS-	--ELER-	-FRDP-	-F ¹ V ² C ³ S ⁴ H ⁵ I ⁶ VHTFAPGV ⁷ TGPDGL ⁸ Y ⁹ G ¹⁰ FV-

NB: GBM, Glucose-binding motif; FBM, Fructose-binding motif.

8. Conclusion

The data presented here reveal that the enzyme active sites are designed not at random, but based on the chemical basis of a particular reaction. Therefore, it is imperative that each enzyme active site is designed based on the molecular interactions between specific amino acids at its active site and its corresponding substrate. The present data also confirm that not only the same active site amino acids are employed for the same type of reactions, but are also conserved and adapted from viruses, bacteria, plants to humans during the long evolutionary process. It is likely that the transferase domain of fructosyl transferases harbouring a similar active site amino acid structure as the nucleotidyl transferases might have been acquired from them during the evolutionary process. After tight binding of an enzyme to its substrate, the movements of the subatomic particles, viz. the proton and electron flow, could take place spontaneously between the active site and substrate molecule, effecting transformation in the substrate. A deeper understanding of the evolution of active sites of enzymes could play an important role in enzyme engineering to design novel enzymes for degrading recalcitrant substrates and also in drug development.

Compliance with ethical standards

Acknowledgments

The author wishes to thank Dr. P. Sathiyamoorthy, Chief Scientist, GEM Research Foundation, Chennai, for the suggestions on the manuscript.

Disclosure of conflict of interest

The author declares no conflicts of interest.

References

- [1] Palanivelu P. DNA Polymerases – An insight into their active sites and mechanism of action, In: Recent Advances in Biological Research, Vol 1, Chapter 2, pp 1-39, SCIENCE DOMAIN International Book Publishers, UK, ISBN: 9788193422441, DOI: 10.9734/bpi/rabr/v1; 2019
- [2] Palanivelu P. Eukaryotic Multi-subunit DNA dependent RNA Polymerases: An Insight into their Active Sites and Catalytic Mechanism. In: Advances and Trends in Biotechnology and Genetics., Vol. 1. Chapter 1, pp 1-38, SCIENCE DOMAIN International Book Publishers, UK, Print ISBN: 978-93-89246-59-9, DOI:10.9734/bpi/atbg/v1;2020.
- [3] Palanivelu P. Single Subunit RNA Polymerases: An Insight into their Active Sites and Catalytic Mechanism, In: Advances and Trends in Biotechnology and Genetics. Chapter 1, pp 1-38, SCIENCE DOMAIN International Book Publishers, UK, ISBN: 978-93-89246-59-9, DOI:10.9734/bpi/atbg/v1; 2019.
- [4] Urey, HC. On the early chemical history of the earth and the Origin of Life. *Proc Natl Acad Sci (USA)*. 1952; 38:351.
- [5] Miller SL. A production of amino acids under possible primitive Earth conditions. *Science (New Series)*. 1953; 117: 528-529.
- [6] Miller SL, Urey HC. (1959). "Organic compound synthesis on the primitive Earth." *Science*. 1959; 130: 245-251.
- [7] Criado-Reyes J, Bizzarri BM, García-Ruiz JM, Saladino R, Mauro ED. The role of borosilicate glass in Miller-Urey experiment. *Sci Rep*. 2021; 11:21009. doi: 10.1038/s41598-021-00235-4.
- [8] Ferris JP, Hill Jr AR, Liu R, Orgel LE. Synthesis of long prebiotic oligomers on mineral surfaces. *Nature*. 1996; 381: 59-61
- [9] Mohajer MA, Basuri P, David G, Zindel D, Miliordos E, Signorell R. Spontaneous formation of urea from carbon dioxide and ammonia in aqueous droplets. *Science*. 2025; 388,1426-1430.
- [10] Service RF. Here's how the first proteins might have assembled, sparking life. *Science*. 2025; doi: 10.1126/science.zlr9qda.
- [11] Lahav N, White D, Chang S. Peptide Formation in the Prebiotic Era: Thermal condensation of glycine in fluctuating clay environments. *Science*. 1978; 201:67-69.
- [12] Choukade R, Kango N. 2021. Production, properties, and applications of fructosyltransferase: a current appraisal. *Crit Rev Biotechnol*. 41: 1178-1193.
- [13] Straathoff AJJ, Kieboom APG, van Bekkum H. Invertase-catalysed fructosyl transfer in water-organic solvent systems. *Carbohydr Res*. 1986; 146:154-159.
- [14] Hidaka H, Hirayama M, Sumi NA. Fructooligosaccharide-producing enzyme from *Aspergillus niger* ATCC 20611. *Agric Biol Chem*. 1988; 52:1181-1187.
- [15] Vijn I, Smeekens S. Fructan: more than a reserve carbohydrate? *Plant Physiol*. 1999; 120:351-360.
- [16] Srikanth R, Reddy CH, Siddartha G, Ramaiah MJ, Uppuluri KB. Review on production, characterization and applications of microbial levan. *Carbohydr Polym*. 2015; 120:102-114.
- [17] Maleux K, Van den Ende W. Levans in excised leaves of *Dactylis glomerata*: effects of temperature, light, sugars, and senescence. *J Plant Biol*. 2007; 50: 671-680.
- [18] Meng G, Fütterer K. Structural framework of fructosyl transfer in *Bacillus subtilis* levansucrase. *Nat Struct Biol*. 2003; 10:935-941. doi: 10.1038/nsb974

[19] Chi ZM, Zhang T, Cao T-S, Liu X-Y, Cui W, Zhao C-H. 2011. Biotechnological potential of inulin for bioprocesses. *Bioresour Technol.* 102: 4295-4303.

[20] Martinez-Fleites C, Ortiz-Lombardia M, Pons T, Tarbouriech N, Taylor EJ, Arrieta JG, Hernandez L, Davies G. Crystal structure of levansucrase from the Gram-negative bacterium *Gluconacetobacter diazotrophicus*. *Biochem J.* 2005; 390:19-27.

[21] Yun J. W. Fructooligosaccharides: occurrence, preparation and application. *Enz Microb Technol.* 1996; 19:107-117.

[22] Kaur N, Gupka A. K. Applications of oligofructose in health and nutrition. *J Biosci.* 2002; 27:703-714.

[23] Livingston DP, Hincha DK, Heyer AG. 2009. Fructan and its relationship to abiotic stress tolerance in plants. *Cell Mol Life Sci.* 66: 2007-2023.

[24] Combie J, Öner ET. From healing wounds to resorbable electronics, Levan can fill bioadhesive roles in scores of markets. *Bioinspir. Biomim.* 2018; 14. doi: 10.1088/1748-3190/aaed92.

[25] Palanivelu, P. Fructofuranosyl hydrolases and transferases: Active sites and mechanism of action. *Indian J Microbiol.* 1998; 38: 177-185

[26] Del Río IM, Andrade MG, Portillo FVL, Carranza CO. Exploring the role of the residues into catalytic cavity of inulosucrase from *Leuconostoc citreum* CW28. *Int J Biol Macromol.* 2024; 279: 135159.

[27] Chuankhayan P, Hsieh CY, Huang YC, Hsieh YY, Guan HH, Hsieh YC, Tien YC, et al. Crystal structures of *Aspergillus japonicus* fructosyltransferase complex with donor/acceptor substrates reveal complete subsites in the active site for catalysis. *J Biol Chem.* 2010; 285: 23251-23264.

[28] Soto M, Rivera M, Rudiño-Piñera E, Olvera C, López-Munguía A. Selected mutations in *Bacillus subtilis* levansucrase semi-conserved regions affecting its biochemical properties. *Protein Eng Des Sel.* 2008; 21:589-595.

[29] Ozimek LK, Kralj S, van der Maarel MJEC, Dijkhuizen L. The levansucrase and inulosucrases enzymes of *Lactobacillus reuteri* 121 catalyse processive and non-processive transglycosylation reactions. *Microbiology.* 2006; 152:1187-1196.

[30] Ozimek LK, Kralj S, Kaper T, van der Maarel MJEC, Dijkhuizen L. Single amino acid residue changes in subsite -1 of inulosucrase from *Lactobacillus reuteri* 121 strongly influence the size of products synthesized. *FEBS J.* 2006; 273:4104-4113.

[31] Marsh PD, Bradshaw DJ. 1995. Dental plaque as a biofilm. *J Ind Microbiol Biotechnol.* 15: 169-175.

[32] Batista FR, Hernández L, Fernández JR, Arrieta J, Menéndez C, Gómez R, Támbara Y, Pons T. Substitution of Asp-309 by Asn in the Arg-Asp-Pro (RDP) motif of *Acetobacter diazotrophicus* levansucrase affects sucrose hydrolysis, but not enzyme specificity. *Biochem J.* 1999; 337:503-506.

[33] Yanase H, Maeda M, Hagiwara E, Yagi H, Taniguchi K, Okamoto K. Identification of functionally important amino acid residues in *Zymomonas mobilis* levansucrase. *J Biochem.* 2002; 132:565-572.

[34] Ozimek LK; van Hijum S; van Koningsveld GA, van der Maarel MJEC, van Geel-Schutten GH, Dijkhuizen L. Site-directed mutagenesis study of the three catalytic residues of the fructosyltransferases of *Lactobacillus reuteri* 121. *FEBS Letters.* 2004; 560: 131-133.

[35] Homann A, Biedendieck R, Gotze S, Jahn D, Seibel J. Insights into polymer versus oligosaccharide synthesis: mutagenesis and mechanistic studies of a novel levansucrase from *Bacillus megaterium*. *Biochem J.* 2007; 407: 189-198.

[36] Polsinelli I, Caliandro R, Demitri N, Benini S. The structure of sucrose-soaked levansucrase crystals from *Erwinia tasmaniensis* reveals a binding pocket for levanbiose. *Int J Mol Sci.* 2020; 21:83 <https://doi.org/10.3390/ijms21010083>.

[37] Pijning T, Anwar MA, Böger M, Dobruchowska JM, Leemhuis H, Kralj S, Dijkhuizen L, Dijkstra BW. Crystal structure of inulosucrase from *Lactobacillus*: insights into the substrate specificity and product specificity of GH68 fructansucrases. *J Mol Biol.* 2011; 412:80-93.

[38] Ghauri K, Pijning T, Munawar N, Ali H, Ghauri MA, Anwar MA Wallis R. Crystal structure of an inulosucrase from *Halalkalicoccus jeotgali* B3T, a halophilic archaeal strain. *FEBS J.* 2021; 288: 5723-5736.

- [39] Strube CP, Homann A, Gamer M, Jahn D, Seibel J, Heinz DW. Polysaccharide synthesis of the levansucrase SacB from *Bacillus megaterium* is controlled by distinct surface motifs. *J Biol Chem.* 2011; 286: 17593-17600
- [40] Chembert R, Petit-Glatron MF, Polymerase and hydrolase activities of *Bacillus subtilis* levansucrase can be separately modulated by site-directed mutagenesis. *Biochem J.* 1991; 279: 35-41.
- [41] Chambert R, Treboul GG. Levansucrase of *Bacillus subtilis* - Characterization of a stabilized Fructosyl-Enzyme complex and identification of an Aspartyl residue as the binding site of the fructosyl group. *Eur J Biochem.* 1976; 71: 493.
- [42] Steitz TA. DNA polymerases: Structural diversity and common mechanisms. *J Biol Chem.* 1999; 274:17395-17398.
- [43] Palanivelu P. An insight into the active sites of the catalytic basic protein subunit PB1 of the RNA polymerase of human influenza viruses. *World J Adv Res Rev.* 17: 625-565 (2023).
- [44] Palanivelu P. An insight into the active sites of the RNA polymerase and proofreading exonuclease of the human respiratory syncytial virus. *World J Adv Res Rev.* 2023; 18: 842-858.
- [45] Palanivelu P. An insight into the replication mechanism of smallpox and monkeypox viruses. *World J Adv Res Rev.* 2024; 23: 2239-2260.