

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

185,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Site-Directed Mutagenesis as Applied to Biocatalysts

Juanita Yazmin Damián-Almazo and Gloria Saab-Rincón

*Departamento de Ingeniería Celular y Biotecnología,
Instituto de Biotecnología, Universidad Nacional Autónoma de México,
Cuernavaca, Morelos,
México*

1. Introduction

Enzymes are biological catalysts responsible for supporting almost all of the chemical reactions in living organisms. Their *activities*, *specificities* and *selectivities* make them attractive as biocatalysts for a wide variety of industries. Examples are agrochemicals, detergents, starch, textiles, personal care, pulp and paper, food processing, and animal feed. The chemo-, enantio- and regioselectivities of biological catalysts are hallmarks that make them especially attractive for use in the synthesis of fine chemicals and pharmaceutical intermediates. They are a viable alternative to chemical synthesis, which is usually characterized by low yield and the accumulation of undesirable secondary products. The imminent decrease in the use of fossil fuels has turned attention to new enzyme-based developments for the production of biofuels (e.g., biodiesel) that use renewable raw materials (Cherry & Fidantsef, 2003).

Nevertheless, natural enzymes are often not optimal for use in industrial conditions. It is usually necessary to change the conditions of the process or, most commonly, to alter one or more of the properties of the enzyme. Desirable changes in the enzyme are those that affect substrate specificity, expression level, solubility, stability, activity, selectivity, or thermal stability. Other desired effects could include tolerance to organic solvents or to extreme pH values (Hibbert et al., 2005; Turner, 2009).

Protein engineering usually involves the modification of amino acid sequences at the DNA sequence level by means of chemical or genetic techniques. The resultant protein is then tested for novel, optimal or improved physical and/or catalytic properties (Ulmer, 1983). There are two different basic approaches to engineer proteins, although it is common to combine both approaches for better results.

- a. *Rational design*. Mutations are introduced at specific places in the protein-encoding gene. Positions to mutagenize are based on the knowledge of possible relationships of sequence, structure, function and/or the catalytic mechanism of the protein. Recently computational predictive algorithms have been developed and used to preselect promising target sites (Bolon & Mayo, 2001; Kaplan & DeGrado, 2004; Kuhlman et al., 2003; Meiler & Baker, 2006; Pavelka et al., 2009; Zanghellini et al., 2006). However, a

deep knowledge of the structure and energy functions is required in order to predict the changes required to modify some parameter of the enzyme. This is especially true if one wishes to change the reaction mechanism.

- b. *Directed evolution.* This approach involves repeated cycles of random mutagenesis of and/or recombination with variants of the gene to create a library of genes with slightly different sequences. The enzyme variants thus obtained are submitted to genetic selection or to high-throughput screening to identify those enzyme variants with improvements in the desired property (Stemmer, 1994a; b; Zhao et al., 1998). Directed evolution has been demonstrated to be a very powerful technique, especially for increasing stability (Giver et al., 1998; Ladenstein & Antranikian, 1998; Song & Rhee, 2000; Uchiyama et al., 2000; Zhao & Arnold, 1999) or to change the specificity of an enzyme (Castle et al., 2004; Cohen et al., 2004; Christians et al., 1999; Joerger et al., 2003; Jurgens et al., 2000; Levy & Ellington, 2001; Matsumura & Ellington, 2001; Sakamoto et al., 2001; Song et al., 2002; Zhang et al., 1997). It is a particularly useful approach, since no structural or mechanistic information is required. In many cases, changes that contribute to the improved properties are far from the active sites. They would not have been targeted by a rational strategy.

Both techniques have strengths, but they also have some limitations. For this reason, it is often common to find rational design work combined with directed evolution. It may be desirable to tune some properties of a designed protein (Savile et al., 2010; Siegel et al., 2010) or randomization may be directed to specific regions of a protein that were identified in the design process.

This review summarizes the most common strategies used to identify possible targets for site-directed mutagenesis to enhance biocatalysis. We included sequence- and structure-based strategies for generating enzymes with desired properties. To illustrate a number of the points discussed above, special attention was paid to the site-directed mutagenesis of glycosyl hydrolases. We also used the modification of alpha amylases as a case study. We have described the sequence-based mutagenesis approach that was used to change the transglycosylation/hydrolysis ratio of alpha amylases. Residues involved in the hydrophobicity and electrostatic environment of the active site were identified by sequence and structural alignments with other glycosyltransferases. As a result, certain residues were targeted for mutagenesis. We also used a multiple sequence alignment and structural information in an approach to reduce the hydrolytic activity of the alpha amylase from *Thermotoga maritima*, while increasing its alcoholytic activity. Unlike the wild-type parent, the modified enzyme was able to synthesize alkyl-glucosides.

2. Approaches for selecting targets to mutagenize

Any biochemical, structural, or protein sequence information may be useful for identifying residues that may influence a desired enzyme property. The information may indicate changes that increase or decrease the overall fitness of the enzyme.

A common approach is to focus on regions or positions that may be directly related to the catalytic property. For example, amino acid residues that alter substrate specificity or selectivity are commonly non-conserved residues. They are often in close contact with catalytic residues in or near the active site, cofactors or substrates (Morley & Kazlauskas,

2005; Paramesvaran et al., 2009; Park et al., 2005). Another approach is the identification of sequence motifs that are thought to have been conserved during evolution (Saravanan et al., 2008). In contrast, residues thought to be involved in thermostabilization are spread throughout the entire sequence. Each such residue is thought to make a small contribution to thermostability. However, the additive effect can be significant. For this reason, random mutagenesis is a powerful tool for achieving protein stabilization. However, some features that are known to contribute to protein stability can be implemented by site-directed mutagenesis strategies. These include the introduction of additional disulfide bridges (Mansfeld et al., 1997); decrease of loop entropy by replacement of some amino acid residues to P or by the shortening of loops (Nagi & Regan, 1997); change of α -helix propensity by mutations to replace G residues (low α -helix propensity) to A residues (high α -helix propensity); or by the introduction of salt bridges to increase electrostatic interactions in the protein (Kumar et al., 2000; Lehmann & Wyss, 2001; Spector et al., 2000).

2.1 Sequence-based mutagenesis

2.1.1 Alanine scanning

Alanine scanning is a method used to determine the contribution of the side-chains of specific residues in a protein. Substitution of residues with alanine removes all side chain atoms past the β -carbon, without introducing additional conformational changes into the protein backbone. Although mutagenesis by alanine scanning can be a laborious method (because each alanine-mutated protein must be constructed, expressed and analyzed separately), it has nevertheless been useful for the study of interactions at protein-protein interfaces or for the identification of residues involved in substrate recognition (Gibbs & Zoller, 1991), protein stability (Blaber et al., 1995), or binding (Ashkenazi et al., 1990; Cunningham & Wells, 1989).

Alternatives to conventional alanine scanning are computational methods for modeling alanine-scanning mutants. This approach has proven to be useful for predicting active-site residues important for activity (Funke et al., 2005) and to identify amino acid residues important in protein-protein interactions (Kortemme et al., 2004).

2.1.2 Protein sequence alignment

Nature has had the opportunity to explore the protein sequence space through millions of years of evolution. Genetic drift is thought to be the driving force that is responsible for the sequence diversity observed today. However, residues that are indispensable to function and/or stability have been maintained by selective pressure. Multiple sequence alignments are useful tools for identifying positions that are unchangeable in a protein. They will also identify those regions with the plasticity to allow multiple changes. Briefly, when residues with a common evolutionary origin or having structural or functional equivalence are arranged so that the highly conserved residues are aligned, their alignment serves as an anchor for the alignment of the sequences in a set. Analysis of position-specific residue usage (residue profiles) gives information about amino acid conservation or variability at each position. When a multiple sequence alignment is combined with phylogenetic information, it is possible to explore ancestral relationships among groups of homologous protein sequences. It is also useful for identifying important amino acids that probably cannot be modified.

2.1.2.1 Correlating amino acid sequence patterns to specific properties

An approach for identifying residues that may be functionally relevant is to correlate an enzyme property with the amino acid patterns observed in a multiple sequence alignment. For example, comparison of more stable proteins with less stable ones is a strategy for identifying possible thermostabilizing residues (Ditursi et al., 2006; Gromiha et al., 1999; Kumar & Nussinov, 2001; Perl et al., 2000).

Sequence patterns can also be used to identify the determinants of specificity. Good examples are the attempts to change cofactor specificity in dehydrogenases to NAD⁺, since NAD⁺ is considerably less energetically demanding for the cell to make than is NADP⁺ (Flores & Ellington, 2005; Kristan et al., 2007; Rodríguez-Zavala, 2008; Rosell et al., 2003).

Even distant mutations can significantly affect the properties of an active site. They may alter slightly the geometry, electrostatic properties or dynamics of amino acids in the active site. Distant residues that are important for their interactions with the active site may be seen as conserved in a multiple sequence alignment. Multiple sequence alignments have also revealed that some residues are infrequent in a sequence; but nevertheless are frequently adjacent. Such cluster-forming residues have probably coevolved. These “protein sectors” are often critical for specific functional roles, including substrate binding, stability, allosteric regulation or catalytic activity (Halabi et al., 2009).

2.1.2.2 Consensus sequence

The method of using a consensus sequence is based on the assumption that, in an amino acid sequence alignment of homologous proteins, the consensus amino acid at a given position contributes more to the stability or the function of the protein than does a non-consensus residue. This assumption is based on the belief that a consensus sequence may closely mimic the sequence of an ancestral protein. One hypothesis posits that many proteins were originally thermophilic or hyperthermophilic (Di Giulio, 2003). Under this premise, the consensus sequence has been used to improve the thermostability of several enzymes. This was achieved by mutation of several residues towards the consensus sequence obtained from a multiple sequence alignment. There are numerous examples in which this approach has been used to increase the thermostability of proteins. Some proteins reconstruct the complete consensus sequence (Lehmann et al., 2000; Sullivan et al., 2011). In others, point mutations were used to identify the residues that increased stability. They were then combined to increase the thermostability of the protein (Maxwell & Davidson, 1998; Nikolova et al., 1998; Yamashiro et al., 2010).

Similarly, Jochens et al. (2010) showed the improvement of an enzyme property by mutagenizing the codon for a residue to codons for those amino acids that appear frequently in natural enzymes at identical positions. Evolution probably selected these residues. They are unlikely to perturb the folding or the function of the protein. In contrast, absent and rarely occurring residues are the ones that are probably not allowed. Their rareness suggests that they may be deleterious to the protein. This approach was used to improve the activity and enantioselectivity of an esterase from *Pseudomonas fluorescens* (PFE). The amino acid distribution at four positions near the active site of PFE, previously reported to influence the enantioselectivity of the enzyme, was determined by a structure-guided multiple-sequence alignment of 171 esterases generated by the 3DM database (Kuipers et al., 2010). A library was created by site-directed mutagenesis of the coding regions for the four

active site positions in PFE. Substitutions were limited to frequently occurring residues. Almost all mutants in the library showed significantly improved activity towards a commonly used esterase substrate. Moreover, one mutant had its specific activity enhanced 240-fold relative to that of the wild-type enzyme. The mutant also exhibited substantially higher enantioselectivity in the hydrolysis of 3-phenyl butyric acid p-nitrophenyl ester ($E=80$) compared to that of the almost nonselective wild-type enzyme ($E=3.2$) (Jochens & Bornscheuer, 2010).

2.1.2.3 Design of ancestral proteins

As mentioned above, one hypothesis suggests that ancestral proteins were able to withstand the harsh conditions prevalent on earth at that time (Di Giulio, 2001). In addition to their thermostability, ancestral enzymes may have been promiscuous with respect to substrates. The evolution theory of proteins holds that current proteins evolved from low-specificity ancestral proteins. Because of their low specificity, the ancestral proteins evolved to become more efficient at using specific substrates. Thus, the reconstruction of ancestral sequences from multiple sequence alignments and phylogenetic trees may provide the opportunity to change enzyme specificity. Several methods based on this approach have been reported. Some of them are given below.

In the *Ancestral Library* method, all residues located close to or within the enzyme's active site are mutated to residues predicted by phylogenetic analysis and ancestral inference. The substitutions are those residues found in the hypothetical proteins at various nodes and branches of the evolutionary trajectories of a given enzyme family. They do not reflect the entire diversity seen in existing family members (Alcolombri et al., 2011).

Alcolombri and coworkers (2011) used serum paraoxonases and cytosolic sulfotransferases (SULTs) as models. In order to promote changes in substrate specificity, they constructed ancestral libraries of enzymes. Their mutagenesis was directed to residues near or within the active site of an enzyme. From a phylogenetic tree, the most probable ancestral sequences were obtained for all nodes. Using these sequences as templates and the three-dimensional structure of the enzyme, residues in and near the active site were located. The ancestral residues were identified, and the relevant altered enzymes constituted a library of mutants. After activity screening, several variants with different activities and specificities were identified. Some mutants had up to 50-fold higher activity than the activity of the starting enzyme.

REAP (Reconstructing Evolutionary Adaptive Paths) analysis uses phylogeny to identify mutations in gene sequences that are thought to have emerged from a common universal ancestor during functional divergence. The findings are used to generate focused and functionally enriched enzyme libraries (Chen et al., 2010).

REAP was implemented to identify differences in the sequence of promiscuous viral polymerases and non-viral polymerases. The differences may be responsible for the functional divergence without loss of catalytic activity. Sequence alignments and a phylogenetic tree of 719 polymerases were constructed. Ancestral proteins sequences were inferred from the collection of sequences at the nodes of the tree. *REAP* identified sites that may have changed during the separation of viral and non-viral polymerases. In one example, mutations of the residues identified by *REAP* analysis for the DNA polymerase of *Thermus aquaticus* yielded 8 mutants that showed a change in the substrate specificity for unnatural dNTP's (Chen et al., 2010).

Similarly, the *Evolutionary Trace* method correlates evolutionary variations within a gene of interest with divergence in the phylogenetic tree of that sequence family. This method has been shown to reveal the functional importance of residues (Lichtarge et al., 1996; Lichtarge et al., 2003).

2.1.2.4 SCA (Statistical Coupling Analysis)

Well-separated mutations can significantly affect the activity, specificity, or enantioselectivity of an enzyme by slightly altering the geometry, electrostatic properties or dynamics of amino acids in the active site. Moreover, it is known that physically contiguous residues form “protein sectors” that can be critical for specific functional roles, including substrate binding, stability, allosteric regulation and catalytic activity (Halabi et al., 2009).

Statistical Coupling Analysis (SCA) is a method that estimates the co-evolution between pairs of amino acids in the multiple sequence alignment of a protein family. SCA shows that proteins can be divided into “protein sectors.” In several different proteins, the sectors correspond to amino acids that are physically contiguous. These amino acids often underlie various aspects of function, allosteric regulation, binding, catalytic specificity, and/or fold stability. An application of SCA revealed networks of small subsets of residues that link distant functional sites and cooperate in allosteric communication (Suel et al., 2003).

2.2 Structure-based mutagenesis

The evolution of proteins involves mutations of single residues, insertions, deletions (Pascarella & Argos, 1992), gene duplications, fusions, exon duplications and shuffling (Grishin, 2001). Such changes, which accumulate over time, make the identification of sequence similarities very difficult. However, structure is more preserved than sequence and can be used as an evidence of homology among proteins. Comparative analyses of protein sequences and structures are important approaches for the identification of structural, evolutionary and functional relationships between proteins.

The rapidly growing number of protein structures in the Protein Database (PDB) and advances in homology modeling are of great value for generating structural alignments. In general, these methods provide a measure of structural similarity between proteins. They also generate an alignment that defines the residues that have structurally equivalent positions in the proteins being compared. Homology modeling can be done even when no sequence similarity is detected. Based on structural alignments, it is possible to identify residues in direct contact with the substrate or near the active-site cavity. A more complex analysis can even look into enzyme locations that are far from the active site, but are part of a network of interactions that hold the active site together. The residues at these locations can be targeted for mutagenesis. There is a server called HotSpot Wizard that combines information from extensive sequence and structure database searches with functional data to create a mutability map for a target protein. This approach was validated by comparing “hot spot” predictions with mutations extracted from the literature (Pavelka et al., 2009).

2.2.1 Site-directed Saturation Mutagenesis (SDSM)

The Site-directed Saturation Mutagenesis (SDSM) approach consists of using all 20 amino acids at a position in a protein. Based on structural knowledge, it may be sufficient to target

the active-site residues (Park et al., 2005; Schmitzer et al., 2004; Wilming et al., 2002; Woodyer et al., 2003). SDSM can be used to complement error-prone PCR. One of the limitations of using error-prone PCR to generate variants of a protein is that the sequence exploration is limited to an average of seven amino acid substitutions per residue. Once positions that seem to be important for improving a property of a protein are identified by error-prone PCR, SDSM will tune the optimization by testing all 20 amino acids in those positions. Multiple positions can be mutagenized simultaneously by SDSM if only a few positions are being explored. However, the number of variants increases exponentially with the number of positions being explored. Therefore, if more than 3 positions are being randomized, it is better to carry out successive targeted randomizations at the positions.

2.2.2 Combinatorial Active-site Saturation Test (CAST)

The Combinatorial Active-site Saturation Test (CAST) was developed to increase the enantioselectivity and/or the substrate specificity of enzymes. The basis of the method is the generation of small libraries of mutant enzymes that are easy to screen for activity. The mutants are produced by simultaneous randomization of sets of two or three spatially close amino acids, whose side chains form part of the substrate-binding pocket.

Application of this methodology allowed the expansion of the substrate specificity of *Pseudomonas aeruginosa* lipase (PAL). Based on the crystal structure, pairs of amino acids surrounding the binding pocket were defined; and the corresponding libraries were created separately by simultaneous saturation mutagenesis at each pair. The libraries were screened for activity with different substrates. The best-performing variants were selected (Reetz et al., 2005). Further optimization was achieved by iterative cycles of CASTing (Reetz et al., 2006a). Mutants that enhanced a given catalytic property were selected. The residue positions thought to be responsible were organized into groups of two or three. Each group was randomized by saturation mutagenesis to create libraries that were subsequently screened. The best hit of those libraries was used as the template for the next round of mutagenesis. Variability at the other sites was introduced by another round of saturation mutagenesis. The process was continued until the desired degree of catalyst improvement had been achieved. Iterative screening has been applied to enhance very different catalytic properties, including thermostability (Reetz et al., 2006b), substrate acceptance and enantioselectivity (Clouthier et al., 2006; Reetz et al., 2006a).

2.2.2.1 B-factor iterative test

The B-factor iterative test is used to modify enzyme thermostability by increasing rigidity at sites to help prevent unfolding. The selection of target residues is made on the basis of crystallographic B-factor data. This value reflects the degree to which the measured electron density for a particular atom spreads out. It is strongly influenced by thermal fluctuations and the mobility of the atom. Residues with the highest B factors have high flexibility. Appropriate mutations lead to enhanced rigidity and, therefore, to higher thermostability. Target sites are chosen as sites for iterative saturation mutagenesis (ISM), in which each of the target residues is mutagenized to saturation. The best mutant of the first screening is then used as the template for a second round of saturation mutagenesis at one of the other selected sites. The cycle is repeated in an iterative manner (Reetz et al., 2006a).

This method has been used to enhance the thermostability and the tolerance to organic solvents of the mesophilic lipase (LipA) from *Bacillus subtilis*. After several rounds of ISM at

residues with high B-factors, a mutant was obtained with the inactivation temperature increased from 48°C to 93°C and with an improved robustness towards organic solvents without affecting the activity of the enzyme (Reetz et al., 2006b; Reetz et al., 2010). Similarly, Jochens and coworkers (Jochens et al., 2010) increased the T_m of the esterase from *Pseudomonas fluorescens* by 9 °C. Smart libraries were guided by the B-factor. By guiding ISM at residues displaying the lowest B-factors, Reetz and coworkers (Reetz et al., 2009) were able to create a lipase from *Pseudomonas aeruginosa* that has a decrease in the T_m value from 71.6 °C to 35.6 °C without affecting the catalytic profile of the enzyme.

2.3 Site-directed homologous recombination

Most proteins are only marginally stable. For this reason, the accumulation of a few mutations is sometimes sufficient to destabilize a protein. The introduction of variability by recombination with the structural gene for a homolog is often less perturbing for folding than is mutagenesis to introduce point mutations. The reason is that some amino acid changes introduced by recombination have already been selected by Nature to give a particular structure and function. Because recombination is more conservative than mutagenesis, several research groups have tried to introduce variability in a sequence by constructing chimeras with genes for homologous proteins. This can be done either randomly (Crameri et al., 1998; Minshull & Stemmer, 1999) or in a site-directed fashion (Landwehr et al., 2007; Li et al., 2007; Pantazes et al., 2007). By substituting a homologous segment, some interactions may be perturbed; and the protein might not be functional. The less perturbing sites for chimeragenesis are thus identified, and a library of recombinants is then constructed by recombining DNA fragments from different homologous genes. The resulting library is screened for a specific property. The properties have included thermostability, activity towards different natural and non-natural substrates, and/or specificity. The power of this technique, compared to random recombination strategies, is that the libraries constructed have a high percentage of folded proteins, thus making it easier to find interesting variants.

2.4 Site-directed loop exchange in proteins

With the same idea as the site-directed chimeragenesis described above, site-directed loop exchange is based on introducing variability only in the binding and/or catalytic loops; the rest of the structure is perturbed very little. The basis of this strategy is that a good portion of catalytic and molecular recognition sites are in loops, while the rest of the structure is the scaffold that maintains the residues and the network of interactions in place to create an environment suitable for catalysis. By exchanging loops, the sequence can be different from that of the parental protein; but the stability and the folding of the protein is maintained. The technique has been widely used, especially for antibodies. It has been recognized that the binding specificity of antibodies relies on the loops of the variable regions (Clark et al., 2009). More recently, import of loops from natural enzymes has been carried out to explore novel activities in a given scaffold (Park et al., 2006). In our laboratory, we have developed a strategy that allows systematic loop exchange from eight different proteins with a TIM barrel fold into a TIM barrel scaffold. (A TIM fold is characterized by a barrel formed by eight parallel β -strands surrounded by seven or eight α -helices. The loops that join the β -strands to the α -helices at the top of the barrel conform the active site in proteins sharing

this fold.) We demonstrated that the libraries generated had a high percentage of properly folded proteins (Ochoa-Leyva et al., 2009).

3. Glycosyl hydrolases

Glycosyl hydrolases (also called glycosidases) constitute a widespread group of enzymes that catalyze the hydrolysis of the glycosidic bond. Glycosyl hydrolases can hydrolyze the glycosidic linkage between two or more carbohydrates or between a carbohydrate and a non-carbohydrate moiety to release smaller sugars. They are classified as exo- and endoglycosidases, depending on their ability to cleave a substrate at the end (the non-reducing end) or in the middle of an oligosaccharide or polysaccharide chain, respectively. Glycosyl hydrolases have been classified into more than 100 families based on amino acid sequence similarities (Davies & Henrissat, 1995; Henrissat, 1991; Henrissat & Bairoch, 1993; 1996; Henrissat & Davies, 1997). This classification system, available on the CAZy (Carbohydrate-Active EnZymes) web site (Cantarel et al., 2009), allows reliable prediction of evolutionary relationships, mechanism (retaining/inverting), active site residues and possible substrates. It is even a reasonable tool for newly sequenced enzymes for which function has not yet been biochemically demonstrated.

3.1 Reactions catalyzed by glycosyl hydrolases

In most cases, the hydrolysis of the glycosidic bond is catalyzed by two amino acid residues - a general acid (proton donor) and a nucleophile/base. Depending on the spatial positions of these catalytic residues, hydrolysis occurs via overall retention or overall inversion of the anomeric configuration (Davies & Henrissat, 1995; McCarter & Withers, 1994; Sinnott, 1990).

3.1.1 Inverting glycosyl hydrolases

Inverting enzymes act by a single-step, acid/base-catalyzed mechanism. Two residues, typically glutamic or aspartic acids located 6-11 Å apart, act as acid and base. The leaving group is directly displaced by the nucleophilic water with a single inversion at the anomeric centre (Fig. 1).

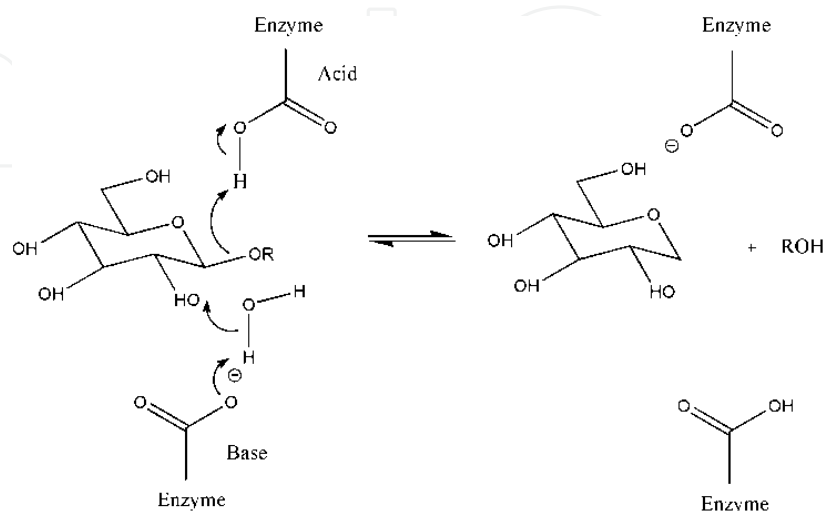


Fig. 1. Inversion hydrolysis mechanism of glycosyl hydrolases.

3.1.2 Retaining glycosyl hydrolases

Retaining glycosidases act through a double-displacement mechanism (each step resulting in inversion at the anomeric centre) involving a covalent glycosyl-enzyme intermediate (Fig. 2). The reaction is catalyzed with acid/base and nucleophilic assistance provided by two amino acid side chains, typically glutamate or aspartate, located 5.5 Å apart. In the first step (glycosylation), one residue plays the role of a nucleophile. It attacks the anomeric centre to displace the aglycon and form a glycosyl enzyme intermediate. At the same time, the other residue functions as an acid catalyst and protonates the glycosidic oxygen as the bond cleaves. In the second step (deglycosylation), the now deprotonated acid-base carboxylate functions as a base to activate the incoming nucleophile (water, saccharide or alcohol) to which the glycosyl is transferred from the enzyme intermediate to give the product.

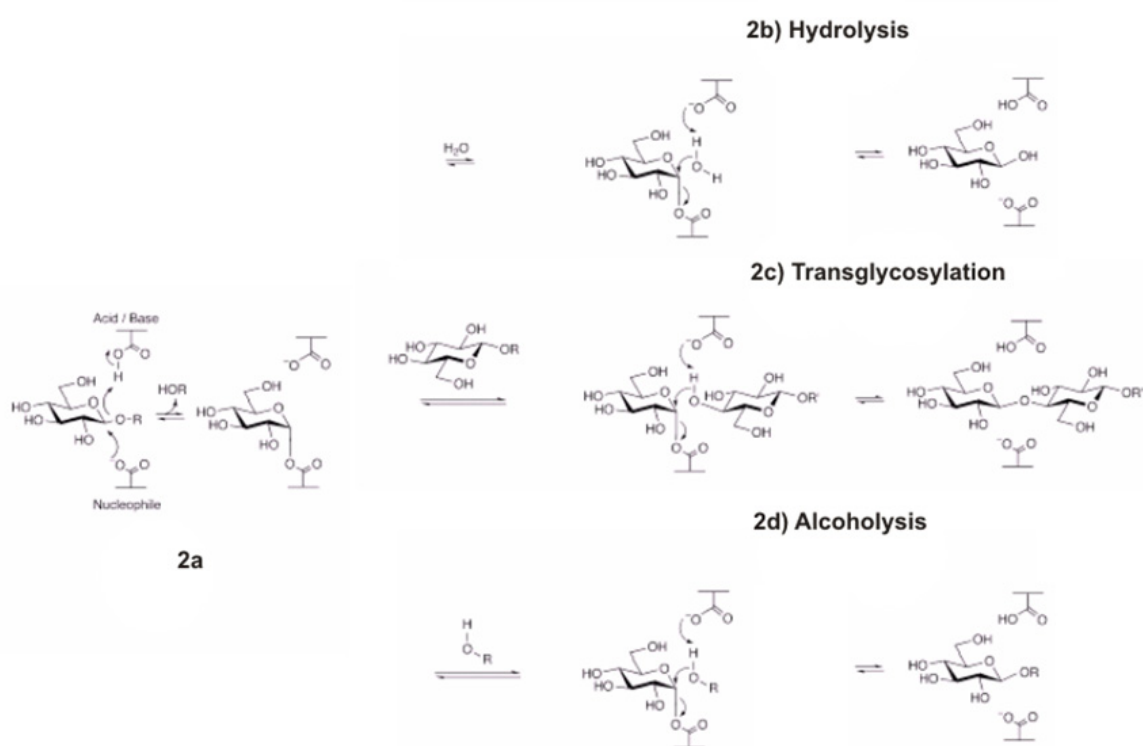


Fig. 2. Retaining mechanism of glycosyl hydrolases.

In a first reaction (a), glycosidic bond breakage of the donor saccharide is carried out and a glycosyl enzyme complex is formed. In (b), the incoming acceptor molecule (water) is activated to promote the release of sugar. If the incoming nucleophile is different from water, the enzyme carries out a transfer reaction, called transglycosylation if the incoming molecule is an oligosaccharide (c), or alcoholysis if the incoming nucleophile is an alcohol (d).

3.2 Industrial uses of glycosyl hydrolases

In nature, glycosyl hydrolases catalyze the degradation of diverse glycosylated polymers, like starch, glucogen, cellulose and hemicellulose. They also participate in anti-bacterial defense strategies (*e.g.*, lysozyme), in pathogenesis mechanisms (*e.g.*, viral neuraminidases)

and in normal cellular functions. Glycosyl hydrolases are also of great importance to industry. For example, in the food industry, enzymes like invertase and amylase are employed for the manufacture of invert sugar or maltodextrins; in the paper and pulp industry, xylanases are used to remove hemicelluloses from paper pulp; cellulases are widely used in the textile industry and in laundry detergents; and recently, cellulases and xylanases have been used in the conversion of lignocellulosic biomass into forms suitable for biofuel production.

However, in most cases, glycosyl hydrolases are not optimal in industrial conditions. It is often necessary to alter their stabilities, catalytic activities and/or substrate specificities by protein engineering methods. One of the most frequently altered enzyme properties is thermostability. It can be a limiting factor in the selection of enzymes for industrial applications due to the elevated temperatures or the extreme pH of many biotechnological processes. The stability of an enzyme can be improved by site-directed mutagenesis (Ben Mabrouk et al., 2011; Ghollasi et al., 2010; Leemhuis et al., 2004; Liu et al., 2008; Yin et al., 2011). One of the most exhaustive efforts was done by Palackal and coworkers (Palackal et al., 2004), who used saturation mutagenesis for each of the 189 amino acid residues of a xylanase. They generated a library of modified enzymes, each altered at single position. This library was then screened for variants with increased thermostability, and nine single amino acid changes that contribute to increased stability were identified. These nine single substitutions were then combinatorially assembled to generate all 512 possible variants. Another round of screening identified eleven enzymes with melting temperatures up to 35°C higher than that of the wild-type enzyme.

Another enzyme property that is desirable to modify is the optimum pH. For example, in soybean β -amylase, the hydrogen bond networks around the catalytic base residue (E380) of the enzyme were removed by point mutations, raising the optimal pH from 5.4 to the more neutral pH range of between 6 and 6.6 (Hirata et al., 2004a; Hirata et al., 2004b).

In vivo, glycosidases catalyze the hydrolysis of glycosidic linkages. However, *in vitro*, they can be used as synthetic catalysts to form glycosidic bonds. This process is called the kinetic approach, and it can be accomplished by reverse hydrolysis or by transglycosylation (Fig. 2 c and d). The utility of glycosidases in the synthesis of glucosides through transglycosylation reactions has been employed to synthesize unusual products that are difficult to obtain by other methods. Several site-directed mutagenesis strategies have been used to increase the transglycosylation activity of glycosidases or to change substrate specificity. For example, rational modification of the β -glycosidase from *Sulfolobus sulfataricus* to accept a wider range of substrates in transglycosylation reactions has been done. Site-directed mutagenesis was used to alter two key residues involved in substrate recognition to provide access to many different glycoside linkages, including the especially problematic β -mannosyl and β -xylosyl linkages (Hancock et al., 2005). We will focus our discussion on the protein engineering work on α -amylases carried out by us and others.

3.3 α -amylases

α -Amylases (EC number 3.2.1.1) are part of the family 13 of glycosyl hydrolases. They catalyze the hydrolysis of internal α -1,4-glycosidic linkages of starch, liberating poly- and oligosaccharides chains of varying lengths. They are found in both eubacteria and

eukaryotes. They have a large number of different substrate specificities, as well as huge variations in both temperature and pH optima (Vihinen & Mantsala, 1989).

α -Amylases are the starting enzymes in the industry of modification and conversion of starch. This is because of their capacity to catalyze reactions under environmentally friendly conditions and without the addition of expensive activated sugars (Buchholz & Seibel, 2008). In the sugar-producing industry (Nielsen & Borchert, 2000), bacterial and fungal α -amylases of family GH13, particularly those of the *Bacillus* species, play a vital role in the starch liquefaction process. Starch from wheat, maize and tapioca is hydrolyzed to produce oligosaccharides by the thermostable α -amylases from *Bacillus licheniformis*. The oligosaccharides are then saccharified to glucose by glucoamylase (Crabb & Shetty, 1999). According to the degree of hydrolysis of starch, α -amylases are divided into two categories: (1) saccharifying α -amylases, which hydrolyze 50 to 60% of the saccharide bonds and (2) liquefying enzymes, which process about 30 to 40% of starch hydrolysis (Fukumoto & Okada, 1963). The enzyme commonly used in the industrial process is the α -amylase from *Bacillus licheniformis*. It has the great advantage of being thermostable. This enzyme thus allows the fast hydrolysis of starch at the high temperatures required to dissolve it, with the consequent decrease in viscosity, before decreasing the temperature for the addition of the next enzyme in the process. Some of the disadvantages of using different enzymes are that the pH or temperature conditions may need to be adjusted during the process, with consequent increase in time, costs, and the introduction of salts (buffers) that will have to be removed from the final product. Thus, several research groups, including ours, have tried to engineer α -amylases to change their product profiles (*i.e.*, to make them more saccharifying) to increase their optimal temperatures and to widen their pH spectra.

All α -amylases consists of three domains, called A, B and C. Domain A contains the catalytic residues and has four conserved sequence regions (numbered I-IV) (Mackay et al., 1985; Nakajima et al., 1986; Rogers, 1985), which have been postulated to be essential for the function of α -amylase. Among α -amylase sequences, the four regions align and are spaced at similar intervals along the proteins. These regions presumably form the active site cleft, the substrate-binding site, and the site for binding the stabilizing calcium ion. Domain B forms a large part of the substrate binding cleft, and it is presumed to be important for the substrate specificity differences observed among α -amylases (MacGregor, 1988). It is the least conserved domain among α -amylases (Guzman-Maldonado & Paredes-Lopez, 1995). Finally, domain C constitutes the C-terminal part of the sequence and seems to be involved in substrate binding. All known α -amylases contain a conserved calcium ion located at the interface between domains A and B (Boel et al., 1990; Kadziola et al., 1998; Machius et al., 1998; Machius et al., 1995). The calcium ion is known to be essential for enzyme stability (Vallee et al., 1959).

Depending on the enzyme, the active site cleft can accommodate from four to ten glucose units, each one bound by amino acid residues that constitute the binding subsite for that glucose unit. Subsites are numbered according to the location of the scissile bond. In α -amylases there are two or three subsites on the reducing end of the scissile bond (subsites +1, +2 and +3). The number of subsites on the non-reducing side of scissile bond varies between two and eleven (subsites -1, -2, ... -11) (Brzozowski et al., 2000; Davies et al., 1997; MacGregor, 1988). The number of subsites and their affinities are some of the determinant factors of the final product profiles of α -amylases (Gyemant et al., 2002; Kandra et al., 2002; Matsui et al., 1992a; b).

3.3.1 Transglycosylation reactions in α -amylases

As other retaining glycosidases, α -amylases, particularly saccharifying amylases, can also catalyze transfer reactions, which are the result of employing molecules other than water (e.g., carbohydrates or alcohols) as glucosyl acceptors (Fig. 2 c and b, respectively). When a high molecular-weight alcohol is used as an acceptor, the products are alkyl-glucosides. These molecules have a high surface tension activity that has important applications in several industries. Although various retaining glycosidases, like β -galactosidase (Moreno-Beltran et al., 1999; Svensson, 1994), β -xylosidase (Shinoyama et al., 1988), β -fructofuranosidase (Rodríguez et al., 1996; Straathof et al., 1988), and β -glucosidase (Chahid et al., 1992; Vulfson et al., 1990), have been used in alcoholysis reactions, the use of a readily available substrate, like starch, gives α -amylases great potential in the catalysis of this type of reaction.

We found a correlation between the efficiency of hydrolysis and the capacity of the enzymes to carry out transglycosylation reactions. A plausible hypothesis is that those enzymes that are able to transglycosylate can recycle intermediate size oligosaccharides produced during hydrolysis to generate longer ones that are better substrates. This would result in a more saccharifying pattern at equilibrium. Transglycosylation activity is not reported in the bacillary α -amylases used in the starch process industry. We decided to introduce this activity by engineering liquefying α -amylases from *Bacillus stearothermophilus* (Saab-Rincon et al., 1999) and *Bacillus licheniformis* (Rivera et al., 2003). We tried to identify residues that could be responsible for transferase activity. Kuriki and coworkers (Kuriki et al., 1996) suggested three residues that are likely to be responsible for controlling the water activity in the active site of the neopullulanase, a natural transferase from *Bacillus stearothermophilus*. When one of these residues (Y377) was mutated to a non-polar residue, the transglycosylation reaction was favored due to a change in the transglycosylation/hydrolysis ratio. We carried out a multiple sequence alignment of α -amylases and cyclodextrin glycosyltransferases (CGTases) and identified a residue (A289 in the *Bacillus stearothermophilus* α -amylase) that is analogous to Y377 in the *Bacillus stearothermophilus* neopullulanase. The *Bacillus stearothermophilus* α -amylase is a liquefying enzyme unable to carry out transglycosylation reactions (Fig. 3). We used site-directed mutagenesis to change the A at residue 289 to Y and F, which are present in natural transferases, like neopullulanases and CGTases. The two mutants that were generated were able to carry out the transfer reaction not only to other saccharides (Fig. 4) but also to alcohols, like methanol, to produce methyl-glucosides. The A289Y mutant was more efficient at catalyzing transfer reactions than was A289F (Fig. 5) (Saab-Rincon et al., 1999). Apparently the hydrophobic nature of the mutant residues and the electrostatic interactions that may affect the geometry of the side chains in the active site are important for the transglycosylation reaction. In contrast, when the same mutations were introduced at the equivalent position (V286) in the α -amylase from *Bacillus licheniformis*, the V286Y mutant showed an increase of hydrolytic activity, whereas the V286F mutant had a higher transglycosylation/hydrolysis ratio (Rivera et al., 2003).

In contrast to bacterial liquefying α -amylases from *B. licheniformis* and *B. stearothermophilus*, several fungal amylases like those from *Aspergillus niger* and *Aspergillus oryzae* have the ability to carry out alcoholysis reactions. These two fungi amylases are responsible for saccharifying enzymes that produce maltose, maltotriose and some glucose. Santamaria and

coworkers (Santamaria et al., 1999) demonstrated that these enzymes were able to carry out alcoholysis reactions in the presence of methanol and starch as substrate, even at high methanol (20%) and starch (15%) concentrations. Although the alcoholysis reaction was reported in α -amylase from *A. oryzae* using aryl-maltoside and either, methanol, ethanol or butanol as substrates (Matsubara, 1961), the alcoholysis reaction with starch as substrate is less efficient (Santamaria et al., 1999).

Enzyme	Region I			Region II			Region III			Region IV		
				*			*			*		
<i>T. maritima</i>	KVIMDLVINHT	129	DGFRIDAAKH	223	VGEVSGN	263	FNFALM	280	FLENHDL	311		
<i>A. oryzae</i>	YLMVDVVANHM	123	DGLRIDTVKHV	211	IGEVLGDG	235	LNYPYI	255	FVENHND	298		
<i>A. niger</i>	YLMVDVVPDHM	123	DGLRIDSVLEV	211	VGEIDNGN	235	LNYPYI	255	FIENHND	298		
<i>H. sapiens sal</i>	RIYVDAVINHM	102	AGFRIDASKHM	202	YQEVIDLG	238	TEFKYG	259	FVDNHDN	301		
<i>H. sapiens pan</i>	RIYVDAVINHM	102	AGFRIDASKHM	202	YQEVIDLG	238	TEFKYG	259	FVDNHDN	301		
<i>Pig Panc</i>	RIYVDAVINHM	102	AGFRIDASKHM	202	FQEVIDLG	238	TEFKYG	259	FVDNHDN	301		
<i>Tenebrio</i>	RIYVDAVINHM	100	AGFRVDAAKHM	190	YQEVIDLG	227	LEFQFG	248	FVDNHDN	288	α -Amylases	
<i>Alteromonas</i>	DIYVDTLINHM	90	KGFRFDASKHV	179	FQEVIDQG	205	TEFKYS	226	FVDNHDN	265		
<i>Barley</i>	KAIADIVINHR	93	DGWRFDFAKGY	184	VAEIWTS	209	FDFTTK	249	FVDNHDN	286		
<i>B. lichen</i>	NVYGDVVINHK	106	DGFRIDAVKHI	236	VAEYQND	266	FDVPLH	289	FVDNHDN	329		
<i>B. amylo</i>	QVYGDVVLNHK	104	DGFRIDAAKHI	236	VAEYQNN	266	FDVPLH	289	FVDNHDN	329		
<i>B. stearo</i>	QVYADVDFDHK	107	DGFRIDAVKHI	239	VGEYWSYD	269	FDVPLH	292	FVDNHDN	332		
<i>B. sub</i>	KVIIDAVINHT	103	DGFRFDAAKHI	180	YGEILQDS	213	TASNYG	232	WVESDHT	270		
<i>B. circ1</i>	KVIIDFAPNHT	141	DGIRMDAVKHM	234	FGEWFLGV	262	LDFRFA	286	FIDNHDM	329		
<i>B. circ2</i>	KVIIDFAPNHT	141	DGIRMDAVKHM	234	FGEWFLGS	262	LDFRFN	286	FIDNHDM	329		
<i>B. specie</i>	KVIIDFAPNHT	141	DGIRMDAVKHM	234	FGEWFLGV	262	LDFRFA	286	FIDNHDM	329	CGTases	
<i>B. stearo</i>	KVIIDFAPNHT	137	DGIRMDAVKHM	230	FGEWFLSE	257	LDFRFG	281	FIDNHDM	325		
<i>Termo sulfu</i>	KVIIDFAPNHT	142	DGIRMDAVKHM	235	FGEWFLGT	263	LDFRES	283	FIDNHDM	330		
<i>Thermus sp</i>	RVMLDAVFNHC	248	DGWRLDVANEI	333	LGEIWHDA	362	MNYPLA	380	LLGSHDT	422		
<i>B. lichen</i>	KIMLDVFNHI	252	DGWRLDVANEV	334	LGEIWHQA	363	MNYPFT	381	LLDSHDT	426	Maltogenic.	
<i>B. stearo</i>	RVMLDAVFNHS	248	DGWRLDVANEV	333	LGEIWHDA	361	MNYPFT	380	LLGSHDT	425		

Fig. 3. Multiple sequence alignment around the four characteristic regions observed in members of the glycoside hydrolase family 13.

The catalytic residues, conserved in all the sequences are marked with asterisks. Aromatic residues involved in the transglycosylation activity of the CGTases are indicated in grey. Residues mutagenized are shown in red. Sequences from α -amylases, GCTases and maltogenic amylases are taken from Damian-Almazo et al. (2008). The abbreviated terms follow: *T. maritima*, *Thermotoga maritima* α -amylase; *A. oryzae*, *Aspergillus oryzae* α -amylase; *A. niger*, *Aspergillus niger* α -amylase; *H. sapiens sal*, *Homo sapiens* salivary α -amylase; *H. sapiens pan*, *Homo sapiens* pancreatic α -amylase; *Pig panc*, *Sus scrofa* pancreatic α -amylase; *Tenebrio*, *Tenebrio molitor* (mealworm) α -amylase; *Alteromonas*, *Pseudoalteromonas haloplanktis* α -amylase; *Barley*, *Hordeum vulgare* α -amylase; *B. lichen*, *Bacillus licheniformis* α -amylase; *B. amylo*, *Bacillus amyloliquefaciens* chimera α -amylase; *B. stearo*, *Bacillus stearothermophilus* α -amylase; *B. sub*, *Bacillus subtilis* 2633 α -amylase; *B. circ1*, *Bacillus circulans* 251 cyclodextrin glycosyltransferase (CGTase); *B. circ2*, *Bacillus circulans* 8 CGTase; *B. specie*, *Bacillus* sp. 1011 CGTase; *B. stearo*, *B. stearothermophilus* CGTase; *Termo sulfu*, *Thermoanaerobacter thermosulfurogenes* CGTase; *Thermus sp.*, *Thermus* sp.; *B. lichen*, *B. licheniformis* maltogenic amylase; *B. stearo*, *B. stearothermophilus* maltogenic amylase.

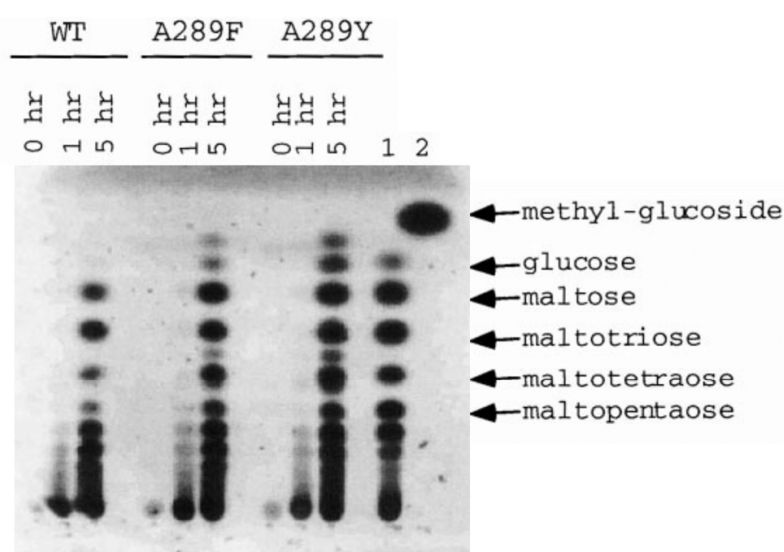


Fig. 4. Product profiles of the alcoholysis reaction with wild-type *Bacillus stearothermophilus* α -amylase and the A289F and A289Y variants at different times of reaction using starch and methanol as substrates.

The product profiles obtained with wild-type (WT) and mutant (A289F and A289Y) *Bacillus stearothermophilus* α -amylases are compared at 0, 1, 5 hours of reaction. We used as standards a mixture of oligosaccharides (1) and methyl-glucoside (2).

Although the wild-type enzyme and the A289F and A289Y mutants showed similar hydrolysis and transglycosylation patterns, the mutants showed products between the glucose and methyl-glucoside standards that could be attributed to alcoholysis reactions. Presumably, those spots for which there are no molecular weight markers correspond to alkyl-oligosaccharides.

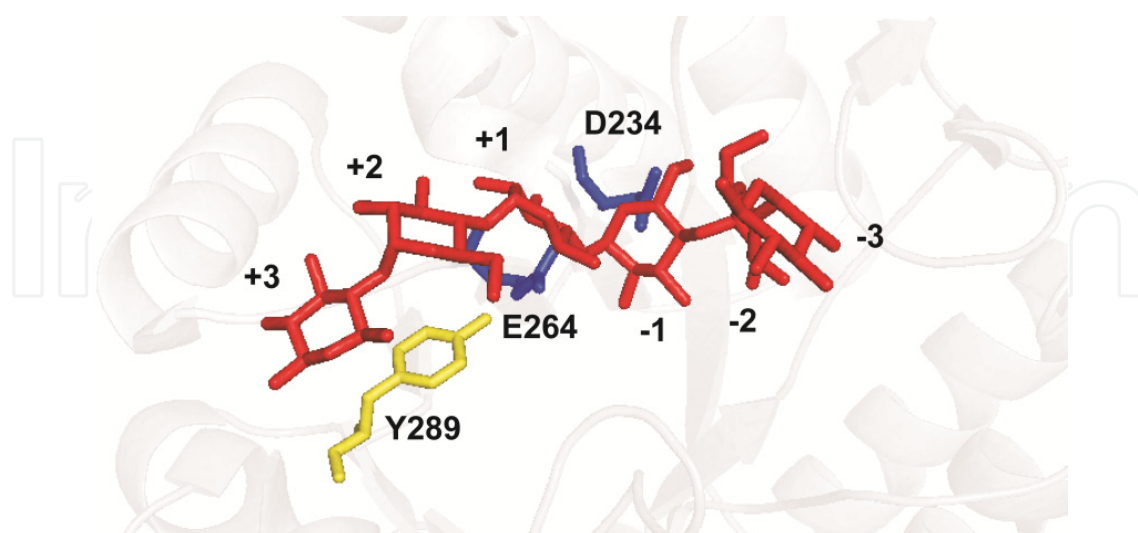


Fig. 5. Structural model of the A289Y mutant of *Bacillus stearothermophilus* α -amylase (PDB code 1HVX) crystallized with the substrate analogue acarbose. (The structure is based on the structure of the α -amylase from *Aspergillus oryzae*, PDB code 7TAA.) Catalytic residues are represented in blue sticks and the mutated residue (Y289) is represented in yellow.

However, the direct use of these enzymes for the production of alkyl-glucosides is precluded by the high temperature required for starch solubilization. The use of a thermophilic saccharifying α -amylase would be attractive, not only in the development of alcoholysis reactions, but also in the starch-processing industry.

Liebl et al. (1997) described an extracellular α -amylase (AmyA) produced by the hyperthermophilic bacterium *Thermotoga maritima* MSB8. The enzyme is a saccharifying amylase with an optimum temperature of 85°C. It can hydrolyze internal α -1-4-glycosidic bonds in various α -glucans, such as starch, amylose, amylopectin and glycogen, to yield mainly glucose and maltose as final products. Because AmyA has the advantage of being a saccharifying enzyme in a stable scaffold, we explored its properties in the transglycosylation and alcoholysis reactions (Damián-Almazo et al., 2008; Damian-Almazo et al., 2008; Moreno et al., 2010). In addition to the characterization reported by Lieb et al., we found that AmyA is capable of using small oligosaccharides (G2 to G7) as substrates for the transglycosylation reactions at short reaction times. This was followed by hydrolysis to yield glucose and maltose as final products. The ability of AmyA to use maltose as a substrate is unusual, as most α -amylases are not capable of using maltose to transfer glucosyl units to other oligosaccharides. Moreover, in the presence of various substrates, AmyA is able to form neotrehalose, a non-reducing disaccharide composed of two glucose molecules joined by α -1, β -1 linkage. It uses 6% maltose as a substrate. Like other saccharifying enzymes, AmyA is capable of transferring glucosyl units to methanol and butanol to produce alkyl-glucosides. When compared to other saccharifying α -amylases, AmyA has a high transfer capacity. The enzyme generates 7.5 mg/ml of methyl-glucoside (Moreno et al., 2010), almost three times the maximum amount found for the *A. niger* α -amylase and almost eight times the maximum amount found for the *A. oryzae* α -amylase (Santamaria et al., 1999).

In order to increase the alcoholytic activity present in AmyA, we constructed a structural homology model based on the structure of the α -amylase from *A. oryzae*. The low sequence identity between these enzymes precluded the use of the automatic modeler function in the Swiss Prot server (Sali et al., 1995; Sanchez & Sali, 1997). Therefore, the sequence alignment of the proteins had to be manually adjusted using as anchors the four highly conserved regions of the α -amylases, as shown in Fig. 3. Once a model was generated, the inhibitor molecule acarbose was placed in the active site using the coordinates of the *A. oryzae* α -amylase (PDB code 7TAA). A close-up of the active site model (Fig. 6) supported our hypothesis of the relationship between the presence of an aromatic residue at the position equivalent to Y377 in neopullulanase and the transglycosylation activity of the enzyme and a saccharifying profile. We identified other residues in subsite +1 that are involved in the transglycosylation activity of other glycosyl hydrolases (Kim et al., 2000; Leemhuis et al., 2004; van der Veen et al., 2001). One of these (H222) is part of the second highly conserved region among glycosyl hydrolases and has also been implicated in calcium ion coordination. In the AmyA model, this residue points toward the sugar moiety at subsite +1. Mutagenesis of the equivalent residue in other amylases has been shown to change transferase activity. In the case of the *B. stearothermophilus* α -amylase, the replacement of the equivalent H238 with aspartic acid generated an enzyme with a reduced hydrolysis rate and a modified final product profile (Vihinen & Mantsala, 1990). We constructed a small library by site-directed mutagenesis to explore the effects of replacing H222 with D, Q and E. All mutants showed a greater amount of methyl-glucoside than did the wild-type enzyme, as a result of a change in the alcoholysis/hydrolysis ratio. Mutant H222Q showed an increase in the alcoholysis

events as a consequence of an increase in alcoholysis and a reduction in hydrolytic activity of almost 30%. The same change was observed in mutants H222D and H222E. The instability of these mutants toward alcohols decreased the final yield of alkyl-glucoside, as shown in Fig. 7 (Damian-Almazo et al., 2008).

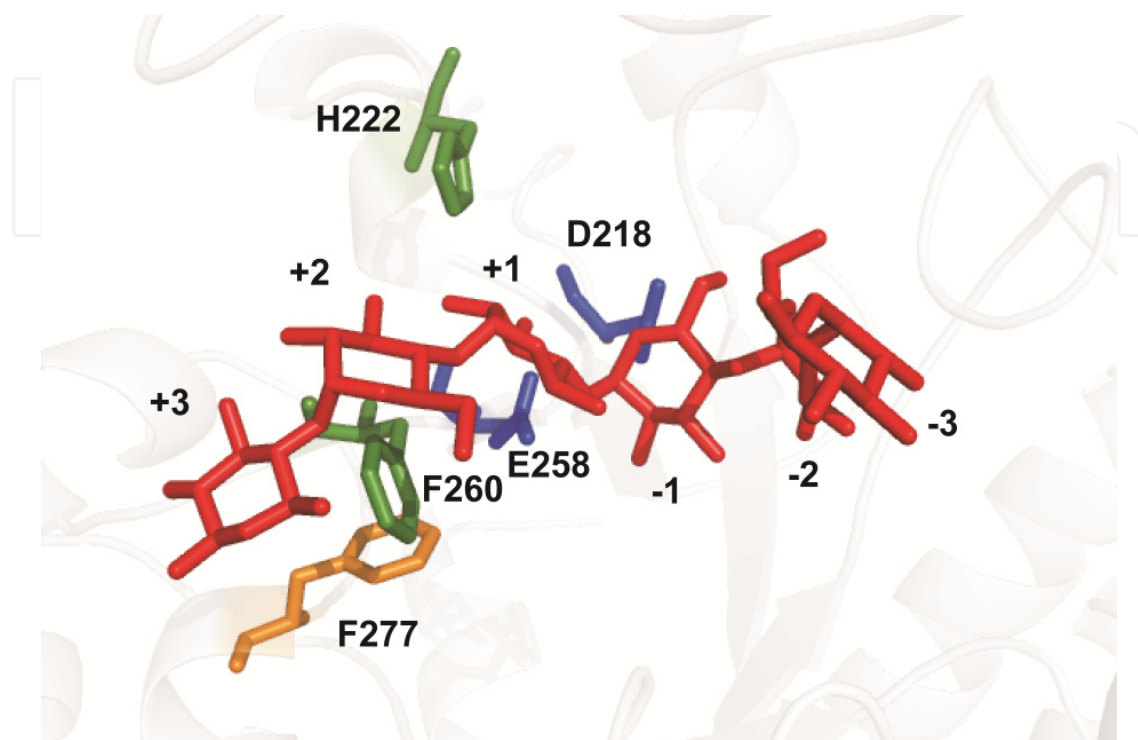


Fig. 6. Homology model obtained for AmyA active site.

The inhibitor acarbose (red) is surrounded by catalytic residues D218 and E258 (blue) and various mutated residues (green). The F277 residue, equivalent to Y377 of neopullulanase, is shown in orange.

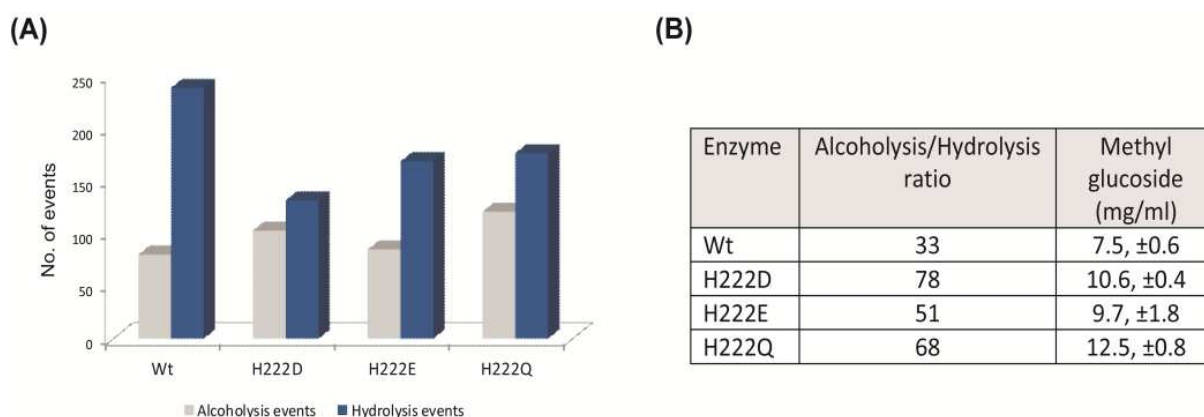


Fig. 7. Alcoholysis reaction yields of the wild-type α -amylase from *Thermotoga maritima* and some of the mutants generated

Quantification of alcoholysis reactions generated by wild-type α -amylase from *Thermotoga maritima* and the H222 residue mutants. (A) Alcoholysis and hydrolysis events from 6% starch - 20% methanol obtained with 20 U/ml of the enzymes shown; (B) alcoholysis/hydrolysis ratios and methyl-glucoside yields.

The comparison of liquefying and saccharifying α -amylases, neopollulanases, CGTases and maltogenic amylases through a multiple sequence alignment (Fig. 3) has also made possible the identification of other residues potentially involved in the transglycosylation activity (Fig. 6). In the CGTase from *Bacillus circulans*, residue F260 has been identified as part of a switch for the transglycosylation and hydrolysis reactions (van der Veen et al., 2001). Mutants formed by changing the equivalent residue in wild-type AmyA (F260) to W and G and the H222Q mutant showed opposite behaviors. In the presence of soluble starch as substrate, mutants H222Q and F260G leave higher amounts of high-molecular weight oligosaccharides, while the wild-type enzyme and mutant F260W show a higher proportion of glucose. These differences were seen as changes in the transglycosylation/hydrolysis ratios. In the double mutant H222Q-F260W, the more transglycosidic pattern of H222Q was recessive, thus eliminating or reducing the presence of longer oligosaccharides (Damián-Almazo et al., 2008).

4. Conclusions

Site-directed mutagenesis is a powerful tool for both the study of protein function and the design of novel proteins. Using several approaches to identify phylogenetically conserved residues or residues involved in binding, it has been possible to modify the properties of enzymes that have industrial and biotechnological applications. In order to increase the transglycosylation reactions carried out by α -amylases, we have applied site-directed mutagenesis to residues close to the active site. Based on multiple sequence alignments of natural transferases, like CGTases, we identified conserved residues involved in the transferase reactions of fungal and bacterial α -amylases. Changes to these residues in α -amylases that originally were unable to perform the transglycosylation reactions altered the product profiles and increased the transglycosylation/hydrolysis ratios. Furthermore, it was possible to increase the alcoholysis reactions in the α -amylase from *Thermotoga maritima*, which was already capable of carrying out this kind of reaction at a low level.

5. Acknowledgements

This work was funded by PAPIIT [Grant number IN206311 to GSR]; and Consejo Nacional de Ciencia y Tecnología [Grant number 154194]

6. References

- Alcolombri, U., Elias, M. & Tawfik, D.S. (2011). Directed evolution of sulfotransferases and paraoxonases by ancestral libraries. *J Mol Biol*, Vol. 411, No. 4, pp. 837-853, 1089-8638 (Electronic) 0022-2836 (Linking).
- Ashkenazi, A., Presta, L.G., Marsters, S.A., Camerato, T.R., Rosenthal, K.A., Fendly, B.M. & Capon, D.J. (1990). Mapping the CD4 binding site for human immunodeficiency virus by alanine-scanning mutagenesis. *Proc Natl Acad Sci U S A*, Vol. 87, No. 18, pp. 7150-7154, 0027-8424 (Print) 0027-8424 (Linking).
- Ben Mabrouk, S., Aghajari, N., Ben Ali, M., Ben Messaoud, E., Juy, M., Haser, R. & Bejar, S. (2011). Enhancement of the thermostability of the maltogenic amylase MAUS149 by Gly312Ala and Lys436Arg substitutions. *Bioresour Technol*, Vol. 102, No. 2, pp. 1740-1746, 1873-2976 (Electronic) 0960-8524 (Linking).

- Blaber, M., Baase, W.A., Gassner, N. & Matthews, B.W. (1995). Alanine scanning mutagenesis of the alpha-helix 115-123 of phage T4 lysozyme: effects on structure, stability and the binding of solvent. *J Mol Biol*, Vol. 246, No. 2, pp. 317-330, 0022-2836 (Print) 0022-2836 (Linking).
- Boel, E., Brady, L., Brzozowski, A.M., Derewenda, Z., Dodson, G.G., Jensen, V.J., Petersen, S.B., Swift, H., Thim, L. & Woldike, H.F. (1990). Calcium binding in alpha-amylases: an X-ray diffraction study at 2.1-A resolution of two enzymes from *Aspergillus*. *Biochemistry*, Vol. 29, No. 26, pp. 6244-6249.
- Bolon, D.N. & Mayo, S.L. (2001). Enzyme-like proteins by computational design. *Proc Natl Acad Sci U S A*, Vol. 98, No. 25, pp. 14274-14279.
- Brzozowski, A.M., Lawson, D.M., Turkenburg, J.P., Bisgaard-Frantzen, H., Svendsen, A., Borchert, T.V., Dauter, Z., Wilson, K.S. & Davies, G.J. (2000). Structural analysis of a chimeric bacterial alpha-amylase. High-resolution analysis of native and ligand complexes. *Biochemistry*, Vol. 39, No. 31, pp. 9099-9107.
- Buchholz, K. & Seibel, J. (2008). Industrial carbohydrate biotransformations. *Carbohydr Res*, Vol. 343, No. 12, pp. 1966-1979, 0008-6215 (Print) 0008-6215 (Linking).
- Cantarel, B.L., Coutinho, P.M., Rancurel, C., Bernard, T., Lombard, V. & Henrissat, B. (2009). The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucleic Acids Res*, Vol. 37, No. Database issue, pp. D233-238, 1362-4962 (Electronic) 0305-1048 (Linking).
- Castle, L.A., Siehl, D.L., Gorton, R., Patten, P.A., Chen, Y.H., Bertain, S., Cho, H.J., Duck, N., Wong, J., Liu, D. & Lassner, M.W. (2004). Discovery and directed evolution of a glyphosate tolerance gene. *Science*, Vol. 304, No. 5674, pp. 1151-1154, 1095-9203 (Electronic) 0036-8075 (Linking).
- Clark, L.A., Boriack-Sjodin, P.A., Day, E., Eldredge, J., Fitch, C., Jarpe, M., Miller, S., Li, Y., Simon, K. & van Vlijmen, H.W. (2009). An antibody loop replacement design feasibility study and a loop-swapped dimer structure. *Protein Eng Des Sel*, Vol. 22, No. 2, pp. 93-101, 1741-0134 (Electronic) 1741-0126 (Linking).
- Clouthier, C.M., Kayser, M.M. & Reetz, M.T. (2006). Designing new Baeyer-Villiger monooxygenases using restricted CASTing. *J Org Chem*, Vol. 71, No. 22, pp. 8431-8437, 0022-3263 (Print) 0022-3263 (Linking).
- Cohen, H.M., Tawfik, D.S. & Griffiths, A.D. (2004). Altering the sequence specificity of HaeIII methyltransferase by directed evolution using in vitro compartmentalization. *Protein Eng Des Sel*, Vol. 17, No. 1, pp. 3-11, 1741-0126 (Print).
- Crabb, W.D. & Shetty, J.K. (1999). Commodity scale production of sugars from starches. *Curr Opin Microbiol*, Vol. 2, No. 3, pp. 252-256, 1369-5274 (Print) 1369-5274 (Linking).
- Crameri, A., Raillard, S.A., Bermudez, E. & Stemmer, W.P. (1998). DNA shuffling of a family of genes from diverse species accelerates directed evolution. *Nature*, Vol. 391, No. 6664, pp. 288-291.
- Cunningham, B.C. & Wells, J.A. (1989). High-resolution epitope mapping of hGH-receptor interactions by alanine-scanning mutagenesis. *Science*, Vol. 244, No. 4908, pp. 1081-1085, 0036-8075 (Print) 0036-8075 (Linking).
- Chahid, Z., Montet, D., Pina, M. & Graille, J. (1992). Effect of water activity on enzymatic synthesis of alkylglycosides. *Biotechnol Lett*, Vol. 14, No., pp. 281-284.

- Chen, F., Gaucher, E.A., Leal, N.A., Hutter, D., Havemann, S.A., Govindarajan, S., Ortlund, E.A. & Benner, S.A. (2010). Reconstructed evolutionary adaptive paths give polymerases accepting reversible terminators for sequencing and SNP detection. *Proc Natl Acad Sci U S A*, Vol. 107, No. 5, pp. 1948-1953, 1091-6490 (Electronic) 0027-8424 (Linking).
- Cherry, J.R. & Fidantsef, A.L. (2003). Directed evolution of industrial enzymes: an update. *Curr Opin Biotechnol*, Vol. 14, No. 4, pp. 438-443, 0958-1669 (Print) 0958-1669 (Linking).
- Christians, F.C., Scapozza, L., Cramer, A., Folkers, G. & Stemmer, W.P. (1999). Directed evolution of thymidine kinase for AZT phosphorylation using DNA family shuffling. *Nat Biotechnol*, Vol. 17, No. 3, pp. 259-264.
- Damián-Almazo, J., López-Munguía, A., Soberón-Mainero, X. & Saab-Rincón, G. (2008). Role of the phenylalanine 260 residue in defining product profile and alcoholic activity of the α -amylase AmyA from *Thermotoga maritima*. *Biologia*, Vol. 63, No. 6, pp. 1035-1043.
- Damián-Almazo, J.Y., Moreno, A., López-Munguía, A., Soberón, X., González-Munoz, F. & Saab-Rincón, G. (2008). Enhancement of the alcoholic activity of α -amylase AmyA from *Thermotoga maritima* MSB8 (DSM 3109) by site-directed mutagenesis. *Appl Environ Microbiol*, Vol. 74, No. 16, pp. 5168-5177, 1098-5336 (Electronic) 0099-2240 (Linking).
- Davies, G. & Henrissat, B. (1995). Structures and mechanisms of glycosyl hydrolases. *Structure*, Vol. 3, No. 9, pp. 853-859, 0969-2126 (Print) 0969-2126 (Linking).
- Davies, G.J., Wilson, K.S. & Henrissat, B. (1997). Nomenclature for sugar-binding subsites in glycosyl hydrolases. *Biochem J*, Vol. 321 (Pt 2), No., pp. 557-559, 0264-6021 (Print) 0264-6021 (Linking).
- Di Giulio, M. (2001). The universal ancestor was a thermophile or a hyperthermophile. *Gene*, Vol. 281, No. 1-2, pp. 11-17.
- Di Giulio, M. (2003). The Universal Ancestor was a Thermophile or a Hyperthermophile: Tests and Further Evidence. *Journal of Theoretical Biology*, Vol. 221, No. 3, pp. 425-436.
- Ditursi, M.K., Kwon, S.J., Reeder, P.J. & Dordick, J.S. (2006). Bioinformatics-driven, rational engineering of protein thermostability. *Protein Eng Des Sel*, Vol. 19, No. 11, pp. 517-524, 1741-0126 (Print) 1741-0126 (Linking).
- Flores, H. & Ellington, A.D. (2005). A modified consensus approach to mutagenesis inverts the cofactor specificity of *Bacillus stearothermophilus* lactate dehydrogenase. *Protein Eng Des Sel*, Vol. 18, No. 8, pp. 369-377, 1741-0126 (Print) 1741-0126 (Linking).
- Fukumoto, J. & Okada, S. (1963). Studies on bacterial amylase. Amylase types of *Bacillus subtilis* species. *J Ferm Technol*, Vol. 41, No., pp. 427.
- Funke, S.A., Otte, N., Eggert, T., Bocla, M., Jaeger, K.E. & Thiel, W. (2005). Combination of computational prescreening and experimental library construction can accelerate enzyme optimization by directed evolution. *Protein Eng Des Sel*, Vol. 18, No. 11, pp. 509-514, 1741-0126 (Print) 1741-0126 (Linking).
- Ghollasi, M., Khajeh, K., Naderi-Manesh, H. & Ghasemi, A. (2010). Engineering of a *Bacillus* α -amylase with improved thermostability and calcium independency. *Appl Biochem Biotechnol*, Vol. 162, No. 2, pp. 444-459, 1559-0291 (Electronic) 0273-2289 (Linking).

- Gibbs, C.S. & Zoller, M.J. (1991). Identification of electrostatic interactions that determine the phosphorylation site specificity of the cAMP-dependent protein kinase. *Biochemistry*, Vol. 30, No. 22, pp. 5329-5334, 0006-2960 (Print) 0006-2960 (Linking).
- Giver, L., Gershenson, A., Freskgard, P.O. & Arnold, F.H. (1998). Directed evolution of a thermostable esterase. *Proc Natl Acad Sci U S A*, Vol. 95, No. 22, pp. 12809-12813.
- Grishin, N.V. (2001). Fold change in evolution of protein structures. *J Struct Biol*, Vol. 134, No. 2-3, pp. 167-185, 1047-8477 (Print) 1047-8477 (Linking).
- Gromiha, M.M., Oobatake, M. & Sarai, A. (1999). Important amino acid properties for enhanced thermostability from mesophilic to thermophilic proteins. *Biophysical Chemistry*, Vol. 82, No. 1, pp. 51-67.
- Guzman-Maldonado, H. & Paredes-Lopez, O. (1995). Amylolytic enzymes and products derived from starch: a review. *Crit Rev Food Sci Nutr*, Vol. 35, No. 5, pp. 373-403.
- Gyemant, G., Hovanszki, G. & Kandra, L. (2002). Subsite mapping of the binding region of alpha-amylases with a computer program. *Eur J Biochem*, Vol. 269, No. 21, pp. 5157-5162.
- Halabi, N., Rivoire, O., Leibler, S. & Ranganathan, R. (2009). Protein sectors: evolutionary units of three-dimensional structure. *Cell*, Vol. 138, No. 4, pp. 774-786, 1097-4172 (Electronic) 0092-8674 (Linking).
- Hancock, S.M., Corbett, K., Fordham-Skelton, A.P., Gatehouse, J.A. & Davis, B.G. (2005). Developing promiscuous glycosidases for glycoside synthesis: residues W433 and E432 in *Sulfolobus solfataricus* beta-glycosidase are important glucoside- and galactoside-specificity determinants. *Chembiochem*, Vol. 6, No. 5, pp. 866-875, 1439-4227 (Print) 1439-4227 (Linking).
- Henrissat, B. (1991). A classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem J*, Vol. 280 (Pt 2), No., pp. 309-316, 0264-6021 (Print) 0264-6021 (Linking).
- Henrissat, B. & Bairoch, A. (1993). New families in the classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem J*, Vol. 293 (Pt 3), No., pp. 781-788, 0264-6021 (Print) 0264-6021 (Linking).
- Henrissat, B. & Bairoch, A. (1996). Updating the sequence-based classification of glycosyl hydrolases. *Biochem J*, Vol. 316 (Pt 2), No., pp. 695-696, 0264-6021 (Print).
- Henrissat, B. & Davies, G. (1997). Structural and sequence-based classification of glycoside hydrolases. *Curr Opin Struct Biol*, Vol. 7, No. 5, pp. 637-644.
- Hibbert, E.G., Baganz, F., Hailes, H.C., Ward, J.M., Lye, G.J., Woodley, J.M. & Dalby, P.A. (2005). Directed evolution of biocatalytic processes. *Biomol Eng*, Vol. 22, No. 1-3, pp. 11-19, 1389-0344 (Print) 1389-0344 (Linking).
- Hirata, A., Adachi, M., Sekine, A., Kang, Y.N., Utsumi, S. & Mikami, B. (2004a). Structural and enzymatic analysis of soybean beta-amylase mutants with increased pH optimum. *J Biol Chem*, Vol. 279, No. 8, pp. 7287-7295, 0021-9258 (Print) 0021-9258 (Linking).
- Hirata, A., Adachi, M., Utsumi, S. & Mikami, B. (2004b). Engineering of the pH optimum of *Bacillus cereus* beta-amylase: conversion of the pH optimum from a bacterial type to a higher-plant type. *Biochemistry*, Vol. 43, No. 39, pp. 12523-12531, 0006-2960 (Print) 0006-2960 (Linking).
- Jochens, H., Aerts, D. & Bornscheuer, U.T. (2010). Thermostabilization of an esterase by alignment-guided focussed directed evolution. *Protein Eng Des Sel*, Vol. 23, No. 12, pp. 903-909, 1741-0134 (Electronic) 1741-0126 (Linking).

- Jochens, H. & Bornscheuer, U.T. (2010). Natural diversity to guide focused directed evolution. *ChemBiochem*, Vol. 11, No. 13, pp. 1861-1866, 1439-7633 (Electronic) 1439-4227 (Linking).
- Joerger, A.C., Mayer, S. & Fersht, A.R. (2003). Mimicking natural evolution in vitro: an N-acetylneuraminase lyase mutant with an increased dihydrodipicolinate synthase activity. *Proc Natl Acad Sci U S A*, Vol. 100, No. 10, pp. 5694-5699.
- Jurgens, C., Strom, A., Wegener, D., Hettwer, S., Wilmanns, M. & Sterner, R. (2000). Directed evolution of a (beta alpha)₈-barrel enzyme to catalyze related reactions in two different metabolic pathways. *Proc Natl Acad Sci U S A*, Vol. 97, No. 18, pp. 9925-9930.
- Kadziola, A., Sogaard, M., Svensson, B. & Haser, R. (1998). Molecular structure of a barley alpha-amylase-inhibitor complex: implications for starch binding and catalysis. *J Mol Biol*, Vol. 278, No. 1, pp. 205-217.
- Kandra, L., Gyemant, G., Remenyik, J., Hovanszki, G. & Liptak, A. (2002). Action pattern and subsite mapping of *Bacillus licheniformis* alpha- amylase (BLA) with modified maltooligosaccharide substrates. *FEBS Lett*, Vol. 518, No. 1-3, pp. 79-82.
- Kaplan, J. & DeGrado, W.F. (2004). De novo design of catalytic proteins. *Proc Natl Acad Sci U S A*, Vol. 101, No. 32, pp. 11566-11570.
- Kim, T.J., Park, C.S., Cho, H.Y., Cha, S.S., Kim, J.S., Lee, S.B., Moon, T.W., Kim, J.W., Oh, B.H. & Park, K.H. (2000). Role of the glutamate 332 residue in the transglycosylation activity of *ThermusMaltogenic* amylase. *Biochemistry*, Vol. 39, No. 23, pp. 6773-6780, 0006-2960 (Print).
- Kortemme, T., Kim, D.E. & Baker, D. (2004). Computational alanine scanning of protein-protein interfaces. *Sci STKE*, Vol. 2004, No. 219, pp. pl2, 1525-8882 (Electronic) 1525-8882 (Linking).
- Kristan, K., Stojan, J., Adamski, J. & LaniÅ;nik RiÅ¼ner, T. (2007). Rational design of novel mutants of fungal 17Î²-hydroxysteroid dehydrogenase. *J Biotechnol*, Vol. 129, No. 1, pp. 123-130.
- Kuhlman, B., Dantas, G., Ireton, G.C., Varani, G., Stoddard, B.L. & Baker, D. (2003). Design of a novel globular protein fold with atomic-level accuracy. *Science*, Vol. 302, No. 5649, pp. 1364-1368, 1095-9203 (Electronic) 0036-8075 (Linking).
- Kuipers, R.K., Joosten, H.J., van Berkel, W.J., Leferink, N.G., Rooijen, E., Ittmann, E., van Zimmeren, F., Jochens, H., Bornscheuer, U., Vriend, G., dos Santos, V.A. & Schaap, P.J. (2010). 3DM: systematic analysis of heterogeneous superfamily data to discover protein functionalities. *Proteins*, Vol. 78, No. 9, pp. 2101-2113, 1097-0134 (Electronic) 0887-3585 (Linking).
- Kumar, S. & Nussinov, R. (2001). How do thermophilic proteins deal with heat? *Cell Mol Life Sci*, Vol. 58, No. 9, pp. 1216-1233, 1420-682X (Print) 1420-682X (Linking).
- Kumar, S., Tsai, C.J. & Nussinov, R. (2000). Factors enhancing protein thermostability. *Protein Eng*, Vol. 13, No. 3, pp. 179-191, 0269-2139 (Print) 0269-2139 (Linking).
- Kuriki, T., Kaneko, H., Yanase, M., Takata, H., Shimada, J., Handa, S., Takada, T., Umeyama, H. & Okada, S. (1996). Controlling substrate preference and transglycosylation activity of neopullulanase by manipulating steric constraint and hydrophobicity in active center. *J Biol Chem*, Vol. 271, No. 29, pp. 17321-17329.
- Ladenstein, R. & Antranikian, G. (1998). Proteins from hyperthermophiles: stability and enzymatic catalysis close to the boiling point of water. *Adv Biochem Eng Biotechnol*, Vol. 61, No., pp. 37-85.

- Landwehr, M., Carbone, M., Otey, C.R., Li, Y. & Arnold, F.H. (2007). Diversification of catalytic function in a synthetic family of chimeric cytochrome p450s. *Chem Biol*, Vol. 14, No. 3, pp. 269-278, 1074-5521 (Print).
- Leemhuis, H., Rozeboom, H.J., Dijkstra, B.W. & Dijkhuizen, L. (2004). Improved thermostability of bacillus circulans cyclodextrin glycosyltransferase by the introduction of a salt bridge. *Proteins*, Vol. 54, No. 1, pp. 128-134, 1097-0134 (Electronic) 0887-3585 (Linking).
- Lehmann, M., Kostrewa, D., Wyss, M., Brugger, R., D'Arcy, A., Pasamontes, L. & van, L.A. (2000). From DNA sequence to improved functionality: using protein sequence comparisons to rapidly design a thermostable consensus phytase. *Protein Eng*, Vol. 13, No. 1, pp. 49-57.
- Lehmann, M. & Wyss, M. (2001). Engineering proteins for thermostability: the use of sequence alignments versus rational design and directed evolution. *Curr Opin Biotechnol*, Vol. 12, No. 4, pp. 371-375.
- Levy, M. & Ellington, A.D. (2001). Selection of deoxyribozyme ligases that catalyze the formation of an unnatural internucleotide linkage. *Bioorg Med Chem*, Vol. 9, No. 10, pp. 2581-2587.
- Li, Y., Drummond, D.A., Sawayama, A.M., Snow, C.D., Bloom, J.D. & Arnold, F.H. (2007). A diverse family of thermostable cytochrome P450s created by recombination of stabilizing fragments. *Nat Biotechnol*, Vol. 25, No. 9, pp. 1051-1056, 1087-0156 (Print).
- Lichtarge, O., Bourne, H.R. & Cohen, F.E. (1996). An evolutionary trace method defines binding surfaces common to protein families. *J Mol Biol*, Vol. 257, No. 2, pp. 342-358, 0022-2836 (Print) 0022-2836 (Linking).
- Lichtarge, O., Yao, H., Kristensen, D.M., Madabushi, S. & Mihalek, I. (2003). Accurate and scalable identification of functional sites by evolutionary tracing. *J Struct Funct Genomics*, Vol. 4, No. 2-3, pp. 159-166, 1345-711X (Print) 1345-711X (Linking).
- Liebl, W., Stemplinger, I. & Ruile, P. (1997). Properties and gene structure of the *Thermotoga maritima* alpha-amylase AmyA, a putative lipoprotein of a hyperthermophilic bacterium. *J Bacteriol*, Vol. 179, No. 3, pp. 941-948.
- Liu, Y.H., Lu, F.P., Li, Y., Yin, X.B., Wang, Y. & Gao, C. (2008). Characterisation of mutagenised acid-resistant alpha-amylase expressed in *Bacillus subtilis* WB600. *Appl Microbiol Biotechnol*, Vol. 78, No. 1, pp. 85-94, 0175-7598 (Print) 0175-7598 (Linking).
- MacGregor, E.A. (1988). Alpha-amylase structure and activity. *J Protein Chem*, Vol. 7, No. 4, pp. 399-415, 0277-8033 (Print) 0277-8033 (Linking).
- Mackay, R.M., Baird, S., Dove, M.J., Erratt, J.A., Gines, M., Moranelli, F., Nasim, A., Willick, G.E., Yaguchi, M. & Seligy, V.L. (1985). Glucanase gene diversity in prokaryotic and eukaryotic organisms. *Biosystems*, Vol. 18, No. 3-4, pp. 279-292, 0303-2647 (Print) 0303-2647 (Linking).
- Machius, M., Declerck, N., Huber, R. & Wiegand, G. (1998). Activation of *Bacillus licheniformis* alpha-amylase through a disorder-->order transition of the substrate-binding site mediated by a calcium- sodium-calcium metal triad. *Structure*, Vol. 6, No. 3, pp. 281-292.
- Machius, M., Wiegand, G. & Huber, R. (1995). Crystal structure of calcium-depleted *Bacillus licheniformis* alpha- amylase at 2.2 Å resolution. *J Mol Biol*, Vol. 246, No. 4, pp. 545-559.

- Mansfeld, J., Vriend, G., Dijkstra, B.W., Veltman, O.R., Van den Burg, B., Venema, G., Ulbrich-Hofmann, R. & Eijssink, V.G. (1997). Extreme stabilization of a thermolysin-like protease by an engineered disulfide bond. *J Biol Chem*, Vol. 272, No. 17, pp. 11152-11156.
- Matsubara, S. (1961). Studies on Taka-amylase A. VII. Transmaltosidation by Taka-amylase A. *J Biochem*, Vol. 49, No. 3, pp. 226-231.
- Matsui, I., Ishikawa, K., Miyairi, S., Fukui, S. & Honda, K. (1992a). Alteration of bond-cleavage pattern in the hydrolysis catalyzed by *Saccharomycopsis* alpha-amylase altered by site-directed mutagenesis. *Biochemistry*, Vol. 31, No. 22, pp. 5232-5236.
- Matsui, I., Ishikawa, K., Miyairi, S., Fukui, S. & Honda, K. (1992b). A mutant alpha-amylase with enhanced activity specific for short substrates. *FEBS Lett*, Vol. 310, No. 3, pp. 216-218.
- Matsumura, I. & Ellington, A.D. (2001). In vitro evolution of beta-glucuronidase into a beta-galactosidase proceeds through non-specific intermediates. *J Mol Biol*, Vol. 305, No. 2, pp. 331-339.
- Maxwell, K.L. & Davidson, A.R. (1998). Mutagenesis of a Buried Polar Interaction in an SH3 Domain: Sequence Conservation Provides the Best Prediction of Stability Effects. *Biochemistry*, Vol. 37, No. 46, pp. 16172-16182.
- McCarter, J.D. & Withers, S.G. (1994). Mechanisms of enzymatic glycoside hydrolysis. *Curr Opin Struct Biol*, Vol. 4, No., pp. 885-892.
- Meiler, J. & Baker, D. (2006). ROSETTALIGAND: protein-small molecule docking with full side-chain flexibility. *Proteins*, Vol. 65, No. 3, pp. 538-548, 1097-0134 (Electronic) 0887-3585 (Linking).
- Minshull, J. & Stemmer, W.P. (1999). Protein evolution by molecular breeding. *Curr Opin Chem Biol*, Vol. 3, No. 3, pp. 284-290.
- Moreno-Beltran, A., Salgado, L., Vazquez-Duhalt, R. & López-Munguía, A. (1999). Modelling the alcoholysis reaction of beta-galactosidase in reverse micelles. *J Mol Catalysis B: Enzymatic*, Vol. 6, No. 1-2, pp. 1-10.
- Moreno, A., Damian-Almazo, J.Y., Miranda, A., Saab-Rincon, G., Gonzalez, F. & Lopez-Munguia, A. (2010). Transglycosylation reactions of *Thermotoga maritima* [alpha]-amylase. *Enzyme and Microbial Technology*, Vol. 46, No. 5, pp. 331-337.
- Morley, K.L. & Kazlauskas, R.J. (2005). Improving enzyme properties: when are closer mutations better? *Trends Biotechnol*, Vol. 23, No. 5, pp. 231-237, 0167-7799 (Print) 0167-7799 (Linking).
- Nagi, A.D. & Regan, L. (1997). An inverse correlation between loop length and stability in a four-helix-bundle protein. *Fold Des*, Vol. 2, No. 1, pp. 67-75, 1359-0278 (Print) 1359-0278 (Linking).
- Nakajima, R., Imanaka, T. & Aiba, S. (1986). Comparison of amino acid sequences of eleven different α -amylases. *Appl Microbiol Biotechnol*, Vol. 23, No. 5, pp. 355-360.
- Nielsen, J.E. & Borchert, T.V. (2000). Protein engineering of bacterial alpha-amylases. *Biochim Biophys Acta*, Vol. 1543, No. 2, pp. 253-274.
- Nikolova, P.V., Henckel, J., Lane, D.P. & Fersht, A.R. (1998). Semirational design of active tumor suppressor p53 DNA binding domain with enhanced stability. *Proc Natl Acad Sci U S A*, Vol. 95, No. 25, pp. 14675-14680, 0027-8424 (Print) 0027-8424 (Linking).
- Ochoa-Leyva, A., Soberon, X., Sanchez, F., Arguello, M., Montero-Moran, G. & Saab-Rincon, G. (2009). Protein design through systematic catalytic loop exchange in the

- (beta/alpha)⁸ fold. *J Mol Biol*, Vol. 387, No. 4, pp. 949-964, 1089-8638 (Electronic) 0022-2836 (Linking).
- Palackal, N., Brennan, Y., Callen, W.N., Dupree, P., Frey, G., Goubet, F., Hazlewood, G.P., Healey, S., Kang, Y.E., Kretz, K.A., Lee, E., Tan, X., Tomlinson, G.L., Verruto, J., Wong, V.W., Mathur, E.J., Short, J.M., Robertson, D.E. & Steer, B.A. (2004). An evolutionary route to xylanase process fitness. *Protein Sci*, Vol. 13, No. 2, pp. 494-503, 0961-8368 (Print) 0961-8368 (Linking).
- Pantazes, R.J., Saraf, M.C. & Maranas, C.D. (2007). Optimal protein library design using recombination or point mutations based on sequence-based scoring functions. *Protein Eng Des Sel*, Vol. 20, No. 8, pp. 361-373, 1741-0126 (Print).
- Paramesvaran, J., Hibbert, E.G., Russell, A.J. & Dalby, P.A. (2009). Distributions of enzyme residues yielding mutants with improved substrate specificities from two different directed evolution strategies. *Protein Eng Des Sel*, Vol. 22, No. 7, pp. 401-411, 1741-0134 (Electronic) 1741-0126 (Linking).
- Park, H.S., Nam, S.H., Lee, J.K., Yoon, C.N., Mannervik, B., Benkovic, S.J. & Kim, H.S. (2006). Design and evolution of new catalytic activity with an existing protein scaffold. *Science*, Vol. 311, No. 5760, pp. 535-538.
- Park, S., Morley, K.L., Horsman, G.P., Holmquist, M., Hult, K. & Kazlauskas, R.J. (2005). Focusing mutations into the *P. fluorescens* esterase binding site increases enantioselectivity more effectively than distant mutations. *Chem Biol*, Vol. 12, No. 1, pp. 45-54, 1074-5521 (Print) 1074-5521 (Linking).
- Pascarella, S. & Argos, P. (1992). Analysis of insertions/deletions in protein structures. *J Mol Biol*, Vol. 224, No. 2, pp. 461-471, 0022-2836 (Print) 0022-2836 (Linking).
- Pavelka, A., Chovancova, E. & Damborsky, J. (2009). HotSpot Wizard: a web server for identification of hot spots in protein engineering. *Nucleic Acids Res*, Vol. 37, No. Web Server issue, pp. W376-383, 1362-4962 (Electronic) 0305-1048 (Linking).
- Perl, D., Mueller, U., Heinemann, U. & Schmid, F.X. (2000). Two exposed amino acid residues confer thermostability on a cold shock protein. *Nat Struct Biol*, Vol. 7, No. 5, pp. 380-383, 1072-8368 (Print) 1072-8368 (Linking).
- Reetz, M.T., Bocola, M., Carballeira, J.D., Zha, D. & Vogel, A. (2005). Expanding the range of substrate acceptance of enzymes: combinatorial active-site saturation test. *Angew Chem Int Ed Engl*, Vol. 44, No. 27, pp. 4192-4196, 1433-7851 (Print) 1433-7851 (Linking).
- Reetz, M.T., Carballeira, J.D., Peyralans, J., Hobenreich, H., Maichele, A. & Vogel, A. (2006a). Expanding the substrate scope of enzymes: combining mutations obtained by CASTing. *Chemistry*, Vol. 12, No. 23, pp. 6031-6038, 0947-6539 (Print).
- Reetz, M.T., Carballeira, J.D. & Vogel, A. (2006b). Iterative saturation mutagenesis on the basis of B factors as a strategy for increasing protein thermostability. *Angew Chem Int Ed Engl*, Vol. 45, No. 46, pp. 7745-7751, 1433-7851 (Print).
- Reetz, M.T., Soni, P. & Fernandez, L. (2009). Knowledge-guided laboratory evolution of protein thermolability. *Biotechnol Bioeng*, Vol. 102, No. 6, pp. 1712-1717, 1097-0290 (Electronic) 0006-3592 (Linking).
- Reetz, M.T., Soni, P., Fernandez, L., Gumulya, Y. & Carballeira, J.D. (2010). Increasing the stability of an enzyme toward hostile organic solvents by directed evolution based on iterative saturation mutagenesis using the B-FIT method. *Chem Commun (Camb)*, Vol. 46, No. 45, pp. 8657-8658, 1364-548X (Electronic) 1359-7345 (Linking).

- Rivera, M.H., Lopez-Munguia, A., Soberon, X. & Saab-Rincon, G. (2003). Alpha-amylase from *Bacillus licheniformis* mutants near to the catalytic site: effects on hydrolytic and transglycosylation activity. *Protein Eng*, Vol. 16, No. 7, pp. 505-514.
- Rodríguez-Zavala, J.S. (2008). Enhancement of coenzyme binding by a single point mutation at the coenzyme binding domain of *E. coli* lactaldehyde dehydrogenase. *Protein Science*, Vol. 17, No. 3, pp. 563-570.
- Rodríguez, M., Gómez, A., González, F., Bárzana, E. & López-Munguía, A. (1996). Selectivity of methyl-fructoside synthesis with β -fructofuranosidase. *Appl Biochem Biotechnol*, Vol. 59, No., pp. 167-175.
- Rogers, J.C. (1985). Conserved amino acid sequence domains in alpha-amylases from plants, mammals, and bacteria. *Biochemical and Biophysical Research Communications*, Vol. 128, No. 1, pp. 470-476.
- Rosell, A., Valencia, E., Ochoa, W.F., Fita, I., Pares, X. & Farres, J. (2003). Complete Reversal of Coenzyme Specificity by Concerted Mutation of Three Consecutive Residues in Alcohol Dehydrogenase. *J. Biol. Chem.*, Vol. 278, No. 42, pp. 40573-40580.
- Saab-Rincon, G., del-Rio, G., Santamaria, R.I., Lopez-Munguia, A. & Soberon, X. (1999). Introducing transglycosylation activity in a liquefying alpha-amylase. *FEBS Lett*, Vol. 453, No. 1-2, pp. 100-106.
- Sakamoto, T., Joern, J.M., Arisawa, A. & Arnold, F.H. (2001). Laboratory evolution of toluene dioxygenase to accept 4-picoline as a substrate. *Appl Environ Microbiol*, Vol. 67, No. 9, pp. 3882-3887.
- Sali, A., Potterton, L., Yuan, F., van, V.H. & Karplus, M. (1995). Evaluation of comparative protein modeling by MODELLER. *Proteins Struct Funct Genet*, Vol. 23, No. 3, pp. 318-326.
- Sanchez, R. & Sali, A. (1997). Evaluation of comparative protein structure modeling by MODELLER-3. *Proteins Struct Funct Genet*, Vol. 1997, No. 50, pp. 50-58.
- Santamaria, R.I., Del Rio, G., Saab, G., Rodriguez, M.E., Soberon, X. & Lopez-Manguia, A. (1999). Alcoholysis reactions from starch with alpha-amylases. *FEBS Lett*, Vol. 452, No. 3, pp. 346-350.
- Saravanan, M., Vasu, K. & Nagaraja, V. (2008). Evolution of sequence specificity in a restriction endonuclease by a point mutation. *Proc Natl Acad Sci U S A*, Vol. 105, No. 30, pp. 10344-10347, 1091-6490 (Electronic) 0027-8424 (Linking).
- Savile, C.K., Janey, J.M., Mundorff, E.C., Moore, J.C., Tam, S., Jarvis, W.R., Colbeck, J.C., Krebber, A., Fleitz, F.J., Brands, J., Devine, P.N., Huisman, G.W. & Hughes, G.J. (2010). Biocatalytic asymmetric synthesis of chiral amines from ketones applied to sitagliptin manufacture. *Science*, Vol. 329, No. 5989, pp. 305-309, 1095-9203 (Electronic) 0036-8075 (Linking).
- Schmitzer, A.R., Lepine, F. & Pelletier, J.N. (2004). Combinatorial exploration of the catalytic site of a drug-resistant dihydrofolate reductase: creating alternative functional configurations. *Protein Eng Des Sel*, Vol. 17, No. 11, pp. 809-819, 1741-0126 (Print) 1741-0126 (Linking).
- Shinoyama, H., Kamiyama, Y. & Yasui, T. (1988). Enzymatic synthesis of alkyl β -xylosides from xylobiose by application of the transxylosyl reaction of *Aspergillus niger* β -xylosidase. *Agric Biol Chem*, Vol. 52, No., pp. 2197-2202.
- Siegel, J.B., Zanghellini, A., Lovick, H.M., Kiss, G., Lambert, A.R., St Clair, J.L., Gallaher, J.L., Hilvert, D., Gelb, M.H., Stoddard, B.L., Houk, K.N., Michael, F.E. & Baker, D. (2010). Computational design of an enzyme catalyst for a stereoselective

- bimolecular Diels-Alder reaction. *Science*, Vol. 329, No. 5989, pp. 309-313, 1095-9203 (Electronic) 0036-8075 (Linking).
- Sinnott, M.L. (1990). Catalytic Mechanisms of Enzymic Glycosyl Transfer. *Chem Rev*, Vol. 90, No., pp. 1171-1202.
- Song, J.K., Chung, B., Oh, Y.H. & Rhee, J.S. (2002). Construction of DNA-shuffled and incrementally truncated libraries by a mutagenic and unidirectional reassembly method: changing from a substrate specificity of phospholipase to that of lipase. *Appl Environ Microbiol*, Vol. 68, No. 12, pp. 6146-6151.
- Song, J.K. & Rhee, J.S. (2000). Simultaneous enhancement of thermostability and catalytic activity of phospholipase A(1) by evolutionary molecular engineering. *Appl Environ Microbiol*, Vol. 66, No. 3, pp. 890-894.
- Spector, S., Wang, M., Carp, S.A., Robblee, J., Hendsch, Z.S., Fairman, R., Tidor, B. & Raleigh, D.P. (2000). Rational modification of protein stability by the mutation of charged surface residues. *Biochemistry*, Vol. 39, No. 5, pp. 872-879, 0006-2960 (Print) 0006-2960 (Linking).
- Stemmer, W.P. (1994a). DNA shuffling by random fragmentation and reassembly: in vitro recombination for molecular evolution. *Proc Natl Acad Sci U S A*, Vol. 91, No. 22, pp. 10747-10751.
- Stemmer, W.P. (1994b). Rapid evolution of a protein in vitro by DNA shuffling. *Nature*, Vol. 370, No. 6488, pp. 389-391.
- Straathof, A.J.J., Vrijenhoef, J.P., Sprangers, E.P.A.T., Bekkum, H.v. & Kieboom, A.P.G. (1988). Enzymic formation of β -D-fructofuranosides from sucrose: activity and selectivity of invertase in mixtures of water and alcohol. *J Carbohydr Chem*, Vol. 7, No., pp. 223-238.
- Suel, G.M., Lockless, S.W., Wall, M.A. & Ranganathan, R. (2003). Evolutionarily conserved networks of residues mediate allosteric communication in proteins. *Nat Struct Biol*, Vol. 10, No. 1, pp. 59-69, 1072-8368 (Print) 1072-8368 (Linking).
- Sullivan, B.J., Durani, V. & Magliery, T.J. (2011). Triosephosphate isomerase by consensus design: dramatic differences in physical properties and activity of related variants. *J Mol Biol*, Vol. 413, No. 1, pp. 195-208, 1089-8638 (Electronic) 0022-2836 (Linking).
- Svensson, B. (1994). Protein engineering in the alpha-amylase family: catalytic mechanism, substrate specificity, and stability. *Plant Mol Biol*, Vol. 25, No. 2, pp. 141-157.
- Turner, N.J. (2009). Directed evolution drives the next generation of biocatalysts. *Nat Chem Biol*, Vol. 5, No. 8, pp. 567-573, 1552-4469 (Electronic) 1552-4450 (Linking).
- Uchiyama, H., Inaoka, T., Ohkuma-Soyejima, T., Togame, H., Shibanaka, Y., Yoshimoto, T. & Kokubo, T. (2000). Directed evolution to improve the thermostability of prolyl endopeptidase. *J Biochem (Tokyo)*, Vol. 128, No. 3, pp. 441-447.
- Ulmer, K.M. (1983). Protein engineering. *Science*, Vol. 219, No. 4585, pp. 666-671, 0036-8075 (Print) 0036-8075 (Linking).
- Vallee, B.L., Stein, E.A., Sumerwell, W.N. & Fischer, E.H. (1959). Metal content of alpha-amylases of various origins. *J Biol Chem*, Vol. 234, No., pp. 2901-2905, 0021-9258 (Print) 0021-9258 (Linking).
- van der Veen, B.A., Leemhuis, H., Kralj, S., Uitdehaag, J.C.M., Dijkstra, B.W. & Dijkhuizen, L. (2001). Hydrophobic Amino Acid Residues in the Acceptor Binding Site Are Main Determinants for Reaction Mechanism and Specificity of Cyclodextrin-glycosyltransferase. *J. Biol. Chem.*, Vol. 276, No. 48, pp. 44557-44562.

- Vihinen, M. & Mantsala, P. (1989). Microbial amylolytic enzymes. *Crit Rev Biochem Mol Biol*, Vol. 24, No. 4, pp. 329-418.
- Vihinen, M. & Mantsala, P. (1990). Conserved residues of liquefying alpha-amylases are concentrated in the vicinity of active site. *Biochem Biophys Res Commun*, Vol. 166, No. 1, pp. 61-65.
- Vulfson, E.N., Patel, R., Beecher, J.E., Andrews, A.T. & Law, B.A. (1990). Glycosidases in organic solvents: I. Alkyl β -glucoside synthesis in a water-organic two-phase system. *Enzyme Microb Technol*, Vol. 12, No., pp. 950-954.
- Wilming, M., Iffland, A., Tafelmeyer, P., Arrivoli, C., Saudan, C. & Johnsson, K. (2002). Examining reactivity and specificity of cytochrome c peroxidase by using combinatorial mutagenesis. *Chembiochem*, Vol. 3, No. 11, pp. 1097-1104, 1439-4227 (Print) 1439-4227 (Linking).
- Woodyer, R., van der Donk, W.A. & Zhao, H. (2003). Relaxing the nicotinamide cofactor specificity of phosphite dehydrogenase by rational design. *Biochemistry*, Vol. 42, No. 40, pp. 11604-11614, 0006-2960 (Print) 0006-2960 (Linking).
- Yamashiro, K., Yokobori, S., Koikeda, S. & Yamagishi, A. (2010). Improvement of *Bacillus circulans* beta-amylase activity attained using the ancestral mutation method. *Protein Eng Des Sel*, Vol. 23, No. 7, pp. 519-528, 1741-0134 (Electronic) 1741-0126 (Linking).
- Yin, Q., Teng, Y., Ding, M. & Zhao, F. (2011). Site-directed mutagenesis of aromatic residues in the carbohydrate-binding module of *Bacillus* endoglucanase EGA decreases enzyme thermostability. *Biotechnol Lett*, Vol. 33, No. 11, pp. 2209-2216, 1573-6776 (Electronic) 0141-5492 (Linking).
- Zanghellini, A., Jiang, L., Wollacott, A.M., Cheng, G., Meiler, J., Althoff, E.A., Rothlisberger, D. & Baker, D. (2006). New algorithms and an in silico benchmark for computational enzyme design. *Protein Sci*, Vol. 15, No. 12, pp. 2785-2794, 0961-8368 (Print) 0961-8368 (Linking).
- Zhang, J.H., Dawes, G. & Stemmer, W.P. (1997). Directed evolution of a fucosidase from a galactosidase by DNA shuffling and screening. *Proc Natl Acad Sci U S A*, Vol. 94, No. 9, pp. 4504-4509.
- Zhao, H. & Arnold, F.H. (1999). Directed evolution converts subtilisin E into a functional equivalent of thermitase. *Protein Eng*, Vol. 12, No. 1, pp. 47-53.
- Zhao, H., Giver, L., Shao, Z., Affholter, J.A. & Arnold, F.H. (1998). Molecular evolution by staggered extension process (StEP) in vitro recombination. *Nat Biotechnol*, Vol. 16, No. 3, pp. 258-261.