ENGINEERING OF GLYCOSIDE HYDROLASES OF BIOTECHNOLOGICAL INTEREST FROM EXTREMOPHILIC ORGANISMS

Fiorella Conte

Dottorato in Scienze Biotecnologiche – XXI ciclo
Indirizzo Biotecnologie molecolari
Università di Napoli Federico II

Dottorato in Scienze Biotecnologiche – XXI ciclo
Indirizzo Biotecnologie molecolari
Università di Napoli Federico II

ENGINEERING OF GLYCOSIDE HYDROLASES OF BIOTECHNOLOGICAL INTEREST FROM EXTREMOPHILIC ORGANISMS

Fiorella Conte

Dottoranda: Fiorella Conte

Relatore: Prof.ssa Simonetta Bartolucci

Correlatore: Dott. Marco Moracci

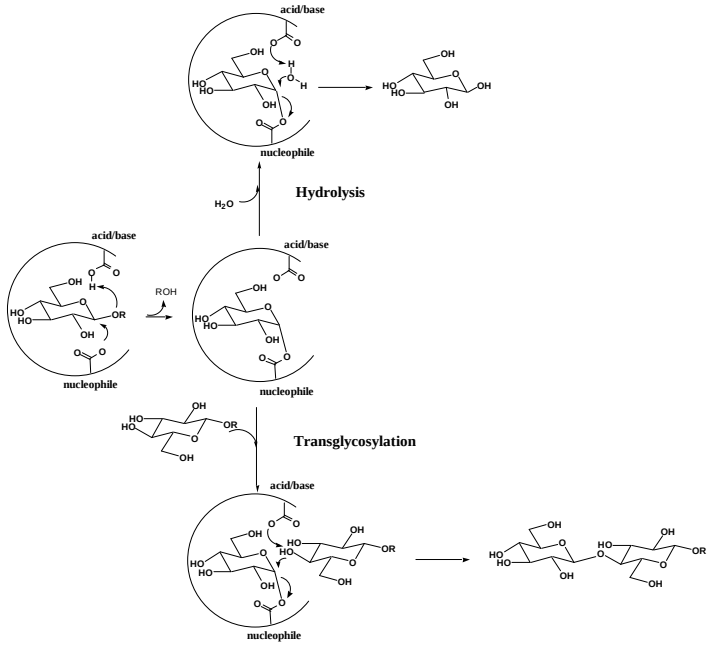
Coordinatore: Prof. Giovanni Sannia

A David e Lorenzo,
perché con loro ho imparato
cosa è importante nella vita

INDICE

Riassunto

I carboidrati sono i biopolimeri più abbondanti in natura e svolgono funzioni di primaria importanza in tutti gli organismi viventi. Infatti, essi costituiscono la fonte di energia per eccellenza per gli organismi chemioeterotrofi; il glicogeno e l'amido, ad esempio, sono polisaccaridi utilizzati dalle cellule come riserva di energia che, una volta idrolizzati a monosaccaridi, entrano nel ciclo metabolico. Inoltre, i carboidrati hanno la capacità di mediare interazioni cellula-cellula sotto forma di glicoconiugati di proteine e lipidi, sono recettori nella trasduzione del segnale, nella risposta antigenica e sono, inoltre, responsabili della compartimentalizzazione delle proteine. Per quanto riguarda la funzione strutturale, la parete cellulare delle piante è costituita da cellulosa, emicellulosa e proteine; la sua modifica e ricostruzione è un prerequisito fondamentale in molti processi della vita delle piante, come la germinazione, la crescita, lo sviluppo dei frutti e la risposta all'attacco di agenti patogeni.All'enorme varietà strutturale dei mono-, oligo- e polisaccaridi presenti in natura corrisponde un gran numero di enzimi in grado di modificarli. Per quanto riguarda la degradazione dei carboidrati, in tutti e tre i domini dei viventi è presente una classe di enzimi, le glicosil idrolasi, che catalizzano l'idrolisi di legami glicosidici. Le glicosil idrolasi (GH) sono divise in due grandi categorie in base al meccanismo di reazione proposto da Koshland [Koshland, 1953] e tutt'oggi ancora accettato. Gli enzimi che idrolizzano i legami glicosidici con l'inversione della configurazione del carbonio anomerico del prodotto rispetto a quella del substrato si definiscono *inverting*; quelli che, invece, effettuano la catalisi mantenendo il carbonio anomerico nella sua configurazione originale vengono detti *retaining*. In entrambi i casi, nel sito catalico, sono presenti due residui amminoacidici (solitamente acido aspartico e/o glutammico) che effettuano direttamente la catalisi e sono definiti nucleofilo e acido-base della reazione. Come mostrato nella figura seguente, nelle glicosil idrolasi di tipo *retaining,* il meccanismo è a due stadi e porta alla formazione di un intermedio covalente glicosil-enzima la cui risoluzione avviene grazie all'intervento di una molecola d'acqua (definita *acceptor*).



Meccanismo di reazione delle glicosidasi *retaining:* idrolisi e transglicosilazione.

Nel caso in cui intervenga un *acceptor* diverso dall'acqua, come per esempio un alcool o uno zucchero, la reazione può portare alla sintesi di un legame glicosidico tra il substrato (detto *donor*) e l'*acceptor*. Questa particolare reazione viene definita di "transglicosilazione".

Lo svantaggio nell'applicazione di questa tecnica alla sintesi di oligosaccaridi è la bassa resa, essendo il prodotto di sintesi comunque substrato dell'enzima. Nel 1998, il gruppo dove ho svolto la tesi, ha dimostrato che, mutando il residuo che funge da nucleofilo della reazione con uno non nucleofilo, l'enzima, con l'aggiunta di un nucleofilo esterno, come lo ione azide o formiato, è in grado di effettuare l'idrolisi di substrati aventi un buon gruppo uscente ovvero un gruppo capace di facilitare l'idrolisi del substrato da parte dell'enzima modificato [Moracci et al., 1998]. In questo caso, il disaccaride prodotto dalla reazione non può essere idrolizzato in quanto presenta un cattivo gruppo uscente che ne ostacola l'idrolisi. Di conseguenza il prodotto di sintesi si accumula nella reazione. Per questo motivo gli enzimi così modificati vengono definiti glicosintasi.

Le glicosintasi termofile, tra quelle riportate in letteratura, sono tra le più interessanti, in quanto permettono di associare alle proprietà utili per la sintesi, anche il vantaggio della stabilità alle alte temperature e, in generale, a condizioni estreme quali l'elevata concentrazione di substrato, di solventi organici (necessari per poter aumentare la solubilità dei substrati ad elevate concentrazioni) e pH acidi (necessari per migliorare lo stato di ionizzazione del residuo che funge da acido/ base nel sito attivo) [Perugino et al., 2004; 2005].

In letteratura, a differenza delle β-glicosintasi, ampiamente caratterizzate, è stata riportata solo una α-glicosintasi derivante da una GH di tipo *retaining*: si tratta del mutante D481G dell'α-glucosidasi da *Schizosaccharomyces pombe* [Okuyama et al., 2002]. La difficoltà nell'ottenere α-glicosintasi è da ricondursi probabilmente alla diversa struttura del sito attivo delle α-glicosidasi che in qualche modo rende più difficile l'approccio. Da qui quindi, la necessità di studiare in dettaglio il meccanismo di reazione delle α-glicosidasi dal punto di vista molecolare al fine di ottenere un'α-glicosintasi. In quest'ottica, nel laboratorio in cui ho svolto il mio lavoro di tesi è stata identificata un'α-L-fucosidasi da *Sulfolobus solfataricus;* questa proteina è codificata da due ORFs separate (SSO11867 e SSO3060) [Cobucci Ponzano et al., 2003a] che codificano per polipeptidi omologhi all' N- ed al C- terminale delle α-L-fucosidasi di famiglia 29, suggerendo che il gene sia espresso *in vivo* attraverso un tipo di controllo traduzionale chiamato *frameshifting* di tipo -1 che fa parte di una classe di meccanismi per la regolazione della traduzione genica detta *recoding* [Baranov et al., 2002]. L'inserzione di una singola base, per mutagenesi sito-diretta, nella regione di sovrapposizione tra SSO11867 e SSO3060, ha permesso di ottenere un' α-fucosidasi termofilica e termostabile che è stata ampiamente caratterizzata. Questo enzima, chiamato *Ss*α-fuc, promuove reazioni di transfucosilazione mediante un meccanismo di reazione di tipo *retaining* [Cobucci Ponzano et al., 2003a]. Esso è stato oggetto di studio della mia tesi. Nel laboratorio in cui ho svolto il mio lavoro erano stati identificati i residui catalici di Ssα-fuc; l'Asp242 svolge il ruolo di nucleofilo ed è stato identificato per la prima volta in un enzima di famiglia 29 [Cobucci Ponzano et al., 2003b]. Mentre i residui che fungono da catalizzatori acido base in

Ssα-fuc sono due: il Glu58 e il Glu292 [Cobucci Ponzano et al., 2005], suggerendo che il meccanismo di reazione coinvolge tre residui catalitici; recentemente tutti i dati sono stati confermati e raccolti in una *review* [Cobucci Ponzano et al., 2008a]. In base a queste caratteristiche, Ssα-fuc è il candidato ideale per lo studio del meccanismo di reazione delle α-glicosidasi e della loro modifica α -glicosintasi.

In parte del mio lavoro di tesi mi sono occupata dello studio dell'espressione di queste due ORF in *E.coli;* utilizzando, infatti, mutanti sito-diretti nella sequenza di sovrapposizione delle due ORF è stato possibile confermare la presenza di una zona di circa sette nucleotidi, detta *slippery sequence,* in cui avviene lo scivolamento della fase di lettura, che viene riconosciuta dall'apparato traduzionale di *E.coli* e porta alla produzione di polipeptide *full-lenght* con un'efficienza del 5%. Sono stati condotti, inoltre, studi volti all'identificazione delle sequenze di regolazione *in vivo* grazie, infatti, ad esperimenti di traduzione *in vitro* con estratti di *S. solfataricus* in collaborazione con il gruppo della Prof.ssa Londei dell'Università di Roma "Sapienza", è stato possibile dimostrare con certezza che la zona di scivolamento è quella predetta [Cobucci Ponzano, 2003a] e che i due microrganismi possono mostrare un diverso comportamento dell'apparato traduzionale in presenza di *slippery sequence* anche con una sola sostituzione nucleotidica; infatti, il mutante in cui è stata modificata un'unica base all'interno della sequenza in oggetto, in *E.coli* produce un polipeptide completo, cosa che invece non avviene in *S. solfataricus*. E' la prima volta che è dimostrata la presenza di questo tipo di controllo della traduzione nel dominio degli Archaea [Cobucci Ponzano et al., 2006].

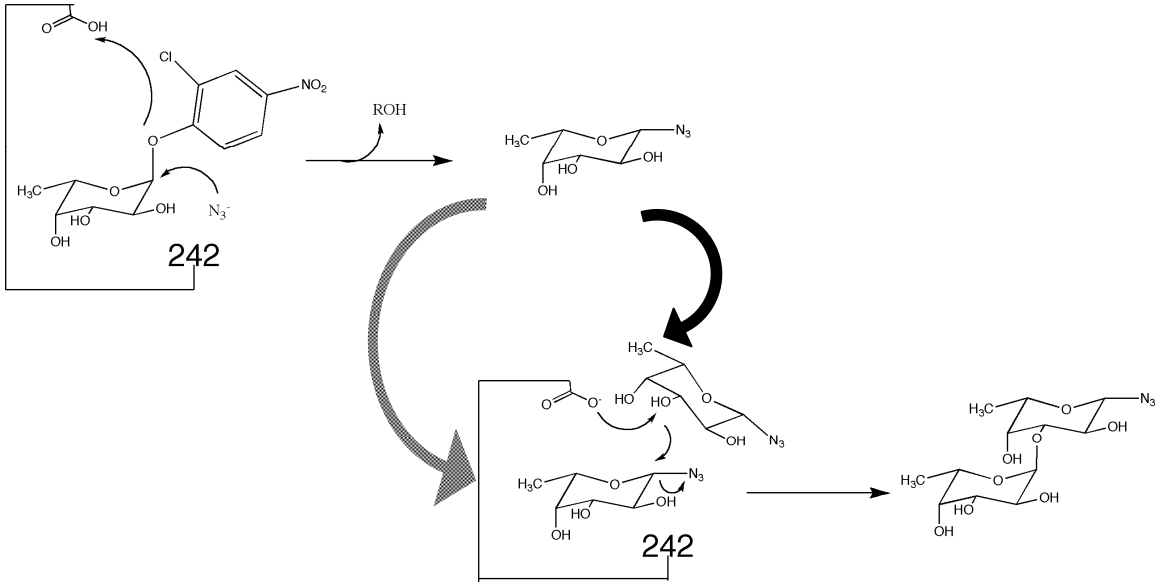
Un altro aspetto, quindi, del mio lavoro di tesi, è consistito nel caratterizzare mutanti nel nucleofilo della reazione di α-fucosidasi termofile, al fine di ottenere fucosintasi.

Le analisi preliminari hanno riguardato il mutante SsD242G in presenza di para-nitrofenil-α-L-fucopiranoside (4Np-α-L-fuc; *donor*) a concentrazioni saturanti, in presenza di nucleofili esterni, quali gli ioni formiato, azide, cloruro e fluoruro a diverse concentrazioni (da 0.1 a 2M) ed a pH compresi tra 4.0 e 6.5. In queste condizioni, anche in presenza di lattosio, pNpβ-D-galattopiranoside e pNpβ-D-glucopiranoside, come *acceptors,* non ho osservato alcuna attività sintetica [Cobucci Ponzano et al., 2008b].

Allo scopo di migliorare le capacità sintetiche dell'enzima ho analizzato l'attività del mutante SsD242G su *donors* alternativi, dotati di un buon gruppo uscente, in quanto il primo stadio della reazione è facilitato [Moracci et al., 1998]. In collaborazione con il gruppo del prof. Parrilli dell'Università degli Studi di Napoli "Federico II", è stato preparato il substrato sintetico 2-cloro-4-nitrofenil-α-L-fucopiranoside (2C4Np-α-L-fuc), che possiede un gruppo uscente migliore del 4Np-α-L-fuc. L'utilizzo di questo composto ha dato i frutti sperati: l'attività dell'enzima, in presenza di 2M azide come nucleofilo esterno, su questo substrato si è rivelata più elevata rispetto al 4Np-α-L-fuc.

Nonostante diversi tentativi in presenza di questo nuovo substrato e diversi nucleofili esterni, sono riuscita ad ottenere un prodotto di sintesi solo in presenza di 0.1 M azide, a pH 6.5, ad una concentrazione di 2C4Np-α-L-fuc pari a 20 mM. Tutte le reazioni sono state analizzate per cromatografia su strato sottile (TLC). Una volta individuati i probabili prodotti sintetizzati, questi sono stati analizzati per NMR e spettrometria di massa in collaborazione con il gruppo della Prof.ssa Corsaro dell'Università di Napoli "Federico II". Da quest'ultima analisi è stato individuato il composto α-L-Fuc-(1-3)-β-L-Fuc-N_3 come prodotto di sintesi.

Per migliorare l'attività sintetica del mutante del nucleofilo dell'α-fucosidasi da *S. solfataricus*, ho sostituito, mediante mutagenesi sito diretta, il residuo di acido aspartico 242 con un residuo di serina, ottenendo il mutante SsD242S. Anche questo mutante in presenza di azide e 2C4Np-α-L-fuc alle stesse condizioni di SsD242G, è in grado di sintetizzare la α-L-Fuc-(1-3)-β-L-Fuc-N_3. Però in queste condizioni il mutante SsD242S ha mostrato un'efficienza di tranfucosilazione maggiore rispetto al mutante SsD242G. Alla luce di questo risultato quindi il meccanismo di reazione proposto potrebbe coinvolgere la molecola di β-L-fucosil azide (β-L-fuc-N_3) ottenuta nel primo passaggio della reazione, che, una volta trovatasi nel sito del *donor* viene attivata e trasferita ad una seconda molecola di β-L-fuc-N_3 che subentra nel sito dell'*acceptor* e formare così il legame glicosidico di tipo α, come illustrato nella figura seguente .



Meccanismo di reazione del mutante SsD242S

Quindi, per valutare quest'ipotesi, ho condotto la reazione a 65℃ a pH 6.5 in presenza della sola β-L-fuc-N_3 sintetizzata dal gruppo del Prof. Parrilli, sia con i mutanti SsD242S che SsD242G: anche in questo caso gli enzimi sono in grado di produrre la α-L-Fuc-(1-3)-β-L-Fuc-N_3, ma senza l'aiuto di un nucleofilo esterno. Questo prodotto così inaspettato conferma che il mutante del nucleofilo è in grado di trasferire il *donor* β-L-fuc-N_3, nonostante si tratti di una specie chimica stabile. Per analizzare inoltre, la capacità del mutante SsD242S di trasferire lo stesso *donor* ad altri *acceptor* l'enzima è stato saggiato in presenza di pNpβ-D-galattopiranoside, pNpN-acetil-β-D-glucosaminide e pNpβ-D-xilopiranoside con risultati positivi. I prodotti di sintesi ottenuti sono stati caratterizzati in collaborazione con il laboratorio del Prof.ssa Corsaro. Per le tre reazioni suddette è stata inoltre determinata l'efficienza di trasferimento del fucosio utilizzando una cromatografia liquida a scambio anionico da alta pressione con rivelatore ad amperometria pulsata (HPAEC-PAD). Per la reazione contenente come *acceptor* pNpβ-D-galattopiranoside l'efficienza è risultata essere pari al 26%, per pNpN-acetil-β-D- glucosaminide è 86% e per pNpβ-D-xilopiranoside è 50%.

Per valutare se questo approccio è di applicabilità generale, abbiamo prodotto il mutante in glicina del residuo nucleofilo dell'α-L-fucosidasi di *Thermotoga maritima* (TmD224G), appartenente alla stessa famiglia delle GH29 [Tarling et al., 2003]. L'enzima mutante è stato caratterizzato ampliamente dal punto di vista biochimico e utilizzato in reazioni di sintesi. A differenza di SsD242S, il mutante TmD224G in presenza del 2C4Np-α-L-fuc e sodio azide, e della sola β-L-fuc-N_3 non dà prodotti disaccaridici. Invece in presenza di β-L-fuc-N_3 come *donor* e di diversi *acceptors*, l'enzima è in grado di trasferire il fucosio. Ciò accade probabilmente perché la β-L-fuc N_3 non è in grado di posizionarsi nel sito dell'*acceptor*, impedendo così la formazione di prodotti di autocondensazione. Anche in questo caso è stata determinata l'efficienza di trasferimento mediante HPAEC-PAD che è risultata essere del 91% nel caso in cui come *acceptor* è utilizzato il pNpβ-D-xilopiranoside. I prodotti sono stati caratterizzati per NMR e spettrometria di massa, ed è stata stabilita l'anomericità e la posizione dei legami neosintetizzati.

In conclusione, durante il mio lavoro di tesi ho ottenuto la prima α-fucosintasi termofila. Ho messo a punto una nuova metodologia per la sintesi di prodotti glicosidici utilizzando come *donor* glicosil azidi aventi configurazione anomerica inversa rispetto a quella del substrato.

L'importanza di questi risultati è data soprattutto dalla possibilità di estendere questa metodologia ad altri mutanti delle glicosil idrolasi di tipo α per l'ottenimento di oligosaccaridi d'interesse biotecnologico

Summary

Carbohydrates are among the structural components of the cell and are considered the most common energy source, but also take part in a variety of molecular recognition processes in intracellular communication and signal transduction such as cell adhesion, differentiation, development and regulation [Sears and Wong 1996; Varki 1993]. Given their importance, oligosaccharides are used as target for several drugs. Enzymatic synthesis of oligosaccharides is considered an interesting approach to control the regio- and stereo-specificiy of the reaction. In biological systems glycosyl transferase (GTs) and glycosidases (GHs) are responsible for the synthesis and catabolism of carbohydrates. Approaches that use the GTs are troublesome because these enzymes are scarce and the relative cost of their substrates is high. Instead, retaining GHs, widespread enzymes using cheap substrates, can synthesize glycosidic bonds transglycosylation. However the products of tranglycosylation reactions can be hydrolyzed by the enzyme because they have the same anomeric configuration of the substrate. This causes reaction yields lower then 40%. Using the nucleophile mutants of retaining GH, in which the nucleophile residue is substituted with a non-nucleophile amino acid as glycine o serine, they can be reactivated in the presence of external nucleophiles (chemical rescue) and they can carry out tranglycosylation reactions, since the mutated enzyme is not capable to hydrolyze the product and it accumulates in the reaction. This particular class of engineering enzymes was named glycosynthases [Moracci et al., 1998]. To date, despite several β-glycosynthase have been described, only one retaining α-glycosynthase is known.

The aim of this work was to study the α-fucosidase from the hyperthermophilic archaeon *Sulfolobus solfataricus.* The final scope was to obtain a novel thermophilic α-fucosynthase. To this aim, I have studied the mechanism of expression *in vivo* of the gene encoding for this enzyme and I demonstrated that its expression is regulated by a mechanism of translational regulation named programmed -1 frameshifting. In addition, I show here that the mutant SsD242S is able use the β-L-fucosyl azide as donor substrate and several arilic compounds as acceptors. This is a novel approach exploiting for the first time a glycoside azide donor for the field of glycosynthases. To validate this new technique, I obtained another α-fucosynthase, from the nucleophile mutant of the α-fucosidase from the thermophilic bacterium *Thermotoga maritima* (TmD224G).

In conclusion the SsD242S and the TmD224G mutants, are the first thermophilic α-fucosynthases that produce glycosidic compounds with high efficiency, from β-L-fucosyl azide like donor and in presence of several acceptors.

Introduction

Glycobiology and its application

Carbohydrates are commonly involved in cellular structure and are considered the most common energy source. However, they are involved in a variety of molecular recognition processes in intracellular communication and signal transduction such as cell adhesion, differentiation, development and regulation [Sears and Wong 1996; Varki, 1993].
These observations motivated the interest for carbohydrate-based drugs; several have already introduced on the market and many others are in various phases of clinical trials. For example anticancer agents, vaccines based on synthetic carbohydrates against HIV and *Haemophilus influenzae* viruses, etc. [Slovin et al., 1999; Allen et al., 2001; Ratner et al., 2002; Geng et al., 2004; Verez-Bencomo e al., 2004].
Among the great variety of oligosaccharides, those containing fucose have an important role in nature. In fact they are involved in a variety of biological events as growth regulators and receptors in signal transduction, cell-cell interaction and antigenic response [Vanhooren et al., 1999]. Moreover fucooligosaccharides are among the most important components of human milk oligosaccharides (HMO) [Bode L., 2006]. Fucosylated oligosaccharides present in HMO, serve as protecting agents of the infant from the virulent attacks of bacteria because they mimic and block the adhesion sites on the cell surface of the bacterial lectines necessaries for the introduction of the guest [Bode L., 2006]. For example, *Campylobacter jujuni*, one of the most predominant causes of diarrhoea worldwide, adheres to intestinal 2'-fucosyllactosammina, which is also present on HMO [Ruiz Palacios et al., 2003]. Fucosylated HMO inhibit *Campylobacter* binding to human intestinal mucosa *ex vivo*, and the incidence of *Campylobacte*r diarrhoea in breast fed infants, is inversely related to the amount of 2'-fucosyllactose in the mother's milk [Morrow et al., 2004]. Only trace amounts of the oligosaccharides are present in mature bovine milk and their identification, characterization and synthesis is of utmost importance for the preparation of bovine milk based infant formula [Bode L., 2006].

The importance of oligosaccharides in many biological systems currently motivates the efforts in developing methods for their large-scale production [Zhou and O'Doherty 2008; Seeberger 2008].

Synthesis of oligosaccharides

Chemical approach

The classical chemical approach involves the careful design of reaction conditions including the choice of the protecting groups, of the catalyst, of the donor leaving groups and of the acceptors. Various excellent chemical methods of obtaining oligosaccharides have been developed, contributing to the elucidation of oligosaccharide biological functions and the development of the potential therapeutic agents [Murata et al., 2006]. In particular, Seeberger and his group have systematically studied all aspects of automated solid-phase synthesis [Seeberger 2008].However, pure classical approach is complicated by the multiple protecting group manipulations to control regioselectivity. Another important aspect is the anomeric control of the reaction: the presence of mixtures of the product, having two

anomeric configurations, implies following purification steps of the species of interest. These procedures often result in time consuming manipulations and low yields, hindering the efficient production of oligosaccharides needed for biological testing [Perugino et al., 2005]. An interesting alternative is the enzymatic synthesis.

Enzymatic approach

Two main classes of biological catalysts are involved in the modification of carbohydrates in nature: glycoside hydrolases (GHs) and glycosyl transferases (GTs). In biological systems they are responsible for the hydrolysis and the synthesis of the sugars respectively. These enzymes are classified in families on the base of their amminoacidic sequence described in a database named CaZy (http://www.cazy.org) [Cantarel et al., 2008]
GTs are classified as belonging to the Leroir and the non-Leroir pathway on the basis of their substrate specificity. Enzymes belonging to the Leloir pathway base their catalysis on sugar nucleotides as donors, while the enzymes of the non-Leroir pathway use glycosyl phosphates as donors. Approaches that use the glycosyl transferase are troublesome because the availability of these enzymes is scarce and the relative cost of their substrates is high [Perugino et al., 2005]
An interesting alternative is enzymatic synthesis by transglycosylation reactions promoted by glycoside hydrolases. These enzymes follow two distinct mechanisms which are termed *inverting* or *retaining* if the enzymatic cleavage of the glycosidic bond liberates a sugar hemiacetal with the opposite or the same anomeric configuration of the glycosidic substrate, respectively. *Inverting* glycosidases operate with a one step, single-displacement mechanism with the assistance of a general acid and a general base group in the active site [McCarter and Withers 1994] (Figure 1).

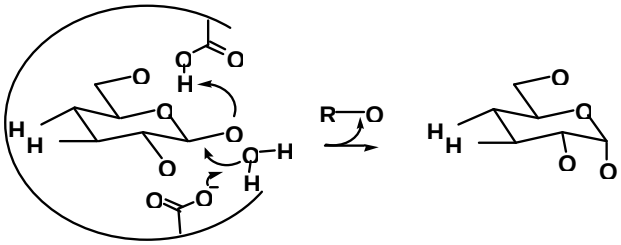


Figure 1: Reaction mechanism of β-Inverting GHs

Instead, *retaining* enzymes follow a two-step mechanism with formation of a covalent glycosyl-enzyme intermediate (Figure 2). The carboxyl group in the enzymatic active site functions as a general acid/base catalyst, and the carboxylate functions as the nucleophile of the reaction [Koshland, 1953]. In the first step (glycosylation step), the nucleophile attacks the anomeric group of the substrate, while the acid/base catalyst, acting in this step as a general acid, protonates the glycosidic oxygen, thereby assisting the leaving of the aglycon moiety. The concerted action of the two amino acids leads to the formation of a covalent glycosyl-enzyme intermediate [Sinnot, 1990; McCarter and Withers, 1994]. In the second step (deglycosylation step), the glycosyl-enzyme intermediate is cleaved by a water molecule that acts as nucleophile being polarized by the general base catalyst. The product of the reaction retained the anomeric configuration of the substrate. When an acceptor different from water, such as an alcohol or a sugar, intercepts the reactive glycosyl-enzyme intermediate, *retaining* enzymes work in transglycosylation mode. In this case, the interglycosidic linkage of the product maintains the same anomeric configuration of

the substrate. This property makes the *retaining* glycosyl hydrolases interesting tools for the synthesis of carbohydrates usually allowing the control of both regio- and stereospecificity. However, yields are not higher than 40% because the products become substrate of the enzyme and can be hydrolyzed.

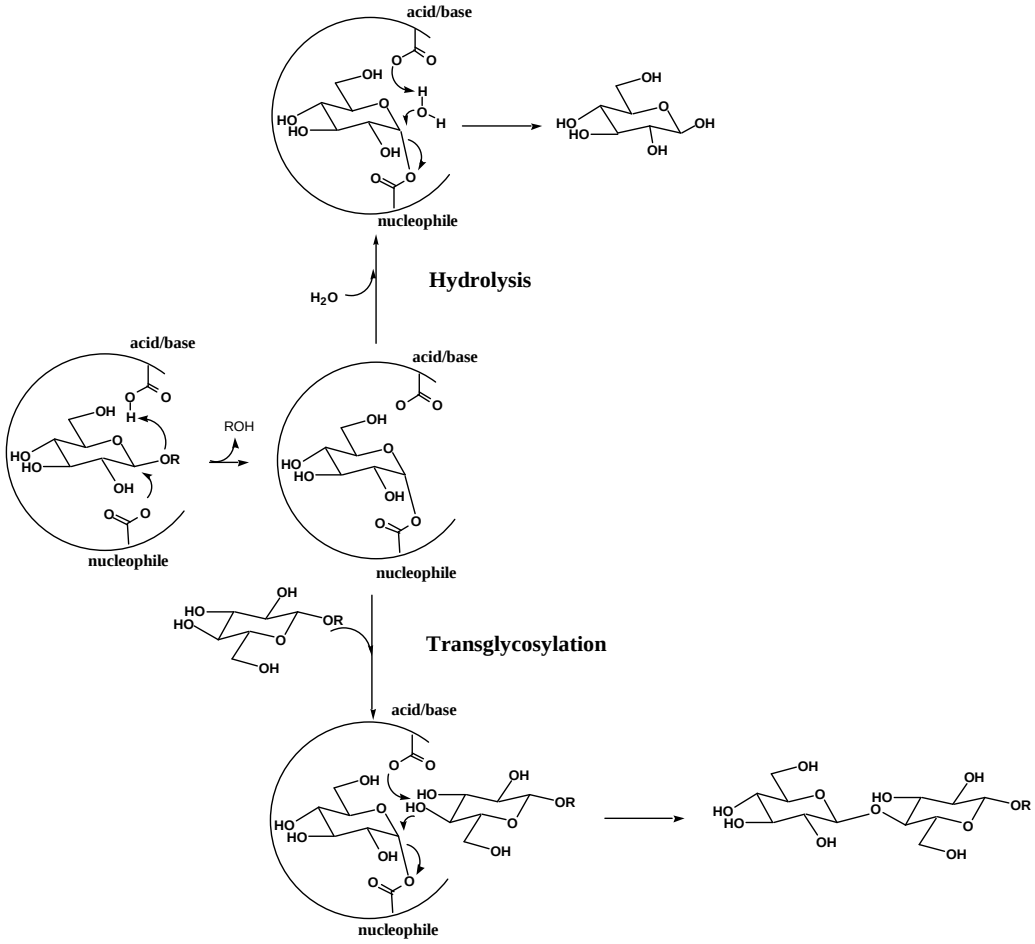


Figure 2:Reaction mechanism of β retaining glycosidases: hydrolysis and tranglycosylation.

Glycosynthetic approach

To improve the reaction yield of transglycosylation and, therefore, to obtain an efficient method for the enzymatic synthesis of oligosaccharides, several years ago was introduced a new class of mutant glycosidases, termed glycosynthases [Mackenzie et al. 1998; Moracci et al. 1998]. Glycosynthases derived from *retaining* glycosidases, in which, the active site nucleophile was replaced with a non-nucleophilic residue in the presence of activated glycosides and suitable reaction conditions, the mutants synthesized oligosaccharides without hydrolysing them. This technology was applied to several *exo*- and *endo*-β-glucosidases. The two main glycosynthase modes of action have been classified into *inverting* and *retaining* on

the base of the anomeric configuration of the donor substrate and the products (Figure 3 and 4).
The mechanism proposed for inverting β-glycosynthases employs α-glycosidic fluorides as donors which mimics the glycosyl enzyme intermediate (Figure 3) [Mackenzie et al., 1998].

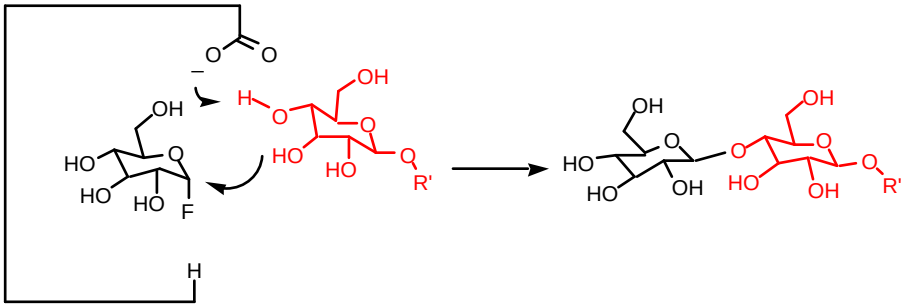


Figure 3: β-inverting glycosynthase [Mackenzie L.F. et al., 1998].

The first reported inverting glycosynthases was a Glu358Ala mutant of the β-glucosidase from *Agrobacterium* sp. [Mackenzie et al., 1998]. The enzyme catalysed a synthetic reaction as described in figure 3 In this case the product of the reaction with the β-anomeric configuration could not be hydrolysed by the mutant enzyme, which was inactivated by mutation. This reaction led to oligosaccharide products with yields typically >80% [Mackenzie et al., 1998]. This invention enabled the production of β-1,4- linked cello-oligosaccharides that can be used as cellulose inhibitors and substrates.

Retaining β-glycosynthases, instead, synthesize oligosaccharides by using activated *donor*-sugars with the same anomeric configuration of the substrate (tipically 2-nitrophenyl- or 2,4-dinitrophenyl-β-glucosides) in the presence of external ions such as sodium formate (Figure 4) [Trincone et al., 2000]. This external nucleophile mimics the natural active site carboxylate promoting the formation of a metastable glycosyl-formate intermediate, which has been isolated in a one case [Viladot et al., 2001]. The products have the same anomeric configuration of the donor, but they cannot be hydrolysed by the enzyme, even in the presence of sodium formate, because of the bad leaving ability of the oligosaccharide groups. Therefore they accumulated in the reaction mixtures leading to quantitative yields.

The first example of retaining β-glycosynthase was obtained from β-glycosidase from *Sulfolobus solfataricus* (Ssβ-gly E387G) [Moracci et al., 1998].

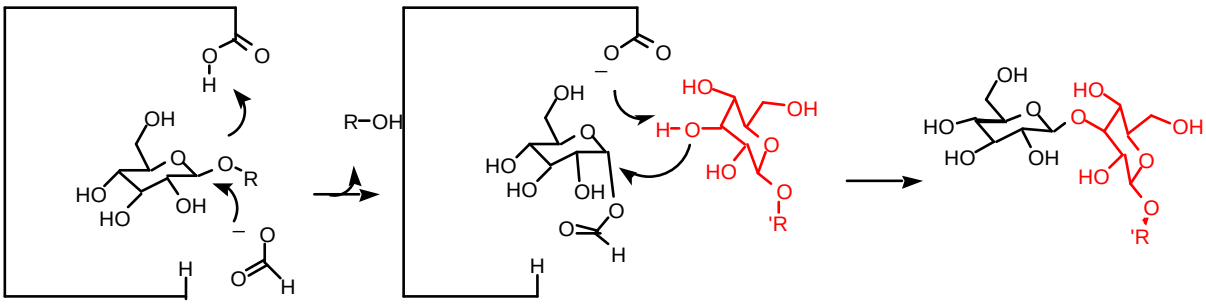


Figure 4: β-retaining glycosynthase. [Moracci et al., 1998].

The Ssβ-gly mutant E387G synthesizes β-1,3- or β-1,6-linked tetrasaccharides, the building blocks of the β-1,3- 1,6- glucans that are recognized as elicitors of the defence response against pathogens in plants and inverterbrates [Perugino et al., 2004].

The approach leading to *inverting* glycosynthases was applied to both exo- and endo-glycosidases belonging to several families of the carbohydrate active enzyme classification, namely, GH1, GH2, GH5, GH7, GH10, GH16, GH17, GH26, GH31, and GH52 [reviewed in Perugino et al. 2005, Hancock et al. 2006] (Table 1).

Table 1

Glycosynthases, thioglycoligases and thioglycosynthases to date

Parent Glycosidase	Glycoside hydrolase family	Catalytic activity	Altered residue(s)	Linkages synthesized	*Endo/exo*
Agrobacterium sp. β-glucosidase (Abg)	GH 1	Glycosynthase	E359	β-1,3/4	*Exo*
		Thioglycoligase	E171	β-1,4	*Exo*
		Thioglycosynthase	E359/E171	β-1,4	*Exo*
S. solfataricus β-glucosidase	GH 1	Glycosynthase	E387	β-1,3/4/6	*Exo*
Thermosphaera aggregans β-glycosidase	GH 1	Glycosynthase	E134	β-1,3/4/6	*Exo*
Pyrococcus furiosus β-glucosidase (CelB)	GH 1	Glycosynthase	E372	β-1,3	*Exo*
T. thermophilus β-glycosidase	GH 1	Glycosynthase	E338	β-1,3	*Exo*
C. fimi β-mannosidase (Man2A)	GH 2	Glycosynthase	E519	β-1,3/4	*Exo*
		Thioglycoligase	E429	β-1,4	*Exo*
E. coli β-galactosidase (LacZ)	GH 2	Glycosynthase	E537	β-1,6	*Exo*
T. maritima β-glucuronidase	GH 2	Glycosynthase	E476	β-1,4	*Exo*
		Thioglycoligase	E383		
Rhodococcus sp. endoglycoceramidase II	GH 5	Glycosynthase	E351	β-1,1	*Endo*
H. insolens cellulase	GH 7	Glycosynthase	E197	β-1,4	*Endo*
B. halodurans reducing-end exo-oligoxylanase	GH 8	Glycosynthase[a]	D263	β-1,4	Endo
C. fimi endo-1,4-β-xylanase	GH 10	Glycosynthase	E235	β–1,4	*Endo*
Bacillus licheniformis 1,3-1,4-β-glucanase	GH 16	Glycosynthase	E143	β-1,3/4	*Endo*
Pyrococcus furiosus laminarinase (LamA)	GH 16	Glycosynthase	E170	β-1,3/4	*Endo*
Hordeum vulgare glucan *endo*-1,3-β-D-glucosidase	GH 17	Glycosynthase	E231	β-1,3	*Endo*
Cellvibrio japonicus mannanase (Man26A)	GH 26	Glycosynthase	E320	β-1,4	*Endo*
Schizosaccharomyces pombe α-glucosidase	GH 31	Glycosynthase	D481	α-1,4	*Exo*
E. coli α-xylosidase (YicI)	GH 31	Thioglycoligase	D482	α-1,4/6	*Exo*
S. solfataricus α-glucosidase	GH 31	Thioglycoligase	D416	α-1,4	*Exo*

[a] Derived from an inverting glycosidase. All other enzymes in this table are generated from retaining glycosidases.

Table 1: Summary of all obtained glycosynthases [Hancock et al., 2006].

Notable examples of glycosynthases include those derived from *Thermus thermophilus* β-glycosidase, which is capable of synthesizing in high yields the β-1,3 glycosidic linkages found in several antigens and in some plant signalling molecules [Dron et al., 2005]. Also of interest is the glycosynthase enzyme developed from *Thermotoga maritima* β-glucuronidase for the synthesis of β-linked glucuronic and galactouronic acid conjugates [Mullegger et al., 2006], which are found in numerous important biological molecules, such as plant and bacterial cell walls and the mammalian glycoaminoglycans, including heparin, heparin sulphate, chondroitin sulphate, and hyaluronan [Whitelock et al., 2005]. Additionaly, the first glycosynthase derived from a endoxylanase (*Cellulmonas fimi* endo β-1,4-xylanase) has been reported, which enables the synthesis of xylo-oligosaccharides [Kim et al., 2006].
Retaining glycosynthases have been obtained only from hyperthermophilic β-glycosidases from family GH1 so far.

Recently, novel glycosynthases have been produced from both *retaining* and *inverting* enzymes, which follow atypical reaction mechanisms. In particular, the *retaining endo*-β-N-acetyl-glucosaminidase from *Mucor hiemalis* (GH85), which follow a substrate-assisted catalysis, was converted into a glycosynthase by mutating the residue involved in the appropriate orientation of the substrate [Umekawa et al. 2008]. Instead, two *inverting* glycosidases, namely the *exo*-oligoxylanase from *B. halodurans* (GH8), and the 1,2-α-L-fucosidase from *Bifidobacterium bifidum* (GH95), acted as glycosynthases once mutated either in the residues working as general base or holding a catalytic water [Honda and Kitaoka 2006; Honda et al. 2008; Wada et al. 2008]. As stated by the authors, these approaches, though interesting, do not provide a common strategy to convert an inverting glycosidase into a glycosynthase [Wada et al. 2008].

Despite the convenience of the method, glycoside hydrolases recalcitrant to become glycosynthases are not uncommon [Ducros et al. 2003; Cobucci-Ponzano et al. 2003b; Perugino et al. 2005]. The most noticeable example is that of retaining α-glycosidases, which are not prone to such approach and only one documented case of α-glucosynthase is available to date [Okuyama et al. 2002]. This is the D481G mutant of α-glucosidase from *Schizosaccharomyces pombe* that leads to products in no more then 41% yields [Okuyama et al. 2002].

I addressed my efforts in developing a novel α-L-fucosynthase for the quantitative production of fucosylated oligosaccharides. The interest in developing synthetic methods is based on the potential applications in biomedicine of fucooligosaccharides [de La Torre et al. 2002; Listinsky et al. 1998; Mori et al. 1998; Noda et al. 1998;Vanhooren and Vandamme 1999].
Transfucosylation reactions utilizing porcine liver and *Alcaligenes* sp. α-L-fucosidases yielded α-L-fucosyl-N-acetyllactosamines products in yields ranging 5-54% [Murata et al. 1999]. In addition, variants obtained by directed evolution of the α-L-fucosidase from *Thermotoga maritima* (Tmα-fuc) allowed to increase the transfucosylation yields from 7% (wild type) to 60%. Noticeably, this result was obtained by mutating three amino acids not directly involved in catalysis [Osanjo et al. 2007].
I have chosen, as model systems, the α-fucosidases from the hyperthermophilic archeon *Sulfolobus solfataricus* and from the thermophilic bacterium *Thermotoga maritima*.

The α-fucosidase from **Sulfolobus solfataricus**

In the laboratory where I carried out my thesis project, it has been identified and characterized the first archaeal α-L-fucosidase [Cobucci-Ponzano et al. 2003a; 2008a]. In fact, the analysis of the genome of the hyperthermophilic archaeon *S. solfataricus* [She et al. 2001] revealed the presence of two open reading frames (ORFs), SSO11867 and SSO3060, encoding for 81 and 426 amino acid polypeptides that are homologous to the *N*- and the *C*-terminal parts, respectively, of full-length bacterial and eukaryal GH29 fucosidases. This gene named *fucA1* contained the regulatory signals of a peculiar translation mechanism of regulation of gene expression named programmed -1 frameshifting [Cobucci Ponzano et al., 2003a]. This type of control of the translation process is part of a class of mechanisms of regulation of gene translation named *recoding* [Cobucci Ponzano et al., 2005b]. *Recoding* is used to increase the diversity in gene expression or for its regulation, and it includes: programmed ribosome frameshifting to a different reading frame, ribosome hopping over nucleotides, and reading of stop codons as sense codons (readthrough). Among recoding events, programmed -1 frameshifting are by far the most prevalent [Farabaugh, 1996]; They have been well characterized in RNA viruses, but the basic molecular mechanism governing these events are almost identical from yeast to human [Hammel et al., 1999].
The programmed -1 frameshifting is triggered by several elements in the mRNA. The slippery sequence, showing the X-XXY-YYZ motif, in which X can be any base, Y is usually A or U, and Z is any base but G, has the function of favouring the tRNA misalignment and it is the site where the shift takes place [Farabaugh 1996, Baranov et al., 2002]. Frameshifting could be further stimulated by other elements flanking the slippery sequence: a codon for a low–abundance tRNA, a stop codon, a Shine-Dalgarno sequence in Bacteria and an mRNA secondary structure. Noticeably, known cases of *recoding* in Archaea [Cobucci Ponzano et al., 2005 b] are limited to termination codon readthrough events that regulate the incorporation of the 21st and 22nd amino acids selenocysteine and pyrrolysine, respectively [Wilting et al., 1997; Hao et al., 2002; Polycarpo et al., 2003; Blight et al., 2004].
No archaeal genes regulated by translational programmed frameshifting and ribosome hopping have been identified experimentally so far; therefore, if compared with the others domains of life, the study of translational recoding in Archaea is still at its down.

In the case of α-L-fucosidase, the two ORFs SSO3060 and SSO11867 found in *S.solfataricus* genome are separated by a -1 frameshift in a 40-base overlap (Figure 5 A). It has been previously reported that the region of overlap between the two ORFs had the characteristic features of the genes expressed by programmed -1 frameshifting [Cobucci Ponzano, et al., 2003a]. The mutation by site-directed mutagenesis the *fucA1* gene exactly in the position predicted from the slippery site, produced a full-length gene, named *fucA1*^{A}, encoding for a polypeptide of 495 amino acids (Figure 5 B). This mutant gene expressed in *E. coli* a fully functional α-L-fucosidase, named Ssα-fuc [Cobucci-Ponzano et al. 2003a; reviewed in Cobucci Ponzano et al., 2008a]. This mutant enzyme is a nonamer of 57 kDa molecular mass subunits in solution and it is highly active and specific for 4nitrophenyl-α-L-fucopyranoside (4NP-α-L-Fuc) at 65℃ [Cobucci-Ponzano et al. 2003a; 2008a,

Rosano.et al. 2004]. Moreover, Ssα-fuc is thermoactive and thermostable, as expected for an enzyme from a hyperthermophilic microorganism. The optimal temperature of the enzyme is 95℃ and it displayed high stability maintaining 60% of the residual activity after 2 h at 80℃ [Cobucci-Po nzano et al. 2003a, 2008a]. For all the subsequent enzymatic characterization I used Ssα-fuc enzyme.

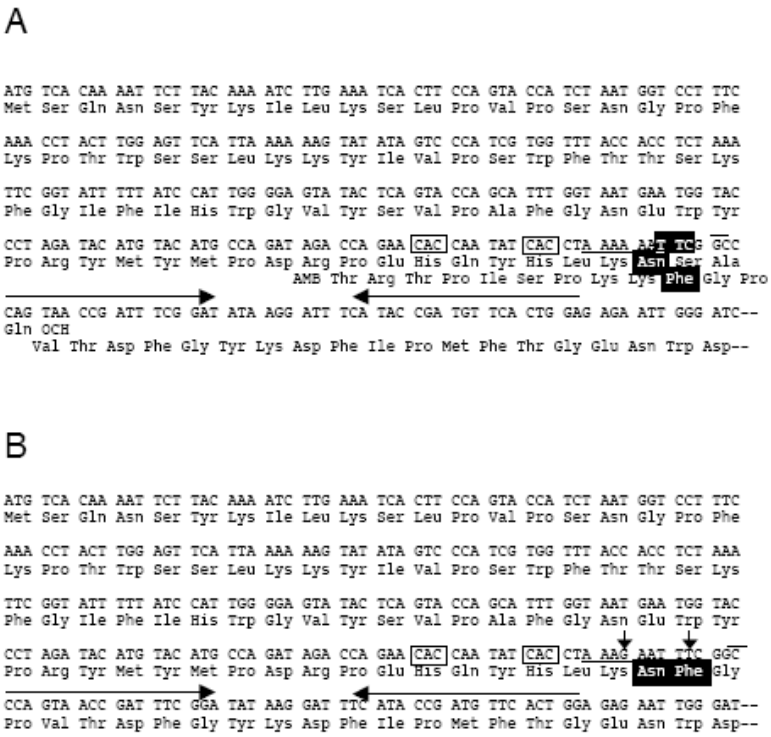


Figure 5: The α-fucosidase gene. (A) Region of overlap in the wild type split *fucA1* gene. The N-terminal SSO11867 ORF is in the zero frame, the C-terminal SSO3060 ORF, for which only a fragment is shown, is in the –1 frame. The slippery heptameric sequence is underlined; the rare codons are boxed and the arrows indicate the stems of the putative mRNA secondary structure. The amino acids involved in the programmed –1 frameshifting and the first codon translated after this event in the –1 frame are shadowed. (B) Fragment of the full-length mutant *fucA1*^{A} gene. The small arrows indicate the mutated nucleotides [Cobucci Ponzano et al., 2006]

Enzymatic Characterization of Ssα-fuc

The reaction mechanism of Ssα-fuc

The reaction mechanism of Ssα-fuc was studied in detail and the residues directly involved in catalysis were identified. The *retaining* reaction mechanism was demonstrated, for the first time in GH29, by using Ssα-fuc. In fact, the enzyme is able to function in transfucosylation mode [Cobucci-Ponzano et al. 2003a, 2008a]. The α-anomeric configuration of the interglycosidic linkages in the products demonstrated that GH29 α–fucosidases follow a *retaining* reaction mechanism (Figure 6) [Cobucci-Ponzano et al. 2003a]. The hydrolytic activity of Ssα-fuc on the disaccharide α-L-Fuc-(1-3)-α-L-Fuc-*O*-4-NP revealed that the enzyme is an exo-glycosyl hydrolase that attacks the substrates from their non-reducing end [Cobucci-Ponzano et al. 2003a, 2008a].

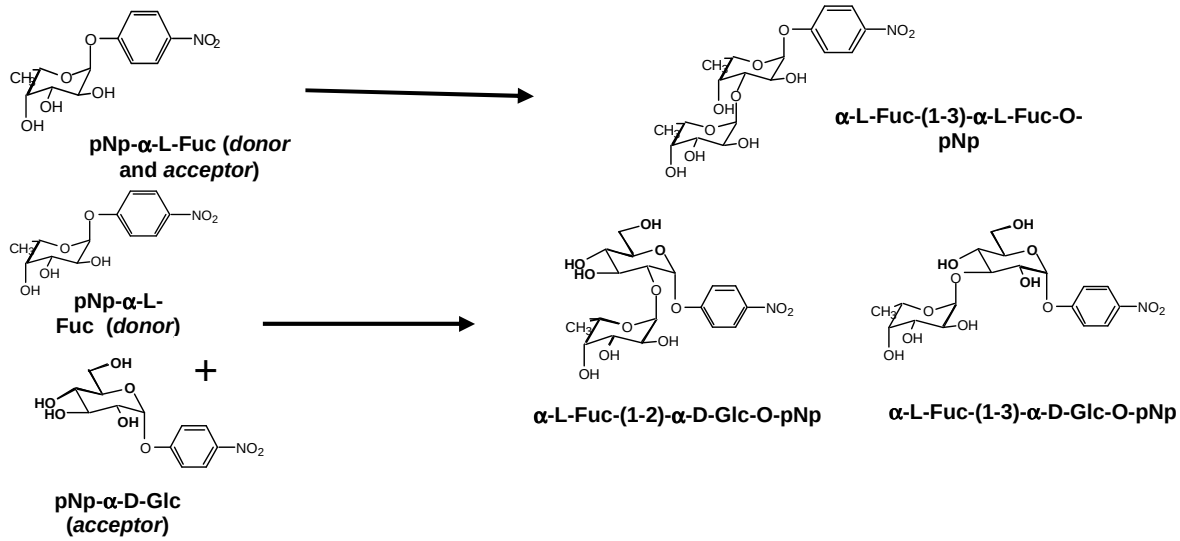


Figure 6: Transfucosylating activity of Ssα-fuc on 4Npα-L-fuc like donor and 4Npα-D-glc like acceptor.

The catalytic residues of Ssα-fuc.

The catalytic residues of glycoside hydrolases can be identified with a variety of methods, including mechanism-based inhibitors labelling the catalytic nucleophile and inspection of X-ray structures [McCarter and Withers 1996; Vocadlo et al. 2000; Vocadlo et al. 2001; Tarling et al. 2003]. A powerful approach often exploited for retaining glycoside hydrolases consists in the mutation of the aspartic/glutamic acid residues identified by sequence analysis and conserved in the family of interest. Mutations of the catalytic residues with non-nucleophilic amino acids lead to the strong reduction or even abolition of the enzymatic activity [Ly and Withers 1999]. However, these mutants can be reactivated in the presence of external nucleophiles such as sodium azide. This methodology is named chemical rescue. The isolation of glycosyl-azide products with an anomeric configuration opposite to that of the substrate allows the identification of the catalytic nucleophile of the reaction (Figure 7) [Ly and Withers 1999]. Instead the isolation of glycosyl-azide products with an

anomeric configuration identical to that of the substrate allows the identification of the catalytic acid/base residue of the reaction (Figure 8). Once the reaction mechanism and the active-site residues of a particular enzyme have been experimentally determined, they can be easily extended to all the homologous enzymes by following the classification in families.

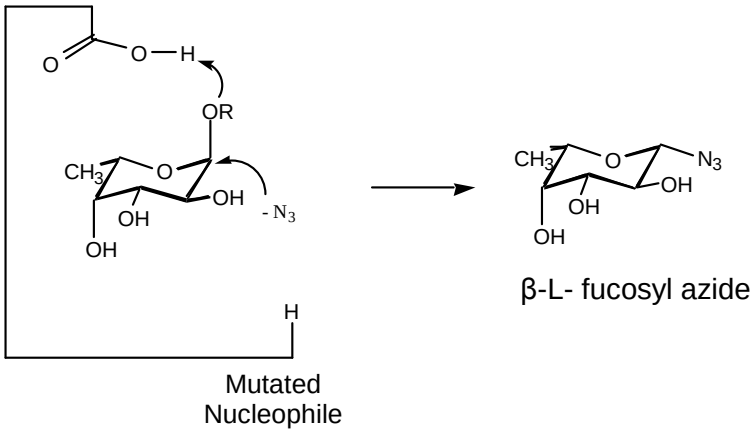


Figure 7: Identification of nucleophile residue of an α-L-fucosidase by chemical rescue

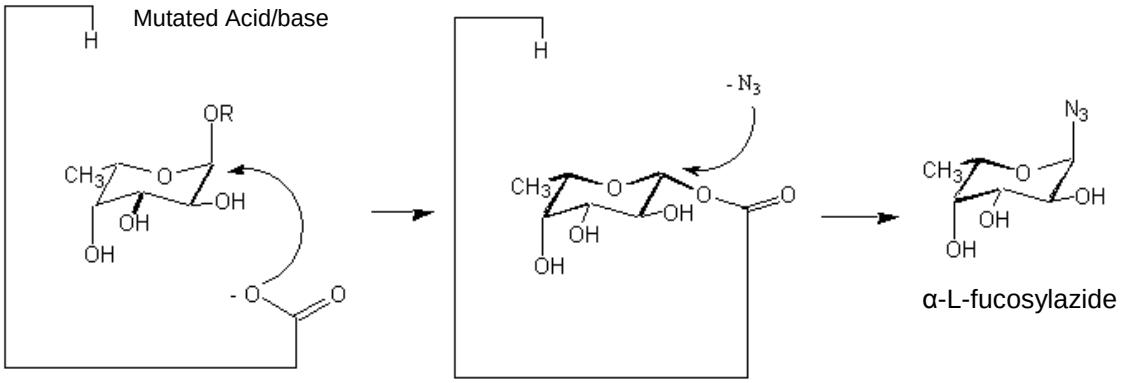


Figure 8: Identification of acid/base residue of an α-L-fucosidase by chemical rescue

By following this methodology, the nucleophile of Ssα-fuc was identified [Cobucci-Ponzano et al. 2003b]. In fact, the fucosyl-azide product obtained by the SsD242G mutant was found in the inverted (β-L) configuration compared with the substrate. This finding allowed, for the first time, the unambiguous assignment of Asp242 in Ssα-fuc and its homologous residues as the nucleophilic catalytic residues of GH29 α-L-fucosidases [Cobucci-Ponzano et al. 2003b, 2008a]. This was the first example of the application of the chemical rescue method to α-(D/L)-glycosidases, demonstrating that chemical rescue could be of general applicability for retaining enzymes.
The chemical rescue approach resulted less straightforward for the identification of the acid /base residue of famly GH29 .
In Ssα-fuc several aminoacids among highly conserved histidine, aspartic, and glutamic acid residues were picked and mutated into glycine; the characterization of the mutants SsE58G and SsE292G, compared to the data collected on the

corresponding residues in the α-fucosidase of thermophilic bacterium *Thermotoga maritima* (Tmα-fuc) (Glu66 and Glu262, respectively), gave unexpected results. The analysis of the 3D structure of the Tmα-fuc and the kinetic characterization of the mutants clearly indicated that, in this enzyme, Glu66 and Glu266 were involved in substrate binding and in the acid/base catalysis, respectively [Sulzenbacher et al. 2004]. Instead kinetic analysis of the mutants, suggested that in Ssα-fuc Glu58 is the acid/base catalyst, [Cobucci Ponzano et al., 2005b; 2008a], while the Glu292 has a relevant role in catalysis presumably modulating the pK_a of the E58. Considering that among the amino acid sequences of GH29 the predicted acid/base residues are not invariant, it would not be surprising that the enzymes show structural differences in the active site explaining the different catalytic machineries and suggesting that in GH29 two catalytic machineries coexist. [Cobucci Ponzano et al., 2008a].

The α-L-fucosidase from Thermotoga maritima

Thermotoga maritima is a thermophilic bacterium producing a wide range of carbohydrate-processing enzymes including a 52 kDa α-L-fucosidase that has 38% sequence identity and 56% similarity to human fucosidases. This enzyme assembles as examer and displays a two domain fold, composed of catalytic $(\beta/\alpha)_8$-like domain and a C-terminal β-sandwiches domain [Sulzenbacher et al., 2004] (Figure 9). The catalytic nucleophile of this enzyme was indentified to be Asp224 by using the mechanism-based covalent inhibitor 2-deoxy-2-fluoro-α-L-fucosyl fluoride and the chemical rescue approach [Tarling et al., 2003].
The structures of a complex enzyme–product and of a covalent glycosyl enzyme intermediate, coupled with kinetic and mutagenesis studies, allowed to identify Glu266 as the acid /base catalyst [Sulzenbacher et al., 2004].

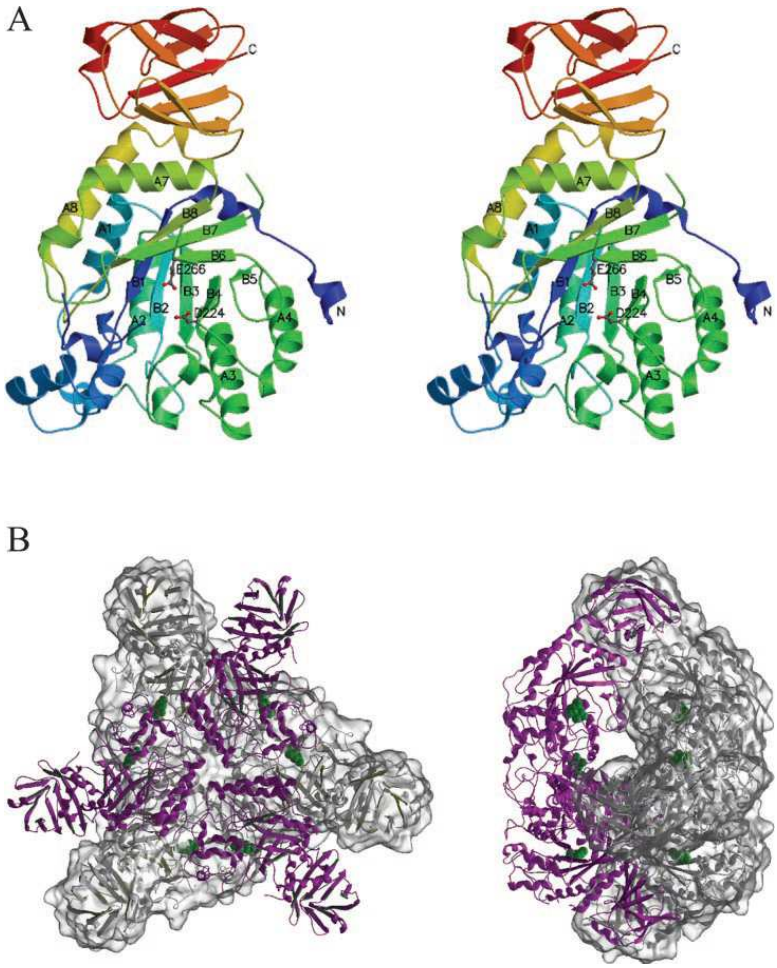


Figure 9: Overall view of TMα-fuc. A) Stereo ribbon diagram and B) the exameric assembly of TMα-fuc as seen from the top (left) and roteed by 90° (right). [Sulzenbacher et al., 2004]

Aim of the thesis

The using of oligosaccharides as target of new drugs is based on their important role in many biological system. The efforts in developing methods for their large-scale production aren't currently available as most of the difficulties arise from the need to perform multiple protecting group manipulations to control regioselectivity in chemical synthesis [Seeberger 2008]. An interesting alternative is enzymatic synthesis and, to this aim, a new class of mutant glycosidases, termed glycosynthases, was developed for oligosaccharide synthesis [Mackenzie et al. 1998; Moracci et al. 1998]. Glycosynthases derive from mutated *retaining* glycosidases, which, in the presence of activated glycosides and suitable reaction conditions, synthesize oligosaccharides without hydrolysing them [Perugino et al. 2005].

I addressed my efforts in studying α-L-fucosidase from *S. solfataricus* (Ssα-fuc). I analyzed the mechanism of expression of this enzyme *in vivo* and in developing a novel α-L-fucosynthase for the quantitative production of fucosylated oligosaccharides using a new approach to reactivate the nucleophile mutants of Ssα-fuc (SsD242S/G).

Once reached this target, it was easily exploited also for the mutant in the nucleophile site (D224) of the *T. maritima* α-L-fucosidase indicating its general applicability, in fact this finding might open new perspectives in the use of new donors for the production of novel α-glycosynthases.

Experimental Procedures

Bacterial Strains

Escherichia coli RB791 *lac I p4000 (lac I ^{q}), lac Z p4008 (lac L8), λ-, IN(rrn D - rrn E)*

Escherichia coli Rosetta DE3 *F^- ompT $hsdS_B$(r_B^- m_B^-) gal dcm (DE3) pRARE (Cam^R)*

Sulfolobus solfataricus **P2**

Culture media

LB (Luria-Bertani Broth):
In 1 liter: 10 g NaCl, 5 g yeast extract, 10 g tryptone.

2YT
In 1 liter: 5 g NaCl, 10 g yeast extract, 16 g tryptone

S. solfataricus **minimal salts medium**
yeast extract (0.1%),
casamino acids (0.1%),
plus carbon source (0.1%)

Chemicals

All commercially available substrates were purchased from Sigma-Aldrich.

Chemical synthesis of 2-chloro-4-nitrophenyl-α-L-fucopyranoside (2C4NPα-L-Fuc)

L-Fucose (700 mg) purchased from CMS Chemicals (Abingdon, UK) was dissolved in 1:1 pyridine/acetic anhydride (8 mL) and stirred overnight at 25℃, after that the volatiles were removed in a rotary evaporator. The residue was diluted with dichloromethane (100 mL) and washed in a separating funnel with 1M HCl (100 mL), then with 1M $NaHCO_3$ (100 mL) and finally with water (100 mL). The organic layer was dried over anhydrous Na_2SO_4, filtered and concentrated to give tetra-acetyl-L-fucose (1.41 g), which was in turn mixed with 2-chloro-4-nitrophenol (1.10 g = 1.5 equivalents). The solid mixture was carefully anhydrified by triple coevaporation with dry toluene (10 mL), then mixed with $FeCl_3$ (755 mg = 1.1 equivalents) under argon atmosphere and suspended at 0℃ in dry dichlorometh ane (25 mL). The mixture was stirred at 0℃ for 4 hours, then diluted with dichl oromethane (200 mL) and washed with 1M $NaHCO_3$ (200 mL). The organic layer was dried over anhydrous Na_2SO_4, filtered and concentrated to give a residue, that after silica-gel column chromatography (6:1:0.01 petroleum ether/ethyl acetate/triethylamine as eluent) afforded crystalline 2-chloro-4-nitrophenyl-2,3,4-triacetyl-α-L-fucopyranoside (470 mg). This compound was dissolved in 4:1:1 methanol/water/triethylamine (18 mL) and the solution was stirred at 25℃ overnight. The n, it was concentrated and the obtained residue was purified by silica-gel column chromatography (99:1:0.01 to 95:5:0.01 dichloromethane/methanol/triethylamine as eluent) to give pure 2C4NP-α-

L-Fuc as yellow crystals (215 mg) (200 MHz 1H-NMR in perdeuterated acetone: δ 8.26 (d, J=2.8 Hz, 1H), 8.20 (dd, J=9.0, 2.8 Hz, 1H), 7.52 (d, J=9.0 Hz, 1H), 5.85 (d, J=3.0 Hz, 1H), 4.06 (m, 2H), 3.99 (dq, J=6.6, 0.9 Hz, 1H), 3.82 (dd, J=2.7, 0.9 Hz, 1H), 1.17 (d, J=6.6 Hz, 3H)).

Chemical synthesis of 2-deoxy-2-fluoro-L-fucosyl fluoride

3,4-Di-*O*-acetyl-L-fucal (193 mg, 0.90 mmol) was dissolved in 1:1 v/v DMF/water (10 mL) and treated with Selectfluor. After stirring at room temperature overnight, AcOEt (130 mL) and water (130 mL) were added to the solution. The organic phase was collected, dried over anhydrous sodium sulphate, filtered and concentrated to give a residue, which was subjected to silica-gel column chromatography (8:1 to 2:1 v/v petroleum ether/ethyl acetate as eluent) to give 3,4-di-*O*-acetyl-2-deoxy-2-fluoro-L-fucose (213 mg, 94%; α/β=1:1) [Burkart, M.D. et al. 1997].
This compound was then dissolved in tetrahydrofuran (7.0 mL) under argon atmosphere and cooled to -30°C. To this solution (diethylamino)sulfur trifluoride (DAST) (136 µL, 1.02 mmol) was added. After few minutes the reaction mixture was heated to room temperature and stirred for 1 hour, after that it was cooled to -30°C again and treated with methanol (2.0 mL) and then concentrated. The residue was dissolved in CH_2Cl_2 (50 mL) and washed with water (50 mL). The organic layer was dried over anhydrous sodium sulphate, filtered and concentrated to give an oily residue that was dissolved in 4:1:1 MeOH/Et_3N/water (6.0 mL). The solution was stirred for 4 hours at room temperature and then concentrated by co-evaporation three times with toluene (5.0 mL). The residue was subjected to silica-gel column chromatography (3:1 to 1:1 v/v petroleum ether/ethyl acetate as eluent) to afford, as first-eluted compound, the desired 2-deoxy-2-fluoro-α-L-fucosyl fluoride (27.8 mg, 19%) as white amorphous crystals. ^{1}H NMR (300 MHz, $(CD_3)_2CO$): δ 5.73 (dd, 1H, $J_{H\text{-}1,F\text{-}1}$=55.2 Hz, $J_{H\text{-}1,H\text{-}2}$=3.0 Hz, H-1), 4.60 (dddd, 1H, $J_{H\text{-}2,F\text{-}2}$=48.9 Hz, $J_{H\text{-}2,F\text{-}1}$=27.6 Hz, $J_{H\text{-}2,H\text{-}3}$=9.6 Hz, $J_{H\text{-}2,H\text{-}1}$=3.0 Hz, H-2), 4.54 (d, 1H, $J_{OH,H\text{-}3}$=8.7 Hz, OH-3), 4.18 (d, 1H, $J_{OH,H\text{-}4}$=3.9 Hz, OH-4), 4.13 (q, 1H, $J_{H\text{-}5,H\text{-}6}$=6.6 Hz, H-5), 4.02 (m, 1H, H-3), 3.86 (q, 1H, $J_{H\text{-}4,H\text{-}3}$=$J_{H\text{-}4,OH\text{-}4}$=3.9 Hz, H-4), 1.24 (d, 3H, $J_{H\text{-}6,H\text{-}5}$=6.6 Hz, H-6); ^{13}C NMR (75 MHz, $(CD_3)_2CO$): δ 106.0 (dd, *J*=223 Hz, *J*=25 Hz), 89.6 (dd, *J*=186 Hz, *J*=25 Hz), 72.9 (d, *J*=2 Hz), 69.9, 68.7 (d, *J*=16 Hz), 16.3.
As second-eluted compound, 2-deoxy-2-fluoro-α-L-fucosyl fluoride (84.5 mg, 58%) was recovered as white amorphous crystals. ^{1}H NMR (300 MHz, $(CD_3)_2CO$): δ 5.29 (ddd, 1H, $J_{H\text{-}1,F\text{-}1}$=54.6 Hz, $J_{H\text{-}1,H\text{-}2}$=7.2 Hz, $J_{H\text{-}1,F\text{-}2}$=4.5 Hz, H-1), 4.60 (d, 1H, $J_{OH,H\text{-}4}$=3.0 Hz, OH-4), 4.39 (dddd, 1H, $J_{H\text{-}2,F\text{-}2}$=53.4 Hz, $J_{H\text{-}2,F\text{-}1}$=16.5 Hz, $J_{H\text{-}2,H\text{-}3}$=9.6 Hz, $J_{H\text{-}2,H\text{-}1}$=7.2 Hz, H-2), 4.20 (d, 1H, $J_{OH,H\text{-}3}$=4.4 Hz, OH-3), 3.92-3.75 (m, 3H, H-3, H-4, H-5), 1.28 (d, 3H, $J_{H\text{-}6,H\text{-}5}$=6.6 Hz, H-6); ^{13}C NMR (75 MHz, $(CD_3)_2CO$): δ 108.1 (dd, *J*=209 Hz, *J*=26 Hz), 92.7 (dd, *J*=179 Hz, *J*=23 Hz), 72.3, 71.8, 63.5, 16.2.
These compounds were synthesized by the group of Prof. M. Parrilli of University of Naples "Federico II".

Site-directed mutagenesis

All mutants present in this thesis are obtained using the GeneTailor Site-directed Mutagenesis System kit (Invitrogen) using as the template plasmids obtained in the laboratory where I performed my thesis [Cobucci Ponzano et al., 2003a].
The mutants of *Thermotoga maritima* α-L-fucosidase are obtained from the template pDEST 17 in which was previously cloned the wild type enzyme [Tarling et al., 2003].
The synthetic oligonucleotides are from PRIMM (Italy), listed in table 2A and B. All mutant genes were sequenced to ascertain the presence of the desired mutation.
This kit relies on the inherent properties of two enzymes, DNA methylase and Mcr BC endonuclease as shown in the workflow diagram below.

The stages of the process are (Figure 10):

1. To methylate plasmid DNA with a methylase;
2. Amplify the plasmid in a mutagenesis reaction with two overlapping primers (forward, that contains the target mutation, and reverse). The product is linear, double stranded DNA containing the mutation.
3. Transform the mutagenesis mixture into wild type *E. coli*. The host cell circularizes the linear mutated DNA, and Mcr BC endonuclease in the host cell digests the methylated template DNA, leaving only unmethilated, mutated product.

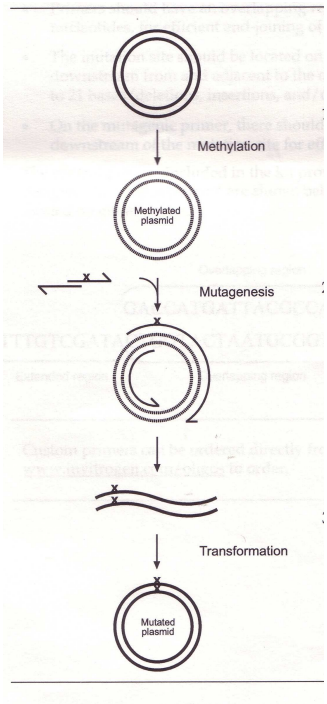


Figure 10: Diagram of mutagenesis kit

The nomenclature used for the different α-fucosidase genes is listed in Table 2A e B.

Regulation of *fuc A1* gene expression

Gene name	Status	Name of recombinant protein	Slippery heptamer	Template	Used Oligonucleotides
*fuc*A1 wild type	-1 frameshifted	-	A-AAA-AAT	-	Cobucci Ponzano et al., 2003 a
*fuc*A1^{A} mutant	Full-lenght	Ssα-fuc	A-AA**G**-AA**T**	pGEX11867/ 3060	Cobucci Ponzano et al., 2003 a
*fuc*A1^{B} mutant	Full-lenght	Ssα-fuc ^{B}	A-AA**G**-AA**G**	pGEX11867/ 3060	FucA1smrev:5'TTTAGGTGATATTGGTGTTCTGGTCTATCT3' Fuc-B: 5'-GAACACCAATATCACCTAAAGAAGTTCGGCCCAGT-3'
*fuc*A1^{sm} mutant	-1 frameshifted	-	A-AA**G**-AAT	pGEX11867/ 3060	FucA1smrev:5'TTTAGGTGATATTGGTGTTCTGGTCTATCT3' FucA1smfwd:5'GAACACCAATATCACCTAAAGAATTCGGCCCA3'
*fuc*A1^{tm} mutant	-1 frameshifted	-	**C**-AA**G**-AA**C**	pGEX11867/ 3060	FucA1tm-rev: 5'-AGGTGATATTGGTGTTCTGGTCTATCTGGC -3' FucA1tmfwd:5'CCAGAACACCAATATCACCTCAAGAACTCGGCCC AGT3'

Table 2A: In site directed mutagenesis experiments, nucleotides modified by substitution and insertion mutations are in red and in green respectively. The mismatched nucleotides are underlined.

.

Glycosynthases

Name of recombinant protein	Template	New Plasmid name	Substitutions	Used Oligonucleotides
SsD242G	pGEX frame-fuc	pGEXFrame fucD242G	D→G GAC→ GGC	D242Gfwd5'GGCCATGGCTAGTCTATTTCGGCTGGTGGATTGC3' D242rev:5'-GAAATAGACTAGCCATGGCCTATACTTTTC-3'
SsD242A	pGEX frame-fuc	pGEXFrame fucD242A	D→A GAC→GGC	D242Afwd:5'GGCCATGGCTAGTCTATTTCGCCTGGTGGATTGC3' D242rev:5'-GAAATAGACTAGCCATGGCCTATACTTTTC-3'
SsD242S	pGEX frame-fuc	pGEXFrame fucD242S	D→S GAC →AGC	D242Sfwd:5'GGCCATGGCTAGTCTATTTCAGCTGGTGGATTGC3' D242rev:5'-GAAATAGACTAGCCATGGCCTATACTTTTC-3'
SsD242G/E58G	pGEXFrame fucD242G	pGEXFrame fucD242G/E58G	E→G GAA→GGA	E58Gfwd: 5'-AGTACCAGCATTTGGTAATGGATGGTACCCTA-3' E58Grev: 5'-CATTACCAAATGCTGGTACTGAGTATACTC-3'
SsD242G/E292G	pGEXFrame fucD242G	pGEXFrame fucD242G/E292G	E→G GAA→GGA	E292Gfwd5'TATACGTTCTTTATTGTTCCTCTTCCAGCTAAATCC3' E292Grev:5'-GGAACAATAAAGAACGTATATCCCTCCACA-3'
TmD224G	pDEST17/wt	pDEST17/D224G	D→G GAC→ GGC	D224Gfwd: 5'TTCCCGACGTTCTCTGGAACGGCATGGGCTGGC3' D224rev: 5'-CCAGGTGGGACTCTCCTTGA-3'
TmD224S	pDEST17/wt	pDEST17/D224S	D→S GAC→TCC	D224Sfwd:5'TTCCCGACGTTCTCTGGAACTCCATGGGCTGGC3' D224rev: 5'-CCAGGTGGGACTCTCCTTGA-3'

Table 2B: Site Directed Mutagenesis Experiments to obtain gycosynthases. The mismatched bases are underlined.

Production in* E.coli *of recombinant α-L-fucosidases

α-L-fucosidases from Sulfolobus solfataricus

The mutants and wild type of α-L-fucosidases from *S. solfataricus* were expressed as fusion proteins with Glutathione-S-transferase (GST). The plasmid containing the mutant gene was used to transform *E. coli* RB791 cells. They were grown in 2 Lt of LB broth at 37℃. Expression of the gene was induce d by the addition of 1 mM IPTG when the culture reached an optical density at 600 nm of 0.6. Growth allowed to proceed for 16 h and cells were harvested by centrifugation at 5,000 x g. The resulting cell pellet was resuspended in PBS 1X (150mM NaCl, 20mM phosphate buffer pH 7.3) with 1% TRYTON X-100 in a ratio1:3 (w:v) and homogenized by French cell pressure treatment. After centrifugation for 30 min at 10,000 x g the crude extract was purified by affinity chromatography on Glutathione Sepharose 4B (Amersham). The active pool is then subjected to thrombin treatment on the resin in order to recover Ssα-fuc mutants free of the fused GST polypeptide.
Thereafter, the soluble GST-free α-fucosidase wild type and mutants were recovered (Figure 11) and a final heating step in a range of 65-80℃ was performed. This procedure is used to remove the residual thrombin from the pure protein solution. This purification procedure was achieved with a matrix that is dedicated only to the purification of each specific mutant, in order to exclude contamination by the wild type enzyme. The enzymes resulted >95% pure by SDS-PAGE. Protein concentration was determined with the method of Bradford [Bradford, 1976], by using bovine serum albumin as standard. The samples stored at 4℃ in PBS buffer were stable for several months.

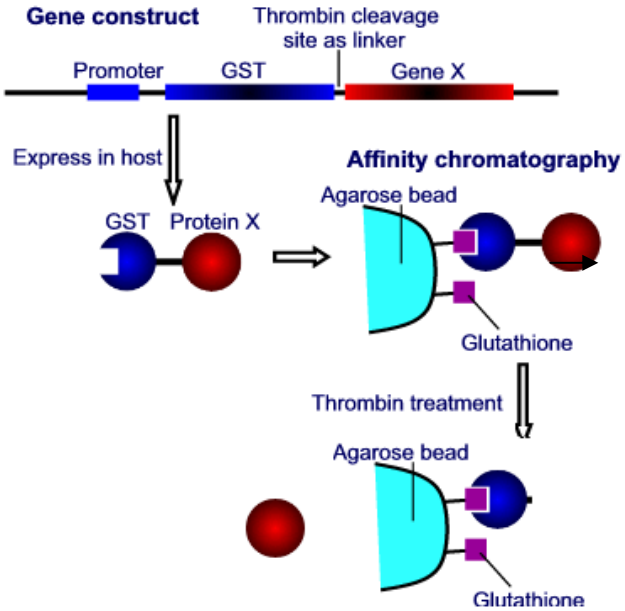


Figure 11: Schematic diagram of GST purification method.

α-L-fucosidases from Thermotoga maritima

The mutants and wild type α-L-fucosidases from *T. maritima* were expressed in the *E. coli* Rosetta DE3 strain grown in 2YT medium containing 50 μg mL^{-1} ampicillin and 30 μg mL^{-1} cloramphenicol at 37℃. Enzyme expression was indu ced with 1 mM IPTG when cell culture reached 0.6 O.D. at 600 nm and cells were harvested after growth at 37℃ for 4 h by centrifugation at 5,000 x g. The purification of the His_6 tagged TmD224G and TmD224S mutants and wild type enzymes was performed by using the Protino Ni-TED 1000 protein purification system (MN). This kit uses an affinity column (containing dry silica-based resin precharged with Ni^{2+} ions) for the binding of His_6-tagged proteins and provides the use of buffers for the lyses of the cells, the equilibration, washing and eluition of the sample from the column. Samples eluted from the column were dialyzed against PBS buffer, as containing 250 mM imidazole. The enzymes, >95% pure by SDS-PAGE, were stable for several months when stored at 4℃.

Kinetic characterization of recombinant α-fucosidases

For all these enzymes, one unit of enzyme activity was defined as the amount of enzyme catalyzing the hydrolysis of 1 μmol of substrate in 1 min at the conditions described. Spontaneous hydrolysis of the substrate was subtracted by using appropriate blank mixtures without enzyme.
All activity data were harvested in UV absorbance with a Varian Cary spectrophotometer equipped, when occurs, with a circulating water bath and they were calculated, plotted and refined with the program *GraFit* [Leatherbarrow 1992].
The coefficient of molar extinction used for all reported assays are listed in table 3.

Conditions	Temperature	λ 4Nitro-phenol / 2C4N-phenol	4Nitro-phenol ε_{mM}	2Cl4Nitro-phenol ε_{mM}
+ Na_2CO_3	r.t.	420 nm	17.2	14
50 mM Phosphate buffer pH 6.5	65℃	405 nm	9.34	17
50 mM Phosphate buffer pH 6.5+ sodium azide	65℃	405 nm	8.0	17
50 mM Phosphate buffer pH 6.5+ sodium formate	65℃	405 nm	8.0	20.6
50 mM Citrate/Phosphate buffer pH 6.0	60℃	400/405 nm	5.3	11.24
50 mM Citrate/Phosphate buffer pH 6.0 + sodium azide	60℃	400/405 nm	4.6	15.65

Table 3: Coefficients of molar extinction for all reported assays.

S. solfataricus *α-fucosidases*

The kinetic constants of the Ssα-fuc^{B} and wild type Ssα-fuc on 4NP-α-L-fuc were measured at 65℃ in 50 mM sodium phosphate buffer pH 6.5, using substrate concentrations in the range 0.005-3mM.
Kinetic constants of wild type Ssαfuc on 2C4NP-α-L-Fuc substrate were measured, at 65℃ in 50 mM sodium phosphate buffer pH 6.5, by us ing substrate concentrations in the range 0.01-0.4 mM. The amount of the enzymes used in the assays was 1 µg. [Cobucci Ponzano et al., 2008b].
Thermal activity of Ssα-fuc^{B} and the wild type were analyzed by assaying the enzyme on 4NP-α-L-fuc substrate concentrations of 1-3 mM in the temperature ranges of 40-65 and 70-90℃, respectively. Thermal stability was tested by incubating pure enzyme in PBS buffer at 80℃ for different times and after by assaying the enzyme in presence of 1 mM 4NP-α-L-fuc at 65℃. The releasing of the 4-nitrophenol is followed at 405 nm. The pH dependence was determined by assaying Ssα-fuc^{B} at 65℃ with 4NP- α-L-fuc substrate at final concentration of 1 mM, in 50 mM of the following buffer systems: sodium citrate (pH 3.0-5.6), sodium phosphate (pH range 6.0-8.0), and glycine/NaOH (pH 8.0-10). The reactions (0.2 mL) were started by adding the enzyme and they were stopped by adding 0.8 mL of iced sodium carbonate 1 M. The optical density of the solution was measured at 420 nm at room temperature.

The chemically rescued activities of SsD242G mutant on 4NP-α-L-Fuc substrate and several acceptors were measured at the indicated conditions. The reactions were incubated for 5-240 min at 65℃ and the releasing o f the 4-nitrophenol is followed at 405 nm at 65℃. Steady-state kinetic parameters of SsD242G mutant on 4NPα-L-Fuc and on 2C4NP-α-L-Fuc substrate were measured by using substrates concentrations ranging from 0.01 to 4 mM. The reactions (0.2 mL) were started by adding the enzyme and they were stopped by adding 0.8 mL of iced sodium carbonate 1 M. The optical density of the solution was measured at 420 nm at room temperature. The amount of the enzyme used in the assays was 10 µg.

The chemically rescued activities of SsD242S mutant on 4NP-α-L-Fuc and 2C4NP-α-L-Fuc substrates were measured at the indicated conditions in the presence of sodium formate (0.1-2.0 M) or sodium azide (0.1-2.0 M). The releasing of the 4-nitrophenol is followed at 405 nm at 65℃.
The kinetic parameters of the SsD242S mutant were measured in 50 mM sodium phosphate buffer (pH 6.5) at 65℃, in 2 M sodium az ide on 4NP-α-L-Fuc and 2C4NP-α-L-Fuc substrates used at concentrations ranging from 0.02 to 4 mM. The amount of the enzyme used in the assays was 3 µg.The reactions (0.2 mL) was started by adding the enzyme and, after 1 min of incubation at 65℃, it was stopped by adding 0.8 mL of iced sodium carbonate 1 M. Successively it was measured the absorbance value at 420 nm at room temperature.

T.maritima *α-fucosidases*

The activity of the TmD224G mutant was measured on 4NP-α-L-Fuc and 2C4NP-α-L-Fuc substrates at the concentrations shown for Ssα-fuc enzymes in 50 mM sodium citrate/phosphate buffer (pH 6.0). The chemically rescued activities of TmD224G mutant on 4NP-α-L-Fuc and 2C4NP-α-L-Fuc substrates in the presence of sodium azide were measured at the indicated conditions. Steady-state kinetic parameters of wild type Tmα-fuc and TmD224G mutant on 4NP-α-L-Fuc and 2C4NP-α-L-Fuc substrates were measured in 50 mM sodium citrate/phosphate buffer (pH 6.0), on 4NP-α-L-Fuc and 2C4NP-α-L-Fuc substrates used at concentrations ranging from 0.005 to 0.8 mM, and, when is indicated, in the presence of 1 M sodium azide. The amount of the enzyme used in the assays was 5 µg. The releasing of the leaving group 4-nitrophenol or 2Cl-4nitrophenol is followed at 400 or 405 nm respectively for all assays at 60℃.

Regulation of fucA1 expression

Western blot

For the Western blot studies, equal amounts of *E. coli* cultures expressing the wild type and mutant *fucA1* genes, normalized for the OD 600 nm, were resuspended in SDS-PAGE loading buffer containing 0.03 M Tris HCl buffer pH 6.8, 3% SDS (w/v), 6.7% glycerol (w/v), 6.7% 2-mercaptoethanol (w/v), and 0.002% blue bromophenol (w/v). The samples were incubated at 100℃ for 5 min (unless otherwise indicated) and were directly loaded onto the gel. Western blot analyses were performed by blotting SDS-PAGEs of the concentrations indicated on Hybond-P polyvinylidenfluorid filters (Amersham Biosciences, Uppsala, Sweden); polyclonal anti-Ssα-fuc antibodies from rabbit (PRIMM, Milan, Italy) and anti-GST antibodies (Amersham Biosciences) were diluted 1:5,000 and 1:40,000, respectively. The filters were washed and incubated with the ImmunoPure anti-rabbit IgG antibody conjugated with the horseradish peroxidase from Pierce Biotechnology (Rockford, IL, USA). Filters were developed with the ECL-plus Western Blotting Detection system (Amersham Biosciences) by following the indications of the manufacturer. The molecular weight markers used in the Western blot analyses were the ECL Streptavidin-HRP conjugate (Amersham Biosciences).
The protein concentration of the samples was measured with the method of Bradford [Bradford, 1976] and the amounts of sample loaded onto the SDS-PAGEs are those indicated. The quantification of the bands identified by Western blot was performed by using the program Quantity One 4.4.0 in a ChemiDoc EQ System (BioRad, Hercules, CA, USA) with the volume analysis tool. The frameshifting efficiency was calculated as the ratio of the intensity of the bands of the frameshifted product / frameshifted product + termination product.

Mass spectrometry analysis for the identification of site -1 frameshifting

Samples of the proteins expressed in *E. coli* from the wild type gene *fucA1* and the mutants *fucA1*^{A} and *fucA1*^{sm}, purified as described, were fractionated on a SDS-PAGE. Protein bands were excised from the gel, washed in 50 mM ammonium bicarbonate pH 8.0 in 50% acetonitrile, reduced with 10 mM DTT at 56℃ for 45 min and alkylated with 55 mM iodoacetamide for 30 min at room temperature in the dark. The gel pieces were washed several times with the buffer, resuspended in 50 mM ammonium bicarbonate and incubated with 100 ng of trypsin for 2 hr at 4℃ and overnight at 37℃. The supernatant containing pepti des was analysed by MALDIMS on an Applied Biosystem Voyager DE-PRO mass spectrometer using α-cyano-4-hydroxycynnamic acid as matrix. Mass calibration was performed by using the standard mixture provided by manufacturer.
Liquid chromatography on-line tandem mass spectrometry (LCMSMS) analyses were performed on a Q-TOF hybrid mass spectrometer (Micromass, Waters, Milford, Massachusetts USA) coupled with a CapLC capillary chromatographic system (Waters). Peptide ions were selected in the collision cell and fragmented. ProteinLynx software, provided by the manufacturers, was used to analyze raw MS and MS/MS spectra and to generate a peak list, which was introduced in the MASCOT software for protein identification.
Mass spectrometry analyses for the identification of site -1 frameshifting were performed by the group of Prof. Piero Pucci of the University of Naples "Federico II"

Experiments of Translation in vitro

Genomic DNA from *S. solfataricus* P2 strain was prepared by Nucleospin Tissue kit (MN, Germany). A DNA fragment of 1,538 nucleotides containing the complete *fucA1* gene, was prepared by polymerase chain reaction (PCR), by using the following synthetic oligonucleotides (Genenco, Florence, Italy): FucA1-fwd and FucA1-rev (listed in table 4). The PCR amplification was performed as described previously [Cobucci Ponzano et al., 2003a] and the amplification products were cloned in the *BssHII* site of the plasmid pBluescript II KS+. The *fucA1* gene was completely re-sequenced to check if undesired mutations were introduced by PCR and the recombinant vector obtained, named pBlu-FucA1, was used for translation *in vitro* experiments.
The plasmids expressing the mutant genes *fucA1*^{A}, *fucA1*^{sm}, and *fucA1*^{tm} for experiments of translation *in vitro* were prepared by substituting the *KpnI-NcoI* wild type fragment, containing the slippery site, with those isolated from the mutants. To check that the resulting plasmids had the correct sequence, the mutant genes were completely re-sequenced.
The mRNAs encoding wild type *fucA1* and its various mutants were obtained by *in vitro* run-off transcription. About 2 µg of each plasmid were linearized with *BssHII* and incubated with 50 units of T7 RNA polymerase for 1h 30 min at 37℃. The transcription mixtures were then treated with 10 units of *DNAseI* (*RNAse* free) for 30 min. The transcribed RNAs were recovered by extracting the samples twice with phenol (pH 4.7) and once with phenol/chloroform 1:1 followed by precipitation with

ethanol. The mRNAs were resuspended in DEPC-treated H_2O at the approximate concentration of 0.6 pmol/µL.
In vitro translation assays were performed essentially as described by Condò et al (1999). The samples (25 µL final volume) contained 5 µL of *S. solfataricus* cell extract, 10 mM KCl, 20 mM Tris/HCl pH 7.0, 20 mM Mg acetate, 3 mM ATP, 1 mM GTP, 5 µg of bulk *S. solfataricus* tRNA, 2 µL of [^{35}S]-methionine (1,200 Ci/mmol at 10 mCi/mL) and approx. 10 pmol of each mRNA. The mixtures were incubated at 70°C for 45 min. After this time, the synthesized proteins were resolved by electrophoresis 12.5 % acrylamide-SDS gels and revealed by autoradiography of the dried gels on an Instant Imager apparatus.
Experiments of Translation in vitro were performed by the group of Prof. Paola Londei of University of Rome "Sapienza".

Trascriptional Analysis of Fuc A1

Cells of *S. solfataricus*, strain P2, were grown in minimal salts culture media supplemented with yeast extract (0.1%), casamino acids (0.1%), plus glucose (0.1%) (YGM) or sucrose (0.1%) (YSM). The extraction of total RNA was performed from *S. solfataricus P2* strain. Cells are lysated by 3 cycles of freeze-thawing (2 min at -70°C; 2 min 37°C) and total RNA was extracted by RNeasy kit (Qiagen, Germany). The Reverse Transcriptase (RT)-PCR experiments was performed by using Titan one tube RT-PCR system (Roche Molecular Biochemicals) and the primers α-fuc-3 and α-fuc-833 (listed in table 4), that allowed the amplification of a region of 833 nucleotides (positions 1-833, in which the A of the first ATG codon is numbered as one) overlapping the ORFs SSO11867 and SSO3060. The PCR program used is as follows: cDNA synthesis and pre-denaturation for 31 min at 50°C and 2 min at 94°C; amplification by 40 cycles of 15 sec at 94°C, 30 sec at 45°C and 2 min at 72°C; final extension of 10 min at 72°C. The cDNA products obtained were directly sequenced with no further purification steps.
For Real Time experiments, total RNA was extensively digested with DNAse (Ambion, Austin, TX, USA) and the absence of DNA was assessed by the lack of PCR amplification with each sets of primers described below. Total cDNA was obtained using the kit Quantitect RT (Qiagen GmbH, Hilden, Germany) from 500 ng of total RNA. cDNA was then amplified in a BioRad LightCycler using the DyNAmo HS Syber Green qPCR Kit (Finnzymes Oy, Espoo, Finland). Synthetic oligonucleotides (PRIMM) used for amplification of a region at the 3' of the ORF SSO3060 were 5'-Real and 3'-Real (table 4) for the gene *fucA1*. In the table 4 are listed also the oligonucleotides used for the 16S rRNA gene (16S1 and 16S2).
Each amplification of the *fucA1* gene contained about 2,500-fold more cDNA than that of the 16S rRNA. Controls with no template cDNA were always included. PCR conditions were: 15 min at 95 °C for initial denaturation, followed by 40 cycles of 10 s at 95 °C, 25 s at 56 °C and 35 s at 72 °C, and a final step of 10 min at 72 °C. Product purity was controlled by melting point analysis of setpoints with 0.5°C temperature increase from 72 to 95°C. PCR products were analysed on 2% agarose gels and visualized by ethidium bromide staining.
The expression values of *fucA1* gene were normalized to the values determined for the 16S rRNA gene. Absolute expression levels were calculated as *fucA1*/16S ratio in YSM and YGM cells, respectively. Relative mRNA expression levels (YSM/YGM

ratio) were calculated as (*fucA1*/16S ratio in YGM cells)/(*fucA1*/16S ratio in YSM). Each cDNA was used in triplicate for each amplification.

Experiments	Oligonucleotides
Translation in vitro	FucA1fwd:5'CTGGAG*GCGCGC***TAATACGACTCACTATAGG**TCAGT TAAATGTCACAAAATTCT-3' FucA1rev:5'GACTTG*GCGCGC*CTATCTATAATCTAGGATAACCCTT AT-3'
RT-PCR	α-fuc-3: 5'-GAGGAAGATCTATGTCACAAAATTCTTACAAAATC-3'; α-fuc-833: 5'-TTGCTTGTAAATTATTACGGG-3'
Real-Time PCR	5'-Real: 5'-TAAATGGCGAAGCGATTTTC-3'; 3'-Real: 5'-ATATGCCTTTGTCGCGGATA-3' 16S1: 5'-GAATGGGGGTGATACTGTCG-3'; 16S2: 5'-TTTACAGCCGGGACTACAGG-3' .

Table 4: Oligonucleotides used in Trascriptional Analysis of Fuc A1 and in Experiments of Translation in vitro. For Translation in vitro experiments, the sequence corresponding to the genome *of S. solfataricus* is underlined. In the FucA1-fwd primer, the sequence of the promoter of the T7 RNA polymerase is in bold and the sequence of the *BssHII* site is shown in italics

Preparation of Fucosynthase

Characterization of Fucosynthase

To purify and characterize the glycosynthetic product, the mutant SsD242G (50 µg) was incubated for 16 h at 65℃ in 0.8 mL of 50 mM s odium phosphate buffer, pH 6.5, in the presence of 0.1 M sodium azide, and 20 mM 2C4NP-α-L-Fuc.

The transfucosylation products of mutant SsD242S from β-L-Fuc-N_3 donor to different acceptors were prepared by incubating 94 µg of enzyme for 16 h at 65℃ in 0.8 mL of 50 mM sodium phosphate buffer, pH 6.5, by using different *donor: acceptor* molar ratios (see Results for details).
The transfucosylation products of mutant TmD242G were prepared by incubating 38 µg of enzyme for 16 h at 70℃ in 0.2 mL of 50 mM sod ium phosphate buffer (pH 6.5) by using 10 mM β-L-Fuc-N_3 donor and 100 mM 4NP-β-D-Xyl (1:10 molar ratio). Blank mixtures without enzyme were also prepared.

Detection of transfucosylation activity by Thin Layer Chromatografy (TLC)

A volume equal to 10-20 µL of the reactions mixtures of enzyme mutants, after the time of incubation, was loaded and separated on a 20x20 silica gel 60 F_{254} TLC using ethyl acetate-methanol-water (70:20:10) as eluent. Compounds were detected with UV light to determine the presence of UV-visible compounds, and by exposure to 4%

α-naphtol in 10% sulfuric acid in ethanol followed by heating (100℃) for the time necessary to signal the emergence of sugar.

Analysis of transfucosylation efficiency of Ssα-fuc and Tmα-fuc mutants

Analysis of transfucosylation efficiency of Ssα-fuc and Tmα-fuc mutants, was performed by use of a High-Performance Anion-Exchange Chromatography with Pulsed Amperometric Detection (HPAEC–PAD) equipped with a PA1 column (Dionex, USA). The reaction mixtures and the blank mixtures described above were diluted (10 to 100-fold) with H_2O and 0.5 nmoles arabinose as internal standard, loaded onto the PA1 column and eluted with 16 mM NaOH. The moles of fucose were determined by integration of the peaks within the chromatograph, based on fucose and arabinose standard curves. The amount of fucose transferred by the enzyme to water was calculated by subtracting the amount of free fucose measured in the blank mixtures from that identified in the reaction mixtures.

To measure the total amount of fucose enzymatically transferred, 1/10 of the reaction mixtures were incubated for 90 min at 65℃ in the presence of 1.2 µg Ssα-fuc wild type. Successively, the solution was treated as described above and run by HPAEC–PAD to measure the total amount of fucose. The efficiency of the transfucosylation reaction was calculated as: total amount of fucose transferred - moles of fucose transferred to water / total amount of fucose transferred x 100.

Labeling, Proteolysis and LC/MS/MS analysis in presence of inhibitor 2-deoxy-2-fluoro-L-fucosyl fluoride

Labelling of Ssα-fuc and SsD242S were performed by incubating 60 µg (1.1 nmol) of each enzyme at 65℃ in sodium phosphate buffer 50 mM (pH 6.5) with 0.8 µg (4.7 nmol) of 2d-2F-Fuc-F for 2 h. After this time, digestion buffer, pH 2.0, containing 0.6 µg of pepsin was added to each reaction mixture; pepsin digestions proceeded for 1 h at room temperature, and the reactions were stopped by injection in an LC/MS/MS system. Mass spectrometric analyses were performed on a system constituted by a nano Acquity UPLC module and a Q-TOF premiere spectrometer, equipped with a nanospray ion source and provided by a lock-mass apparatus to perform a real-time calibration correction mode. Chromatographic separations of peptides mixtures were achieved using a C_{18} Proteo (2.1 x 150 mm) column (Phenomenex, Torrance, CA, USA) using a 60 minutes linear gradient from 2% to 60% of 1% formic acid in CH_3CN, flow rate 250 nl/min. Mass spectra were acquired over the *m/z* range 400-2000, and dependent scan collision induced fragmentations of the most intense ions were performed. Mass calibration was performed using a mixture of angiotensin and insulin as external standard, and [Glu]-Fibrinopeptide B human as lock mass standard.

To identify the modified peptide(s), MS/MS data were elaborated scanning all those peptide undergoing a neutral loss of 149.06 mass unit during the fragmentation process.

Analysis of inhibition with Mass spectrometry was carried out by Dott. Fabrizio Del Piaz of the University of Salerno.

Analysis of glycosynthetic reaction products

The reaction mixtures were separately freeze dried and the purification of each sample was obtained by reverse phase chromatography (Polar-RP 80A, Phenomenex, 4 µ, 250 x 10 mm) on a Agilent HPLC instrument 1100 series, using H_2O/CH_3OH 6/4 as eluant. The eluted products were first analysed by positive ions reflectron MALDI-TOF mass spectrometry.

The disaccharide **2** (Table…)showed a pseudomolecular ion $(M+Na)^+$ at 358.15 *m/z* (calculated mass 358.12 u) and the methylation analysis indicated the presence of terminal fucose and 3-substituted fucose. ^{1}H NMR (600 MHz, D_2O): δ 5.04 (d, 1H, $J_{H-1,H-2}$=3.7 Hz, H-1_B), 4.69 (d, 1H, $J_{H-1,H-2}$=8.7 Hz, H-1_A), 4.30 (q, 1H, $J_{H-5,H-6}$=6.6 Hz, H-5_B), 3.99 (d, 1H, $J_{H-4,H-3}$=3.1 Hz, H-4_A), 3.93 (dd, 1H, $J_{H-3,H-2}$=10.3 Hz, $J_{H-3,H-4}$=3.1 Hz, H-3_B), 3.87 (q, 1H, $J_{H-5,H-6}$=6.4 Hz, H-5_A), 3.82 (d, 1H, $J_{H-4,H-3}$=3.1 Hz, H-4_B), 3.79 (dd, 1H, $J_{H-2,H-3}$=10.3 Hz, $J_{H-2,H-1}$=3.7 Hz, H-2_B), 3.70 (dd, 1H, $J_{H-2,H-3}$=9.5 Hz, $J_{H-3,H-4}$=3.1 Hz, H-3_A), 3.58 (t, 1H, $J_{H-2,H-3}$=9.5 Hz, H-2_A), 1.27 (d, 3H, $J_{H-6,H-5}$=6.4 Hz, H-6_A), 1.19 (d, 3H, $J_{H-6,H-5}$=6.6 Hz, H-6_B); ^{13}C NMR (150 MHz, D_2O): δ 96.7 (C-1_B), 91.4 (C-1_A), 78.8 (C-3_A), 72.7 (C-5_A), 72.6 (C-2_A), 72.9 (C-4_B), 70.2 (C-3_B), 68.9 (C-2_B), 68.6 (C-4_A), 67.8 (C-5_B), 16.2 (C-6_B), 16.0 (C-6_A).

Product **3** (Table…): MALDI-TOF-MS $(M+Na)^+$ 440.93 *m/z* (calculated mass 440.13 u). Methylation analysis: terminal-fucose and 3-substituted xylose; ^{1}H-NMR: $^3J_{H1,H2}$ Xyl = 7.6 Hz, $^3J_{H1,H2}$ Fuc = 3.3 Hz.

Product **4** (Table…): MALDI-TOF-MS: $(M+Na)^+$ 439.90 *m/z* (calculated mass 440.13 u). Methylation analysis: terminal-fucose and 4-substituted xylose; ^{1}H-NMR: $^3J_{H1,H2}$ Xyl = 7.3 Hz, $^3J_{H1,H2}$ Fuc = 3.5 Hz.

Product **5** (Table…): MALDI-TOF-MS: $(M+Na)^+$ 585.90 *m/z* (calculated mass 586.18 u). Methylation analysis: terminal-fucose and 3,4-substituted xylose. ^{1}H-NMR: $^3J_{H1,H2}$ Xyl = 6.2 Hz, $^3J_{H1,H2}$ Fuc = 2.4 Hz.

The disaccharides **6**, **7** and **8** (Table…) showed a pseudomolecular ion $(M+Na)^+$ at 469.86 *m/z* (calculated mass 470.13 u), whereas that of trisaccharide **9** (Table…) was 615.96 *m/z* (calculated mass 616.19 u). The methylation analysis showed the presence of 6-substituted, 4-substituted, 3-substituted and 2,3-disubstituted galactose for **6**, **7**, **8** and **9** (Table…), respectively. In addition all the oligosaccharides showed the presence of terminal fucose. The anomeric configuration for all the monosaccharides was deduced by ^{1}H-NMR analysis of each product by measuring $^3J_{H1,H2}$ coupling constants: $^3J_{H1,H2}$ Gal = 7.4 Hz, $^3J_{H1,H2}$ Fuc = 3.5 Hz.

Product **10** (Table…): MALDI-TOF-MS: $(M+Na)^+$ 511.06 *m/z* (calculated mass 511.16 u). Methylation analysis: terminal-fucose and 3-substituted *N*-acetyl-glucosamine. ^{1}H-NMR: $^3J_{H1,H2}$ GlcNAc = 8.3 Hz, $^3J_{H1,H2}$ Fuc = 3.3 Hz.

These analyses were performed by the group of Prof. M. Michela Corsaro of the University of Naples "Federico II".

Results

***Translation mechanism of fucA1 gene in* S.solfataricus**

***Expression of fucA1 in* E. coli**

The analysis of the genome of the hyperthermophilic archaeon *S. solfataricus* [She et al. 2001] revealed that *fucA1* is expressed from two open reading frames (ORFs), SSO11867 and SSO3060, encoding for 81 and 426 amino acid polypeptides. It was previously reported that the region of overlap between the two ORFs contained signals putatively regulating the expression of the gene by programmed -1 frameshifting (Figure 12).

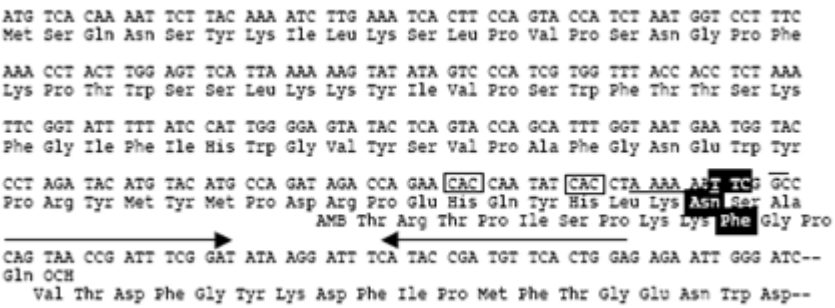


Figure 12: The α-fucosidase gene. Region of overlap in the wild type split *fucA1* gene. The N-terminal SSO11867 ORF is in the zero frame, the C-terminal SSO3060 ORF, for which only a fragment is shown, is in the –1 frame. The signals putatively regulating the programmed -1 frameshifting are: the slippery heptameric sequence (underlined), the rare codons (boxed) and the stems of the putative mRNA secondary structure arrows. The amino acids involved in the programmed –1 frameshifting and the first codon translated after this event in the –1 frame are shadowed [Cobucci Ponzano et al., 2006].

To analyze the expression of the wild type *fucA1* gene in *E.coli*, it was cloned in pGEX vector [Cobucci Ponzano et al., 2003a] and expressed in *E. coli* as a Glutathione S-Transferase (GST)-fused protein. The gene produced trace amounts of α-fucosidase activity (2.3×10^{-2} units mg^{-1} after removal of GST), suggesting that a programmed -1 frameshifting may occur also in *E. coli*. The enzyme was purified from a culture of 4 lt of transformed *E.coli* RB791/pGEX11867/3060 cells by using the GST purification system and SDS-PAGE revealed a major protein band (Figure 13).

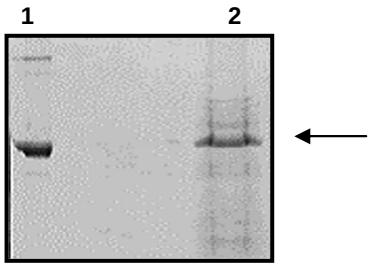


Figure 13: Coomassie stained 7% SDS-PAGE showing (line 1) the recombinant Ssα-fuc (3 μg) and (line 2) the purified products of the wild type split *fucA1* gene, indicated by arrow [Cobucci Ponzano et al., 2006].

The sample and control bands were excised from the gel, digested *in situ* with trypsin and directly analysed by matrix-assisted laser desorption/ionization mass spectrometry (MALDIMS). As shown in Figure 14A and B both spectra revealed the occurrence of an identical mass signal at m/z 1244.6 (arrows) corresponding to a peptide encompassing the overlapping region of the two ORFs. This result was confirmed by liquid chromatography on-line tandem mass spectrometry (LCMSMS) analysis of the peptide mixtures. The fragmentation spectra of the two signals showed the common sequence Asn-Phe-Gly-Pro-Val-Thr-Asp-Phe-Gly-Tyr-Lys (Peptide A) in which the amino acid from the ORF SSO11867 is underlined. These results unequivocally demonstrate that the protein containing the Peptide A is produced in *E. coli* by a frameshifting event that occurred exactly within the slippery heptamer predicted from the analysis of the DNA sequence in the region of overlap between the ORFs SSO11867 and SSO3060.

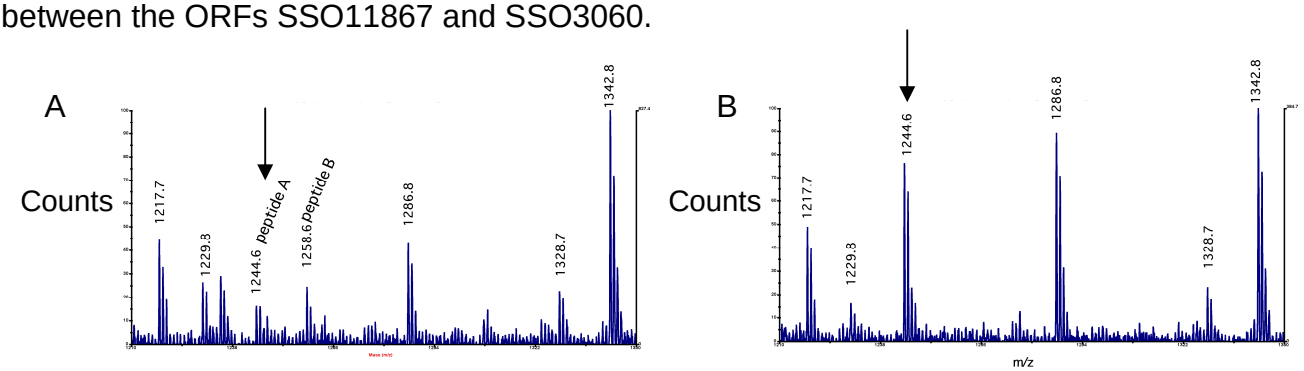


Figure 14: MALDIMS of the purified products of the wild type *fucA1* gene and of Ssα-fuc are shown in (A) and (B), respectively [Cobucci Ponzano et al., 2006].

Remarkably, the MALDIMS analysis of the products of the wild type *fucA1* gene revealed the presence of a second Peptide B at m/z 1258.6 that is absent in the spectra of the Ssα-fuc control protein (Figure 14A and B). The sequence of Peptide B obtained by LCMSMS (Figure 15 A) was Lys-Phe-Gly-Pro-Val-Thr-Asp-Phe-Gly-Tyr-Lys. This sequence differs only by one amino acid (boxed in figure 15B) from Peptide A demonstrating that the interrupted gene *fucA1* expresses in *E. coli* two full-length proteins originated by different -1 frameshifting events. Polypeptide A results from a shift in a site A and it is identical to Ssα-fuc prepared by site directed mutagenesis [Cobucci Ponzano et al., 2003a] while polypeptide B, named Ssα-fuc^{B}, is generated by frameshifting in a second site B (Figure 15).

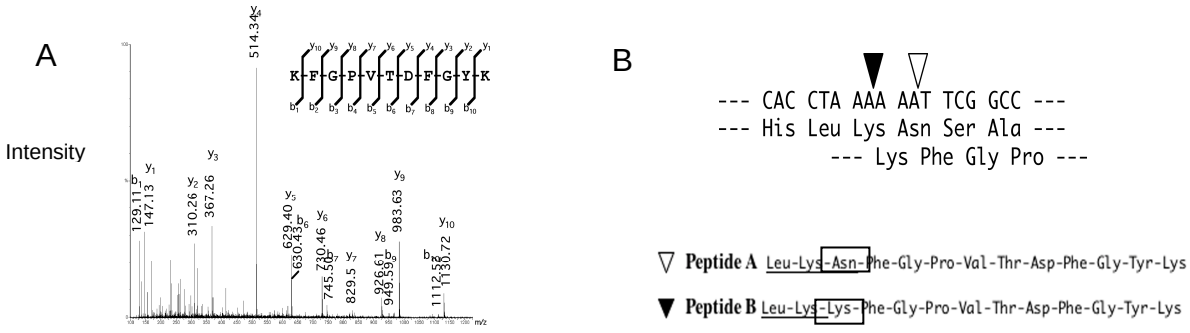


Figure 15: (A) LCMSMS analysis of Peptide B. (B) The proposed frameshifting sites in the *fucA1* gene. The open and the closed arrows indicate the shifting sites A and B, respectively (see the text for details) .

To measure the global efficiency of frameshifting in the two sites of the wild type gene *fucA1,* I analysed the total extracts of *E. coli* by Western blot using anti-GST antibodies (Figure 16A). Two bands with marked different electrophoretic mobility were observed: the polypeptide of 78.7±1.1 kDa migrated like GST-Ssα-fuc fusion and was identified as originated from frameshifting in either site A or B of *fucA1*. The protein of 38.1±1.2 kDa, which is not expressed by the mutant gene *fucA1*^{A} (not shown), had a electrophoretic mobility compatible with GST fused to the polypeptide encoded by the ORF SSO11867 solely (27 kDa and 9.6 kDa, respectively). This polypeptide originated from the translational termination of the ribosome at the OCH codon of the *fucA1* N-terminal ORF (Figure 16B). The calculated ratio of frameshifting to the termination products was 5%, indicating that frameshifting occurs very efficiently *in vivo*.

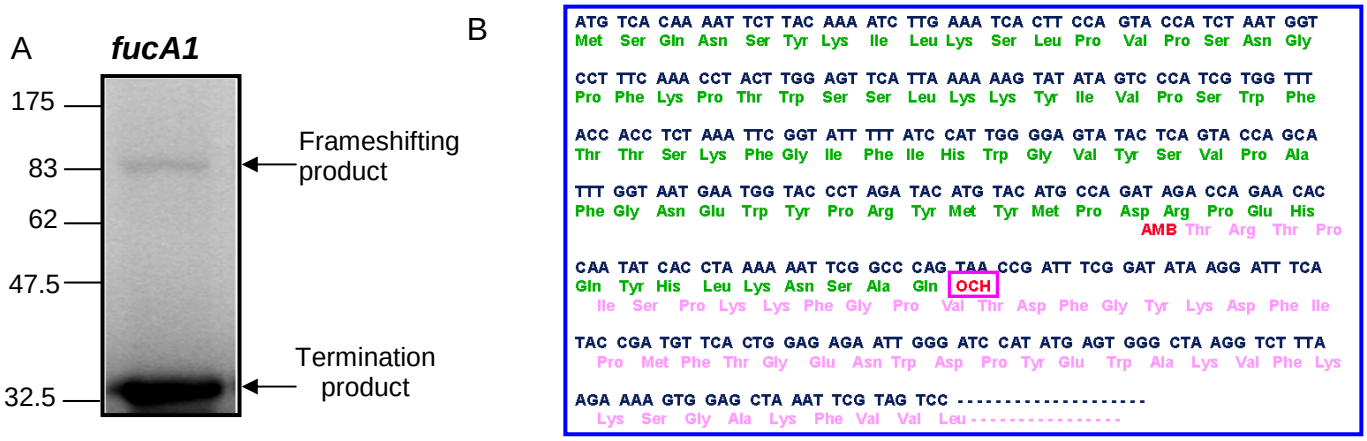


Figure 16: Western blot of *E. coli* cellular extracts expressing the wild type *fucA1* gene. The blot was probed with anti-GST antibodies. The pre-stained molecular weight markers were: β-galactosidase (175,000), paramyosin (83,000), glutamic dehydrogenase (62,000), aldolase (47,500), and triosephosphate isomerase (32,500) [Cobucci Ponzano et al., 2006]. B) Overlapping sequence of *fucA1* gene: the OCH stop codon is boxed .

Preparation and characterization of Ssα-fuc^B

To test if the full-length α-fucosidase produced by the -1 frameshifting event in site B is active (Ssα-fuc^B), I prepared the enzyme by site-directed mutagenesis. The slippery sequence in *fucA1* A-AAA-AAT was mutated in A-AA<u>G</u>-AA<u>G</u>-T where mutations are underlined. The new mutant gene was named *fucA1^B*. The first G, producing the conservative mutation AAA → AAG, was made to disrupt the slippery sequence and hence reducing the shifting efficiency.
The second G was inserted to produce *in vitro* the frameshifting observed in the amino acid sequence of Peptide B. Therefore, the sequence of the two full-length mutant genes *fucA1^A* and *fucA1^B* differs only in the region of the slippery sequence (italics): *A-AA<u>G</u>-AAT*-T<u>T</u>C-GGC and *A-AA<u>G</u>-AA<u>G</u>*-TTC-GGC, respectively (the mutations are underlined (Table 5).

Gene name	Status	Name of recombinant protein	Slippery heptamer
*fuc*A1 wild type	-1 frameshifted	-	A-AAA-AAT
*fuc*A1^A mutant	Full-lenght	Ssα-fuc	A-AA**G**-AA<u>T</u>
*fuc*A1^B mutant	Full-lenght	Ssα-fuc ^{B}	A-AA**G**-AA<u>G</u>

Table 5: Nomenclature and characteristics of the α-fucosidase gene. Nucleotides modified by substitution and insertion mutations are boldface and underlined respectively.

The recombinant Ssα-fuc^B was purified up to about 95%. Gel filtration chromatography demonstrated that in native conditions Ssα-fuc^B had the same nonameric structure of Ssα-fuc with an identical molecular weight of 508 kDa (not shown). In addition, Ssα-fuc^B had the same high substrate selectivity of Ssα-fuc. The two enzymes have high affinity for 4NPα-L-Fuc substrate at 65°C; the K_M is identical within the experimental error (0.0287+/-0.005 vs 0.0287+/-0.004 of Ssα-fuc^B and the wild type respectively) while the k_{cat} of Ssα-fuc^B is about 48% of that of Ssα-fuc (137+/-5.7 vs 287+/-11 of Ssα-fuc^B and the wild type respectively). In addition, 4-nitrophenyl-α-L-arabinoside, -rhamnoside, 4-nitrophenyl-α-D-glucoside, -xyloside, -galactoside, and -mannoside were not substrate of Ssα-fuc^B as previously shown for Ssα-fuc [Cobucci Ponzano et al., 2003a]. This suggests that the different amino acid sequence did not significantly affect the active site.
Both enzymes showed an identical profile of specific activity *vs.* temperature with an optimal temperature higher than 95°C (not shown). At 80°C, the optimal growth temperature of *S. solfataricus*, the half life of Ssα-fuc^B is 45 min, almost 4-fold lower than that of Ssα-fuc. The two enzymes showed different behaviour at pH below 6.0 at which Ssα-fuc^B is only barely active and stable; however, the two enzymes showed similar values of specific activity at pHs above 6.0, which is close to the intracellular pH of *S. solfataricus.*

Characterization of the slippery sequence of *fucA1* in *E. coli*

The experimental data reported above indicate that the predicted slippery heptanucleotide in the region of overlap between the ORFs SSO11867 and SSO3060 of the wild type gene *fucA1* could regulate in *cis* the frameshifting events

observed in *E. coli.* To test this hypothesis, I mutated the sequence A-AAA-AAT into A-AAG-AAT and C-AAG-AAC (mutations are underlined) obtaining the *fucA1* single mutant (*fucA1*^{sm}) and triple mutant (*fucA1*^{tm}) genes, respectively. It is worth noting that the mutations disrupt the slippery sequence, but they maintain the -1 frameshift between the two ORFs (Figure 17).

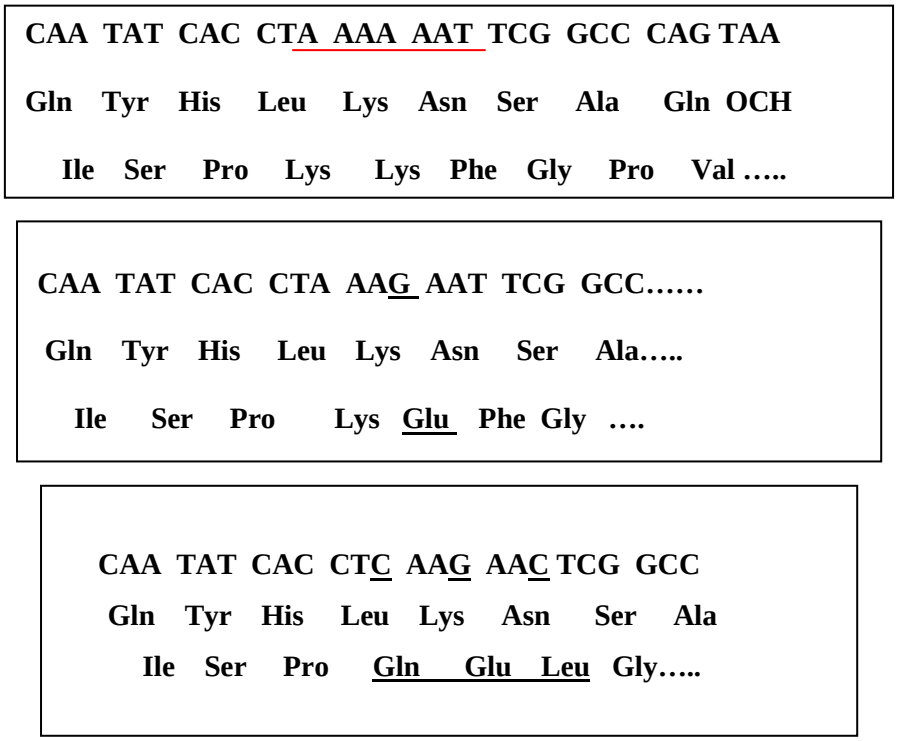


Figure 17: In panel A is indicated the sequence of gene the *fucA1 wild type* with the slippery sequence (underlined). In the panel B and C instead are underlined the mutation of fucA1^{sm} and fucA1^{tm} respectively.

Surprisingly, the expression of *fucA1*^{sm} in *E. coli* produced a full-length polypeptide that, after purification by affinity chromatography and removal of the GST protein, showed the same electrophoretic migration of Ssα-fuc and Ssα-fuc^{B} (Figure 18).

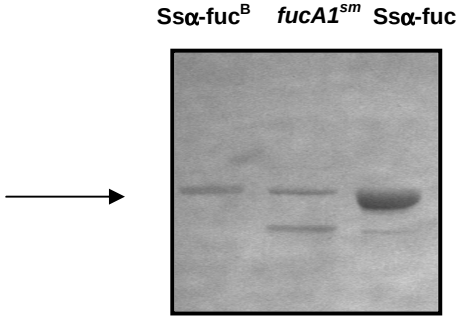


Figure 18: Analysis of the expression in *E. coli* of the mutants in the slippery sequence. Coomassie stained 7% SDS-PAGE showing the purified recombinant Ssα-fuc^{B} (1.2 µg), the product of the gene *fucA1*^{sm} (2 µg), and Ssα-fuc (4 µg). The bands with faster electrophoretic mobility result from the proteolytic cleavage of the full-length protein [Cobucci Ponzano et al., 2003a; 2006].

This protein was then characterised by mass spectrometry analyses following *in situ* tryptic digestion. Interestingly, the MALDI spectra revealed the presence of a single peptide encompassing the overlapping region between the two ORFs with a mass value of 1259.7 Da (Peptide C, Figure 19 A). The sequence of Peptide C, determined from the fragmentation spectra obtained by LCMSMS analysis, was Glu-Phe-Gly-Pro-Val-Thr-Asp-Phe-Gly-Tyr-Lys (Figure 19 B).

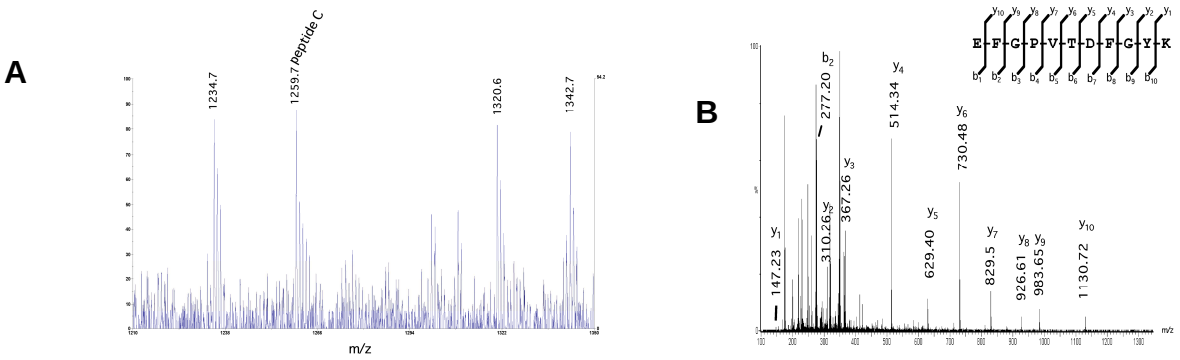


Figure 19 (A) Partial MALDIMS spectrum of the tryptic digest from mutant *fucA1*^{sm} expressed in *E. coli*. The mass signal corresponding to Peptide C encompassing the overlapping region is indicated. (B) LCMSMS analysis of Peptide C [Cobucci Ponzano et al., 2006].

Remarkably, apart for the Glu residue, this sequence is identical to that of Peptide B produced from *fucA1*. Instead Peptide A (Figure 14) was not observed. This indicates that in the mutant gene *fucA1*^{sm} only one of the two frameshifting events observed in the wild type *fucA1* gene had occurred. The presence of a Glu instead of Lys was not unexpected. The mutation A-AAA-AAT $\rightarrow$ A-AA<u>G</u>-AAT in *fucA1*^{sm} was conservative in the zero frame of the ORF SSO11867 (AAA $\rightarrow$ AAG, both encoding Lys), but it produced the mutation AAA $\rightarrow$ GAA (Lys $\rightarrow$ Glu) in the -1 frame of the ORF SSO3060. It is worth noting that the frameshifting efficiency of the gene *fucA1*^{sm}, calculated by Western blot as described above, was two-folds higher (10%) if compared to *fucA1* (5%) (Figure 16A). This indicates that the mutation cancelled the frameshifting site A and, in the same time, enhanced the frameshifting efficiency of site B.

In contrast, the triple mutant *fucA1*^{tm} produced in *E. coli* only the low molecular weight band resulting from translational termination (Figure 20 A). No full-length protein could be detected in Western blots probed with either anti-GST (Figure 20 A) or anti-Ssα-fuc antibodies (Figure 20 B). These data show that the disruption of the heptameric slippery sequence completely abolished the frameshifting in *E. coli* confirming that this sequence has a direct role in controlling the frameshifting *in vivo*. [Cobucci Ponzano et al., 2006].

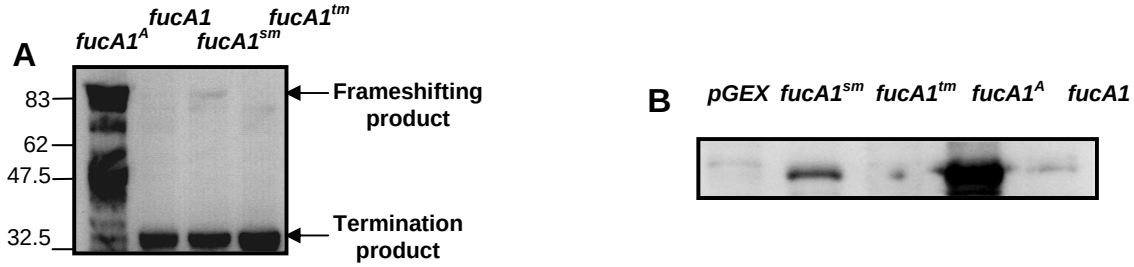


Figure 20: (A) Western blot of *E. coli* cellular extracts expressing *fucA1*^{A}, the wild type *fucA1*, *fucA1*^{sm}, and *fucA1*^{tm} genes (see Materials and Methods). The blot was probed with anti-GST antibodies. (B) Western blot of partially purified protein samples expressed in *E. coli* fused to GST from wild type and mutant *fucA1* genes. Cellular extracts were loaded on GST-Sepharose matrix. After washing, equal amounts of slurries (30 µL of 300 µL) were denaturated and loaded on a 8% SDS-PAGE. Extracts of *E. coli* cells expressing the parental plasmid pGEX-2TK, were used as the negative control (pGEX). The blot was probed with anti-Ssα-fuc antibodies [Cobucci Ponzano et al., 2006].

Expression of *fucA1* in *S. solfataricus*

To test whether *fucA1* is expressed in *S. solfataricus* we analysed the extracts of cells grown on yeast extract, sucrose, and casaminoacids medium (YSM). Accurate assays showed that *S. solfataricus* extracts contained 3.4 x10^{-4} units mg^{-1} of α-fucosidase activity. These very low amounts hampered the purification of the enzyme. The extracts of *S. solfataricus* cells grown on YSM revealed by Western blot a band of a molecular mass greater than 97 kDa and no signals were detected with the pre-immune serum confirming the specificity of the anti-Ssα-fuc antibodies (Figure 21 A). The different molecular mass may result from post-translational modifications occurred in the archaeon or from the incomplete denaturation of a protein complex. In particular, the latter event is not unusual among enzymes from hyperthermophilic archaea [Moracci et al. 1992]. To test which hypotheses were appropriate, cellular extracts of *S. solfataricus* were analysed by Western blot extending the incubation at 100℃ to 2 h. Interesti ngly, this treatment shifted the high-molecular mass band to 67.6±1.2 kDa (Figure 21B and 21C), which still differs from that of the recombinant Ssα-fuc, 58.9±1.2 kDa, leaving the question on the origin of this difference unsolved. To try to shed some light we immunoprecipitated extracts of *S. solfataricus* with anti-Ssα-fuc antibodies and we analysed the major protein band by MALDIMS. Unfortunately, we could not observe any peptide compatible with the fucosidase because the heavy IgG chain co-migrated with the band of the expected molecular weight (not shown).
To test if the scarce amounts of the α-fucosidase in *S. solfataricus* extracts was the result of reduced expression at transcriptional level, it is performed a Northern blot analysis of total RNA extracted from cells grown either on YSM or YGM media. We

could not observe any signal by using probes matching the 3' of the ORF SSO3060 (not shown). These results suggest that *fucA1* produced a rare transcript, therefore, we analysed the level of mRNA by RT-PCR and by Real Time-PCR. A band corresponding to the region of overlap between the ORFs SSO11867 and SSO3060 was observed in the RNA extracted from cells grown on YSM and YGM media, demonstrating that at these conditions the two ORFs were co-transcribed (Figure 22 A).

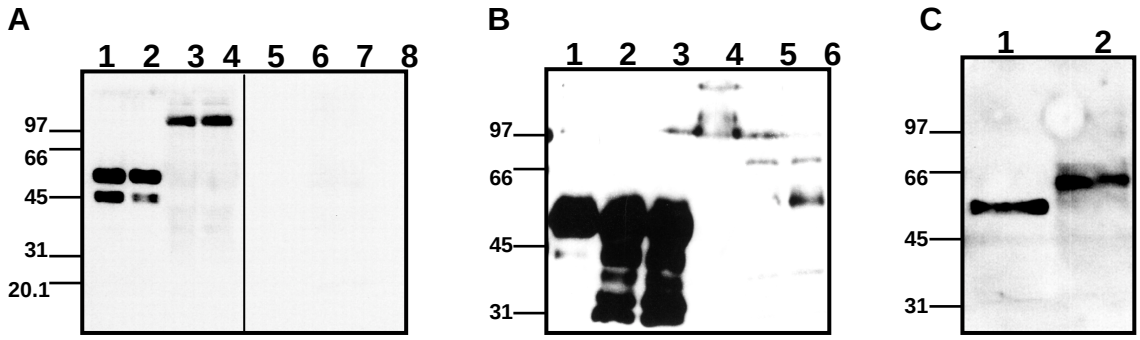


Figure 21: Analysis of the expression of the α-fucosidase in *S. solfataricus*. (A) Western blot analysis of recombinant Ssα-fuc (lanes 1, 2, 5, and 6, 0.14 µg) and of extracts of *S. solfataricus* cells grown on YSM (lanes 3, 4, 7, and 8, 153 µg). Samples in lanes 1, 3, 5 and 7 were not denaturated before loading. The left panel shows the blot probed with anti-Ssα-fuc antibodies; the right panel was probed with the pre-immune serum diluted 1:5,000. (B) Western blot analysis: recombinant Ssα-fuc (lanes 1, 2, and 3, 0.5 µg) incubated at 100℃ for 5 min, 1 h, and 2 h, respectively; extracts of *S. solfataricus* cells (lanes 4, 5, and 6, 1 mg) incubated at 100℃ for 5 min, 1 h, and 2 h, respectively. (C) Western blot analysis of recombinant Ssα-fuc (lane 1, 0.1 µg) incubated at 100℃ for 5 min, and of extracts of *S. solfataricus* cells (lane 2, 1 mg) incubated at 100℃ for 2 h, respectively. The molecular weight markers were: phosphorylase b (97,000), albumin (66,000), ovalbumin (45,000), carbonic anhydrase (31,000), and trypsin inhibitor (20,100).

The experiments of Real Time-PCR shown in Figure 22 B demonstrated that *fucA1* mRNA was amplified after about 38 cycles, 2.2-fold more than the rRNA16S. In addition, considering that it used about 2,500-fold more cDNA for the amplification of *fucA1*, it was calculated that this gene is transcribed about 5,000-fold less efficiently than the rRNA 16S. No significant differences in the fucA1 mRNA level were observed in cells grown in YSM or YGM media, suggesting that the transcription of the gene is not regulated at these conditions. This is further confirmed by the analysis by Western blot of the extracts of the same cells of *S. solfataricus* used to prepare the total RNAs, which revealed equal amounts of α-fucosidase in the two extracts (Figure 22 C). Therefore, the low α-fucosidase activity observed at the conditions tested is the result of the poor transcription of the *fucA1* gene.

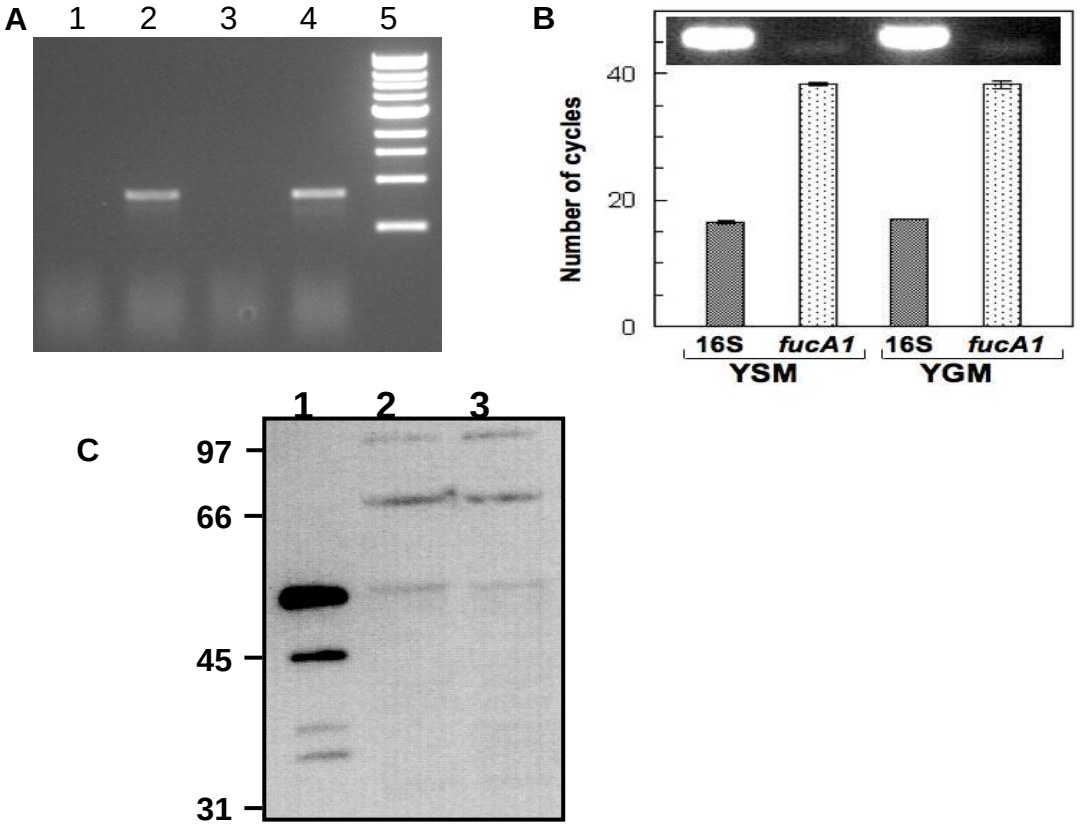


Figure 22: Analysis of the expression of *fucA1* in different media. (A) Agarose gel showing the products of RT-PCR encompassing the ORFs SSO11867 and SSO3060 by using total cellular RNA extracted from cells grown in YSM (lanes 1 and 2) and YGM (lanes 3 and 4); lanes 1 and 3, control (amplification of total RNA supplemented with Taq and without reverse transcriptase enzyme); lane 2 and 4, *fucA1*. (B) Comparison of the *fucA1* mRNA levels in YSM and YGM by Real Time-PCR. The inset shows the corresponding products found in the Real Time-PCR visualized by ethidium bromide staining. (C) Western blot of *S. solfataricus* extracts of cells grown in YSM (lane 2, 80 µg) and YGM (lane 3, 80 µg). Lane 1 recombinant Ssα-fuc (0.2 µg).

Analysis of the expression of *fucA1* in *S. solfataricus* by *in vitro* translation

To determine whether, *S. solfataricus* ribosomes performed the -1 frameshifting, in collaboration with the group of Prof. Londei of "Sapienza" University of Rome, *in vitro* translation experiments were performed as described by Condò et al. [Condò et al., 1999] on mRNAs produced by *in vitro* transcription from the different α-fucosidase constructs. Autoradiography of a SDS-PAGE of the translation products (Figure 23) revealed that the wild type *fucA1* transcript produced a tiny but clear band whose molecular weight corresponded to that of the full-length Ssα-fuc obtained by site-directed mutagenesis [Cobucci Ponzano et al., 2003a]; the latter was translated quite efficiently in the cell-free system. Judging from the relative intensity of the signals given by the translation products of the wild type *fucA1* and the full-length mutant *fucA1*^{A}, the efficiency of the -1 frameshifting in the homologous system was about 10%. No signals corresponding to the polypeptides expected from the separated ORFs SSO11867 and SSO3060 (9.6 kDa and 46.5 kDa, respectively) were observed. However, it should be noted that the product of SSO11867, even if synthesized, is too small to be detected in the gel system employed for this experiment. The larger product of ORF SSO3060, on the other hand, is certainly absent. These data unequivocally demonstrate that the ribosomes of *S. solfataricus* can decode the split *fucA1* gene by programmed -1 frameshifting with considerable efficiency producing a full-length polypeptide from the two ORFs SSO11867 and SSO3060 [Cobucci Ponzano et al., 2006].
Remarkably, under the same conditions at which *fucA1* drives the expression of the full-length protein, we could not observe any product from the *fucA1*^{sm} and *fucA1*^{tm} constructs. These data demonstrate that the integrity of the heptanucleotide is essential for the expression of the *fucA1* gene in *S. solfataricus,* thus further confirming that the gene is decoded by programmed -1 frameshifting in this organism. In addition, the lack of expression of *fucA1*^{sm} by translation *in vitro* in *S. solfataricus* contrasts with the efficient expression of this mutant in *E. coli*, indicating that the two organisms recognize different sequences regulating the translational frameshifting [Cobucci Ponzano et al., 2006].

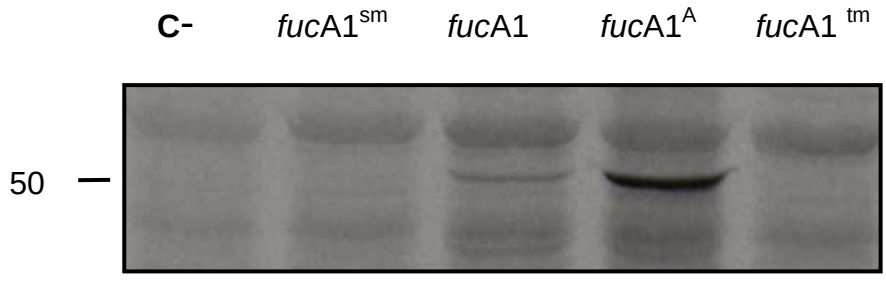


Figure 23: *In vitro* translation. Of each sample 15μL was loaded on 12.5% acrylamide-SDS gel and newly synthesized proteins were revealed by autoradiography. Lane 1, no mRNA added [Cobucci Ponzano et al., 2006].

Preparation of the α-fucosynthase

Glycosynthetic trials with SsD242G

The approach developed by the group in which I performed my thesis for the production of thermophilic glycosynthase consists in the incubation of a retaining glycoside hydrolase mutated in the nucleophilic residue, in the presence of an external ion and a substrate with a good leaving group (Figure 24) [Moracci et al. 1998; Trincone et al. 2000; Perugino et al. 2003; Trincone et al. 2003; Perugino et al. 2006].

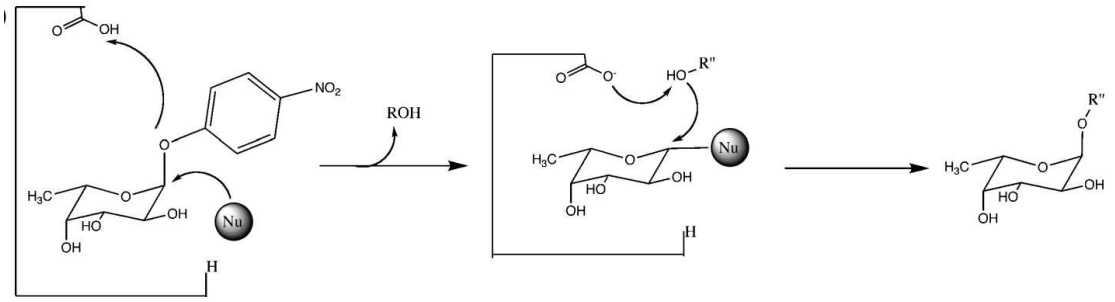


Figure 24: Proposed strategy for the production of a fucosynthase the external nucleophile is indicated by a grey sphere; the R'' group in the fucosylated product is the novel leaving group (see text for the details).

However, it was previously reported that this approach was not successful with the mutant of nucleophile residue of Ssα-fuc (SsD242G) [Cobucci-Ponzano et al. 2003b]. The absence of synthetic products could be due to the nature of the ion used in these experiments (sodium formate) which did not allow the formation of a nucleophyl-glycoside intermediate, sufficiently stable to transfer the glycon mojety to an acceptor [Moracci et al, 1998]. To test this hypothesis I have analysed the activity of the SsD242G mutant in a variety of conditions, including sodium formate or sodium acetate as buffers at different pHs, sodium formate, sodium acetate or sodium chloride as different nucleophile agents in sodium phosphate buffer, in the presence of four glycoside acceptors, and at different concentrations of 4NP-α-L-Fuc substrate donor (Table 6) [Cobucci Ponzano et al., 2008b]. All the assays were performed at 65℃. As a comparison, I reported the specific acti vity of the SsD242G mutant at standard conditions without external nucleophiles. Table 6 shows that, at most of the reactivation conditions tested, the SsD242G mutant showed an increased hydrolytic activity. In particular, sodium formate was, by far, the best nucleophile tested while higher concentrations of 4NP-α-L-Fuc did not improved the activity (entries 2-9). Moreover, the addition in the reaction mixture of either α- or β-D-gluco, and -galactosides, did not increased the activity (entries 10-19); instead, it even hampered the activity chemically rescued (compare entry 2 with entries 10-13). [Cobucci Ponzano et al., 2008b]

Entry	4NP-Fuc donor (mM)	Acceptor	Buffer	pH	External nucleophile	Specific activity (U/mg)	% of maximal activity
1	2	–	50 mM sodium phosphate	6.5	–	8×10^{-4}	–
2	2	–	50 mM sodium phosphate	6.5	2 M sodium formate	0.24	100
3	40	–	50 mM sodium phosphate	6.5	2 M sodium formate	1.4×10^{-3}	0.6
4	40	–	50 mM sodium phosphate	6.5	0.5 M NaCl	2.0×10^{-3}	0.8
5	2	–	0.1 M sodium formate	4.0	–	0.11	46
6	40	–	0.1 M sodium formate	4.0	–	0.05	21
7	2.5	–	0.5 M sodium formate	5.2	–	0.21	87
8	40	–	0.5 M sodium formate	5.2	–	0.08	33
9	40	–	0.1 M sodium acetate	4.0	–	0.01	4
10	1	5mM 4NP-β-D-Gal	50 mM sodium phosphate	6.5	2 M sodium formate	0.05	21
11	1	3mM 4NP-α-D-Gal	50 mM sodium phosphate	6.5	2 M sodium formate	2.2×10^{-3}	0.9
12	1	5 mM 4NP-β-D-Glc	50 mM sodium phosphate	6.5	2 M sodium formate	0.03	12
13	1	3 mM 4NP-α-D-Glc	50 mM sodium phosphate	6.5	2 M sodium formate	1.5×10^{-3}	0.6
14	1	5 mM 4NP-β-D-Gal	50 mM sodium phosphate	6.5	0.5 M NaCl	6.6×10^{-3}	3
15	1	5 mM 4NP-β-Glc	50 mM sodium phosphate	6.5	0.5 M NaCl	3.3×10^{-3}	1
16	1	5 mM 4NP-β-D-Gal	0.5 M sodium formate	5.2	–	2.0×10^{-3}	0.8
17	1	3 mM 4NP-α-D-Gal	0.5 M sodium formate	5.2	–	2.9×10^{-3}	1
18	1	5 mM 4NP-β-D-Glc	0.5 M sodium formate	5.2	–	1.9×10^{-3}	0.8
19	1	2 mM 4NP-α-D-Glc	0.5 M sodium formate	5.2	–	3.4×10^{-3}	1

Table 6: Chemically rescued activity of the SsD242G mutant in presence of several external nucleophiles and several buffers [Cobucci Ponzano et al., 2008 b]

The inspection by thin layer chromatography (TLC) of the reaction mixtures incubated for 16 h at 65°C reported in Table 6, did not revealed any product (data not shown), confirming that the mutant did not worked as a glycosynthase. A possible explanation of this result could be that the leaving ability of the aglycon group in the substrate used, 4-nitrophenol, was insufficient to perform glycosynthesis. In fact, hyperthermophilic β-glycosynthases efficiently produced oligosaccharides by using 2-nitrophenyl-β-D-glycoside substrates, in which 2-nitrophenol, though showing a pK_a similar to 4-nitrophenol (pK_a 7.22 and 7.18, respectively), can form a chelate ring by hydrogen bonding and thereby increasing the leaving ability upon protonation [Perugino et al., 2003].
Unfortunately, no α-L-fucosides substrates showing aglycons with leaving ability better than 4-nitrophenol are commercially available. To overcome these problems, the group of Prof. Parrilli of the University of Naples "Federico II" synthesized the substrate 2-chloro-4-nitrophenyl-α-L-fucopyranoside (2C4NPα-L-Fuc) (Figure 25). In this compound the aglycon 2-chloro-4-nitrophenol shows a pK_a of 5.45 [Tehan et al. 2002], which is noticeably lower than the 4-nitrophenol one (7.18); this difference makes the 2-chloro-4-nitrophenyl a much better leaving group and, conceivably, this would improve the first step of the reaction. In fact, the good leaving ability of the aglycon would require less acidic assistance by the acid/base residue and thereby increasing the efficiency of the external ion [Cobucci Ponzano et al., 2008b].

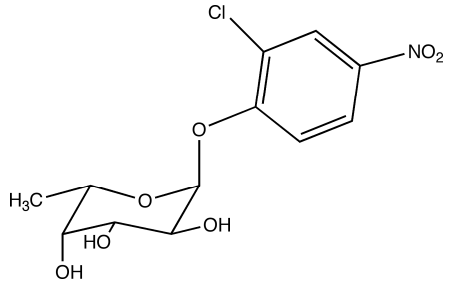


Figure 25: 2-Cl-4NO_2-phenyl-α-L-fucopyranosyde [Cobucci Ponzano et al, 2008b]

The steady-state kinetic constant at 65°C of the mutant SsD242G on 4NP- and 2C4NP α-L-Fuc were measured in 50 mM sodium phosphate buffer, pH 6.5, in the presence of 2 M sodium azide and were compared with the kinetic constants of the wild type Ssα-fuc without external ions (Table 7).

	k_{cat} (sec^{-1})	K_M (mM)	k_{cat}/K_M (sec^{-1} mM^{-1})
wild type			
4Np-Fuc	287±11	0.028±0.004	10,250
2C4Np-Fuc	157±9	0.013±0.004	11,602
SsD242G+ NaN_3 2M			
4Np-Fuc	9.66±0.28	0.19±0.02	51.55
2C4Np-Fuc	387±18.3	0.13±0.02	2,790

Table 7: Steady-state kinetic constants of wild type and SsD242G mutant. [Cobucci Ponzano et al., 2008b]

The specificity constant k_{cat}/K_M of the mutant on 2C4NPα-L-Fuc was more than 7-fold higher than that found on the commercial substrate 4NPα-L-Fuc as expected for an aryl glycoside substrate with a better leaving group [Kempton and Withers 1992]. This activation reflects mainly the increment in k_{cat} while the K_M values remain almost unaltered, indicating that the affinity of the SsD242G mutant was the same on 4NP- and 2C4NP-α-L-Fuc. Interestingly, the wild type on the latter substrate produced kinetic constants similar to those obtained with 4NPα-L-Fuc (Table 7), indicating that the different leaving abilities of the aglycons in the two substrates did not change the reaction rates. These data show that, for the wild type, the deglycosylation step of the reaction is the limiting one, since the better leaving ability of the 2-chloro-4-nitrophenol aglycon, most important for the first, glycosylating step of the reaction, did not make any difference. By contrast, the presence of the activated leaving group greatly enhanced the activity of the mutant showing that the SsD242G mutation affect the first step of the reaction.
To test if this activation reflects a better synthetic activity we analysed the reaction products of the mutant incubated in the presence of 2.5 mM 2C4NPα-L-Fuc, 0.5 M sodium formate buffer pH 5.2 at 65°C for 16 h; unfortunately, again, no oligosaccharide product was observed by TLC. However, the specific activity of the SsD242G mutant at these conditions was 2-fold higher than that measured, at the same conditions, on 4NPα-L-Fuc (0.5 U/mg and 0.21 U/mg on 2C4NP- and 4NP-α-L-Fuc, respectively). The activity observed in the SsD242G with the new substrate in the presence of formate was 9-fold lower than that found with sodium azide. These data show that the impaired catalytic activity of the mutant in the nucleophile of the reaction of the α-L-fucosidase from *S. solfataricus* can be significantly improved in the presence of external ions and of an aryl-α-L-fucoside substrate prepared by chemical synthesis [Cobucci Ponzano et al., 2008b].

Construction of Ssα-fuc nucleophile mutants and kinetic characterization

To test if the nature of the mutation introduced in the α-L-fucosidase from *S. solfataricus* could be responsible for the absence of synthesis observed in the mutant SsD242G [Cobucci-Ponzano et al. 2008b], I have prepared two other mutants of the enzyme, namely SsD242A and SsD242S. It is well known that non-nucleophilic residues such as Gly, Ala, and Ser are among the best acting amino acids in the preparation of glycosynthases. In particular, for serine, the polar character of the side chain can form hydrogen bonds to the external nucleophile and therefore promoting the nucleophilic attack to the anomeric centre of the substrate [Mayer et al. 2000].
SsD242A and SsD242S mutants were produced as previously described and were purified by exploiting the Glutathione S-Transferase tag in a single purification step yielding about 1.2 mg of pure mutant enzyme from 2 lt of *E. coli* culture.
As expected from the essential role played in catalysis by the nucleophile of the reaction D242 the activity of both enzymes was severely hampered (Table 8). Incubation of the mutant SsD242S with sodium formate or sodium azide in 50 mM sodium phosphate buffer (pH 6.5) at 65°C resulted in increased activity on both 4NP-α-L-Fuc and 2C4NP-α-L-Fuc substrates as a function of the concentration of the nucleophile (Figures 26-27). The rescue of activity in sodium formate was similar for the two substrates (Figure 26) whereas, in sodium azide, SsD242S mutant showed an about 3-fold higher specific activity on 2C4NP-α-L-Fuc if compared to 4NP-α-L-Fuc substrate (Figure 27). In contrast, the SsD242A mutant was not reactivated at all the conditions tested (Table 8) and it was not further characterized.
The steady-state kinetic parameters at 65°C for SsD242S on the 4NP-α-L-Fuc and 2C4NP-α-L-Fuc substrates were obtained by incubating at the fixed concentration of 2 M sodium azide (saturating concentration) and varying the concentration of the substrates. The apparent k_{cat} and K_M values of SsD242G and SsD242S mutant are compared to the wild type Ssα-fuc (Table 8). The specificity constant on 4NP-α-L-Fuc is almost 5-fold higher than that of the mutant SsD242G on the same substrate in the presence of 2M sodium azide; instead, at the same conditions, for the 2C4NP-α-L-Fuc substrate the k_{cat} of the SsD242G/S mutants and the Ssα-fuc wild type are similar (Table 8).

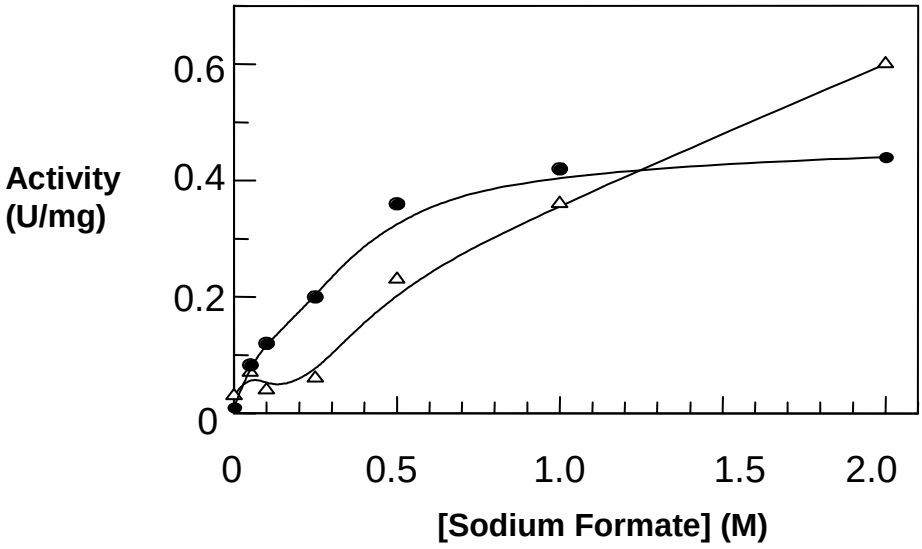


Figure 26: Specific activity of the SsD242S mutant in the presence of sodium formate. Assays were performed by using SsD242S mutant (5 µg) at 65℃ in 50 mM sodium phosphate buffer (pH 6.5), in the presence of 2 mM 4NP-α-L-Fuc (closed circles), and 0.4mM 2C4NP-α-L-Fuc (open triangles) .

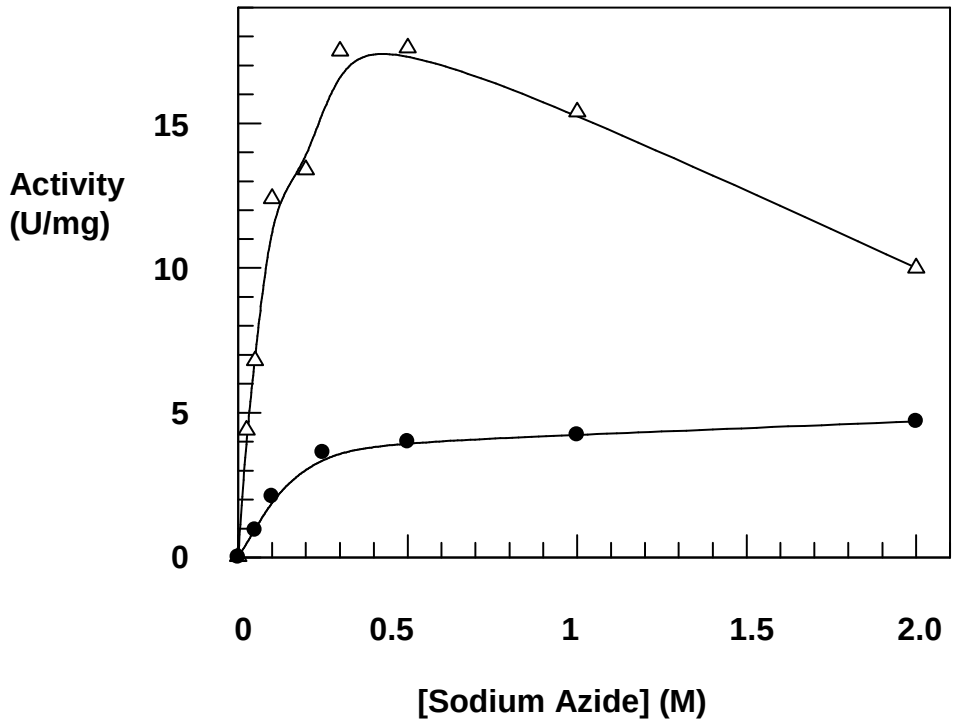


Figure 27: Specific activity of the SsD242S mutant in the presence of sodium azide. Assays were performed by using SsD242S mutant (5 µg) at 65℃ in 50 mM sodium phosphate buffer (pH 6.5), in the presence of 2 mM 4NP-α-L-Fuc (closed circles) 3 mM 2C4NP-α-L-Fuc (open triangles).

	k_{cat} (s^{-1})	K_M (mM)	k_{cat}/K_M (s^{-1} mM^{-1})
Wild type			
4Np-fuc	287±11	0.028±0.004	10250
2C4Np-fuc	157±9	0.013±0.004	11602
SsD242A			
4Np-fuc	0.50[a]	--	--
+ sodium azide 2 M			
4Np-fuc	0.67[a]	--	--
+ sodium formate 2 M			
4Np-fuc	0.05[a]	--	--
SsD242G			
4Np-fuc	0.24[a]	--	--
+ sodium azide 2 M			
4Np-fuc	9.66±0.28	0.19±0.02	51.55
2C4Np-fuc	387±18	0.13±0.02	2790
SsD242S			
4Np-fuc	0.08[a]	--	--
2C4Np-fuc	0.25[a]	--	--
+ sodium azide 2 M			
4Np-fuc	47±1	0.19±0.02	247
2C4Np-fuc	286±35	0.29±0.13	987

Table 8: Kinetic parameters of Ssα-fuc mutants in presence of 4Np-α-L-fuc and 2C4Np-α-L-fuc. [a]k_{cat} was determined from the initial velocity at saturating concentration of substrate

Products of the rescued activity of Ssα-fuc mutants

The incubation of the SsD242G and SsD242S mutants (6 µg each) in 50 mM sodium phosphate buffer (pH 6.5), 0.1 M sodium azide, and 4NP-α-L-Fuc or 2C4NP-α-L-Fuc substrates used at the maximal concentration obtained (2 and 20 mM, respectively) yielded newly formed transfucosylation products identified by thin-layer chromatography (TLC) (Figure 28). One of these compounds is not unexpected, in fact, we reported previously that the SsD242G mutant, reactivated with sodium azide, produce β-L-fucopyranosyl azide based on the inverting mechanism of a α-L-fucosidase mutated in the nucleophilic catalytic residue with this external nucleophile [Cobucci Ponzano et al. 2003b]. However, the two mutants at the same conditions, but only in the presence of the 2C4NP-α-L-Fuc substrate, produced, very unexpectedly, a more polar compound, which was not UV-visible on the TLC. Remarkably, after incubation with the wild type Ssα-fuc, this product was completely hydrolysed, demonstrating that it contained an α-L-anomeric bond (Figure 28B). To isolate and characterize the products of the reaction of the mutants in the presence of the 2C4NP-α-L-Fuc substrate, we scaled up the reaction by incubating the SsD242G mutant (48 µg) in 50 mM sodium phosphate buffer (pH 6.5), 0.1 M sodium azide, and 20 mM 2C4NP-α-L-Fuc substrate. This allowed us to unequivocally define the structure of products ***1*** and ***2*** by combined use of nuclear magnetic resonance and mass spectrometry (Figure 28C). Compound ***1*** was, as expected, β-L-fucopyranosyl azide (β-L-Fuc-N_3) while, surprisingly, compound ***2*** was the disaccharide α-L-fucopyranosyl(1-3)-β-L-fucopyranosyl azide (α-L-Fuc-(1-3)-α-L-Fuc-N_3). It is worth noting that both mutants could not produce compound ***2*** from the 4NP-α-L-Fuc substrate, but only from the 2C4NP-α-L-Fuc (Figure 28A). Moreover, the inspection of the TLC in Figure 28B showed that the SsD242S produced compound ***2*** more efficiently than SsD242G mutant. These results suggest that the better leaving ability of the 2-chloro-4-nitrophenol aglycone, if compared to the 4-nitrophenol aglycone, and the mutation D242S were decisive elements for the improvement of the transfucosylation reaction. It is worth noting that the disaccharide ***2*** contained the β-L-Fuc-N_3 group at the non-reducing end (Figure 28C) and no autocondensation products of the 2C4NP-α-L-Fuc donor could be isolated, suggesting that the substrate was an efficient donor, but a poor acceptor. However, the most surprising result is that the α-L-fucosidase from *S. solfataricus* mutated in the nucleophilic catalytic residue promoted the formation of *both* β-L- and α-L-bonds by transfucosylation reactions when its activity was chemically rescued in sodium azide.

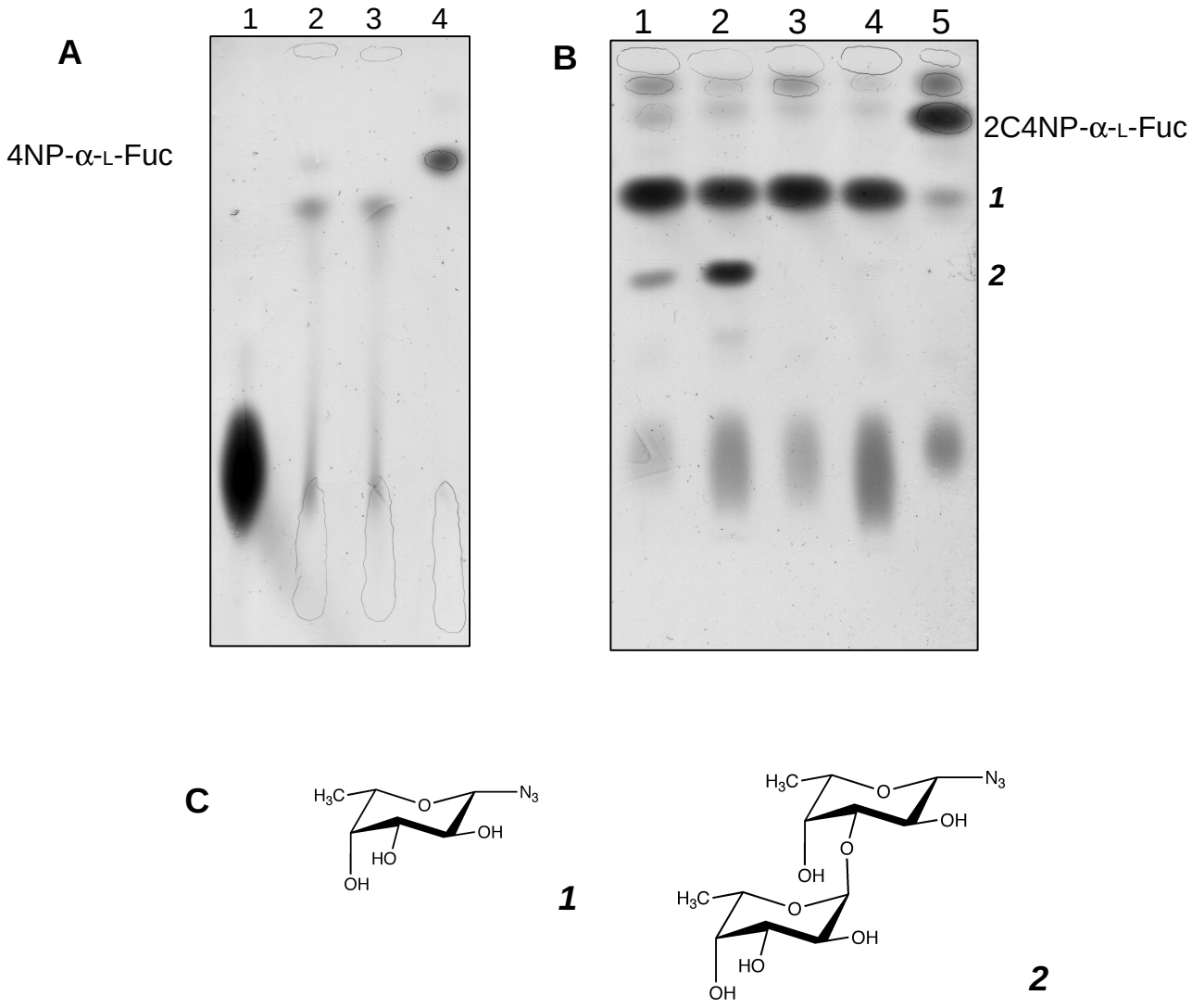
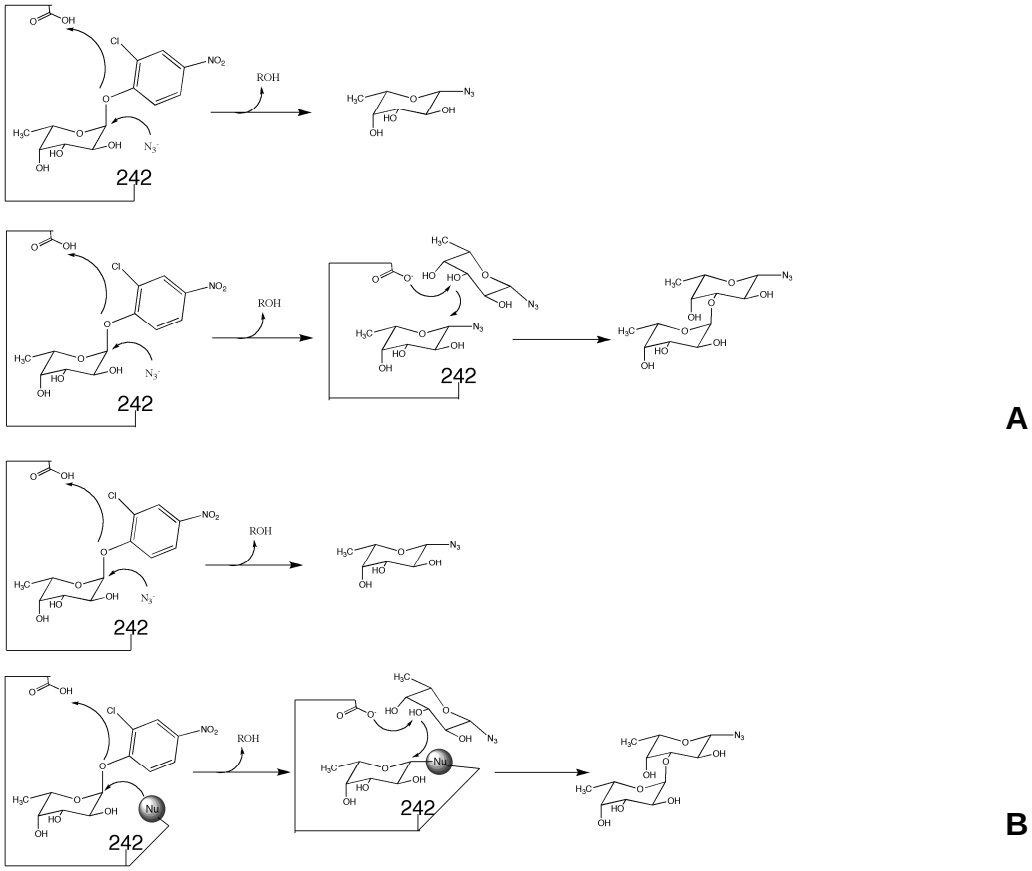


Figure 28: TLC detection of tranfucosylation activity of SsD242G and S mutants in the presence of sodium azide. The mutants enzymes SsD242G and SsD242S (6 µg) were incubated for 16 h at 65℃ in 50 mM sodium phosphat e buffer, pH 6.5, in the presence of sodium azide and substrate at the concentrations indicated below. The products were separated on TLC. Panel A. 2 M Sodium azide and 2 mM 4NP-α-L-Fuc; lane 1: fucose marker; lane 2: SsD242G reaction; lane 3: SsD242S reaction; lane 4: blank with no enzyme. Panel B. 0.1 M Sodium azide and 20 mM 2C4NP-α-L-Fuc: lane 1: SsD242G reaction; lane 2: SsD242S reaction; lanes 3 and 4: the same samples shown in lanes 1 and 2, respectively, after incubation in the presence of wild type Ssα-fuc (1.2 µg) at 65℃ for 2 hrs; lane 5: blank with no enzyme. Panel C. Based on later product analysis ***1*** and ***2*** turned out to be β-L-Fuc-azide and α-L-Fuc-(1,3)-β-L-Fuc-azide, respectively.

Characterization of the reaction mechanism of Ssα-fuc mutants

To explain how the Ssα-fuc mutants promoted the synthesis of the compounds shown in Figure 28 C, here I proposed two alternative reaction mechanisms (Scheme 1), which share a common preliminary phase. In fact, the first phase of the reaction (Scheme 1 A and B upper diagrams, respectively), consists in the nucleophilic attack of the azide to the anomeric centre of the substrate leading to the reactivation of the enzyme and to the production of a compound with *inverted* anomeric configuration (compound ***1***: β-L-Fuc-N_3 from the 2C4NP-α-L-Fuc substrate). The reaction shown in this phase is not novel and it is widely exploited to identify the nucleophile of the reaction in *retaining* glycosidases [Ly and Withers 1999]. On the basis of the anomeric configurations of the substrate and the product of the reaction (from 2C4NP-α-L-Fuc to β-L-Fuc-N_3) it is commonly stated that the mutant follow an *inverting* mechanism. Following to this phase, we propose that as a result of its accumulation, the β-L-Fuc-N_3 product acts as a novel acceptor that is activated by the catalytic general base residue of the enzyme (Scheme 1 A and B lower diagrams, respectively). Therefore, compound ***1*** is transferred to another β-L-Fuc-N_3 molecule leading to the α-L-Fuc-(1-3)-β-L-Fuc-N_3 disaccharide ***2*** containing the α-L-bond. It is worth noting that the substrate is 2C4NP-α-L-Fuc, then, in the phase two of the reaction (Scheme 1A and B lower diagrams, respectively), the enzyme catalyse the formation of the product by following a *retaining* mechanisms. In conclusion, we propose that the Ssα-fuc mutants, in the presence of 2C4NP-α-L-Fuc substrate and sodium azide, follow a novel reaction mechanism in two phases in which the *retaining* mechanism follows the *inverting* mechanism.



Scheme 1: Two different hypotesis to explain the mechanism of synthetic reaction of SsD242S.

For the mechanicistic reasons stated above the *inverting* mechanism by which is produced compound ***1*** can be easily foreseen; instead, two alternative mechanisms might explain the formation of compound ***2***. In Scheme 1A (lower diagram), we propose that, once the β-L-Fuc-N_3 molecule is formed, it is not released, but it persists in the active site of the enzyme as a metastable fucosyl-azide intermediate. Therefore, the activation of another β-L-Fuc-N_3 molecule acting as acceptor determines the autocondensation reaction leading to the α-L-Fuc-(1-3)-β-L-Fuc-N_3 disaccharide. This mechanism would imply that azide could work as leaving group. Though not impossible, this possibility would be certainly surprising in a reaction promoted by an enzyme mutated in the catalytic nucleophile. Therefore, as an alternative, it is hypothesized in Scheme 1B (lower diagram) that a third residue in the active site of the enzyme, not playing a direct role in catalysis, might gain the function of nucleophile of the reaction by replacing the D242 catalytic nucleophile present in the wild type enzyme. This residue might form the covalent fucosyl-enzyme intermediate which is successively transferred to the β-L-Fuc-N_3 acceptor. In this second hypothesis, the claiming of an additional nucleophile residue is not purely speculative. We have previously reported that two glutamic acid residues, namely Glu58 and Glu292, cooperate with D242 in catalysis [Cobucci-Ponzano et al. 2005], thus, one of these residues might have assumed a nucleophilic character in the mutant SsD242S.

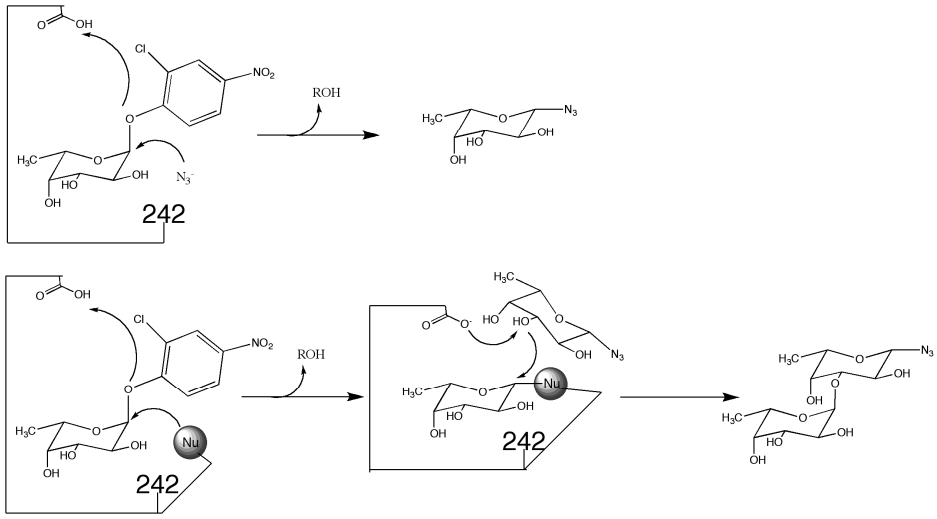
Double Mutants Experiments

To test the validity of the reaction mechanism hypothesized in Scheme 1B, I prepared the double mutants SsD242G/E58G and SsD242G/E292G. The mutants were expressed and purified as described above by using the GST-tag. For the two mutants the procedure yielded about 1.2 mg of pure protein from 2 lt of *E. coli* cultures. Prolonged incubation of the SsD242G/E58G and SsD242G/E292G mutants at 65℃ with 0.1 M sodium azide and 2C4NP- α-L-Fuc substrate (20 mM) led to the synthesis of β-L-Fuc-N_3, but no disaccharide product was observed by TLC (data not shown). Noticeably, most of the substrate remained unreacted indicating that both mutants were very inefficient catalysts. This experiment, though confirming the importance of both Glu58 and Glu292 residues in the catalytic machinery of Ssα-fuc, did not help in understanding their mechanicistic role in the SsD242G mutant.

Inhibition analysis of wild type Ssα-fuc and SsD242S

Thus, to identify a possible residue acting as additional nucleophile in the SsD242G and SsD242S mutants it is analyzed the effect of the mechanism based inhibitor 2-deoxy-2-fluoro-α-L-fucosyl fluoride (2d-2F-Fuc-F). The preparation of this compound was carried out by the group of Prof. Parilli from the University of Naples “Federico II”. The 2d-2F-Fuc-F compound allowed the unequivocal identification of the catalytic nucleophile of the α-L-fucosidase from *Thermotoga maritima* (Tmα-fuc). In this enzyme, the formation of a sufficiently stable 2-deoxy-2-fluoro-fucosyl enzyme complex allowed the easy identification of the catalytic nucleophile by HPLC/electrospray mass spectrometry [Tarling et al. 2003]. In our case, we postulated that a residue acting as additional catalytic nucleophile in SsD242S might react with the mechanism-based inhibitor. To test this hypothesis, it was incubated at 65℃ in sodium phosphate buffer 50 mM (pH 6.5) 60 µg (1.1 nmol) of SsD242S

mutant with 0.8 µg (4.7 nmol) of 2d-2F-Fuc-F (enzyme active site : inhibitor 1:4.5). As positive control, the wild type Ssα-fuc was incubated with the inhibitor at the same conditions. As previously reported for Tmα-fuc, the mechanism-based inhibitor 2d-2F-Fuc-F did not lead to time-dependent inhibition of the activity of wild type Ssα-fuc. This was explained by the rapid turn-over via transglycosylation to 4NP-α-L-Fuc during the activity assays [Tarling et al. 2003]. However, following pepsin digestion of the labelled Ssα-fuc, in collaboration of the group of Dr. Del Piaz of Salerno University using a LC/MS/MS system, they could isolated a peptide showing a particular molecular weight, attributed to the fragment of the protein carrying a 2-deoxy-2-fluorofucosyl moiety, possibly bound to Asp242. The identification of the modified peptide was achieved both on the basis of the measured molecular weight, and on the basis of information on the sequence of the peptide obtained by tandem MS experiments. The unmodified peptide showing an experimental MW similar to theoretical MW was also detected, In contrast, the SsD242S mutant did not produced any labelled peptide indicating that this protein did not react with the inhibitor. Though this negative result is not conclusive, our data strongly indicate that no additional catalytic nucleophile was produced by mutating Asp242 in Ssα-fuc, making questionable the validity of the reaction mechanism shown in Scheme 2 below.



Scheme 2: Second nucleophile hypothesis for SsD242S mechanism.

Proposed reaction mechanism of SsD242S

To validate the reaction mechanism described in Scheme 1 A, I tested if the mutant SsD242S was able to use directly the β-L-Fuc-N_3 substrate as donor in the absence of the sodium azide as external nucleophile. The β-L-Fuc-N_3 was chemically synthesized by the group of Prof.Parrilli. The SsD242S mutant was incubated for 16 h at 65℃ in 50 mM sodium phosphate (pH 6.5) in the presence of 20 mM β-L-Fuc-N_3; the products of the reaction were analysed by TLC (Figure 29). Remarkably, I observed the formation of a product with the same polarity of the disaccharide ***2***. At the conditions shown not all the donor was converted and the transfucosylation efficiency, defined as the amount of fucose transferred to an acceptor different from water, was 40%. However, the regioselectivity was remarkably high with exclusive formation of the α-(1-3)-bond; no other regioisomers were detectable. Similar results were obtained with the mutant SsD242G, but the inspection of the reaction products

by TLC indicated lower efficiency. As a control, we incubated the wild type Ssα-fuc at the same conditions, but the β-L-Fuc-N_3 substrate remained unreacted (data not shown), ruling out possible artefacts produced by wild type contaminations. However, the incubation of the wild type Ssα-fuc for 1 h at 65℃ in 50 mM sodium phosphate (pH 6.5) in the presence of 10 mM β-L-Fuc-N_3 and 10 mM 2C4NP-β-L-Fuc yielded on TLC two spots showing the same polarity of compounds ***1*** and ***2***.

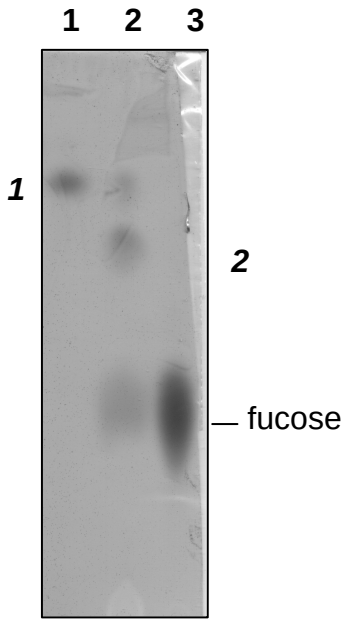
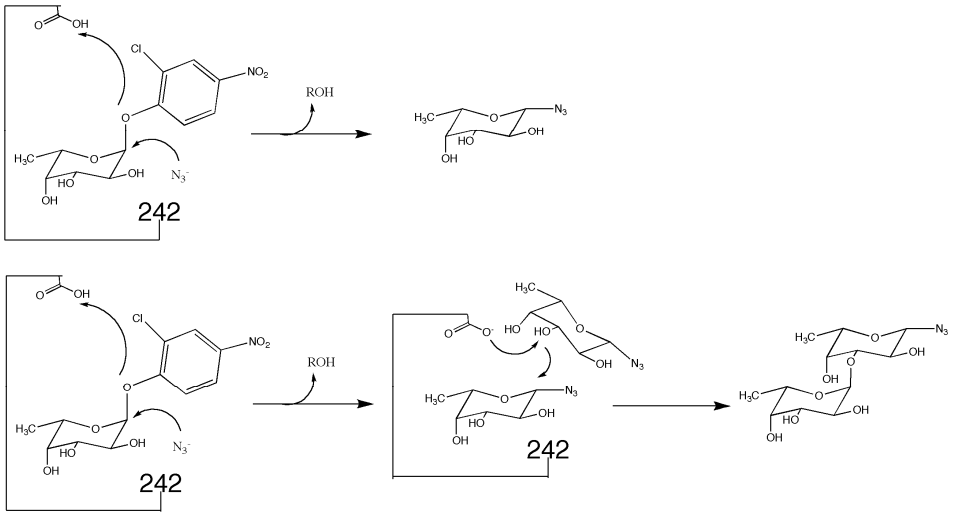


Figure 29: TLC detection of transfucosylation activity of SsD242S mutants in the presence of β-L-Fuc-azide donor. The mutant SsD242S (6 µg) was incubated for 16 h at 65℃ in 50 mM sodium phosphate buffer (pH 6.5) in the presence of 20 mM β-L-Fuc-azide as substrate. Lane 1: blank with no enzyme; lane 2: SsD242S reaction; lane 3: fucose marker.

These results indicated that Ssα-fuc mutants SsD242G and SsD242S can efficiently exploit β-L-Fuc-N_3 as donor in promoting transfucosylation reactions leading to α-L-Fuc-(1-3)-β-L-Fuc-N_3 by autocondensation. This confirmed the reaction mechanism described in Scheme 3 in which the β-L-Fuc-N_3, produced by the mutants from the 2C4NP-α-L-Fuc substrate, become a donor during the course of the reaction.



Scheme 3: Proposed mechanism of the α-L-fucosynthase.

Oligosaccharide synthesis by Ssα-fuc mutants

The demonstration of the mechanism described in Scheme 3 indicates that the mutant SsD242S has excellent chances to be exploited for the synthesis of fucosylated oligosaccharides. To this aim I performed a series of synthetic trials by using the β-L-Fuc-N_3 donor and different glycosides as substrate acceptors. The mutant SsD242S (6 μg) was incubated for 16 h at 65°C in sodium phospha te buffer (pH 6.5) in the presence of β-L-Fuc-N_3 as donor (at concentration ranges 1.5-20 mM) in the presence of different aryl-glycosides (concentrations range 6-34 mM) as well as monosaccharide (3-40 mM), and disaccharide (12-20 mM) acceptors. The concentration used for the acceptor depended on the solubility of the different compounds and the concentration of the β-L-Fuc-N_3 donor was adjusted consequently so that we had final molar ratios between 1:2 and 1:3.4 donor: acceptor. Incubations with all the monosaccharides (fucose, xylose, arabinose, galactose, glucose, and mannose), the disaccharides lactose, cellobiose, *N*-acetyl-lactosammine, and 2-*O*-α-L-fucosyl-D-galactose, and with some aryl-glycosides (4-aminophenyl-α-L-fucopyranoside, 4-aminophenyl-β-D-glucopyranoside, 4NP-α-D-Man, 2NP-β-D-laminaribioside, 2NP-β-D-cellobioside), and uridine 5'-diphospho-N-acetyl-glucosamine), gave no transfucosylation products. In all these cases, but 4-aminophenyl-α-L-fucopyranoside only the α-L-Fuc-(1-3)-β-L-Fuc-N_3 autocondensation product was observed; presumably, 4-aminophenyl-α-L-fucopyranoside competed with β-L-Fuc-N_3 for the donor site. However, interestingly, the SsD242S mutant promoted the formation of several transfucosylation products, identified by TLC, on a large variety of aryl-glycosides summarized in Table 9.

Entry	Donor	Acceptor	Acceptor concentration (mM)	Number of UV-visible synthetic products observed by TLC
1	β-Fuc-N_3	4NP-α-L-Fuc	6-15	1
2		4NP-α-D-Glc	15	2
3		4NP-α-D-Xyl	10 - 15	3
4		4NP-α-L-Ara	10	3
5		4NP-β-D-Glc	20 - 34	2
6		2NP-β-D-Xyl	20	2
7		4NP-β-D-Xyl	20 - 34	3
8		4NP-β-D-Gal	20 - 34	2
9		4NP-β-D-GlcNAc	10 - 15	2
10		2NP-β-D-Fuc	20 - 34	3
11		4NP-β-D-Man	20 - 34	2
12		4NP-β-L-Fuc	20	2
13		MU-β-D-Glc	5	1
14		MU-α-L-Fuc	5	1
15		4NP-α-D-Mal	20	3
16		4NP-β-D-Lac	10	2

Table 9: Glycosynthetic trials of SsD242S with several acceptors.

Remarkably, the mutant enzyme promoted transfucosylation reactions on 2- and 4NP-glycosides of either exoses (α–L, β-L, and β-D-Fuc, α- and β-D-Glc, β-D-Gal, and β-D-Man), pentoses (α- and β-D-Xyl and α-L-Ara), *N*-acetyl-glucosammine, disaccharides (Mal and Lac), and on methylumbelliferyl-fuco- and glucosides. Moreover, SsD242S took acceptors containing either α/β-L/D anomeric bonds. In all cases we observed by TLC analysis the formation of autocondensation products.
These results showed that the mutant has wide specificity for the acceptor molecule, however, since we did not further characterize the reaction products, we do not know if the transfucosylation products are due to a rather relaxed regioselectivity or to the production of longer oligosaccharides. Notwithstanding these limitations, it is worth noting that most of the donor was converted and all the products accumulated in the reaction even after prolonged incubation as expected for a functional fucosynthase.

Construction and characterization of Tmα-fuc nucleophile mutants

I aimed to exploit the transfucosylating activity of fucosidase mutants in the catalytic nucleophile using β-L-fucosyl azide donor as a novel general strategy for the production of an efficient fucosynthase. To test this possibility I prepared mutants of the α-L-fucosidase from the thermophilic bacterium *Thermotoga maritima*. This enzyme, though belonging to family GH29, has only 25% amino acid sequence identity to Ssα-fuc. In addition, the Tmα-fuc active site residues have been experimentally identified [Tarling et al. 2003], and 3D-structure is currently available [Sulzenbacher et al. 2004]; thus, it is an ideal model system to test if our strategy can be generally applied to fucosidases.
By analogy with Ssα-fuc, we mutated the catalytic nucleophile Asp224 of Tmα-fuc in Gly and Ser residues obtaining the mutants TmD224G and TmD224S. The mutant proteins were purified to homogeneity by exploiting the His-tag; about 0.3 mgs of pure protein were obtained from 2 lt of *E. coli* culture. The His-tag was not removed as it did not affect the activity and the stability of the mutants at all the conditions tested.
As expected, the mutation of the catalytic nucleophile severely affected the fucosidase activity at 60℃ in 50 mM sodium citrate /phosphate buffer (pH 6.0) in the presence of 1 mM 4NP-α-L-Fuc and 1 mM 2C4NP-α-L-Fuc. At these conditions the activity of the TmD224S mutant could not be detected and it was not rescued in sodium azide or sodium formate, each tested at a concentration range of 0-2 M, thus, this mutant was not analysed any further. Instead, the residual activity of the TmD224G mutant (0.05 and 0.46 s^{-1} on 4NP-α-L-Fuc and 2C4NP-α-L-Fuc, respectively) was rescued by sodium azide in 50 mM sodium citrate/phosphate buffer (pH 6.0) at 60℃ on both substrates as a function of the concentration of the nucleophile (Figure 30). The specific activity of the mutant was higher on 2C4NP-α-L-Fuc as observed for the SsD242S mutant in sodium azide (compare Figure 30 and Figure 27). Instead, the activity of this TmD224G mutant could not be rescued by using sodium formate as external nucleophile.
The steady-state kinetic constants of the fucosidase activity rescued in sodium azide of the TmD224G mutant are compared to those of the wild type Tmα-fuc in Table 10. The specificity constant of the mutant at a saturating concentration of sodium azide was 4- and 18-fold lower than that of the wild type on 4NP-α-L-Fuc and 2C4NP-α-L-Fuc, respectively, however, the affinity for both the substrates was not significantly changed by mutation remaining in the μM range for both the enzymes.

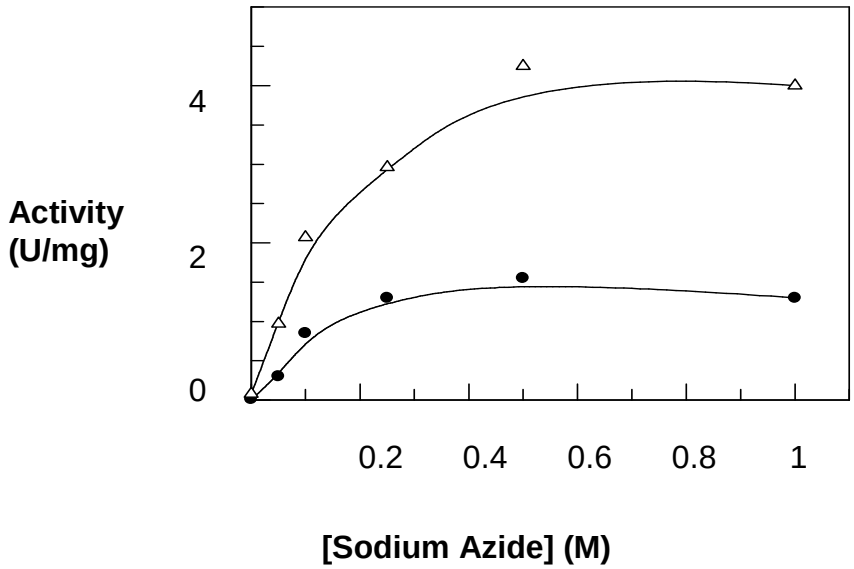


Figure 30: Specific activity of the TmD224G mutant in the presence of sodium azide. Assays were perfomed at 60℃ incubating 6 and 4 μg of TmD224G mutant (for 4Np-α-L-Fuc and 2C4Np-α-L-Fuc, respectively) in 50 mM sodium citrate/phosphate buffer (pH 6.0), sodium azide, and 0.1 mM 4Np-α-L-Fuc and 2C4Np-α-L-Fuc (closed squares and open triangles, respectively

	k_{cat} (s^{-1})	K_M (mM)	k_{cat}/K_M (s^{-1} mM^{-1})
Wild type			
4Np-fuc	79.6±3.0	0.033±0.005	2426
2C4Np-fuc	88.5±5.2	0.007±0.002	12287
TmD224G			
4Np-fuc	0.05[a]	--	--
2C4Np-fuc	0.46[a]	--	--
+ sodium azide 2 M			
4Np-fuc	9.2±0.3	0.015±0.002	597
2C4Np-fuc	26.7±0.7	0.040±0.005	681

[a] k_{cat} was determined from the initial velocity at saturating concentration of substrate

Table10: Steady-state kinetic constants of wild type Tmα-fuc and TmD224G at 60℃

Encouraged by these results, I analysed by TLC the reaction products of the TmD224G mutant after incubation at 60℃ for 16 hrs in sodium azide concentrations ranging 0.05-1.0 M, 50 mM sodium citrate/phosphate buffer (pH 6.0), and 20 mM 2C4NP-α-L-Fuc. The single transfucosylation product observed by TLC showed the

same polarity of β-L-Fuc-N_3 standard while no disaccharide products were formed (data not shown). More surprisingly, by TLC analysis we could not observe any transfucosylation product when the mutant TmD224G was incubated at 60℃ for 16 hrs in 50 mM sodium citrate/phosphate buffer (pH 6.0) and 20 mM β-L-Fuc-N_3 donor (data not shown). However, incubations of the mutant, at the same conditions in the presence of 10 mM β-L-Fuc-N_3 donor and several different aryl-glycosides , lactose, and *N*-acetyl-lactosammine acceptors used at 1:10 donor: acceptor molar ratios revealed by TLC analysis several transfucosylation products (data not shown). As shown previously for SsD242S mutant (Table 9), the TmD224G was not active on lactose and *N*-acetyl-lactosammine while it recognized as acceptors 2- and 4NP-β-D-Xyl, 2NP-β-D-Fuc, and 4NP-β-D-Glc. Instead, in contrast with SsD242S, TmD224G was not active on 4NP-β-L-Fuc and 4NP-β-D-lactose.

Oligosaccharide synthesis by hyperthermophilic fucosynthases

The synthetic ability of the mutants SsD242S and TmD224G was analysed by performing preparative reactions with the two enzymes. In particular, 94 µg of SsD242S mutant were incubated for 16 h at 65℃ in 5 0 mM sodium phosphate buffer (pH 6.5) with 10 mM of β-L-Fuc-N_3 donor and 4NP-β-D-Xyl or 4NP-β-D-Gal (34 mM, and 20 mM, respectively, in reactions I and II in Table 11). In reaction III 5 mM of donor were incubated with 15 mM of 4NP-β-D-GlcNAc. Instead, TmD224G mutant (38 µg), was incubated in the same buffer for 16 h at 70℃ with 10 mM of β-L-Fuc-N_3 donor and 100 mM 4NP-β-D-Xyl as acceptor (reaction IV in Table 11). The transfucosylation efficiency, defined as the amount of fucose transferred to an acceptor different from water, was measured by HPAEC-PAD and calculated as described in the Experimental procedures, SsD242S mutant showed the highest transfucosylation efficiency to the 4NP-β-D-GlcNAc acceptor (86%) followed by 4NP-β-D-Xyl (50%), and 4NP-β-D-Gal (26%) (Figure 31), instead the TmD224G mutant promoted the transfer of fucose to the 4NP-β-D-Xyl acceptor with the highest efficiency (92%) (Figure 32).

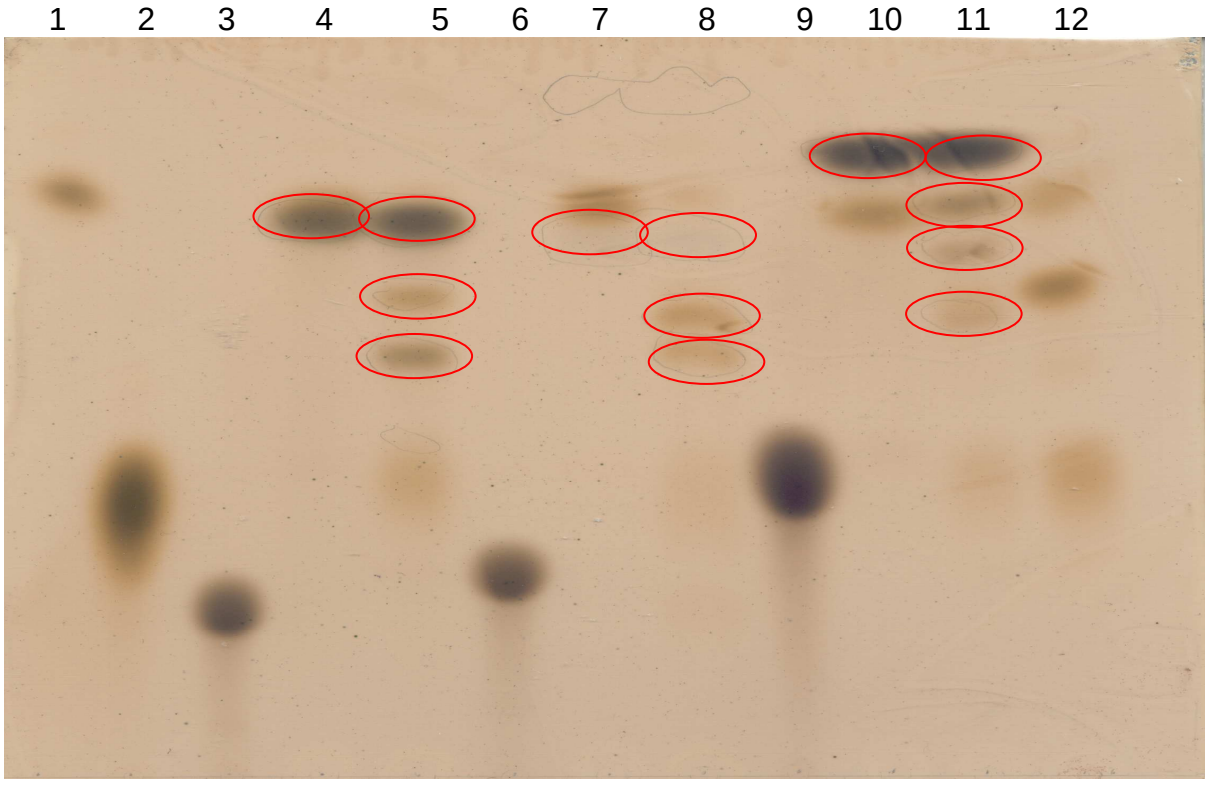


Figure 31: TLC of SsD242S synthesis analyzed reactions. Line 1: standard β-L-fuc azide (loaded 3µL from a 40 mM solution); Line 2: standard fucose (loaded 5µL from a solution 200mM). Line 3-6-9: standard galactose, glucose, xylose respectively; Line 4-7-10: blanks of the reaction with 4NP-β-D-Gal, 4NP-β-D-GlcNAc and 4NP-β-D-Xyl respectively (loaded 20µL). Line 5-8-11: Reaction with 4NP-β-D-Gal, 4NP-β-D-GlcNAc and 4NP-β-D-Xyl respectively (loaded 20µL). Line12: reaction with only β-L-Fuc-N_3 (loaded 20µL). UV-visible signals are indicated in red.

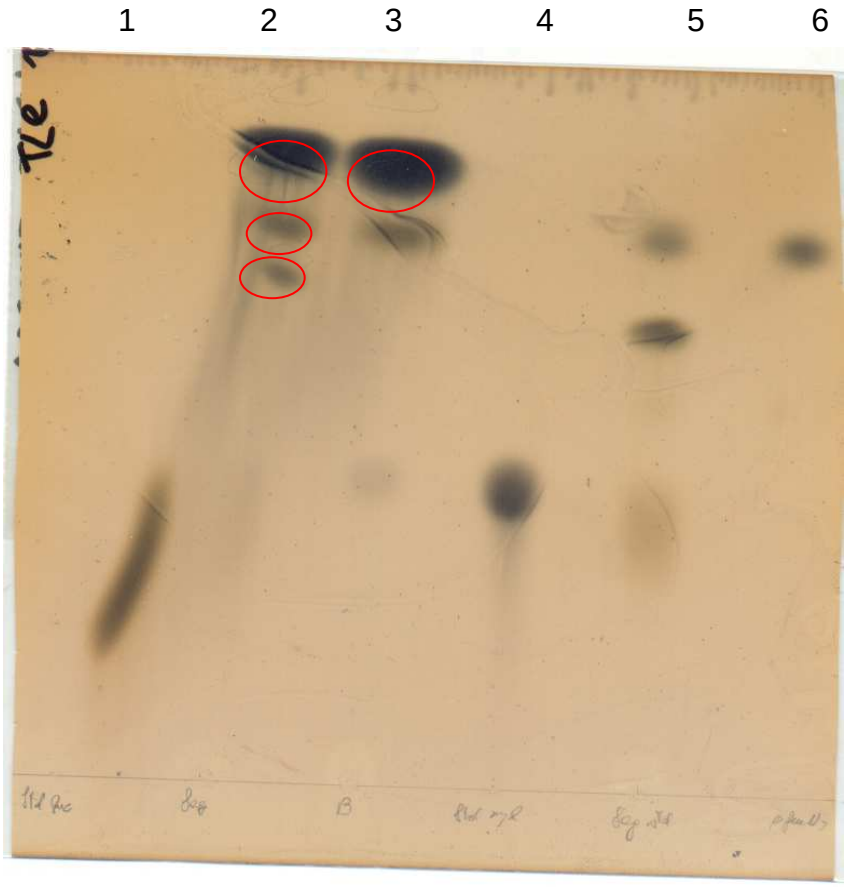


Figure 32: TLC of TMD224G synthesis analyzed reaction. Line 1-4: standard fucose and xylose respectively (loaded 5µL from a solution 200mM). Line 2: TMD224G reaction with 4NP-β-D-Xyl (loaded 20 µL).Line 3: Blank of the TMD224G reaction with 4NP-β-D-Xyl (loaded 20µL). Line 5: SsD242S reaction with only β-L-Fuc-N_3 (loaded 20µL). Line 6: standard β-L-fuc azide (loaded 3µL from a 40 mM solution). UV-visible signals are indicated in red.

To analyse the regioselectivity of the two mutants, the products of the reaction were isolated by reverse phase HPLC and each product was identified by MALDI-TOF mass spectrometry, methylation analysis and ^{1}H-NMR spectroscopy. The relative amounts of the different transfucosylation products for each acceptor were obtained from the signals integration of the HPLC chromatogram. the SsD242S mutant always produced the autocondensation product ***2*** and α-L-Fuc-N_3 found in trace amounts. This compound is rather unexpected, and, presumably it results from contaminations resulting from the preparation of the donor.
In these experiments I used donor: acceptor molar ratios ranging between 1:2 to 1:3.4 (see above); inverted ratios, such as 3:1 β-L-Fuc-N_3 : 4NP-β-D-Xyl led to better total transfucosylation efficiency (76% vs 51%), but to higher amounts of α-L-Fuc-(1-3)-β-L-Fuc-N_3 product, therefore, excess of donor was considered detrimental for synthetic purposes.
The SsD242S mutant showed relaxed regioselectivity: in reaction I, the enzyme produces mainly the α-(1-3) linkage while, in reaction II with 4NP-β-D-Gal acceptor, the regioselectivity is switched to the formation of a α-(1-6) bond. Instead, in reaction III the α-(1-3) regioselectivity is very high with the formation of a single transglycosylation compound to the acceptor (***10***).

Remarkably, the SsD242S mutant, with 4NP-β-D-Xyl and 4NP-β-D-Gal acceptors, catalysed the formation of trisaccharides, namely compounds ***5*** and ***9***. The former, presumably, is formed once compound ***3*** compete with ***1*** in the acceptor subsite +1 [Davies et al. 1997], whereas the formation of the trisaccharide ***9***, which showed α-(1-3) and α-(1-2), linkages, might occur when compound ***8*** acts as acceptor. It is worth noting that compound ***6***, which is the most abundant isolated product in reaction II (78%) and contained α-(1-6) linkages, did not lead to trisaccharides. In contrast, compound ***8,*** which is present as only 8% of the transfucosylation products, could act as acceptor leading to the trisaccharide ***9***. These results seem to indicate that the +1 acceptor binding site of the enzyme has strict selectivity for α-(1-3) bond disaccharides, though the autocondensation product ***2*** was not an acceptor of the mutant. Moreover, transfucosylation on disaccharide acceptors always led to branched products as we never observed the transfer of the fucosyl moiety to the fucose at the non-reducing end of the acceptor (compounds ***5*** and ***9***). Presumably the regioselectivity on the disaccharide acceptor, α-(1-4) and α-(1-2) in ***5*** and ***9***, respectively, depends on the nature of the sugar at the reducing end.

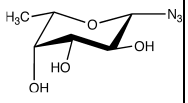
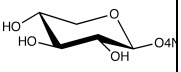
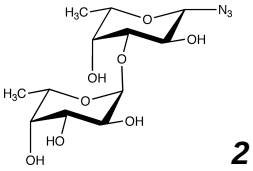
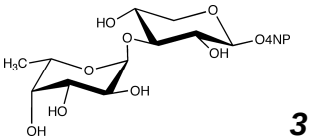
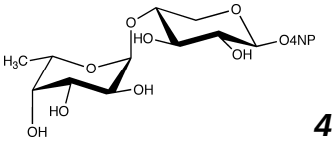
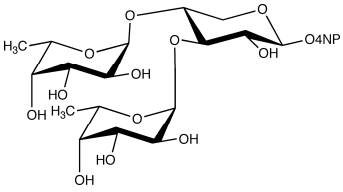
Reaction	Donor	Acceptor	Products	Relative ratios of transfucosylation products (%)
I	 β-L-Fuc-N_3	 4NP-β-D-Xyl	 ***2*** α-L-Fuc-(1-3)-β-L-Fuc-N_3	--
			 ***3*** α-Fuc-(1-3)-β-xyl-4Np	53
			 ***4*** α-Fuc-(1-4)-β-xyl-4Np	31
			 ***5*** α-Fuc-(1-3)-β-xyl-4Np 4 α-fuc(1	16

Table 11 Reaction I: Summary table of the synthetic reaction of SsD242S in presence of β-L-fuc-N_3 like donor and 4NP-β-D-Xyl like acceptor.

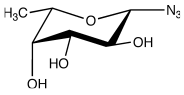
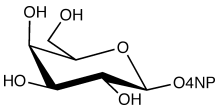
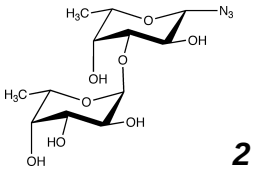
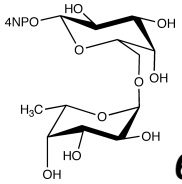
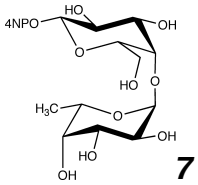
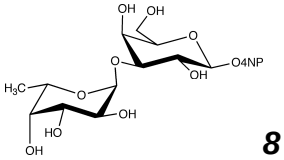
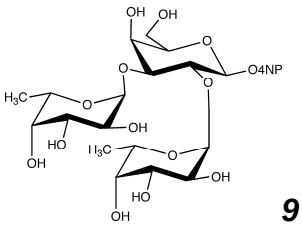
Reaction	Donor	Acceptor	Products	Relative ratios of transfucosylation products (%)
II	 β-L-Fuc-N_3	 4NP-β-D-gal	 ***2*** α-L-Fuc-(1-3)-β-L-Fuc-N_3  ***6*** α-fuc-(1-6)-β-gal-4NP  ***7*** α-fuc-(1-4)-β-gal-4NP  ***8*** α-fuc-(1-3)-β-gal-4NP  ***9*** α-fuc-(1-3)-β-gal-4NP 2 α-fuc-(1	-- 78 10 8 4

Table 11 Reaction II: Summary table of the synthetic reaction of SsD242S in presence of β-L-fuc-N_3 like donor and 4NP-β-D-Gal like acceptor.

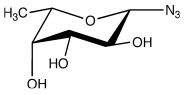
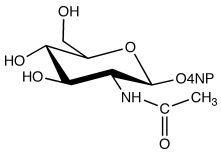
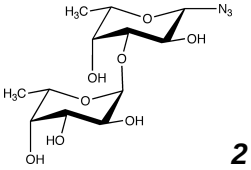
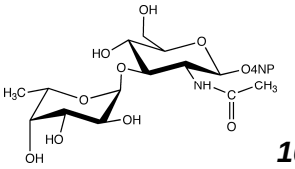
Reaction	Donor	Acceptor	Products	Relative ratios of transfucosylation products (%)
III	 β-L-Fuc-N_3	 4NP-β-D-GlucNAc	 ***2*** α-L-Fuc-(1-3)-β-L-Fuc-N_3  ***10*** α-fuc-(1-3)β-GlcNAc-4NP	-- 100

Table 11 Reaction III: Summary table of the synthetic reaction of SsD242S in presence of β-L-fuc-N_3 like donor and 4NP-β-D-GlcNAc

The TmD224G mutant synthetic activity, tested on 4NP-β-D-Xyl acceptor (reaction IV), led to α-(1-4) and α-(1-3) linkages formation in about 1:1 ratio. However, this mutant did not catalyse autocondensation reactions, and promote the transfer of the fucosyl mojety with extremely high efficiency (92%).

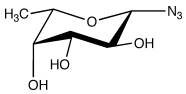
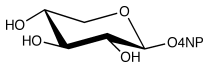
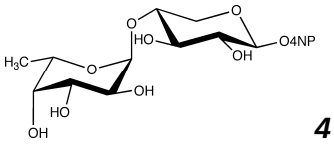
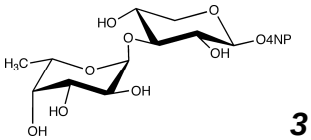
Reaction	Donor	Acceptor	Products	Relative ratios of transfucosylation products (%)
IV	 β-L-Fuc-N_3	 4NP-β-D-Xyl	 ***4*** α-Fuc-(1-4)-β-xyl-4Np  ***3*** α-Fuc-(1-3)-β-xyl-4Np	55 45

Table 11 Reaction IV: Summary table of the synthetic reaction of TmD224G in presence of β-L-fuc-N_3 like donor and 4NP-β-D-Xyl like acceptor.

Discussion

Analysis of expression mechanism of the α-L-fucosidase from S. solfataricus

The identification of genes whose expression is regulated by recoding events is often serendipitous. In the framework of the studies on glycosidases from hyperthermophiles carried out in the laboratory in which I have performed my thesis, we identified in the genome of the archaeon *S. solfataricus* a split gene encoding a putative α-fucosidase, which could be expressed through programmed -1 frameshifting [Cobucci Ponzano et al., 2003a, 2008a]. I tackled this issue by studying the expression of *fucA1* in *S. solfataricus* and in *E. coli* to overcome the problems connected to the scarcity of expression of the α-fucosidase gene and to the manipulation of hyperthermophiles. As already reported by others, in fact, it is a common strategy to study recoding events from different organisms in *E. coli* [Polycarpo et al., 2003].

The expression in *E. coli* of the wild type split gene *fucA1* led to the production by frameshifting of two full-length polypeptides with an efficiency of 5%. This is a value higher than that observed in other genes expressed by translational frameshifting in a heterologous system such as the proteins gpG and gpGT (0.3-3.5%) [Xu et al., 2004]

The gene *fucA1* is expressed in *S. solfataricus* at very low level at the conditions tested. In particular, the transcriptional analysis of the gene revealed that it is approxymately 5,000-fold less abundant than 16S rRNA in both YSM and YGM media. Similarly, no differences in the two media could be found by Western blot probed with anti-Ssα-fuc antibodies indicating that the low expression of the enzyme in *S. solfataricus* is the result of scarce transcription rather than suppressed translation.

Western blots allowed us to identify a specific band about 8.7 kDa heavier than that of the recombinant Ssα-fuc and experiments of translation *in vitro* showed that the wild type gene expresses a full-length polypeptide exhibiting the same molecular mass of the recombinant protein. This demonstrates that the translational machinery of *S. solfataricus* is fully competent to perform programmed frameshifting. It seems likely that the observed discrepancy in molecular mass might arise from post-translational modifications that cannot be produced by the translation *in vitro.* Further experiments are required to characterize the α-L-fucosidase identified in *S. solfataricus*.

MALDIMS and LCMSMS analyses of the products in *E. coli* of the wild type split gene *fucA1* demonstrated that two independent frameshifting events occurred *in vivo* in the proposed slippery site [Cobucci Ponzano et al, 2006]. In particular, the sequences obtained by LCMSMS demonstrate that peptide A results from a simultaneous backward slippage of both the P- and the A-site tRNAs (Figure 33A). Instead, the sequence of peptide B is the result of the re-positioning on the -1 frame of only the P-site tRNA, in fact, the next incorporated amino acid is specified by the codon in the new frame (Figure 33B).

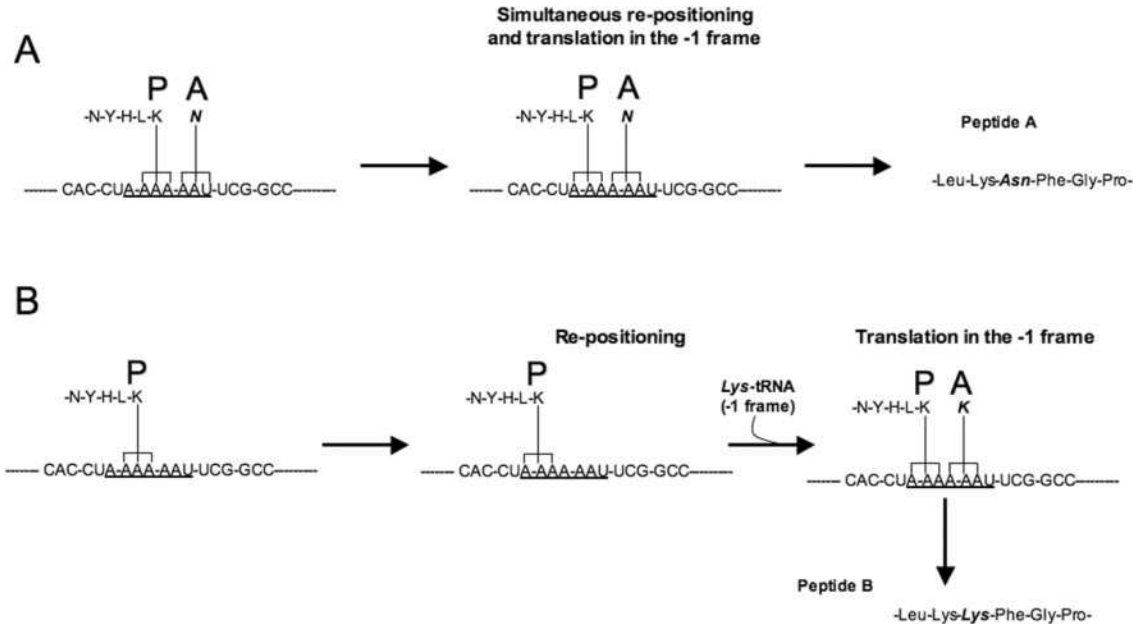


Figure 33: *Proposed mechanisms of programmed -1 frameshifting of fucA1***.** (A) Simultaneous P- and A-site slippage; (B) P-site slippage. For the seek of clarity, the amino acids bound to the tRNA and in the peptides identified by LCMSMS are reported with the one- and the three-letters code, respectively. The aminoacids bound to the tRNAs shown in the A-site are highlighted in bold and italics [Cobucci Ponzano et al, 2006].

Therefore, the expression by -1 frameshifting of the wild type gene *fucA1* in *E. coli* follows the models proposed for ribosomal frameshifting [Baranov et al. 2002]. I confirmed the significance of the slippery heptanucleotide in promoting the programmed frameshifting *in vivo* by mutating the putative regulatory sequence. The triple mutant *fucA1*^{tm} gave no full-length products; presumably, the mutations in both the P- and in the A-site of the slippery sequence dramatically reduced the efficiency of the -1 frameshifting as previously observed in metazoans [Plant and Dinman 2006]. This result confirms that the intact slippery sequence in the wild type gene *fucA1* is absolutely necessary for its expression in *E. coli*. In contrast, surprisingly, the single mutant *fucA1*^{sm} showed an even increased frequency of frameshifting (10%) if compared to the wild type and produced only one polypeptide by shifting specifically in site B [Cobucci Ponzano et al., 2006]. We explained this result observing that the mutation in the P-site of the slippery sequence A-AAA-AAT $\rightarrow$ A-AAG-AAT created a novel slippery sequence A-AAG identical to that controlling the expression by programmed -1 frameshifting of a transposase gene in *E. coli* [Sekine at al., 1994]. Therefore, apparently, the single mutation inactivated the simultaneous P- and A-site tRNA re-positioning and, in the same time, fostered the shifting efficiency of the tRNA in the P-site. It is worth noting that, instead, in *S. solfataricus*, only the simultaneous slippage is effective (Figure 33B) and even the single mutation in the slippery sequence of *fucA1*^{sm} completely annulled the expression of the gene. This indicates that this sequence is essential in the archaeon and that programmed frameshifting in *S. solfataricus* and *E. coli* exploits different mechanisms. Furthermore, since the only difference between the enzymes produced by the frameshifting sites A and B, Ssα-fuc and Ssα-fuc^{B}, respectively, is the stability at 80℃, which is the *S. solfataricus* physiological temperature, the functionality of Ssα-fuc^{B} in the archaeon appears questionable.

The reason why *fucA1* is regulated by programmed -1 frameshifting in not known. However, the physiological significance of programmed frameshifting has been assigned to a minority of the cellular genes while for most of them it is still uncertain [Namy et al., 2004]. This mechanism of recoding is exploited to set the ratio of two polypeptides such as the τ and γ subunits of the DNA polymerase III holoenzyme in *E. coli* [Tsuchihashi et al., 1990]. Alternatively, programmed frameshifting balances the expression of a protein as the bacterial translational release factor 2 and the eukaryotic ornithine decarboxylase antizyme [Namy et al., 2004; Hammell et al., 1999]. In the case of *fucA1,* the polypeptide encoded by the smaller ORF SSO11867 could never be detected by Western blots analyses. In addition, the modelling of Ssα-fuc on the high resolution crystal structure of the α-L-fucosidase from *Thermotoga maritima* [Rosano et al., 2004; Sulzenbacher et al., 2004] showed that the *fucA1* N-terminal polypeptide is not an independent domain. Moreover, it is shown that SSO11867 includes essential catalytic residues [Cobucci Ponzano, et al 2005], excluding the possibility that a functional α-fucosidase can be obtained from the ORF SSO3060 alone. Therefore, several lines of evidence allow us to exclude that programmed -1 frameshifting is used to set the ratio of two polypeptides of the α-fucosidase from *S. solfataricus*. More probably, this translational mechanism might be required to control the expression level of *fucA1*.

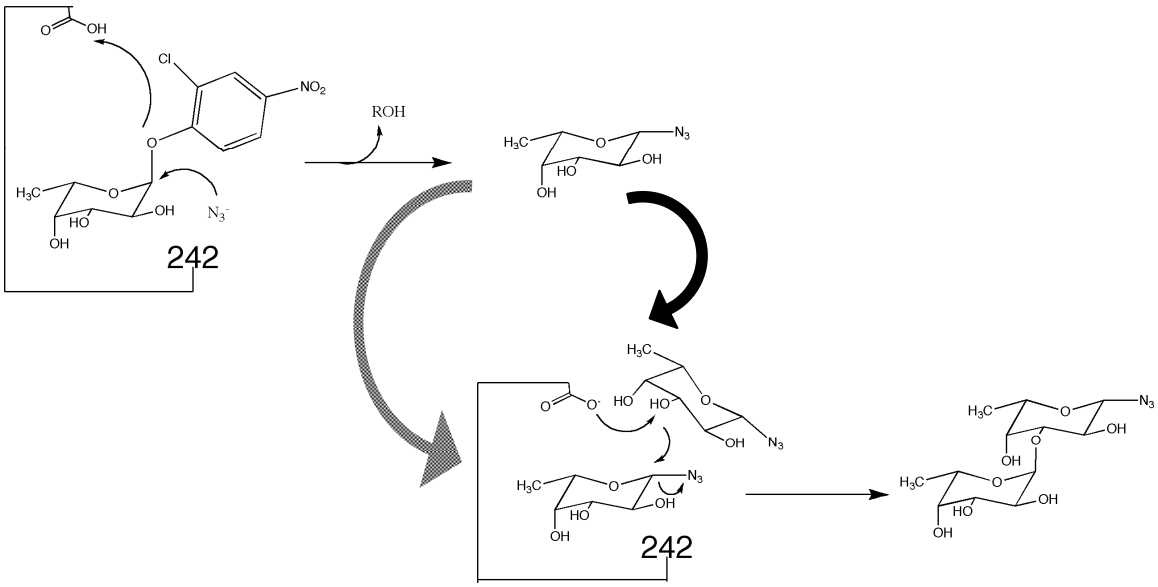
Noticeably, this is the only fucosidase gene regulated by programmed -1 frameshifting. Among carbohydrate active enzymes, the only example of expression through this recoding mechanism is that reported for a gene encoding for a α(1,2)-fucosyltransferase from *Helicobacter pylori* that is interrupted by a -1 frameshifting [Wang et al.,1999]. In this case, the expression by programmed frameshifting would lead to a functional enzyme synthesizing components of the surface lipopolysaccharides to evade the human immune defensive system. It is hard to parallel this model to *fucA1*. Nevertheless, the monosaccharide fucose is involved in a variety of biological functions [Staudacher et al., 1999]. Therefore, the α-L-fucosidase might play a role in the metabolism of fucosylated oligosaccharides; experiments are currently in progress to knockout the wild type *fucA1* gene and to insert constitutive functional mutants of this gene in *S. solfataricus*.

FucA1 is the only archaeal α-L-fucosidase gene identified so far; hence, it is probably the result of a horizontal gene transfer event in *S. solfataricus*. However, since there are no α-fucosidases genes regulated by programmed frameshifting in Bacteria and Eukarya, it is tempting to speculate that this sophisticated mechanism of translational regulation pre-existed in *S. solfataricus* and it was applied to the fucosidase gene for physiological reasons. The identification of other genes interrupted by -1 frameshifts in *S. solfataricus* would open the possibility that they are regulated by programmed -1 frameshifting. Recently, the computational analysis of prokaryotic genomes revealed that seven Archaea harbour interrupted coding sequences, but *S. solfataricus* is not included in this study [Perrodou et al., 2006]. A computational analysis on several archaeal genomes revealing that 34 interrupted genes are present in the genome of *S. solfataricus*, 11 of these genes are composed by two ORFs separated by -1 frameshifting and could be expressed by recoding (B. Cobucci Ponzano., M. Rossi, and M. Moracci, manuscript in preparation).

In this work has shown, for the first time, that programmed -1 frameshifting is present in the Archaea domain. This finding is the missing piece in the puzzle of the phylogenetic distribution of programmed frameshifting demonstrating that this mechanism is universally conserved.

Realization of the first thermophilic α-L-fucosynthase

I have described here the preparation of new mutants in the nucleophile of the reaction of two α-L-fucosidases from the hyperthermophilic prokaryotes *S. solfataricus* and *T. maritima* , namely SsD242S and TmD224G, and their exploitation as efficient α-L-fucosynthases. The approach described here uses, for the first time, β-L-fucopyranosyl azide (β-L-Fuc-N_3) as donor substrate, which allows the transfer of the fucosyl moiety to several glycosides with efficiencies in the range 26-91%. This finding was the result of the reactivation of the mutants in the presence of an activated substrate (2C4NP-α-L-Fuc) and sodium azide.
This ion usually rescues the enzymatic activity of GH mutants in the nucleophile of the reaction by acting as external nucleophile and producing glycosyl-azide. These compounds are usually very stable and can be isolated easily from the reaction mixture and structurally characterized. In particular, the inspection of the anomeric configuration of the glycosyl-azide product, which is inverted if compared to the substrate used, is considered one of the most reliable approaches to assign the role of nucleophile of the reaction to GHs [Ly and Withers 1999; Withers 2001; Zechael and Withers 2001] and it has allowed the unequivocal identification of the active site nucleophile of the reaction in both α- and β-glycosidases [Ly and Withers 1999; Cobucci-Ponzano et al. 2003b]. In alternative to sodium azide, sodium formate, despite its lesser nucleophilicity relative to azide, can also be used to rescue the activity of the mutant. In this case, the glycosyl-formate ester, which is usually much more unstable than glycosyl-azide bond, frequently encounters hydrolysis [Ly and Withers 1999]. However, it was reported in the past that the instability of the glycosyl-formate could be exploited for the synthesis of oligosaccharides by *retaining* β-glycosynthases [Moracci et al. 1998]. In fact, I showed that β-glycosidases mutated in the nucleophile, in the presence of sodium formate and an activated substrate, formed glycosyl-formate, whose stability was sufficient to act as the intermediate of the reaction and allowing the transfer of the glycosyl moiety to a sugar acceptor.
With this in mind, my initial attempts to produce a novel α-fucosynthase aimed to restore the hydrolytic activity of Ssα-fuc mutants in the nucleophile of the reaction in the presence of sodium formate and activated substrates [Cobucci Ponzano et al., 2008b]. However, in these conditions, the formation of free fucose could only be observed with no synthesis of oligosaccharides, suggesting that the β-L-fucosyl formate product expected from the reactivation was rapidly hydrolyzed. In contrast, the chemical rescue of the enzymatic activity of SsD242G/S mutants in the presence of sodium azide and 2C4NP-α-L-Fuc substrate led to the synthesis of β-L-Fuc-N_3 (***1***) and, more importantly, of the disaccharide α-L-Fuc-(1-3)-β-L-Fuc-N_3 (***2***) (Figure 28). In addition, the SsD242S mutant, incubated in the presence of β-L-Fuc-N_3 catalysed the autocondensation reaction yielding the same disaccharide ***2*** (Figure 29). The data reported show that the SsD242S mutant acted as α-fucosynthase by following the mechanism described in Scheme 4. The typical azide rescue of the nucleophile mutant and the mechanism of the *inverting* glycosynthases are shown in the upper and the lower parts, respectively.



Scheme 4: Reaction mechanism of SsD242S

I propose that, when SsD242G/S mutants are incubated in the presence of sodium azide and 2C4NP-α-L-Fuc substrate, the mutants are able to perform the reactions in one pot. Instead, in the presence of β-L-Fuc-N_3 alone, this compound can bind to the active site and it can be ligated by the mutant to a suitable acceptor sugar bound to the aglycon pocket has usually happen for *inverting* glycosynthases (Scheme 4 lower part). The novelty is that this mechanism occurs with fucoside azide while in conventional glycosynthases donor substrates are glycosides activated with fluoride, which acted as small aglycon leaving group [for a review see Williams and Withers 2000].

To the best of our knowledge glycosyl azides have never been used before for glycosynthases. These compounds acted as efficient donors in transglycosylation reactions catalysed by *retaining* GHs, offering the advantage, with respect to conventional aryl-glycosides, of being easily chemically synthesized in high yields, highly water soluble, and leading to the facile purification of the transglycosylation products because azide ion is easily removable [Day and Withers 1986; Fialová et al. 2005]. However, the reactivity of glycosyl-azides is not surprising for fully-functional glycosidases, which promote the classical double displacement mechanism with general acid-dase catalytic assistance. Instead, the molecular reasons behind the ability maintained by the Ssα-fuc mutated in the nucleophile of the reaction in promoting the transfer of the fucose moiety from the β-L-Fuc-N_3 to different acceptors, can be less easily foreseen. A possible explanation can be that the anomeric configuration of fucosyl-azide used here increases its reactivity. It is well known that heteroatomic substituents adjacent to a heteroatom within a cyclohexane ring prefer the *axial* rather than the *equatorial* orientation[1]; this thermodynamic preference is named *anomeric effect* and can explain the higher stability of α-

[1] IUPAC. Compendium of Chemical Terminology, 2nd ed. (the "Gold Book"). Compiled by A. D. McNaught and A.Wilkinson. Blackwell Scientific Publications, Oxford (1997). XML on-line corrected version: http://goldbook.iupac.org (2006-) created by M. Nic, J. Jirat, B. Kosata; updates compiled by A. Jenkins. ISBN 0-9678550-9-8. doi:10.1351/goldbook.

anomers *vs* β-ones. This property supports the experimental observation that α-glycosyl fluorides have much longer half-lives in aqueous solutions if compared to the β-anomers; for instance, the α-D-glucoside fluoride has a 12 days half-life while for β-D-Glc-F is stable for 30 min [Albert et al. 2000]. The stability of these compounds allowed their use as donors in reactions catalysed by β- and α-glucosynthases, respectively [Mackenzie at al. 1998; Okuyama et al. 2002], however, for fucosides, expecially the β-L-fucoside anomers, the fluoride derivatives are much more instable. Therefore, though the bound azide is more stable than fluoride, β-L-Fuc-N_3 might be reactive enough to function as donor in the transfucosylation reactions catalysed by the SsD242S mutant, requiring less acid/base assistance than expected.

I have shown here that also a mutant in the nucleophile of the reaction of the GH29 α-L-fucosidase from *Thermotoga maritima* had α-fucosynthase activity. This confirmed that the approach described is of general application. Nevertheless, the nature of the catalyst is of decisive importance in the preparation of a fucosynthase: with respect to SsD242S, the TmD224G mutant never catalysed autocondensation reactions and the TmD224S mutant was inactive at the conditions tested. In particular, in the enzymes from *S. solfataricus* and *T. maritima*, only Ser and Gly residues, respectively replacing the natural Asp, led to a α-fucosynthase; mutations with other non-nucleophilic amino acids were not effective. These observations indicate that the structure of the catalytic and of the acceptor binding site are also very important in determining the activity of a glycosynthase. Without a detailed 3D-structure of these enzymes it is very difficult to explain these results. The difference in behaviour of the enzymes from the two sources suggests that, though they belong to the same GH family, the architecture of their catalytic sites is different. In addition to this, the two enzymes have also different affinities for β-L-Fuc-N_3 in the acceptor binding site: in SsD242S this compound binds efficiently and the autocondensation reactions occurr also in the presence of other acceptors (see Table 11) while TmD224G mutant binds to β-L-Fuc-N_3 only in the donor site and I never observed the formation of autocondensation products. The differences observed between two enzymes from the same family once again confirm that the production of a novel glycosynthase is a very tricky operation that is very difficult to plan in advance. The exploration of different strategies in order to identify the best donor/acceptor substrates in a particular catalytic context is necessary.

When I scaled-up the fucosynthetic reactions of the SsD242S mutant on selected acceptor glycosides I found that the regiospecificity and the transfucosylating efficiency, in terms of the ability in transferring fucose to the acceptor relative to water, depended on the acceptor used. The highest efficiency was observed on 4NP-β-D-GlcNAc acceptor with exclusive selectivity for the oxydryl group in the C3. It is worth noting that the glycoside produced, which is the disaccharide Fuc-α-L-(1,3)-L-GlcNac rather common in glycoproteins and found in the Lewis^x, Lewis^y, and Sialyl Lewis^x antigens, was obtained with no efforts to improve the yields by testing different reaction conditions. Therefore, this method has great margins of improvement. The synthetic ability of TmD224G was tested by using 4NP-β-D-Xyl as a model acceptor substrate. The transfucosylating efficiency was >90% and about equal amounts of Fuc-α-L-(1,3)/(1,4)-Xyl products were purified from the reaction mixtures.

In conclusion, I reported here for the first time the preparation of two novel *retaining* α-fucosynthases, which followed the classical glycosynthetic approach proposed about ten years ago for β-glycosynthases [Mackenzie et al. 1998], but, for the first time, utilize β-fucosyl azide as donor substrate. Our finding might open new

perspectives in the use of azide derivatives for the production of novel α-glycosynthases. The only α-glucosynthase known so far utilizes as donor the β-Glc-F [Okuyama et al. 2002], which, as stated above, has an half-life of 29-30 min in several aquaeos buffer systems. However, the half-lives of other glycoside fluoride derivatives can be less convenient. For instance, β-D-galactopyranoside- and β-D-mannopyranoside fluorides at the same conditions have half-lives ranging between 6-8 min and 11-17 min, respectively [Albert et al. 2000]. Therefore, the use of fluorinated substrates in enzymatic α-galactosynthetic and α-mannosynthetic reactions would be not feasible for the large spontaneous hydrolysis at the operational conditions. In these regards, azide derivatives of β-D-Gal, –Man, or other sugars might show the right balance between stability and reactivity to work as suitable donors in α-glycosynthetic reactions. Future work in the development of novel glycosynthases should take into account the relative stability of donors derivatised with different chemical groups.

References

Albert M, Repetschnigg W, Ortner J, Gomes J, Paul BJ, Illaszewicz C, Weber H, Steiner W, Dax K. (2000) Simultaneous detection of different glycosidase activities by 19F NMR spectroscopy. *Carbohydr Res*. Aug 7;**327**(4):395-400.

Allen JR, Harris CR, Danishefsky SJ. (2001) Pursuit of optimal carbohydrate-based anticancer vaccines: preparation of a multiantigenic unimolecular glycopeptide containing the Tn, MBr1, and Lewis(y) antigens.*J Am Chem Soc*. Mar 7;**123** (9):1890-97.

Baranov P.V. Gesteland R.F., Atkins J.F. (2002) Recoding: translational bifurcations in gene expression *Gene* **286**, 187-201

Blight, S.K., Larue R.C., Mahapatra A., Longstaff D.G., Chang E., Zhao G., Kang P.T., Green Church K.B., Chan M.K., Krzycki J.A., (2004) Direct charging of tRNA (CUA) with pyrrolysine in vitro and *in vivo*. *Nature*, **431**, 333-35

Bode L., 2006 Recent advances on structure Metabolism and Function of Human milk Oligosaccharides, ***J.Nutr*** *136*: 2127-30.

Bradford, M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding, *Anal. Biochem.* **72**, 248-254.

Burkart, M.D.; Zhang, Z.; Hung, S.-C.; Wong, C.-H. (1997), A new method for the synthesis of fluoro-carbohydrates and glycosides using Selectfluor *J. Am. Chem. Soc.* ***119*****,** 11743-11746

Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B. (2008) The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucleic Acids Res.* Oct 5.

Cobucci Ponzano B., Mazzone.M, Rossi M., Moracci M. (2005) Probing the catalytically essential residues of the α-L-fucosidase from hyperthermophilic archeon Sulfolobus solfataricus *Biochemistry*, vol. **44** No. 16,pagg 6331-42.

Cobucci-Ponzano B, **Conte F**, Benelli D, Londei P, Flagiello A, Monti M, Pucci P, Rossi M, Moracci M (2006) The gene of an archaeal alpha-L-fucosidase is expressed by translational frameshifting. *Nucleic Acids Res* **34**:4258-4268

Cobucci-Ponzano B, **Conte F**, Rossi M, Moracci M. (2008a) The alpha-L-fucosidase from *Sulfolobus solfataricus*. *Extremophiles*. Jan;**12**(1):61-8.

Cobucci-Ponzano B, **Conte F**, Mazzone M, Bedini E, Corsaro MM, Rossi M, Moracci M. (2008b) Design of new reaction conditions for characterization of a mutant thermophilic α-L-fucosidase. *Biocat. Biotrans*, **26**, 18-24.

Cobucci-Ponzano B, Rossi M, Moracci M. (2005b) Recoding in archaea. *Mol Microbiol*. Jan;**55**(2):339-48.

Cobucci-Ponzano B, Trincone A, Giordano A, Rossi M, Moracci M. (2003b) Identification of the catalytic nucleophile of the family 29 alpha-L-fucosidase from *Sulfolobus solfataricus* via chemical rescue of an inactive mutant. *Biochemistry.* Aug 19;**42**(32):9525-31.

Cobucci-Ponzano, B., Trincone, A., Giordano, A., Rossi, M., and Moracci, M. (2003a) Identification of an archaeal alpha-L-fucosidase encoded by an interrupted gene. Production of a functional enzyme by mutations mimicking programmed –1 frameshifting. *J. Biol. Chem.* **278**, 14622-14631.

Condò I., Ciammarucconi A., Benelli D., Ruggero D., Londei P., (1999) Cis acting signals controlling tranlational initiation in the thermophilic archaeon Sulfolobus solfataricus. *Mol. Microbiol*., **34**, 377-84

Day AG, Withers SG. (1986) The purification and characterization of a beta-glucosidase from Alcaligenes faecalis. *Biochem Cell Biol*.;**64**(9):914-22.

Davies GJ, Wilson KS, Henrissat B. (1997) Nomenclature for sugar-binding subsites in glycosyl hydrolases. *Biochem J*. Jan 15;**321** (Pt 2):557-9.

de La Torre F, Sampedro J, Zarra I, Revilla G. (2002) AtFXG1, an Arabidopsis gene encoding alpha-L-fucosidase active against fucosylated xyloglucan oligosaccharides. Plant Physiol 128:247–255.

Dron J., Feng HY., Tellier C., Hoffmann L., Tran ., Robiller C., Dion M (2005) *Thermus thermophilus* glycosyntahase for the efficient synthetis of galactosyl and glucosyl β-1,3-glycosides. *Eur . J. Org. Chem*. 1977-83

Ducros V.M., Tarling C.A., Zechel D.L., Brzozowski A.M., Frandsen T.P., von Ossowski I., Schulein M., Withers S.G., Davies G.J.,(2003) Anatomy of glycosynthesis: structure and kinetics of the Humicola insolens Cel7B E197A and E197S glycosynthase mutants. ,*Chem. Biol. **10*****,*** 619-628.

Farabaugh P.J., (1996) Programmed translational frameshifting *Annu. Rev., Genet*. **30**, 507-528

Fialová, P, Carmona, A.T., Robina I:, Ettrich, R, Sedmera P., Prikrylová, Petrásková-Husaková, L., Kren, V. (2005) Glycosyl azide – a novel substrate for enzymatic transglycosylations. *Tet. Letters*, **46,** 8715-8718.

Geng X, Dudkin VY, Mandal M, Danishefsky SJ. (2004) In pursuit of carbohydrate-based HIV vaccines, part 2: The total synthesis of high-mannose-type gp120 fragments--evaluation of strategies directed to maximal convergence. *Angew Chem Int Ed Engl*. May 3;**43**(19):2562-5.

Hammel A. B., Taylor R. C. Peltz S. W., and Dinman J.D., (1999) Programmed translational frameshifting *Genome Res*. **5**, 417-427

Hancock SM, Vaughan MD, Withers SG. (2006) Engineering of glycosidases and glycosyltransferases. *Curr Opin Chem Biol.* Oct;10, **5:**509-19.

Hao B., Gong W., Ferguson T. K., James C. M., Krzycki J. A., Chan M. K.,(2002) A new UAG-encoded residue in the structure of a methanogen methyltrasferase. *Science* **296**, 1462-66

Honda Y, Fushinobu S, Hidaka M, Wakagi T, Shoun H, Taniguchi H, Kitaoka M. (2008) Alternative strategy for converting an inverting glycoside hydrolase into a glycosynthase. *Glycobiology*. Apr;18 **(4):**325-30.

Honda Y., Kitaoka M., (2006) The first glycosynthase derived from an inverting glycoside hydrolase. *JBC,* **281** (3) , 1426-31

Kempton JB, Withers SG. (1992) Mechanism of Agrobacterium beta-glucopsidase:kiunetic studies. *Biochemistry* **31**: 9961-69

Kim Y-M, Fox DT., Hekmat O., Kantner T, McIntosh LP., Withers SG (2006): Glycosynthase-based synthesis of xylooligosaccharides using an engineered retaining xylanase from Cellulomonas fimi. *Org Biomol Chem;* **4,** 2025-32.

Koshland D.E., (1953) Streochemistry and the mechanism of enzymatic reactions, *Biol. Rev. Camb. Philos.Soc.* **28**, 416-36
Le pendu J (2004) Histo-blood group antigen and human milk oligosaccharides. Gnetic polymorphism and risk of intectious diseases, *Adv. Exp. Med. Biol*. **554**, 135-43

Leatherbarrow, R.J. (1992) GraFit version 3.0, Erithacus Software Ltd., Staines, U.K

Listinsky JJ, Siegal GP, Listinsky CM. (1998). Alpha-L-fucose: a potentially critical molecule in pathologic processes including neoplasia. *Am J Clin Pathol* **110:**425–440.

Ly HD, Withers SG. (1999) Mutagenesis of glycosidases. *Annu Rev Biochem.*;**68**:487-522

Mackenzie L.F.et al., (1998) Glycosynthase: mutant glycosidases for oligosaccaride synthesis. *J. Am. Chem. Soc*. **120**, 5583-84

Mayer C, Zechel DL, Reid SP, Warren RA, Withers SG. (2000) The E358S mutant of Agrobacterium sp. beta-glucosidase is a greatly improved glycosynthase. *FEBS Lett.* Jan 21;**466**(1):40-4.

Mc Carter J.D., Withers S.G., (1994) Mechanism of enzymatic glycoside hydrolisis. *Curr Opin Struct Biol* **4**: 885-892.

McCarter JD, Withers SG (1996) Unequivocal identification of Asp-214 as the catalytic nucleophile of *Saccharomyces cerevisiae* alpha-glucosidase using 5-fluoro glycosyl fluorides. *J Biol Chem* **271**:6889-6894

Moracci M, Trincone A, Perugino G, Ciaramella M, Rossi M. (1998) Restoration of the activity of active-site mutants of the hyperthermophilic beta-glycosidase from

Sulfolobus solfataricus: dependence of the mechanism on the action of external nucleophiles. *Biochemistry* **37:**17262-17270.

Moracci M., La Volpe A., Pulitzer J.F., Rossi M., Ciaramella M., (1992) Expression of thermostable beta-galactosidase gene from the archeobacterium Sulfolobus solftaricus in Saccharomyces cerevisiae and characterization of a new inducible promoter for heterologous expression. *J. Bacteriol*., **174**, 873-82

Mori E, Hedrick JL,Wardrip NJ, Mori T, Takasaki S. (1998) Occurrence of reducing terminal N-acetylglucosamine 3-sulfate and fucosylated outer chains in acidic N-glycans of porcine zona pellucida glycoproteins. *Glycoconj J* **15**:447–456.

Morrow AL, Ruiz-Palacios GM, Altaye M, Jiang X, Guerrero ML, Meinzen-Derr JK, Farkas T, Chaturvedi P, Pickering LK, Newburg DS. (2004) Human milk oligosaccharides are associated with protection against diarrhea in breast-fed infants. *J Pediatr*. Sep;**145**(3):297-303

Mullegger J, Chen H., Chan WY., Reid SP., Jahn M., Warren RAJ, Selleh HM, Withers SG (2006) Thermostable glycosynthase and thioglycoligases derived from thermotoga maritima β-glucuroinidase. *Chem Bio Chem* **7**: 1028-30.

Murata M., Morimoto S., Zeng X., Watanabe S., Usui T., (1999) Enzymatic Synthesis of α-L-fucosidasyl-N-acellactosamines and 3'-O-α-L-fucosyllactose utilizing α-L-fucosidases. *Carbohydr.Res*. **320** (3-4): 192-9

Murata T., Usui T., 2006 Enzymatic Synthesis of Oligosaccharides and Neoglycoconjugates, *Biosci. Biotechnol. Biochem* **70** (5) 1049-59

Namy O., Rousset J.P., Naptine S., Brierley I., (2004) Reprogrammed genetic decoding in cellular gene expression. *Mol. Cell.* **13,** 157-168.

Noda K, Miyoshi E, Uozumi N, Gao CX, Suzuki K, Hayashi N, Hori M, Taniguchi N. (1998). High expression of alpha-1-6 fucosyltransferase during rat hepatocarcinogenesis. *Int J Cancer* **75:**444–4450.

Okuyama M., Mori H., Watanabe K., Kimura A., Chiba S., (2002) α-glucosidase mutant catalyzes α-glyconthase type reaction. *Biosci. Biotechnol. Biochem*. **66**(4) 928-33

Osanjo G, Dion ., Drone J., Solleux C., Tran V., Rabiller C., Tellier C., (2007) Directed evolution of the α-L-fucosidase from Thermotoga maritima into α-L-transfucosidase *Biochemistry* **46**, 1022-33

Perrodou, E., Deshayes, C., Muller, J., Schaeffer, C., Van Dorsselaer, A., Ripp, R., Poch, O., Reyrat, J.M., and Lecompte, O. (2006) ICDS database: interrupted CoDing sequences in prokaryotic genomes. *Nucleic Acids Res.* **34**, 338-343.

Perugino G, Cobucci-Ponzano B, Rossi, M, Moracci, M, (2005) Recent advances in the oligosaccharide synthesis promoted by catalytically engineered glycosidases. *Ad. Synth. Catal*., **347**, 941-950.

Perugino G, Falcicchio P, Corsaro MM, Matsui I, Parrilli M, Rossi M, Moracci M. (2006). Preparation of a glycosynthase from the β-glycosidase of the hyperthermophilic Archaeon Pyrococcus horikoshii*. Biocat Biotrans* **24**:23-29.

Perugino G, Trincone A, Giordano A, van der Oost J, Kaper T, Rossi M, Moracci M. (2003). Activity of hyperthermophilic glycosynthases is significantly enhanced at acidic pH. *Biochemistry* **42**:8484-8493.

Perugino G., Trincone A., Rossi M., Moracci M., (2004) Oligosaccharides synthesis by glycosynthase *Trends in Biotechnology* **22** 1; 31-37

Plant, E.P. and Dinman, J.D. (2006) Comparative study of the effects of heptameric slippery site composition on -1 frameshifting among different eukaryotic systems. *RNA* **12**, 666-673.

Polycarpo C., Ambrogelly A., Ruan B., Tumbulka- Hansen D., Ataide S.F., Ishitani R., Yokoyama S., Nureki O., Ibba M., Soll D., (2003) Activation of the pyrrolysine suppressor tRNA requires formation of a ternary complex with class I and II lysyl-tRNA sinthetases. *Mol. Cell*., **12,** 287-294

Ratner D; Plante OJ; and Seeberger P.H., (2002) A Linear Synthesis of Branched High-Mannose Oligosaccharides from the HIV-1 Viral Surface Envelope Glycoprotein gp120. *Eur. J. Org. Chem* **5**, 826-33

Rosano, C., Zuccotti, S., Cobucci-Ponzano, B., Mazzone, M., Rossi, M., Moracci, M., Petoukhov, M.V., Svergun, D.I., and Bolognesi, M. (2004) Structural characterization of the nonameric assembly of an Archaeal alpha-L-fucosidase by synchrotron small angle X-ray scattering. *Biochem. Biophys. Res. Commun.* **320**, 176-182.

Ruiz-Palacios GM, Cervantes LE, Ramos P, Chavez-Munguia B, Newburg DS. (2003) Campylobacter jejuni binds intestinal H(O) antigen (Fuc alpha 1, 2Gal beta 1, 4GlcNAc), and fucosyloligosaccharides of human milk inhibit its binding and infection. *J Biol Chem*. Apr 18;**278** (16):14112-20.

Sears P. Wong CH (1996) Interventation of carbohydrate recognition by proteins and nucleic acids. *Proc Natl Acad Sci Sa* **93**: 12086-93

Seeberger P.H., (2008) Automated carbohydrate synthesis as platform to address fundamental aspects of glycobiology-current status and future challenges ; *Carbohydrate Research* **343**: 1889-96

Sekine Y., Eisaki N., Ohtsubo E., (1994) Translational control in production of transposase nd in transposition of insertion sequence IS3. *J.Mol. Biol*., **235**, 1406-1420

She Q, Singh RK, Confalonieri F, Zivanovic Y, Allard G, Awayez MJ, Chan-Weiher CC, Clausen IG, Curtis BA, De Moors A, Erauso G, Fletcher C, Gordon PM, Heikamp-de Jong I, Jeffries AC, Kozera CJ, Medina N, Peng X, Thi-Ngoc HP, Redder P, Schenk ME, Theriault C, Tolstrup N, Charlebois RL, Doolittle WF, Duguet

M, Gaasterland T, Garrett RA, Ragan MA, Sensen CW, Van der Oost J**;** (2001) The complete genome of the crenarchaeon Sulfolobus solfataricus P2; *PNAS* July, 3, **98**, No. 14 pp.7835-40

Sinnot ML (1990) Catalytic mechanism of enzymic glycosyl transfer. *Chem Rev* **90**:1171–1202

Slovin SF, Ragupathi G, Adluri S, Ungers G, Terry K, Kim S, Spassova M, Bornmann WG, Fazzari M, Dantis L, Olkiewicz K, Lloyd KO, Livingston PO, Danishefsky SJ, Scher HI. (1999) Carbohydrate vaccines in cancer: immunogenicity of a fully synthetic globo H hexasaccharide conjugate in man.*Proc Natl Acad Sci* U S A. May 11;**96**(10):5710-5

Staudacher, E., Altmann, F., Wilson, I.B., and Marz, L. (1999) Fucose in N-glycans: from plant to man. *Biochim. Biophys. Acta* **1473**, 216-236.

Sulzenbacher G, Bignon C, Nishimura T, Tarling CA, Withers SG, Henrissat B, Bourne Y. (2004) Crystal structure of Thermotoga maritima alpha-L-fucosidase. Insights into the catalytic mechanism and the molecular basis for fucosidosis. *J Biol Chem.* Mar 26;**279**(13):13119-28.

Tarling CA, He S, Sulzenbacher G, Bignon C, Bourne Y, Henrissat B, Withers SG. (2003) Identification of the catalytic nucleophile of the family 29 alpha-L-fucosidase from Thermotoga maritima through trapping of a covalent glycosyl-enzyme intermediate and mutagenesis. *J Biol Chem*. Nov 28;**278**(48):47394-9.

TehanBG, Lloyd FJ, Wong MG, Pitt WR Montana JG, Manallack DT, Gancia E. (2002). Estimation of pKa using semiempirical molecular orbital methods part 1: applicationto phenols and carboxylic acids. *Quant Struct-Act Relat* **21**: 457-72

Trincone A, Giordano A, Perugino G, Rossi M, Moracci M. (2003). Glycosynthase-catalysed syntheses at pH below neutrality. *Bioorg Med Chem Lett* **13**:4039-4042

Trincone A, Perugino G, Rossi M, Moracci M. (2000). A novel thermophilic glycosynthase that effects branching glycosylation. *Bioorg Med Chem Lett* **10**:365-368.

Tsuchihashi, Z., and Kornberg, A. (1990) Translational frameshifting generates the gamma subunit of DNA polymerase III holoenzyme. *Proc. Natl. Acad. Sci. U. S. A.* **87**, 2516-2520.

Umekawa M, Huang W, Li B, Fujita K, Ashida H, Wang LX, Yamamoto K. (2008) Mutants of Mucor hiemalis endo-beta-N-acetylglucosaminidase show enhanced transglycosylation and glycosynthase-like activities. *J Biol Chem.* Feb 22;**283**(8):4469-79

VanhoorenPT and Vandamme EJ (1999) L-fucose occurrence, physiological role , chemical, ernzymatic and microbial synthesis *J. Chem. Technol. Biotechnol*. **74**: 479-497

Varki A. (1993) Biological roles of oligosaccharides: all of theories are correct. *Glycobiology* **2:** 97-130

Verez-Bencomo V, Fernández-Santana V, Hardy E, Toledo ME, Rodríguez MC, Heynngnezz L, Rodriguez A, Baly A, Herrera L, Izquierdo M, Villar A, Valdés Y, Cosme K, Deler ML, Montane M, Garcia E, Ramos A, Aguilar A, Medina E, Toraño G, Sosa I, Hernandez I, Martínez R, Muzachio A, Carmenates A, Costa L, Cardoso F, Campa C, Diaz M, Roy R. (2004) A synthetic conjugate polysaccharide vaccine against Haemophilus influenzae type b. *Science*. Jul 23;**305**(5683):522-5

Viladot JL, Canals F, Batllori X, Planas A. (2001) Long-lived glycosyl-enzyme intermediate mimic produced by formate re-activation of a mutant endoglucanase lacking its catalytic nucleophile. *Biochem J.* Apr 1;**355**(Pt 1):79-86.

Vocadlo D.J., Davies G. J., Laine R, Withers SG (2001) Catalysis by hen egg-white lysozyme proceeds via a covalen intermediate. *Nature* **412**: 835-38

Vocadlo DJ, Mayer C, He S, Withers SG (2000) Mechanism of action and identification of Asp242 as the catalytic nucleophile of *Vibrio furnisii N*-acetyl-beta-D-glucosaminidase using 2-acetamido-2-deoxy-5-fluoro-alpha-L-idopyranosyl fluoride. *Biochemistry* **1**:117-126.

Wada J, Honda Y, Nagae M, Kato R, Wakatsuki S, Katayama T, Taniguchi H, Kumagai H, Kitaoka M, Yamamoto K. 1,2-alpha-l-Fucosynthase: (2008) Oct 6 A glycosynthase derived from an inverting alpha-glycosidase with an unusual reaction mechanism. *FEBS Lett.*

Wang, G., Rasko, D.A., Sherburne, R., and Taylor, D.E. (1999) Molecular genetic basis for the variable expression of Lewis Y antigen in *Helicobacter pylori*: analysis of the alpha (1,2) fucosyltransferase gene. *Mol. Microbiol.* **31**, 1265-1274.

Whitelock JM, Iozzo R.V., (2005) Heparan sulphate: a complex polymer charged with biological activity *Chem Rev* **105**: 2745-64

Williams SJ, Withers SG. (2000) Glycosyl fluorides in enzymatic reactions. *Carbohydr Res*. Jul 10;**327**(1-2):27-46.

Wilting R., Schorling S., Perssons B.C., Bock A., (1997) Selenoprotein synthesis in archaea: identification of an mRNA element of Methanococcus Jannashii probably directing selenocysteine Insertion. *J. Mol. Biol*., **266,** 637-41

Withers, S.G. (2001) Mechanisms of glycosyl transferases and hydrolases. *Carb. Polymers*, **44**, 325-337.

Xu, J., Hendrix, R.W., and Duda, R.L. (2004) Conserved translational frameshift in dsDNA bacteriophage tail assembly genes. *Mol. Cell* **16**, 11-21.

Zechel DL, Withers SG. Dissection of nucleophilic and acid-base catalysis in glycosidases. *Curr Opin Chem Biol*. 2001

Zhou M, O'Doherty G. (2008) The de novo synthesis of oligosaccharides: application to the medicinal chemistry SAR-study of digitoxin. *Curr Top Med Chem*.;**8**(2):114-25.

Acknowledgements

In questi tre anni ho vissuto un'esperienza "particolare" in tutte le sue possibili accezioni. Sono stati ricchi di gioia ma nello stesso momento ho attraversato momenti di grande sconforto e amarezza. Penso di aver superato quest'ultimi sia, con l'appoggio e il sostegno di tante persone e nonostante la presenza di altre, che sicuramente hanno, a loro modo, contribuito alla crescita della mia persona. Ringrazio la **Prof.ssa Simonetta Bartolucci**, per la sua disponibilità e per i suoi preziosi consigli. Ringrazio il **Prof. Mosè Rossi** per avermi dato la possibilità di lavorare nel suo Istituto. Devo in particolar modo ringraziare il **Dott. Marco Moracci**, sia dal punto di vista umano ma soprattutto per quello professionale. Penso di aver imparato da lui cosa vuol dire aver abnegazione e passione per il proprio lavoro ed il suo esempio, qualsiasi sia la mia strada, mi guiderà sempre.
Ringrazio il mio gruppo di lavoro, la **Dott.ssa Beatrice Cobucci Ponzano, Federica Santoro** ed in particolar modo **Andrea**, ops Dott.Andrea Strazzulli un caro amico e un prezioso collega, compagno di tante avventure; per me è sempre stata una spalla a cui appoggiarmi in tutti i momenti (belli e brutti) che io abbia trascorso. Grazie di cuore! Ringrazio inoltre **Francesca**, la mia prima (ed unica?) studentessa ma anche e soprattutto amica alla quale ho cercato di trasmettere ciò che, a mia volta e non molto tempo fa, mi è stato insegnato ed insieme abbiamo imparato tante altre nuove cose… ed anche **Anna** di cui ho sempre ammirato la grande intraprendenza e autonomia. Non vi dimenticherò mai!
Vorrei ringraziare ora tutti i miei **AMICI**, quelli storici che rimarranno per sempre, coloro che pur non sapendo il brutto momento che stavo attraversando mi sono stati vicino: ringrazio **Alessandra, Alessia, Annalisa, Daria, Davide, Diana, Francesca, Ilaria, Marianna, Massimo, Michele, Noemi, Peppe, Sabrina, Sandro & Sandra, Sara, Simona, Valeria e tutti i ragazzi dell'AC.**
Ringrazio **le nonne Anna e Carmela, gli zii e cugini, zia Rita, Valerio** per il supporto in questa mia vita per così dire "scolastica". Dedico inoltre questo mio lavoro ai **miei nonni Eduardo e Giovanni** e mio **zio Corrado**: penso che sarebbero stati contenti di me!
Per i miei genitori e Fiammetta va un ringraziamento particolare: non ci sono parole per descrivere l'affetto e la riconoscenza che provo per loro, posso solo dire GRAZIE!
…e come dicono gli inglesi *the last but not least*, vorrei dedicare questo lavoro alla persona con cui spero inizierà la mia nuova vita: **David**, grazie mille per tutto! Per aver gioito con me nei momenti belli (tanti) e per non avermi mai fatto mai mancare il tuo affetto e il tuo amore soprattutto nei momenti più bui.

Nel cuore dell'inverno ho finalmente imparato
che dentro di me è sopita un'estate invincibile.
A.Camus

List of Publications

Cobucci Ponzano B., **Conte F**., Londei P., Flagiello A., Monti M., Pucci P., Rossi M., Moracci M**.,** "The gene encoding the α-L-fucosidase from archaeon *Sulfolobus solfataricus* is expressed by translational frameshifting" *Nucleic Acid Research* 2006, 34(15):4258-68*.*

Cobucci Ponzano B., **Conte F**., Rossi M., Moracci M., "The α-L-fucosidase from *Sulfolobus solfataricus*"; *Extremophiles* 2008 Jan;12(1):61-8.;

Cobucci-Ponzano B., **Conte F**., Mazzone M., Bedini E., Corsaro M.M., Rossi M., Moracci M., "Design of new reaction conditions for the characterization of a mutant thermophilic α-L-fucosidase" *Biocat Biotrans Jan-Feb 2008, 26(1-2): 18-24*

Meeting Communications

Pucci P., Monti M., Flagiello A., Cobucci Ponzano B.**, Conte F.** and Moracci M. (2005) "The Elucidation of the Mechanism of the expression in *E. coli* of an Archaeal Interrupted Gene by a Proteomic Approach". HUPO 4th Annual World congress. August 28-September 1, Munchen (Germany)

Cobucci Ponzano B., **Conte F**., Benelli D., Londei P., Flagello A., Monti M., Pucci P., Rossi M., Moracci M. (2006) "The first archaeal gene expressed by translational frameshifting" Proteine 2006, Novara 1-3 giugno, P16, p 80

Cobucci Ponzano B, **Conte F**, Rossi M., Moracci M., (2006) "Characterization of the reaction mechanism of first α-L-fucosidase" Extremophiles 2006 17-21 Sept, Brest, France, L17, p 70

Moracci M., Cobucci Ponzano B., **Conte F**., Benelli D., Londei P., Flagello A., Monti M., Pucci P., Rossi M., (2006) "The first evidence of -1 programmed frameshifting in archaea" Extremophiles 2006, 17-21 Sept, Brest, France, P179, p 306

Cobucci Ponzano B., **Conte F**., Capasso C., Rossi M., Moracci M., (2007) "Structure/function analysis of the first GH38 α-mannosidase from archaea" 7th Carbohydrate Bioengineering Meeting; 2007, 22-25 April , Braunschweig, Germany, P88, p.153.

4258–4268 *Nucleic Acids Research, 2006, Vol. 34, No. 15* *Published online 18 August 2006*
doi:10.1093/nar/gkl574

The gene of an archaeal α-L-fucosidase is expressed by translational frameshifting

Beatrice Cobucci-Ponzano[1], Fiorella Conte[1], Dario Benelli[2], Paola Londei[2], Angela Flagiello[3], Maria Monti[3], Piero Pucci[3], Mosè Rossi[1,4] and Marco Moracci[1,*]

[1]Institute of Protein Biochemistry—Consiglio Nazionale delle Ricerche, Via P. Castellino 111, 80131 Naples, Italy, [2]Dipartimento di Biochimica Medica e Biologia Medica (DIBIM), Università di Bari—Piazzale Giulio Cesare, 70124 Bari, Italy, [3]CEINGE Biotecnologie Avanzate s.c.a r.l., Dipartimento di Chimica Organica e Biochimica, Università di Napoli Federico II, Complesso Universitario di Monte S. Angelo, via Cinthia 4, 80126 Napoli, Italy and [4]Dipartimento di Biologia Strutturale e Funzionale, Università di Napoli 'Federico II', Complesso Universitario di Monte S. Angelo, Via Cinthia 4, 80126 Naples, Italy

Received April 19, 2006; Revised July 21, 2006; Accepted July 22, 2006

ABSTRACT

The standard rules of genetic translational decoding are altered in specific genes by different events that are globally termed recoding. In Archaea recoding has been unequivocally determined so far only for termination codon readthrough events. We study here the mechanism of expression of a gene encoding for a α-L-fucosidase from the archaeon *Sulfolobus solfataricus* (*fucA1*), which is split in two open reading frames separated by a −1 frameshifting. The expression in *Escherichia coli* of the wild-type split gene led to the production by frameshifting of full-length polypeptides with an efficiency of 5%. Mutations in the regulatory site where the shift takes place demonstrate that the expression *in vivo* occurs in a programmed way. Further, we identify a full-length product of *fucA1* in *S.solfataricus* extracts, which translate this gene *in vitro* by following programmed −1 frameshifting. This is the first experimental demonstration that this kind of recoding is present in Archaea.

INTRODUCTION

Translation is optimally accurate and the correspondence between the nucleotide and the protein sequences are often considered as an immutable dogma. However, the genetic code is not quite universal: in certain organelles and in a small number of organisms the meaning of different codons has been reassigned and all the mRNAs are decoded accordingly. More surprisingly, the standard rules of genetic decoding are altered in specific genes by different events that are globally termed recoding (1). In all cases, translational recoding occurs in competition with normal decoding, with a proportion of the ribosomes not obeying to the 'universal' rules. Translational recoding has been identified in both prokaryotes and eukaryotes. It has crucial roles in the regulation of gene expression and includes stop codon readthrough, ribosome hopping and ±1 programmed frameshifting [for reviews see (2–4)].

In stop codon readthrough a stop codon is decoded by a tRNA carrying an unusual amino acid rather than a translational release factor. Specific stimulatory elements downstream to the stop codon regulate this process (5). Hopping, in which the ribosome stops translation in a particular site of the mRNA and re-start few nucleotides downstream, is a rare event and it has been studied in detail only in the bacteriophage T4 (6). In programmed frameshifting, ribosomes are induced to shift to an alternative, overlapping reading frame 1 nt 3′-wards (+1 frameshifting) or 5′-wards (−1 frameshifting) of the mRNA. This process is regulated and its frequency varies in different genes. The ±1 programmed frameshifting has been studied extensively in viruses, retrotransposons and insertion elements for which many cases are documented (7–9). Instead, this phenomenon is by far less common in cellular genes. A single case of programmed +1 frameshifting is known in prokaryotes (10,11) while in eukaryotes, including humans, several genes regulated by this recoding event have been described previously [(4) and references therein]. Compared to +1 frameshifting, −1 frameshifting is less widespread with only two examples in prokaryotes (12–14) and few others in eukaryotes (15–17).

The programmed −1 frameshifting is triggered by several elements in the mRNA. The slippery sequence, showing the X-XXY-YYZ motif, in which X can be any base, Y is usually A or U, and Z is any base but G, has the function of favouring the tRNA misalignment and it is the site where the shift takes place (3,18). Frameshifting could be further stimulated by other elements flanking the slippery sequence: a codon for a low-abundance tRNA, a stop codon, a Shine–Dalgarno sequence and an mRNA secondary structure. It has been

*To whom correspondence should be addressed. Tel: +39 081 6132271; Fax: +39 081 6132277; Email: m.moracci@ibp.cnr.it

reported that these elements, alone or in combination, enhance frameshifting by pausing the translating ribosome on the slippery sequence (4,18).

Noticeably, known cases of recoding in Archaea [recently reviewed in (19)] are limited to termination codon readthrough events that regulate the incorporation of the 21st and 22nd amino acids selenocysteine and pyrrolysine, respectively (20–23).

No archaeal genes regulated by translational programmed frameshifting and ribosome hopping have been identified experimentally so far; therefore, if compared with the others domains of life, the study of translational recoding in Archaea is still at its dawn.

We showed that the α-L-fucosidase gene from the crenarchaeon *Sulfolobus solfataricus* is putatively expressed by programmed −1 frameshifting (24). This gene, named *fucA1*, is organized in the open reading frames (ORFs) SSO11867 and SSO3060 of 81 and 426 amino acids, respectively, which are separated by a −1 frameshifting in a 40 base overlap (Figure 1A). We have reported previously that the region of overlap between the two ORFs had the characteristic features of the genes expressed by programmed −1 frameshifting including a slippery heptanucleotide A-AAA-AAT (codons are shown in the zero frame) flanked by a putative stem–loop and the rare codons CAC (Figure 1A) resembling the prokaryotic stem–loops/hairpins and the Shine–Dalgarno-like sites (24). We showed that the frameshifting, obtained by mutating by site-directed mutagenesis the *fucA1* gene exactly in the position predicted from the slippery site, produced a full-length gene, named *fucA1*^{A}, encoding for a polypeptide of 495 amino acids (Figure 1B). This mutant gene expressed in *Escherichia coli* a fully functional α-L-fucosidase, named Ssα-fuc, which was thermophilic, thermostable and had an unusual nonameric structure (24,25). More recently, we determined the reaction mechanism and the function of the residues of the active site of the mutant enzyme (26,27).

The functionality of the product of the mutant gene *fucA1*^{A} does not provide direct experimental evidence that programmed −1 frameshifting occurs *in vivo* and in *S.solfataricus*. To address these issues, we report here the study of the expression of the wild-type split gene *fucA1* and of its mutants in the slippery sequence. We demonstrate here that *fucA1* is expressed by programmed −1 frameshifting in both *E.coli* and *S.solfataricus*. This is the first experimental demonstration that this kind of recoding is present in the Archaea domain of life. The relevance of programmed −1 frameshifting in Archaea is also discussed.

Figure 1. The α-fucosidase gene. (**A**) Region of overlap in the wild-type split *fucA1* gene. The N-terminal SSO11867 ORF is in the zero frame, the C-terminal SSO3060 ORF, for which only a fragment is shown, is in the −1 frame. The slippery heptameric sequence is underlined; the rare codons are boxed and the arrows indicate the stems of the putative mRNA secondary structure. The amino acids involved in the programmed −1 frameshifting and the first codon translated after this event in the −1 frame are shadowed. (**B**) Fragment of the full-length mutant *fucA1*^{A} gene. The small arrows indicate the mutated nucleotides.

MATERIALS AND METHODS

Analysis of the α-fucosidase expression

S.solfataricus cells were grown, and cell extracts obtained, as described previously (24,28).

The expression in the *E.coli* strain BL21(RB791) of the wild-type gene *fucA1* and of the mutant genes $fucA1^{A}$ [previously named FrameFuc in (24)], $fucA1^{B}$, $fucA1^{sm}$ and $fucA1^{tm}$ as fusions of glutathione *S*-transferase (GST) and the purification of the recombinant proteins were performed as reported previously (23). The nomenclature used in this paper for the different α-fucosidase genes is listed in Table 1.

For the western blot studies, equal amounts of *E.coli* cultures expressing the wild-type and mutant *fucA1* genes, normalized for the OD_{600}, were resuspended in SDS–PAGE loading buffer containing 0.03 M Tris–HCl buffer, pH 6.8, 3% SDS (w/v), 6.7% glycerol (w/v), 6.7% 2-mercaptoethanol (w/v) and 0.002% blue bromophenol (w/v). The samples were incubated at 100°C for 5 min (unless otherwise indicated) and were directly loaded on to the gel. Western blot analyses were performed by blotting SDS–PAGEs of the concentrations indicated on Hybond-P polyvinylidenfluorid filters (Amersham Biosciences, Uppsala, Sweden); polyclonal anti-Ssα-fuc antibodies from rabbit (PRIMM, Milan, Italy) and anti-GST antibodies (Amersham Biosciences) were diluted 1:5000 and 1:40 000, respectively. The filters were washed and incubated with the ImmunoPure anti-rabbit IgG antibody conjugated with the horseradish peroxidase (HRP) from Pierce Biotechnology (Rockford, IL, USA). Filters were developed with the ECL-plus Western Blotting Detection system (Amersham Biosciences) by following the manufacturer's indications. The molecular weight markers used in the western blot analyses were the ECL streptavidin–HRP conjugate (Amersham Biosciences).

The protein concentration of the samples was measured with the method of Bradford (29) and the amounts of sample loaded on to the SDS–PAGEs are those indicated. The quantification of the bands identified by western blot was performed by using the program Quantity One 4.4.0 in a ChemiDoc EQ System (Bio-Rad, Hercules, CA, USA) with the volume analysis tool. The frameshifting efficiency was calculated as the ratio of the intensity of the bands of the frameshifted product/frameshifted product + termination product.

The mutants in the slippery sequence of the wild-type gene *fucA1* were prepared by site-directed mutagenesis from the vector pGEX-11867/3060, described previously (24,27). The synthetic oligonucleotides used (PRIMM) were the following: FucA1sm-rev, 5′-TTTAGGTGATATTGGTGTTCTGGTCTATCT-3′; FucA1sm-fwd, 5′-GAACACCAATATCACCTAAAGAATTCGGCCCA-3′; FucA1tm-rev, 5′-AGGTGATATTGGTGTTCTGGTCTATCTGGC-3′; FucA1tm-fwd, 5′-CCAGAACACCAATATCACCTCAAGAACTCGGCCCAGT-3′, where the mismatched nucleotides in the mutagenic primers are underlined. Direct sequencing identified the plasmids containing the desired mutations and the mutant genes, named $fucA1^{sm}$ and $fucA1^{tm}$, were completely re-sequenced.

Expression and characterization of Ssα-fuc^{B}

The mutant Ssα-fuc^{B} was prepared by site-directed mutagenesis from the vector pGEX-11867/3060, by using the same site-directed mutagenesis kit described above. The synthetic oligonucleotides used were FucA1sm-rev (described above) and the following mutagenic oligonucleotide: Fuc-B, 5′-GAACACCAATATCACCTAAAGAAGTTCGGCCCAGT-3′, where the mismatched nucleotides are underlined. Direct sequencing identified the plasmid containing the desired mutations and the mutant gene, named $fucA1^{B}$, was completely re-sequenced. The enzymatic characterization of Ssα-fuc^{B} was performed as described previously (24,27).

Mass spectrometry experiments

Samples of the proteins expressed in *E.coli* from the wild-type gene *fucA1* and the mutants $fucA1^{A}$ and $fucA1^{sm}$, purified as described, were fractionated on an SDS–PAGE. Protein bands were excised from the gel, washed in 50 mM ammonium bicarbonate, pH 8.0, in 50% acetonitrile, reduced with 10 mM DTT at 56°C for 45 min and alkylated with 55 mM iodoacetamide for 30 min at room temperature in the dark. The gel pieces were washed several times with the buffer, resuspended in 50 mM ammonium bicarbonate and incubated with 100 ng of trypsin for 2 h at 4°C and overnight at 37°C. The supernatant containing peptides was analysed by MALDIMS on an Applied Biosystem Voyager DE-PRO mass spectrometer using α-cyano-4-hydroxycynnamic acid as matrix. Mass calibration was performed by using the standard mixture provided by manufacturer.

Liquid chromatography online tandem mass spectrometry (LCMSMS) analyses were performed on a Q-TOF hybrid mass spectrometer (Micromass, Waters, Milford, MA, USA) coupled with a CapLC capillary chromatographic system (Waters). Peptide ions were selected in the collision cell and fragmented. Analysis of the daughter ion spectra led to the reconstruction of peptide sequences.

Experiments of translation *in vitro*

Genomic DNA from *S.solfataricus* P2 strain was prepared as described previously (24). A DNA fragment of 1538 nt containing the complete *fucA1* gene, was prepared by PCR, by using the following synthetic oligonucleotides (Genenco, Florence, Italy): FucA1-fwd, 5′-CTGGAG*GCGCGC***TAATACGACTCACTATAGG**TCAGTTAAATGTCACAAAATTCT-3′; FucA1-rev, 5′-GACTTG*GGCGCGC*CTATCTATAATCTAGGATAACCCTTAT-3′, in which the sequence corresponding to the genome *of S.solfataricus* is underlined. In the FucA1-fwd primer, the sequence of the promoter of the T7 RNA polymerase is in boldface and the sequence of the BssHII site is shown in italics. The PCR amplification was performed as described previously (24) and the amplification products were cloned in the BssHII site of the plasmid pBluescript II KS+. The *fucA1* gene was completely re-sequenced to check if undesired mutations were introduced by PCR and the recombinant vector obtained, named pBluFucA1, was used for translation *in vitro* experiments.

The plasmids expressing the mutant genes $fucA1^{A}$, $fucA1^{sm}$ and $fucA1^{tm}$ for experiments of translation *in vitro* were prepared by substituting the KpnI–NcoI wild-type fragment, containing the slippery site, with those isolated from the mutants. To check that the resulting plasmids had the correct sequence, the mutant genes were completely re-sequenced.

The mRNAs encoding wild-type *fucA1* and its various mutants were obtained by *in vitro* run-off transcription. About 2 μg of each plasmid was linearized with BssHII and incubated with 50 U of T7 RNA polymerase for 1 h 30 min at 37°C. The transcription mixtures were then treated with 10 U of DNAseI (RNAse free) for 30 min. The transcribed RNAs were recovered by extracting the samples twice with phenol (pH 4.7) and once with phenol/chloroform 1:1 followed by precipitation with ethanol. The mRNAs were resuspended in DEPC-treated H_2O at the approximate concentration of 0.6 pmol/μl.

In vitro translation assays were performed essentially as described by Condò *et al*. (28). The samples (25 μl final volume) contained 5 μl of *S.solfataricus* cell extract, 10 mM KCl, 20 mM Tris–HCl, pH 7.0, 20 mM Mg acetate, 3 mM ATP, 1 mM GTP, 5 μg of bulk *S.solfataricus* tRNA, 2 μl of [^{35}S]methionine (1200 Ci/mmol at 10 mCi/ml) and ~10 pmol of each mRNA. The mixtures were incubated at 70°C for 45 min. After this time, the synthesized proteins were resolved by electrophoresis 12.5% acrylamide–SDS gels and revealed by autoradiography of the dried gels on an Instant Imager apparatus.

Transcriptional analysis of *fucA1*

Cells of *S.solfataricus*, strain P2, were grown in minimal salts culture media supplemented with yeast extract (0.1%), casamino acids (0.1%), plus glucose (0.1%) (YGM) or sucrose (0.1%) (YSM). The extraction of total RNA was performed as reported previously (24). Total RNA was extensively digested with DNAse (Ambion, Austin, TX, USA) and the absence of DNA was assessed by the lack of PCR amplification with each sets of primers described below. The RT–PCR experiments were performed as reported previously (24) by using the primers described previously that allowed the amplification of a region of 833 nt (positions 1–833, in which the A of the first ATG codon is numbered as one) overlapping the ORFs SSO11867 and SSO3060 (24).

For real-time PCR experiments total cDNA was obtained using the kit Quantitect RT (Qiagen GmbH, Hilden, Germany) from 500 ng of the same preparation of RNA described above. cDNA was then amplified in a Bio-Rad LightCycler using the DyNAmo HS Syber Green qPCR Kit (Finnzymes Oy, Espoo, Finland). Synthetic oligonucleotides (PRIMM) used for the amplification of a region at the 3′ of the ORF SSO3060 were as follows: 5′-Real: 5′-TAAATGGCGAAGCGATTTTC-3′; 3′-Real: 5′-ATATGCCTTTGTCGCGGATA-3′ for the gene *fucA1*. 5′-GAATGGGGGTGATACTGTCG-3′ and 5′-TTTACAGCCGGGACTACAGG-3′ for the 16S rRNA gene.

For each amplification of the *fucA1* gene was used ~2500-fold more cDNA than that used for the amplification of the 16S rRNA. Controls with no template cDNA were always included. PCR conditions were 15 min at 95°C for initial denaturation, followed by 40 cycles of 10 s at 95°C, 25 s at 56°C and 35 s at 72°C, and a final step of 10 min at 72°C. Product purity was controlled by melting point analysis of setpoints with 0.5°C temperature increase from 72 to 95°C. PCR products were analysed on 2% agarose gels and visualized by ethidium bromide staining.

The expression values of *fucA1* gene were normalized to the values determined for the 16S rRNA gene. Absolute expression levels were calculated as *fucA1*/16S ratio in YSM and YGM cells, respectively. Relative mRNA expression levels (YSM/YGM ratio) were calculated as (*fucA1*/16S ratio in YGM cells)/(*fucA1*/16S ratio in YSM). Each cDNA was used in triplicate for each amplification.

RESULTS

Expression of *fucA1* in *E.coli*

The wild-type *fucA1* gene, expressed in *E.coli* as a GST-fused protein, produced trace amounts of α-fucosidase activity (2.3×10^{-2} units mg^{-1} after removal of GST), suggesting that a programmed −1 frameshifting may occur in *E.coli* (24). The enzyme was then purified by using the GST purification system and analysed by SDS–PAGE revealing a major protein band (Figure 2A). The sample and control bands were excised from the gel, digested *in situ* with trypsin and directly analysed by matrix-assisted laser desorption/ionization mass spectrometry (MALDIMS). As shown in Figure 2B and C, both spectra revealed the occurrence of an identical mass signal at *m*/*z* 1244.6 corresponding to a peptide (Peptide A) encompassing the overlapping region of the two ORFs. This result was confirmed by liquid chromatography online tandem mass spectrometry (LCMSMS) analysis of the peptide mixtures. The fragmentation spectra of the two signals showed the common sequence Asn-Phe-Gly-Pro-Val-Thr-Asp-Phe-Gly-Tyr-Lys in which the amino acid from the ORF SSO11867 is underlined. These results unequivocally demonstrate that the protein containing the Peptide A is produced in *E.coli* by a frameshifting event that occurred exactly within the slippery heptamer predicted from the analysis of the DNA sequence in the region of overlap between the ORFs SSO11867 and SSO3060 (Figure 1A).

Remarkably, the MALDIMS analysis of the products of the wild-type *fucA1* gene revealed the presence of a second Peptide B at *m*/*z* 1258.6 that is absent in the spectra of the Ssα-fuc control protein (Figure 2B and C). The sequence of Peptide B obtained by LCMSMS (Figure 2D) was Lys-Phe-Gly-Pro-Val-Thr-Asp-Phe-Gly-Tyr-Lys. This sequence differs only by one amino acid from Peptide A demonstrating that the interrupted gene *fucA1* expresses in *E.coli* two full-length proteins originated by different −1 frameshifting events. Polypeptide A results from a shift in a site A and it is identical to Ssα-fuc prepared by site-directed mutagenesis (24), suggesting that the expression occurred with the simultaneous P- and A-site slippage. Instead, polypeptide B, named Ssα-fuc^B, is generated by frameshifting in a second site B as the result of a single P-site slippage (Figure 2E).

To measure the global efficiency of frameshifting in the two sites of the wild-type gene *fucA1* we analysed the total extracts of *E.coli* by western blot using anti-GST antibodies (Figure 2F). Two bands with marked different electrophoretic mobility were observed: the polypeptide of 78.7 ± 1.1 kDa migrated like GST-Ssα-fuc fusion and was identified as originated from frameshifting in either site A or B of *fucA1*. The protein of 38.1 ± 1.2 kDa, which is not expressed by the mutant gene *fucA1*^{A} (not shown), had an electrophoretic mobility compatible with GST fused to the polypeptide encoded by the

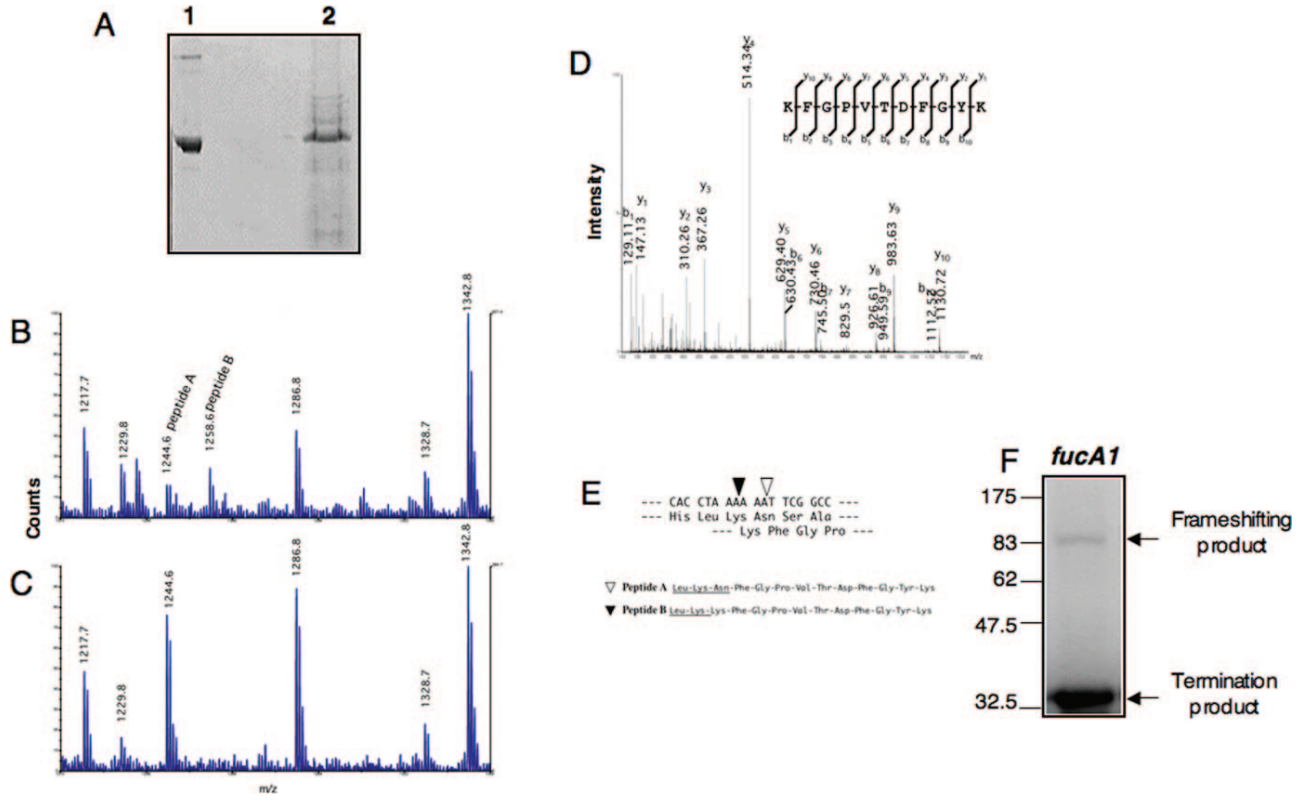


Figure 2. Analysis of the expression of *fucA1* in *E.coli*. (**A**) Coomassie stained 7% SDS–PAGE showing (line 1) the recombinant Ssα-fuc (3 μg) and (line 2) the purified products of the wild-type split *fucA1* gene. The protein concentration of the latter sample could not be quantified because of the scarcity of the purification yields. MALDIMS of the purified products of the wild-type *fucA1* gene and of Ssα-fuc are shown in (**B**) and (**C**), respectively. Peptide A and B are indicated. (**D**) LCMSMS analysis of peptide B. (**E**) The proposed frameshifting sites in the *fucA1* gene. The open and the closed arrows indicate the shifting sites A and B, respectively (for details see text). The sequence of the peptides A and B are also indicated. (**F**) Western blot of *E.coli* cellular extracts expressing the wild-type *fucA1* gene. The blot was probed with anti-GST antibodies. The pre-stained molecular weight markers were β-galactosidase (175 000), paramyosin (83 000), glutamic dehydrogenase (62 000), aldolase (47 500) and triosephosphate isomerase (32 500).

ORF SSO11867 solely (27 and 9.6 kDa, respectively). This polypeptide originated from the translational termination of the ribosome at the OCH codon of the *fucA1* N-terminal ORF (Figure 1A). The calculated ratio of frameshifting to the termination products was 5%.

Preparation and characterization of Ssα-fuc^B

To test if the full-length α-fucosidase produced by the −1 frameshifting event in site B (Ssα-fuc^B), resulting from the single P-site slippage has different properties from Ssα-fuc, whose sequence arises from the simultaneous P- and A- site slippage, we prepared the enzyme by site-directed mutagenesis. The slippery sequence in *fucA1* A-AAA-AAT was mutated in A-AAG-AAG-T where mutations are underlined. The new mutant gene was named *fucA1^B*. The first G, producing the conservative mutation AAA→AAG, was made to disrupt the slippery sequence and hence reducing the shifting efficiency. The second G was inserted to produce the frameshifting that results in the amino acid sequence of Peptide B. Therefore, the sequence of the two full-length mutant genes *fucA1^A* and *fucA1^B* differs only in the region of the slippery sequence: A-AAG-AAT-**TTC-GGC** and A-AAG-AAG-**TTC-GGC**, respectively (the mutations are underlined, the nucleotides in boldface were originally in the −1 frame) (Table 1).

Table 1. Nomenclature and characteristics of the α-fucosidase genes

Gene name	Status	Name of the recombinant protein	Slippery heptamer[a]
fucA1 wild type	−1 frameshifted	—	A-AAA-AAT
fucA1^A mutant	Full-length	Ssα-fuc	A-AAG-AA**T**
fucA1^B mutant	Full-length	Ssα-fuc^B	A-AA**G**-AA**G**
fucA1^{sm} mutant	−1 frameshifted	—	A-AAG-AAT
fucA1^{tm} mutant	−1 frameshifted	—	C-AAG-AAC

[a]Nucleotides modified by substitution and insertion mutations are underlined and in boldface, respectively.

The recombinant Ssα-fuc^B was purified up to ∼95% (Materials and Methods). Gel filtration chromatography demonstrated that in native conditions Ssα-fuc^B had the same non-americ structure of Ssα-fuc with an identical molecular weight of 508 kDa (data not shown). In addition, Ssα-fuc^B had the same high substrate selectivity of Ssα-fuc. The two enzymes have high affinity for 4-nitrophenyl-α-L-fucoside (4NP-Fuc) substrate at 65°C; the K_M is identical within the experimental error (0.0287 ± 0.005 mM) while the k_{cat} of Ssα-fuc^B (137 ± 5.7 s^{-1}) is ∼48% of that of Ssα-fuc (287 ± 11 s^{-1}). In addition, 4-nitrophenyl-α-L-arabinoside, -rhamnoside, 4-nitrophenyl-α-D-glucoside, -xyloside, -galactoside and -mannoside were not substrates of Ssα-fuc^B as shown previously for Ssα-fuc (24). This suggests that the different amino acid sequence did not

significantly affect the active site. Both enzymes showed an identical profile of specific activity versus temperature with an optimal temperature higher than 95°C (data not shown). The heat stability and the pH dependence of Ssα-fuc and Ssα-fuc^B are reported in Figure 3. At 80°C, the optimal growth temperature of *S.solfataricus*, the half-life of Ssα-fuc^B is 45 min, almost 4-fold lower than that of Ssα-fuc (Figure 3A). The two enzymes showed different behaviour at pH <6.0 at which Ssα-fuc^B is only barely active and stable (Figure 3B); however, the two enzymes showed similar values of specific activity at pHs above 6.0, which is close to the intracellular pH of *S.solfataricus (30)*.

Characterization of the slippery sequence of *fucA1* in *E.coli*

The experimental data reported above indicate that the predicted slippery heptanucleotide in the region of overlap between the ORFs SSO11867 and SSO3060 of the wild-type gene *fucA1* could regulate in *cis* the frameshifting events observed in *E.coli*. To test this hypothesis, we mutated the sequence A-AAA-AAT into A-AAG-AAT and C-AAG-AAC (mutations are underlined) obtaining the *fucA1* single mutant (*fucA1^{sm}*) and triple mutant (*fucA1^{tm}*) genes, respectively. It is worth noting that the mutations disrupt the slippery sequence, but they maintain the −1 frameshift between the two ORFs (Table 1).

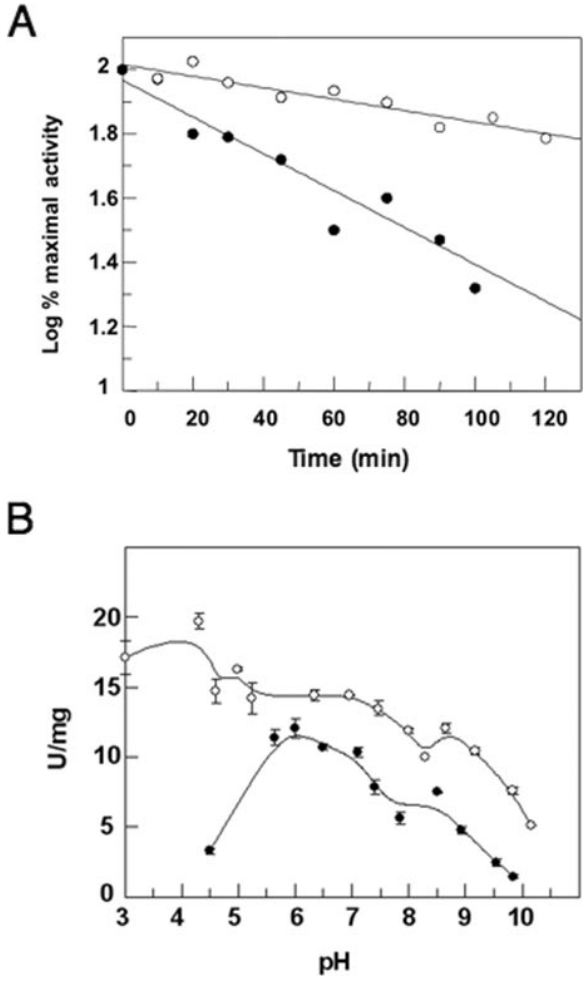


Figure 3. Comparison of the stability and pH dependence of Ssα-fuc and Ssα-fuc^B. (**A**) Thermal stability of Ssα-fuc (open circles) and Ssα-fuc^B (closed circles) at 80°C. (**B**) pH dependence of Ssα-fuc (open circles) and Ssα-fuc^B (closed circles) at 65°C.

Surprisingly, the expression of *fucA1^{sm}* in *E.coli* produced a full-length polypeptide that, after purification by affinity chromatography and removal of the GST protein, showed the same electrophoretic migration of Ssα-fuc and Ssα-fuc^B (Figure 4A). This protein was then characterized by mass spectrometry analyses following *in situ* tryptic digestion. Interestingly, the MALDI spectra revealed the presence of a single peptide encompassing the overlapping region between the two ORFs with a mass value of 1259.7 Da (peptide C; Figure 4B). The sequence of peptide C, determined from the fragmentation spectra obtained by LCMSMS analysis, was Glu-Phe-Gly-Pro-Val-Thr-Asp-Phe-Gly-Tyr-Lys (Figure 4C). Remarkably, apart from the Glu residue, this sequence is identical to that of peptide B produced from *fucA1*, indicating that in the mutant gene *fucA1^{sm}* only one of the two frameshifting events observed in the wild-type *fucA1* gene had occurred. The presence of a Glu instead of Lys was not unexpected. The mutation A-AAA-AAT→A-AAG-AAT in *fucA1^{sm}* was conservative in the zero frame of the ORF SSO11867 (AAA→AAG, both encoding Lys), but it produced the mutation AAA→GAA (Lys→Glu) in the −1 frame of the ORF SSO3060.

It is worth noting that the frameshifting efficiency of the gene *fucA1^{sm}*, calculated by western blot as described above, was 2-folds higher (10%) if compared to *fucA1* (5%) (Figure 4D). This indicates that the mutation cancelled the frameshifting site A and, in the same time, enhanced the frameshifting efficiency of site B.

In contrast, the triple mutant *fucA1^{tm}* produced in *E.coli* only the low molecular weight band resulting from translational termination (Figure 4D). No full-length protein could be detected in western blots probed with either anti-GST (Figure 4D) or anti-Ssα-fuc antibodies (Figure 4E). These data show that the disruption of the heptameric slippery sequence completely abolished the frameshifting in *E.coli* confirming that this sequence has a direct role in controlling the frameshifting *in vivo*.

Expression of fucA1 in *S.solfataricus*

To test whether *fucA1* is expressed in *S.solfataricus* we analysed the extracts of cells grown on yeast extract, sucrose and casaminoacids medium (YSM). Accurate assays showed that *S.solfataricus* extracts contained 3.4×10^{-4} units mg^{-1} of α-fucosidase activity. These very low amounts hampered the purification of the enzyme. The extracts of *S.solfataricus* cells grown on YSM revealed by western blot a band of a molecular mass >97 kDa and no signals were detected with the pre-immune serum confirming the specificity of the anti-Ssα-fuc antibodies (Figure 5A). The different molecular mass may result from post-translational modifications occurred in the archaeon or from the incomplete denaturation of a protein complex. In particular, the latter event is not unusual among enzymes from hyperthermophilic archaea (31,32). To test which hypotheses were appropriate, cellular extracts of *S.solfataricus* were analysed by western blot extending the incubation at 100°C to 2 h. Interestingly, this

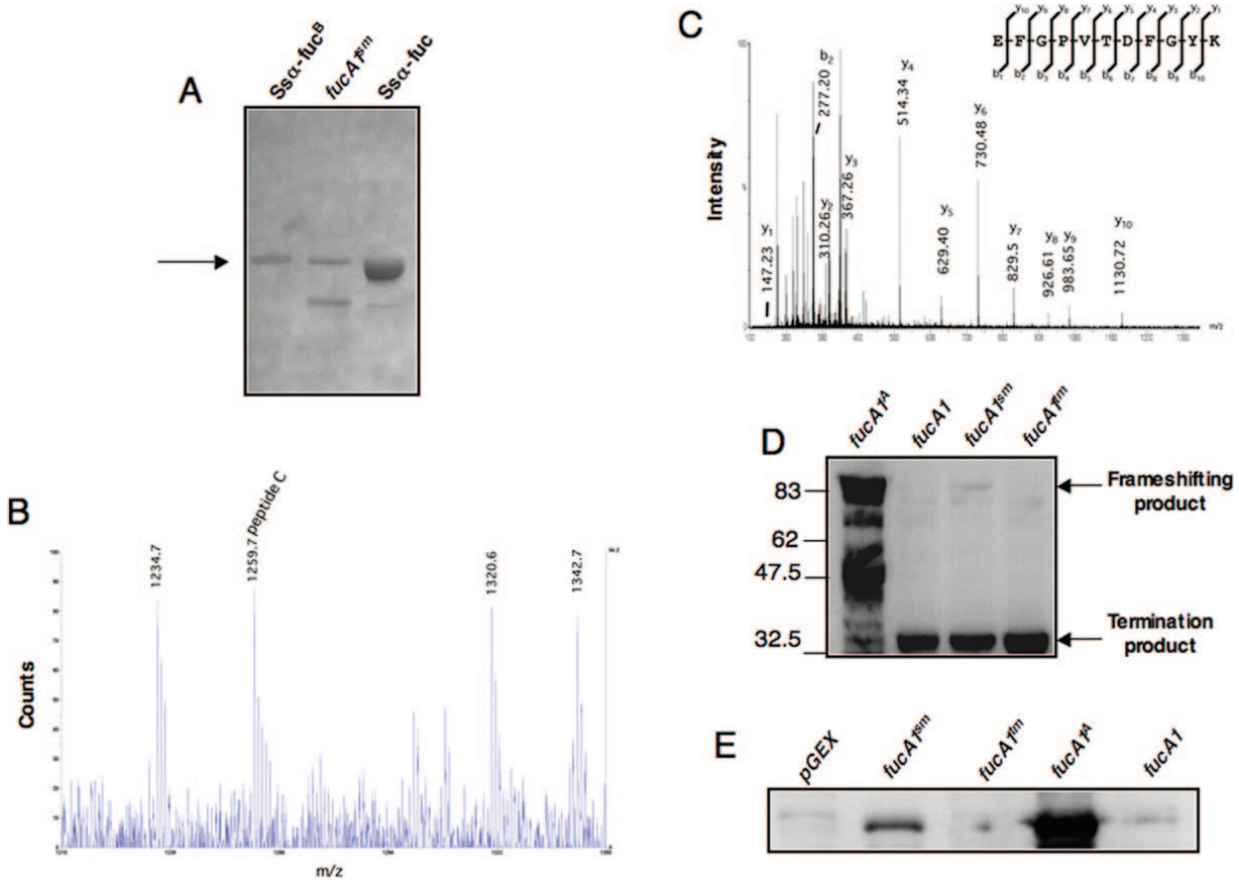


Figure 4. Analysis of the expression in *E.coli* of the mutants in the slippery sequence. (**A**) Coomassie stained 7% SDS–PAGE showing (arrow) the purified recombinant Ssα-fuc^{B} (1.2 μg), the product of the gene *fucA1^{sm}* (2 μg), and Ssα-fuc (4 μg). The bands with faster electrophoretic mobility result from the proteolytic cleavage of the full-length protein (25). (**B**) Partial MALDIMS spectrum of the tryptic digest from mutant *fucA1^{sm}* expressed in *E.coli*. The mass signal corresponding to peptide C encompassing the overlapping region is indicated. (**C**) LCMSMS analysis of peptide C. The amino acid sequence inferred from fragmentation spectra is reported. (**D**) Western blot of *E.coli* cellular extracts expressing *fucA1^{A}*, the wild-type *fucA1*, *fucA1^{sm}* and *fucA1^{tm}* genes (Materials and Methods). The blot was probed with anti-GST antibodies. (**E**) Western blot of partially purified protein samples expressed in *E.coli* fused to GST from wild-type and mutant *fucA1* genes. Cellular extracts were loaded on GST–Sepharose matrix. After washing, equal amounts of slurries (30 μl of 300 μl) were denaturated and loaded on a 8% SDS–PAGE. Extracts of *E.coli* cells expressing the parental plasmid pGEX-2TK were used as the negative control (pGEX). The blot was probed with anti-Ssα-fuc antibodies.

treatment shifted the high-molecular mass band to 67.6 ± 1.2 kDa (Figure 5B and C), which still differs from that of the recombinant Ssα-fuc, 58.9 ± 1.2 kDa, leaving the question on the origin of this difference unsolved. To try to shed some light we immunoprecipitated extracts of *S.solfataricus* with anti-Ssα-fuc antibodies and we analysed the major protein band by MALDIMS. Unfortunately, we could not observe any peptide compatible with the fucosidase because the heavy IgG chain co-migrated with the band of the expected molecular weight (data not shown).

To test if the scarce amounts of the α-fucosidase in *S.solfataricus* extracts was the result of reduced expression at transcriptional level, we performed a northern blot analysis of total RNA extracted from cells grown either on YSM or YGM media. We could not observe any signal by using probes matching the 3′ of the ORF SSO3060 (data not shown). These results suggest that *fucA1* produced a rare transcript; therefore, we analysed the level of mRNA by RT–PCR and by real-time PCR. A band corresponding to the region of overlap between the ORFs SSO11867 and SSO3060 was observed in the RNA extracted from cells grown on YSM

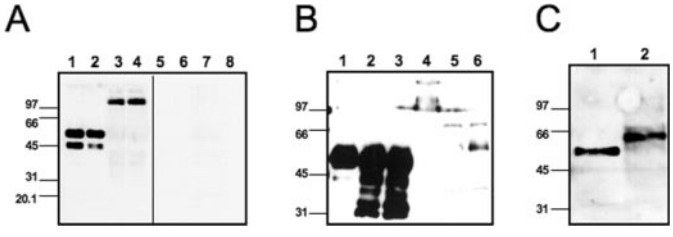


Figure 5. Analysis of the expression of the α-fucosidase in *S.solfataricus*. (**A**) Western blot analysis of recombinant Ssα-fuc (lanes 1, 2, 5 and 6, 0.14 μg) and of extracts of *S.solfataricus* cells grown on YSM (lanes 3, 4, 7 and 8, 153 μg). Samples in lanes 1, 3, 5 and 7 were not denaturated before loading. The left panel shows the blot probed with anti-Ssα-fuc antibodies; the right panel was probed with the pre-immune serum diluted 1:5 000. (**B**) Western blot analysis: recombinant Ssα-fuc (lanes 1, 2 and 3, 0.5 μg) incubated at 100°C for 5 min, 1 h and 2 h, respectively; extracts of *S.solfataricus* cells (lanes 4, 5 and 6, 1 mg) incubated at 100°C for 5 min, 1 h and 2 h, respectively. (**C**) Western blot analysis of recombinant Ssα-fuc (lane 1, 0.1 μg) incubated at 100°C for 5 min and of extracts of *S.solfataricus* cells (lane 2, 1 mg) incubated at 100°C for 2 h, respectively. The molecular weight markers were: phosphorylase b (97 000), albumin (66 000), ovalbumin (45 000), carbonic anhydrase (31 000) and trypsin inhibitor (20 100).

and YGM media, demonstrating that under these conditions the two ORFs were co-transcribed (Figure 6A).

The experiments of real-time PCR shown in Figure 6B demonstrated that rRNA16S was amplified after ~17 cycles while the amplification of *fucA1* mRNA was observed after 38 cycles, despite the fact that we used ~2500-fold more cDNA for the amplification of *fucA1*. This indicates that the gene *fucA1* is transcribed at very low level. No significant differences in the *fucA1* mRNA level were observed in cells grown in YSM or YGM media. This is further confirmed by the analysis by western blot of the extracts of the same cells of *S.solfataricus* used to prepare the total RNAs, which revealed equal amounts of α-fucosidase in the two extracts (Figure 6C). Therefore, the low α-fucosidase activity observed under the conditions tested is the result of the poor transcription of the *fucA1* gene.

Analysis of the expression of *fucA1* in *S.solfataricus* by *in vitro* translation

To determine whether, and with what efficiency, the −1 frameshifting could be performed by *S.solfataricus* ribosomes, mRNAs obtained by *in vitro* transcription of the cloned wild-type *fucA1* gene and the mutants thereof were used to program an *in vitro* translation system prepared as described by Condò *et al*. (28). To this aim, a promoter of T7 polymerase was inserted ahead of the gene of interest to obtain RNA transcripts endowed with the short 5′-untranslated region of 9 nt observed for the natural *fucA1* mRNA (24). Autoradiography of an SDS–PAGE of the translation products (Figure 7) revealed that the wild-type *fucA1* transcript produced a tiny but clear band whose molecular weight corresponded to that of the full-length Ssα-fuc obtained by site-directed mutagenesis (24); the latter was translated quite efficiently in the cell-free system in spite of being encoded by a quasi-leaderless mRNA. Judging from the relative intensity of the signals given by the translation products of the wild-type *fucA1* and the full-length mutant $fucA1^{A}$, the efficiency of the −1 frameshifting in the homologous system was ~10%. No signals corresponding to the polypeptides expected from the separated ORFs SSO11867 and SSO3060 (9.6 and 46.5 kDa, respectively) were observed. However, it should be noted that the product of SSO11867, even if synthesized, is too small to be detected in the gel system employed for this experiment. The larger product of ORF SSO3060, on the other hand, is certainly absent. These data unequivocally demonstrate that the ribosomes of *S.solfataricus* can decode the split *fucA1* gene by programmed −1 frameshifting with considerable efficiency producing a full-length polypeptide from the two ORFs SSO11867 and SSO3060.

Remarkably, under the same conditions at which *fucA1* drives the expression of the full-length protein, we could not observe any product from the $fucA1^{sm}$ and $fucA1^{tm}$ constructs. These data demonstrate that the integrity of the heptanucleotide is essential for the expression of the *fucA1* gene in *S.solfataricus,* thus further confirming that the gene is decoded by programmed −1 frameshifting in this organism. In addition, the lack of expression of $fucA1^{sm}$ by translation *in vitro* in *S.solfataricus* contrasts with the efficient expression of this mutant in *E.coli*, indicating that the two

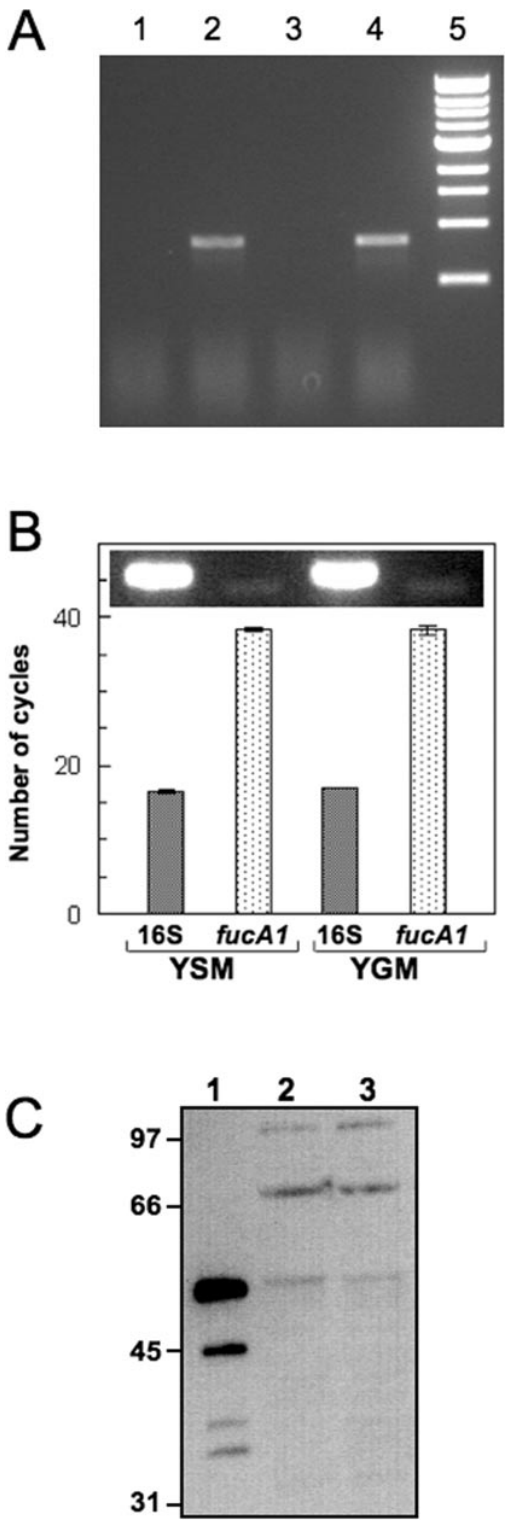


Figure 6. Analysis of the expression of *fucA1* in different media. (**A**) Agarose gel showing the products of RT–PCR encompassing the ORFs SSO11867 and SSO3060 by using total cellular RNA extracted from cells grown in YSM (lanes 1 and 2) and YGM (lanes 3 and 4); lanes 1 and 3, control (amplification of total RNA supplemented with *Taq* and without reverse transcriptase enzyme); lane 2 and 4, *fucA1*. (**B**) Comparison of the *fucA1* mRNA levels in YSM and YGM by real-time PCR. The inset shows the corresponding products found in the real-time PCR visualized by ethidium bromide staining. (**C**) Western blot of *S.solfataricus* extracts of cells grown in YSM (lane 2, 80 μg) and YGM (lane 3, 80 μg). Lane 1 recombinant Ssα-fuc (0.2 μg).

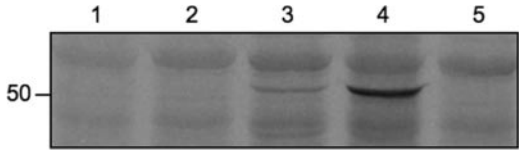


Figure 7. *In vitro* translation. Of each sample 15 µl was loaded on 12.5% acrylamide–SDS gel and the newly synthesized proteins were revealed by autoradiography. Lane 1, no mRNA added; lane 2, *fucA1*^{sm}; lane 3, wild-type *fucA1*; lane 4, full-length *fucA1*^{A}; lane 5, *fucA1*^{tm}.

organisms recognize different sequences regulating the translational frameshifting.

DISCUSSION

The identification of genes whose expression is regulated by recoding events is often serendipitous. In the framework of our studies on glycosidases from hyperthermophiles, we identified in the genome of the archaeon *S.solfataricus* a split gene encoding a putative α-fucosidase, which could be expressed through programmed −1 frameshifting (24). We tackled this issue by studying the expression of *fucA1* in *S.solfataricus* and in *E.coli* to overcome the problems connected to the scarcity of expression of the α-fucosidase gene and to the manipulation of hyperthermophiles. As already reported by others, in fact, it is a common strategy to study recoding events from different organisms in *E.coli* (23,33).

The expression in *E.coli* of the wild-type split gene *fucA1* led to the production by frameshifting of two full-length polypeptides with an efficiency of 5%. This is a value higher than that observed in other genes expressed by translational frameshifting in a heterologous system such as the proteins gpG and gpGT (0.3–3.5%) (33).

The gene *fucA1* is expressed in *S.solfataricus* at very low level under the conditions tested. In particular, the transcriptional analysis of the gene revealed that it is expressed at very low level in both YSM and YGM media. Similarly, no differences in the two media could be found by western blot probed with anti-Ssα-fuc antibodies, indicating that the low expression of the enzyme in *S.solfataricus* is the result of scarce transcription rather than suppressed translation.

Western blots allowed us to identify a specific band ~8.7 kDa heavier than that of the recombinant Ssα-fuc and experiments of translation *in vitro* showed that the wild-type gene expresses a full-length polypeptide exhibiting the same molecular mass of the recombinant protein. This demonstrates that the translational machinery of *S.solfataricus* is fully competent to perform programmed frameshifting. It seems likely that the observed discrepancy in molecular mass might arise from post-translational modifications that cannot be produced by the translation *in vitro*. Further experiments are required to characterize the α-L-fucosidase identified in *S.solfataricus*.

MALDIMS and LCMSMS analyses of the products in *E.coli* of the wild-type split gene *fucA1* demonstrated that two independent frameshifting events occurred *in vivo* in the proposed slippery site. In particular, the sequences obtained by LCMSMS demonstrate that peptide A results from a simultaneous backward slippage of both the P- and the A-site tRNAs (Figure 8A). Instead, the sequence of peptide B is the result of the re-positioning on the −1 frame of only the P-site tRNA; in fact, the next incorporated amino acid is specified by the codon in the new frame (Figure 8B). Therefore, the expression by −1 frameshifting of the wild-type gene *fucA1* in *E.coli* follows the models proposed for ribosomal frameshifting (34). We confirmed the significance of the slippery heptanucleotide in promoting the programmed frameshifting *in vivo* by mutating the putative regulatory sequence. The triple mutant *fucA1*^{tm} gave no full-length products; presumably, the mutations in both the P- and in the A-site of the slippery sequence dramatically reduced the efficiency of the −1 frameshifting as observed previously in metazoans (35). This result confirms that the intact slippery sequence in the wild-type gene *fucA1* is absolutely necessary for its expression in *E.coli*. In contrast, surprisingly, the single mutant *fucA1*^{sm} showed an even increased frequency of frameshifting (10%) if compared to the wild-type and produced only one polypeptide by shifting specifically in site B. We explained this result observing that the mutation in the P-site of the slippery sequence A-AAA-AAT→A-AAG-AAT created a novel slippery sequence A-AAG identical to that controlling the expression by programmed −1 frameshifting of a transposase gene in *E.coli* (36). Therefore, apparently, the single mutation inactivated the simultaneous P- and A-site tRNA re-positioning and, in the same time, fostered the shifting efficiency of the tRNA in the P-site. It is worth noting that, instead, in *S.solfataricus*, only the simultaneous slippage is effective (Figure 8B) and even the single mutation in the slippery sequence of *fucA1*^{sm} completely annulled the expression of the gene. This indicates that this sequence is essential in the archaeon and that programmed frameshifting in *S.solfataricus* and *E.coli* exploits different mechanisms. Furthermore, since the only difference between the enzymes produced by the frameshifting sites A and B, Ssα-fuc and Ssα-fuc^{B}, respectively, is the stability at 80°C, which is the *S.solfataricus* physiological temperature, the functionality of Ssα-fuc^{B} in the archaeon appears questionable.

The reason why *fucA1* is regulated by programmed −1 frameshifting is not known. However, the physiological significance of programmed frameshifting has been assigned to a minority of the cellular genes while for most of them it is still uncertain [see (4) and reference therein; (16)]. This mechanism of recoding is exploited to set the ratio of two polypeptides such as the τ and γ subunits of the DNA polymerase III holoenzyme in *E.coli* (12). Alternatively, programmed frameshifting balances the expression of a protein, as the bacterial translational release factor 2 and the eukaryotic ornithine decarboxylase antizyme [see (4) and (18) and references therein]. In the case of *fucA1,* the polypeptide encoded by the smaller ORF SSO11867 could never be detected by western blots analyses. In addition, the modelling of Ssα-fuc on the high-resolution crystal structure of the α-L-fucosidase from *Thermotoga maritima* (25,37) showed that the *fucA1* N-terminal polypeptide is not an independent domain. Moreover, we have shown recently that SSO11867 includes essential catalytic residues (27), excluding the possibility that a functional α-fucosidase can be obtained from the ORF SSO3060 alone. Therefore, several lines of evidence allow us to exclude that

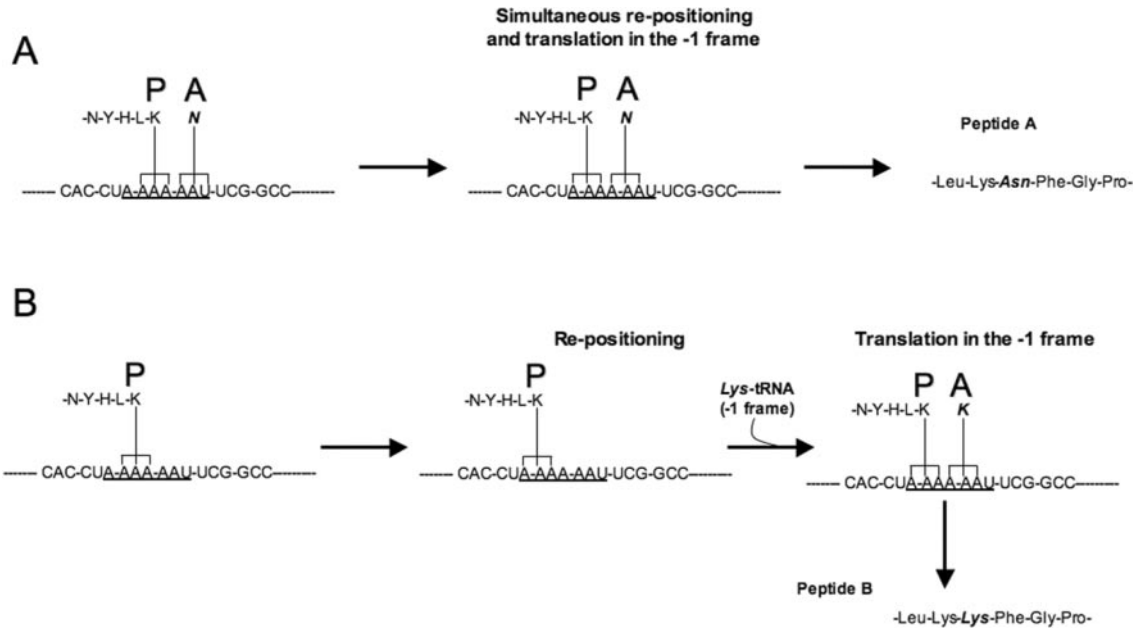


Figure 8. Proposed mechanisms of programmed −1 frameshifting of *fucA1*. (**A**) Simultaneous P- and A-site slippage; (**B**) P-site slippage. For the sake of clarity, the amino acids bound to the tRNA and in the peptides identified by LCMSMS are reported with the one- and the three-letters codes, respectively. The amino acids bound to the tRNAs shown in the A-site are highlighted in boldface and italics.

programmed −1 frameshifting is used to set the ratio of two polypeptides of the α-fucosidase from *S.solfataricus*. More probably, this translational mechanism might be required to control the expression level of *fucA1*.

Noticeably, this is the only fucosidase gene expressed by programmed −1 frameshifting. Among carbohydrate active enzymes, the only example of expression through this recoding mechanism is that reported for a gene encoding for a α(1,2)-fucosyltransferase from *Helicobacter pylori* that is interrupted by a −1 frameshifting (38). In this case, the expression by programmed frameshifting would lead to a functional enzyme synthesizing components of the surface lipopolysaccharides to evade the human immune defensive system. It is hard to parallel this model to *fucA1*. Nevertheless, the monosaccharide fucose is involved in a variety of biological functions (39). Therefore, the α-L-fucosidase might play a role in the metabolism of fucosylated oligosaccharides; experiments are currently in progress to knockout the wild-type *fucA1* gene and to insert constitutive functional mutants of this gene in *S.solfataricus*.

FucA1 is the only archaeal α-L-fucosidase gene identified so far; hence, it is probably the result of a horizontal gene transfer event in *S.solfataricus*. However, since there are no α-fucosidases genes regulated by programmed frameshifting in Bacteria and Eukarya, it is tempting to speculate that this sophisticated mechanism of translational regulation preexisted in *S.solfataricus* and it was applied to the fucosidase gene for physiological reasons. The identification of other genes interrupted by −1 frameshifts in *S.solfataricus* would open the possibility that they are regulated by programmed −1 frameshifting. Recently, the computational analysis of prokaryotic genomes revealed that seven Archaea harbour interrupted coding sequences, but *S.solfataricus* is not included in this study (40). A computational analysis on several archaeal genomes revealed that 34 interrupted genes are present in the genome of *S.solfataricus*, 11 of these genes are composed by two ORFs separated by −1 frameshifting and could be expressed by recoding (B. Cobucci-Ponzano, M. Rossi and M. Moracci, manuscript in preparation).

We have experimentally shown here, for the first time, that programmed −1 frameshifting is present in the Archaea domain. This finding is the missing piece in the puzzle of the phylogenetic distribution of programmed frameshifting demonstrating that this mechanism is universally conserved.

ACKNOWLEDGEMENTS

We are grateful to Maria Carmina Ferrara for technical assistance in the real-time PCR experiments and to Maria Ciaramella and Massimo Di Giulio for useful discussion. We thank the TIGEM-IGB DNA sequencing core for the sequencing of the clones. The IBP-CNR belongs to the Centro Regionale di Competenza in Applicazioni Tecnologico-Industriali di Biomolecole e Biosistemi. P.P., M. R. and M. M. were supported by MIUR project 'Folding di proteine: l'altra metà del codice genetico' RBAU015B47. This work was partially supported by the ASI project MoMa n. 1/014/06/0. Funding to pay the Open Access publication charges for this article was provided by MIUR project RBAU015B47.

Conflict of interest statement. None declared.

REFERENCES

1. Gesteland,R.F., Weiss,R.B. and Atkins,J.F. (1992) Recoding: reprogrammed genetic decoding. *Science*, **257**, 1640–1641.
2. Gesteland,R.F. and Atkins,J.F. (1996) Recoding: dynamic reprogramming of translation. *Annu. Rev. Biochem.*, **65**, 741–768.
3. Farabaugh,P.J. (1996) Programmed translational frameshifting. *Annu. Rev. Genet.*, **30**, 507–528.

4. Namy,O., Rousset,J.P., Napthine,S. and Brierley,I. (2004) Reprogrammed genetic decoding in cellular gene expression. *Mol. Cell*, **13**, 157–168.
5. Bertram,G., Innes,S., Minella,O., Richardson,J. and Stansfield,I. (2001) Endless possibilities: translation termination and stop codon recognition. *Microbiology*, **147**, 255–269.
6. Herr,A.J., Wills,N.M., Nelson,C.C., Gesteland,R.F. and Atkins,J.F. (2004) Factors that influence selection of coding resumption sites in translational bypassing: minimal conventional peptidyl-tRNA: mRNA pairing can suffice. *J. Biol. Chem.*, **279**, 11081–11087.
7. Engelberg-Kulka,H. and Schoulaker-Schwarz,R. (1994) Regulatory implications of translational frameshifting in cellular gene expression. *Mol. Microbiol.*, **1**, 3–8.
8. Sung,D. and Kang,H. (2003) Prokaryotic and eukaryotic translational machineries respond differently to the frameshifting RNA signal from plant or animal virus. *Virus Res.*, **92**, 165–170.
9. Dos Ramos,F., Carrasco,M., Doyle,T. and Brierley,I. (2004) Programmed −1 ribosomal frameshifting in the SARS coronavirus. *Biochem. Soc. Trans.*, **32**, 1081–1083.
10. Craigen,W.J. and Caskey,C.T. (1986) Expression of peptide chain release factor 2 requires high-efficiency frameshift. *Nature*, **322**, 273–275.
11. Baranov,P.V., Gesteland,R.F. and Atkins,J.F. (2002) Release factor 2 frameshifting sites in different bacteria. *EMBO Rep.*, **4**, 373–377.
12. Tsuchihashi,Z. and Kornberg,A. (1990) Translational frameshifting generates the gamma subunit of DNA polymerase III holoenzyme. *Proc. Natl Acad. Sci. USA*, **87**, 2516–2520.
13. Tsuchihashi,Z. and Brown,P.O. (1992) Sequence requirements for efficient translational frameshifting in the *Escherichia coli* dnaX gene and the role of an unstable interaction between tRNA(Lys) and an AAG lysine codon. *Genes Dev.*, **6**, 511–519.
14. Mejlhede,N., Atkins,J.F. and Neuhard,J. (1999) Ribosomal −1 frameshifting during decoding of *Bacillus subtilis* cdd occurs at the sequence CGA AAG. *J. Bacteriol.*, **181**, 2930–2937.
15. Shigemoto,K., Brennan,J., Walls,E., Watson,C.J., Stott,D., Rigby,P.W. and Reith,A.D. (2001) Identification and characterisation of a developmentally regulated mammalian gene that utilises −1 programmed ribosomal frameshifting. *Nucleic Acids Res.*, **29**, 4079–4088.
16. Wills,N.M., Moore,B., Hammer,A., Gesteland,R.F. and Atkins,J.F. (2006) A functional −1 ribosomal frameshift signal in the human paraneoplastic Ma3 gene. *J. Biol. Chem.*, **281**, 7082–7088.
17. Hammell,A.B., Taylor,R.C., Peltz,S.W. and Dinman,J.D. (1999) Identification of putative programmed −1 ribosomal frameshift signals in large DNA databases. *Genome Res.*, **9**, 417–427.
18. Baranov,P.V., Gesteland,R.F. and Atkins,J.F. (2002) Recoding: translational bifurcations in gene expression. *Gene*, **286**, 187–201.
19. Cobucci-Ponzano,B., Rossi,M. and Moracci,M. (2005) Recoding in archaea. *Mol. Microbiol.*, **55**, 339–348.
20. Wilting,R., Schorling,S., Persson,B.C. and Bock,A. (1997) Selenoprotein synthesis in archaea: identification of an mRNA element of *Methanococcus jannaschii* probably directing selenocysteine insertion. *J. Mol. Biol.*, **266**, 637–641.
21. Hao,B., Gong,W., Ferguson,T.K., James,C.M., Krzycki,J.A. and Chan,M.K. (2002) A new UAG-encoded residue in the structure of a methanogen methyltransferase. *Science*, **296**, 1462–1466.
22. Polycarpo,C., Ambrogelly,A., Ruan,B., Tumbula-Hansen,D., Ataide,S.F., Ishitani,R., Yokoyama,S., Nureki,O., Ibba,M. and Söll,D. (2003) Activation of the pyrrolysine suppressor tRNA requires formation of a ternary complex with class I and class II lysyl-tRNA synthetases. *Mol. Cell*, **12**, 287–294.
23. Blight,S.K., Larue,R.C., Mahapatra,A., Longstaff,D.G., Chang,E., Zhao,G., Kang,P.T., Green-Church,K.B., Chan,M.K. and Krzycki,J.A. (2004) Direct charging of tRNA(CUA) with pyrrolysine *in vitro* and *in vivo*. *Nature*, **431**, 333–335.
24. Cobucci-Ponzano,B., Trincone,A., Giordano,A., Rossi,M. and Moracci,M. (2003) Identification of an archaeal alpha-L-fucosidase encoded by an interrupted gene. Production of a functional enzyme by mutations mimicking programmed −1 frameshifting. *J. Biol. Chem.*, **278**, 14622–14631.
25. Rosano,C., Zuccotti,S., Cobucci-Ponzano,B., Mazzone,M., Rossi,M., Moracci,M., Petoukhov,M.V., Svergun,D.I. and Bolognesi,M. (2004) Structural characterization of the nonameric assembly of an Archaeal alpha-L-fucosidase by synchrotron small angle X-ray scattering. *Biochem. Biophys. Res. Commun.*, **320**, 176–182.
26. Cobucci-Ponzano,B., Trincone,A., Giordano,A., Rossi,M. and Moracci,M. (2003) Identification of the catalytic nucleophile of the family 29 alpha-L-fucosidase from *Sulfolobus solfataricus* via chemical rescue of an inactive mutant. *Biochemistry*, **42**, 9525–9531.
27. Cobucci-Ponzano,B., Mazzone,M., Rossi,M. and Moracci,M. (2005) Probing the catalytically essential residues of the alpha-L-fucosidase from the hyperthermophilic archaeon *Sulfolobus solfataricus*. *Biochemistry*, **44**, 6331–6342.
28. Condò,I., Ciammaruconi,A., Benelli,D., Ruggero,D. and Londei,P. (1999) *Cis*-acting signals controlling translational initiation in the thermophilic archaeon *Sulfolobus solfataricus*. *Mol. Microbiol.*, **34**, 377–384.
29. Bradford,M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.*, **72**, 248–254.
30. Lübben,M. and Schäfer,G. (1989) Chemiosmotic energy conversion of the archaebacterial thermoacidophile *Sulfolobus acidocaldarius*: oxidative phosphorylation and the presence of an F0-related *N, N′*-dicyclohexylcarbodiimide-binding proteolipid. *J. Bacteriol.*, **171**, 6106–6116.
31. Moracci,M., La Volpe,A., Pulitzer,J.F., Rossi,M. and Ciaramella,M. (1992) Expression of the thermostable beta-galactosidase gene from the archaebacterium *Sulfolobus solfataricus* in *Saccharomyces cerevisiae* and characterization of a new inducible promoter for heterologous expression. *J. Bacteriol.*, **174**, 873–882.
32. Tachibana,Y., Kuramura,A., Shirasaka,N., Suzuki,Y., Yamamoto,T., Fujiwara,S., Takagi,M. and Imanaka,T. (1999) Purification and characterization of an extremely thermostable cyclomaltodextrin glucanotransferase from a newly isolated hyperthermophilic archaeon, a *Thermococcus* sp. *Appl. Environ. Microbiol.*, **65**, 1991–1997.
33. Xu,J., Hendrix,R.W. and Duda,R.L. (2004) Conserved translational frameshift in dsDNA bacteriophage tail assembly genes. *Mol. Cell*, **16**, 11–21.
34. Baranov,P.V., Gesteland,R.F. and Atkins,J.F. (2004) P-site tRNA is a crucial initiator of ribosomal frameshifting. *RNA*, **10**, 221–230.
35. Plant,E.P. and Dinman,J.D. (2006) Comparative study of the effects of heptameric slippery site composition on −1 frameshifting among different eukaryotic systems. *RNA*, **12**, 666–673.
36. Sekine,Y., Eisaki,N. and Ohtsubo,E. (1994) Translational control in production of transposase and in transposition of insertion sequence IS3. *J. Mol. Biol.*, **235**, 1406–1420.
37. Sulzenbacher,G., Bignon,C., Nishimura,T., Tarling,C.A., Withers,S.G., Henrissat,B. and Bourne,Y. (2004) Crystal structure of *Thermotoga maritima* alpha-L-fucosidase. Insights into the catalytic mechanism and the molecular basis for fucosidosis. *J. Biol. Chem.*, **279**, 13119–13128.
38. Wang,G., Rasko,D.A., Sherburne,R. and Taylor,D.E. (1999) Molecular genetic basis for the variable expression of Lewis Y antigen in *Helicobacter pylori*: analysis of the alpha (1,2) fucosyltransferase gene. *Mol. Microbiol.*, **31**, 1265–1274.
39. Staudacher,E., Altmann,F., Wilson,I.B. and Marz,L. (1999) Fucose in *N*-glycans: from plant to man. *Biochim. Biophys. Acta*, **1473**, 216–236.
40. Perrodou,E., Deshayes,C., Muller,J., Schaeffer,C., Van Dorsselaer,A., Ripp,R., Poch,O., Reyrat,J.M. and Lecompte,O. (2006) ICDS database: interrupted CoDing sequences in prokaryotic genomes. *Nucleic Acids Res.*, **34**, 338–343.

Extremophiles (2008) 12:61–68
DOI 10.1007/s00792-007-0105-y

REVIEW

The α-L-fucosidase from *Sulfolobus solfataricus*

Beatrice Cobucci-Ponzano · Fiorella Conte · Mosè Rossi · Marco Moracci

Received: 6 November 2006 / Accepted: 17 July 2007 / Published online: 9 August 2007

Abstract Glycoside hydrolases form hyperthermophilic archaea are interesting model systems for the study of catalysis at high temperatures and, at the moment, their detailed enzymological characterization is the only approach to define their role in vivo. Family 29 of glycoside hydrolases classification groups α-L-fucosidases involved in a variety of biological events in *Bacteria* and *Eukarya*. In *Archaea* the first α-L-fucosidase was identified in *Sulfolobus solfataricus* as interrupted gene expressed by programmed −1 frameshifting. In this review, we describe the identification of the catalytic residues of the archaeal enzyme, by means of the chemical rescue strategy. The intrinsic stability of the hyperthermophilic enzyme allowed the use of this method, which resulted of general applicability for β and α glycoside hydrolases. In addition, the presence in the active site of the archaeal enzyme of a triad of catalytic residues is a rather uncommon feature among the glycoside hydrolases and suggested that in family 29 slightly different catalytic machineries coexist.

Keywords Glycoside hydrolase · Chemical rescue · Nucleophile · Acid/base · Catalytic triad

Communicated by D.A. Cowan.

B. Cobucci-Ponzano (✉) · F. Conte · M. Rossi · M. Moracci
Institute of Protein Biochemistry,
Consiglio Nazionale delle Ricerche,
Via P. Castellino 111, 80131 Naples, Italy
e-mail: b.cobucciponzano@ibp.cnr.it

M. Rossi
Dipartimento di Biologia Strutturale e Funzionale,
Università di Napoli "Federico II",
Complesso Universitario di Monte S. Angelo,
Via Cinthia 4, 80126 Naples, Italy

Introduction

Carbohydrates serve as structural components and energy source of the cell and are involved in a variety of molecular recognition processes in intercellular communication (Varki 1993; Sears and Wong 1996). Consequently, glycoside hydrolases play important roles in biological systems ranging from the degradation of polysaccharides as food source through to the modification of glycoconjugates on the surfaces of proteins and cells. Among the available glycoside hydrolases, the enzymes from hyperthermophiles are of particular interest for both basic and applied research. In fact, the function of the glycoconjugates identified in hyperthermophiles and of the enzymes involved in their synthesis and degradation is still largely unknown (Lower and Kennelly 2002). In addition, hyperthermophilic glycosidases are interesting model systems in basic research for the study of protein adaptation to heat and, since they catalyze single substrate reactions by following well-known mechanisms, they are the ideal candidates for the study of catalysis at high temperatures. Furthermore, they are particularly appealing for industrial applications as they show peculiar enzymological properties and can withstand the harsh operational conditions adopted in industrial applications. Beside this, the unique substrate specificities or reduced substrate/product inhibition allow the synthesis of new products that are not produced by their mesophilic counterparts (Fischer et al. 1996). On the other hand, the harsh conditions of growing of these organisms have hindered microbiological and genetic studies in vivo; therefore, the isolation of the genes encoding for hyperthermophilic glycosidases and the detailed enzymological characterization of these enzymes is the only approach to define their role in vivo.

Glycoside hydrolases are classified on the basis of their amino acid sequence similarity in about 100 families (http://www.cazy.org) and 14 clans showing conserved three-dimensional (3D) structures. These enzymes follow two distinct mechanisms which are termed *retaining* or *inverting* if the enzymatic cleavage of the glycosidic bond liberates a sugar hemiacetal with the same or the opposite anomeric configuration of the glycosidic substrate, respectively. *Inverting* glycoside hydrolases operate with a one step, single-displacement mechanism with the assistance of a general acid and a general base group in the active site (McCarter and Withers 1994). Instead, *retaining* enzymes follow a two-step mechanism with formation of a covalent glycosyl-enzyme intermediate (Fig. 1). The carboxyl group in the active centre functions as a general acid/base catalyst, and the carboxylate functions as the nucleophile of the reaction (Koshland 1953). In the first step (glycosylation step), the nucleophile attacks the anomeric group of the substrate, while the acid/base catalyst, acting in this step as a general acid, protonates the glycosidic oxygen, thereby assisting the leaving of the aglycon moiety. The concerted action of the two amino acids leads to the formation of a covalent glycosyl-enzyme intermediate (Sinnot 1990; McCarter and Withers 1994). In the second step (deglycosylation step), the glycosyl-enzyme intermediate is cleaved by a water molecule that acts as nucleophile being polarized by the general base catalyst. The product of the reaction retained the anomeric configuration of the substrate. When an acceptor different from water, such as an alcohol or a sugar, intercepts the reactive glycosyl-enzyme intermediate, *retaining* enzymes work in transglycosylation mode; the interglycosidic linkage of the product maintains the same anomeric configuration of the substrate. This property makes the *retaining* glycosyl hydrolases interesting tools for the synthesis of carbohydrates. Despite the differences, the two mechanisms show significant similarities: they both operate via transition states with substantial oxocarbenium ion character and have a catalytic dyad with a pair of carboxylic acids directly involved in catalysis.

The determination of the reaction mechanism and the identification of key active-site residues in glycoside hydrolases are crucial issues to allow the classification of these enzymes (Henrissat and Bairoch 1993; Henrissat and Davies 1997), to unravel the catalytic machinery (McCarter and Withers 1994; Zechel and Withers 2000), and to produce enzymes with novel characteristics (Perugino et al. 2004).

Family 29 of glycoside hydrolases (GH29) groups α-L-fucosidases (EC 3.2.1.51) from plants, vertebrates, and pathogenic microbes of plants and humans (Henrissat 1991). α-L-fucosidases are exo-glycosidases capable of cleaving α-linked L-fucose residues from glycoconjugates, in which the most common linkages are α-(1-2) to galactose and α-(1-3), α-(1-4), and α-(1-6) to *N*-acetyl-glucosamine residues. These compounds are involved in a variety of biological events as growth regulators and as the glucidic part of receptors in signal transduction, cell–cell interactions, and antigenic response (Vanhooren and Vandamme 1999). The central role of fucose derivatives in biological processes explains the interest in α-L-fucosidase and fucosyl-transferase activities. α-L-fucosidases in higher plants and in mammals are associated with different mechanisms of cell growth and regulation, since they are involved in the modification of fucosylated glucans (Staudacher et al. 1999). In plants, α-L-fucosylated oligosaccharides derived from xyloglucan have been shown to regulate auxin- and acid pH-induced growth (de La Torre et al. 2002). In mammals, oligosaccharides containing fucose are reported to play important roles in a variety of physiological and pathological events (Xiang and Bernstein 1992; Wiese et al. 1997; Listinsky et al. 1998; Mori et al. 1998; Noda et al. 1998; Russell et al. 1998; Michalski and Klein 1999; Rapoport and Pendu 1999).

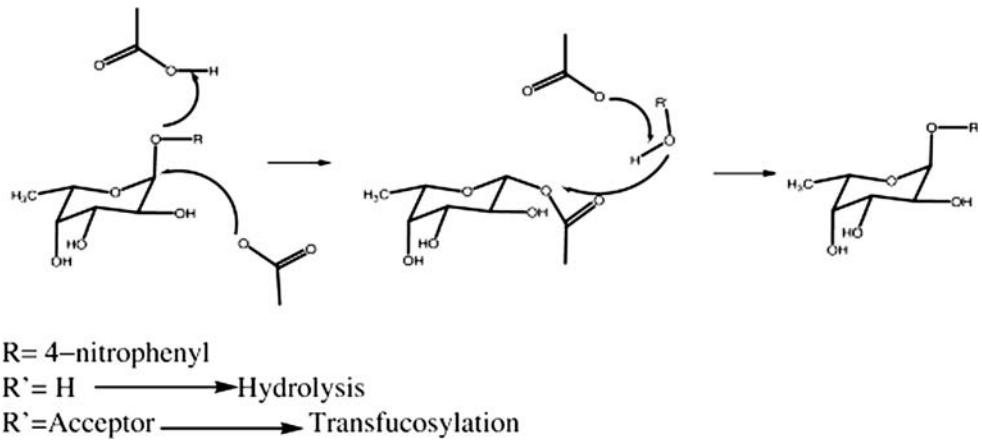


Fig. 1 *Retaining* reaction mechanism of α-L-fucosidases

Here, the characterization of the reaction mechanism and the identification of the catalytic residues of the first archaeal α-L-fucosidase identified in the hyperthermophile *Sulfolobus solfataricus* are briefly reviewed.

General features of the α-L-fucosidase

The first archaeal α-L-fucosidase has been identified and characterized recently (Cobucci-Ponzano et al. 2003a). The analysis of the genome of the hyperthermophilic archaeon *S. solfataricus* (She et al. 2001) revealed the presence of two open reading frames (ORFs), SSO11867 and SSO3060, encoding for 81 and 426 amino acid polypeptides that are homologous to the *N*- and the *C*-terminal parts, respectively, of full-length bacterial and eukaryal GH29 fucosidases (Henrissat 1991). The two ORFs are separated by a −1 frameshift and, to produce a single polypeptide, a single base was inserted by site-directed mutagenesis in the region of overlap between SSO11867 and SSO3060, restoring a single reading frame between the ORFs. The single ORF obtained was used to express the enzyme in *Escherichia coli* (Cobucci-Ponzano et al. 2003a). The recombinant enzyme, Ssα-fuc, is a nonamer of 57 kDa molecular mass subunits in solution and is highly active and specific for 4NP-α-L-fucoside (4-NP-α-L-Fuc) at 65°C (Cobucci-Ponzano et al. 2003a; Rosano et al. 2004). Moreover, Ssα-fuc is thermoactive and thermostable, as expected for an enzyme from a hyperthermophilic microorganism. The optimal temperature of Ssα-fuc is 95°C and the enzyme displayed high stability maintaining 60% of the residual activity after 2 h at 80°C (Cobucci-Ponzano et al. 2003a). It is worth noting that the mutation inserted to obtain the recombinant Ssα-fuc was designed on the basis of a mechanism of regulation of gene expression known as programmed −1 frameshifting (Farabaugh 1996). Very recently it was found that the two ORFs express in vivo a full length protein by programmed −1 frameshifting, demonstrating, for the first time, that this mechanism of gene expression, known so far only in *Eukarya* and *Bacteria* (Baranov et al. 2001) is used to regulate the expression of this gene in *S. solfataricus* (Cobucci-Ponzano et al. 2006).

In the framework of our mechanistic studies on glycoside hydrolases, the reaction mechanism of Ssα-Fuc was studied in detail and the residues directly involved in catalysis were identified. The *retaining* reaction mechanism was demonstrated, for the first time in GH29, by using Ssα-fuc. In fact, the enzyme is able to function in transfucosylation mode as reported for several mesophilic α-fucosidases (Murata et al. 1999; Farkas et al. 2000); its synthetic ability was demonstrated by using 4-NP-α-L-Fuc and 4-NP-α-D-glucoside (4-NP-α-D-Glc) as donor and acceptor, respectively. The fucosylated products were disaccharides of the acceptor in which the α-L-fucose moiety of the donor is attached at positions 2 and 3 of Glc (α-L-Fuc*p*-(1-2)-α-D-Glc-*O*-4-NP and α-L-Fuc*p*-(1-3)-α-D-Glc-*O*-4-NP) (Cobucci-Ponzano et al. 2003a). The α-anomeric configuration of the interglycosidic linkages in the products demonstrated that GH29 α-fucosidases follow a *retaining* reaction mechanism (Cobucci-Ponzano et al. 2003a). The hydrolytic activity of Ssα-fuc on the disaccharide α-L-Fuc-(1-3)-α-L-Fuc-*O*-4-NP revealed that the enzyme is an exo-glycosyl hydrolase that attacks the substrates from their non-reducing end (Cobucci-Ponzano et al. 2003a).

Identification of the nucleophile of the reaction

The active-site residues of retaining α- and β-glycosidases have been identified with a variety of methods, including mechanism-based inhibitors labelling the catalytic nucleophile and inspection of X-ray structures (McCarter and Withers 1996; Vocadlo et al. 2000, 2001; Tarling et al. 2003). An alternative approach often exploited for retaining glycoside hydrolases consists in the mutation of aspartic/glutamic acid residues identified by sequence analysis and conserved in the family of interest. Mutations of the catalytic residues with non-nucleophilic amino acids lead to the strong reduction or even abolition of the enzymatic activity (Ly and Withers 1999). However, these mutants can be reactivated in the presence of external nucleophiles such as sodium azide. The isolation of glycosyl-azide products with an anomeric configuration opposite to that of the substrate allows the identification of the catalytic nucleophile of the reaction (Fig. 2) (Ly and Withers 1999). This methodology is termed chemical rescue. Once the reaction mechanism and the active-site residues of a particular enzyme have been experimentally determined, they can be easily extended to all the homologous enzymes by following the classification in families.

The nucleophile of GH29 α-L-fucosidases was identified, for the first time, by reactivating the Ssα-fuc D242G inactive mutant in the presence of sodium azide and by analyzing the anomeric configuration of the fucosyl-azide product (Cobucci-Ponzano et al. 2003b). The D242G mutant showed a turnover number (k_{cat}) of 0.24 s^{-1} on 4-NP-α-L-Fuc, which is 1.2×10^{-3} times that of the wild type activity (287 s^{-1}), indicating that the D242G mutation affected a residue involved in catalysis in Ssα-fuc. In the presence of 2 M sodium azide the mutant revealed a k_{cat} value of 9.66 s^{-1}, indicating a 40-fold reactivation by azide (Table 1). The fucosyl-azide product obtained by the D242G mutant was found in the inverted

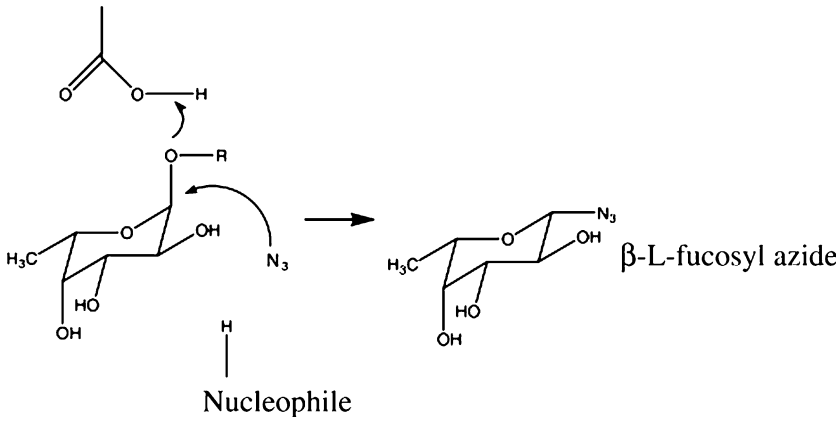


Fig. 2 Identification, by chemical rescue, of the nucleophile residue in *retaining* α-fucosidases. The α-L-fucoside substrate is converted in β-L-fucosyl-azide

Table 1 Steady-state kinetic constants of wild type and D242G mutant at 65°C

	k_{cat} (s^{-1})	K_M (mM)	k_{cat}/K_M (s^{-1} mM^{-1})
Wild type	287 ± 11	0.028 ± 0.004	10,250
D242G	NA[a]	NA[a]	–
+NaN_3 2M	9.7 ± 0.3	0.19 ± 0.02	51.5
+NaCOOH 0.1 M	5.9 ± 0.2	1.0 ± 0.1	5.8

[a] *NA* (not measurable activity) means that, using 20 μg of enzyme in the assay, the rates of change in absorbance at substrate concentration below 0.1 mM did not vary in the experimental conditions

(β-L) configuration when compared with the substrate (Fig. 2). This finding allowed, for the first time, the unambiguous assignment of Asp242 and its homologous residues as the nucleophilic catalytic residues of GH29 α-L-fucosidases (Cobucci-Ponzano et al. 2003b). The activity of the mutant D242G was also rescued on 4-NP-α-L-Fuc in the presence of sodium formate. The steady-state kinetic constants of D242G, determined in the presence of external nucleophiles, revealed that sodium azide and sodium formate produced about 0.5 and 0.056% of reactivation of the mutant, respectively, which was calculated by taking as 100% the k_{cat}/K_M of the wild type (Table 1) (Cobucci-Ponzano et al. 2003b). The higher nucleophilicity of sodium azide, when compared with formate, explains the higher reactivation produced by the former. The specific activity of the mutant increases with temperature up to 80°C, indicating that, despite the mutation, the reactivated enzyme maintains its thermophilicity.

This was the first example of the application of the chemical rescue method to α-(D/L)-glycosidases as it has been used so far only for β-D-glycosidases. Later, by following a similar approach, the corresponding residue was also identified in the α-fucosidase from *Thermotoga maritima* (Tmα-fuc) (Tarling et al. 2003). These results on GH29 enzymes demonstrated that chemical rescue could be of general applicability for retaining enzymes.

Identification of the acid/base catalyst of the reaction

The approach utilized for the identification of the acid/base catalyst of retaining glycosyl hydrolases is less straightforward if compared to the nucleophile. In fact, the use of specific inhibitors for the acid/base catalyst is still elusive and successful results are less common (Tull et al. 1996; Vocadlo et al. 2002). For these reasons, the acid/base catalyst of several retaining glycosidases was identified through 3D structure inspection and detailed characterization of mutants in which conserved aspartic and glutamic acid residues have been replaced by isosteric and non-ionizable amino acids as asparagine, glutamine, alanine, or glycine (Ly and Withers 1999). Replacing the acid/base catalyst with the small non-ionizable glycine residue generally reduces dramatically the activity of the mutant and modifies its pH profile. In fact, when the acid/base is removed, the basic limb of the typical bell-shaped pH dependence curve is severely affected (Ly and Withers 1999). The chemical rescue of the activity of the inactive mutant is also a useful tool. In fact, as described above for the mutant in the residue acting as the nucleophile of the reaction, the presence of the glycine generates a room in the active site allowing the access of a small nucleophilic ion. However, this time, the external nucleophile (i.e. azide) occupies the cavity formed by mutation after the formation of the glycosyl-enzyme intermediate. In these cases, the rate enhancement and the isolation of a glycosyl-azide product with the same anomeric configuration of the substrate resulted in the most effective method to unequivocally identify the acid/base catalyst (MacLeod et al. 1996; Viladot et al. 1998; Ly and Withers 1999; Vallmitjana et al. 2001; Debeche et al. 2002; Li et al. 2002; Rydberg et al. 2002; Shallom et al. 2002; Vocadlo et al. 2002; Bravman et al. 2003; Zechel et al. 2003; Paal et al. 2004; Sulzenbacher et al. 2004).

By following this line of approach, in Ssα-fuc several aminoacids among highly conserved histidine, aspartic,

and glutamic acid residues were picked and mutated into a glycine; the mutants H46G, E58G H123G, D124G, D146G were characterized in detail. Furthermore, the mutant E292G was added to this collection since the corresponding residue was identified as the acid/base catalyst in Tmα-fuc. Surprisingly, this residue falls in a region of alignment scarcely conserved in GH29 (Sulzenbacher et al. 2004). The preliminary kinetic characterization of the Ssα-fuc mutants on 4-NP-α-L-Fuc, reported in Table 2, revealed that D124 and D146 were not directly involved in catalysis since the activity was not significantly affected by the mutations. Instead, the affinity for the substrate of H46G and H123G was remarkably different from that of the wild type: the mutation of His46 produced a 607-fold increase in the K_M, while no saturation was observed with the mutant H123G. Also the mutation of the residues Glu58 and Glu292 affected catalysis severely (Table 2); again, no saturation could be observed with the former residue, while E292G showed unchanged affinity for 4-NP-α-L-Fuc but a 154-fold reduction in the turnover number (Cobucci-Ponzano et al. 2005).

This preliminary characterization indicated that the residues His46, His123, Glu58, and Glu292 are involved in substrate binding or in catalysis; however, experiments of chemical rescue of the enzymatic activity on the mutants H46G and H123G allowed us to exclude their involvement in catalysis (Cobucci-Ponzano et al. 2005). Furthermore, the inspection of the crystal structure of Tmα-fuc suggested that His46 and His122, which correspond to His34 and His128 in Tmα-fuc, respectively, stabilize the 4-hydroxyl group of fucose.

The characterization of the mutants E58G and E292G, compared to the data collected on the corresponding residues in Tmα-fuc (Glu66 and Glu262, respectively), gave unexpected results, suggesting that in GH29 two catalytic machineries coexist. The analysis of the 3D structure of the Tmα-fuc and the kinetic characterization of the mutants clearly indicated that, in this enzyme, Glu66 and Glu266 were involved in substrate binding and in the acid/base catalysis, respectively (Sulzenbacher et al. 2004). In particular, the mutation in the residue Glu66 produced a tenfold drop of activity while the mutation in Glu266 determined the absence of saturation by the substrate. Intriguingly, in Ssα-fuc the mutant E58G mirrored the behaviour of the Tmα-fuc mutant in the Glu266 residue (lack of saturation) while the Ssα-fuc mutant E292G showed a marked inactivation as observed for the Tmα-fuc mutant in Glu66 (Table 2). These results suggested that in Ssα-fuc Glu58 is the acid/base catalyst. This conclusion was further supported by the analysis of the pH dependence of wild type and mutants Ssα-fuc. Wild type enzyme has a peculiar pH dependence, which is not bell-shaped, but shows a reproducible increase of activity at pH 8.6 (Fig. 3), suggesting that more than two ionizable groups are involved in catalysis (Debeche et al. 2002). The pH dependence of the mutant E292G is similar to that of the wild type enzyme; instead, this pH profile was dramatically changed in the E58G mutant, producing a typical bell-shaped curve with a pH optimum at 4.6 sharper than that of the wild type (3.0–5.0) (Fig. 3). These data suggested that the removal of Glu58 unmasked the ionization of a group responsible for the basic limb (p*K*a 5.3) and possibly increased the p*K*a of the nucleophile of the reaction mainly determining the acidic limb. Unfortunately, the pH dependence of the wild type and mutant Tmα-fuc was not reported, precluding a detailed comparison.

To try to better define the nature of the acid/base catalyst in Ssα-fuc the chemical rescue, which, as described above, is one of the most definite tools to assign this role in glycosidases, was exploited. Remarkably, E58G was activated by more than 70-fold in the presence of sodium azide, formate and acetate (Table 3). In addition, it was observed

Table 2 Steady-state kinetic constants of wild type and mutants at 65°C

	k_{cat} (s^{-1})	K_M (mM)	k_{cat}/K_M (s^{-1} mM^{-1})
Wild type	287 ± 11	0.028 ± 0.004	10,250
D124G	240 ± 7	0.09 ± 0.01	2,740
D146G	224 ± 5	0.033 ± 0.003	6,791
H46G	419 ± 99	17.0 ± 5.7	25
H123G	ND[a]	ND[a]	3.5
E58G	ND[a]	ND[a]	1.9
E292G	1.86 ± 0.09	0.06 ± 0.01	33

The specificity constants were calculated by using 4-NP-α-L-Fuc in the range 1–7 mM and by plotting (velocity) vs. (S)

[a] *ND* not determined. No saturation was observed on up to 25 mM 4-NP-α-L-Fuc

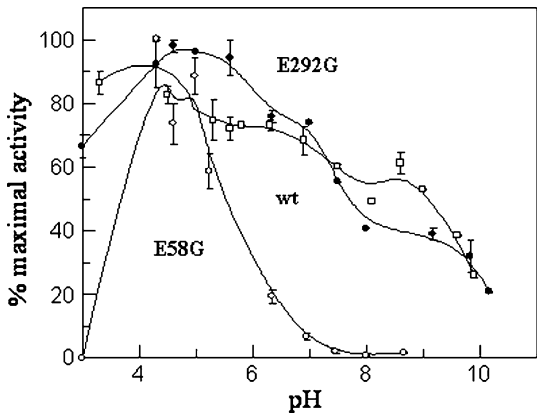


Fig. 3 pH Dependence of wild type, E58G and E292G. Data are reported as percent of maximal activity on 4-NP-α-L-Fuc at 65°C

Table 3 Steady-state kinetic constants of wild type, E58G and E292G mutants in different reaction conditions

	k_{cat} (s^{-1})	K_M (mM)	k_{cat}/K_M (s^{-1} mM^{-1})	pH	Reaction condition[a]
Wild type	287 ± 11	0.028 ± 0.004	10,250	6.3	Sodium phosphate
	430 ± 49	0.26 ± 0.09	1,624	5.0	Sodium citrate
E58G	ND[b]	ND[b]	1.9	6.3	Sodium phosphate
	143 ± 8	1.6 ± 0.3	89	4.6	Sodium citrate
	586 ± 43	0.6 ± 0.2	950	4.6	Sodium acetate
	846 ± 46	1.1 ± 0.2	790	4.6	Sodium formate
	679 ± 82	2.9 ± 0.4	229	6.3	Sodium phosphate + NaN_3[c]
E292G	1.86 ± 0.09	0.06 ± 0.01	33	6.3	Sodium phosphate
	1.86 ± 0.17	0.09 ± 0.02	21	4.6	Sodium acetate

The specificity constants were calculated by using 4-NP-α-L-Fuc in the range 1–7 mM and by plotting (velocity) vs. (S)

[a] Assays were performed at 65°C in the reported buffers in 0.05 M concentration on 4-NP-α-L-Fuc

[b] Not determined; no saturation was observed with up to 25 mM 4-NP-α-L-Fuc

[c] Sodium azide was used at the concentration of 0.15 M

that the sodium azide anion activated E58G only in the presence of larger ions (phosphate and citrate) while this effect was much reduced in acetate and formate, which already activate the mutant. Presumably, in phosphate and citrate buffers, azide has full access to the small cavity created by the mutation in the active site. Instead, acetate and formate, occupying this space, could activate the reaction hampering the access of azide. In striking contrast, the activity of E292G could not be rescued by any of the nucleophiles used. These results made it very unlikely that Glu292 is the acid/base catalyst of Ssα-fuc, and allowed to assign this role to Glu58.

These data demonstrated that the Glu58 is the acid-base catalyst and suggested that the Glu292 has a relevant role in catalysis presumably modulating the p*K*a of the latter, thereby affecting the pH optimum of the enzyme. Intriguingly, the behaviour of the catalytic residues of Ssα-fuc is different from that of Tmα-fuc (Sulzenbacher et al. 2004). Nevertheless, considering that among the amino acid sequences of GH29 the predicted acid/base residues are not invariant, it would not be surprising that the enzymes show structural differences in the active site explaining the different catalytic machineries.

Conclusions

The first archaeal α-L-fucosidase was identified in *S. solfataricus* and is encoded by an interrupted gene. The recombinant enzyme Ssα-fuc, obtained by site directed mutagenesis, is fully active and thermostable and allowed the first detailed study on the *retaining* reaction mechanism of GH29 glycoside hydrolases. Interestingly, the inspection of the catalytic machinery of Ssα-fuc revealed the presence in the active site of a triad of catalytic residues, namely Asp242, Glu58, and Glu292. This is a rather uncommon feature among the glycoside hydrolases and the comparison with the Tmα-fuc suggested that in GH29 slightly different catalytic machineries coexist. It is worth noting that the use of the chemical rescue method at harsh pHs and ionic strengths was possible because of the intrinsic stability of Ssα-fuc and resulted of general applicability for β and α glycoside hydrolases.

The body of this work demonstrates that the α-L-fucosidase from *S. solfataricus* is an interesting model system to uncover new mechanisms of gene expression in *Archaea* and to study the reaction mechanisms of glycoside hydrolases. In addition, the transfucosylating activity of Ssα-fuc and the availability of several mutants in the active site could be the starting points for the biotechnological exploitation of this enzyme in the synthesis of fucosylated oligosaccharides.

Acknowledgments This work was supported by MIUR project "Folding di proteine: l'altra metà del codice genetico" RBAU015B47_006. The IBP-CNR belongs to the Centro Regionale di Competenza in Applicazioni Tecnologico-Industriali di Biomolecole e Biosistemi.

References

Baranov PV, Gurvich OL, Fayet O, Prere MF, Miller WA, Gesteland RF, Atkins JF, Giddings MC (2001) RECODE: a database of frameshifting, bypassing and codon redefinition utilized for gene expression. Nucleic Acids Res 29:264–267

Bravman T, Belakhov V, Solomon D, Shoham G, Henrissat B, Baasov T, Shoham Y (2003) Identification of the catalytic residues in family 52 glycoside hydrolase, a beta-xylosidase from *Geobacillus stearothermophilus* T-6. J Biol Chem 278:26742–26749

Cobucci-Ponzano B, Trincone A, Giordano A, Rossi M, Moracci M (2003a) Identification of an archaeal alpha-L-fucosidase encoded by an interrupted gene. Production of a functional enzyme by mutations mimicking programmed −1 frameshifting. J Biol Chem 278:14622–14631

Cobucci-Ponzano B, Trincone A, Giordano A, Rossi M, Moracci M (2003b) Identification of the catalytic nucleophile of the family 29 alpha-L-fucosidase from *Sulfolobus solfataricus* via chemical rescue of an inactive mutant. Biochemistry 42:9525–9531

Cobucci-Ponzano B, Mazzone M, Rossi M, Moracci M (2005) Probing the catalytically essential residues of the alpha-L-fucosidase from the hyperthermophilic archaeon *Sulfolobus solfataricus*. Biochemistry 44:6331–6342

Cobucci-Ponzano B, Conte F, Benelli D, Londei P, Flagiello A, Monti M, Pucci P, Rossi M, Moracci M (2006) The gene of an archaeal alpha-L-fucosidase is expressed by translational frameshifting. Nucleic Acids Res 34:4258–4268

Debeche T, Bliard C, Debeire P, O'Donohue MJ (2002) Probing the catalytically essential residues of the alpha-L-arabinofuranosidase from *Thermobacillus xylanilyticus*. Protein Eng 1:21–28

de La Torre F, Sampedro J, Zarra I, Revilla G (2002) AtFXG1, an *Arabidopsis* gene encoding alpha-L-fucosidase active against fucosylated xyloglucan oligosaccharides. Plant Physiol 128:247–255

Farabaugh PJ (1996) Programmed translational frameshifting. Microbiol Rev 60:103–134

Farkas E, Thiem J, Ajisaka K (2000) Enzymatic synthesis of fucose-containing disaccharides employing the partially purified alpha-L-fucosidase from *Penicillium multicolor*. Carbohydr Res 328:293–299

Fischer L, Bromann R, Kengen SWM, de Vos WM, Wagner F (1996) Catalytical potency of beta-glucosidase from the extremophile *Pyrococcus furiosus* in glucoconjugate. Biotechnology 14:88–91

Henrissat B (1991) A classification of glycosyl hydrolases based on amino acid sequence similarities. Biochem J 280:309–316

Henrissat B, Bairoch A (1993) New families in the classification of glycosyl hydrolases based on amino acid sequence similarities. Biochem J 293:781–788

Henrissat B, Davies G (1997) Structural and sequence-based classification of glycoside hydrolases. Curr Opin Struct Biol 7:637–644

Koshland DE (1953) Stereochemistry and the mechanism of enzymatic reactions. Biol Rev Camb Philos Soc 28:416–436

Li YK, Chir J, Tanaka S, Chen FY (2002) Identification of the general acid/base catalyst of a family 3 beta-glucosidase from *Flvobacterium meningosepticum*. Biochemistry 41:2751–2759

Listinsky JJ, Siegal GP, Listinsky CM (1998) Alpha-L-fucose: a potentially critical molecule in pathologic processes including neoplasia. Am J Clin Pathol 110:425–440

Lower BH, Kennelly PJ (2002) The membrane-associated protein-serine/threonine kinase from *Sulfolobus solfataricus* is a glycoprotein. J Bacteriol 184:2614–2619

Ly HD, Withers SG (1999) Mutagenesis of glycosidases. Annu Rev Biochem 68:487–522

MacLeod AM, Tull D, Rupitz K, Warren RA, Withers SG (1996) Mechanistic consequences of mutation of active site carboxylates in a retaining beta-1,4-glycanase from *Cellulomonas fimi*. Biochemistry 35:13165–13172

McCarter JD, Withers SG (1994) Mechanisms of enzymatic glycoside hydrolysis. Curr Opin Struct Biol 4:885–892

McCarter JD, Withers SG (1996) Unequivocal identification of Asp-214 as the catalytic nucleophile of *Saccharomyces cerevisiae* alpha-glucosidase using 5-fluoro glycosyl fluorides. J Biol Chem 271:6889–6894

Michalski JC, Klein A (1999) Glycoprotein lysosomal storage disorders: alpha- and beta-mannosidosis, fucosidosis and alpha-*N*-acetylgalactosaminidase deficiency. Biochim Biophys Acta 1455:69–84

Mori E, Hedrick JL, Wardrip NJ, Mori T, Takasaki S (1998) Occurrence of reducing terminal *N*-acetylglucosamine 3-sulfate and fucosylated outer chains in acidic *N*-glycans of porcine zona pellucida glycoproteins. Glycoconj J 15:447–456

Murata T, Morimoto S, Zeng X, Watanabe S, Usui T (1999) Enzymatic synthesis of alpha-L-fucosyl-*N*-acetyllactosamines and 3′-*O*-alpha-L-fucosyllactose utilizing alpha-L-fucosidases. Carbohydr Res 320:192–199

Noda K, Miyoshi E, Uozumi N, Gao CX, Suzuki K, Hayashi N, Hori M, Taniguchi N (1998) High expression of alpha-1-6 fucosyltransferase during rat hepatocarcinogenesis. Int J Cancer 75:444–450

Paal K, Ito M, Withers SG (2004) *Paenibacillus* sp. TS12 glucosylceramidase Kinetic studies of a novel sub-family of family 3 glycosidases and identification of the catalytic residues. Biochem J 378:141–149

Perugino G, Trincone A, Rossi M, Moracci M (2004) Oligosaccharide synthesis by glycosynthases. Trends Biotechnol 1:31–37

Rapoport E, Pendu JL (1999) Glycosylation alterations of cells in late phase apoptosis from colon carcinomas. Glycobiology 9:1337–1345

Rosano C, Zuccotti S, Cobucci-Ponzano B, Mazzone M, Rossi M, Moracci M, Petoukhov MV, Svergun DI, Bolognesi M (2004) Structural characterization of the nonameric assembly of an Archaeal alpha-L-fucosidase by synchrotron small angle X-ray scattering. Biochem Biophys Res Commun 320:176–182

Russell L, Waring P, Beaver JP (1998) Increased cell surface exposure of fucose residues is a late event in apoptosis. Biochem Biophys Res Commun 250:449–453

Rydberg EH, Li C, Maurus R, Overall CM, Brayer GD, Withers SG (2002) Mechanistic analyses of catalysis in human pancreatic alpha-amylase: detailed kinetic and structural studies of mutants of three conserved carboxylic acids. Biochemistry 41:4492–4502

Sears P, Wong CH (1996) Intervention of carbohydrate recognition by proteins and nucleic acids. Proc Natl Acad Sci USA 93:12086–12093

Shallom D, Belakhov V, Solomon D, Gilead-Gropper S, Baasov T, Shoham G, Shoham Y (2002) The identification of the acid-base catalyst of alpha-arabinofuranosidase from *Geobacillus stearothermophilus* T-6, a family 51 glycoside hydrolase. FEBS Lett 514:163–167

She Q, Singh RK, Confalonieri F, Zivanovic Y, Allard G, Awayez MJ, Chan-Weiher CC, Clausen IG, Curtis BA, De Moors A, Erauso G, Fletcher C, Gordon PM, Heikamp-de Jong I, Jeffries AC, Kozera CJ, Medina N, Peng X, Thi-Ngoc HP, Redder P, Schenk ME, Theriault C, Tolstrup N, Charlebois RL, Doolittle WF, Duguet M, Gaasterland T, Garrett RA, Ragan MA, Sensen CW, Van der Oost J (2001) The complete genome of the crenarchaeon *Sulfolobus solfataricus* P2. Proc Natl Acad Sci USA 98:7835–7840

Sinnot ML (1990) Catalytic mechanism of enzymic glycosyl transfer. Chem Rev 90:1171–1202

Staudacher E, Altmann F, Wilson IB, Marz L (1999) Fucose in *N*-glycans: from plant to man. Biochim Biophys Acta 1473:216–236

Sulzenbacher G, Bignon C, Nishimura T, Tarling CA, Withers SG, Henrissat B, Bourne Y (2004) Crystal structure of *Thermotoga maritima* α-L-fucosidase. Insights into the catalytic mechanism and the molecular basis for fucosidosis. J Biol Chem 279:13119–13128

Tarling CA, He S, Sulzenbacher G, Bignon C, Bourne Y, Henrissat B, Withers SG (2003) Identification of the catalytic nucleophile of the family 29 α-L-fucosidase from *Thermotoga maritima* through trapping of a covalent glycosylenzyme intermediate and mutagenesis. J Biol Chem 278:47394–47399

Tull D, Burgoyne DL, Chow DT, Withers SG, Aebersold R (1996) A mass spectrometry-based approach for probing enzyme active sites: Identification of Glu 127 in *Cellulomonas fimi* exoglycanase as the residue modified by *N*-bromoacetyl cellobiosylamine. Anal Biochem 234:119–125

Vallmitjana M, Ferrer-Navarro M, Planell R, Abel M, Ausin C, Querol E, Planas A, Perez-Pons JA (2001) Mechanism of the family 1 beta-glucosidase from *Streptomyces* sp: catalytic residues and kinetic studies. Biochemistry 40:5975–5982

Vanhooren PT, Vandamme EJ (1999) L-fucose: occurrence, physiological role, chemical, enzymatic and microbial synthesis. J Chem Technol Biotechnol 74:479–497

Varki A (1993) Biological roles of oligosaccharides: all of the theories are correct. Glycobiology 2:97–130

Viladot JL, de Ramon E, Durany O, Planas A (1998) Probing the mechanism of Bacillus 1,3-1,4-beta-D-glucan 4-glucanohydrolases by chemical rescue of inactive mutants at catalytically essential residues. Biochemistry 37:11332–11342

Vocadlo DJ, Mayer C, He S, Withers SG (2000) Mechanism of action and identification of Asp242 as the catalytic nucleophile of *Vibrio furnisii N*-acetyl-beta-D-glucosaminidase using 2-acetamido-2-deoxy-5-fluoro-alpha-L-idopyranosyl fluoride. Biochemistry 1:117–126

Vocadlo DJ, Davies GJ, Laine R, Withers SG (2001) Catalysis by hen egg-white lysozyme proceeds via a covalent intermediate. Nature 412:835–838

Vocadlo DJ, Wicki J, Rupitz K, Withers SG (2002) A case for reverse protonation: identification of Glu160 as an acid/base catalyst in *Thermoanaerobacterium saccharolyticum* beta-xylosidase and detailed kinetic analysis of a site-directed mutant. Biochemistry 41:9736–9746

Wiese TJ, Dunlap JA, Yorek MA (1997) Effect of L-fucose and D-glucose concentration on L-fucoprotein metabolism in human Hep G2 cells and changes in fucosyltransferase and alpha-L-fucosidase activity in liver of diabetic rats. Biochim Biophys Acta 1335:61–72

Xiang J, Bernstein IA (1992) Differentiative changes in fucosyltransferase activity in newborn rat epidermal cells. Biochem Biophys Res Commun 189:27–32

Zechel DL, Withers SG (2000) Glycosidase mechanisms: anatomy of a finely tuned catalyst. Acc Chem Res 33:11–18

Zechel DL, Reid SP, Stoll D, Nashiru O, Warren RA, Withers SG (2003) Mechanism, mutagenesis, and chemical rescue of a beta-mannosidase from *Cellulomonas fimi*. Biochemistry 42:7195–7204

Biocatalysis and Biotransformation
Publication details, including instructions for authors and subscription information:
http://www-intra.informaworld.com/smpp/title~content=t713454445

Design of new reaction conditions for characterization of a mutant thermophilic **α**-l-fucosidase

B. Cobucci-Ponzano [a]; F. Conte [a]; M. Mazzone [a]; E. Bedini [b]; M. M. Corsaro [b]; M. Rossi [ac]; M. Moracci [a]
[a] Institute of Protein Biochemistry, Naples, Italy
[b] Dipartimento di Chimica Organica e Biochimica, Università di Napoli Federico II, Naples, Italy
[c] Dipartimento di Biologia Strutturale e Funzionale, Università di Napoli "Federico II", Naples, Italy

First Published on: 18 February 2008
To cite this Article: Cobucci-Ponzano, B., Conte, F., Mazzone, M., Bedini, E., Corsaro, M. M., Rossi, M. and Moracci, M. (2008) 'Design of new reaction conditions for characterization of a mutant thermophilic **α**-l-fucosidase', Biocatalysis and Biotransformation, 26:1, 18 - 24
To link to this article: DOI: 10.1080/10242420701792209
URL: http://dx.doi.org/10.1080/10242420701792209

Biocatalysis and Biotransformation, January–February 2008; 26(1–2): 18–24

ORIGINAL ARTICLE

Design of new reaction conditions for characterization of a mutant thermophilic α-L-fucosidase

B. COBUCCI-PONZANO[1], F. CONTE[1], M. MAZZONE[1], E. BEDINI[2], M. M. CORSARO[2], M. ROSSI[1,3], & M. MORACCI[1]

[1]*Institute of Protein Biochemistry,* [2]*Dipartimento di Chimica Organica e Biochimica, Università di Napoli Federico II and* [3]*Dipartimento di Biologia Strutturale e Funzionale, Università di Napoli "Federico II", Naples, Italy*

(Received 5 May 2007; accepted 20 July 2007)

Abstract
Glycosynthases are mutant glycosidases, which in the presence of activated glycosides and suitable reaction conditions, synthesize oligosaccharides without hydrolysing them. This feature makes these catalysts promising tools for the large scale synthesis of carbohydrates. However, despite the popularity of the glycosynthetic approach, the number of enzymes effecting glycosynthetic reactions is still limited. We report here on the design of novel reaction conditions for a thermophilic α-L-fucosidase mutant, which might provide a route for the production of novel glycosynthases.

Keywords: *Glycosidase reaction mechanism, glycosynthase, oligosaccharide synthesis, site-directed mutagenesis, fucosylation*

Introduction

The chemical synthesis of carbohydrates is an extremely powerful tool that, in principle, would allow the formation of a large variety of compounds and even combinatorial oligosaccharide libraries (Schweizer & Hindsgaul 1999). However, the problems of controlling the stereo- and the regioselectivity of the reaction are demanding and hamper efficient production of oligosaccharides (Crout & Vic 1998). This explains why enzymatic and chemoenzymatic approaches have aroused so much interest in recent years (Flitsch 2000; Karst and Linhardt 2003; Moracci et al. 2001; Murata and Usui 2006; Sears & Wong 2001; Wymer & Toone 2000). The enzyme-based methods for the synthesis of carbohydrates available so far are based on two major classes of catalysts: glycosyl transferases and glycosidases. The former enzymes allow quantitative yields, but their use is hampered by the high costs of the substrates and the scarcity of the enzymes. Alternatively, glycosidases, promoting synthesis through either equilibrium-controlled (reverse hydrolysis) or kinetically-controlled (transglycosylation) processes, are cheap enzymes, but provide lower yields (10–40%), which are generally uneconomical for large-scale synthesis.

A novel approach, based on the modification of glycosidases by site-directed mutagenesis in the residue acting as the nucleophile of the reaction and their subsequent reaction with activated substrates, was first proposed for the β-glycosidase from *Agrobacterium* sp. (Mackenzie et al. 1998; Wang et al. 1998) and subsequently exploited in several laboratories worldwide (for a review see Perugino et al. 2004). Consequently, these mutant enzymes, which synthesize oligosaccharides without hydrolysing them, were named *glycosynthases*.

Although the preparation of a glycosynthase from a specific glycosidase is, in principle, a simple strategy, this has not led to a large variety of different enzymes and the glycosynthases available so far are limited to glycoside families 1, 2, 5, 7, 8, 10, 16, 17, 26 and 31 out of the more than 100 glycosidase families described (http://www.cazy.org/). In fact, there are reports of several enzymes which are recalcitrant to becoming glycosynthases: for instance, only one α-glycosynthase is known so far (Okuyama et al. 2002), and enzymes from families

Correspondence: Dr M. Moracci Institute of Protein Biochemistry – Consiglio Nazionale delle Ricerche, Via P. Castellino 111, 80131, Naples, Italy. Tel. +39-081-6132271. Fax +39-081-6132277. E-mail: m.moracci@ibp.cnr.it

ISSN 1024-2422 print/ISSN 1029-2446 online
DOI: 10.1080/10242420701792209

29, 31, 35 and 39 modified to this end, did not yield glycosynthases (Cobucci-Ponzano et al. 2003b; Perugino et al. 2005). Therefore, the glycosynthetic approach is still problematic and more studies are required to understand the molecular basis of these difficulties.

We tackled this issue by studying a thermophilic α-L-fucosidase from the hyperthermophilic archaeon *Sulfolobus solfataricus* (Ssαfuc). Fucose derivatives play a central role in biological processes including the regulation of plant growth (de La Torre et al. 2002) and several physiological and pathological events in mammals (Listinsky et al. 1998; Mori et al. 1998; Noda et al. 1998). Therefore, the efficient synthesis of fucosylated oligosaccharides has potential applications in biomedicine (Vanhooren & Vandamme 1999).

Ssαfuc, belonging to glycoside family 29, was produced by merging, with a mutation that produced a single frame, the two Open Reading Frames (ORF) which encode for the enzyme in *S. Solfataricus* (Cobucci-Ponzano et al. 2003a), but are otherwise seperated by a-1 frameshift. The enzyme is expressed *in vivo* by a mechanism of translational regulation of gene expression called programmed -1 frameshifting (Cobucci-Ponzano et al. 2006). The recombinant enzyme was characterized in detail, and the catalytic triad in the active site was identified by site-directed mutagenesis and enzymological studies (Cobucci-Ponzano et al. 2003b, 2005).

Attempts to produce a fucosynthase from Ssαfuc were not successful. As expected, mutation of the residue acting as the nucleophile of the reaction with a non-nucleophilic residue (Asp242Gly) produced a completely inactive enzyme. However, incubation of the mutant with sodium formate acting as external nucleophile and the substrate 4-nitrophenyl α-L-fucopyranoside (4NP-Fuc) produced a 16-fold reactivation, but no fucosylated oligosaccharide production (Cobucci-Ponzano et al. 2003b). Here, we describe the kinetic characterization of the SsαfucD242G mutant reactivated in the presence of different nucleophiles and a chemically-synthesized activated substrate. The results indicate that engineering the substrate, the enzyme and the reaction conditions together is a promising strategy for the production of a novel fucosynthase.

Materials and methods

Chemicals

All commercially available substrates were of analytical grade from Sigma (USA).

Expression and purification of D242G mutant of Ssαfuc

The mutant enzyme was expressed and purified as previously described (Cobucci-Ponzano et al. 2003b). Briefly, the D242G mutant was expressed from *E. coli* BL21(RB791)/pGEX-SsαfucD242G as a fusion protein with glutathione-S-transferase (GST). GST binding was performed by adding 3 mL of the glutathione sepharose 4B matrix (Amersham Biotech) to the crude extract and incubating overnight at 4°C. After binding, the matrix was washed and incubated overnight at 4°C with 60 units of thrombin solution. Thereafter, the soluble, GST-free α-fucosidase mutant was recovered and a final heating step at 80°C was performed. This purification procedure was achieved with a matrix that is dedicated only to the purification of the mutant, excluding contamination by the wild type enzyme from external sources.

Chemical synthesis of 2-chloro-4-nitrophenyl α-L-fucopyranoside (2C4NP-Fuc)

L-Fucose (700 mg) purchased from CMS Chemicals (Abingdon, UK) was dissolved in 1:1 pyridine/acetic anhydride (8 mL) and stirred overnight at 25°C, after which the volatiles were removed in a rotary evaporator. The residue was diluted with dichloromethane (100 mL) and washed in a separating funnel with 1 M HCl (100 mL), then with 1 M $NaHCO_3$ (100 mL) and finally with water (100 mL). The organic layer was dried over anhydrous Na_2SO_4, filtered and concentrated to give tetra-acetyl-L-fucose (1.41 g), which was, in turn, mixed with 2-chloro-4-nitrophenol (1.10 g = 1.5 equivalents). The solid mixture was carefully dried by triple co-evaporation with dry toluene (10 mL), then mixed with $FeCl_3$ (755 mg = 1.1 equivalents) under an argon atmosphere and suspended at 0°C in dry dichloromethane (25 mL). The mixture was stirred at 0°C for 4 h, then diluted with dichloromethane (200 mL) and washed with 1 M $NaHCO_3$ (200 mL). The organic layer was dried over anhydrous Na_2SO_4, filtered and concentrated to give a residue that, after silica-gel column chromatography (6:1:0.01 petroleum ether/ethyl acetate/triethylamine as eluent), afforded crystalline 2-chloro-4-nitrophenyl 2,3,4-triacetyl-α-L-fucopyranoside (470 mg). This compound was dissolved in 4:1:1 methanol/water/triethylamine (18 mL) and the solution was stirred at 25°C overnight. Then, it was concentrated and the obtained residue was purified by silica-gel column chromatography (99:1:0.01 to 95:5:0.01 dichloromethane/methanol/triethylamine as eluent) to give pure 2C4NP-Fuc as yellow crystals [215 mg; 200 MHz 1H-NMR in perdeuterated

acetone: δ 8.26 (*d*, $\mathcal{J}$ = 2.8 Hz, 1H), 8.20 (*dd*, $\mathcal{J}$ = 9.0, 2.8 Hz, 1H), 7.52 (*d*, $\mathcal{J}$ = 9.0 Hz, 1H), 5.85 (*d*, $\mathcal{J}$ = 3.0 Hz, 1H), 4.06 (*m*, 2H), 3.99 (*dq*, $\mathcal{J}$ = 6.6, 0.9 Hz, 1H), 3.82 (*dd*, $\mathcal{J}$ = 2.7, 0.9 Hz, 1H), 1.17 (*d*, $\mathcal{J}$ = 6.6 Hz, 3H)].

Elemental analysis

Calculated: C 45.08, H 4.41, N 4.38.
Found: C 45.28, H 4.42, N 4.32.

Enzyme characterization and glycosynthetic trials

The activity of the Ssαfuc wild type and D242G mutant under standard conditions was measured at 65°C in 50 mM sodium phosphate buffer at pH 6.5, with 4NP-Fuc substrate at a final concentration of 1 mM, using up to 20 μg of enzyme, as reported previously (Cobucci-Ponzano et al. 2003a). Spontaneous hydrolysis of the substrate was subtracted by using appropriate blank mixtures without enzyme. One unit of enzyme activity was defined as the amount of enzyme catalyzing the hydrolysis of 1 μmol of substrate in 1 min under the conditions described.

The chemically rescued activities of the D242G mutant on 4NP-Fuc substrate were measured under the conditions described. The reaction (0.2 mL) was started by adding the enzyme and, after 5–240 min of incubation at 65°C, stopped by adding 0.8 mL of ice-cold 1 M sodium carbonate. The optical density of the solution was measured at 420 nm at room temperature. Spontaneous hydrolysis of the substrate was subtracted by using appropriate blank mixtures without enzyme. The molar extinction coefficient of para-nitrophenol measured at 420 nm, at room temperature in 1 M sodium carbonate buffer is 17.2 mM^{-1} cm^{-1}.

Identical reaction mixtures prepared as described above were incubated for 16 h at 65°C, stopped by freezing and loaded onto a silica gel TLC (Ethyl acetate/CH_3OH/H_2O, 7:2:1, v/v/v).

Steady-state kinetic parameters of wild type and the D242G mutant using 4NP-Fuc as substrate have been described previously (Cobucci-Ponzano et al. 2003b). The kinetic parameters of the D242G mutant on 2C4NP-Fuc substrate were measured in 50 mM sodium phosphate buffer pH 6.5 at 65°C, with 2C4NP-Fuc concentrations ranging from 0.01 to 1 mM, by following the method described above. The amount of the enzyme used in the assays was 10 μg. The molar extinction coefficient of 2-chloro-4-nitrophenol, measured at 420 nm, at room temperature in 1M sodium carbonate buffer is 14 mM^{-1} cm^{-1}. Kinetic constants of wild type Ssαfuc on 2C4NP-Fuc substrate were measured, at 65°C in 50 mM sodium phosphate buffer pH 6.5 using substrate concentrations in the range 0.01–0.4 mM. The amount of enzyme used in the assays was 1 μg. The molar extinction coefficient of 2-chloro-4-nitrophenol is 17 mM^{-1} cm^{-1} measured at 405 nm, at 65°C, in 50 mM sodium phosphate buffer pH 6.5. All kinetic data were calculated as the average of at least two experiments and were plotted and refined with the program *GraFit* (Leatherbarrow 1992).

Results and discussion

Retaining glycosidases, leading to products with the same anomeric configuration as the substrate, follow a two-step mechanism with the formation of a covalent glycosyl-enzyme intermediate (Koshland 1953). In the first step (glycosylation step), the nucleophilic group attacks the anomeric centre of the substrate while the acid/base residue, acting in this step as a general acid, assists the leaving of the aglycon moiety. The concerted action of the two amino acids leads to the formation of a covalent glycosyl-enzyme intermediate (Figure 1a). In the second step (deglycosylation step), the glycosyl-enzyme intermediate is cleaved by a water molecule that acts as nucleophile, being polarized by the general base catalyst. The product maintains the same anomeric configuration as the substrate. When an acceptor different from water, such as an alcohol or a sugar, intercepts the reactive glycosyl-enzyme intermediate, retaining enzymes work in transglycosylation mode. From this mechanism it is clear that mutation of the residue acting as the nucleophile in the reaction to a non-nucleophile residue produces a completely inactive enzyme. However, we have demonstrated that these mutants in hyperthermophilic glycosidases can be exploited to prepare glycosynthases; in fact, when the activity of the mutant enzyme is chemically rescued in the presence of an activated substrate and an external nucleophile, such as sodium formate, the synthesis of oligosaccharides is promoted (Moracci et al. 1998). These products result from the action of the external ion mimicking the nucleophile function eliminated by mutation and allows the formation of a glycosyl-nucleophile intermediate (Figure 1b). Subsequently, a sugar acceptor, activated by the residues acting as a general base, reacts with the intermediate leading to the final product. The disaccharide formed contains a poor leaving group (the sugar that acted as acceptor) and accumulates in the reaction. This method has been successful for several glycosynthases from hyperthermophilic archaea and exploited to produce complex oligosaccharides (Moracci et al. 1998; Trincone et al. 2000, 2003;

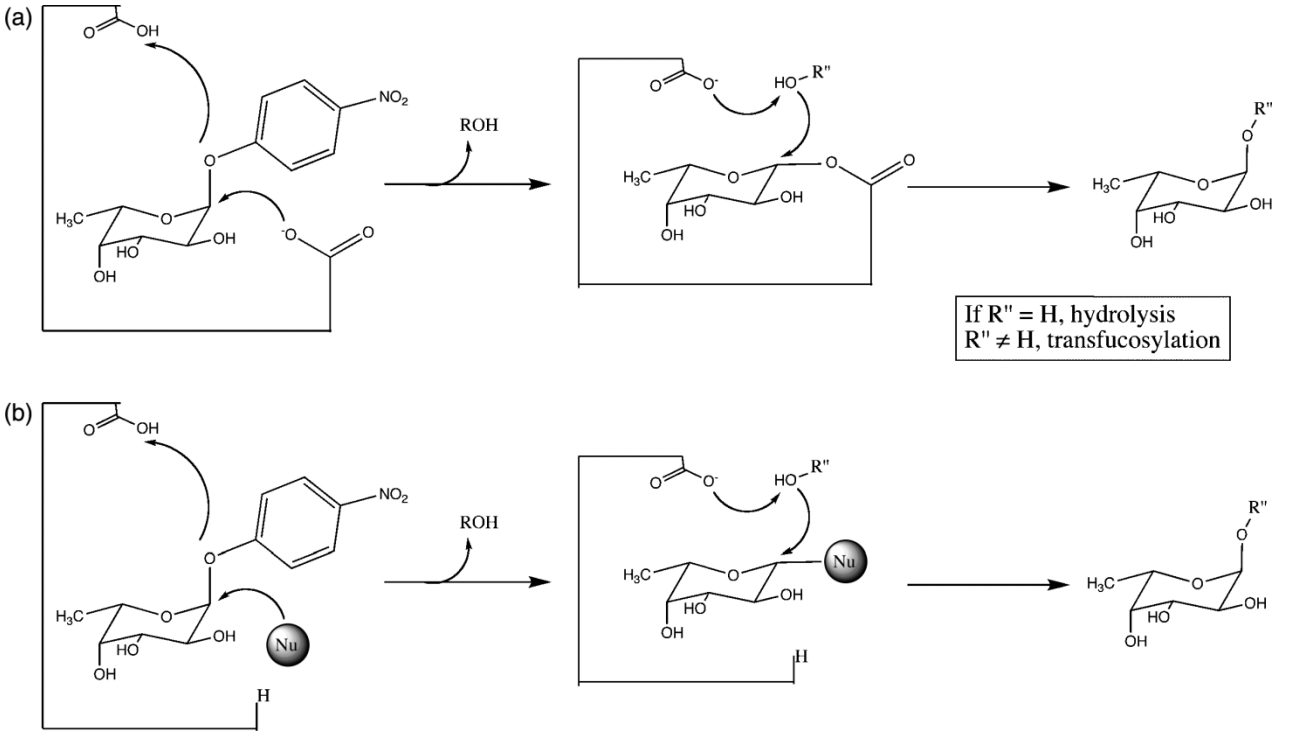


Figure 1. Reaction mechanism of a retaining α-L-fucosidase (a) Proposed strategy for the production of a fucosynthase (b) the external nucleophile is indicated by a grey sphere; the R″ group in the fucosylated product is the novel leaving group (see text for the details).

Perugino et al. 2003, 2006); therefore, we decide to apply it to the nucleophile mutant of the α-L-fucosidase from *S. solfataricus* (Ssαfuc).

We have previously reported that the D242G mutant of Ssαfuc, although 16-fold more active on 4NP-Fuc when assayed in 100 mM sodium formate buffer pH 4.0, did not catalyse the formation of oligosaccharides, indicating that the mutant did not work as a α-L-fucosynthase (Cobucci-Ponzano et al. 2003b). The absence of synthetic products could be due to the nature of the ion used in these experiments or to the acidic pH exploited; in fact, for the success of our glycosynthetic approach, the formation of a glycosyl-nucleophile intermediate that is sufficiently stable to transfer the glycon moiety to an acceptor, is crucial. In this regard, 0.1 M sodium formate buffer at pH 4.0, though leading to reactivation, might represent poor conditions for glycosynthase action. To test this hypothesis we analysed the activity of the SsαfucD242G mutant under a variety of conditions, including sodium formate or sodium acetate buffers at different pHs, sodium formate, sodium acetate or sodium chloride as different nucleophiles in sodium phosphate buffer, in the presence of four glycoside acceptors, and at different concentrations of 4NP-Fuc substrate donor (Table I). All the assays were performed at 65°C. As a comparison, we reported the specific activity of the D242G mutant under standard conditions without external nucleophiles. Remarkably, the activity of the enzyme was much lower than we had shown previously (0.03 U/mg) (Cobucci-Ponzano et al. 2003b). Detailed characterization of the mutant demonstrated that the specific activity previously published was overestimated due to the presence of traces (0.3 mM) of sodium azide, used as an antibacterial during storage, in the old enzyme preparations. As a consequence, the highest reactivation observed with mutant preparations depleted of sodium azide was now 300-fold (Table I, entry 2) and more than 137-fold based on the conditions previously reported (entry 5) (Cobucci-Ponzano et al. 2003b). Table I shows that, under most of the reactivation conditions tested, the Ssαfuc D242G mutant showed increased hydrolytic activity. In particular, sodium formate was by far the best nucleophile tested while higher concentrations of 4NP-Fuc did not improve the activity (entries 2–9). Moreover, addition of either α- or β-D-gluco, or -galactosides to the reaction mixture did not increase the activity (entries 10–19); instead, it hampered the chemically rescued activity (compare entry 2 with entries 10–13). Therefore, the D-glycosides used were poor acceptors and probably compete with 4NP-Fuc in the donor site.

Analysis of the reaction mixtures by thin layer chromatography (TLC) even after 16 h incubations, did not reveal any product (data not shown), confirming that the mutant did not work as a glycosynthase. A possible explanation for this could

Table I. Chemically rescued activity of the SsαfucD242G mutant.

Entry	4NP-Fuc donor (mM)	Acceptor	Buffer	pH	External nucleophile	Specific activity (U/mg)	% of maximal activity
1	2	–	50 mM sodium phosphate	6.5	–	8×10^{-4}	–
2	2	–	50 mM sodium phosphate	6.5	2 M sodium formate	0.24	100
3	40	–	50 mM sodium phosphate	6.5	2 M sodium formate	1.4×10^{-3}	0.6
4	40	–	50 mM sodium phosphate	6.5	0.5 M NaCl	2.0×10^{-3}	0.8
5	2	–	0.1 M sodium formate	4.0	–	0.11	46
6	40	–	0.1 M sodium formate	4.0	–	0.05	21
7	2.5	–	0.5 M sodium formate	5.2	–	0.21	87
8	40	–	0.5 M sodium formate	5.2	–	0.08	33
9	40	–	0.1 M sodium acetate	4.0	–	0.01	4
10	1	5mM 4NP-β-D-Gal	50 mM sodium phosphate	6.5	2 M sodium formate	0.05	21
11	1	3mM 4NP-α-D-Gal	50 mM sodium phosphate	6.5	2 M sodium formate	2.2×10^{-3}	0.9
12	1	5 mM 4NP-β-D-Glc	50 mM sodium phosphate	6.5	2 M sodium formate	0.03	12
13	1	3 mM 4NP-α-D-Glc	50 mM sodium phosphate	6.5	2 M sodium formate	1.5×10^{-3}	0.6
14	1	5 mM 4NP-β-D-Gal	50 mM sodium phosphate	6.5	0.5 M NaCl	6.6×10^{-3}	3
15	1	5 mM 4NP-β-Glc	50 mM sodium phosphate	6.5	0.5 M NaCl	3.3×10^{-3}	1
16	1	5 mM 4NP-β-D-Gal	0.5 M sodium formate	5.2	–	2.0×10^{-3}	0.8
17	1	3 mM 4NP-α-D-Gal	0.5 M sodium formate	5.2	–	2.9×10^{-3}	1
18	1	5 mM 4NP-β-D-Glc	0.5 M sodium formate	5.2	–	1.9×10^{-3}	0.8
19	1	2 mM 4NP-α-D-Glc	0.5 M sodium formate	5.2	–	3.4×10^{-3}	1

Assays were performed at 65°C as described in the Materials and methods.

be that the leaving ability of the aglycon group in 4-nitrophenol, was insufficient to perform glycosynthesis. Hyperthermophilic β-glycosynthases efficiently produced oligosaccharides when using 2-nitrophenyl β-D-glycoside substrates, in which 2-nitrophenol, although having a pK_a similar to that of 4-nitrophenol (pK_a 7.22 and 7.18, respectively), can form a chelate ring by hydrogen bonding, thereby increasing the leaving ability upon protonation (Perugino et al. 2003).

Unfortunately, no α-L-fucosides containing aglycons with a leaving ability better than 4-nitrophenol are commercially available. Therefore, we synthesized the substrate 2-chloro-4-nitrophenyl α-L-fucopyranoside (2C4NP-Fuc; Figure 2), using a modification of the existing methods (Kasai et al. 1992; Gu et al. 2003). Commercially-available L-fucose was firstly peracetylated and then treated with 2-chloro-4-nitrophenol in the presence of $FeCl_3$ to give 2-chloro-4-nitrophenyl 2,3,4-tri-acetyl-α-L-fucopyranoside. Final deacetylation with triethylamine in 4:1 methanol/water afforded pure 2C4NP-Fuc after silica-gel column chromatography.

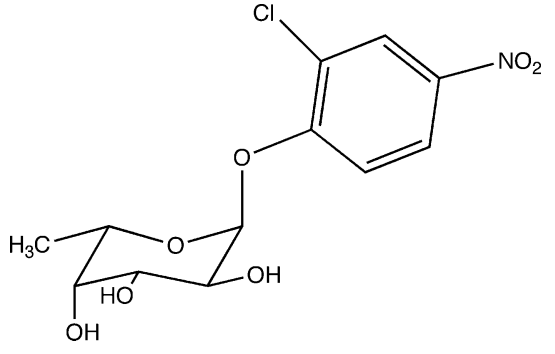


Figure 2. Chemical structure of the z-chloride-4-nitrophenyl.

The aglycon 2-chloro-4-nitrophenol has a pK_a of 5.45 (Tehan et al. 2002), which is noticeably lower than that of 4-nitrophenol (7.18); this difference making the 2-chloro-4-nitrophenol a much better leaving group. Conceivably, this could accelerate the first step of the reaction as seen for the α-L-fucosidase from human placenta (Wang & Cao 2004).

The steady-state kinetic constants at 65°C of the mutant SsαfucD242G on 4NP- and 2C4NP-Fuc were measured in 50 mM sodium phosphate buffer, pH 6.5, in the presence of 2 M sodium azide and were compared with the kinetic constants of the wild type Ssαfuc without external ions (Table II). The specificity constant k_{cat}/K_M of the mutant on 2C4N-Fuc was more than 7-fold higher than that found on the commercial substrate 4NP-Fuc as expected for an aryl-glycoside substrate with a better leaving

Table II. Steady-state kinetic constants of wild type and D242G mutant Ssαfuc.

	k_{cat} (sec^{-1})	K_M (mM)	k_{cat}/K_M (sec^{-1} mM^{-1})
Wild type			
4Np-Fuc[1]	287±11	0.028±0.004	10,250
2C4Np-Fuc	157±9	0.013±0.004	11,602
D242G+ 2M NaN_3			
4Np-Fuc	9.66±0.28	0.19±0.02	51.55
2C4Np-Fuc	55.3±2.7	0.14±0.02	384.29

Assays were performed at 65°C as described in Materials and methods.
[1]Data are from (Cobucci-Ponzano et al. 2003a).

group (Kempton & Withers 1992). This activation reflects mainly the increment in k_{cat}, while the K_M values remained almost unaltered, indicating that the affinity of the D242G mutant was the same for 4NP- and 2C4NP-Fuc. Interestingly, the wild type on the latter substrate produced kinetic constants similar to those obtained with 4NP-Fuc (Table II), indicating that the different leaving abilities of the aglycons in the two substrates did not change the reaction rates. This suggests that, for the wild type, deglycosylation is the limiting step, since the better leaving ability of the 2-chloro-4-nitrophenol aglycon, which is important for the first, glycosylating step of the reaction, did not make any difference. However, more detailed kinetic characterization is needed to confirm this hypothesis. In contrast, the presence of the activated leaving group greatly enhanced the activity of the mutant showing that the D242G mutation affects the first step of the reaction.

To test if this activation resulted in better synthetic activity we analysed the reaction products of the mutant incubated in the presence of 2.5 mM 2C4NP-Fuc, 0.5 M sodium formate buffer pH 5.2 at 65°C for 16 h, but as before, no oligosaccharide product was observed by TLC. However, the specific activity of the Ssαfuc D242G mutant under these conditions was more than 2-fold higher than that measured, under the same conditions, with 4NP-Fuc (0.5 U/mg and 0.21 U/mg on 2C4NP- and 4NP-Fuc, respectively). The activity observed with SsαfucD242G using the new substrate in the presence of formate was 9-fold lower than that found with sodium azide.

We have shown here that the impaired catalytic activity of the α-L-fucosidase nucleophile mutant can be significantly improved in the presence of external ions and of a synthetic aryl-α-L-fucoside substrate. However, the enzyme did not synthesize products and, hence, was not a glycosynthase, suggesting that a successful glycosynthetic reaction is the result of a delicate balance between the most suitable characteristics of the mutant enzyme, the substrate, and external nucleophiles. Experiments are currently underway to mutate the Asp242 and other residues in the active site, which may be important for catalytic activity or substrate recognition in Ssαfuc.

Acknowledgements

This work was supported by the project MoMa n. 1/014/06/0 of the Agenzia Spaziale Italiana. The IBP-CNR belongs to the Centro Regionale di Competenza in Applicazioni Tecnologico-Industriali di Biomolecole e Biosistemi.

References

Cobucci-Ponzano B, Trincone A, Giordano A, Rossi M, Moracci M. 2003a. Identification of an archaeal alpha-L-fucosidase encoded by an interrupted gene. Production of a functional enzyme by mutations mimicking programmed -1 frameshifting. J Biol Chem 278:14622–14631.

Cobucci-Ponzano B, Trincone A, Giordano A, Rossi M, Moracci M. 2003b. Identification of the catalytic nucleophile of the family 29 alpha-L-fucosidase from *Sulfolobus solfataricus* via chemical rescue of an inactive mutant. Biochemistry 42:9525–9531.

Cobucci-Ponzano B, Mazzone M, Rossi M, Moracci M. 2005. Probing the catalytically essential residues of the alpha-L-fucosidase from the hyperthermophilic archaeon *Sulfolobus solfataricus*. Biochemistry 44:6331–6342.

Cobucci-Ponzano B, Conte F, Benelli D, Londei P, Flagiello A, Monti M, Pucci P, Rossi M, Moracci M. 2006. The gene of an archaeal alpha-L-fucosidase is expressed by translational frameshifting. Nucleic Acids Res 34:4258–4268.

Crout DH, Vic G. 1998. Glycosidases and glycosyl transferases in glycoside and oligosaccharide synthesis. Curr Opin Chem Biol 2:98–111.

de La Torre F, Sampedro J, Zarra I, Revilla G. 2002. AtFXG1, an Arabidopsis gene encoding alpha-L-fucosidase active against fucosylated xyloglucan oligosaccharides. Plant Physiol 128:247–255.

Flitsch SL. 2000. Chemical and enzymatic synthesis of glycopolymers. Curr Opin Chem Biol 4:619–625.

Gu G, Du Y, Hu H, Jin C. 2003. Synthesis of 2-chloro-4-nitrophenyl α-L-fucopyranoside: a substrate for α-L-fucosidase (AFU). Carbohydr Res 338:1603–1607.

Karst NA, Linhardt RJ. 2003. Recent chemical and enzymatic approaches to the synthesis of glycosaminoglycan oligosaccharides. Curr Med Chem 10:1993–2031.

Kasai K, Okada K, Yamatsugu N. 1994. Preparation of p-nitrophenyl α-L-fucopyranoside derivative, reagents for determination of α-L-fucosidase activity containing the same as the active ingredient, and method for determination of α-L-fucosidase activity by using the reagent Pat Jan Appl 92353689, 15 December 1992, Jpn Kokai Tokyo Koho JP 06179690.

Kempton JB, Withers SG. 1992. Mechanism of *Agrobacterium* beta-glucosidase: kinetic studies. Biochemistry 31:9961–9969.

Koshland DE. 1953. Stereochemistry and the mechanism of enzymatic reactions. Biol Rev Camb Philos Soc 28:416–436.

Leatherbarrow RJ. 1992 GraFit version 3.0. Staines, Erithacus Software Ltd., UK.

Listinsky JJ, Siegal GP, Listinsky CM. 1998. Alpha-L-fucose: a potentially critical molecule in pathologic processes including neoplasia. Am J Clin Pathol 110:425–440.

Mackenzie LF, Wang Q, Warren RAJ, Withers SG. 1998. Glycosynthases: mutant glycosidases for oligosaccharide synthesis. J Am Chem Soc 120:5883–5884.

Moracci M, Trincone A, Perugino G, Ciaramella M, Rossi M. 1998. Restoration of the activity of active-site mutants of the hyperthermophilic beta-glycosidase from *Sulfolobus solfataricus*: dependence of the mechanism on the action of external nucleophiles. Biochemistry 37:17262–17270.

Moracci M, Trincone A, Cobucci-Ponzano B, Perugino G, Ciaramella M, Rossi M. 2001. Enzymatic synthesis of oligosaccharides by two glycosyl hydrolases of *Sulfolobus solfataricus*. Extremophiles 5:145–152.

Mori E, Hedrick JL, Wardrip NJ, Mori T, Takasaki S. 1998. Occurrence of reducing terminal N-acetylglucosamine 3-sulfate and fucosylated outer chains in acidic N-glycans of porcine zona pellucida glycoproteins. Glycoconj J 15:447–456.

Murata T, Morimoto S, Zeng X, Watanabe S, Usui T. 1999. Enzymatic synthesis of alpha-L-fucosyl-N-acetyllactosamines and 3′-O-alpha-L-fucosyllactose utilizing alpha-L-fucosidases. Carbohydr Res 320:192–199.

Noda K, Miyoshi E, Uozumi N, Gao CX, Suzuki K, Hayashi N, Hori M, Taniguchi N. 1998. High expression of alpha-1-6 fucosyltransferase during rat hepatocarcinogenesis. Int J Cancer 75:444–4450.

Okuyama, H. Mori, K. Watanabe, A. Kimura, S. Chiba. 2002. α-glucosidase mutant catalyzes 'α-glycosynthase'-type reaction. Biosci Biotechnol Biochem 66:928–933.

Perugino G, Trincone A, Giordano A, van der Oost J, Kaper T, Rossi M, Moracci M. 2003. Activity of hyperthermophilic glycosynthases is significantly enhanced at acidic pH. Biochemistry 42:8484–8493.

Perugino G, Trincone A, Rossi M, Moracci M. 2004. Oligosaccharide synthesis by glycosynthases. Trends Biotechnol 1:31–37.

Perugino G., Cobucci-Ponzano B, Rossi M, Moracci M. 2005. Recent advances in the oligosaccharide synthesis promoted by catalytically engineered glycosidases. Adv Synth Catal 347:941–950.

Perugino G, Falcicchio P, Corsaro MM, Matsui I, Parrilli M, Rossi M, Moracci M. 2006. Preparation of a glycosynthase from the β-glycosidase of the hyperthermophilic Archaeon *Pyrococcus horikoshii*. Biocat Biotrans 24:23–29.

Schweizer F, Hindsgaul O. 1999. Combinatorial synthesis of carbohydrates. Curr Opin Chem Biol 3:291–298.

Sears P, Wong CH. 2001. Toward automated synthesis of oligosaccharides and glycoproteins. Science 291:2344–2350.

Tehan BG, Lloyd EJ, Wong MG, Pitt WR, Montana JG, Manallack DT, Gancia E. 2002. Estimation of pKa using semiempirical molecular orbital methods. part 1: application to phenols and carboxylic acids. Quant Struct-Act Relat 21:457–472.

Trincone A, Perugino G, Rossi M, Moracci M. 2000. A novel thermophilic glycosynthase that effects branching glycosylation. Bioorg Med Chem Lett 10:365–368.

Trincone A, Giordano A, Perugino G, Rossi M, Moracci M. 2003. Glycosynthase-catalysed syntheses at pH below neutrality. Bioorg Med Chem Lett 13:4039–4042.

Vanhooren P, Vandamme EJ. 1999. L-Fucose: occurrence, physiological role, chemical, enzymatic and microbial synthesis. J Chem Technol Biotechnol 74:479–497.

Wang JJ, Cao E-H. 2004. Rapid kinetic rate assay of the serum α-L-fucosidase in patients with hepatocellular carcinoma by using a novel substrate. Clin Chim Acta 347:103–109.

Wang Q, Graham RW, Trimbur D, Warren RAJ, Withers SG. 1994. Changing enzymatic reaction mechanisms by mutagenesis: conversion of a retaining glucosidase to an inverting enzyme. J Am Chem Soc 116:11594–11595.

Wymer N, Toone EJ. 2000. Enzyme-catalyzed synthesis of carbohydrates. Curr Opin Chem Biol 4:110–119.