



**HAL**  
open science

# Mise au point de NéoLectines spécifiques des furanosides : bioingénierie, diagnostic et imagerie

Mateja Senicar

## ► To cite this version:

Mateja Senicar. Mise au point de NéoLectines spécifiques des furanosides : bioingénierie, diagnostic et imagerie. Sciences agricoles. Université d'Orléans, 2019. Français. NNT : 2019ORLE3037 . tel-03131503

**HAL Id: tel-03131503**

**<https://theses.hal.science/tel-03131503>**

Submitted on 4 Feb 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

**ÉCOLE DOCTORALE**

**(SANTÉ, SCIENCES BIOLOGIQUES ET CHIMIE DU VIVANT)**

**Institut de Chimie Organique et Analytique, UMR 7311, CNRS**

**Centre de Biophysique Moléculaire, UPR 4301, CNRS**

**THÈSE** présentée par:

**Mateja SENICAR**

Soutenue le: 13 décembre 2019

Pour obtenir le grade de:

**Docteur de l'Université d'Orléans**

Discipline/Spécialité: Biologie – Biochimie

**Mise au point de NéoLectines spécifiques des furanosides:  
bioingénierie, imagerie et diagnostic**

**THÈSE dirigée par:**

**Richard DANIELLOU**  
**Svetlana V. ELISEEVA**

Professeur, Université d'Orléans – directeur de Thèse  
Chargée de recherche, CNRS – co-directrice de Thèse

**RAPPORTEURS:**

**Laurent SALMON**  
**Pascale DELANGLE**

Professeur, Université Paris-Sud – Université Paris-Saclay  
Directrice de recherche, Université Grenoble-Alpes

**JURY:**

**Laurent SALMON**  
**Pascale DELANGLE**  
**Tom DESMET**  
**Stéphane PETOUD**  
**Ludovic LANDEMARRE**  
**Richard DANIELLOU**  
**Svetlana V. ELISEEVA**

Professeur, Université Paris-Sud – Université Paris-Saclay (Président du jury)  
Directrice de recherche, Université Grenoble-Alpes  
Professeur, Université de Gent  
Professeur, Directeur de recherche INSERM  
Directeur d'entreprise GLYcoDiag  
Professeur, Université d'Orléans – directeur de Thèse  
Chargée de recherche, CNRS – co-directrice de Thèse

*Multi enim sunt vocati, pauci vero electi.*

*Matthaeus 22:14*

## **Acknowledgements**

*This doctoral thesis was part of the NeoLect project and the work was supported by APR IR, Région Centre-Val de Loire, France.*

*I thank Prof. Laurent Salmon (Université Paris-Sud – Université Paris-Saclay) and Prof. Pascale Delangle (Directrice de recherche, Université Grenoble-Alpes) for accepting to review this thesis, as well as to Prof. Tom Desmet (Professeur, Université de Gent) and Prof. Stéphane Petoud (Professeur, Directeur de recherche INSERM) for accepting to be part of this thesis defence committee.*

*Thesis is life experience and job at the same like no other. I sincerely thank Prof. Richard Daniellou for giving me opportunity to be part of his Glycobio research team and to learn from him. He literally travelled on the other side of the world (Japan) to start this project. Thank You for trusting me with this thesis. I really enjoyed working on it. I hope that I was up to the expectations. Special thanks goes to Dr. Pierre Lafite and his support, advices and enthusiasm. Coming to lab every day felt like going to another family and I thank you all for that.*

*I thank Prof. Stéphane Petoud for welcoming me in his team and giving me opportunity to work with him. I specially thank Dr. Svetlana V. Eliseeva, my co-directrice, for all the help and support and for introducing me into the bright world of lanthanides. Many thanks to Dr Guillaume Collet for his help and advices during imaging experiments and for driving me to Patàpain! Thank to Anton, my CBM bureau coloc, and Keyron, bureau voisin, for the coffee breaks and fun.*

*Thought these three years many past and present members of GlycoBio team have contributed to unique work and high-spirit atmosphere. I specially thank Philippe for our daily small talks and his support to my French lessons. Many thanks to Jihen (Jedi) and her friendship during our long manips in the lab, her cheerful and energetic music choices, ensuring enthusiasm even during the most depressing days, Wednesday evenings in V&Bs and volley-ball tournaments (when my head injuries started 😊).*

*During these three years, I have shared my office, moods, ups-and downs and hopefully knowledge with many colocataires. So here it goes, season 2016/2017 (cast of the first year): Yasmina, Emilie (my best French teachers ever!) and Hyuna; season 2017/2018 (cast of the*

*second year): Pauline, Alexis, Ferdinand, Magda and Thomas; season 2018/2019 (the cast of the final season, third year): Julien, Alexis, Martyna and Tomoki (by far, the craziest combination of team members. Thanks to you, I can be assured that none day will be like the other and your jokes, humor and fun will keep my spirits high!) and of course past and present members of GlycoDiag team, Blanka and Benoit and director Mr. Ludovic Landemarre.*

*To Arnaud, my colocataire de bureau, and Damien (Dam-Dam, Benjamin, Padawan or also known as "call me by my name" 🤪), I will only say: "Le cerveau n'est pas une partie vitale, vise les bouboules plutôt" (This is exclusively internal joke, and should not be taken out of the context). See you guys!*

*I also thank my BTS stagiaires Antoine, Florian and Boris as well as my Edifice students (les petites) Martin, Alexis, Lucas, Chloé and Juliette for their participation and interest they showed during our teaching sessions.*

*Hélène and Eugénie, these three years definitely would not be the same without your company! You're absolutely fabulous! Our petite pause clope, soirée picole, petits messages "coucou" à ma voisine préférée, #Hélènechallenge and many other little inventions of ours were great fun and joy!*

*I thank Mélanie, for sharing her family time during Christmas holidays and introducing me to Saran where I enjoyed lovely Saturday mornings. I hope that mon poney préféré, JoJo, now officially retired, is happily grazing on pasture and is far from finishing as saucisson (bonne retraite à toi!).*

*Finally, gros bisous to mama & papa and their ever-present support!*

*It seems that I'm about to become a doctor!* 

*Cheers!*

*Mateja*

*October 2019*

### **Abbreviations**

<b>6xHis-tag</b> – hexahistidine fusion tag	<b>IUBMB</b> – Union of Biochemistry and Molecular Biology
<b>Araf</b> – arabinofuranose	<b>LPGs</b> – lipophosphoglycans
<b>Araf-ase</b> – arabinofuranosidase	<b>LPS</b> – lipopolysaccharide
<b>BF</b> – bright field	<b>Manf</b> – mannofuranose
<b>bp</b> – base pair	<b>MC</b> – metallacrown
<b>CAZy</b> – Carbohydrate-Active Enzymes	<b>MRI</b> – magnetic resonance imaging
<b>EC</b> – Enzyme Commission	<b>NIR</b> – near-infrared
<b>EDTA</b> – ethylenediaminetetraacetic acid	<b>NMR</b> – nuclear magnetic resonance
<b>ELISA</b> – enzyme-linked immunosorbent assay	<b>NPs</b> – nanoparticles
<b>EPS</b> – exopolysaccharides	<b>NTA-Ni<sup>2+</sup></b> – Ni <sup>2+</sup> immobilized on a nitrilotriacetic acid matrix
<b>FAD</b> – flavine adenine dinucleotide	<b>ORF</b> – open reading frame
<b>Fru<sub>f</sub></b> – fructofuranose	<b>PAMPs</b> – pathogen-associated molecular patterns
<b>Fuc<sub>f</sub></b> – fucofuranose	<b>PCR</b> – polymerase chain reaction
<b>Gal<sub>f</sub></b> – galactofuranose	<b>PDB</b> – Protein Data Bank
<b>Gal<sub>f</sub>-ase</b> – galactofuranosidase	<b>PET</b> – positron emission tomography
<b>Gal<sub>f</sub>T</b> – galactofuranosyltransferase	<b>PET/MR</b> – positron emission tomography/ magnetic resonance
<b>GFP</b> – green fluorescent protein	<b>pNP</b> – <i>para</i> -nitrophenol
<b>GHS</b> – glycoside hydrolases	<b>pNP-β-D-Gal<sub>f</sub></b> – <i>para</i> -nitrophenyl β-D-galactofuranose
<b>GIPLs</b> – glycoinositolphospholipids	<b>quinHA</b> – quinaldichydroxamic acid
<b>Glc<sub>f</sub></b> – glucofuranose	<b>SDS-PAGE</b> – sodium dodecyl sulfate-polyacrylamide gel electrophoresis
<b>GM</b> – galactomannan	<b>SPECT</b> – single-photon emission computed tomography
<b>hIntL</b> – human intelectin → intelectin-1	<b>t<sub>R</sub></b> – retention time
<b>HPLC</b> – high performance thin layer chromatography	<b>UDP-Gal<sub>f</sub></b> – uracil diphosphate galactofuranose
<b>HRMS</b> – high-resolution mass spectrometry	<b>UDP-Galp</b> – UDP-galactopyranose
<b>IMAC</b> – immobilized metal-affinity chromatography	<b>UGM</b> – UDP-galactopyranose mutase
<b>IPTG</b> – isopropyl-β-D-1-thiogalactopyranoside	<b>XCT</b> – X-ray computed tomography
<b>ITNL1</b> – intelectin-1 → human intelectin	

## Chapter I Bibliographical overview

<b>I. General introduction</b> .....	<b>1</b>
<b>II. Carbohydrates – essential molecules of life</b> .....	<b>6</b>
1. Natural glycofuranoses .....	6
2. Galactofuranose .....	7
2.1. Occurrence in nature.....	7
2.1. In bacteria.....	8
2.2. In fungi.....	10
2.3. In protozoa .....	11
<b>2.4. In other organisms</b> .....	<b>13</b>
3. Medical relevance & functional impact .....	14
3.1. Galactofuranose antigens – therapeutic & diagnostic target.....	14
3.2. Aspects of enzymatic biosynthesis & metabolism .....	16
<i>UDP-Galactopyranose Mutase (UGM)</i> .....	16
<i>Galactofuranosyltransferase (GalT)</i> .....	17
<b>II. Carbohydrate-degrading enzymes</b> .....	<b>19</b>
2. Glycosidases – a general overview.....	19
2.1. Glycofuranoside hydrolases .....	22
2.2. Galactofuranosidases .....	22
2.2.1. <i>Sources &amp; activities – the early reports</i> .....	22
2.2.2. <i>Cloning &amp; characterization of the first galactofuranosidase</i> .....	26
<b>III. Lanthanide (III) complexes &amp; nanoparticles – near-infrared emitting probes for biological imaging</b> .....	<b>27</b>
1. Optical imaging in the near-infrared domain .....	27
2. Lanthanide (III) coordination compounds – brief overview.....	28
<b>IV. Objectives of the thesis</b> .....	<b>30</b>
<b>Bibliography</b> .....	<b>36</b>

## Chapter II Galactofuranosidase – an original and specific glycoside hydrolase

<b>I. Recombinant DNA technology in protein production.....</b>	<b>46</b>
1. The basic principles of gene and plasmid manipulations.....	46
2. Expression systems for recombinant protein production .....	48
<b>II. Obtaining a desired Galf-ase plasmid construct – the applications of gene (sub)cloning techniques .....</b>	<b>50</b>
1. Plasmid construct pET-50b(+) – screening of Galf-ase gene insert bearing plasmid .....	50
2. Galf-ase gene subcloning – preparing pET-28a(+) plasmid construct.....	52
<b>III. Production and characterization of recombinant Galf-ase.....</b>	<b>54</b>
1. Overexpression and purification .....	54
2. Optimum conditions and kinetic parameters determination .....	56
2.1. Substrate specificity .....	56
2.2. Temperature and pH properties .....	57
2.3. Freeze-thaw stability.....	58
2.4. Kinetic analysis .....	60
3. Inhibition studies.....	62
4. Crystallization trials .....	64
<b>Résumé.....</b>	<b>66</b>
<b>Bibliography.....</b>	<b>67</b>

## Chapter III Tailoring galactofuranosidase activity – potential platform for new applications

<b>I. Glycosidase engineering.....</b>	<b>69</b>
1. Thioglycosides – versatile tools in glycobiology .....	69
2. Thioglycoligases.....	72
<b>II. Design and production of Galf-ase E464X mutants .....</b>	<b>74</b>
1. Selection and creation of mutagenesis points .....	74
2. Overexpression and purification of Galf-ase E464X mutants.....	76
3. Galf-ase E464X mutants kinetic parameters determination – Identification of the acid/base catalyست.....	<b>Erreur ! Signet non défini.</b>
<b>III. Application of Galf-ase E464X mutants in thioglycosides synthesis.....</b>	<b>79</b>
1. Selection of appropriate Galf-ase E464X mutant.....	79
2. Galf-ase E464Q mutant catalysed synthesis of thiogalactofuranosides.....	80
<b>IV. Conclusion .....</b>	<b>84</b>

<b>Résumé</b> .....	<b>86</b>
<b>Bibliography</b> .....	<b>86</b>
 <b>Chapter IV Gal<math>\alpha</math>-specific &amp; lanthanide (III) complex loaded nanoparticles for NIR imaging – concept validation</b>	
<b>I. Engineering neolectins from glycosidases</b> .....	<b>91</b>
1. Application in biotechnology, diagnostic and imaging .....	91
<b>II. Construction of luminescent nanoparticle scaffold for neolectin bioconjugation &amp; imaging applications</b> .....	<b>93</b>
1. Construction and characterization of luminescent nanoparticle scaffold .....	93
1.1. Encapsulation procedure.....	95
1.2. Surface functionalization .....	96
1.3. Photophysical properties of YbZn <sub>16</sub> (quinoHA) <sub>16</sub> and corresponding NPs .....	98
2. Bioconjugation of luminescent nanoparticle scaffold .....	101
2.1. Gal $\alpha$ -specific near-infrared imaging with Gal $\alpha$ -ase functionalized luminescent nanoparticles	103
<b>III. Conclusion</b> .....	<b>104</b>
<b>Résumé</b> .....	<b>105</b>
<b>Bibliography</b> .....	<b>106</b>
 <b>Chapter V Conclusion &amp; perspectives</b>	
<b>V. Conclusion &amp; perspectives</b> .....	<b>109</b>
 <b>Chapter VI Experimental part</b>	
<b>6.1. Materials and methods</b> .....	<b>117</b>
<b>6.2. Molecular biology</b> .....	<b>117</b>
6.2.1. <i>In silico</i> sequence analysis .....	117
6.2.2. Polymerase chain reaction (PCR).....	117
6.2.3. DNA analysis and purification .....	121
6.2.4. Bacterial strains and plasmids.....	122
<b>6.3. Biochemistry</b> .....	<b>123</b>
6.3.1. Protein overexpression .....	123
6.3.2. Protein analysis.....	123
6.3.3. Protein purification.....	125
6.3.4. Enzymatic activity assays .....	126
6.3.5. Furanthioglycosylation .....	130
<b>6.4. Coordination chemistry</b> .....	<b>131</b>

6.4.1. Synthesis of Yb <sup>III</sup> /Zn <sup>II</sup> metallacrowns: LnZn <sub>16</sub> (quinoHA) <sub>16</sub> .....	131
<b>6.5. Nanotechnology.....</b>	<b>132</b>
6.5.1. Encapsulation of Yb <sup>III</sup> /Zn <sup>II</sup> metallacrowns (LnZn <sub>16</sub> (quinoHA) <sub>16</sub> ) in polystyrene nanoparticles: LnX@PS/COOH .....	132
6.5.2. Surface modification LnX@PS/COOH to PS/NTA and LnX@PS/NTA.....	132
6.5.3. Synthesis of PS/NTA-Ni <sup>2+</sup> and LnX@PS/NTA-Ni <sup>2+</sup> .....	133
6.5.4. Bioconjugation of 6xHis-tagged recombinant Gal $\beta$ -ase to PS/NTA-Ni <sup>2+</sup> and LnX@PS/NTA-Ni <sup>2+</sup> .....	133
6.5.5. Characterization of PS/NTA-Ni <sup>2+</sup> & LnX@PS/NTA-Ni <sup>2+</sup> nanoparticles.....	134

## **Chapter VII Appendix**

<b>7.1. Additional figures.....</b>	<b>136</b>
7.1.1. Sequence alignment.....	136
7.1.2. Structural formulas of <i>p</i> NP-linked substrates .....	138
7.1.3. The HPLC chromatograms and ESI-HRMS spectras of thioglycosylation reactions by E464Q.....	139
7.1.5. Short communication published in the journal Carbohydrate research .....	147

## Chapter I Bibliographical overview

<b>Figure I-1.</b> Glycoside hydrolases as platform for versatile applications.....	<b>1</b>
<b>Figure I-2.</b> The carbohydrate layer on the cell surface is composed of the glycans (oligosaccharides) attached to membrane glycolipids and membrane glycoproteins. <sup>11</sup> .....	<b>6</b>
<b>Figure I-3.</b> The equilibrium representation between cyclic forms of D-galactose in solution. The rapid interconverted of cyclic, Gal $\beta$ and Gal $\alpha$ , forms through acyclic form and favorable equilibrium for Gal $\alpha$ are indicated by the arrows. <sup>19</sup> .....	<b>7</b>
<b>Figure I-4.</b> A schematic representation of <i>Mycobacterium tuberculosis</i> cell wall structure. <b>A)</b> Depiction of the major cell wall features, including peptidoglycan, arabinogalactan and mycolic acids. <b>B)</b> Detailed structure of mycobacterial arabinogalactan composed of galactan and arabinan domains. <sup>28,29</sup> .....	<b>9</b>
<b>Figure I-5.</b> A schematic representation of <i>Aspergillus fumigatus</i> cell wall structure. <b>A)</b> Depiction of the major cell wall features, including chitin, glucan and galactomannan. <b>B)</b> Structural fragment of fungal galactomannan composed of mannan and galactan units. <sup>29,38</sup> .....	<b>11</b>
<b>Figure I-6.</b> A schematic representation of main Gal $\beta$ -containing glycoconjugates present at the cell surface of leishmanial promastigote. <sup>42</sup> .....	<b>12</b>
<b>Figure I-7.</b> Chemical structures of Gal $\beta$ -containing glycoconjugates. <b>A)</b> The structure of toxin prymnesin-1, produced by <i>Prymnesiou parvum</i> . <b>B)</b> The structure of agelagalastatin, glycosphingolipid produced by <i>Agelas sp.</i> <sup>25</sup> .....	<b>13</b>
<b>Figure I-8.</b> Schematic representation of the Gal $\beta$ -glycan biosynthesis and the involvement of the three main enzymes, UDP-galactopyranose mutase (UGM), galactofuranosyltransferase (Gal $\beta$ T) and galactofuranosidase (Gal $\beta$ -ase).....	<b>16</b>
<b>Figure I-9.</b> UGM-catalyzed izomerization of UDP-Gal $\alpha$ and UDP-Gal $\beta$ .....	<b>16</b>
<b>Figure I-10.</b> Biosynthesis of Gal $\beta$ -glycan catalyzed by Gal $\beta$ T.....	<b>18</b>
<b>Figure I-8.</b> Schematic representation of <b>A)</b> inverting and <b>B)</b> retaining GH catalytic mechanis.	<b>21</b>
<b>Figure I-9.</b> Ranges of wavelengths usually used for fluorescence optical imaging. <sup>113</sup> .....	<b>27</b>
<b>Figure I-10.</b> Schematic representation of the general architecture of the luminescent Ln <sup>3+</sup> complexes and the antenna effect. <b>A)</b> Pendant antenna. <b>B)</b> Chelating antenna.....	<b>29</b>
<b>Figure I-11.</b> A graphical abstract of the main thesis objectives.....	<b>32</b>

## Chapter II Galactofuranosidase – an original and specific glycoside hydrolase

<b>Figure II-1.</b> Schematic representation of recombinant DNA cloning procedure.....	<b>47</b>
<b>Figure II-2.</b> Schematic representation of pET-28a(+) vector backbone.....	<b>48</b>

**Figure II-3.** Schematic representation of the pET vector and *E. coli* Rosetta™ (DE3) host expression system the control elements.<sup>6</sup>.....50

**Figure II-4.** Schematic representation of pET-50b(+) plasmid construct cloning/expression region coding for *N*-terminal Nus protein and double 6xHis-tagged *Galf*-ase.....51

**Figure II-5.** PCR and restriction enzyme analysis of pET-50b(+) plasmid construct. **A)** Agarose-gel electrophoresis of the PCR reaction indicating the amplification efficiency of *Galf*-ase insert gene between T7 and T7term. **B)** Agarose-gel electrophoresis of pET-50b(+) plasmid construct after EcoRI and HindIII double digestion reaction.....52

**Figure II-6.** Schematic representation of *Galf*-ase gene subcloning procedure. **A)** Preparation of pET-28a(+) plasmid construct, bearing *Galf*-ase gene insert. **B)** pET-28a(+) plasmid construct confirmation on agarose-gel electrophoresis after EcoRI and HindIII restriction enzyme analysis.53

**Figure II-7.** Confirmation of EcoRI restriction site mutagenesis on agarose-gel electrophoresis after pET-28a(+) plasmid construct EcoRI and HindIII restriction enzyme analysis.....54

**Figure II-8.** The design and production of *N*-terminal fused *Galf*-ase protein. **A)** Schematic representation of *N*-terminal single peptide 6xHis-tagged *Galf*-ase fusion protein. **B)** The SDS-PAGE analysis of *Galf*-ase after IPTG induction and NTA-Ni<sup>2+</sup>chromatography purification.....55

**Figure II-9.** Schematic representation of pET-28a(+) plasmid construct cloning/expression region coding for *N*-terminal single peptide 6xHis-tagged *Galf*-ase.....56

**Figure II-10.** Schematic representation of hydrolytic reaction between *Galf*-ase and a panel of *pNP*-linked pyranosyl and furanosyl monosaccharide.....56

**Figure II-11.** Substrate specificity screening of *Galf*-ase. **A)** Substrate specificity of *Galf*-ase tested towards 21 *pNP*-linked monosaccharide substrates. **B)** Structural formulas of *pNP*-β-D-*Galf* (1) and the *pNP*-α-L-*Araf* (2).....57

**Figure II-12.** Determination of optimum activity conditions for *Galf*-ase. **A)** Influence of pH on *Galf*-ase hydrolytic activity towards *pNP*-β-D-*Galf*. **B)** Influence of temperature on *Galf*-ase hydrolytic activity towards *pNP*-β-D-*Galf*.....58

**Figure II-13.** Effect of freeze and thaw cycles and storage conditions on *Galf*-ase residual hydrolytic activity towards *pNP*-β-D-*Galf*.....59

**Figure II-14.** Michaelis-Menten plot of *pNP*-β-D-*Galf* hydrolysis reaction catalysed by *Galf*-ase. Mean values and SD error bars were calculated from three independent experiments and from a single protein preparation.....60

**Figure II-15.** Michaelis-Menten plot of *pNP*-α-L-*Araf* hydrolysis reaction catalysed by *Galf*-ase. Mean values and SD error bars were calculated from three independent experiments and from a single protein preparation.....61

**Figure II-16.** Effects of concentration of inhibitors 3-6 on the hydrolytic activity of *Galf*-ase towards *pNP*-β-D-*Galf*. Mean values and SD error bars were calculated from three independent experiments and from a single protein preparation.....62

**Figure II-17.** Lineweaver–Burk plot of competitive inhibition of Gal $f$ -ase hydrolytic activity towards  $p$ NP- $\beta$ -D-Gal $f$  by **4**. Mean values and SD error bars were calculated from two independent experiments and from a single protein preparation.....**63**

**Figure II-14.** Photographs of the crystals obtained during Gal $f$ -ase crystallization trials. **A)** Crystals submitted to X-ray crystallography proved to be of salt origin. **B)** One of the examples of the salt crystals obtained during Gal $f$ -ase crystallography trials.....**64**

### **Chapter III Tailoring galactofuranosidase activity – potential platform for new applications**

**Figure III-1.** Structures of some naturally occurring thiosugars sharing sulfur-containing heterocycle structure.....**69**

**Figure III-2.** Structures of some naturally occurring thioglycosides.....**70**

**Figure III-3.** Proposed general mechanism for retaining thioglycoligase.....**73**

**Figure III-4.** Alignment of partial ORF Gal $f$ -ase (*Streptomyces spp.*) amino acid sequence and corresponding regions of its homologs. The conserved E464 and E573 amino acid residues are indicated with asterisks.....**74**

**Figure III-5.** Confirmation of pET-28a(+) plasmid construct integrity, after E464X mutant variants generation, on agarose-gel electrophoresis by EcoRI and HindIII restriction enzyme analysis.....**76**

**Figure III-6.** SDS-PAGE analysis of E464X Gal $f$ -ase mutant variants after NTA-Ni<sup>2+</sup> chromatography purification.....**76**

**Figure III-7.** Structural formulas of amino acids selected for the generation of E464X Gal $f$ -ase mutant variants.....**77**

**Figure III-8.** Michaelis-Menten plot of  $p$ NP- $\beta$ -D-Gal $f$  hydrolysis reaction catalysed by Gal $f$ -ase E464S, E464C and E464Q mutant variants. Mean values and SD error bars were calculated from three independent experiments (\*except E464C) and from a single protein preparation.....**78**

**Table III-2.** Comparison of Michaelis-Menten kinetics constants for hydrolysis of  $\beta$ -D-Gal $f$  by E464X Gal $f$ -ases mutant variants.....**79**

**Figure III-9.** Influence of pH on E464Q Gal $f$ -ase mutant variant hydrolytic activity towards  $p$ NP- $\beta$ -D-Gal $f$ .....**80**

**Figure III-10.** Structural formulas of thiol acceptors used for E464Q Gal $f$ -ase catalyzed furanothioglycoligase reactions.....**80**

**Figure III-11.** Proposed furanothioglycoligation mechanism catalyzed by E464Q Gal $f$ -ase mutant variant.....**81**

**Figure III-12.** HPLC chromatograms of representative thioglycosylation reactions by E464Q. **A)** Chromatogram of crude thioglycosylation reaction with thioaryl acceptor. **B)** Chromatogram of purified product fraction from thioglycosylation reaction.....**82**

<b>Figure III-13.</b> Structural formulas of synthesized furanothioglycosides.....	<b>83</b>
<b>Figure III-14.</b> The measured HRMS-ESI spectrum of thioaryl galactofuranosyl ion in negative mode. The analysed sample is the fraction purified on C18 chromatography resin (eluent solvent mixture: <b>A:B=6:4</b> ; <b>A=0.1%</b> formic acid (v/v) in water, <b>B=0.1%</b> formic acid (v/v) in acetonitrile) and corresponds to the peak ( $t_R=9.19$ min) detected on HPLC ( <b>Figure III-12. B</b> ).....	<b>83</b>
 <b>Chapter IV Galf-specific &amp; lanthanide (III) complex loaded nanoparticles for NIR imaging – concept validation</b>	
<b>Figure IV-1. A)</b> X-ray crystal structure of $DyZn_{16}(quinHA)_{16}$ along the (left) <i>a</i> -axis and (right) <i>c</i> -axis. <b>B)</b> Emission spectra of $LnZn_{16}(quinHA)_{16}$ ( $Ln = Yb, Nd, Er$ , quinHA = quinaldichydroxamic acid). <sup>12</sup> .....	<b>93</b>
<b>Figure IV-2.</b> Absorption spectra of $YbZn_{16}(HA)_{16}$ MCs (HA = pyzHA, quinHA, quinoHA) in methanol or water solutions at room temperature.....	<b>94</b>
<b>Figure IV-3.</b> Schematic representation of the three main steps in which internal and external functionalization of NPs can be achieved.....	<b>94</b>
<b>Figure IV-4.</b> Schematic representation of the encapsulation step . <b>A)</b> Top-down view of the crystal structure of $YbZn_{16}(quinoHA)_{16}$ MC obtained by X-ray diffractometry on a single crystal, done at the Department of Chemistry of the University of Michigan by Tu N. Nguyen and Jeff W. Kampf (colour code: green, Yb; yellow, Zn; red, O; blue, N; grey, C; hydrogen atoms and solvents have been omitted for clarity). <b>B)</b> Encapsulation of $YbZn_{16}(quinoHA)_{16}$ complex inside PS/COOH NPs followed by subsequent colour change to orange.....	<b>95</b>
<b>Figure IV-5.</b> Schematic representation of synthetic route for preparation of PS/NTA NPs: activation of NPs carboxylic groups with EDC/sulfo-NHS to create the amide bond with AIDA.....	<b>96</b>
<b>Figure IV-6.</b> Photographs of the pellets of PS/NTA- $Ni^{2+}$ NPs <b>(A)</b> before and <b>(B)</b> after incubation with EDTA. The change of colour indicates the release of $Ni^{2+}$ from the surface.....	<b>97</b>
<b>Figure IV-7.</b> Corrected and normalized (left) excitation ( $\lambda_{em} = 980$ nm) and (right) emission spectra of $YbZn_{16}(quinoHA)_{16}$ in DMF solution ( $50 \mu M$ , $\lambda_{ex} = 420$ nm) or for aqueous suspensions of $YbZn_{16}(quinoHA)_{16}@PS$ NPs ( $6$ mg/mL, $\lambda_{ex} = 510$ nm), room temperature.....	<b>98</b>
<b>Figure IV-8.</b> Changes in emission intensity at 980 nm for a solution of $YbZn_{16}(quinoHA)_{16}$ in DMF ( $50 \mu M$ ) or an aqueous suspension of $YbZn_{16}(quinoHA)_{16}@PS/NTA-Ni^{2+}$ NPs ( $6$ mg/mL) under continuous illumination at 405 nm.....	<b>100</b>
<b>Figure IV-9.</b> Investigation of the activity of the bioconjugated Galf-ase. <b>A)</b> The specific <i>pNP-β-D-Galf</i> activity of Galf-ase obtained after separate bioconjugation with NTA- $Ni^{2+}$ and NTA surface functionalized NPs compared to free enzyme as a control. <b>B)</b> The SDS-PAGE analysis of imidazole released Galf-ase from NTA- $Ni^{2+}$ surface functionalized NPs upon activity assay.....	<b>102</b>
<b>Figure IV-10.</b> Immobilization of 6xHis-tagged Galf-ase on NTA- $Ni^{2+}$ functionalized surface of NPs and the coordination bonds formed between NTA carboxylic groups and imidazole rings with nickel cation.....	<b>103</b>

## Chapter V Conclusion & perspectives

**Figure V-1.** A graphical abstract of the main short-term and long-term perspectives for the future work.....**112**

## Chapter I Bibliographical overview

<b>Table I-1.</b> Comparative properties of $\beta$ -D-galactofuranosidases.....	<b>25</b>
<b>Table I-2.</b> Comparative properties of recombinant $\beta$ -D-galactofuranosidases.....	<b>26</b>

## Chapter II Galactofuranosidase – an original and specific glycoside hydrolase

<b>Table II-1.</b> Comparison of Michaelis-Menten kinetics constants for hydrolysis of $\alpha$ -L-Araf (1) and $\beta$ -D-Galf (2) by different Araf-ases and Galf-ases.....	<b>61</b>
<b>Table II-2.</b> Comparison of the inhibitory effects of different galactofuranosides (3-6) towards Galf-ase and their respective $IC_{50}$ and $K_i$ values.....	<b>63</b>

## Chapter III Tailoring galactofuranosidase activity – potential platform for new applications

<b>Table III-1.</b> Codon changes and respective amino acid changes introduced at 464 position at wild-type Galf-ase amino acid sequence by site-directed mutagenesis.....	<b>75</b>
--	-----------

## Chapter IV Galf-specific & lanthanide (III) complex loaded nanoparticles for NIR imaging – concept validation

<b>Table IV-1.</b> Luminescence lifetimes ( $\tau_{obs}$ ) and absolute quantum yields ( $Q_{Yb}^L$ ) of $YbZn_{16}(\text{quinoHA})_{16}$ in DMF solution (50 $\mu\text{M}$ ) or for aqueous suspensions of $YbZn_{16}(\text{quinoHA})_{16}@PS$ NPs (6 mg/mL). <sup>a</sup> .....	<b>99</b>
---	-----------

## Chapter VI Experimental part

<b>Table VI-1.</b> Primers used for wild type Galf-ase gene amplification.....	<b>118</b>
<b>Table VI-2.</b> Cycling parameters for wild type Galf-ase gene amplification.....	<b>118</b>
<b>Table VI-3.</b> Primers used for plasmid construct pET-28a(+) gene sequencing.....	<b>118</b>
<b>Table VI-4.</b> Primers used for site-directed mutagenesis.....	<b>119</b>
<b>Table VI-5.</b> Cycling parameters for site-directed mutagenesis.....	<b>119</b>
<b>Table VI-6.</b> Primers used for deletion mutagenesis.....	<b>120</b>
<b>Table VI-7.</b> Cycling parameters for deletion mutagenesis reaction.....	<b>120</b>

# **CHAPTER I**

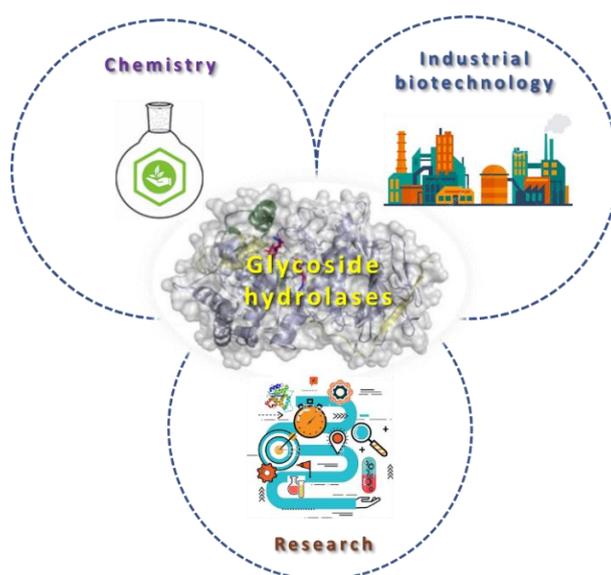
## **Bibliographical overview**

<b>I. General introduction</b> .....	<b>1</b>
<b>II. Carbohydrates – essential molecules of life</b> .....	<b>6</b>
1. Natural glycofuranoses .....	6
2. Galactofuranose .....	7
2.1. Occurrence in nature.....	7
2.1. In bacteria .....	8
2.2. In fungi.....	10
2.3. In protozoa .....	11
<b>2.4. In other organisms</b> .....	<b>13</b>
3. Medical relevance & functional impact .....	14
3.1. Galactofuranose antigens – therapeutic & diagnostic target.....	14
3.2. Aspects of enzymatic biosynthesis & metabolism .....	16
<i>UDP-Galactopyranose Mutase (UGM)</i> .....	16
<i>Galactofuranosyltransferase (GalT)</i> .....	17
<b>II. Carbohydrate-degrading enzymes</b> .....	<b>19</b>
2. Glycosidases – a general overview .....	19
2.1. Glycofuranoside hydrolases .....	22
2.2. Galactofuranosidases .....	22
2.2.1. <i>Sources &amp; activities – the early reports</i> .....	22
2.2.2. <i>Cloning &amp; characterization of the first galactofuranosidase</i> .....	26
<b>III. Lanthanide (III) complexes &amp; nanoparticles – near-infrared emitting probes for biological imaging</b> .....	<b>27</b>
1. Optical imaging in the near-infrared domain .....	27
2. Lanthanide (III) coordination compounds – brief overview.....	28
<b>IV. Objectives of the thesis</b> .....	<b>30</b>
<b>Bibliography</b> .....	<b>36</b>

## I. General introduction

Outside their role inside living organisms, the enzymes have a long history of use in human life dating back to prehistorical times when food culture began to emerge and enzyme-fermented products from various food sources were introduced.<sup>1</sup> Since then, our dependence on the enzymes and enzyme-related products has not changed, rather it is relevant more than ever. Nowadays, when environmental concerns over pollution and industrialization are growing, they offer alternative to traditional chemical treatments in production of goods.<sup>2</sup>

The impact of modern biotechnology and recombinant DNA technology has made enzymes available and economically feasible. It enabled a whole new diversity of enzymes to be accessed in the field of glycoenzymes. In parallel carbohydrate-based materials have emerged in an increasing number of applications in the food and feed, pharmaceutical or other industries. To access these carbohydrate structures in high demand, natural glycoenzyme catalysts, have provided over recent decades an alternative to chemical synthesis. Utilization of glycoside hydrolases brings all the important advantages of an enzyme catalyst. They greatly contribute to the development of novel synthetic tools and are strongly beneficial when looking for very specific sugar structure(s). Their engineering capacity gives freedom to expand their repertoire for diverse applications. To develop such innovative and versatile tools, tight interactions between scientific community and industry has to be considered (**Figure I-1.**)<sup>3,4</sup>



**Figure I-12.** Glycoside hydrolases as platform for versatile applications.

A current area of interest is broadening the search spectrum to rare glycan substrate specific glycosyl hydrolases, preferably, from narrow group of organisms involved in pathological conditions, in order to create altered characteristics, various function and application possibilities within protein engineering.

In the context of interdisciplinary approach, the NeoLect project, supported by APR IR Région Centre-Val de Loire and this doctoral thesis have been initiated. The main idea and objectives of this thesis can be outlined from its title "*Design of galactofuranosyl-specific NeoLectins: bioengineering, imaging and diagnosis*". Research work has been carried out in collaboration between two groups, Enzymology and Glycobiology group, which is part of Organic Synthesis – GlycoBio & Chemistry team at Institute of Organic and Analytical Chemistry (ICOA), and Luminescent lanthanide compounds, optical spectroscopy and bioimaging group, which is part of Chemistry, Imaging and Exobiology team at the Centre for Molecular Biophysics (CBM). Both research groups are gathered under the French National Centre for Scientific Research (CNRS).

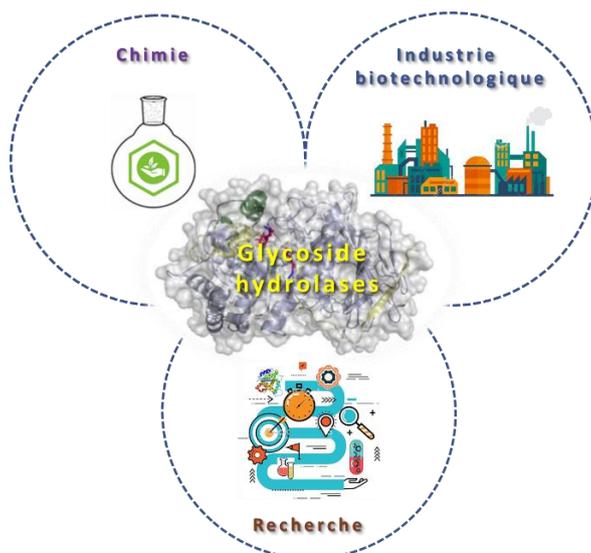
The Enzymology and Glycobiology research group of Pr. Richard Daniellou and Dr. Pierre Lafite was founded in 2012 and the group's main research interest is towards understanding the mechanisms and structures of enzymes involved in synthesis (glycosyltransferases) or cleavage (glycoside hydrolases) of glycoconjugates. The group has well established multidisciplinary procedure and uses enzymes as biocatalysts to synthesize original carbohydrate-containing compounds, mainly thioglycosides. Since different experimental techniques, including molecular biology, biochemistry and structural studies, are employed in this laboratory thus allowing us to realize the first part of the thesis objective, characterization and bioengineering of galactofuranosidase towards neolectins.

The Luminescent lanthanide compounds, optical spectroscopy and bioimaging group, led by Prof. Stéphane Petoud, aims to provide new tools for biological analysis and imagery. This specifically includes the development of near-infrared (NIR) emitting lanthanide compounds and novel sensitizing antenna. This aspect of their expertise provided us with opportunity to realize second major objective of thesis, development of NIR luminescent probe amenable for functionalization with enzymes.

## I. Introduction générale

En dehors de leur rôle dans les organismes vivants, les enzymes ont une longue histoire d'utilisation dans la vie humaine qui remonte à la préhistoire, lorsque la culture alimentaire a commencé à émerger et que des produits fermentés par enzymes provenant de diverses sources alimentaires ont été introduits.<sup>1</sup> Depuis lors, notre dépendance à l'égard des enzymes et des produits apparentés n'a pas changé, mais elle est plus pertinente que jamais. Aujourd'hui, alors que les préoccupations environnementales liées à la pollution et à l'industrialisation ne cessent de croître, elles offrent une alternative aux traitements chimiques traditionnels dans la production de biens.<sup>2</sup>

L'impact de la biotechnologie moderne et de la technologie de l'ADN recombinant a rendu les enzymes disponibles et économiquement viables. Elle a permis d'accéder à une toute nouvelle diversité d'enzymes dans le domaine des glycoenzymes. Parallèlement, des matériaux à base de glucides sont apparus dans un nombre croissant d'applications dans l'alimentation humaine et animale, l'industrie pharmaceutique ou d'autres industries. Pour accéder à ces structures glucidiques en forte demande, des catalyseurs naturels à base de glycoenzymes ont fourni au cours des dernières décennies une alternative à la synthèse chimique. L'utilisation des glycosides hydrolases apporte tous les avantages importants d'un catalyseur enzymatique. Elle contribue grandement au développement de nouveaux outils synthétiques et est très bénéfique pour la recherche de structures sucrières très spécifiques. Leur capacité d'ingénierie leur donne la liberté d'élargir leur répertoire pour des applications diverses. Pour mettre au point des outils aussi novateurs et polyvalents, il faut tenir compte des interactions étroites entre la communauté scientifique et l'industrie (**Figure I-1.**)<sup>3,4</sup>



**Figure I-1.** Les glycosides hydrolases comme plate-forme pour des applications polyvalentes.

Un domaine d'intérêt actuel est l'élargissement du spectre de recherche aux glycoside hydrolases spécifiques des substrats de glycanes rares, de préférence formant un groupe étroit d'organismes impliqués dans des conditions pathologiques, afin de créer des caractéristiques modifiées, diverses fonctions et possibilités d'application dans le génie protéique.

Dans le cadre d'une approche interdisciplinaire, le projet NeoLect, soutenu par APR IR Région Centre-Val de Loire et cette thèse de doctorat ont été initiés. L'idée principale et les objectifs de cette thèse peuvent être résumés à partir de son titre " *Mise au point de NéoLectines spécifiques des furanosides: bioingénierie, imagerie et diagnostic* ". Les travaux de recherche ont été menés en collaboration entre deux groupes, le groupe Enzymologie et Glycobiologie, qui fait partie de l'équipe Synthèse organique - GlycoBio & Chimie de l'Institut de Chimie Organique et Analytique (ICOA), et le groupe Composés luminescents de lanthanide, spectroscopie optique et bioimagerie, qui fait partie de l'équipe Chimie, imagerie et exobiologie du Centre de Biophysique Moléculaire (CBM). Les deux groupes de recherche sont regroupés au sein du Centre National de la Recherche Scientifique (CNRS).

Le groupe de recherche en Enzymologie et Glycobiologie du Pr. Richard Daniellou et Dr Pierre Lafite a été créé en 2012 et le principal intérêt de recherche du groupe est de comprendre les mécanismes et structures des enzymes impliquées dans la synthèse (glycosyltransférases) ou le clivage (glycosides hydrolases) des glycoconjugués. Le

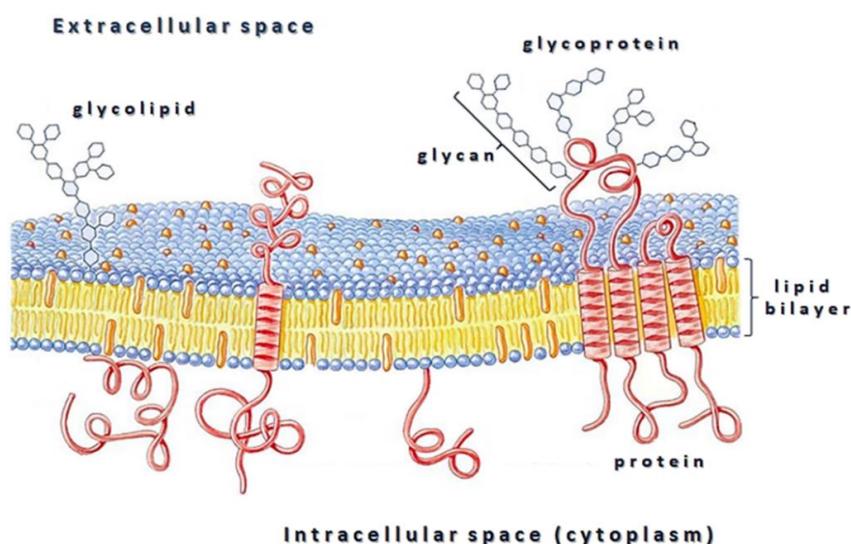
groupe a une approche multidisciplinaire bien établie et utilise des enzymes comme biocatalyseurs pour synthétiser des composés originaux contenant des glucides, principalement des thioglycosides. Ainsi différentes techniques expérimentales, incluant la biologie moléculaire, la biochimie et les études structurales, sont utilisées dans ce laboratoire, nous permettant ainsi de réaliser la première partie de l'objectif de la thèse, la caractérisation et la bioingénierie de la galactofuranosidase vers les néolectines.

Le groupe sur les composés luminescents de lanthanides, la spectroscopie optique et la bio-imagerie, dirigé par le Pr. Stéphane Petoud, vise à fournir de nouveaux outils d'analyse et d'imagerie biologiques. Il s'agit notamment de la mise au point de composés lanthanides émettant dans le proche infrarouge et d'une nouvelle antenne sensibilisatrice. Cet aspect de leur expertise nous a permis de réaliser un deuxième objectif majeur de thèse le développement d'une sonde luminescente NIR apte à la fonctionnalisation par enzymes.

## II. Carbohydrates – essential molecules of life

### 1. Natural glycofuranoses

Glycans, oligosaccharide moieties that originate from a small set of monosaccharides, make general patterns that are typically found in conjugated glycoproteins and glycolipids and usually are exposed to the extracellular side of membrane where they form exopolysaccharide or glycocalyx (**Figure I-2.**)<sup>5-8</sup> The most common monomeric constituents of glycans are hexoses, given that the monosaccharides with less than four (trioses, tetroses) or more than six carbon atoms (heptoses, octoses) are relatively rare in nature.<sup>9,10</sup>



**Figure I-13.** The carbohydrate layer on the cell surface is composed of the glycans (oligosaccharides) attached to membrane glycolipids and membrane glycoproteins.<sup>11</sup>

Putative set of monosaccharide building blocks, i.e. glucose, mannose, galactose, fructose, ribose, xylose and arabinose, constitutes glycans in organisms.<sup>12</sup> In practice, they exist in a cyclic six-membered pyranose or five-membered furanose forms, among which one, the most thermodynamically stable pyranosyl moiety, is nearly always preferred. The curiosity of glycan structures is that furanoses, with the exception of (deoxy)ribofuranose backbone of DNA and RNA, are absent in mammals, including humans.<sup>13</sup> The hexofuranose galactofuranose (Gal<sub>f</sub>), the pentofuranose arabinofuranose (Ara<sub>f</sub>), and the hexulose fructofuranose (Fru<sub>f</sub>) represent three most common furanose sugars found in glycans of plants and microorganisms.<sup>14</sup> Beside these furanoses, glucofuranose (Glc<sub>f</sub>), fucofuranose (Fuc<sub>f</sub>) and mannofuranose (Man<sub>f</sub>) are more scarcely encountered in plants or bacteria. Exploiting the differences between the human glycome and those of

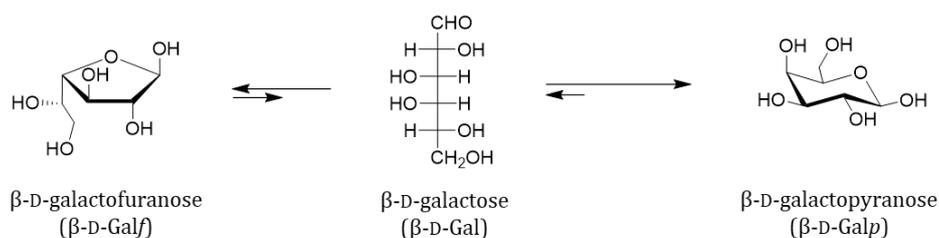
pathogenic microorganisms is a promising target for biomedical research.<sup>15,16</sup> Hence, the focus is on one particular furanose, namely galactofuranose.

## 2. Galactofuranose

### 2.1. Occurrence in nature

D-Galactose is a hexoaldose, C-4 epimer of glucose, widely distributed in nature while the existence of less common L-galactose was only found in some plant, algae and snail glycans and it can be found in both pyranosic and furanose configurations.<sup>17</sup>

Chemically, in the solid state and in the solutions, hexoses exist in a cyclic hemiacetal form and occupy either six-membered pyranose or five-membered furanose forms, where sterically and thermodynamically favoured is pyranose ring. Both cyclic forms of galactose, Gal $f$  and Gal $p$ , are interconverted through the linear form, which is present only to an instantaneous extent (**Figure I-3**).<sup>18</sup>



**Figure I-14.** The equilibrium representation between cyclic forms of D-galactose in solution. The rapid interconverted of cyclic, Gal $f$  and Gal $p$ , forms through acyclic form and favorable equilibrium for Gal $p$  are indicated by the arrows.<sup>19</sup>

Nevertheless, galactofuranose is by far the most widespread hexofuranose in nature. This thermodynamically disfavoured hexofuranose is absent in mammalian glycoconjugates, but is present in other domains of life ranging from archaea and bacteria, to protozoa, fungi, and plants where it forms glycosidic linkages mainly in  $\beta$ -anomeric configuration, although several examples in which  $\alpha$ -anomeric configuration occurs can be found. Distinctly, in mammals, only galactopyranose is found as a ubiquitous form. Although exceptionally it can be present as free monosaccharide, in nature, it is commonly linked to other molecules like glucose, building simple disaccharide structures (lactose or milk sugar), or as constituent of glycans in complex glycoproteins (ABO blood group antigens) and glycolipids.<sup>17</sup>

Interest for furanoses dates back to the beginning of the 20th century and the very starts of carbohydrate chemistry when the first synthetic methods for selective glycofuranoside

synthesis were developed.<sup>20</sup> Galactofuranose was first identified in 1937 as a component of fungal extracellular polysaccharide, galactocarlose, produced by *Penicillium charlesii*.<sup>21,22</sup> Only year later, synthesis of the first  $\beta$ -D-galactofuranosides was reported.<sup>23</sup> Over the decades, galactofuranose has been found in many naturally occurring molecules originating from variety of organisms, not necessarily pathogenic. Lately, interest for this unusual carbohydrate has not diminished due to findings of its xenobiotic and immunogenic properties. In the following sections occurrence of  $\beta$ -D-Galf units will be presented. Here, the intention is not to be exhaustive and give a complete overview of all natural Galf-containing structures, but rather to focus on major structures, particularly found in prokaryotes and two different classes of eukaryotes, fungi and protozoa, and to illustrate them with representative and most studied examples. Thereby, the focus will mainly be on pathogenic organisms such are bacteria *Mycobacterium tuberculosis*, fungus *Aspergillus fumigatus* and protozoa *Leishmania major*.

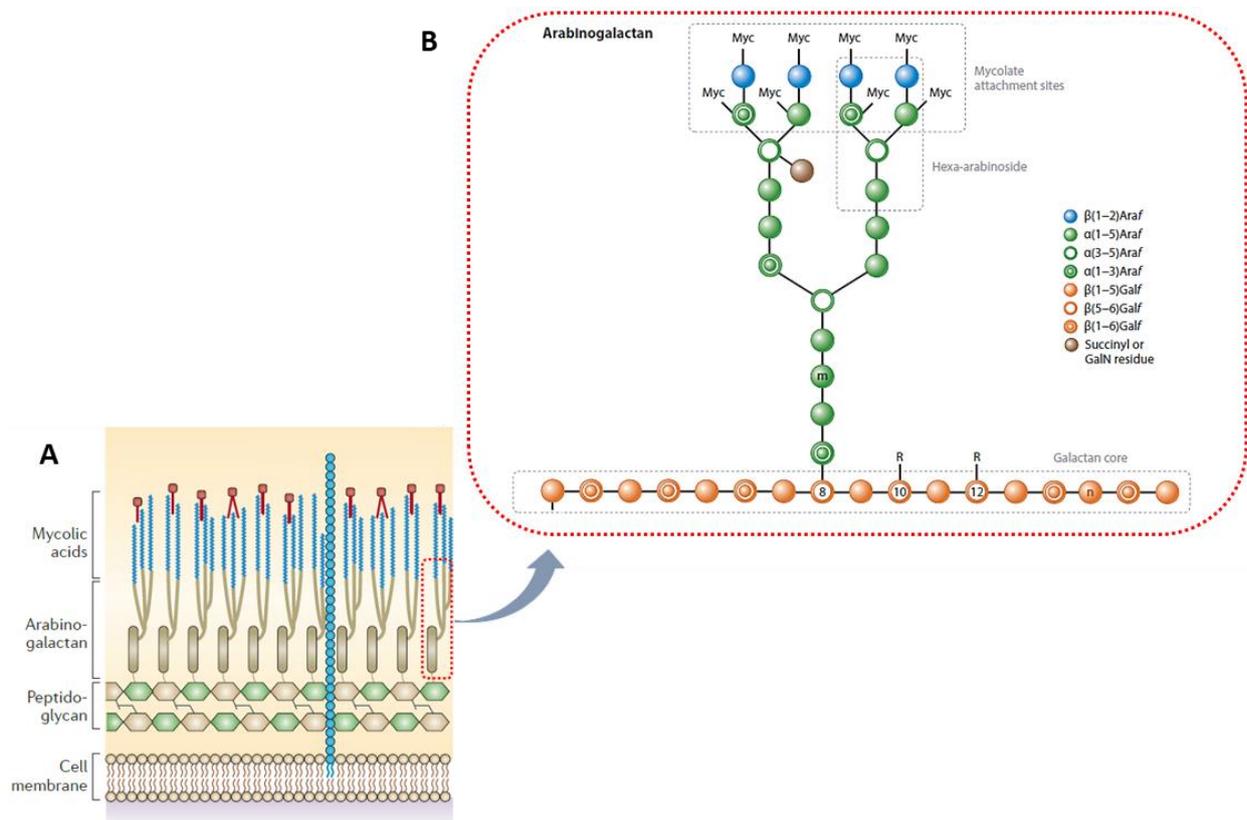
### 2.1. In bacteria

In both Gram-positive and Gram-negative bacteria, Galf is either part of a homopolymer, with a Galf disaccharide as a repetitive unit, or part of a heteropolymer, linked to another monosaccharide frequently galactopyranose, forming a regular glycans that are sometimes branched.<sup>24</sup> These glycans, composed of Galf units, are usually part of complex glycoconjugates structures that are building bacterial cell wall. Some of these bacteria are highly pathogenic and presence of Galf-containing conjugates appears to be essential for their virulence. These include lipopolysaccharide (LPS) O-antigens of *Escherichia coli* K-12 and enteroinvasive *E. coli* O164 as well as the galactan-I repeating unit of O-antigen from *Klebsiella pneumoniae*.<sup>25</sup>

The notable example of highly complex structure composed largely of rare carbohydrates and lipids is the mycobacterial cell envelope where Galf has crucial constructive role.<sup>26</sup> After several decades of successful chemotherapeutic treatment and vaccination prevention, *Mycobacterium tuberculosis* has emerged as multidrug resistant and tuberculosis become one of major causes of mortality worldwide.<sup>27</sup> The cell envelope of *Mycobacterium tuberculosis* is very thick, hydrophobic structure displaying very limited permeability. It consists of three main structural components, typical prokaryotic plasma membrane, cell wall and outer membrane, also referred as mycomembrane.<sup>28</sup> The entire

complex of cell wall is noncovalently attached to plasma membrane with the bottom peptidoglycan layer, which *via* disaccharide linker connects to the central glycan core, highly branched arabinogalactan polysaccharide. In addition, two-thirds of arabinans are terminated by a cluster of mycolic acids that form outer membrane layer and capsular segment containing variety of loosely attached proteins, lipids and polysaccharides (**Figure I-4. A**).<sup>28</sup>

Arabinogalactan is cell wall core assembly, composed of arabinofuranose and galactofuranose that make two distinct structural motifs, arabinan and galactan. The galactan is linear polymer and consists of approximately 30 alternating  $\beta$ -D-Galf-(1 $\rightarrow$ 5) and  $\beta$ -D-Galf-(1 $\rightarrow$ 6) residues that bear highly branched arabinan chains which inner backbone contains about 35 (1 $\rightarrow$ 5)- $\alpha$ -D-Araf residues (**Figure I-4. B**).



**Figure I-15.** A schematic representation of *Mycobacterium tuberculosis* cell wall structure. **A)** Depiction of the major cell wall features, including peptidoglycan, arabinogalactan and mycolic acids. **B)** Detailed structure of mycobacterial arabinogalactan composed of galactan and arabinan domains.<sup>28,29</sup>

Mycobacterial cell wall is essential for its viability and largely is responsible for the ability to survive and escape the hostile environment of the macrophage. It is evident that  $\beta$ -D-

Galf is an anchoring platform for arabinogalactan complex, therefore it remains a central focus as drug target.<sup>27,30</sup>

On the other hand, the motif  $\beta$ -D-Galf-(1→6)- $\alpha$ -D-Galf was found in exopolysaccharides (EPS) of non-pathogenic strains such as *Lactobacillus rhamnosus*, isolated from human intestinal flora,<sup>23</sup> *Bifidobacterium catenulatum*<sup>31</sup> or *Streptococcus thermophilus*, produced in skimmed milk.<sup>32</sup> It is also interesting to note that the sulfated polysaccharide-peptidoglycan complex, produced by soil bacteria *Arthrobacter sp.*, contains a repeating trisaccharidic motif and shows antitumor activity.<sup>33</sup>

Although  $\alpha$ -D-Galf is rare, it is sometimes present in previously mentioned structures that are incorporated in glycan core. This is the case in the anaerobic eubacteria *Clostridium thermocellum* and *Bacteroides cellulosolvens*, which produce a high molecular-mass cellulose binding protein complex, cellulosome. The latter is glycosylated and within tetrasaccharide units,  $\alpha$ -D-Galf is incorporated.<sup>24</sup>

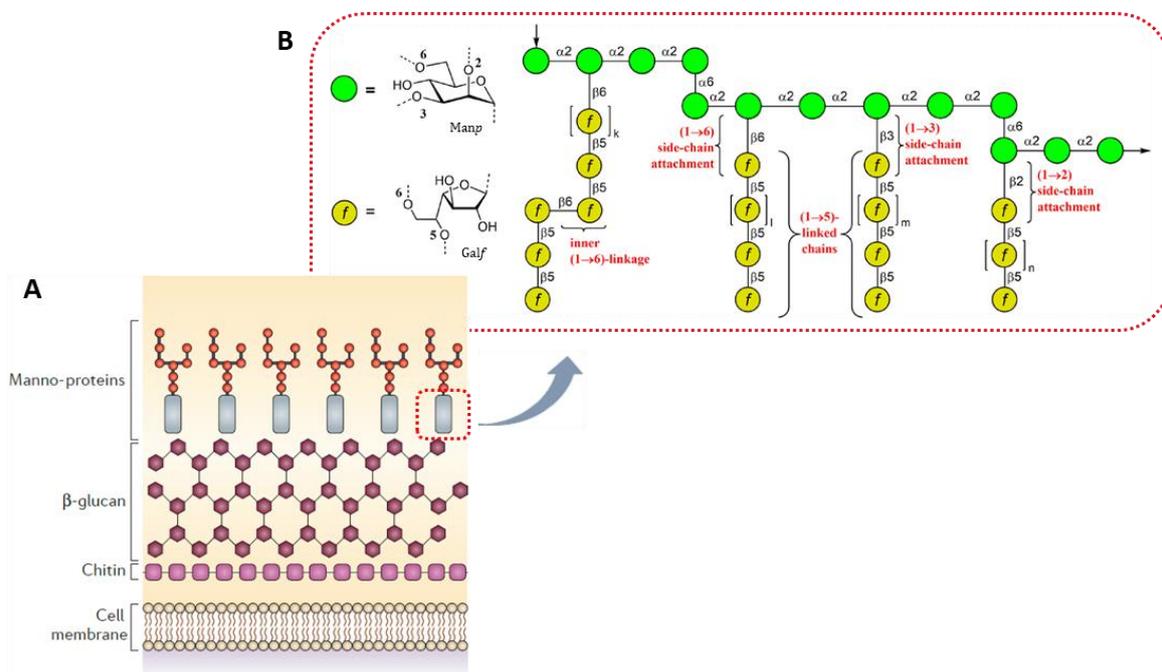
## 2.2. In fungi

Fungal cell wall structural organization and composition is generally layered. It consists of the inner, relatively conserved chitin core followed by branched glucan, which is bound to the outer, more heterogeneous layer composed of proteins and/or other polysaccharides whose composition depends on the fungal species (**Figure I-5. A**).<sup>34</sup>

In pathogenic fungi *Penicillium*, *Histoplasma*, *Cryptococcus* and *Aspergillus species* Galf has been found as part of cell wall glycan.<sup>35</sup>

*Aspergillus fumigatus* causes aspergillosis, one of the main of fungal infections in humans, especially in immunodeficient patients. It is one of the most studied fungus due to the presence of  $\beta$ -D-Galf units. Interestingly, single Galf is essential for immunoreactivity of the glycan and therefore they have been studied as *in vitro* markers for the early diagnosis of the invasive aspergillosis.<sup>24</sup> It is remarkable since 5% of the dry weight of *A. fumigatus* consists of Galf, and it is quite abundant monosaccharide found and described in at least four different molecules, galactomannan and its secreted form, glycoproteins, and within glycosphingolipids. Galactomannan is a main exopolysaccharide structural form of the cell wall, composed of (1→2)- and (1→6)- $\alpha$ -D-mannopyranoside core chain with around four to ten units of (1→5)- $\beta$ -D-Galf residues (**Figure I-5. B**). This galactomannan is also

secreted and can be found as a free, soluble polysaccharide in the medium. In addition,  $\beta$ -Gal $f$ -(1 $\rightarrow$ 5)- $\beta$ -Gal $f$  was identified in both *O*- and *N*-linked glycans from the peptidogalactomannan and glycoinositolphosphoceramide.<sup>36,37</sup>



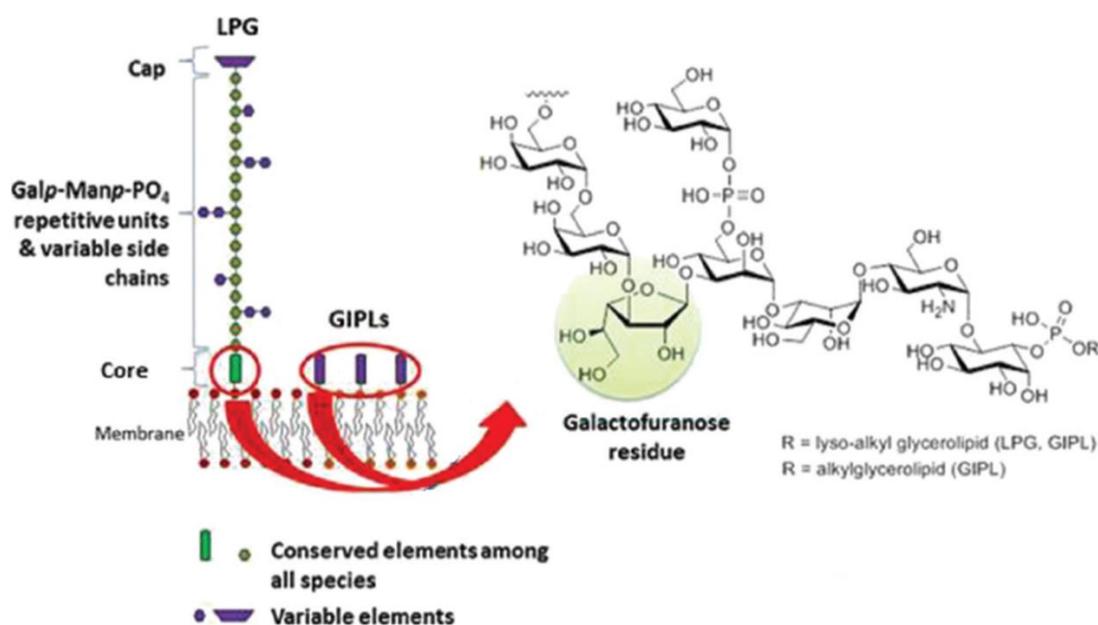
**Figure I-16.** A schematic representation of *Aspergillus fumigatus* cell wall structure. **A)** Depiction of the major cell wall features, including chitin, glucan and galactomannan. **B)** Structural fragment of fungal galactomannan composed of mannan and galactan units.<sup>29,38</sup>

Other, noteworthy examples of  $\beta$ -Gal $f$ -containing molecules present in other fungal species, include less common  $\alpha$ - $\beta$ -Gal $f$  found together with  $\beta$ - $\beta$ -Gal $f$  in varianose, extracellular polysaccharide produced by *Penicillium varians*.<sup>39</sup> The same repeating units of varianose have been found in the cell wall exopolysaccharides of *P. vermiculatum* and *Talaromyces flavus* while *Apodus deciduus* contains (1 $\rightarrow$ 2)-Gal $f$  disaccharide only in  $\alpha$ -configuration.<sup>24</sup>

### 2.3. In protozoa

The presence of Gal $f$  in eukaryotic protozoa was reported in the early 1980s in family of *Trypanosomatidae*, which includes for humans and animals parasitic genera *Trypanosoma* and *Leishmania*. The glycoconjugates in these parasites have been in detail characterized and apparently, no other examples of protozoa except trypanosomatids *Crithidia spp.* and *Endotrypanum schaudinni* containing Gal $f$  have been reported.<sup>40,41</sup>

Protozoan parasites of the genus *Leishmania* have two different life cycle stages, flagellate promastigote stage in insects and amastigote stage in mammalian phagocytic cells, during which many morphological and structural changes of membrane glycoalyx occur. The related Gal $\beta$  glycosylated membrane macromolecules are well studied in *Leishmania major*, causative agent of leishmaniosis, and include two types of molecules, lipophosphoglycans (LPGs) and glycositolphospholipids (GIPLs). LPG is a major glycoconjugate that covers entire surface of promastigote and contains internal hexasaccharide  $\beta$ -Gal $\beta$ (1 $\rightarrow$ 3)-Man $\beta$  core (**Figure I-6.**), which is conserved among all *Leishmania* species. The GIPLs are present at the membrane surface ten times more than LPGs, and contain the same repeating unit, only externally positioned. Both glycolipid complexes LPG and GIPL are involved in the virulence and survival of the parasite, and their unconventional hexasaccharide core has become target for the search for new drugs.<sup>42,43</sup>

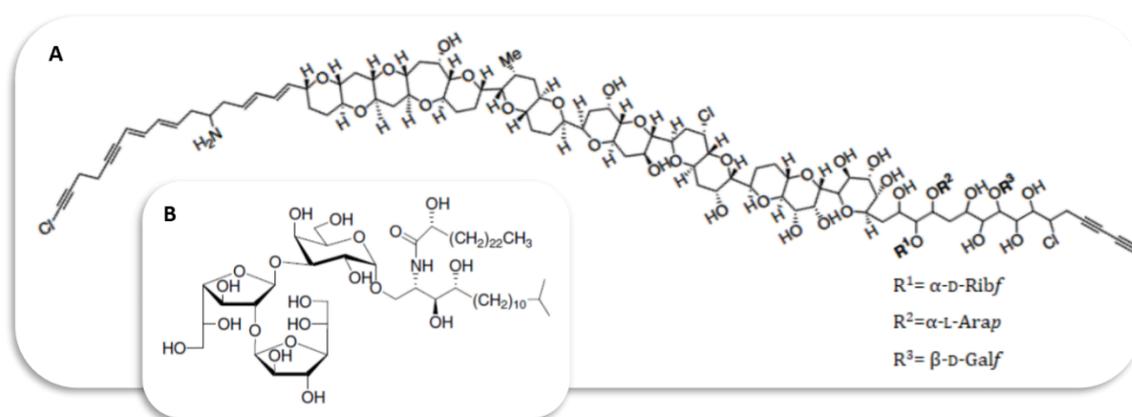


**Figure I-17.** A schematic representation of main Gal $\beta$ -containing glycoconjugates present at the cell surface of leishmanial promastigote.<sup>42</sup>

## 2.4. In other organisms

Beside previously mentioned groups of organisms, other eukaryotic organisms expressing different Gal $\beta$ -containing surface molecules include certain plants, lichens and marine organisms. Even though rare in terrestrial plants, Gal $\beta$  was first described in 1981 as part of the cell wall glycoproteins in the unicellular green algae *Chlamydomonas reinhardtii*.<sup>44</sup> Molecularly unique, prymnesin-1 toxin, isolated from red tide alga *Prymnesiou parvum*, contains  $\beta$ -D-Gal $\beta$  residue linked to C<sub>90</sub> unbranched carbon chain (**Figure I-7. A**).<sup>45</sup>

Among marine organism, especially sponges of the genus *Agelas sp.* glycosphingolipid agelagalastatin was described to contain the trisaccharide  $\alpha$ -Gal $\beta$ (1 $\rightarrow$ 2)- $\beta$ -Gal $\beta$ (1 $\rightarrow$ 3)-Gal $\beta$  and it showed antitumoral activity (**Figure I-7. B**).<sup>46</sup> Steroidal glycosides and ganglioside containing Gal $\beta$  were isolated from the starfish *Anthenea chinensis* and *Achanthaster planci* respectively.<sup>25,24</sup> The presence of Gal $\beta$  was also reported in the nematode *Caenorhabditis elegans*, one of the model organisms in molecular and animal biology studies but no reports about the presence in the mammals, including humans, has been reported so far.<sup>47</sup>



**Figure I-18.** Chemical structures of Gal $\beta$ -containing glycoconjugates. **A)** The structure of toxin prymnesin-1, produced by *Prymnesiou parvum*. **B)** The structure of agelagalastatin, glycosphingolipid produced by *Agelas sp.*.<sup>25</sup>

### 3. Medical relevance & functional impact

#### 3.1. Galactofuranose antigens – therapeutic & diagnostic target

The search for molecules unique to pathogenic microorganisms, pathogen-associated molecular patterns (PAMPs), preferably surface-exposed, conserved in pathogens, and absent in host organism, to circumvent potential risk of interference, has seemingly led to one potential target, galactofuranose.<sup>48,49</sup>

Biomolecules involving galactofuranose have attracted interest since they fulfil these requirements and their presence in many pathogens suggests to be advantageous element of survival and is considered essential for their virulence.<sup>25</sup> Hence, galactofuranose is as a new interesting candidate for a target in medical or biotechnological applications.

Possibilities to exploit the absence of this unusual monosaccharide in mammals, arises from its biological significance in these pathogens. Impact of galactofuranose deficiency on cell morphology and growth and its role in virulence was focus of numerous research studies, predominately done on eukaryotic pathogens of *Aspergillus* and *Leishmania* species. The modifications of the cell surface were most evident in fungi and resulted in aberrant morphological changes and growth reduction, leading to hypersensitivity to drugs and osmotic stress. The lack of Galf had variable impact on the virulence capacity of *Leishmania*, *Trypanosoma* and *Aspergillus* species. While in *Leishmania major* Galf-deficient mutants presented only the initial delay in infection onset, in *L. mexicana* infectivity was not attenuated and in the related parasite *T. cruzi*, Galf-containing strains were less infective than those expressing Galp. On the other hand *A. fumigatus* displayed doubtly attenuated virulence, which might be dose-related.<sup>19</sup>

Interestingly, the virulence effects vary within the observed organisms, indicating that Galf involvement cannot be generalized and rather should be considered separately, related to specific pathogen species. Understanding the exact role and contribution of individual Galf-containing glycoconjugates and Galf itself on morphology, survival and virulence is not unanimous as well as its role in the immune response and remains to be clarified.

Therefore, medical application of Galf-based therapeutic or mimetics is still challenging and relies on elucidation of the Galf biosynthesis pathways.

However, the importance of *Gal*f as diagnostic target showed to be very useful. The presence of *Aspergillus* exoantigens of galactomannan (GM) origin, known to be secreted by the fungus during the growth *in vitro* and *in vivo*, is specific indicator of this invasive disease and become detection target.<sup>50</sup> Monoclonal antibody detection methods for early serological diagnosis of galactomannan antigens, thus invasive pulmonary aspergillosis, have been experimentally developing since 1980s. It was only until 1995 that a double-direct sandwich enzyme-linked immunosorbent assay (ELISA), that employs a rat anti-GM monoclonal antibody, EB-A2, directed against the  $\beta$ -Gal*f*(1→3)- $\beta$ -Gal*f* epitopes of GM was developed. Today, two GM antigen detection kits are commercially available, the Pastorex *Aspergillus* and the Platelia™ *Aspergillus*. The Pastorex, latex agglutination test, has been for the most part replaced by the Platelia™ *Aspergillus*, EIA, which has been available in Europe for more than 20 years and in the USA since 2003. Apparently, these are the only commercially available tests specifically based on the detection of *Gal*f epitopes.<sup>51,52</sup>

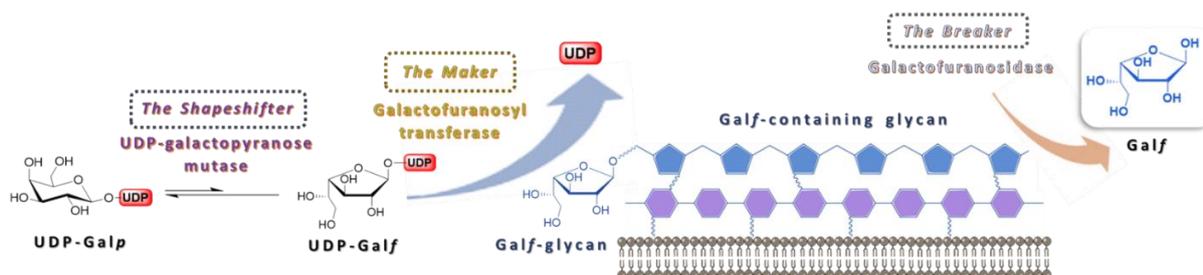
In parallel, a novel experimental procedure for non-invasive detection of *Aspergillus* lung infection, based on antibody-guided positron emission tomography and magnetic resonance (immunoPET/MR) imaging, has been developed and tested. In 2016 the prototype version, [<sup>64</sup>Cu]DOTA-mJF5 tracer showed that a mouse monoclonal antibody (mJF5) specifically binds to mannoprotein antigen, pathogen related only, and that the antibody-labeled, radionucleotide <sup>64</sup>Cu and DOTA chelator complex, allows PET imaging. This highly specific [<sup>64</sup>Cu]DOTA-mJF5 tracer allows repeated imaging and distinguishes aspergillosis from pulmonary inflammation and from bacterial lung infections.<sup>43</sup>

Only a year later, the same team reported the development of a humanised version of the JF5 antibody (hJF5). This new, [<sup>64</sup>Cu]NODAGA-hJF5 tracer showed not only improved imaging capabilities but also high specificity towards Gal*f*(1→3)- $\beta$ -Gal*f* epitopes, present in a mannoprotein antigen released by *Aspergillus* during lung infection. This was the first time that *Gal*f-specific, antibody-guided *in vivo* imaging has been used for non-invasive pre-clinical diagnosis of a fungal lung disease.<sup>54</sup>

These recent experimental techniques based on *Gal*f specific antibodies are still developing and are greatly contributing towards more specific targeting of epitope patterns. The related imaging and diagnostic aspects are yet to be explored.

### 3.2. Aspects of enzymatic biosynthesis & metabolism

The uniqueness of *Galf* as a component of important surface glycoconjugates of human pathogen has led to increased interest for the elaboration of its biosynthesis. The detailed mechanism of its biosynthesis and metabolism have been difficult to elucidate due to the instability of *Galf* itself and the fact that few of these enzymes have been isolated and studied in purified form. From the studies available so far, it is evident that two putative enzymes, UDP-galactopyranose mutase (UGM) and galactofuranosyltransferase (*GalfT*) catalyse two synthesis steps and give a rise to a galactofuran conjugates, while the involvement of galactofuranosidase (*Galf*-ase), as degrading enzyme, remains unspecified (**Figure I-8**).<sup>55</sup> In addition to the confirmation of canonical metabolism scheme and absence of alternative pathways, identification of three classes of enzymes involved and coding genes, the efforts to understand their *in vivo* interaction and importance for viability or virulence, possibly in one organism, are still ongoing.

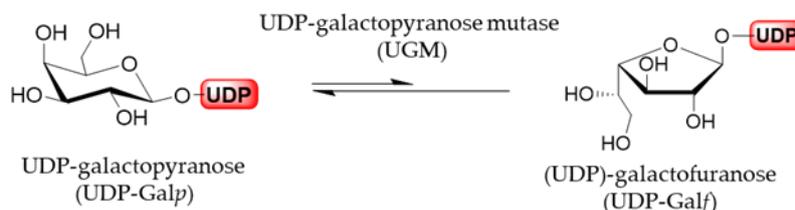


**Figure I-8.** Schematic representation of the *Galf*-glycan biosynthesis and the involvement of the three main enzymes, UDP-galactopyranose mutase (UGM), galactofuranosyltransferase (*GalfT*) and galactofuranosidase (*Galf*-ase).

#### **UDP-Galactopyranose Mutase (UGM)**

Generally, furanoses, as nucleotide activated monomer donors, are transferred to the acceptors, growing oligomers or polymers, by glycofuranosyltransferases. *In vivo*, these furanose donors are generated from corresponding pyranoses by the activity of pyranose mutases. The sole source of all *Galf* residues is uracil diphosphate (UDP)-galactofuranose (UDP-Galf) which originates directly from UDP-galactopyranose (UDP-Galp). In addition, UDP-Galp is biosynthesized in all species from glucose, originating from *de novo* synthesis pathway, or in some organisms it may be formed from free galactose by galactose salvage pathway. The activated UDP-Galp is interconverted, through reversible ring-contraction, into UDP-Galf by the cytosolic enzyme UDP-Galp mutase. The equilibrium of the UGM-

catalyzed reaction greatly favours thermodynamically more stable UDP-Galp by the 11:1 ratio (**Figure I-9**).<sup>56</sup>



**Figure I-9.** UGM-catalyzed isomerization of UDP-Galp and UDP-Galf.

It was in 1971 first established in *Salmonella typhimurium* that UDP-Galp is precursor of UDP-Galf. More than twenty years later, in 1996, the *glf* gene encoding for a UGM was first identified and cloned from *E. coli*.<sup>57</sup> Soon after the existence of homologue genes and UGMs was confirmed and characterized in several other prokaryotes, bacteria *Klebsiella pneumoniae*, *Mycobacterium tuberculosis* and *Campylobacter jejuni*, as well as in protozoan eukaryotes *Leishmania major* and *Trypanosoma cruzi*, fungus *Aspergillus fumigatus* and nematode *Caenorhabditis elegans*.

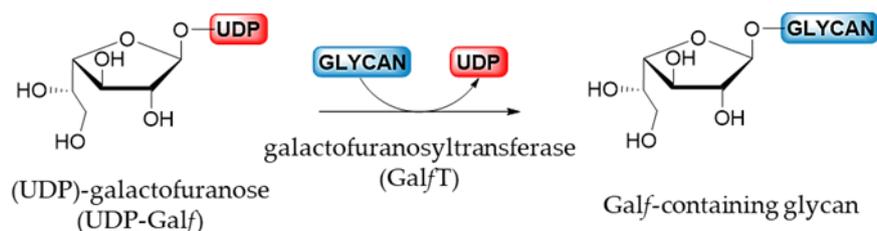
In 2001, the first crystal structure of UGM from *E. coli* was reported. Currently, a total of 59 crystal structures of UGM from nine organisms have been deposited in the Protein Data Bank (PDB). Some of these structures have been crystalized in both, the active (reduced) and the inactive (oxidized) state, as well as complexed with different substrates. The structures gave insight into its tertiary structure and revealed that overall architecture of active site consists of conserved amino acid residues and that generally, UGMs are flavoenzymes working by the unique mechanism involving flavine adenine dinucleotide (FAD) cofactor in its reduced form.<sup>58,59</sup>

Since the UGM is at the centre of *Galf* biosynthesis, to date mutagenesis or deletion of genes encoding it continues to facilitate the characterization of the impact of *Galf* absence on *in vivo* organism integrity.

### ***Galactofuranosyltransferase (GalfT)***

After the isomerization catalyzed by UGM, the newly produced UDP-Galf is transported, *via* UDP-Galf transporter, only in eukaryotes, from the cytosol into the Golgi apparatus where the glycosylation by galactofuranosyltransferases takes place. *GalfT*s are the final enzymes involved in the biosynthesis of *Galf*-containing molecules by catalyzed reaction

that mechanistically corresponds to the nucleophilic substitution of an acceptor on the electrophilic anomeric position of an activated sugar donor (**Figure I-10.**). Probably due to the high costs and difficulties in obtaining the activated donor combined with the limited access of the relevant enzymes, Gal $f$ Ts were less studied than the mutases. However, at the moment only a few Gal $f$ Ts of prokaryotic, exclusively bacterial, and eukaryotic origin have been cloned and characterized.



**Figure I-10.** Biosynthesis of Gal $f$ -glycan catalyzed by Gal $f$ T.

By far the most studied prokaryotic Gal $f$ Ts are Gal $f$ T1 and Gal $f$ T2, two essential transferases for the biosynthesis of *Mycobacterium tuberculosis* galactan. Interestingly, they share a low sequence homology and are coded by different genes, Gal $f$ T1 is encoded by *Rv3782* and Gal $f$ T2 by *Rv3808c* gene respectively. Both transferases have been expressed as recombinant enzymes and characterized. The subsequent studies elucidated the role of each transferase and demonstrated that Gal $f$ T1 is responsible for the transfer of the two first Gal $f$  on the acceptor substrate, whereas Gal $f$ T2 pursues the polymerization of the resulting acceptor by introducing approximately 30 remaining monosaccharides.

Since both use UDP- $\alpha$ -D-galactofuranose as the donor and the change from  $\alpha$ -stereochemistry in UDP-Galf to the  $\beta$ -stereochemistry of the newly synthesized glycan, indicates that catalysis follows an inverting mechanism. They are also characterized as bifunctional since they synthesize both  $\beta$ -Gal $f$ -(1 $\rightarrow$ 5)-Gal $f$  and  $\beta$ -Gal $f$ -(1 $\rightarrow$ 6)-Gal $f$  linkages between Gal $f$  residues. Gal $f$ T2 has been more extensively studied than Gal $f$ T1 therefore it is the only Gal $f$ T with a determined crystal structure and confirmed presence<sup>50,51</sup> of only one catalytic site.<sup>60,61</sup>

Few other well-characterized prokaryotic Gal $f$ Ts include WbbI from *Escherichia coli* K-12 that is able to transfer  $\beta$ -(1 $\rightarrow$ 6)-Gal $f$  to  $\alpha$ -glucose<sup>62</sup> and WbbO from *Klebsiella pneumoniae*, another bifunctional transferase that couples  $\beta$ -Gal $f$ -(1 $\rightarrow$ 6)-Gal $p$ .<sup>63</sup>

Until recently, LPG1 from *Leishmania major* was the first and only described eukaryotic Gal $f$ T. In the extensive study from 2018, all the all four putative transferases (LPG1,

LPG1G, LPG1L and LPG1R) encoded in the *L. major* genome were cloned, overexpressed and their kinetic parameters determined, thus demonstrating that they are able to use both UDP-Galf and UDP-Galp as donor substrates.<sup>64</sup>

Surprisingly another two known eukaryotic GalT's are of fungal origin, GfsA from *Aspergillus nidulans* and AfgfsA from *A. fumigatus*. As it was shown, both enzymes are localized in the Golgi apparatus and use UDP-Galf as sugar donor.<sup>65,66</sup> Moreover galactofuranosyltransferase activity was demonstrated in *P. fellatum* through incorporation of radiolabelled Galf into peptidophosphogalactomannan prepared from its membrane but no further investigations were taken.<sup>67</sup>

## II. Carbohydrate-degrading enzymes

### 2. Glycosidases – a general overview

Given large variety of monosaccharide structures and types of molecules, they can form linkages with, there is equally great number of enzymes acting on these most structurally diverse set of substrates. Collectively designated as carbohydrate-active enzymes, they catalyze the breakdown, biosynthesis and/or modification of complex carbohydrates and glycoconjugates. Reflecting their enzymatic activity, they are recognised as glycoside hydrolases, glycosyltransferases, polysaccharide lyases, carbohydrate esterases and glycoside phosphorylases.<sup>68,69</sup>

Glycoside hydrolases (glycosidases or carbohydrases, GHs) comprise a widespread group of enzymes that selectively hydrolyse glycosidic bonds in glycans (carbohydrate-carbohydrate moiety) and glycoconjugates (carbohydrate-noncarbohydrate moiety) leading to the formation of a carbohydrate hemiacetal or hemiketal and the corresponding free aglycon.<sup>70</sup> They can catalyze hydrolysis of *O*-, *N*- and *S*-linked glycosides, among which the *O*-glycoside hydrolases are the most studied. The *N*-glycoside hydrolases (EC 3.2.2) are mainly enzymes of nucleic acid metabolism and only known *S*-glycoside hydrolase, thioglucosidase (EC 3.2.1.147) has a wide specificity for different thioglycosides.<sup>71</sup>

The *O*-glycosidic bond, is one of the most stable of the covalent linkages within naturally occurring biopolymers used by organisms for long-term storage of energy and information. The half-lives for spontaneous hydrolysis, in the acidic conditions, of a simple glycoside is approximately 5 million years. Taken the stability of these polymers in

physiological pH values, the glycosidic bonds appears to be even more stable to spontaneous hydrolysis than the phosphodiester bonds in DNA or peptide bonds in proteins, proofing GHs to be one the most proficient catalysts by accelerating the hydrolytic reaction rate up to  $10^{17}$  fold.<sup>72,73</sup>

Since they are key enzymes of carbohydrate metabolism that are widely distributed in Nature and are found in almost all living organisms with the exception of some Archaea and a few unicellular parasitic eukaryotes. For most organisms, 1% of their genes encode glycoside hydrolases, including some viruses.<sup>74</sup> Therefore, their function or dysfunction has been implicated in a number of different diseases such as lactose intolerance and lysosomal storage diseases.

Due to the great variety and complexity of their natural carbohydrate substrates, many GHs exhibit broad substrate specificity and activity, making their classification difficult. Therefore, different nomenclature and classification approaches are currently in use that combine specific features of amino acid sequence similarities, protein structure features, substrate or product specificity, reaction stereochemistry mechanism or mode of attack/action (exo and endo enzymes).<sup>75</sup>

The simplest classification of GHs is based on their substrate specificities and is used by the International Union of Biochemistry and Molecular Biology (IUBMB)<sup>†</sup> where GHs are given an Enzyme Commission (EC) number code EC 3.2.1. This classification approach does not reflect the structural characteristics of GHs and enzymes generally, which are now apparent.<sup>75</sup> As a result, in September 1998 Carbohydrate-Active Enzymes (CAZy) data base was created gathering GHs in families based on amino acid sequence and protein structures folding similarities.<sup>76,77</sup> Currently, 165 families have been identified and their number is continuously updated.<sup>††</sup> The key benefit of CAZy classification is insight into catalytic mechanism because active-site residues are essentially conserved within given GH family.<sup>76,78</sup>

GHs catalyse glycosidic bond hydrolysis through retaining or inverting mechanism depending on the change in the anomeric carbon stereochemistry during the reaction.

---

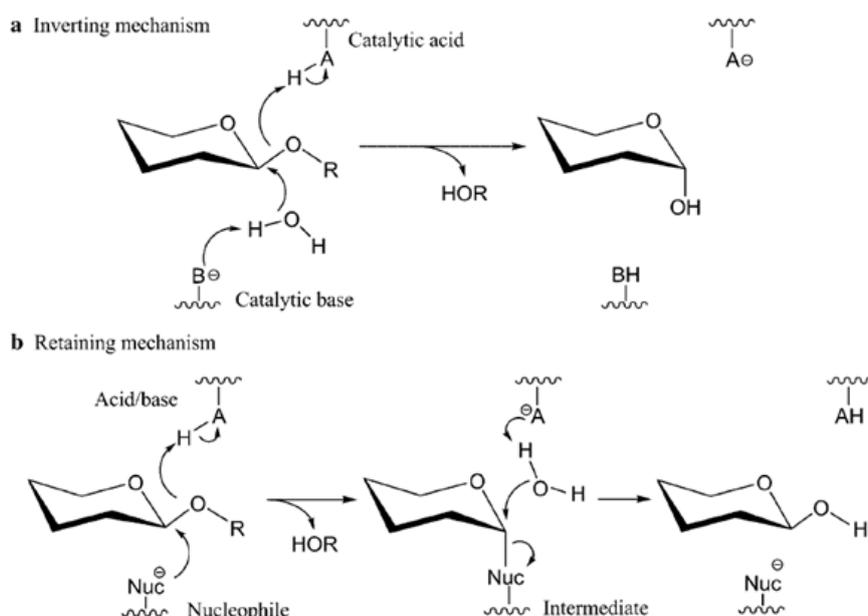
<sup>†</sup> <http://www.iubmb.org/>

<sup>††</sup> <http://www.cazy.org/> (accessed 22/07/2019)

Both mechanisms involve oxacarbenium ion-like transition states and a pair of carboxylic acids at the active site.<sup>79,80</sup>

In inverting GHs, the two active residues are a catalytic acid residue and a catalytic base residue which are approximately 10 Å apart. The reactions are carried out by a single-displacement mechanism (S<sub>N</sub>1). The catalytic acid residue donates a proton to the glycosyl oxygen while the catalytic base residue removes a proton from a water molecule, increasing its nucleophilicity and facilitating attack on the anomeric centre (**Figure I-11. A**). The greater distance between the catalytic residues is necessary to accommodate the nucleophilic water molecule between the catalytic general base residue and the anomeric carbon.<sup>81-83</sup>

In retaining GHs, the two active site residues are a general acid/base residue and a nucleophile which are approximately 5.5 Å apart and the reaction proceeds in a double-displacement mechanism (S<sub>N</sub>2). A catalytic general acid/base residue acts in two steps. First in glycosylation, as an acid, facilitating departure of the leaving group by donating a proton to the glycosyl oxygen atom while the nucleophile residue forms the covalent glycosyl-enzyme intermediate. In the second deglycosylation step, the deprotonated acid/base residue, acts as a general base to activate a water molecule which attacks the covalent glycosyl-enzyme intermediate with the two inversion steps resulting in retention of anomeric centre configuration (**Figure I-11. B**).<sup>81-83</sup>



**Figure I-19.** Schematic representation of **A**) inverting and **B**) retaining GH catalytic mechanism.

One of these catalytic mechanisms is conserved in most GHs families with some mechanistic exceptions in GH-23, GH-321, GH-83 and GH-97 family and one fundamentally different mechanism, discovered in recent years.<sup>84</sup>

The families also display similar active-site topologies, divided in three general groups, that include pocket or crater, cleft or groove and tunnel topology.<sup>85</sup> The shape of the active-site cleft reflects their mode of action which classifies them into two distinctive groups, *endo*-glycosidases and *exo*-glycosidases, whether the enzyme cleaves internal or external glycosidic bond of a polysaccharide.<sup>75</sup>

### 2.1. Glycofuranoside hydrolases

Among glycoside hydrolases, glycopyranosidases are extensively studied and only a few glycofuranosidases have been described including  $\beta$ -D-fructofuranosidase (EC 3.2.1.26),  $\alpha$ -L-arabinofuranosidase (EC 3.2.1.55),  $\beta$ -L-arabinofuranosidase (EC 3.2.1.185) and  $\beta$ -D-galactofuranosidase (EC 3.2.1.146).<sup>86</sup> Among listed, galactofuranosidases remained relatively unexplored and small number of studies regarding the isolation and characterization of this enzyme are available to date. Here, the review summarizes the data related to the research of galactofuranosidases, beginning from the first reference to galactofuranosidase activity to the only one cloned since then.

### 2.2. Galactofuranosidases

#### 2.2.1. Sources & activities – the early reports

The glycobiology of galactofuranose and enzymes involved in its metabolism are attractive targets for the development of antimicrobial agents since  $\beta$ -D-galactofuranosyl residues are constituents of pathogenic microorganisms, but are absent in mammal glycoconjugates.<sup>24</sup> There are three main enzymes involved in the biosynthesis and the metabolism of the *Galf*-containing molecules, UDP-galactopyranose mutase, galactofuranosyltransferase and galactofuranosidase. Although the metabolism of  $\beta$ -D-galactofuranosides has been extensively studied, mostly in the infectious microorganisms such as *Mycobacteria*, *Trypanosoma*, *Leishmania* and *Aspergillus*, there are only few reports related to these enzymes, especially *Galf*-ase.

Over the past forty years, *Galf*-ase has been identified as responsible for the degradation of the D-*Galf* containing glycoconjugates.<sup>24,55</sup> In 1977 a specific extracellular *exo*- $\beta$ -D-galactofuranosidase (*exo*- $\beta$ -D-*Galf*-ase) was the first one isolated and partially purified

from *Penicillium fellutanum* (ex type of *Penicillium charlesii*) culture filtrates. The enzyme catalysed hydrolysis of  $\beta$ -galactofuranosides only and was incapable of cleaving  $\alpha$ -L-arabinofuranosides, the pentosyl homologs. The enzyme was stable to freezing and thawing and obtained optimum activity between pH 3 and 4 at the temperature of 47 °C.<sup>87</sup>

Later, extracellular exo- $\beta$ -D-Galf-ases have been described and/or purified from culture medium of filamentous fungi such as *Helminthosporium sacchari*,<sup>88,89</sup> *Trichoderma harzianum*<sup>90</sup>, *Penicillium* and *Aspergillus* species<sup>91</sup>, among which the one from *P. fellutanum* have been the most studied,<sup>87,92-94</sup> *Aspergillus niger*<sup>95</sup> and protozoa *Trypanosoma cruzi*<sup>96</sup>. All the known  $\beta$ -Galf-ases were exclusively exo enzymes and mostly have been of fungal origin.

There are only two reported endo- $\beta$ -D-Galf-ases, isolated from fungus *Penicillium oxalicum*<sup>97</sup> and from bacteria *Bacillus sp.*<sup>98</sup> The first endo- $\beta$ -D-Galf-ase, purified in 1992 from supernatants of *P. oxalicum* autolysed cultures, hydrolysed specifically  $\beta$ -D-(1 $\rightarrow$ 5)-linked galactofuranose residues. This was the first endo- $\beta$ -D-Galf-ase that optimum pH, stability properties, substrate specificity and kinetic characteristics were determined.<sup>97</sup>

Only a few specific  $\beta$ -D-Galf-ases have been reported and mostly detected in fungal species as an extracellular enzyme, secreted into the medium in low quantities.

In the absence of gene, associated with these enzymes, being identified and expressed, its production was challenging. In order to induce higher enzyme secretion, the synthetic growth medium was supplemented with various monosaccharides, disaccharides or polysaccharides<sup>92</sup>, sometimes with those containing galactofuranoside residues<sup>95,98</sup> or the media from natural sources, like apple juice<sup>91</sup> was used. Production of Galf-ase in *A. niger* were induced to significant levels in the presence of Galf containing glycoconjugates (mycelial wall extracts)<sup>95</sup> and in *P. fellutanum* and *A. fumigatus* when the medium was depleted of glucose<sup>87,91,92,99</sup> or when glucose was replaced by galactose as the carbon source.<sup>100</sup>

Mostly, Galf-ases were purified from *P. fellutanum* culture filtrates and have become a model enzymes in characterization studies<sup>87,88</sup>, but were also commercially available from crude enzyme preparations,<sup>90,101</sup> or from *T. cruzi* cell lysates<sup>96</sup>.

To detect low levels of *Galf*-ase in culture media, to isolate it and possibly purify, an array of *Galf*-containing glycoconjugates, acting like potential substrates or inhibitors, has been synthesized and assayed. Some proved to be good inhibitors, as alkyl, aryl and heteroaryl 1-thio- $\beta$ -D-galactofuranosides.<sup>100,102</sup> An affinity chromatography system was also developed using two inhibitors, 4-aminophenyl thio- $\beta$ -D-galactofuranoside and D-galactono-1,4-lactone, as the immobilized and eluting ligands.<sup>93</sup>

These early studies presented pioneering work in *Galf*-ase research. In the most cases, properties were determined using partially purified or even crude enzyme preparations and one of the main problems was the lack of a simple and sensitive quantitative method for the detection of their catalytic activity as well as standardized substrate. The activity was usually determined by measuring released galactose during hydrolysis of galactofuranose-containing exopolysaccharide preparations of natural origin or methyl  $\beta$ -D-*Galf* by the galactose oxidase method. Already experimentally established use of nitrophenyl glycosides, widely used as substrates for estimating the activity, kinetics and specificity of glycosidases, was extended to *para*-nitrophenyl  $\beta$ -D-galactofuranose (*p*NP- $\beta$ -D-*Galf*)<sup>103,91</sup> which became standardized and commercially available<sup>95</sup> substrate for assaying galactofuranosidases in general.

A colorimetric assay, with *p*NP- $\beta$ -D-*Galf* as a substrate was extensively used, however, this substrate was not recognised by endo- $\beta$ -D-*Galf*-ase from *Bacillus sp.*<sup>98</sup> and exo- $\beta$ -D-*Galf*-ase from *T. cruzi*<sup>96</sup> and could be attributed to particular substrate specificity, relating to aglycone.

Diversity of organisms, experimental conditions and substrates were employed during these research studies making it very difficult to unanimously make general conclusion. However, they showed reliable evidence that suggested some common characteristics, shared by those reported *Galf*-ases, which can be outlined (**Table I-1.**). Substrate specificity, regarding glycon moiety from either natural or artificial substrates was exclusively towards  $\beta$ -D-*Galf*. The enzymes proved to be stable over a longer period of time, even as crude preparations, and showed optimal activity in acidic conditions, usual for the enzymes of invertebrates, mostly between pH 4-5, and the stability at the temperature up to 40 °C.

**Table I-3.** Comparative properties of  $\beta$ -D-galactofuranosidases.

Enzyme	Species	Substrate	pH	T (°C)	M (kDa)	$K_M$ (mM) <sup>†</sup>	Year	Reference
extracellular exo- $\beta$ -D-Galf-ase	<i>Penicillium charlesii</i> (fungus)	pPGM <sup>a</sup>	4	47	—	—	1977	87
$\beta$ -Galf-ase	<i>Helminthosporium sacchari</i> (fungus)	1- <i>O</i> -methyl- $\beta$ -Galf	4.2 & 5.2	38	—	—	1983	88
$\beta$ -D-Galf-ase	<i>H. sacchari</i> (fungus)	HS toxin <sup>b</sup>	4.6	37	—	—	1983	89
extracellular $\beta$ -D-Galf-ase	<i>Penicillium spp.</i> <i>Aspergillus spp.</i> (fungi)	pNP- $\beta$ -D-Galf	5	30	—	—	1989	91
endo- $\beta$ -Galf-ase	<i>Penicillium oxalicum</i> (fungus)	$\beta$ -(1→5)-galactofuran	5	37	77	—	1992	97
exo- $\beta$ -D-Galf-ase	<i>Trichoderma harzianum</i> (fungus)	EPS <sup>c</sup>	4 – 4.5	35 – 40	35	—	1992	90
endo- $\beta$ -Galf-ase	<i>Bacillus sp.</i> (bacteria)	—	6	37	67	—	1995	98
exo- $\beta$ -D-Galf-ase	<i>Penicillium fellutanum</i> (fungus)	pNP- $\beta$ -D-Galf	3 – 6	37	70	0.3	1999	93
extracellular $\beta$ -Galf-ase	<i>Aspergillus niger</i> (fungus)	pNP- $\beta$ -D-Galf	3 – 4	37	90	4	2001	95
exo- $\beta$ -D-Galf-ase	<i>P. fellutanum</i> (fungus)	1- <i>O</i> -methyl- $\beta$ -Galf	4 – 4.5	40	70	2.6	2001	94
exo- $\beta$ -D-Galf-ase	<i>Trypanosoma cruzi</i> (protozoa)	LPPG <sup>d</sup>	—	—	55	—	2003	96

<sup>†</sup> $K_M$  values determined only for pNP- $\beta$ -D-Galf and its derivatives as a substrate.

<sup>a</sup>pPGM – peptidophosphogalactomannan from *Penicillium fellutanum*

<sup>b</sup>HS toxin – host-selective toxin from *Helminthosporium sacchari*

<sup>c</sup>EPS – extracellular polysaccharides from *Penicillium digitatum*

<sup>d</sup>LPPG – lipopeptidophosphoglycan from *Trypanosoma cruzi*

### 2.2.2. Cloning & characterization of the first galactofuranosidase

Although the first *Galf*-ase was described back in 1977, and several exo- and endo-*Galf*-ases were purified from the culture supernatants and cell lysates of filamentous fungi, bacteria and protozoa, the genes encoding these enzymes have not been identified and expressed earlier, nor its amino acid sequence determined.

Recently, in 2015, based on the draft genome sequence analysis of soil, Gram-positive bacteria *Streptomyces sp.*, strain JHA19, an open reading frame that encoded *Galf* specific enzyme was identified. Based on the sequencing results, the genome size was 7.7 Mb and the entire *Galf*-ase open reading frame (ORF) fragment contained 2361 base pairs (bp), which encoded 786 amino acids.<sup>104</sup>

The *Galf*-ase gene fragment was cloned, expressed and purified as recombinant Nus and double 6xHis-tagged fusion protein. The enzyme was described as exo-type *Galf*-ase that hydrolysed galactomannan, naturally occurring *Galf*-containing oligosaccharide, extracted from *A. fumigatus* cell wall, as well as the artificial substrate *pNP*- $\beta$ -D-*Galf*. No activity was observed with other *pNP* furanosyl and pyranosyl glycoconjugates, *pNP*- $\alpha$ -L-*Araf* including. The optimum pH was found to be 5.5 and the enzyme was stable at the temperatures up to 40 °C with the  $K_M$  value of 4.4 mM (**Table I-2.**). This was the first report of identification and cloning of gene coding for *Galf*-specific *Galf*-ase enzyme that does not also exhibit arabinofuranosidase (*Araf*-ase) activity.<sup>105</sup>

Another *Galf*-ase gene was later also found in the genome of *Streptomyces sp.*, strain JHA26, coding for 869 amino acids. The *Galf*-ase specific enzyme was expressed, purified and characterized with the highest activity at pH 4.5 and temperature stability up to 45 °C and with  $K_M$  of 6.8 mM for *pNP*- $\beta$ -D-*Galf* as a substrate (**Table I-2.**).<sup>106,107</sup>

**Table I-4.** Comparative properties of recombinant  $\beta$ -D-galactofuranosidases.

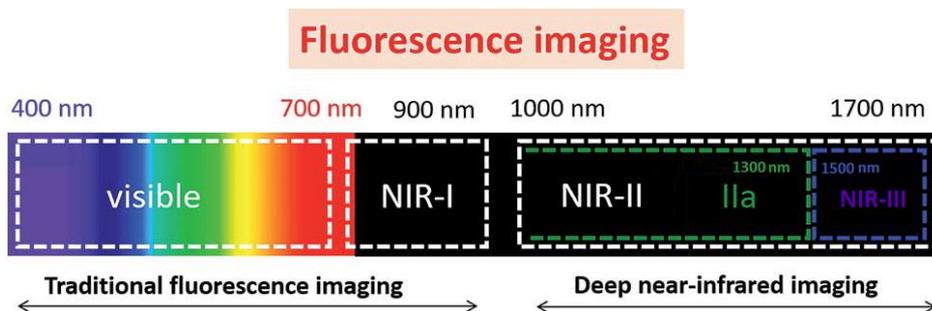
<i>Galf</i> -ase	Substrate	pH	T (°C)	$K_M$ (mM)	Year	Reference
<i>Streptomyces sp. (JHA19)</i>	<i>pNP</i> - $\beta$ -D- <i>Galf</i>	5.5	50	4.4	2015	105
<i>Streptomyces sp. (JHA26)</i>	<i>pNP</i> - $\beta$ -D- <i>Galf</i>	4.5	45	6.8	2017	106

### III. Lanthanide(III) complexes & nanoparticles – near-infrared emitting probes for biological imaging

#### 1. Optical imaging in the near-infrared domain

Biological imaging is essential to the advancement of various disciplines of life sciences and medicine since it can provide the detailed visualization of different biological entities and the monitoring of a wide range of physiological and pathological processes.<sup>108</sup> The field of biological imaging is very broad and includes techniques such as positron emission tomography (PET), single-photon emission computed tomography (SPECT), X-ray computed tomography (XCT), magnetic resonance imaging (MRI) and optical imaging, to name a few. The latter technique relies on the use of light for excitation and when emission is detected the imaging is referred to as luminescence or fluorescence optical imaging that can be performed in the visible or in the near-infrared (NIR) ranges of wavelengths.<sup>109</sup>

The near-infrared region of the electromagnetic spectrum, which by broad definition encompasses wavelength between 700 and 1700 nm, is highly advantageous for optical imaging applications.<sup>110</sup> Non-invasive luminescence optical imaging typically operates in so-called NIR biological diagnostic windows, which is divided into the first NIR window (NIR-I) region from 650/700 to 900 nm and the second NIR window (NIR-II) region from 1000 to 1700 nm.<sup>111-113</sup> Recently, a novel definition of NIR imaging windows has been introduced that suggested the division into three regions, NIR-I from 650 to 1000 nm, NIR-II from 1000 to 1300 nm and NIR-III from 1500 to 1800 nm, respectively (**Figure I-12**).<sup>114</sup>



**Figure I-20.** Ranges of wavelengths usually used for fluorescence optical imaging.<sup>113</sup>

There are several advantages for the use of NIR light for fluorescence optical imaging in comparison to the visible one. Biological materials have low autofluorescence in the NIR window, which allows facile discrimination between the desired signal of the NIR-

emitting reporter and the background, leading to an enhanced signal-to-noise ratio and improved detection sensitivity. Additionally, the NIR light can penetrate deeper in the tissue (>1 cm depth) than visible light and due to the reduced light scattering and the lower absorption an increased optical imaging resolution can be achieved. Finally, NIR light interacts less with biological material, thus minimizing the risk of disturbing or damaging the biological systems that are being observed.<sup>108,115,116</sup>

Numerous advantages of optical bioimaging in the NIR window have triggered the development of the suitable probes. A wide range of NIR-emitting probes and materials of different origins has been suggested including small organic molecules, fluorescent proteins, carbon nanotubes and inorganic nanoparticles. In addition, since several lanthanide(III) cations emit in the NIR range and provide complementary advantages with respect to other materials, NIR-emitting lanthanide(III)-based molecular probes have also been created.<sup>115,116</sup>

### **2. Lanthanide(III) coordination compounds – brief overview**

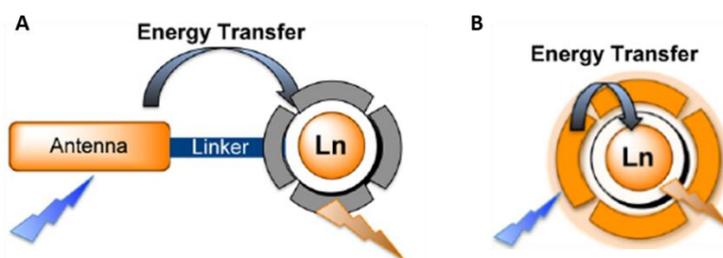
The study and development of lanthanide chemistry dates back to 18th century. In the last few decades, lanthanide-based compounds have found numerous applications in military, chemical industry, agriculture and biomedicine.<sup>117</sup> In particular, special attention has been drawn to the probes containing lanthanide(III) ions ( $\text{Ln}^{3+}$ ), especially NIR-emitting ones such as  $\text{Yb}^{3+}$ ,  $\text{Nd}^{3+}$ ,  $\text{Sm}^{3+}$ ,  $\text{Dy}^{3+}$ ,  $\text{Ho}^{3+}$ ,  $\text{Pr}^{3+}$ ,  $\text{Tm}^{3+}$  and  $\text{Er}^{3+}$ . The lanthanides belong to the f-block of the periodic table series that consists of 15 elements, from lanthanum (La) to lutetium (Lu), with atomic numbers ranging from 57 to 71, respectively. They have general electronic configuration  $[\text{Xe}]4f^n5d^16s^2$  ( $n=0$  (La) to 14 (Lu)) that turns into  $[\text{Xe}]4f^n$  for  $\text{Ln}^{3+}$  as the most common oxidation state. All of the elements in the group have a close chemical resemblance and due to intraconfigurational 4f-4f transitions, all  $\text{Ln}^{3+}$  ions except  $\text{La}^{3+}$  and  $\text{Lu}^{3+}$  are luminescent and are exhibiting emission in the UV, visible and/or NIR ranges.<sup>118-120</sup>

This specific electronic configuration of  $\text{Ln}^{3+}$  is responsible for the unique magnetic and optical properties which are particularly important for bioanalytical and imaging applications. More specifically, 4f-4f transitions are sharp emission bands compared to organic fluorophores and semiconductor nanocrystals, and their emission wavelengths are insignificantly affected by the environment (varied pH, temperature and biological

environment), allowing them to be easily discriminated and used in a broad range of conditions, including biological environments.<sup>115,116,121</sup>

However, free  $\text{Ln}^{3+}$  have very low molar absorption coefficients (in the range  $0.1\text{--}10\text{ M}^{-1}\text{ cm}^{-1}$ ) due to the forbidden nature of most of the 4f-4f transitions. Therefore, they must be indirectly sensitized through a light-harvesting antenna to circumvent this limitation. To achieve this goal  $\text{Ln}^{3+}$  are usually surrounded by appropriate highly absorbing organic chromophores which can efficiently absorb the light and transfer the resulting energy to the accepting electronic levels of the  $\text{Ln}^{3+}$ . For biological imaging,  $\text{Ln}^{3+}$  are often chelated with multidentate ligands that can form highly thermodynamically or kinetically stable complexes and thus limit the potential toxicity induced by free  $\text{Ln}^{3+}$ . Organic ligands should also protect  $\text{Ln}^{3+}$  from interactions with solvent molecules since overtones of high-energy vibrations (O-H, N-H) can quench  $\text{Ln}^{3+}$  luminescence, particularly in the NIR range.

Light-harvesting antennae able to sensitize a characteristic  $\text{Ln}^{3+}$  emission can be attached through a linker to a chelating unit (**Figure I-13. A**) or already possess groups and directly coordinate the  $\text{Ln}^{3+}$  (**Figure I-13. A**). Polyaminocarboxylates (diethylenetriaminepentaacetic (DTPA) or ethylenediaminetetraacetic (EDTA) acids) and macrocycles (1,4,7,10-tetraazacyclododecane-1,4,7,10-tetraacetic acid (DOTA)), often serve as chelating units<sup>121-124</sup>



**Figure I-21.** Schematic representation of the general architecture of the luminescent  $\text{Ln}^{3+}$  complexes and the antenna effect. **A**) Pendant antenna. **B**) Chelating antenna.

## IV. Objectives of the thesis

The main objective of this thesis is the ***creation of galactofuranose-specific luminescent probe for near-infrared optical imaging***. In order to reach this objective, experimental work was strategically divided into four phases (**Figure I-14.**):

- i) To fully characterize galactofuranosidase, naturally occurring galactofuranose-specific hydrolytic enzyme, to date the only one with gene identified and available as cloned construct.*
- ii) To bioengineer neolectins from wild-type galactofuranosidase and to characterize their thioglycoligase potential.*
- iii) To create a probe that will be able to exhibit NIR lanthanide luminescence and carry functional groups to allow attachment of generated neolectins.*
- iv) Final phase will lead to organisation of protocol to test created galactofuranose-specific luminescent probe for near-infrared optical imaging of appropriate target microorganism and to deliver a proof-of-concept validation.*

The thesis is presented in seven chapters. The main structure is organized around first four chapters, with the first one giving bibliographical overview and theoretical context of the subject thesis is focused upon. Experimental work and results are later elaborated and discussed through chapters two, three and four. Each of this chapters starts with a brief introduction, relevant to presented experimental work, and ends with central results summarized in corresponding conclusion. Chapter five brings final conclusion, followed by proposition of further short-term and long-term work perspectives. Under chapters six and seven, detailed experimental procedures are described as well as all the additional data not included in the precedent chapters.

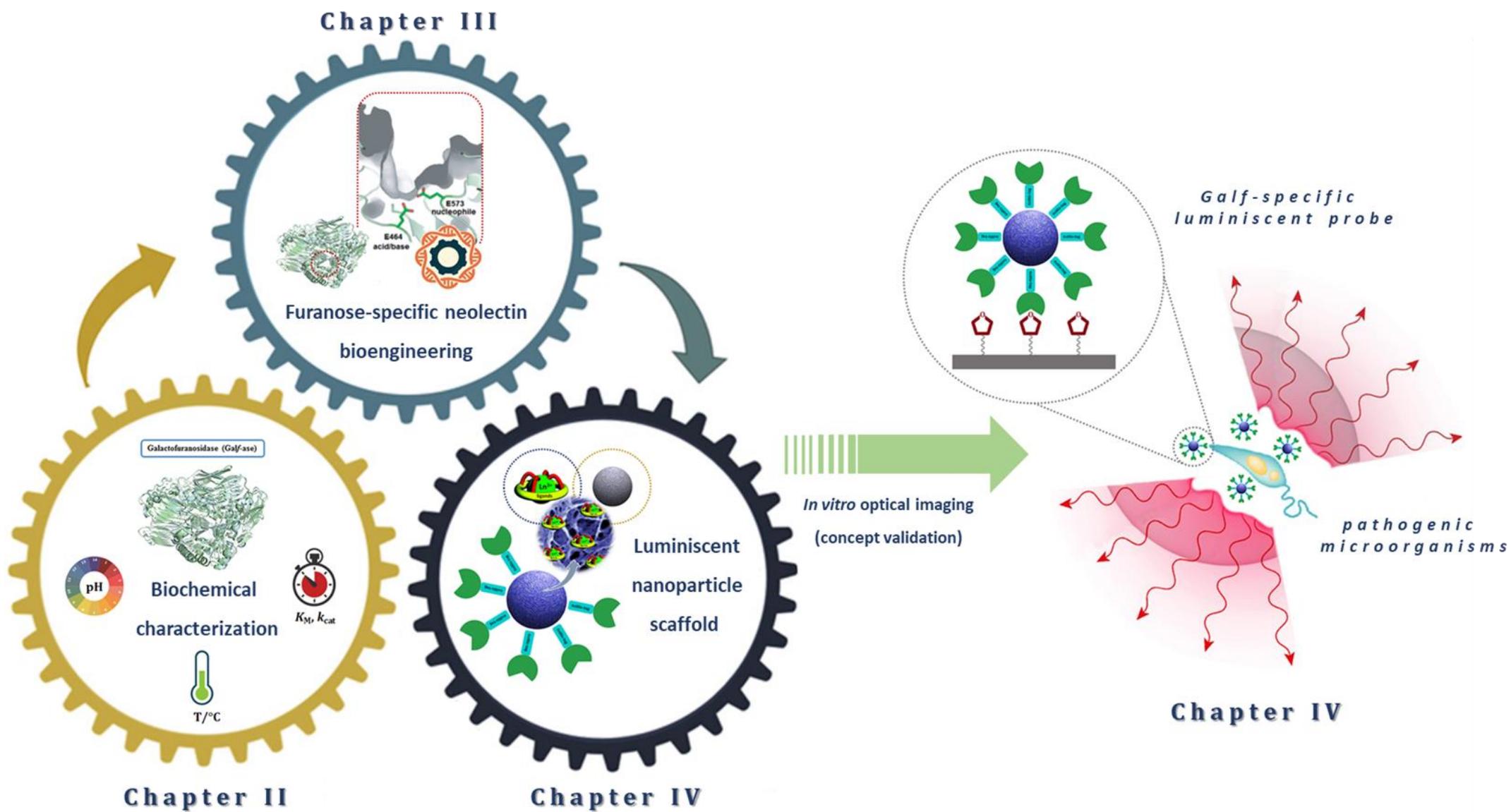
Detailed thesis objectives can be perceived within work provided in chapters two, three and four.

Chapter II, brings detailed characterization of galactofuranosidase, in order to assess its potential to be used in several other biotechnological approaches. Detailed work, including application of gene (sub)cloning techniques and recombinant enzyme production and characterization will be taken. First part is to analyse original plasmid construct, bearing Gal $\beta$ -ase insert gene, and to choose appropriate (sub)cloning technique approach and recipient expression vector. Second part will include overexpression and

purification of recombinant enzyme which hydrolytic activity will be in detail biochemically characterized to determine substrate specificity, optimum pH and temperature conditions, storage stability as well as kinetic parameters and potential inhibitors. Also screening for adequate crystallization conditions to yield crystals for an X-ray diffraction investigation will be carried out.

Chapter III, is focused on *GalF*-ase bioengineering to create neolectins and to explore furanothioglycoligase potential. Mutagenesis points, targeting catalytic amino acid residues, will be selected and created through site-directed mutagenesis. The engineered *GalF*-ase mutants will be overexpressed, purified and their kinetic activity characterized. Hydrolytically inactive variants or variants with attenuated catalytic activity will be tested for furanothioglycoside synthesis. For this purpose, several thiophenol and 2-mercaptothiazoline derivatives will be screened as aglycone acceptors to appropriate glycosyl donor.

Chapter IV, presents final framework of this thesis which brings together neolectin potential of *GalF*-ase mutants, to act as specificity tools, and lanthanide luminescence for near-infrared imaging. Within this chapter several aspects will be considered. Firstly, the choice of near-infrared emitting lanthanide-based system that will be compatible for enzyme bioconjugation. Second part describes bioconjugation techniques and determinations of enzyme activity after the coupling with near-infrared emitting probe. Finally, enzyme labelled luminescent probe will be tested, in a manner of a concept validation, for near-infrared optical imaging.



**Figure I-22.** A graphical abstract of the main thesis objectives.

## IV. Objectifs de la thèse

L'objectif principal de cette thèse est **la création d'une sonde luminescente spécifique au galactofuranose pour l'imagerie optique proche infrarouge**. Afin d'atteindre cet objectif, les travaux expérimentaux ont été stratégiquement divisés en quatre phases (**Figure I-14.**):

- i) Caractériser complètement la galactofuranosidase, enzyme hydrolytique naturelle spécifique du galactofuranose, à ce jour la seule dont nous possédons un gène identifié et disponible comme construction clonée.*
- ii) Concevoir des néolectines issues de galactofuranosidase de type sauvage et caractériser leur potentiel en thioglycoligase.*
- iii) Créer une sonde capable de montrer la luminescence des lanthanides NIR et de porter des groupes fonctionnels pour permettre la fixation des néoélectines générées.*
- iv) Organiser un protocole pour tester la sonde luminescente spécifique au galactofuranose créée pour l'imagerie optique dans le proche infrarouge du micro-organisme cible approprié et pour fournir une preuve de concept.*

La thèse est présentée en sept chapitres. La structure principale s'articule autour de quatre premiers chapitres, le premier donnant un aperçu bibliographique et le contexte théorique de la thèse. Les travaux expérimentaux et les résultats sont ensuite développés et discutés dans les chapitres deux, trois et quatre. Chacun de ces chapitres commence par une brève introduction, pertinente aux travaux expérimentaux présentés, et se termine par des résultats centraux résumés dans la conclusion correspondante. Le chapitre cinq apporte la conclusion finale, suivie d'une proposition de nouvelles perspectives de travail à court et à long terme. Les chapitres six et sept décrivent les procédures expérimentales détaillées ainsi que toutes les données supplémentaires qui ne figurent pas dans les chapitres précédents.

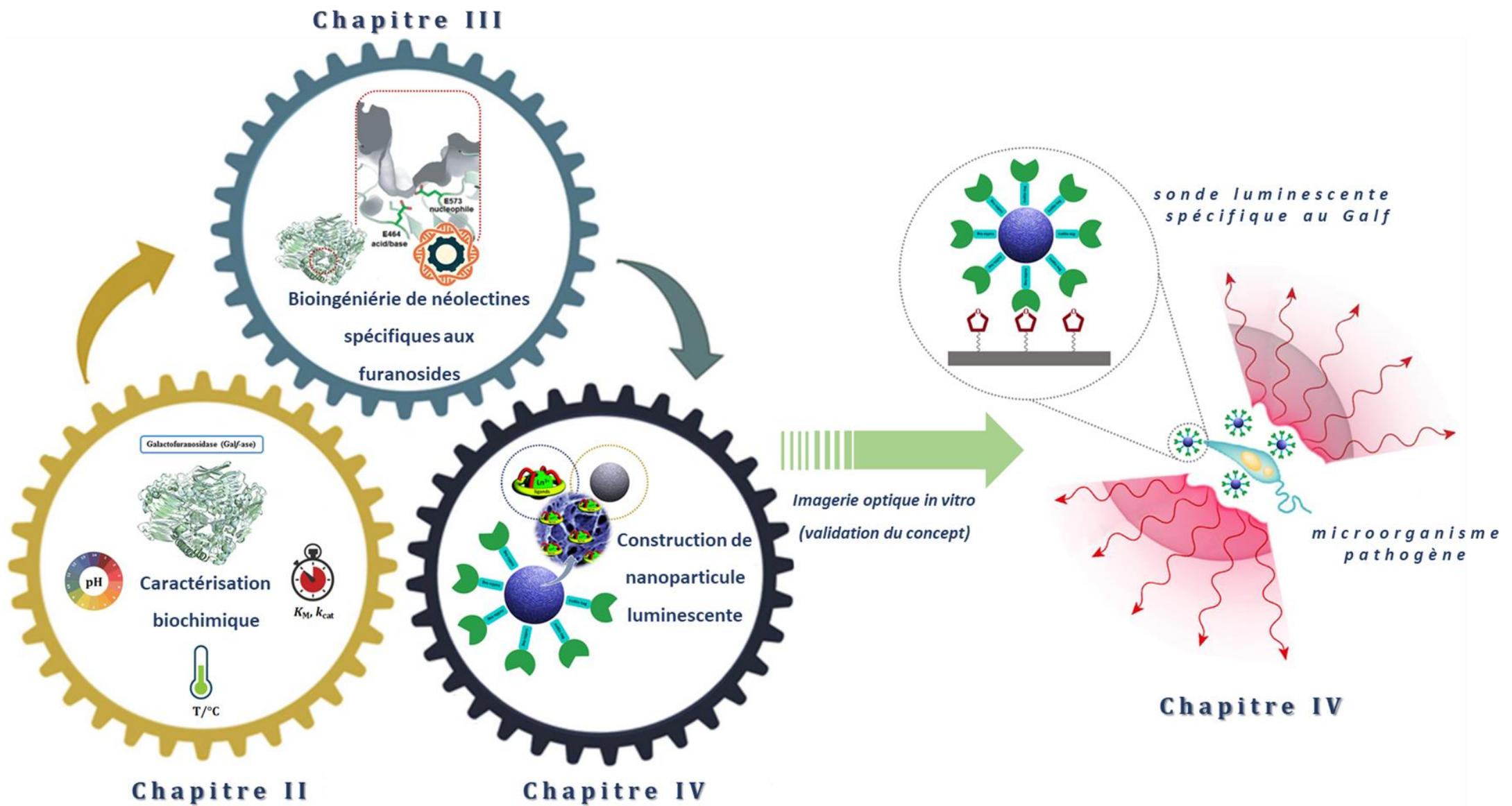
Les objectifs détaillés de la thèse peuvent être perçus dans le cadre des travaux fournis dans les chapitres deux, trois et quatre.

Le chapitre II présente une caractérisation détaillée de la galactofuranosidase, afin d'évaluer son potentiel à être utilisée dans plusieurs autres approches biotechnologiques. Des travaux détaillés, y compris l'application de techniques de (sous-)clonage de gènes et la production et la caractérisation d'enzymes recombinantes, seront effectués. La

première partie consiste à analyser la construction plasmidique originale, porteuse du gène d'insertion de *Galf*-ase, et à choisir l'approche (sous-)technique de clonage et le vecteur d'expression approprié. La deuxième partie comprendra la surexpression et la purification de l'enzyme recombinante, dont l'activité hydrolytique sera caractérisée biochimiquement en détail pour déterminer la spécificité du substrat, les conditions optimales de pH et de température, la stabilité au stockage ainsi que les paramètres cinétiques et les inhibiteurs potentiels. On procédera également à un criblage pour déterminer des conditions de cristallisation adéquates pour produire des cristaux en vue d'une étude de diffraction des rayons X.

Le chapitre III est axé sur la bioingénierie de la *Galf*-ase pour créer des néolectines et explorer le potentiel de la furanothioglycoligase. Des points de mutagenèse, ciblant les résidus d'acides aminés catalytiques, seront sélectionnés et créés par mutagenèse dirigée. Les mutants modifiés de *Galf*-ase seront surexprimés, purifiés et leur activité cinétique sera caractérisée. Des variantes inactives du point de vue hydrolytique ou des variantes à activité catalytique atténuée seront testées pour la synthèse des furanothioglycosides. À cette fin, plusieurs dérivés du thiophénol et de la 2-mercaptothiazoline feront l'objet d'un dépistage comme accepteurs de l'aglycone chez les donneurs de glycosyle appropriés.

Le chapitre IV, présente le cadre final de cette thèse qui réunit le potentiel néoélectinique des mutants de *Galf*-ase, pour agir comme outils de spécificité, et la luminescence des lanthanides pour l'imagerie proche infrarouge. Dans ce chapitre, plusieurs aspects seront examinés. Premièrement, le choix d'un système à base de lanthanide émettant dans le proche infrarouge qui sera compatible pour la bioconjugaison enzymatique. La deuxième partie décrit les techniques de bioconjugaison et la détermination de l'activité enzymatique après le couplage avec la sonde émettrice dans le proche infrarouge. Enfin, une sonde luminescente marquée par une enzyme sera testée, dans un but de validation de concept, pour l'imagerie optique dans le proche infrarouge.



**Figure I-14.** Un résumé graphique des principaux objectifs de la thèse.

## Bibliography

1. Ulusu, N. N. Curious Cases of the Enzymes. *J. Med. Biochem.* **34**, 271–281 (2015).
2. Cherry, J. R. & Fidantsef, A. L. Directed evolution of industrial enzymes: an update. *Curr. Opin. Biotechnol.* **14**, 438–443 (2003).
3. Gloster, T. M. Development of inhibitors as research tools for carbohydrate-processing enzymes. *Biochem. Soc. Trans.* **40**, 913–928 (2012).
4. Cipolla, L. *et al.* Discovery and design of carbohydrate-based therapeutics. *Expert Opin. Drug Discov.* **5**, 721–737 (2010).
5. Thisbe K. Lindhorst., A. Essentials of Carbohydrate Chemistry and Biochemistry. *Angew. Chem. Int. Ed.* **47**, 1360–1361 (2008).
6. *Essentials of Glycobiology.* (Cold Spring Harbor Laboratory Press, 2009).
7. Campo, V. L., Aragão-Leoneti, V., Teixeira, M. B. M. & Carvalho, I. Carbohydrates and Glycoproteins: Cellular Recognition and Drug Design. in (2010).
8. Taylor, M. E. & Drickamer, K. *Introduction to Glycobiology.* (Oxford University Press, 2011).
9. Comprehensive Glycoscience | ScienceDirect.
10. Beerens, K., Desmet, T. & Soetaert, W. Enzymes for the biocatalytic production of rare sugars. *J. Ind. Microbiol. Biotechnol.* **39**, 823–834 (2012).
11. Alberts, B. *Molecular biology of the cell.* (Garland Science, Taylor and Francis Group, 2015).
12. Lichtenthaler, F. W. Carbohydrates: Occurrence, Structures and Chemistry. in *Ullmann's Encyclopedia of Industrial Chemistry* (ed. Wiley-VCH Verlag GmbH & Co. KGaA) a05\_079.pub2 (Wiley-VCH Verlag GmbH & Co. KGaA, 2010).
13. Imamura A. & Lowary T. Chemical Synthesis of Furanose Glycosides. *Trends Glycosci. Glycotechnol.* **23**, 134–152 (2011).
14. Lafite, P. & Daniellou, R. Rare and unusual glycosylation of peptides and proteins. *Nat. Prod. Rep.* **29**, 729 (2012).
15. Poulin, M. B. & Lowary, T. L. Methods to Study the Biosynthesis of Bacterial Furanosides. in *Methods in Enzymology* vol. 478 389–411 (Elsevier, 2010).
16. Van Vranken, D. *Introduction to Bioorganic Chemistry and Chemical Biology.* (Garland Science, 2012).

17. Marino, C. & de Lederkremer, R. M. Galactose configurations in Nature with emphasis on the biosynthesis of Galactofuranose in glycans. 29.
18. P. Collins and R. Ferrier. Rantwijk. Monosaccharides. Their Chemistry and Their Roles in Natural Products. *Recl. Trav. Chim. Pays-Bas* **115**, 420–420 (1996).
19. Tefsen, B., Ram, A. F., van Die, I. & Routier, F. H. Galactofuranose in eukaryotes: aspects of biosynthesis and functional impact. *Glycobiology* **22**, 456–469 (2012).
20. Haworth, W. N. & Porter, C. R. Isolation of crystalline  $\alpha$ - and  $\beta$ -ethylglucofuranosides ( $\gamma$ -ethylglucosides) and other crystalline derivatives of glucofuranose. *J. Chem. Soc. Resumed* 2796–2806 (1929) doi:10.1039/JR9290002796.
21. Clutterbuck, P. W., Haworth, W. N., Raistrick, H., Smith, G. & Stacey, M. Studies in the biochemistry of micro-organisms. *Biochem. J.* **28**, 94–110 (1934).
22. Haworth, W. N., Raistrick, H. & Stacey, M. Polysaccharides synthesised by micro-organisms. *Biochem. J.* **31**, 640–644 (1937).
23. Green, J. W. & Pacsu, E. Glycofuranosides and Thioglycofuranosides. III. New Crystalline Furanosides of d-Galactose and l-Arabinose. *J. Am. Chem. Soc.* **60**, 2056–2057 (1938).
24. Marino, C., Gallo-Rodriguez, C. & de Lederkremer, R. M. Galactofuranosyl-containing Glycans: Occurrence, Synthesis and Biochemistry. 62.
25. Peltier, P., Euzen, R., Daniellou, R., Nugier-Chauvin, C. & Ferrières, V. Recent knowledge and innovations related to hexofuranosides: structure, synthesis and applications. *Carbohydr. Res.* **343**, 1897–1923 (2008).
26. Lowary, T. L. Twenty Years of Mycobacterial Glycans: Furanosides and Beyond. *Acc. Chem. Res.* **49**, 1379–1388 (2016).
27. Thanna, S. & Sucheck, S. J. Targeting the trehalose utilization pathways of *Mycobacterium tuberculosis*. *MedChemComm* **7**, 69–85 (2016).
28. Jankute, M., Cox, J. A. G., Harrison, J. & Besra, G. S. Assembly of the Mycobacterial Cell Wall. *Annu. Rev. Microbiol.* **69**, 405–423 (2015).
29. Brown, L., Wolf, J. M., Prados-Rosales, R. & Casadevall, A. Through the wall: extracellular vesicles in Gram-positive bacteria, mycobacteria and fungi. *Nat. Rev. Microbiol.* **13**, 620–630 (2015).
30. Abrahams, K. A. & Besra, G. S. Mycobacterial cell wall biosynthesis: a multifaceted antibiotic target. *Parasitology* **145**, 116–133 (2018).

31. Nagaoka, M. *et al.* Structure of a galactan from cell walls of *Bifidobacterium catenulatum* YIT4016. *Carbohydr. Res.* **281**, 285–291 (1996).
32. Faber, E. J., van den Haak, M. J., Kamerling, J. P. & Vliegenthart, J. F. G. Structure of the exopolysaccharide produced by *Streptococcus thermophilus* S3. *Carbohydr. Res.* **331**, 173–182 (2001).
33. Yamazaki, K., Suzuki, M., Inukai, K., Kuga, H. & Korenaga, H. Structural Study on a Sulfated Polysaccharide-peptidoglycan Complex Produced by *Arthrobacter* sp. *Biosci. Biotechnol. Biochem.* **62**, 2138–2144 (1998).
34. Gow, N. A. R., Latge, J.-P. & Munro, C. A. The Fungal Cell Wall: Structure, Biosynthesis, and Function. *Microbiol. Spectr.* **5**, (2017).
35. Latge, J.-P. Galactofuranose containing molecules in *Aspergillus fumigatus*. *Med. Mycol.* **47**, S104–S109 (2009).
36. Latge, J.-P. *Aspergillus fumigatus* and Aspergillosis. *CLIN MICROBIOL REV* **12**, 41 (1999).
37. Oka, T. Biosynthesis of galactomannans found in filamentous fungi belonging to Pezizomycotina. *Biosci. Biotechnol. Biochem.* **82**, 183–191 (2018).
38. Matveev, A. L. *et al.* Novel mouse monoclonal antibodies specifically recognize *Aspergillus fumigatus* galactomannan. *PLOS ONE* **13**, e0193938 (2018).
39. Jansson, P.-E. & Lindberg, B. Structural studies of varianose. *Carbohydr. Res.* **82**, 97–102 (1980).
40. Lederkremer, R. M. de, Casal, O. L., Alves, M. J. M. & Colli, W. Evidence for the presence of D-galactofuranose in the lipopeptidophosphoglycan from *Trypanosoma cruzi*. *FEBS Lett.* **116**, 25–29 (1980).
41. de Lederkremer, R. M. & Colli, W. Galactofuranose-containing glycoconjugates in trypanosomatids. *Glycobiology* **5**, 547–552 (1995).
42. Cabezas, Y. *et al.* Leishmania cell wall as a potent target for antiparasitic drugs. A focus on the glycoconjugates. *Org. Biomol. Chem.* **13**, 8393–8404 (2015).
43. Oppenheimer, M., Valenciano, A. L. & Sobrado, P. Biosynthesis of Galactofuranose in Kinetoplastids: Novel Therapeutic Targets for Treating Leishmaniasis and Chagas' Disease. *Enzyme Res.* **2011**, 1–13 (2011).
44. O'Neill, M. A. & Roberts, K. Methylation analysis of cell wall glycoproteins and glycopeptides from *Chlamydomonas reinhardtii*. *Phytochemistry* **20**, 25–28 (1981).

45. Igarashi, T., Satake, M. & Yasumoto, T. Structures and Partial Stereochemical Assignments for Prymnesin-1 and Prymnesin-2: Potent Hemolytic and Ichthyotoxic Glycosides Isolated from the Red Tide Alga *Prymnesium parvum*. *J. Am. Chem. Soc.* **121**, 8499–8511 (1999).
46. Pettit, G. R. *et al.* Antineoplastic agents. Part 395.1 Isolation and structure of agelagalastatin from the Papua New Guinea marine sponge *Agelas* sp. *Chem. Commun.* 915–916 (1999)
47. Novelli, J. F. *et al.* Characterization of the *Caenorhabditis elegans* UDP-galactopyranose mutase homolog glf-1 reveals an essential role for galactofuranose metabolism in nematode surface coat synthesis. *Dev. Biol.* **335**, 340–355 (2009).
48. Bishop, J. R. & Gagneux, P. Evolution of carbohydrate antigens—microbial forces shaping host glycomes? *Glycobiology* **17**, 23R-34R (2007).
49. Heesemann, L. *et al.* Studies on galactofuranose-containing glycostructures of the pathogenic mold *Aspergillus fumigatus*. *Int. J. Med. Microbiol.* **301**, 523–530 (2011).
50. Stynen, D. *et al.* Rat monoclonal antibodies against *Aspergillus* galactomannan. *Infect. Immun.* **60**, 2237–2245 (1992).
51. Verdaguer, V., Walsh, T. J., Hope, W. & Cortez, K. J. Galactomannan antigen detection in the diagnosis of invasive aspergillosis. *Expert Rev. Mol. Diagn.* **7**, 21–32 (2007).
52. Marino, C., Rinflerch, A. & de Lederkremer, R. M. Galactofuranose antigens, a target for diagnosis of fungal infections in humans. *Future Sci. OA* **3**, FSO199 (2017).
53. Rolle, A.-M. *et al.* ImmunoPET/MR imaging allows specific detection of *Aspergillus fumigatus* lung infection in vivo. *Proc. Natl. Acad. Sci.* **113**, E1026–E1033 (2016).
54. Davies, G. *et al.* Towards Translational ImmunoPET/MR Imaging of Invasive Pulmonary Aspergillosis: The Humanised Monoclonal Antibody JF5 Detects *Aspergillus* Lung Infections *In Vivo*. *Theranostics* **7**, 3398–3414 (2017).
55. Eppe, G., Bkassiny, S. E. & Vincent, S. P. Chapter 9:Galactofuranose Biosynthesis: Discovery, Mechanisms and Therapeutic Relevance. in *Carbohydrates in Drug Design and Discovery* 209–241 (2015).
56. Houseknecht, J. B. & Lowary, T. L. Chemistry and biology of arabinofuranosyl- and galactofuranosyl-containing polysaccharides. *Curr. Opin. Chem. Biol.* **5**, 677–682 (2001).

57. Nassau, P. M. *et al.* Galactofuranose biosynthesis in *Escherichia coli* K-12: identification and cloning of UDP-galactopyranose mutase. *J. Bacteriol.* **178**, 1047–1052 (1996).
58. Kizjakina, K., J. Tanner, J. & Sobrado, P. Targeting UDP-Galactopyranose Mutases from Eukaryotic Human Pathogens. *Curr. Pharm. Des.* **19**, 2561–2573 (2013).
59. Tanner, J. J., Boechi, L., Andrew McCammon, J. & Sobrado, P. Structure, mechanism, and dynamics of UDP-galactopyranose mutase. *Arch. Biochem. Biophys.* **544**, 128–141 (2014).
60. Poulin, M. B., Zhou, R. & Lowary, T. L. Synthetic UDP-galactofuranose analogs reveal critical enzyme–substrate interactions in GlfT2-catalyzed mycobacterial galactan assembly. *Org. Biomol. Chem.* **10**, 4074 (2012).
61. Wheatley, R. W., Zheng, R. B., Richards, M. R., Lowary, T. L. & Ng, K. K. S. Tetrameric Structure of the GlfT2 Galactofuranosyltransferase Reveals a Scaffold for the Assembly of Mycobacterial Arabinogalactan. *J. Biol. Chem.* **287**, 28132–28143 (2012).
62. Wing, C., Errey, J. C., Mukhopadhyay, B., Blanchard, J. S. & Field, R. A. Expression and initial characterization of WbbI, a putative D -Galf:α- D -Glc β-1,6-galactofuranosyltransferase from *Escherichia coli* K-12. *Org. Biomol. Chem.* **4**, 3945–3950 (2006).
63. Guan, S., Clarke, A. J. & Whitfield, C. Functional Analysis of the Galactosyltransferases Required for Biosynthesis of d-Galactan I, a Component of the Lipopolysaccharide O1 Antigen of *Klebsiella pneumoniae*. *J. Bacteriol.* **183**, 3318–3327 (2001).
64. Ati, J. *et al.* The LPG1x family from *Leishmania major* is constituted of rare eukaryotic galactofuranosyltransferases with unprecedented catalytic properties. *Sci. Rep.* **8**, 17566 (2018).
65. Komachi, Y. *et al.* *gfsA* encodes a novel galactofuranosyltransferase involved in biosynthesis of galactofuranose antigen of O-glycan in *Aspergillus nidulans* and *Aspergillus fumigatus*. *Mol. Microbiol.* **90**, 1054–1073 (2013).
66. Katafuchi, Y. *et al.* *GfsA* is a β1,5-galactofuranosyltransferase involved in the biosynthesis of the galactofuran side chain of fungal-type galactomannan in *Aspergillus fumigatus*. *Glycobiology* **27**, 568–581 (2017).

67. Mariño, K., Marino, C., Lima, C., Baldoni, L. & Lederkremer, R. M. de. The First Chemical Synthesis of UDP[6-3H]- $\alpha$ -D-galactofuranose. *Eur. J. Org. Chem.* **2005**, 2958–2964 (2005).
68. Cantarel, B. L. *et al.* The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucleic Acids Res.* **37**, D233–D238 (2009).
69. The CAZypedia Consortium. Ten years of CAZypedia: a living encyclopedia of carbohydrate-active enzymes. *Glycobiology* **28**, 3–8 (2018).
70. Guillotin, L., Lafite, P. & Daniellou, R. Chapter 10. Enzymatic thioglycosylation: current knowledge and challenges. in *Carbohydrate Chemistry* (eds. Pilar Rauter, A., Lindhorst, T. & Queneau, Y.) vol. 40 178–194 (Royal Society of Chemistry, 2014).
71. Naumoff, D. G. Hierarchical classification of glycoside hydrolases. *Biochem. Mosc.* **76**, 622–635 (2011).
72. Wolfenden, R., Lu, X. & Young, G. Spontaneous Hydrolysis of Glycosides. *J. Am. Chem. Soc.* **120**, 6814–6815 (1998).
73. Zechel, D. L. & Withers, S. G. Glycosidase Mechanisms: Anatomy of a Finely Tuned Catalyst. *Acc. Chem. Res.* **33**, 11–18 (2000).
74. Davies, G. J., Gloster, T. M. & Henrissat, B. Recent structural insights into the expanding world of carbohydrate-active enzymes. *Curr. Opin. Struct. Biol.* **15**, 637–645 (2005).
75. Henrissat, B. & Davies, G. Structural and sequence-based classification of glycoside hydrolases. *Curr. Opin. Struct. Biol.* **7**, 637–644 (1997).
76. Henrissat, B. A classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem. J.* **280**, 309–316 (1991).
77. Davies, G. J. & Sinnott, M. L. Sorting the diverse: the sequence-based classifications of carbohydrate-active enzymes. in (2008). doi:10.1042/bj20080382.
78. Bourne, Y. & Henrissat, B. Glycoside hydrolases and glycosyltransferases: families and functional modules. *Curr. Opin. Struct. Biol.* **11**, 593–600 (2001).
79. Chakladar, S. *et al.* A mechanism-based inactivator of glycoside hydrolases involving formation of a transient non-classical carbocation. *Nat. Commun.* **5**, 5590 (2014).
80. Cerqueira, N., Brs, N., Joo, M. & Alexandrino, P. Glycosidases – A Mechanistic Overview. in *Carbohydrates - Comprehensive Studies on Glycobiology and Glycotechnology* (ed. Chang, C.-F.) (InTech, 2012). doi:10.5772/52019.

81. Koshland, D. E. Stereochemistry and the Mechanism of Enzymatic Reactions. *Biol. Rev.* **28**, 416–436 (1953).
82. Rye, C. S. & Withers, S. G. Glycosidase mechanisms. *Curr. Opin. Chem. Biol.* **4**, 573–580 (2000).
83. Vocadlo, D. J. & Davies, G. J. Mechanistic insights into glycosidase chemistry. *Curr. Opin. Chem. Biol.* **12**, 539–555 (2008).
84. Vuong, T. V. & Wilson, D. B. Glycoside hydrolases: Catalytic base/nucleophile diversity. *Biotechnol. Bioeng.* **107**, 195–205 (2010).
85. Davies, G. & Henrissat, B. Structures and mechanisms of glycosyl hydrolases. *Structure* **3**, 853–859 (1995).
86. Chlubnova, I. *et al.* Specific and non-specific enzymes for furanosyl-containing conjugates: biosynthesis, metabolism, and chemo-enzymatic synthesis. *Carbohydr. Res.* **356**, 44–61 (2012).
87. Rietschel-Berst, M. *et al.* Extracellular exo-beta-galactofuranosidase from *Penicillium charlesii*: isolation, purification, and properties. *J. Biol. Chem.* **252**, 3219–3226 (1977).
88. Daley, L. S. & Strobel, G. A.  $\beta$ -Galactofuranosidase activity in *Helminthosporium sacchari* and its relationship to the production of helminthosporoside. *Plant Sci. Lett.* **30**, 145–154 (1983).
89. Livingston, R. S. & Scheffer, R. P. Conversion of *Helminthosporium sacchari* Toxin to Toxoids by  $\beta$ -Galactofuranosidase from *Helminthosporium*. *Plant Physiol.* **72**, 530–534 (1983).
90. Lugt, A. W. V. B.-V. D. *et al.* New structural features of the antigenic extracellular polysaccharides of *Penicillium* and *Aspergillus* species revealed with exo-beta-D-galactofuranosidase. *J. Bacteriol.* **174**, 6096–6102 (1992).
91. Cousin, M. A., Notermans, S., Hoogerhout, P. & Van Boom, J. H. Detection of beta-galactofuranosidase production by *Penicillium* and *Aspergillus* species using 4-nitrophenyl beta-D-galactofuranoside. *J. Appl. Bacteriol.* **66**, 311–317 (1989).
92. Pletcher, C. H., Lomar, P. D. & Gander, J. E. Factors affecting the accumulation of exocellularexo- $\beta$ -d-galactofuranosidase and other enzymes from *Penicillium charlesii*. *Exp. Mycol.* **5**, 133–139 (1981).

93. Miletti, L. C., Marino, C. & Colli, W. Immobilized 4-aminophenyl 1-thio- $\beta$ -D-galactofuranoside as a matrix for affinity purification of an exo- $\beta$ -D-galactofuranosidase. *Carbohydr. Res.* **7** (1999).
94. Tuekam, B. A., Park, Y.-I., Unkefer, C. J. & Gander, J. E. Relationship of Exo- $\beta$ -D-Galactofuranosidase Kinetic Parameters to the Number of Phosphodiester in *Penicillium fellutanum* Peptidophosphogalactomannan: Enzyme Purification and Kinetics of Glycopeptide and Galactofuran Chain Hydrolysis. *Appl. Environ. Microbiol.* **67**, 4648–4656 (2001).
95. Wallis, G. L. F., Hemming, F. W. & Peberdy, J. F. An extracellular L-galactofuranosidase from *Aspergillus niger* and its use as a tool for glycoconjugate analysis. *Biochim. Biophys. Acta* **10**.
96. Miletti, L. C. *et al.* Evidence for exo  $\beta$ -D-galactofuranosidase in *Trypanosoma cruzi*. *Mol. Biochem. Parasitol.* **127**, 85–88 (2003).
97. Reyes, F. *et al.* Purification of a new galactanase from *Penicillium oxalicum* catalysing the hydrolysis of -(1-5)-galactofuran linkages. **281**, 4 (1992).
98. Ramli, N., Fujinaga, M., Tabuchi, M., Takegawa, K. & Iwahara, S. Isolation and Characterization of a Novel Endo-  $\beta$  -galactofuranosidase from *Bacillus* sp. *Biosci. Biotechnol. Biochem.* **59**, 1856–1860 (1995).
99. Mennink-Kersten, M. A. S. H., Ruegebrink, D., Wasei, N., Melchers, W. J. G. & Verweij, P. E. In Vitro Release by *Aspergillus fumigatus* of Galactofuranose Antigens, 1,3- $\beta$ -D-Glucan, and DNA, Surrogate Markers Used for Diagnosis of Invasive Aspergillosis. *J. Clin. Microbiol.* **44**, 1711–1718 (2006).
100. Mariño, K., Lima, C., Maldonado, S., Marino, C. & de Lederkremer, R. M. Influence of exo beta-D-galactofuranosidase inhibitors in cultures of *Penicillium fellutanum* and modifications in hyphal cell structure. *Carbohydr. Res.* **337**, 891–897 (2002).
101. Dubourdieu, D., Desplanques, C., Villetaz', J.-C. & Ribereau, P. INVESTIGATIONS OF AN INDUSTRIAL P\_D-GLUCANASE FROM *Tri.* **11**.
102. Marino, C. *et al.* 1-Thio- $\beta$ -D-galactofuranosides: synthesis and evaluation as  $\beta$ -D-galactofuranosidase inhibitors. *Glycobiology* **8**, 901–904 (1998).
103. Varela, O., Marino, C. & de Lederkremer, R. M. Synthesis of p-nitrophenyl  $\beta$ -D-galactofuranoside. A convenient substrate for  $\beta$ -galactofuranosidase. *Carbohydr. Res.* **155**, 247–251 (1986).

104. Matsunaga, E. *et al.* Draft Genome Sequence of *Streptomyces* sp. JHA19, a Strain That Possesses  $\beta$ -D-Galactofuranosidase Activity. *Genome Announc.* **2**.
105. Matsunaga, E. *et al.* Identification and Characterization of a Novel Galactofuranose-Specific  $\beta$ -D-Galactofuranosidase from *Streptomyces* Species. *PLOS ONE* **10**, e0137230 (2015).
106. Matsunaga, E. *et al.* Characterization of a PA14 domain-containing galactofuranose-specific  $\beta$ -D-galactofuranosidase from *Streptomyces* sp. *Biosci. Biotechnol. Biochem.* **81**, 1314–1319 (2017).
107. Matsunaga, E., Higuchi, Y., Mori, K., Tashiro, K. & Takegawa, K. Draft Genome Sequence of *Streptomyces* sp. JHA26, a Strain That Harbors a PA14 Domain Containing  $\beta$ -d-Galactofuranosidase. *Genome Announc.* **5**, e00190-17 (2017).
108. Kenry, Duan, Y. & Liu, B. Recent Advances of Optical Imaging in the Second Near-Infrared Window. *Adv. Mater.* **30**, 1802394 (2018).
109. Pansare, V. J., Hejazi, S., Faenza, W. J. & Prud'homme, R. K. Review of Long-Wavelength Optical and NIR Imaging Materials: Contrast Agents, Fluorophores, and Multifunctional Nano Carriers. *Chem. Mater.* **24**, 812–827 (2012).
110. Hong, G., Antaris, A. L. & Dai, H. Near-infrared fluorophores for biomedical imaging. *Nat. Biomed. Eng.* **1**, 0010 (2017).
111. Weissleder, R. A clearer vision for in vivo imaging. *Nat. Biotechnol.* **19**, 316–317 (2001).
112. Frangioni, J. V. In vivo near-infrared fluorescence imaging. *Curr. Opin. Chem. Biol.* **7**, 626–634 (2003).
113. Ding, F., Zhan, Y., Lu, X. & Sun, Y. Recent advances in near-infrared II fluorophores for multifunctional biomedical imaging. *Chem. Sci.* **9**, 4370–4380 (2018).
114. Hemmer, E. *et al.* Upconverting and NIR emitting rare earth based nanostructures for NIR-bioimaging. *Nanoscale* **5**, 11339–11361 (2013).
115. Foucault-Collet, A. *et al.* Lanthanide near infrared imaging in living cells with Yb<sup>3+</sup> nano metal organic frameworks. *Proc. Natl. Acad. Sci.* **110**, 17199–17204 (2013).
116. Martinić, I., Eliseeva, S. V. & Petoud, S. Near-infrared emitting probes for biological imaging: Organic fluorophores, quantum dots, fluorescent proteins, lanthanide(III) complexes and nanomaterials. *J. Lumin.* **189**, 19–43 (2017).
117. Dong, H. *et al.* Lanthanide Nanoparticles: From Design toward Bioimaging and Therapy. *Chem. Rev.* **115**, 10725–10815 (2015).

118. Bünzli, J.-C. G. Benefiting from the Unique Properties of Lanthanide Ions. *Acc. Chem. Res.* **39**, 53–61 (2006).
119. Bünzli, J.-C. G. Lanthanide Luminescent Bioprobes (LLBs). *Chemistry Letters* <https://infoscience.epfl.ch/record/131021> (2009) doi:10.1246/cl.2009.104.
120. Eliseeva, S. V. & Bünzli, J.-C. G. Lanthanide luminescence for functional materials and bio-sciences. *Chem. Soc. Rev.* **39**, 189–227 (2009).
121. Heffern, M. C., Matosziuk, L. M. & Meade, T. J. Lanthanide Probes for Bioresponsive Imaging. *Chem. Rev.* **114**, 4496–4539 (2014).
122. Faulkner, S., Pope, S. J. A. & Burton-Pye, B. P. Lanthanide Complexes for Luminescence Imaging Applications. *Appl. Spectrosc. Rev.* **40**, 1–31 (2005).
123. Hemmilä, I. & Laitala, V. Progress in Lanthanides as Luminescent Probes. *J. Fluoresc.* **15**, 529–542 (2005).
124. Lanthanide Chemistry: From Coordination in Chemical Complexes

## **CHAPTER II**

### **Galactofuranosidase – an original and specific glycoside hydrolase**

<b>I. Recombinant DNA technology in protein production.....</b>	<b>46</b>
1. The basic principles of gene and plasmid manipulations.....	46
2. Expression systems for recombinant protein production .....	48
<b>II. Obtaining a desired Galf-ase plasmid construct – the applications of gene (sub)cloning techniques .....</b>	<b>50</b>
1. Plasmid construct pET-50b(+) – screening of Galf-ase gene insert bearing plasmid.....	50
2. Galf-ase gene subcloning – preparing pET-28a(+) plasmid construct.....	52
<b>III. Production and characterization of recombinant Galf-ase.....</b>	<b>54</b>
1. Overexpression and purification .....	54
2. Optimum conditions and kinetic parameters determination .....	56
2.1. Substrate specificity .....	56
2.2. Temperature and pH properties .....	57
2.3. Freeze-thaw stability.....	58
2.4. Kinetic analysis .....	60
3. Inhibition studies.....	62
4. Crystallization trials .....	64
<b>Résumé.....</b>	<b>66</b>
<b>Bibliography.....</b>	<b>67</b>

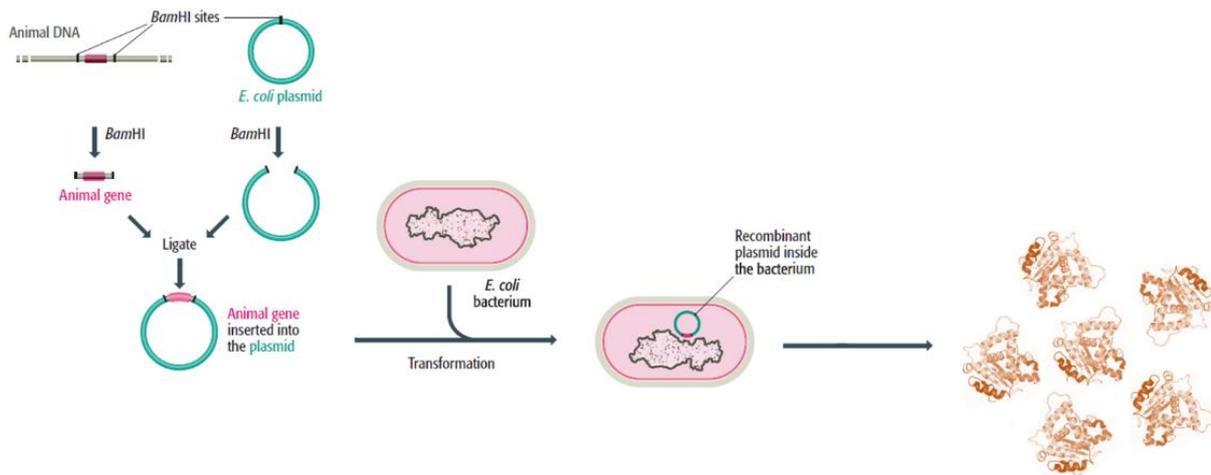
## I. Recombinant DNA technology in protein production

### 1. The basic principles of gene and plasmid manipulations

In almost every area of biochemical or biotechnological research, obtaining a sufficient quantity of the substance of interest is a major concern.<sup>1</sup> When focusing on investigation of proteins (enzymes), only few are abundant enough to be isolated from their native hosts, since *in vivo* they are present in low quantities that complement the needs of the cell, or are found in sources inconvenient for isolation to yield pure protein, free of contaminants, needed for proper characterization.<sup>2</sup>

These difficulties have been largely eliminated through the development of molecular cloning techniques, also referred to as genetic engineering or recombinant DNA technology, which have, since the late 1970s, enormously contributed to the progress in biochemistry and gave a rise to the biotechnology industry.<sup>1</sup> These techniques encompasses a number of experimental protocols leading to the transfer, cloning of a structural gene from one organism to another.<sup>3</sup> Today, these techniques are standardized and available as commercial products, known as kits that are routinely used in almost any molecular biology laboratory.<sup>4</sup> The production of large quantities of scarce and/or novel proteins using recombinant DNA technology is a relatively straightforward manner that requires three components: a gene, a recipient vector, and an expression host that will produce the desired protein.<sup>3,5</sup> The following section will briefly highlight main points of recombinant protein production.

The initial procedure is to obtain a clone (copy) of a single specific protein coding fragment (insert/target gene) of DNA from a donor organism. The sequence must be identified, selected and extracted from the complete DNA of an organism (genome) by PCR or enzymatic digestion into clonable elements. Once the specific clone that carries the target DNA has been cleaved (cut, or digested) or amplified, it has to be joined (ligated) to another DNA entity (a cloning vector) to form a new, recombined DNA molecule (cloning vector– insert DNA construct or DNA construct) (**Figure II-1.**)<sup>3,5</sup>



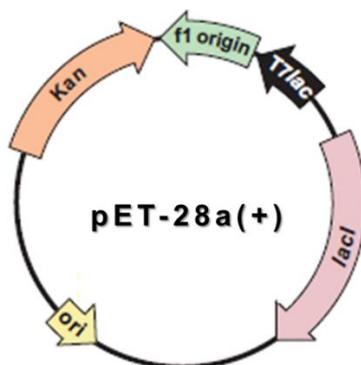
**Figure II-1.** Schematic representation of recombinant DNA cloning procedure.

Expression of the protein of interest as soluble, active and easy to purify, requires the use of particular vector. The plasmids, bacterial circular extrachromosomal DNA, that are autoreplicative and relatively small, less than 10 kb, meet all the criteria to serve as recipient DNA vectors.<sup>5</sup>

The pET systems are the plasmid synthetic vectors, designed as both cloning and expression vectors. They were originally constructed by Studier and colleagues in the late 1980s and presents the most powerful system for the cloning and expression of recombinant proteins in *E. coli*.<sup>6</sup> This system permits easier cloning, expression, detection, and purification of target proteins and has been greatly expanded and includes over 35 vector types.<sup>6,7</sup>

Commonly used pET-28a(+) vector contains 5 369 pb and has primary characteristics common to most plasmids vectors as pBR322 origin of replication, kanamycin resistance selection marker and multiple cloning region/site (**Figure II-2.**). Multiple cloning region has restriction enzyme sequences that enable usage of multiple (sub)cloning strategies. It carries an *N*-terminal and an optional *C*-terminal sequences adjacent upstream and downstream to the cloning site that encodes 6xHis-tag peptide fusion tags. The method of cloning permits to determine whether or not these tags or other additional amino acids from the vector will be fused to the protein of interest. Also multiple cloning region contains properly positioned transcriptional and translational protein expression control sequences which are under strong T7 promotor control as well as T7 transcription terminator. Promotor T7 and terminator T7 term are also convenient sequencing points

for directed double strand forward and reverse sequencing verification of the cloning site.<sup>6</sup>



**Figure II-2.** Schematic representation of pET-28a(+) vector backbone.

### 2. Expression systems for recombinant protein production

The following phase in production of protein of interest is finding an appropriate host organism (expression system) that allows selective overexpression of a cloned gene. Heterologous (recombinant) proteins from cloned DNA originating from a wide variety of organisms are usually successfully produced using both prokaryotic and eukaryotic host organisms such as bacteria, yeasts, fungi, plants or insects and mammalian cells. Among prokaryotic microorganism host systems that are available, *Escherichia coli* is used when additional posttranslational protein processing, like protein glycosylation, characteristic for eukaryotes is not necessary and produced proteins are used for functional and structural studies.<sup>3,8</sup>

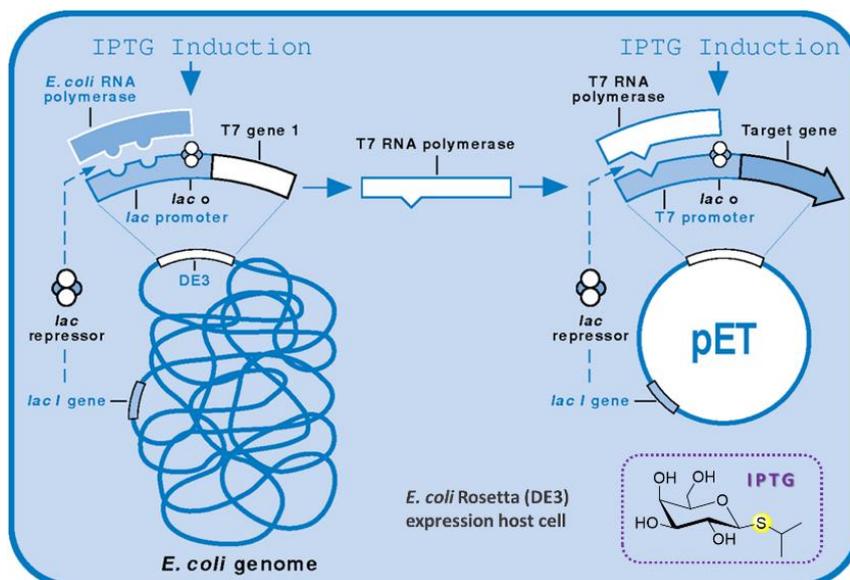
*E. coli* is approximately 0.5  $\mu\text{m}$  in diameter and 1.0–3.0  $\mu\text{m}$  in length Gram-negative, non-spore forming, facultative anaerobe bacilli that has one circular chromosome coding about 4000 genes. It is motile bacteria with surface carries about 10 flagella and thousands of pili that allow the cells to attach to surfaces. The earliest description and laboratory isolation, originally called *Bacterium coli commune*, was from the feces of a child in 1885 by the Austrian pediatrician Theodor Escherich. It is a common commensal inhabitant of the terminal small intestine and large intestine (colon) of mammals, including humans, and normally is harmless. The pathogenic strains are either causing disease inside the intestinal tract or at extra-intestinal sites.<sup>9,10</sup>

As bacteria it provides many advantages to be cultured in the laboratory and used as model organism for research since it is easy to maintain at low cost and under well-

defined laboratory conditions. Its main advantage is the rapid growth followed by high cell density which provides large numbers of genetically identical cells (clones) as they are derived by binary division from a single cell. Moreover, it can be grown in liquid culture or as solid culture on agar plates and can be stored for a short (weeks) or long period (years). The small size of the *E. coli* genome has been sequenced in entirety and provides obvious advantages for easy genetic analysis and engineering.<sup>11-13</sup>

The extensive use of *E. coli* is the result of vast amount of research that has been carried out on its genetics, molecular biology, biochemistry, and physiology. Therefore it is one of the best understood and characterized living organisms. The laboratory model strains such as *E. coli* K-12 and BL21 (DE3) and their derivatives have been well adapted to laboratory manipulation methods over the past seventy years and are widely used.<sup>14</sup>

Suitable *E. coli* hosts for pET system, that is pET-28a(+) vector that has the bacteriophage T7 promoter, to control the transcription start site and the expression of cloned sequences genes, are DH5 $\alpha$  and Rosetta™ (DE3) (Novagen®) strains commonly used as host cells for cloning and for protein expression. Rosetta (DE3) strain is a BL21 (DE3) strain derivative, that carries a chromosomal copy of the T7 RNA polymerase gene under the control of the *lac* repressor as well as pRARE plasmid encoding tRNA genes for all the rare eukaryotic codon bias. Since *E. coli* RNA polymerase does not recognize the T7 promoter, there is no transcription of the target gene in the absence of a T7 RNA polymerase under repressive conditions. After induction by isopropyl- $\beta$ -D-1-thiogalactopyranoside (IPTG), a structural non-metabolizable analogue of allolactose, and in the presence of T7 RNA polymerase, which is extremely selective for T7 promoter and which allows high expression levels of cloned genes, the production of desired gene can reach more than 50% of the total cell protein content (**Figure II-3.**). This system is extremely powerful for controlled production at optimum growth conditions to feature satisfactory quantity of protein.<sup>6,8,15</sup>



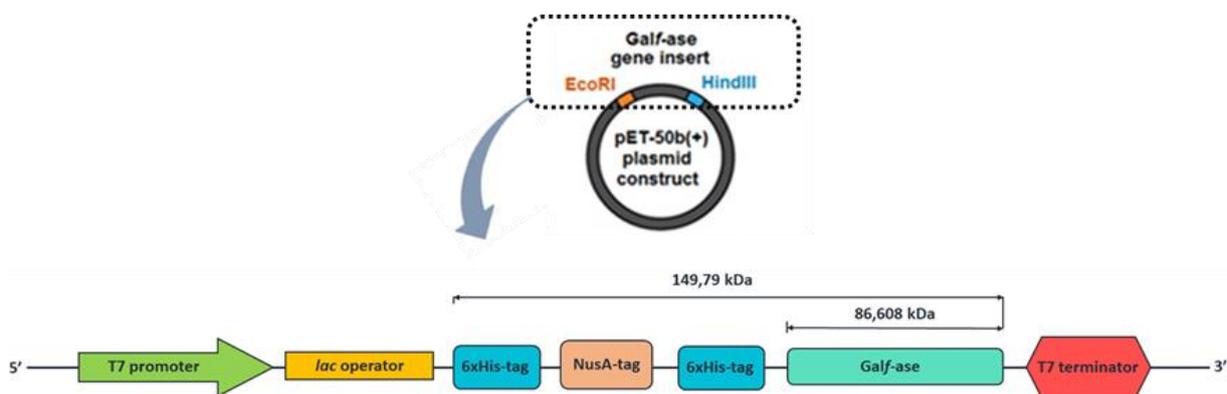
**Figure II-3.** Schematic representation of the pET vector and *E. coli* Rosetta™ (DE3) host expression system the control elements.<sup>6</sup>

## II. Obtaining a desired Gal $\beta$ -ase plasmid construct – the applications of gene (sub)cloning techniques

### 1. Plasmid construct pET-50b(+) – screening of Gal $\beta$ -ase gene insert bearing plasmid

The first Gal $\beta$ -ase-specific enzyme and gene encoding it have been identified in *Streptomyces spp.* strain JHA19 and cloned into pET-50b(+) plasmid vector. According to literature available information, the entire Gal $\beta$ -ase ORF fragment contains 2 361 bp, which encode Gal $\beta$ -ase protein of 786 amino acids with the predicted molecular mass of 86.6 kDa and belonging to GH2 family.<sup>16,17</sup>

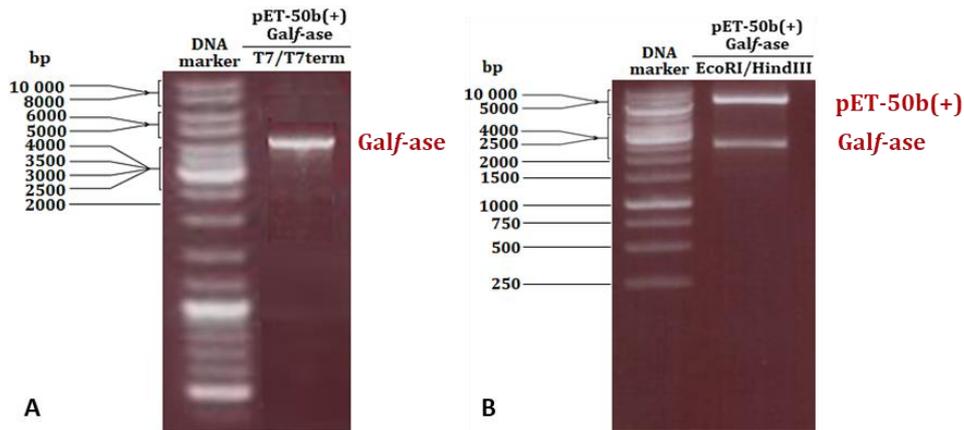
In our hands, pET-50b(+) plasmid construct bearing Gal $\beta$ -ase gene insert, kindly provided by Prof. K. Takegawa, was the starting point to reach given objectives of this thesis. However this first cloned Gal $\beta$ -ase protein was co-expressed with a large fusion NusA-tag<sup>18</sup> (*N*-utilization substance A) as solubility enhancer and two 6xHis-tags for affinity purification purpose, thus leading to a fusion large protein of around 150 kDa (**Figure II-4.**).



**Figure II-4.** Schematic representation of pET-50b(+) plasmid construct cloning/expression region coding for *N*-terminal Nus protein and double 6xHis-tagged Gal-f-ase.

These multiple fusion tags led to observation that they might hamper i) the overexpression of the Gal-f-ase leading to low yield, and ii) the physico-chemical and enzymatic properties of this original biocatalyst. In general, pET-50b(+) plasmid construct and the character of fused Gal-f-ase enzyme was incompatible with obtaining the biochemical and structural characteristics as closest possible to the performance of a natural enzyme. Therefore, the confirmation of the presence of Gal-f-ase gene insert and mapping its exact position in pET-50b(+) vector was necessary in order to conduct further molecular (sub)cloning. Targeted polymerase chain reaction (PCR) amplification of Gal-f-ase gene and restriction enzyme analysis were undertaken as preparatory analysis steps leading to preferable choice of recipient vector which was to provide appropriately fused Gal-f-ase enzyme.

Briefly, recombinant plasmid construction has been screened by PCR using the universal primers, T7 and T7term, targeting plasmid vector DNA flanking the Gal-f-ase insert gene. The agarose-gel electrophoresis of the amplified region showed that the synthesis of the full-length Gal-f-ase gene was successful, giving strong band signal at 4.3 kb, confirming the presence of the gene insert at predicted multiple cloning site (**Figure II-5. A**), thus allowing the double digestion with EcoRI and HindIII restriction enzymes. Restriction enzyme analysis revealed two strong band signals at 6.5 kb and 2.5 kb corresponding to expected pET-50b(+) vector and Gal-f-ase gene size (**Figure II-5. B**) and allowed us to obtain the Gal-f-ase fragment as closest as possible to its ORF fragment size of 2.3 kb.

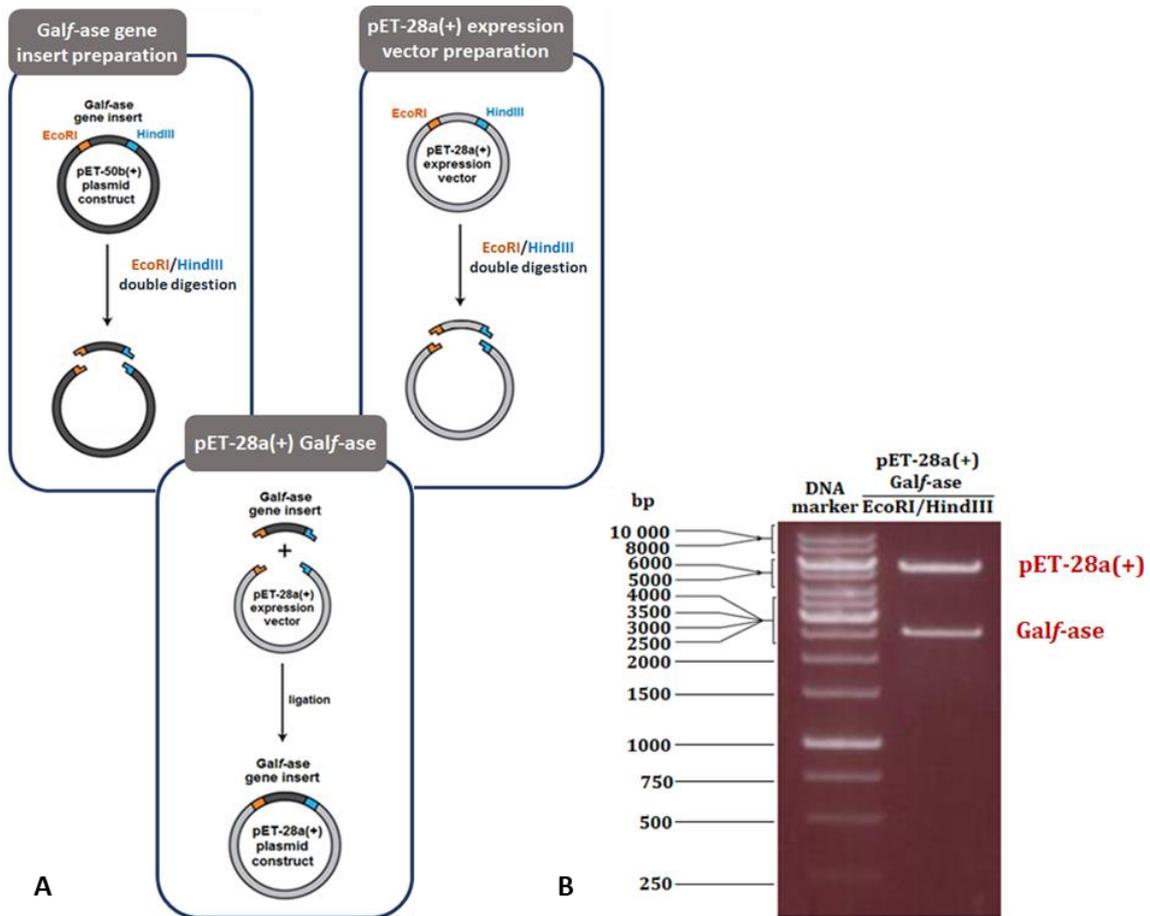


**Figure II-5.** PCR and restriction enzyme analysis of pET-50b(+) plasmid construct. **A)** Agarose-gel electrophoresis of the PCR reaction indicating the amplification efficiency of *Galf-ase* insert gene between T7 and T7term. **B)** Agarose-gel electrophoresis of pET-50b(+) plasmid construct after *EcoRI* and *HindIII* double digestion reaction.

## 2. *Galf-ase* gene subcloning – preparing pET-28a(+) plasmid construct

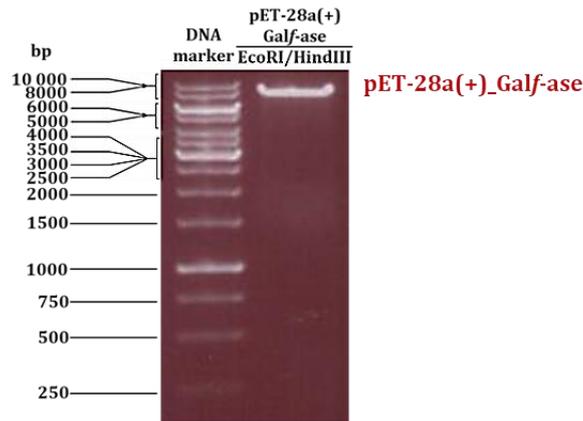
To avoid the presence of large NusA-tag at the *N*-terminal end of *Galf-ase* protein, and to gain the desired functionality, subcloning methodology was undertaken. For this purpose, another expression vector was chosen, pET-28a(+), using a cloning strategy that will add only one *N*-terminal 6xHis-tag for purification and detection (Western blot) purposes.

Plasmid construct pET-50b(+), containing *Galf-ase* insert gene, was double digested with *EcoRI* and *HindIII* restriction enzymes and the resulting 2.3 kb *Galf-ase* gene fragment was gel excised and purified. After ligation to *EcoRI* and *HindIII*-cut pET-28a(+) vector, newly obtained plasmid construct, containing ligated *Galf-ase* gene, (**Figure II-6. A**) was introduced into competent *E. coli* DH5 $\alpha$  cells. Double digestion with *EcoRI* and *HindIII* restriction enzymes gave two strong band signals at 5.3 kb and 2.3 kb corresponding to expected pET-28a(+) vector and *Galf-ase* gene size, confirming the success of ligation reaction (**Figure II-6. B**).



**Figure II-6.** Schematic representation of Galf-ase gene subcloning procedure. **A)** Preparation of pET-28a(+) plasmid construct, bearing Galf-ase gene insert. **B)** pET-28a(+) plasmid construct confirmation on agarose-gel electrophoresis after EcoRI and HindIII restriction enzyme analysis.

In-frame ligation of Galf-ase gene was not possible in one step (sub)cloning reaction. To ensure that the ligated Galf-ase gene was translated in its correct reading frame, single deletion mutation, at the beginning of ORF, was introduced into EcoRI restriction site sequence that precede the Galf-ase gene sequence. Obtained plasmid construct was used for competent *E. coli* DH5 $\alpha$  cells transformation. Double digestion with EcoRI and HindIII restriction enzymes confirmed the integrity of Galf-ase gene and its correct ORF since the EcoRI restriction site was no longer entirely preserved after deletion, only one band signal at 7.7 kb was present corresponding to linearized size of pET-28a(+) plasmid construct (**Figure II-7.**).



**Figure II-7.** Confirmation of EcoRI restriction site mutagenesis on agarose-gel electrophoresis after pET-28a(+) plasmid construct EcoRI and HindIII restriction enzyme analysis.

Subcloning methodology including insert gene and vector preparation, ligation reaction and single deletion mutation was undertaken since the traditional cloning approach involving PCR insert gene amplification and ligation reaction revealed to be unsuccessful.

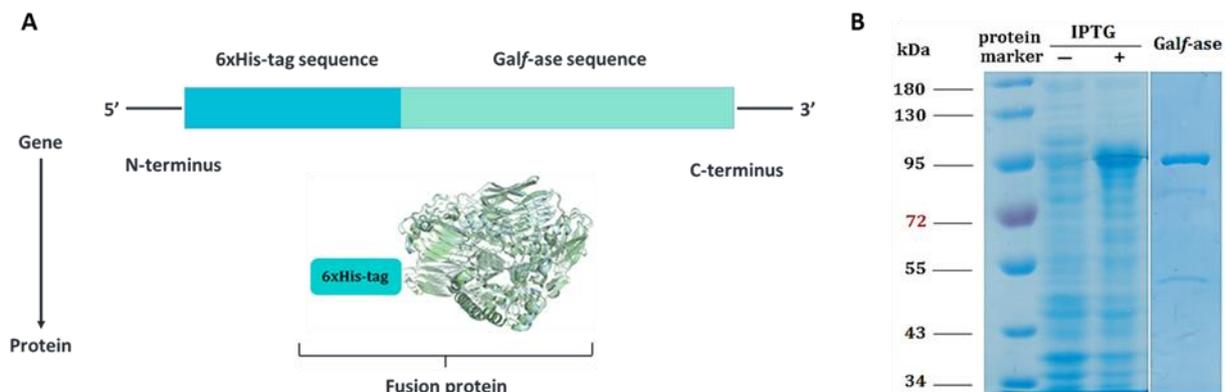
The integrity and intact of Galf-ase insert gene sequence was confirmed by DNA sequencing, performed by Eurofins Genomics, and the verified recombinant pET-28a(+) plasmid construct was used for Rosetta (DE3) protein expression strain transformation.

### III. Production and characterization of recombinant Galf-ase

#### 1. Overexpression and purification

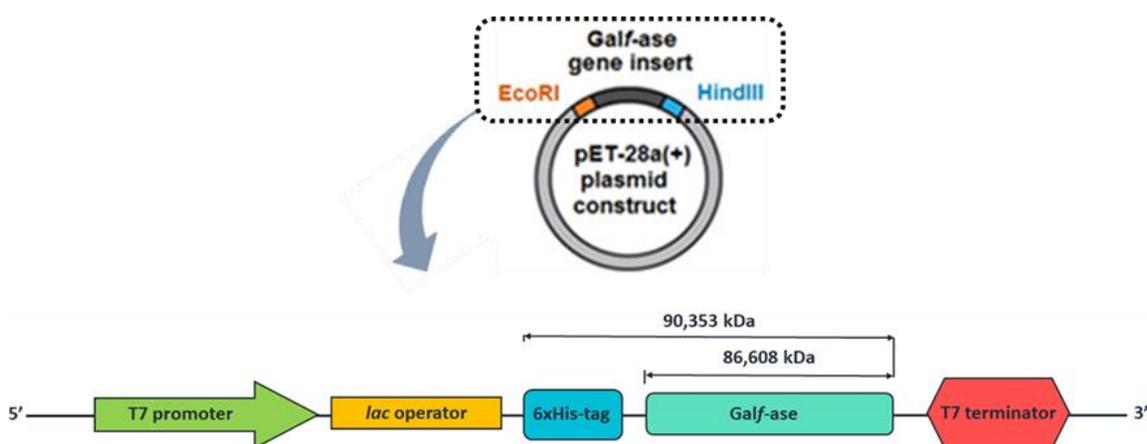
After successful subcloning, Galf-ase gene was part of expression vector pET-28a(+) and was transformed into *E. coli* Rosetta™ (DE3) (Novagen®) expression system. Galf-ase was expressed as soluble recombinant protein bearing only *N*-terminal peptide hexahistidine fusion tag (6xHis-tag) (**Figure II-8. A**). 6xHis-tag facilitated selective recombinant protein purification from the bacterial cellular environment by immobilized metal-affinity chromatography (IMAC) and, due to its small size, was less likely to interfere with protein activity and structure.<sup>8</sup> The 6xHis-tagged Galf-ase was expressed and purified by IMAC with a Ni<sup>2+</sup> immobilized on a nitrilotriacetic acid matrix (NTA-Ni<sup>2+</sup>) with a yield of 0.5 mg/litre of bacterial culture.<sup>19,20</sup> If required, size-exclusion chromatography was applied in order to additionally purify collected protein fraction. The purified enzyme displayed an intense band on the sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) at an apparent molecular weight (90.3 kDa), consistent with that theoretically

predicted for the *Gal*f-ase fusion protein (**Figure II-8. B**). The *Gal*f-ase of this purity was further used for biochemical characterization.



**Figure II-8.** The design and production of *N*-terminal fused *Gal*f-ase protein. **A**) Schematic representation of *N*-terminal single peptide 6xHis-tagged *Gal*f-ase fusion protein. **B**) The SDS-PAGE analysis of *Gal*f-ase after IPTG induction and NTA-Ni<sup>2+</sup>chromatography purification.

In this plasmid vector cloning region is under T7 promoter and inducible *lac* operator control and Rosetta™ (DE3) protein expression system is also carrying chromosomal copy of the T7 RNA polymerase gene (**Figure II-9.**). Therefore, protein expression is regulated and induced by IPTG. The inducible *lac* operator is one of the most commonly used operators for heterologous protein expression in *E. coli*. and IPTG is currently the most efficient molecular inducer used to regulate its transcriptional activity. However, during bacterial growth prior to induction, *Gal*f-ase expression was observed. This phenomenon of leaky-basal expression, limited bacterial growth and was repressed by the addition of 0.5 % of glucose to the medium.<sup>21</sup> Induction of bacterial culture at mid-exponential growth phase (OD<sub>600</sub> 0.6) at lower concentration of IPTG (0.1 mM) and prolonged, overnight incubation at low temperature (15 °C) proved optimal to yield the soluble protein. Despite this measures, *Gal*f-ase was obtained in low yields (0.5 mg/L) and overexpression was not fully accomplished since enzyme formed insoluble inclusion bodies if induction conditions varied, hence indicating signs of protein toxicity. Nevertheless, established induction conditions provided sufficient quantity and satisfactory purity, to secure experiments continuation and were not altered.



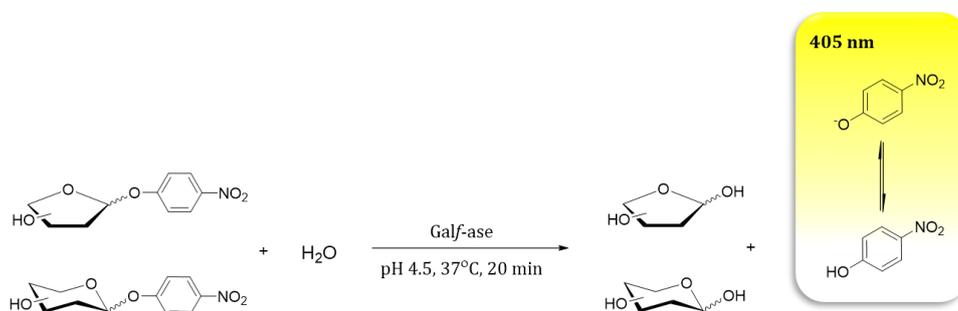
**Figure II-9.** Schematic representation of pET-28a(+) plasmid construct cloning/expression region coding for *N*-terminal single peptide 6xHis-tagged Gal-f-ase.

## 2. Optimum conditions and kinetic parameters determination

The detailed study of newly obtained Gal-f-ase enzyme was recently published as short communication in the journal *Carbohydrate Research*, Volume 480, 1 July 2019. In this section, selected paragraphs from the article *Galactofuranosidase from JHA19 Streptomyces sp.: subcloning and biochemical characterization*, covering general properties and kinetic characterization are here presented. The article, in its entirety, is enclosed in **Appendix 7.1.5**.

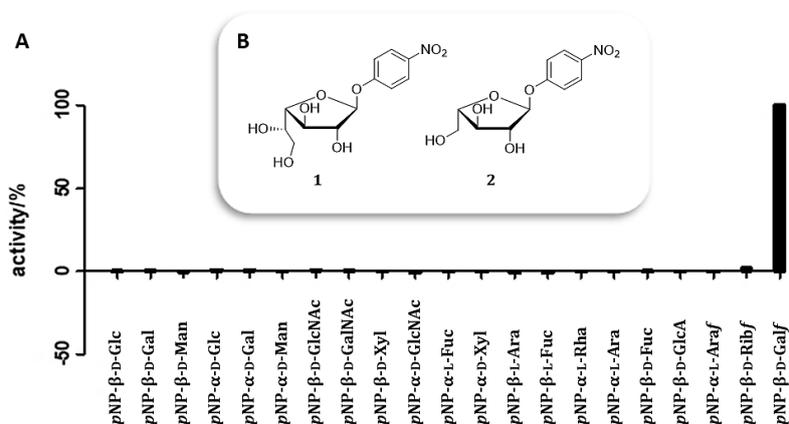
### 2.1. Substrate specificity

A Simple one-step colorimetric assay which commonly employs synthetic *para*-nitrophenol (*p*NP) sugars as artificial substrates was used for the assessment of Gal-f-ase activity. The assay directly correlates the release of formed *p*NP with the hydrolytic activity of tested enzyme and is quantified spectrophotometrically at 405 nm (**Figure II-10**).



**Figure II-10.** Schematic representation of hydrolytic reaction between Gal-f-ase and a panel of *p*NP-linked pyranosyl and furanosyl monosaccharides.

The hydrolytic activity and substrate specificity of Galf-ase was screened against 21 commercially available *p*NP-pyranosyl and -furanosyl substrates at 1 mM (**Appendix 7.1.2. Figure VII-2.**). The only hydrolysed substrate was *p*NP- $\beta$ -D-Galf and the Galf-ase exhibited no hydrolytic activity towards any other of the 20 tested *p*NP-sugars (**Figure II-11. A**). It is also noteworthy that, in these experimental conditions, no activity was observed with the *p*NP- $\alpha$ -L-Araf, stereochemical analogue of *p*NP- $\beta$ -D-Galf differing only by absence of the extra C-6 hydroxymethyl group (**Figure II-11. B**), nor the *p*NP- $\beta$ -D-Ribf. This result is in total agreement with the previous observation by the team of Takegawa.<sup>17</sup>



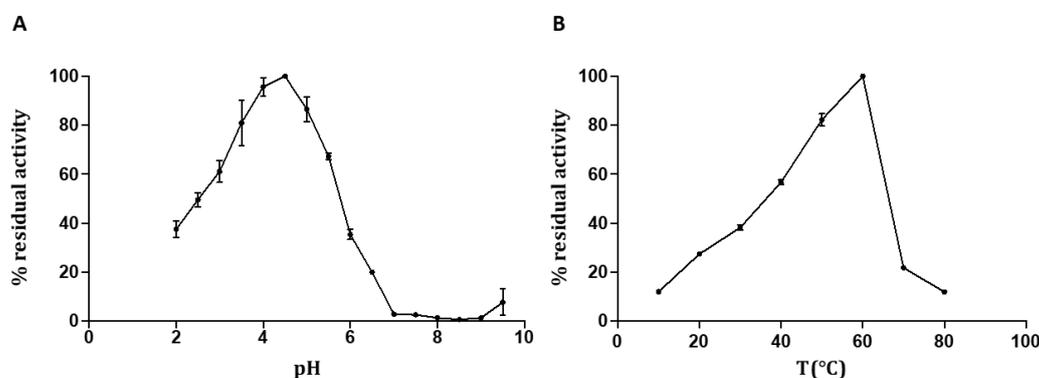
**Figure II-11.** Substrate specificity screening of Galf-ase. **A)** Substrate specificity of Galf-ase tested towards 21 *p*NP-linked monosaccharide substrates. **B)** Structural formulas of *p*NP- $\beta$ -D-Galf (1) and the *p*NP- $\alpha$ -L-Araf (2).

## 2.2. Temperature and pH properties

The biochemical properties of purified Galf-ase were investigated with *p*NP- $\beta$ -D-Galf as a substrate in different buffer systems covering the pH intervals between pH 2 and 9.5 and temperature intervals between 10 and 80 °C. The optimum pH was comprised in a narrow range between 3.5 and 5, significantly decreasing above pH 5. The peak of maximum activity is present at pH 4.5 (**Figure II-12. A**). This reflects the presence of the two catalytic glutamic acid residues, ubiquitous in most GHs, which are facing each other from opposite sides of the active site with their side residues strategically positioned about 5 Å from each other. One residue is serving as an acid/base catalyst and the other as a nucleophile. The protonation states of this key residues implies that, in the optimum reactant state, the one serving as general acid/base must be protonated and neutral, and the nucleophile must be deprotonated and negatively charged. Therefore they are

sensitive to variation in pH and exhibiting characteristic bell-shaped curve centered at the optimum pH.<sup>22</sup>

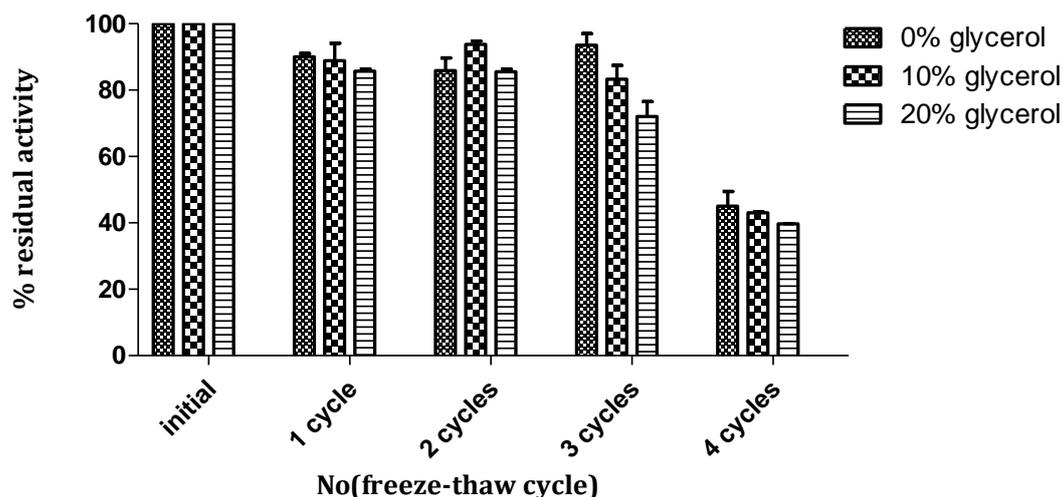
Once the optimum pH conditions were established, the influence of different temperatures on activity was also measured. For 20 min reactions, the optimal temperature was at 60 °C with a clear peak of activity (**Figure II-12. B**). Temperatures above the optimum value resulted in dramatic loss of activity. However, due to practical constraints, all the following activity assays were assessed at the physiological temperature of 37 °C. Still, this current Gal $f$ -ase biocatalyst exhibits different and improved physico-chemical properties toward the previously reported one, with a maximum pH shift from 5.5 to 4.5 and a maximum temperature from 40 °C to 60 °C, thus demonstrating the drawback of large fusion tags.



**Figure II-12.** Determination of optimum activity conditions for Gal $f$ -ase. **A)** Influence of pH on Gal $f$ -ase hydrolytic activity towards  $p$ NP- $\beta$ -D-Gal $f$ . **B)** Influence of temperature on Gal $f$ -ase hydrolytic activity towards  $p$ NP- $\beta$ -D-Gal $f$ .

### 2.3. Freeze-thaw stability

Gal $f$ -ase stability towards multiple freeze and thaw cycles was assayed over a different storage conditions, without and in the presence of glycerol (10% and 20%), and throughout 16 days. Percent of activity was normalized to an initial activity measured for each storage condition sample. Gal $f$ -ase exhibited stability across all storage conditions tested and maintained minimally 70% activity after three cycles (**Figure II-13.**). After four cycles, Gal $f$ -ase demonstrated low activity (40%) over all storage conditions and exhibited higher sensitivity to inactivation due to multiple freezing. No cryoprotectant was included in this assay to test the stability of crude Gal $f$ -ase sample which showed to be resistant to inactivation to multiple freeze-thaw cycles and retaining 80% activity after three cycles.

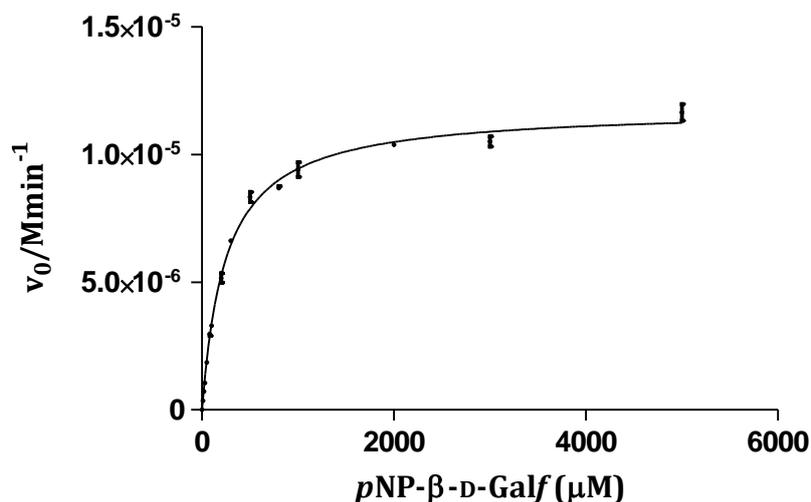


**Figure II-13.** Effect of freeze and thaw cycles and storage conditions on Galf-ase residual hydrolytic activity towards *p*NP- $\beta$ -D-Galf.

The addition of glycerol did not significantly improve or conserve enzyme stability when compared to one observed in the absence of a cryoprotectant. Glycerol is a common cryoprotectant which protects proteins from inactivation during freezing and thawing by inhibiting the formation of ice crystals upon freezing.<sup>23</sup> Also, it is worth mentioning that the purified enzyme remains fairly active over a minimum of three weeks when stored at 4°C without presence of any additives such are cryoprotectants, protease inhibitors, reducing agents, metal chelators or antimicrobial agents. All together these experiments demonstrate the good stability of the Galf-ase under storage condition, which will be highly convenient in order to use it during chemo-enzymatic synthesis.

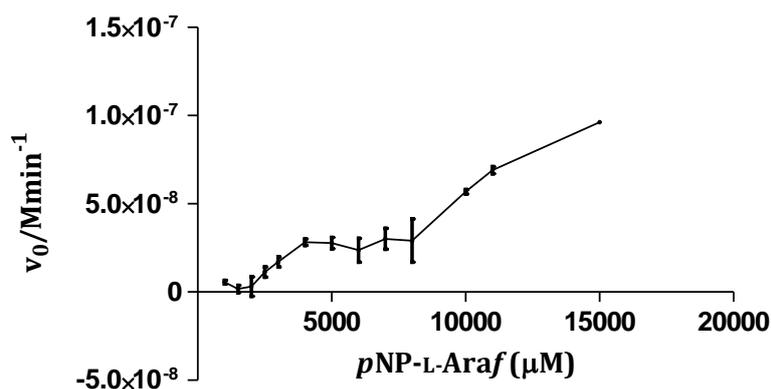
## 2.4. Kinetic analysis

To determine catalytic parameters, Gal $f$ -ase activity was further investigated with *p*NP- $\beta$ -D-Gal $f$  as a substrate at optimum pH and at appropriate temperature. The Gal $f$ -ase exhibited typical Michaelis-Menten kinetics, with  $K_M$  value of 250  $\mu$ M and with  $k_{cat}$  value of 3.5  $s^{-1}$  giving a  $k_{cat}/K_M$  of 14  $mM^{-1}s^{-1}$  (**Figure II-14.** & **Table II-1.**).



**Figure II-14.** Michaelis-Menten plot of *p*NP- $\beta$ -D-Gal $f$  hydrolysis reaction catalysed by Gal $f$ -ase. Mean values and SD error bars were calculated from three independent experiments and from a single protein preparation.

Investigation of kinetic parameters was also extended to *p*NP- $\alpha$ -L-Ara $f$ , despite displaying no hydrolysis when screened for substrate specificity. This was to further investigate specificity of Gal $f$ -ase since the previous studies with AbfD3  $\alpha$ -L-arabinofuranosidase from *Thermobacillus xylanolyticus* showed that, despite being a poor substrate for AbfD3-catalysed hydrolysis, *p*NP- $\beta$ -D-Gal $f$  was recognised as a substrate in transglycosylation reactions catalysed by this enzyme.<sup>24,25</sup> The *p*NP- $\alpha$ -L-Ara $f$  did not appear to show saturation kinetics with concentrations as high as 15 mM (**Figure II-15.**). The assay did not yield interpretable kinetic parameters since the hydrolysis rate was too low to give accurate values.



**Figure II-15.** Michaelis-Menten plot of *pNP-α-L-Araf* hydrolysis reaction catalysed by Galf-ase. Mean values and SD error bars were calculated from three independent experiments and from a single protein preparation.

These observations emphasized that Galf-ase has a really poor Araf-ase activity as it badly recognised the  $\alpha$ -L-arabinofuranose moiety, thus demonstrating the crucial role of the C-6 hydroxymethyl group in the specific recognition by the Galf-ase. In addition, when compared to the recombinant  $\alpha$ -L-Araf hydrolases AbfD3 and Araf51 owing dual specificity for  $\alpha$ -L-Araf and  $\beta$ -D-Galf, the specificity of Galf-ase is incredibly increased toward the hexofuranosyl moiety by a factor of about 100-fold (**Table II-1.**). When compared to the only two reported cloned Galf-ases, from *Streptomyces sp.* strains JHA19 and JHA26, the values are not in the same order of magnitude. For Galf-ase JHA19 and JHA26,  $K_M$  value ranges from 4.4 to 6.8 mM respectively,<sup>17,26</sup> yielding at least a 18-fold lower substrate affinity compared to Galf-ase's  $K_M$  which is in the  $\mu$ M level. Finally, the Galf-ase parameters described herein are the best reported to date and fully compatible with the development of the specific and efficient biocatalyzed incorporation of Galf entities into more complex glycoconjugates.

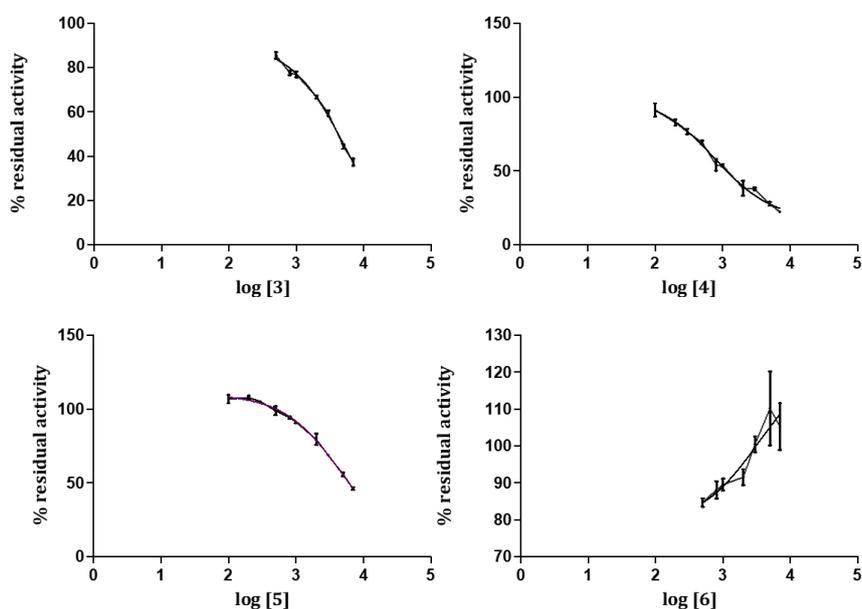
**Table II-1.** Comparison of Michaelis-Menten kinetics constants for hydrolysis of  $\beta$ -D-Galf (**1**) and  $\alpha$ -L-Araf (**2**) by different Araf-ases and Galf-ases.

Enzyme	$K_M$ for <b>1</b> (mM)	$k_{cat}$ ( $s^{-1}$ )	$k_{cat}/K_M$ ( $s^{-1}\cdot mM^{-1}$ )	$K_M$ for <b>2</b> (mM)	$k_{cat}$ ( $s^{-1}$ )	$k_{cat}/K_M$ ( $s^{-1}\cdot mM^{-1}$ )
AbfD3	> 50	15	0.13	3.5	425	122
Araf51	53	12	0.23	0.25	103	412
JHA 19 & 26	4.4 & 6.8	only $V_M$	–	–	–	–
Galf-ase	0.25	3.5	14	nd <sup>a</sup>	nd	nd

<sup>a</sup> nd not determined

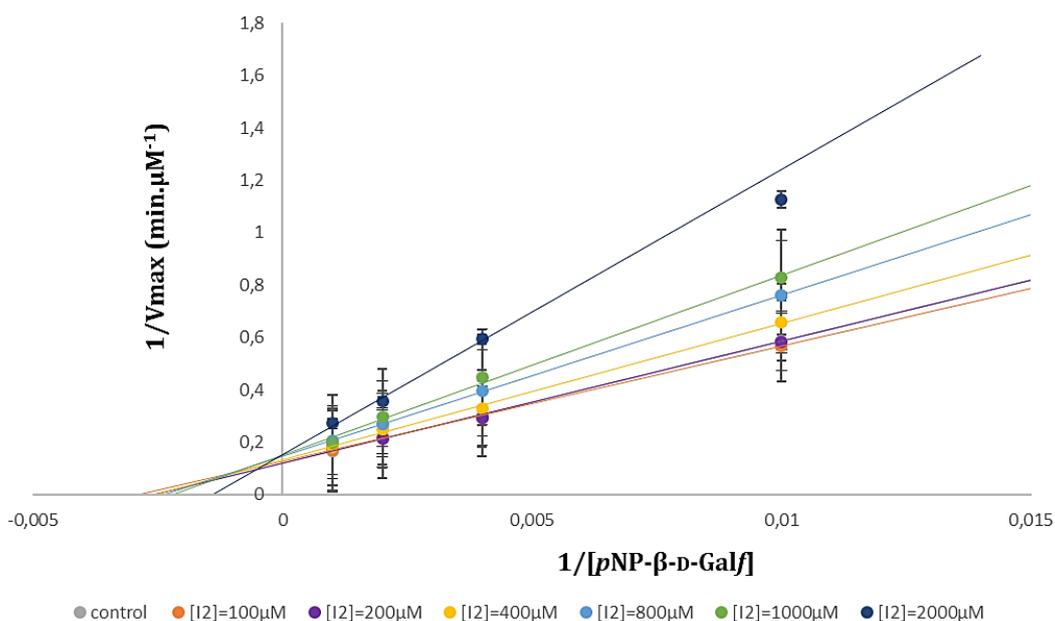
### 3. Inhibition studies

The inhibitory properties of several compounds, i.e. alkyl, thioheteroaryl and thioimidoyl  $\beta$ -D-galactofuranosides **3-6**, were investigated by incubating the enzyme, Galf-ase, in the presence of both the substrate *p*NP- $\beta$ -D-Galf and compounds **3-6** (**Figure II-16.**). These compounds were tested as inhibitors because (i) the glycone part, that is  $\beta$ -D-Galf, is recognized by the active site, (ii) the aglycone part is mimicking a carbohydrate moiety and (iii) to probe the stability of the thioglycosidic bond.<sup>27</sup> It has been reported previously that alkyl, aryl and heteroaryl 1-thio- $\beta$ -D-galactofuranoside derivatives are good inhibitors of exo  $\beta$ -D-galactofuranosidase, isolated from *P. fellutanum* culture media, and D-galactono-1,4-lactone has been evaluated as a reference inhibitor with an IC<sub>50</sub> value of 0.02 mM.<sup>28</sup>



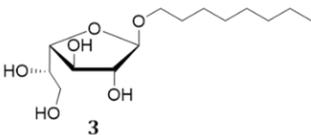
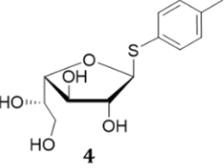
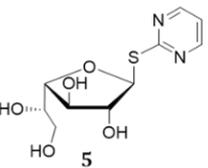
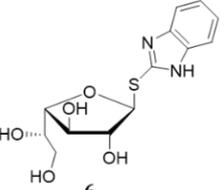
**Figure II-16.** Effects of concentration of inhibitors **3-6** on the hydrolytic activity of Galf-ase towards *p*NP- $\beta$ -D-Galf. Mean values and SD error bars were calculated from three independent experiments and from a single protein preparation.

In our hands, compounds **3**, **5** and **6** showed weak inhibitory activity with IC<sub>50</sub> values of 6, 5 and 3 mM respectively, while Galf-ase was more sensitive to the inhibition by thiogalactofuranoside **4** (**Table II-2.**). This latter showed the best inhibition activity, with IC<sub>50</sub> value of 0.8 mM. The analysis of the Lineweaver–Burk plot (**Figure II-17.**) indicated that the compound **4** is a competitive inhibitor of the Galf-ase with *K<sub>i</sub>* value of 0.4 mM. Nevertheless, compared to the only reference L-arabino-1,4-lactone inhibitor previously probed on Galf-ase,<sup>17</sup> **4** exhibited a 100 times increased in inhibition, that can be mostly attributed to its structural analogy to the artificial substrate.



**Figure II-17.** Lineweaver–Burk plot of competitive inhibition of Galf-ase hydrolytic activity towards *pNP-β-D-Galf* by **4**. Mean values and SD error bars were calculated from two independent experiments and from a single protein preparation.

**Table II-2.** Comparison of the inhibitory effects of different galactofuranosides (**3-6**) towards Galf-ase and their respective  $IC_{50}$  and  $K_i$  values.

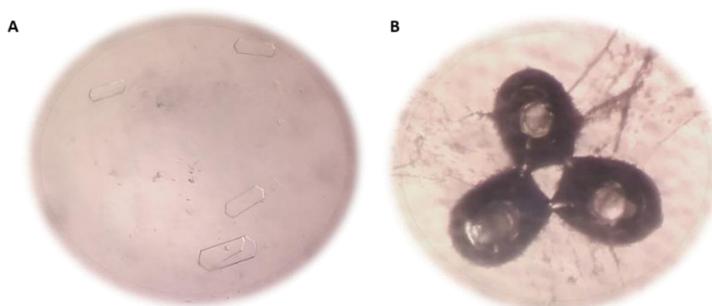
Compound	$IC_{50}$	$K_i$ (mM)
 3	6	3.0
 4	0.8	0.4
 5	5	2.5
 6	3	1.5

#### 4. Crystallization trials

$\beta$ -D-Galactofuranosidase has been only recently cloned and no structural study is available yet. To conduct a preliminary crystallization trials, enzyme was successfully expressed as the soluble 6xHis-tagged and was purified by coupling two consecutive chromatographic techniques, an initial affinity column chromatography and size-exclusion chromatography, subsequently used to achieve a purity of greater than 95% as determined by SDS-PAGE. Freshly concentrated protein (10 mg/ $\mu$ L) was used for initial crystallization screening by applying the on plate sitting-drop vapour-diffusion method with the BScreen Classic HTS I+II (Jena Bioscience GbmH, Jena, Germany) commercially available screening kits covering 192 different crystallization conditions.

Crystals of native Galf-ase, in ligand-free form, have grown after one day and reached their final size after three weeks in several conditions. In particular, tetragonal crystals (**Figure II-14. A**) obtained in the presence of potassium sodium tartarate revealed promising as diffraction-quality crystals of Galf-ase while other crystals immediately appear to be salt precipitants (**Figure II-14. B**). Diffraction properties of selected Galf-ase crystals were analysed by the X-ray beam and unfortunately diffraction data revealed crystals to be of salt origin and no data set was collected.

The commercial kit allowed to screen initial conditions in order to find hit points of crystallization conditions that may be later optimized. Protein crystallization is affected by many parameters and in this crystallization trials the quantity of obtained Galf-ase enzyme most certainly limited the successful protein crystallization as the recommended minimum concentration of 10 mg/mL was barely reached after the lysis of 30 L of induced bacterial culture.



**Figure II-14.** Photographs of the crystals obtained during Galf-ase crystallization trials. **A)** Crystals submitted to X-ray crystallography proved to be of salt origin. **B)** One of the examples of the salt crystals obtained during Galf-ase crystallography trials.

#### IV. Conclusion

In this chapter, we provided a complete biochemical and kinetic characterization of a subcloned recombinant Gal $f$ -ase since the recent publication of the first ever cloned one. The work comprised of preparation of appropriate plasmid construct pET-28a(+) and production of single 6xHis-tag fused enzyme in prokaryotic *E. coli* Rosetta expression system. The purified soluble enzyme was subsequently thoroughly characterized and was submitted to crystallization trials.

This enzyme was obtained with a reasonable yield of 0.5 mg/liter of culture, and optimally active at pH 4.5 and at 60°C. It exhibited substrate specificity exclusively towards *p*NP- $\beta$ -D-Gal $f$ , giving the best  $K_M$  value of 250  $\mu$ M, and is inhibited by substrate's structural thio analogue ( $IC_{50}$  = 0.8 mM), which proved to be a moderate competitive inhibitor ( $K_I$  = 1 mM). Although, as mentioned before, this yield was reasonable for characterization studies, it presents limiting factor for crystallization screening assays. Observed insufficient overexpression and consequently low yields, present difficulty for finding promising crystallization conditions especially when it comes to yield protein crystals suitable for X-ray diffraction.

Since to date, only a few furanosidases have been involved in the chemoenzymatic approaches, it was of great interest to explore and evaluate the potential synthetic ability of this enzyme to act as an efficient thioligase after site-directed mutagenesis. Therefore, its biocatalyst potential for diverse furanothioglycoside synthesis was screened and results are presented in following **Chapter III**.

## Résumé

Dans ce chapitre, nous avons fourni une caractérisation biochimique et une cinétique complète d'une Gal $f$ -ase recombinante sous-clonée. Le travail comprenait la préparation d'une construction plasmidique appropriée pET-28a(+) et la production d'une enzyme fusionnée 6xHis-tag unique dans le système d'expression procaryote de *E. coli* Rosetta. L'enzyme soluble purifiée a ensuite été soigneusement caractérisée et soumise à des essais de cristallisation.

Cette enzyme a été obtenue avec un rendement raisonnable de 0,5 mg/litre de culture et une activité optimale à pH 4,5 et à 60°C. Elle a montré une spécificité de substrat exclusivement vers pNP- $\beta$ -D-Gal $f$ , donnant la meilleure valeur  $K_M$  de 250  $\mu$ M. Elle est inhibée par l'analogie thio structurel du substrat ( $IC_{50}$  = 0,8 mM), qui s'est avéré être un inhibiteur compétitif modéré ( $K_I$  = 1 mM). Bien que, comme nous l'avons mentionné précédemment, ce rendement était raisonnable pour les études de caractérisation, il constitue un facteur limitant pour les essais de criblage par cristallisation. La surexpression insuffisante observée et par conséquent les faibles rendements obtenus, sont des difficultés pour trouver des conditions de cristallisation prometteuses. Notamment lorsqu'il s'agit de cristaux de protéines à rendement élevé convenant à la diffraction des rayons X.

Jusqu'à présent, seules quelques furanosidases ont été impliquées dans les approches chimioenzymatiques. Il était donc très intéressant d'explorer et d'évaluer la capacité synthétique potentielle de cette enzyme à agir comme une thioligase efficace après une mutagenèse dirigée par le site. Par conséquent, son potentiel biocatalyseur pour la synthèse de divers furanothioglycosides a été examiné et les résultats sont présentés au **Chapitre III** suivant.

## Bibliography

1. Voet, D. & Voet, J. G. *Biochemistry*. (John Wiley & Sons, 2011).
2. Hartley, J. L. Cloning technologies for protein expression and purification. *Curr. Opin. Biotechnol.* **17**, 359–366 (2006).
3. Glick, B. R., Pasternak, J. J. & Patten, C. L. *Molecular biotechnology: principles and applications of recombinant DNA*. (ASM Press, 2010).
4. Weiner, M. P. & Slatko, B. E. Kits and their unique role in molecular biology: a brief retrospective. *BioTechniques* **44**, 701–704 (2008).
5. Brown, T. A. *Genomes 3*. (Garland Science Pub, 2007).
6. Novagen pET system book.pdf.
7. Structural Genomics Consortium *et al.* Protein production and purification. *Nat. Methods* **5**, 135–146 (2008).
8. Rosano, G. L. & Ceccarelli, E. A. Recombinant protein expression in *Escherichia coli*: Advances and challenges. *Front. Microbiol.* **5**, 1–17 (2014).
9. Welch, R. A. The Genus *Escherichia* in *The Prokaryotes* 60–71 (Springer New York, 2006).
10. Daegelen, P., Studier, F. W., Lenski, R. E., Cure, S. & Kim, J. F. Tracing Ancestors and Relatives of *Escherichia coli* B, and the Derivation of B Strains REL606 and BL21(DE3). *J. Mol. Biol.* **394**, 634–643 (2009).
11. Shiloach, J. & Fass, R. Growing *E. coli* to high cell density—A historical perspective on method development. *Biotechnol. Adv.* **23**, 345–357 (2005).
12. Sezonov, G., Joseleau-Petit, D. & D’Ari, R. *Escherichia coli* Physiology in Luria-Bertani Broth. *J. Bacteriol.* **189**, 8746–8749 (2007).
13. Clark, D. P. & Pazdernik, N. J. *Biotechnology*. (Elsevier/AP Cell, Academic Cell is an imprint of Elsevier, 2016).
14. Joseph, B. C., Pichaimuthu, S. & Srimeenakshi, S. An Overview of the Parameters for Recombinant Protein Expression in *Escherichia coli*. *J. Cell Sci. Ther.* **06**, (2015).
15. Vethanayagam, J. G. & Flower, A. M. [No title found]. *Microb. Cell Factories* **4**, 3 (2005).
16. Matsunaga, E. *et al.* Draft Genome Sequence of *Streptomyces* sp. JHA19, a Strain That Possesses  $\beta$ -D-Galactofuranosidase Activity. *Genome Announc.* **2**.
17. Matsunaga, E. *et al.* Identification and characterization of a novel galactofuranose-specific  $\beta$ -D-galactofuranosidase from *Streptomyces* species. *PLoS ONE* **10**, 1–16 (2015).

18. Davis, G. D., Elisee, C., Newham, D. M. & Harrison, R. G. New fusion protein systems designed to give soluble expression in *Escherichia coli*. *Biotechnol. Bioeng.* **65**, 382–8 (1999).
19. Winzerling, J. J., Berna, P. & Porath, J. How to use immobilized metal ion affinity chromatography. *Methods* **4**, 4–13 (1992).
20. Gaberc-porekar, V. & Menart, V. Perspectives of immobilized-metal affinity chromatography. (2001).
21. Briand, L. *et al.* A self-inducible heterologous protein expression system in *Escherichia coli*. *Sci. Rep.* **6**, 33037 (2016).
22. Wan, Q. *et al.* Direct determination of protonation states and visualization of hydrogen bonding in a glycoside hydrolase with neutron crystallography. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 12384–12389 (2015).
23. Alves, N. J., Turner, K. B., Medintz, I. L. & Walper, S. A. Protecting enzymatic function through directed packaging into bacterial outer membrane vesicles. *Sci. Rep.* **6**, (2016).
24. Euzen, R. *et al.* A Chemoenzymatic Approach for the Synthesis of Unnatural Disaccharides Containing D-Galacto- or D-Fucofuranosides. *Eur. J. Org. Chem.* **2005**, 4860–4869 (2005).
25. Rémond, C., Plantier-Royon, R., Aubry, N. & O'Donohue, M. J. An original chemoenzymatic route for the synthesis of  $\beta$ -D-galactofuranosides using an  $\alpha$ -L-arabinofuranosidase. *Carbohydr. Res.* **340**, 637–644 (2005).
26. Matsunaga, E. *et al.* Characterization of a PA14 domain-containing galactofuranose-specific  $\beta$ -D-galactofuranosidase from *Streptomyces* sp. *Biosci. Biotechnol. Biochem.* **81**, 1314–1319 (2017).
27. Lopez, G., Daniellou, R., O'Donohue, M., Ferrières, V. & Nugier-Chauvin, C. Thioimidoyl furanosides as first inhibitors of the  $\alpha$ -L-arabinofuranosidase AbfD3. *Bioorg. Med. Chem. Lett.* **17**, (2007).
28. Repetto, E., Marino, C. & Varela, O. Synthesis of the (1→6)-linked thiodisaccharide of galactofuranose: Inhibitory activity against a  $\beta$ -galactofuranosidase. *Bioorg. Med. Chem.* **21**, 3327–3333 (2013).

## **CHAPTER III**

### **Tailoring galactofuranosidase activity – potential platform for new applications**

<b>I. Glycosidase engineering</b> .....	<b>69</b>
1. Thioglycosides – versatile tools in glycobiology .....	69
2. Thioglycoligases.....	72
<b>II. Design and production of Galf-ase E464X mutants</b> .....	<b>74</b>
1. Selection and creation of mutagenesis points .....	74
2. Overexpression and purification of Galf-ase E464X mutants.....	76
3. Galf-ase E464X mutants kinetic parameters determination – Identification of the acid/base catalyst.....	<b>Erreur ! Signet non défini.</b>
<b>III. Application of Galf-ase E464X mutants in thioglycosides synthesis</b> .....	<b>79</b>
1. Selection of appropriate Galf-ase E464X mutant.....	79
2. Galf-ase E464Q mutant catalysed synthesis of thiogalactofuranosides .....	80
<b>IV. Conclusion</b> .....	<b>84</b>
<b>Résumé</b> .....	<b>86</b>
<b>Bibliography</b> .....	<b>86</b>

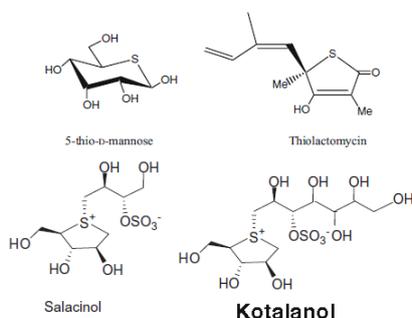
## I. Glycosidase engineering

### 1. Thioglycosides – versatile tools in glycobiology

Thiosugars are a class of sulfur-containing carbohydrates with either having a sulfur as heteroatom, replacing the endocyclic oxygen, or as a replacement of anomeric (glycosidic) oxygen or any of the other hydroxyl groups' oxygen. Structurally, they are oxygen analogues but demonstrate specific biological activities and physiochemical properties that are directly attributed by the replacement of the oxygen atom by sulfur. The differences of those two atoms, sharing the same group in periodic table of the elements, are evident when sulfur is incorporated into the sugar ring. It is larger, more polarizable than oxygen and therefore the carbon-sulfur bond is longer (C-S  $\sim 1,8 \text{ \AA}$ ; C-O  $\sim 1,4 \text{ \AA}$ ), less polar (electronegative) than the bond between carbon and oxygen, giving smaller endocyclic C-S-C angle than the corresponding oxygen-containing analogue.<sup>1</sup>

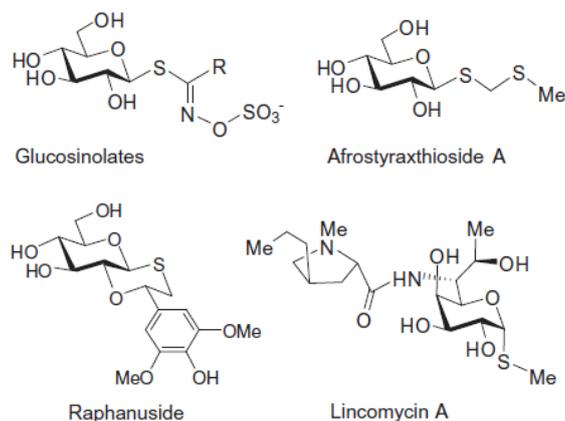
Only few naturally occurring carbohydrate components containing sulfur have been isolated from plants, marine organisms or terrestrial sources although sulphur is an essential element found ubiquitously in all living systems. There are only two classes of natural thiosugars, thioanhydrosugars and thioglycosides.<sup>2</sup>

Thioanhydrosugars occur scarcely in nature and generally all share sulfur-containing heterocycle structure, sometimes also as thioanhydro-bridge. The related molecules include the first naturally occurring free thiosugar, 5-thio-D-mannose, isolated in 1987 from the marine sponge *Clathria pyramida*, as well as salacinol and kotalanol, isolated from *Salacia reticulata*, plant used in Indian Ayurvedic traditional medicine for the treatment of diabetes, thiolactomycin, which has antiparasitic and antibacterial activity and tagetitoxin, bacterial phytotoxin produced by the plant pathogenic bacteria *Pseudomonas syringae* pv. *tagetis* that is potentially useful herbicide (**Figure III-1**).<sup>2,3</sup>



**Figure III-1.** Structures of some naturally occurring thiosugars sharing sulfur-containing heterocycle structure.

Thioglycosides are glycoconjugates in which a sulfur atom has replaced the glycosidic oxygen atom and are the most abundant family of naturally occurring thiosugars. Through sulfur atom they make thioglycosidic linkage with vast array of compounds such as thiodisaccharides and thioglycoproteins. The most of thioglycosides belong to the family of glucosinolates, a well-defined group of secondary plant metabolites with a characteristic structure and biochemistry. There is more than 130 characterized glucosinolates that largely originate from *Brassicaceae* family and just four species of plants, *Barbarea vulgaris*, *Arabidopsis thaliana*, *Eruca sativa* and *Isatis tinctoria*. In plants, as secondary metabolites, they serve as the defence molecules against herbivores, insects or microbial attacks. Upon tissue disruption, glucosinolates are rapidly hydrolysed by myrosinase ( $\beta$ -thioglycosidase) to isothiocyanates, nitriles and thiocyanates, which are repellents or sometimes toxic to predators but contribute to the highly distinctive flavour and odor of these plants which are used as flavour components in aliment.<sup>4</sup> Some of the representatives are sinigrin, found in the seeds of black mustard and the other thioglycosides include afrostyraxthioside A from *Afrostyrax lepidophyllus*, raphanuside from the seeds of *Raphanus sativus* and lincomycin A and its derivatives which are used as thioglycoside-based antibiotics (**Figure III-2.**)<sup>3</sup>



**Figure III-2.** Structures of some naturally occurring thioglycosides.

Generally, thioglycosides have unique physicochemical properties that combine their water solubility, as glycomolecules and structural analogues of widespread *O*-glycosides, and therefore their bioavailability, tolerated by most biological systems because *S*-glycosidic linkage is more resistant towards enzymatic cleavage or any other targeted acid/base degradation. The relevance of thioglycosides is evident in synthetic and medicinal chemistry where they become useful tools: i) in organic synthesis as routinely applied glycosyl donors for various glycosylation reactions, ii) as glycosidase inhibitors,

iii) metabolically stable substrates for X-ray crystallographic structural and mechanistic studies or purification probes for glycosidases or iv) as a scaffolds for carbohydrate based-therapeutics and vaccines.<sup>2,5</sup>

Since the synthesis of 5-thio-D-xylopyranose, the first sulfur-containing sugar analogue in 1961, numerous synthetic routes for thioglycosides and their derivatives have been developed and advanced, mostly involving S<sub>N</sub>2-type reactions between halide glycosyl donors and activated thiol acceptors.<sup>1,6</sup> In addition, alternative and ecologically acceptable strategy has been introduced using *S*-glycosyltransferases (*S*-GTs), the only enzymes that naturally catalyse the formation of the thioglycosidic linkages. These enzymes have been successfully purified from plant extracts and genes encoding identified. Among studied, *A. thaliana S*-GT, UGT74B1 is one of the only examples of natural *S*-transferases that has been extensively studied and characterized as versatile biocatalyst for *in vitro* synthesis of thioglycosides.<sup>7,8</sup>

A wide variety of methods has been explored and developed for the preparation of thioglycopyranosides exclusively and are thoroughly documented. However, literature reports on the synthesis of their thioglycofuranoside counterparts are scarce regarding either organic synthesis methods or chemoenzymatic approaches.

The works that include thioglycofuranoside synthesis, *Araf* and *Galf* as furanoses in particular, has been mainly developed by the laboratory teams of *Rosa M. de Lederkremer & Carla Marino* (Argentina), *Todd L. Lowary* (Canada) and *Vincent Ferrières* (France) who have long interest in glycobiology related to these furanoses. They developed acknowledged synthetic routes for preparation of substrates and inhibitors, specifically for *Galf*-ase, including alkyl, thioheteroaryl and thiomidoyl β-D-galactofuranosides or furanothiodisaccharides.<sup>9-11</sup>

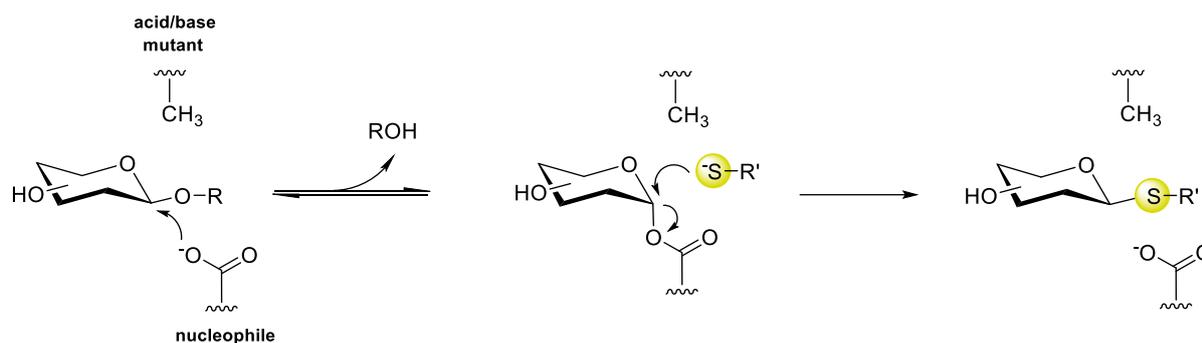
Since existence of natural *S*-furanoglycosyltransferase has not been evidenced so far, recently a novel methodology for the synthesis of *S*-furanoglycosidic linkages that explores furanothioglycoligases, retaining GHs that lack the catalytic acid/base residue, has been presented.<sup>12</sup>

## 2. Thioglycoligases

An alternative to classical organic chemistry synthesis methods of specific glycoconjugates is the use of enzymes. In glycochemistry, glycosidic bond formation can be accomplished using two groups of enzymes, GTs that naturally perform synthesis of glycosidic bonds, and GHs. In carbohydrate synthesis by GHs, currently, there are two different approaches. One employs the use of preferential transglycosylases, GHs that in addition to hydrolysis, also possess the ability to catalyse the formation of glycosides either by exploiting their capacity to perform reverse hydrolysis in high substrate concentrations (thermodynamic control approach) or by employing activated substrates (kinetic control approach) and other is implementing enzyme engineering to create mutated GHs in order to enhance their synthesis activity.<sup>13,14</sup>

Based on mechanistic grounds, two key active site amino acid residues, the catalytic nucleophile and the catalytic acid/base, have been specifically replaced thus two classes of engineered glycosidases have been developed. Generally, they can be divided on the glycosynthases that are able to catalyse *O*-glycosidic bond formation (catalytic nucleophile residue mutants) and thioglycoligases (catalytic acid/base or/and nucleophile residue mutants) that are able to catalyse *S*-glycosidic linkage formation.<sup>15,16</sup>

This engineering strategy was first introduced in 1998 and since then it has generated many glycosynthases that contributed to the high-yield enzymatic synthesis of *O*-glycoconjugates. The potential of this strategy was further extended to chemoenzymatic synthesis of *S*-glycosidic linkages.<sup>17-19</sup> According to the double displacement mechanism in retaining GHs, use of a donor sugar with an activated leaving group and accompanied by a highly nucleophilic acceptor compensates the missing acid/base catalyst and results in thioglycosidic linkage formation (**Figure III-3**). For this purpose, first thioglycoligases were generated using alanine acid/base mutants of two retaining  $\beta$ -glycosidases,  $\beta$ -glucosidase Abg from *Agrobacterium sp.* and the  $\beta$ -mannosidase Man2A from *Cellulomonas fimi*. Use of dinitrophenyl glycosides with thiol acceptors yielded around 70% of corresponding thioglycosides.<sup>15,20</sup>



**Figure III-3.** Proposed general mechanism for retaining thioglycoligase.

Later, several other thioglycoligases have been generated in the same manner from diverse retaining  $\alpha$ - and  $\beta$ -glycosidases and the choice of optimal acid/base substitution residue as well as use of activated glycoside donors and thiol acceptor was further explored.<sup>17</sup> In addition first thioglycosynthases, double-mutant of  $\beta$ -glucosidase Abg from *Agrobacterium sp.*, which lacks both the catalytic nucleophile and the catalytic acid/base residue, efficiently catalysed thiodisaccharide formation from glycosyl fluoride donor and thiosugar acceptors.<sup>21</sup>

Thioglycoligases and thioglycosynthases represent the only reliable enzymatic pathway to yield thioglycosides, notably thioglycopyranosides, using glycosyl hydrolases. In general, this approach demonstrated to be valuable for the specific thiooligosaccharides and thioglycoconjugates synthesis. Hence it was extended to date solely example of furanothioglycoligase obtained from  $\alpha$ -L-arabinofuranosidase Araf51 from *Clostridium thermocellum*.

The thioglycoligases potency of the Araf51 E173A mutant was demonstrated through moderate to high yields (50-98%) preparation of several thioaryl glycosides and one arabinofuranosylcluster, by using 1-thioimidoyl arabinofuranoside as an efficient glycosyl donor.<sup>22,23</sup>

Although, the gene of the first  $\beta$ -D-galactofuranosidase has been identified, isolated and expressed in 2015, but the use of it as galactofuranothioliase has not been described yet for the synthesis of S-galactofuranosides. Interestingly CtAraf51, an  $\alpha$ -L-arabinofuranosidase from *Ruminiclostridium thermocellum*, was artificially converted into galactofuranosidase and finally into thioglycoligase, after several mutagenesis steps, thus allowing to improve catalytic efficiency of this enzyme towards pNP-Galf for the synthesis of various S-galactofuranoconjugates.<sup>24</sup>

## II. Design and production of Galf-ase E464X mutants

### 1. Selection and creation of mutagenesis points

Locating and identification of catalytic residues at  $\beta$ -D-Galf-ase active site was essential to proceed towards generation of novel thioglycoligase mutants. To date, no Galf-ase crystals with ligands or substrates is available nor resolved crystal structure reported, thus hindering identification of the catalytic residues. Therefore, prediction of catalytic residues was done *in silico* by Galf-ase amino acid sequence analysis and alignment with suitable homologous sequences of other characterized enzymes.<sup>25</sup>

Since the deduced Galf-ase amino acid sequence analysis, as previously reported<sup>26</sup>, revealed that it belongs to the GH2 family and its substrate specificity was exclusively towards galactofuranose<sup>27</sup>, the closest homologous enzymes to Galf-ase were searched among  $\beta$ -galactosidases from GH2 family with determined crystal structure. The selected homologous amino acid sequences from *Bacillus circulans*, *Bacteroides fragilis* and *Bacteroides vulgatus*, sharing between 24% and 26% identities, were analysed along with the Galf-ase amino acid sequence from *Streptomyces sp.* The sequence alignment revealed highly conserved glutamic acid residues in all of the aligned sequences. Two positive matches at position 464, enabled to identify glutamic acid residue (E464) as catalytic acid/base, and at position 573, to identify the second glutamic acid residue (E573) as catalytic nucleophile (**Figure III-4.; Appendix 7.1.1. Figure VII-1.**).



**Figure III-4.** Alignment of partial ORF Galf-ase (*Streptomyces spp.*) amino acid sequence and corresponding regions of its homologs. The conserved E464 and E573 amino acid residues are indicated with asterisks.

Identified glutamic acid residues were expected to be the catalytic ones. As an extension to this to identification, their actual functional role in the catalytic site had to be examined

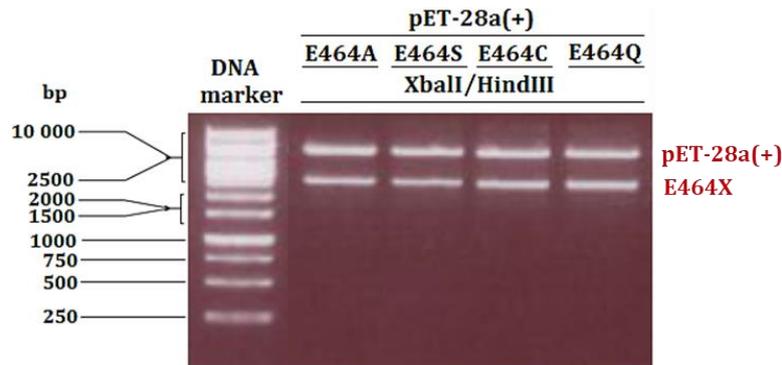
experimentally, by the appropriate site-specific mutations. Therefore, using site-directed mutagenesis method, the acid/base catalytic residue (E464) has been specifically substituted to alanine (E464A), serine (E464S), cysteine (E464C) and glutamine (E464Q) in four separate mutagenesis reactions. In order to minimally interfere in the *Galf*-ase gene nucleotide sequence, the preservation of the glutamic acid GAG codon is done by taking the advantage of the plasticity of the genetic code where one amino acid can be encoded by multiple codons. Single point substitution of a minimum number of bases was introduced, which lead to a change in desired amino acids (**Table III-1.**)

**Table III-1.** Codon changes and respective amino acid changes introduced at 464 position at wild-type *Galf*-ase amino acid sequence by site-directed mutagenesis.

Mutant variant	Codon change	Amino acid change
E464 <b>A</b>	<b>GAG</b> → <b>GCG</b>	Glu (E) → Ala (A)
E464 <b>S</b>	<b>GAG</b> → <b>AGC</b>	Glu (E) → Ser (S)
E464 <b>C</b>	<b>GAG</b> → <b>TGC</b>	Glu (E) → Cys (C)
E464 <b>Q</b>	<b>GAG</b> → <b>CAG</b>	Glu (E) → Gln (Q)

The acid/base carboxylic acid residue has been replaced by the amino acids that have the residues of specific interest, briefly have no negative charge and subsequently the hydrolytic mechanism rate should be affected in favour of thioglycoligase activity while the nucleophilic residue remain intact, and in addition to explore their potential role to function as neolectins.

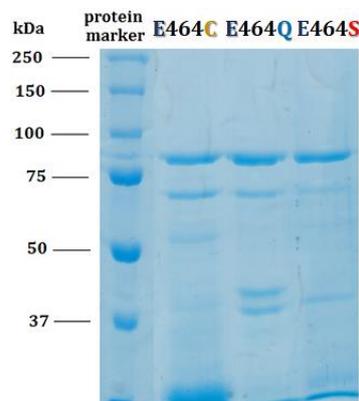
Plasmid construct pET-28a(+), containing wild type *Galf*-ase gene inset, served as a template to generate desired E464X *Galf*-ase mutants. The four newly obtained constructs, encoding *Galf*-ase E464X mutants, were checked for integrity by double digestion with XbaI and HindIII restriction enzymes which gave two strong band signals at 5.3 kb and 2.3 kb corresponding to expected pET-28a(+) vector and *Galf*-ase gene size (**Figure III-5.**). The all four mutagenized plasmid constructs were sequenced to ensure that the appropriate mutations were introduced and that the rest of the *Galf*-ase gene sequence remained intact.



**Figure III-5.** Confirmation of pET-28a(+) plasmid construct integrity, after E464X mutant variants generation, on agarose-gel electrophoresis by EcoRI and HindIII restriction enzyme analysis.

## 2. Overexpression and purification of Galf-ase E464X mutants

After successful site-directed mutagenesis, plasmid constructs were transformed into *E. coli* Rosetta™ (DE3) (Novagen®) expression system and E464X mutants were expressed under similar conditions previously described for the wild-type Galf-ase. Three mutant Galf-ase variants, E464S, E464C and E464Q respectively, were expressed as soluble recombinant proteins bearing only *N*-terminal peptide 6xHis-tag, necessary for selective protein purification using NTA-Ni<sup>2+</sup> affinity chromatography. The E464A was the solely mutant variant which was not obtained in the soluble form indicating that the mutation probably affected the folding and stability and caused the protein to form insoluble aggregates. The corresponding recombinant mutants were provided in modest yields typically ranging from 0.13 mg to 0.3 mg of protein per culture litre for E464S, E464C and E464Q respectively. The purity of all enzymes was determined by SDS-PAGE gel, showing a band with apparent molecular weight of 90.3 kDa, and was satisfactory enough to further use them for biochemical characterization (**Figure III-6.**).

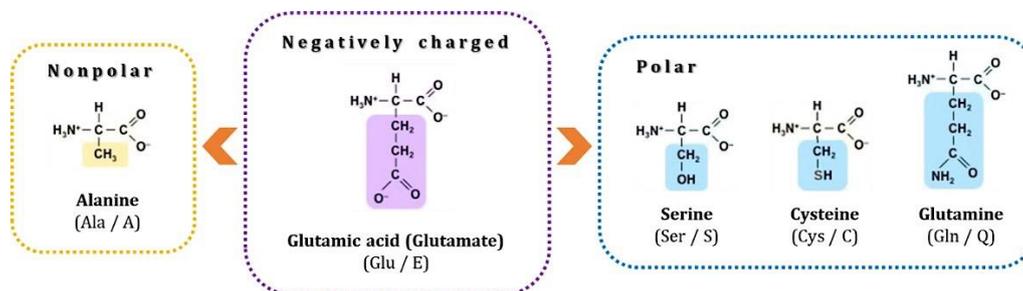


**Figure III-6.** SDS-PAGE analysis of E464X Galf-ase mutant variants after NTA-Ni<sup>2+</sup> chromatography purification.

### 3. Galf-ase E464X mutants kinetic parameters determination – Identification of the acid/base catalyst

To assign catalytic acid/base function to selected individual glutamic acid residue at 464 position, the three mutant enzymes were assayed for kinetic parameters in Michaelis-Menten assay. As a result of single mutation in catalytic acid/base residue of retaining GHs, it is expected to exhibit reduced but detectable hydrolytic activity of attributed substrate. This reflected the choice of amino acids which were used as mutants. Guided by previous literature available examples for the thioglycoligases, where alanine and glutamine were found to be the best options, we also extended choice to serine and cysteine.

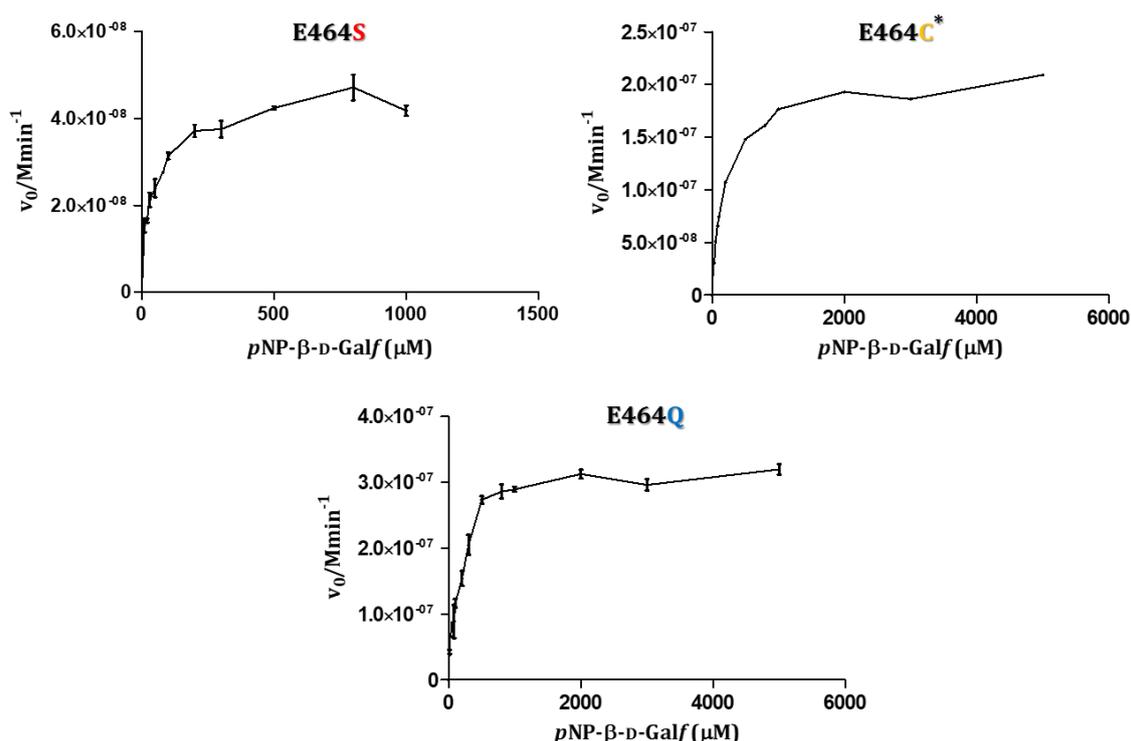
This enzyme engineering was implemented to simultaneously attenuate hydrolysis and to preserve general architecture of the active site, mainly referring to specificity for hydrolysis substrate (*pNP-β-D-Galf*) and that the vacant cavity created by mutation of the acid/base does not considerably increase and disrupt the formation of hydrogen bond between the two catalytic residues. Therefore, introduced amino acids, E464S, E464C and E464Q, have polar that is polarizable groups, for cysteine respectively, that can participate in hydrogen bond formation, with exception of E464A, which has nonpolar side chain, but are not strong enough Brønsted acids/bases to efficiently serve as activators during deglycosylation step compared to glutamic acid. Thus E464Q is the amide analogue to substituted glutamic acid that does not vary in the aliphatic chain length, whereas serine and cysteine have shorter side chain length and the latter has sulfur as larger analogue of oxygen (**Figure III-7.**).



**Figure III-7.** Structural formulas of amino acids selected for the generation of E464X Galf-ase mutant variants.

The capacity of the Galf-ase mutant variants E464S, E464C and E464Q, successfully obtained as soluble and purified fractions, to hydrolyse natural substrate *pNP-β-D-Galf* and to characterize their kinetics and the hydrolysis rate-limiting steps, the Michaelis-

Menten assay was performed under the similar conditions used for wild type enzyme. All the mutants gave typical Michaelis-Menten saturation curves (**Figure III-8.**) and exhibited slightly decreased  $K_M$  value than the one determined for wild-type, apart from E464S which stands out by 7 fold decrease in the  $K_M$  value (**Table III-2.**). Altogether obtained values are in the same order of magnitude. Indeed, the replacement of glutamic acid residue to serine, cysteine and glutamine has resulted in a 2 000 orders of magnitude decrease in the hydrolytic activity compared to the wild-type enzyme, confirming that this residue is essential for catalysis.



**Figure III-8.** Michaelis-Menten plot of  $p\text{NP-}\beta\text{-D-Galf}$  hydrolysis reaction catalysed by Galf-ase E464S, E464C and E464Q mutant variants. Mean values and SD error bars were calculated from three independent experiments (\*except E464C) and from a single protein preparation.

The deglycosylation step in hydrolysis is the rate-limiting and therefore more sensitive to acid/base replacement. Since the general-acid residue has been removed, deglycosylation remains very slow because there is no basic residue to activate the nucleophilic water molecule. This consequently reflected on the significant decrease in catalytic efficiency by a factor of 3 000-1 400 times for E464S and E464C mutants respectively, and 56 times lower for E464Q mutant. This kinetic observations are in accordance with the other data obtained in similar assays used for the identification of the catalytic residues in retaining and several inverting GHs or to probe their thioglycoligase catalysis.<sup>28-30</sup>

**Table III-2.** Comparison of Michaelis-Menten kinetics constants for hydrolysis of  $\beta$ -D-Galf by E464X Galf-ases mutant variants.

Enzyme	$K_M$ (mM)	$k_{cat}$ ( $\text{min}^{-1}$ )	$k_{cat}/K_M$ ( $\text{min}^{-1}\cdot\text{mM}^{-1}$ )
E464	0.25	213	852
E464 <b>S</b>	0.036	>0.01	0.27
E464 <b>C</b>	0.17	0.1	0.58
E464 <b>Q</b>	0.166	2.5	15

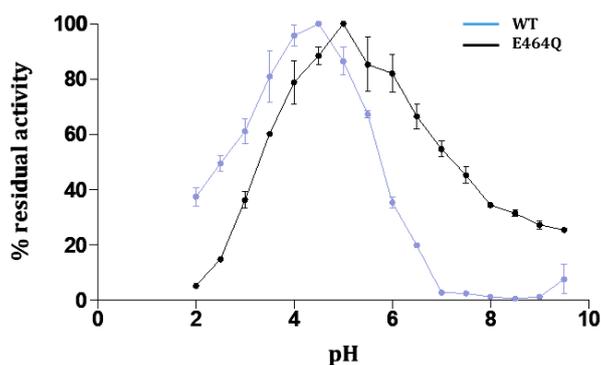
As expected this mutations served a double purpose, to confirm proposed residue E464 as the catalytic one and to alter hydrolytic activity. In this site-directed mutagenesis approach, a glycosidase, Galf-ase, has been modified into neoelectin as hydrolytic activity was attenuated and recognition of synthetic substrate has been conserved. The obtained three E464X mutant variants are therefore promising candidates to be tested as carbohydrate-binding proteins for Galf-specific imaging, presented in **Chapter IV**. Also, this engineering approach simultaneously allowed to explore both, neoelectin and thioglycoligase potential of generated mutants.

### III. Application of Galf-ase E464X mutants in thioglycosides synthesis

#### 1. Selection of appropriate Galf-ase E464X mutant

Generally, all the mutants showed similar diminution changes in hydrolytic activity, however the E464Q featured only a minor change in the  $K_M$  value, 166  $\mu\text{M}$  compared to 250  $\mu\text{M}$ , from the wild type, and preserved relatively good catalytic efficiency. The advantage of this mutant was, although low yet sufficient quantity of 0.3 mg/L and taken together with the more likely theoretic possibility that the amide could provide better assistance to the deglycosylation step, possibly *via* hydrogen bonding with the incoming thiol acceptor, therefore this mutant variant was selected to perform preliminary thioglycoside synthesis studies.

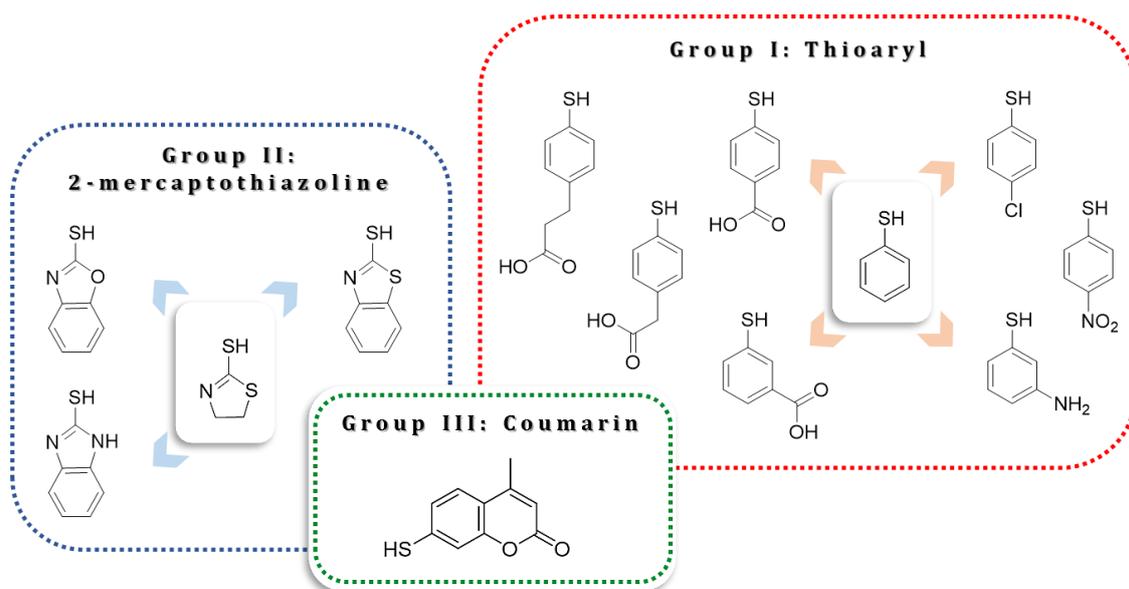
In addition, the pH profile of the glutamine residue was examined (**Figure III-9.**). As expected, the pH dependence of the mutant enzyme is a typical bell-shaped curve, however it has more extended basic shift which is reflecting the gradual reduction of activity at high pH values. This suggests that the protonated group, the acid/base catalyst, has been removed and the optimal pH has moderately been changed from the pH 4,5 of the wild-type enzyme and is slightly higher for about 0.5 pH unit for E464Q mutant.



**Figure III-9.** Influence of pH on E464Q Galf-ase mutant variant hydrolytic activity towards *pNP-β-D-Galf*.

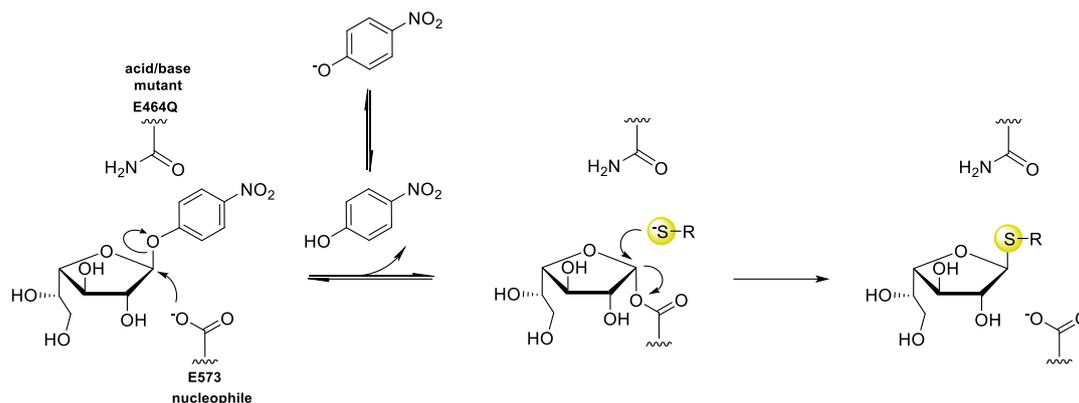
## 2. Galf-ase E464Q mutant catalysed synthesis of thiogalactofuranosides

Having characterized the Galf-ase E464X mutant variants, insight into synthetic utility of chosen E464Q mutant was undertaken. E464Q biocatalyzed-thioligation reactions were carried out in the presence of donor sugar and various commercially available thiol acceptors, 13 in total, which are arranged into three groups, for ease of comparison, based on structural similarities of the thiol carrying moiety as follows: I) thioaryl, II) 2-mercaptothiazoline and III) coumarin (**Figure III-10.**). The selected acceptors had two important characteristics for thioglycoligase approach, their aromatic moiety is known to be a good mimic of sugar entities and they could be easily visualized by standard analytical methods. To ensure synthetic activity, the reactions were first performed at the analytical scale for screening setup, and then if proved useful, selected and scaled up for preparative scale.



**Figure III-10.** Structural formulas of thiol acceptors used for E464Q Galf-ase catalyzed furanothioglycoligase reactions.

The donor sugar, already an actual substrate for the enzyme, which is conveniently activated at anomeric position with *p*NP group, relatively good leaving group that in synergy with an array of aglycone thiol acceptors, strong nucleophiles, can undergo useful transfer to form thioglycoconjugate in the absence of base catalytic assistance (**Figure III-11**). This strategy was previously shown to be successfully applicable to Araf51 arabinofuranosidase, closest substrate analogue to galactofuranose enzyme and the first furanosidase efficiently converted into thioglycoligase.

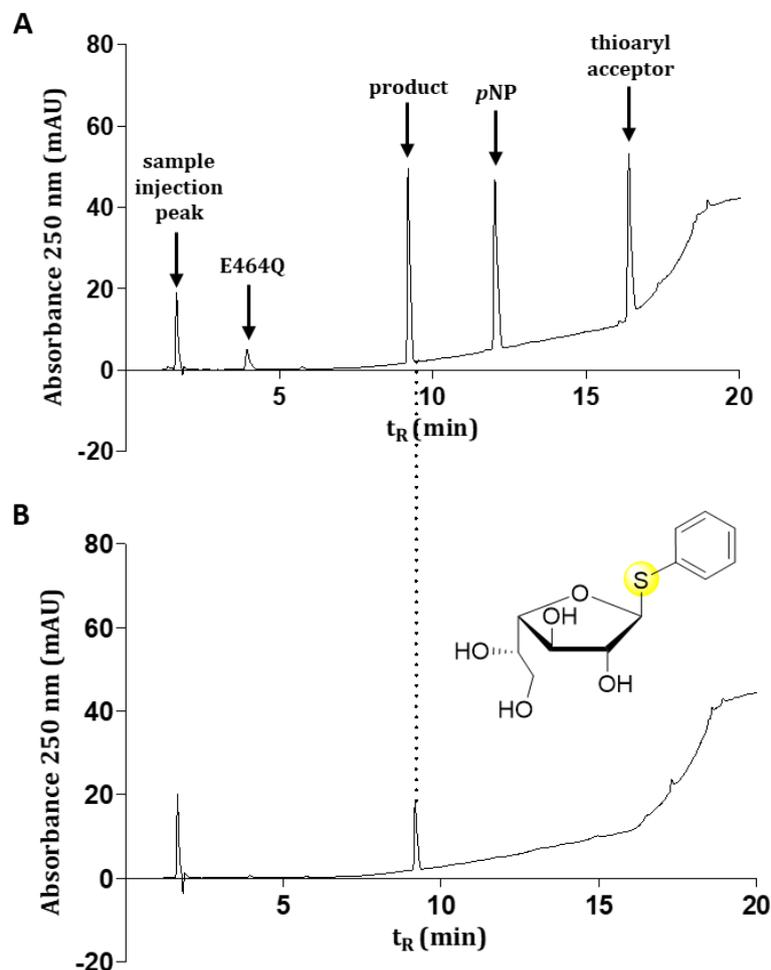


**Figure III-11.** Proposed furanothioglycosylation mechanism catalyzed by E464Q Galf-ase mutant variant.

To screen E464Q mutant ability to catalyze transglycosylation, individual analytical-scale reactions were performed starting from *p*NP-β-D-Gal donor and each of thio acceptors. Firstly, a reaction with thioaryl acceptor was set up as a model for all biocatalyzed coupling. Reaction was left to proceed in buffered (pH 4.5) aqueous solution at the physiological temperature (37 °C), overnight (16 h) to assure complete or highest possible consumption of the donor sugar. Owing the aromatic character of both, released *p*NP aglycone from the donor sugar, and the acceptor, substrate consumption progress could be determined, as well as the formation of the thioglycoligase product. Upon completion, the formation of the thioglycoligase product was analyzed by high performance thin layer chromatography (HPLC), and the novel compound could be detected by UV-absorbance spectrum compared with free uncoupled thiol acceptor.

In the present model reaction with thioaryl acceptor, based upon the HPLC chromatogram, identification of reactants peaks was done by comparison of retention times ( $t_R$ ) and UV absorbance spectra to individual control compounds (**Figure III-12. A**). This allowed identification of formed thioglycoside product, which after chromatographic

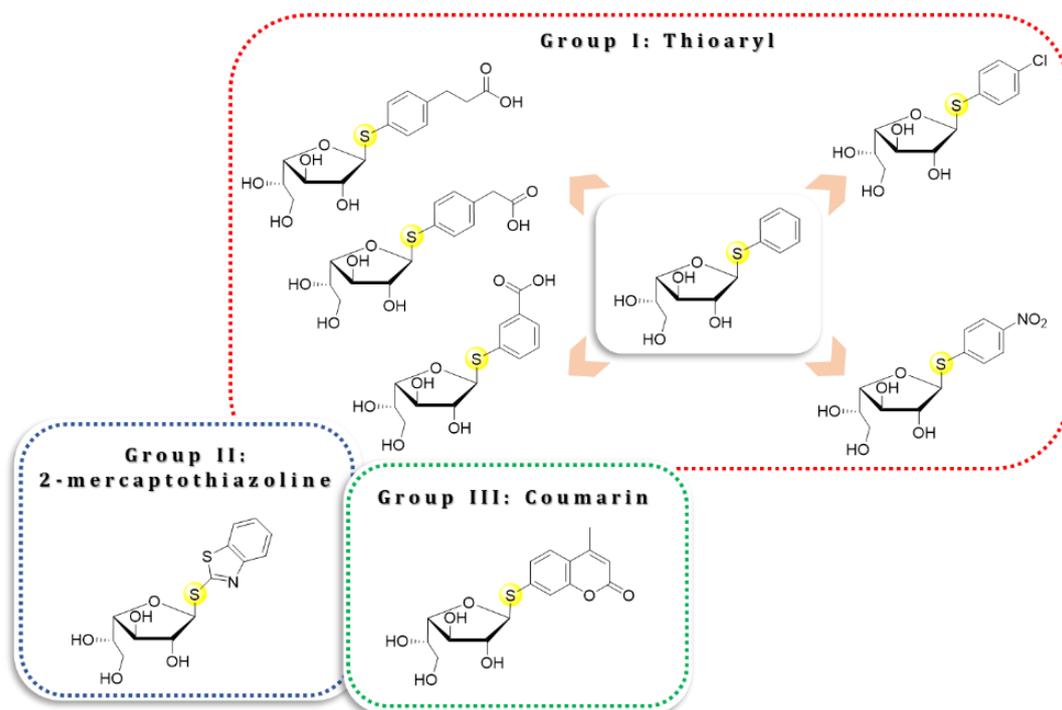
purification was isolated and determined  $t_R$  of 9.19 min was aligned to one observed in the initial reaction (**Figure III-12. B**).



**Figure III-12.** HPLC chromatograms of representative thioglycosylation reactions by E464Q. **A)** Chromatogram of crude thioglycosylation reaction with thioaryl acceptor. **B)** Chromatogram of purified product fraction from thioglycosylation reaction.

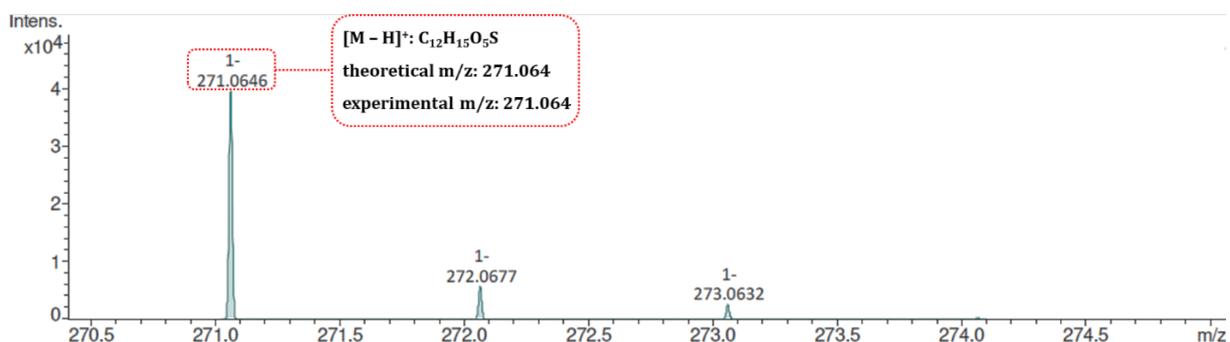
Generally the majority of screened acceptors were able to act as substrates and led to formation of corresponding thioglycoconjugate compounds. The enzyme was able to tolerate variety of thiol acceptors as well as series of different substituents within each group. The only exception for thioaryl group was the 3-aminothiophenol, which did not show reactivity when tested on analytical scale while 4-mercaptobenzoic acid product formation was detected on analytical scale but was not yielded in sufficient quantity to be characterized. As of 2-mercaptothiazole and its benzoxazole and benzimidazole derivatives, the formation of the desired thioglycosides lacked in all three cases, leaving benzothiazole as only active representative substituent of this group. Surprisingly, enzyme was able to catalyse transfer of coumarin aglycone, which by far presents the

largest of all screened thiol aglycones with its two aromatic units, and yield corresponding thioglycoconjugate (**Figure III-13.**).



**Figure III-13.** Structural formulas of synthesized furanothioglycosides.

Reactions with selected acceptors were upscaled and resulting products were qualitatively isolated after one pass by reverse-phase chromatography in appropriate solvent mixture. The high-resolution electrospray ionization mass spectrometry (HRMS-ESI) analysis indicated that the new compounds have the molecular weight of expected thiogalactofuranosides (**Figure III-14., Appendix 7.1.3. Figure VII-3. – Figure VII-10.**).



**Figure III-14.** The measured HRMS-ESI spectrum of thioaryl galactofuranosyl ion in negative mode. The analysed sample is the fraction purified on C18 chromatography resin (eluent solvent mixture: **A**:**B**=6:4; **A**=0.1% formic acid (v/v) in water, **B**=0.1% formic acid (v/v) in acetonitrile) and corresponds to the peak ( $t_R=9.19$  min) detected on HPLC (**Figure III-12. B**).

The isolated products were not of satisfactory purity nor quantity to be unambiguously confirmed by nuclear magnetic resonance (NMR) spectroscopy. It should be noted that compounds were not obtained in quantitative yields and therefore it is not possible to hypothesise about correlation of different parameters, such as  $pK_a$  and nucleophilicity of the thiols or variation of substituent position and electronic properties, which could influence the yields.

To address this issue, it should be noted that in this context it was screen work which allowed us to monitor thiogalactofuranoside product formation and accordingly to attribute thioglycoligase activity to E464Q. Taking into account obtained results, in this preliminary study, this is indeed confirmation that the E464Q mutant is a promising novel furanothioglycoligase and it appears that within the active site it can optimally accommodate the variety of thiol acceptors.

#### IV. Conclusion

To further expand Gal $f$ -ase hydrolytic activity towards novel catalytic activities, the furanothioglycoligase approach was used as a test system. Firstly, the amino acid residues, important for catalysis, were predicted by homology sequence analysis of relevant enzymes, constructed for Gal $f$ -ase from *Bacillus circulans*, *Bacteroides fragilis* and *Bacteroides vulgatus*, respectively. This analysis revealed two conserved glutamic acid residues at position E464 and E573, proposed as the catalytic acid/base and as catalytic nucleophile.

The identification of catalytic acid–base residues in glycosidases is based mainly on the replacement of the putative carboxylic acid with a non-ionizable amino acid, and characterizing the mutant enzymes. Therefore, an array of four mutant variants, E464X containing alanine, serine, cysteine and glutamine residues were constructed by site-directed mutagenesis and introduced into the same pET-28a(+) vector and *E. coli* expression system, as described in previous chapter.

Three mutants E464S, E464C and E464Q, were obtained as soluble 6xHis-tagged enzymes. The Michaelis-Menten kinetic characterization of mutants distinguished E464 as an actual catalytic residue from other residues. As expected, the  $k_{cat}$  values measured for these mutants, in particular for E464S and E464C, were considerably reduced, and were 2 000-fold lower than the wild-type's. This gave focus on one particular mutant,

E464Q which in terms of its activity and quantity was selected as best candidate for thioglycoligation reaction screening and due to amide chain could provide some assistance to thioglycoligase catalysis.

The initial screening for thioglycoligase activity with E464Q was performed on the panel of 13 thiol acceptors, differing at their aryl, mercapto and coumarin moiety. The yielded eight corresponding furanothioglycoside products which were characterized by HPLC and HRMS-ESI.

This, to our knowledge is the first use of natural galactofuranosidase as thioglycoligase, and the second furanosidase that successfully catalyzed furanothioglycoside synthesis. The E464Q mutant variant proved to be active furanothioglycoligase biocatalyst, capable to transfer variety of aromatic thio acceptors to activated sugar donor. Altogether, it showed that  $\beta$ -D-galactofuranosidase from *Streptomyces sp.* can be used as enzyme platform for general concept glycosidase engineering.

## Résumé

Pour étendre d'avantage l'activité hydrolytique de la Gal $\beta$ -ase vers de nouvelles activités catalytiques, l'approche furanothioglycoligase a été utilisée. L'identification des résidus acides-bases catalytiques dans les glycosidases repose principalement sur le remplacement de l'acide carboxylique supposé par un autre acide aminé, et sur la caractérisation des enzymes mutantes. Par conséquent, un tableau de quatre variantes mutantes, E464X contenant des résidus d'alanine, de sérine, de cystéine et de glutamine, a été construit par mutagenèse dirigée et introduit dans le même vecteur pET-28a(+) et le même système d'expression de *E. coli*, comme décrit au chapitre précédent.

Trois mutants E464S, E464C et E464Q, ont été obtenus sous forme d'enzymes solubles marquées 6xHis-tag. La caractérisation cinétique des mutants de Michaelis-Menten a permis de distinguer le E464 en tant que résidu catalytique réel. L'accent a ainsi été mis sur un mutant particulier, le E464Q, qui, en termes d'activité et de quantité, a été sélectionné comme le meilleur candidat pour le criblage par réaction de thioglycoligation et qui, en raison de sa chaîne amide, pourrait apporter une aide à la catalyse thioglycoligase.

Le screening initial de l'activité de la thioglycoligase avec E464Q a été effectué sur le panel de 13 accepteurs de thiols, différents par leur groupement aryle, mercapto et coumarine. Les huit produits de furanothioglycosides correspondants, caractérisés par HPLC et HRMS-ESI, ont été obtenus.

Il s'agit, à notre connaissance, de la première utilisation de la galactofuranosidase naturelle comme thioglycoligase. Le variant mutant E464Q s'est avéré être un biocatalyseur furanothioglycoligase actif, capable de transférer une variété de thioaccepteurs aromatiques au donneur de sucre activé. Dans l'ensemble, il a montré que  $\beta$ -D-galactofuranosidase de *Streptomyces sp.* peut être utilisée comme plateforme enzymatique pour l'ingénierie de la glycosidase de concept général.

## Bibliography

1. Jose G. fernandez-Bolanos, Najim A. L Al-Masoudi, Inés Maya. *Advances in Carbohydrate Chemistry and Biochemistry-Sugar derivatives having sulfur in the ring*. (Elsevier, 2002).
2. Witczak, Z. J. & Culhane, J. M. Thiosugars: new perspectives regarding availability and potential biochemical and medicinal applications. *Appl. Microbiol. Biotechnol.* **69**, 237–244 (2005).
3. Lian, G., Zhang, X. & Yu, B. Thioglycosides in Carbohydrate Research. *Carbohydr. Res.* **403**, 13–22 (2015).
4. Agerbirk, N. & Olsen, C. E. Glucosinolate structures in evolution. *Phytochemistry* **77**, 16–45 (2012).
5. Chen, H.-M. & Withers, S. G. Syntheses of p-nitrophenyl 3- and 4-thio- $\beta$ -D-glycopyranosides. *Carbohydr. Res.* **345**, 2596–2604 (2010).
6. Pachamuthu, K. & Schmidt, R. R. Synthetic Routes to Thiooligosaccharides and Thioglycopeptides. *Chem. Rev.* **106**, 160–187 (2006).
7. Marroun, S. *et al.* UGT74B1 from *Arabidopsis thaliana* as a versatile biocatalyst for the synthesis of desulfoglycosinolates. *Org. Biomol. Chem.* **14**, 6252–6261 (2016).
8. Lafite, P. *et al.* S-glycosyltransferase UGT74B1 can glycosylate both S- and O-acceptors: mechanistic insights through substrate specificity. *Mol. Catal.* **479**, 110631 (2019).
9. Marino, C. *et al.* 1-Thio- $\beta$ -D-galactofuranosides: synthesis and evaluation as  $\beta$ -D-galactofuranosidase inhibitors. *Glycobiology* **8**, 901–904 (1998).
10. Mariño, K. & Marino, C. Synthesis of heteroaryl 1-thio- $\beta$ -D-galactofuranosides and evaluation of their inhibitory activity towards a  $\beta$ -D-galactofuranosidase. **11** (2005).
11. Repetto, E., Marino, C. & Varela, O. Synthesis of the (1→6)-linked thiodisaccharide of galactofuranose: Inhibitory activity against a  $\beta$ -galactofuranosidase. *Bioorg. Med. Chem.* **21**, 3327–3333 (2013).
12. Pennec, A., Daniellou, R., Loyer, P., Nugier-Chauvin, C. & Ferrières, V. Araf51 with improved transglycosylation activities: one engineered biocatalyst for one specific acceptor. *Carbohydr. Res.* **402**, 50–55 (2015).
13. Wang, L.-X. & Huang, W. Enzymatic transglycosylation for glycoconjugate synthesis. *Curr. Opin. Chem. Biol.* **13**, 592–600 (2009).

14. Bissaro, B., Monsan, P., Fauré, R. & O'Donohue, M. J. Glycosynthesis in a waterworld: new insight into the molecular basis of transglycosylation in retaining glycoside hydrolases. *Biochem. J.* **467**, 17–35 (2015).
15. Jahn, M. & Withers, S. G. New Approaches to Enzymatic Oligosaccharide Synthesis: Glycosynthases and Thioglycoligases. *Biocatal. Biotransformation* **21**, 159–166 (2003).
16. Wang, L.-X. Expanding the Repertoire of Glycosynthases. *Chem. Biol.* **16**, 1026–1027 (2009).
17. Danby, P. M. & Withers, S. G. Advances in Enzymatic Glycoside Synthesis. *ACS Chem. Biol.* **11**, 1784–1794 (2016).
18. Bojarová, P. & Křen, V. Glycosidases: a key to tailored carbohydrates. *Trends Biotechnol.* **27**, 199–209 (2009).
19. Hancock, S., Vaughan, M. & Withers, S. Engineering of glycosidases and glycosyltransferases. *Curr. Opin. Chem. Biol.* **10**, 509–519 (2006).
20. Jahn, M., Marles, J., Warren, R. A. J. & Withers, S. G. Thioglycoligases: Mutant Glycosidases for Thioglycoside Synthesis. *Angew. Chem. Int. Ed.* **42**, 352–354 (2003).
21. Jahn, M. *et al.* Thioglycosynthases: double mutant glycosidases that serve as scaffolds for thioglycoside synthesis. *Chem Commun* 274–275 (2004) doi:10.1039/B313155F.
22. Almendros, M. *et al.* Exploring the synthetic potency of the first furanothioglycoligase through original remote activation. *Org. Biomol. Chem.* **9**, 8371 (2011).
23. Mélanie Almendros, Marc François-Heude, Laurent Legentil, & Caroline Nugier-Chauvin, Richard Daniellou and Vincent Ferrières. Chemo-enzymatic synthesis of an original arabinofuranosyl cluster: optimization of the enzymatic conditions. *Arkivoc* **2013**, 123 (2012).
24. Pavic, Q., Pillot, A., Tasseau, O., Legentil, L. & Tranchimand, S. Improvement of the versatility of an arabinofuranosidase against galactofuranose for the synthesis of galactofuranoconjugates. *Org. Biomol. Chem.* **17**, 6799–6808 (2019).
25. Zvelebil, M. J. J. M. & Sternberg, M. J. E. Analysis and prediction of the location of catalytic residues in enzymes. *Protein Eng. Des. Sel.* **2**, 127–138 (1988).
26. Matsunaga, E. *et al.* Identification and Characterization of a Novel Galactofuranose-Specific  $\beta$ -D-Galactofuranosidase from *Streptomyces* Species. *PLOS ONE* **10**, e0137230 (2015).

27. Seničar, M. *et al.* Galactofuranosidase from JHA 19 *Streptomyces* sp.: subcloning and biochemical characterization. *Carbohydr. Res.* **480**, 35–41 (2019).
28. Shallom, D. *et al.* 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 (2002).
29. Shallom, D. *et al.* Biochemical Characterization and Identification of the Catalytic Residues of a Family 43  $\beta$ -D-Xylosidase from *Geobacillus stearothermophilus* T-6 †. *Biochemistry* **44**, 387–397 (2005).
30. McGrath, C. E., Vuong, T. V. & Wilson, D. B. Site-directed mutagenesis to probe catalysis by a *Thermobifida fusca* -1,3-glucanase (Lam81A). *Protein Eng. Des. Sel.* **22**, 375–382 (2009).

## **CHAPTER IV**

### **Galf-specific & lanthanide (III) complex loaded nanoparticles for NIR imaging – concept validation**

<b>I. Engineering neolectins from glycosidases.....</b>	<b>91</b>
1. Application in biotechnology, diagnostic and imaging .....	91
<b>II. Construction of luminescent nanoparticle scaffold for neolectin      bioconjugation &amp; imaging applications.....</b>	<b>93</b>
1. Construction and characterization of luminescent nanoparticle scaffold .....	93
1.1. Encapsulation procedure.....	95
1.2. Surface functionalization .....	96
1.3. Photophysical properties of YbZn <sub>16</sub> (quinoHA) <sub>16</sub> and corresponding NPs .....	98
2. Bioconjugation of luminescent nanoparticle scaffold .....	101
2.1. Gal $\beta$ -specific near-infrared imaging with Gal $\beta$ -ase functionalized luminescent nanoparticles	103
<b>III. Conclusion.....</b>	<b>104</b>
<b>Résumé.....</b>	<b>105</b>
<b>Bibliography.....</b>	<b>106</b>

## **I. Engineering neolectins from glycosidases**

### **1. Application in biotechnology, diagnostic and imaging**

The broad potential of glycoside hydrolase engineering has been used either to improve or alter their catalytic activity thus enabling access to diverse carbohydrates and their derivatives, often new to natural ones. Among the engineered carbohydrates, the representative groups used for biocatalyzed synthesis are glycosynthases, thioglycoligases and thioglycosynthases.<sup>1</sup> Apart from their synthetic usage, another potential based on their specific enzyme-substrate interactions can be explored. In a wide sense, glycoside hydrolases can be considered as specific carbohydrate-binding proteins and their enzyme-substrate interactions are comparable to lectin-carbohydrate or antigen-antibody interactions in terms of specificity.<sup>2</sup> Therefore, if hydrolytic activity of these carbohydrate-processing enzymes is inactivated without changing their carbohydrate binding and recognition properties, they could be converted into lectins.<sup>2</sup> Lectins are carbohydrate-binding proteins of non-immune origin that recognize and non-covalently bind simple or complex carbohydrates without enzymatically altering them.<sup>3,4</sup>

Historically, pioneering work of inactivation of hydrolytic activity began in the late 1960s on a hen egg-white lysozyme as a model enzyme. It should be mentioned, however, at that time these inactivation studies were not meant or considered to be done in the context of lectin generation but rather to elucidate the hydrolytic mechanism.<sup>5</sup> Nevertheless, the observation that these derivatives are enzymatically inactive but still retain their binding capability to carbohydrate substrate was significant. Therefore, in 1994 the selective inactivation of one of the catalytic residues of the lysozyme was the first report of intentional conversion of glycoside hydrolase to lectin-like protein. This kind of engineered glycoside hydrolase was denoted 'neolectin' as relation to carbohydrate-binding characteristics to natural lectins.<sup>6</sup>

The recombinant DNA technology and gene engineering enabled neolectin generation from a diverse repertoire of glycoside hydrolases through site-directed or random mutagenesis.<sup>4</sup> To the best of our knowledge exclusively pyranose-specific glycoside hydrolases were used as templates to create neolectins. Currently, the catalytically inactive glycoside hydrolases were mainly developed as microbial diagnostic and imaging probes. The notable example includes endoglycosidase, chitinase of which was

extensively used as template enzyme in several neolectin studies. Binding affinities for the catalytic acid/base (E172Q) inactive chitinase, Chit42 from *Trichoderma harzianum*, were altered when additional mutations at substrate binding-site domain were introduced thus broadening binding specificity to non-chitinous oligosaccharides.<sup>7</sup> Bacterial chitinase from *Vibrio parahaemolyticus* was catalytically inactivated and used as fungal or bacterial detection systems.<sup>2</sup> Catalytically inactive endosialidase, which could bind but not degrade polysialic acid, was used as both fused with the green fluorescent protein (GFP)<sup>8</sup> and as bacteriophage-bound neolectin for detection of bacterial and eukaryotic polysialic acid.<sup>2</sup>

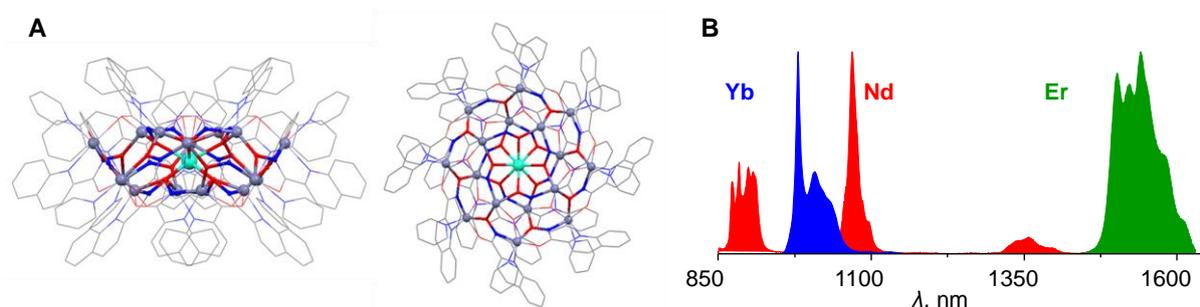
The thioglycosylase engineering of furanosidases,  $\alpha$ -L-arabinofuranosidase Ara $\beta$ 51 from *Clostridium thermocellum*, in particular, was successfully accomplished and expanded existing repertoire of engineered glycoside hydrolases to another group of carbohydrate structure forms.<sup>9</sup> Nevertheless, presently, none of the existing characterized furanosidases was submitted to engineering with the purpose of creating corresponding neolectins. However, there was only one report that described galactose-binding lectin. In 2015 a human lectin named human intelectin (hIntL) was identified and its cDNA cloned. This natural lectin was purified from human placenta and was characterized as a secretory glycoprotein consisting of 295 amino acids with basic structural unit being a 120 kDa homotrimer.<sup>10</sup> This new type of lectin has (i) an affinity to galactofuranosyl residues, (ii) can interact with multiple glycan epitopes exclusively related to microbes, such is arabinogalactan, and (iii) has a well resolved crystal structure.<sup>11</sup> Today, hIntL also known as intelectin-1 (ITNL1) or omentin, is commercially available through different kits as recombinant protein and is used for research purposes in ELISA-based assays with corresponding anti-INTL1 antibodies of different origin.

## II. Construction of luminescent nanoparticle scaffold for neolectin bioconjugation & imaging applications

### 1. Construction and characterization of luminescent nanoparticle scaffold

To create galactofuranose-specific lanthanide-based luminescent probe for near-infrared (NIR) optical imaging, several aspects have to be considered. First, a NIR-emitting, photostable and biocompatible (soluble or able to form stable suspension in water) lanthanide-based system that would allow its functionalization and attachment of Galf-ase has to be chosen and created. Second, *in vitro* imaging object has to be selected and the corresponding microscopy experiment designed. Mainly first aspect is addressed in this chapter. The experiments concerning second aspect are currently in progress.

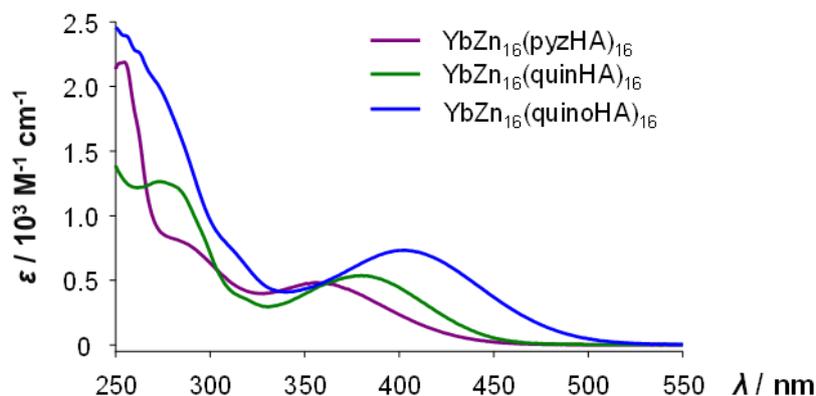
Among different classes of NIR-emitting  $\text{Ln}^{3+}$  complexes, our attention was turned to metallacrowns (MCs). Indeed, recently, it has been demonstrated that  $\text{Ln}^{3+}/\text{Zn}^{2+}$  MCs with encapsulated sandwich structure (**Figure IV-1. A**) can well protect and efficiently sensitize characteristic emission of  $\text{Yb}^{3+}$ ,  $\text{Nd}^{3+}$  and  $\text{Er}^{3+}$  in the NIR range with top values of quantum yields.<sup>12</sup> Derivatives of  $\text{Ln}^{3+}/\text{Zn}^{2+}$  MCs assembled using pyrazinehydroxamic acid (pyzHA) are water soluble and can be used as NIR-emitting probes for the selective labelling of necrotic HeLa cells<sup>13</sup> or for the simultaneous cell fixation and labelling<sup>14</sup>.



**Figure IV-1. A)** X-ray crystal structure of  $\text{DyZn}_{16}(\text{quinHA})_{16}$  along the (left)  $a$ -axis and (right)  $c$ -axis. **B)** Emission spectra of  $\text{LnZn}_{16}(\text{quinHA})_{16}$  ( $\text{Ln} = \text{Yb}, \text{Nd}, \text{Er}$ ,  $\text{quinHA} = \text{quinaldichydroxamic acid}$ ).<sup>12</sup>

Despite numerous advantages reported for  $\text{Ln}^{3+}/\text{Zn}^{2+}$  MCs with pyzHA and quinHA hydroxamic acids, these complexes have absorption bands ending at 450 nm (**Figure IV-2.**). To shift absorption more towards the visible range  $\text{Ln}^{3+}/\text{Zn}^{2+}$  MCs have been assembled using quinoxalinehydroxamic acid (quinoHA). The resulting MCs possess the same as previously reported  $\text{LnZn}_{16}(\text{HA})_{16}$  ( $\text{HA} = \text{pyzHA}, \text{quinHA}$ )<sup>12,13</sup> encapsulated

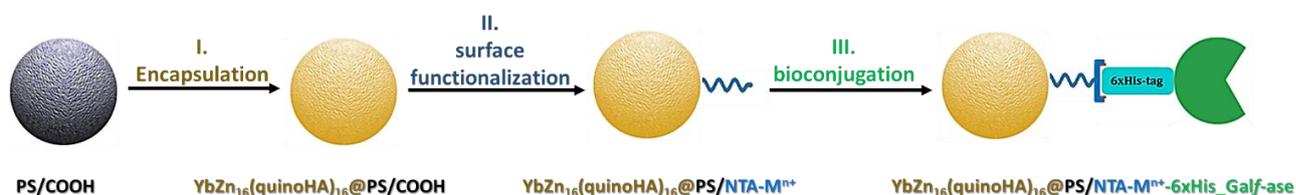
sandwich structure (**Figure IV-4. A**) while absorption bands are extended to 500 nm (**Figure IV-2.**).



**Figure IV-2.** Absorption spectra of  $\text{YbZn}_{16}(\text{HA})_{16}$  MCs (HA = pyzoHA, quinHA, quinoHA) in methanol or water solutions at room temperature.

Synthetic strategies were developed to introduce functional groups (alkyl or maleimide) onto MC scaffold.<sup>15</sup> However, to avoid long synthetic procedures and potentially improve detection sensitivity by combining a large number of  $\text{Ln}^{3+}$  complexes within one entity, we decided to encapsulate hydrophobic  $\text{YbZn}_{16}(\text{quinoHA})_{16}$  MCs inside 100 nm polystyrene (PS) nanoparticles (NPs). Such NPs are biocompatible, form stable suspension in water, can encapsulate hydrophobic molecules *via* easy swelling procedures and, finally, possess functional groups (PS/COOH) on the surface for further coupling with Galf-ase.

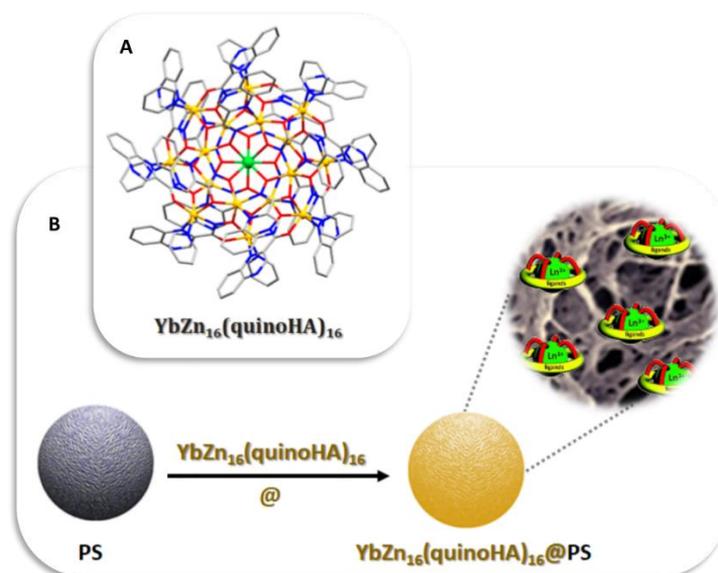
Therefore, the necessary modifications can be achieved in three main steps (**Figure IV-3.**): (I) encapsulation of  $\text{Ln}^{3+}$  MCs ( $\text{YbZn}_{16}(\text{quinoHA})_{16}@\text{PS}/\text{COOH}$ ) followed by (II) a subsequent surface-functionalization ( $\text{YbZn}_{16}(\text{quinoHA})_{16}@\text{PS}/\text{NTA}-\text{M}^{n+}$ ), which allowed final modification step, (III) conjugation of an enzyme onto the surface of NPs ( $\text{YbZn}_{16}(\text{quinoHA})_{16}@\text{PS}/\text{NTA}-\text{M}^{n+}-6\text{xHis\_Galf-ase}$ ).



**Figure IV-3.** Schematic representation of the three main steps in which internal and external functionalization of NPs can be achieved.

### 1.1. Encapsulation procedure

For the proof-of-a-concept creation and investigation of galactofuranose-specific luminescent probe for near-infrared optical imaging, the structurally and photophysically well defined Yb<sup>III</sup>/Zn<sup>II</sup> metallocrown (MC) sandwich complex<sup>12,16</sup> was preferred choice as luminescent agent. Due to literature reported data they show NIR luminescence and high quantum yield values for this type of Yb<sup>3+</sup> containing MC compounds and were previously explored as stains for *in vitro* whole-cell visualisation by NIR optical microscopy.<sup>14</sup> Moreover, in these structures, a single Yb<sup>3+</sup> cation is sandwiched between two [12-MC<sub>Zn(II),quinHA-4</sub>] MCs and is further encapsulated by an additional [24-MC<sub>Zn(II),quinHA-8</sub>] MC, leading to an overall stoichiometry of YbZn<sub>16</sub>(quinHA)<sub>16</sub>, in which the MC framework is formed by the self-assembled organization of Zn<sup>2+</sup> ions and a tetradentate chromophoric ligand based on quinaldichydroxamic acid (quinHA) (**Figure IV-4. A**).<sup>12</sup> The hydrophobic character of this YbZn<sub>16</sub>(quinHA)<sub>16</sub> complex makes it suitable for the encapsulation inside, as well hydrophobic, cavities of polystyrene NPs and therefore minimises the leaking of the complex into aqueous solution and undesired interactions with biomolecules in the case of *in vivo* application.

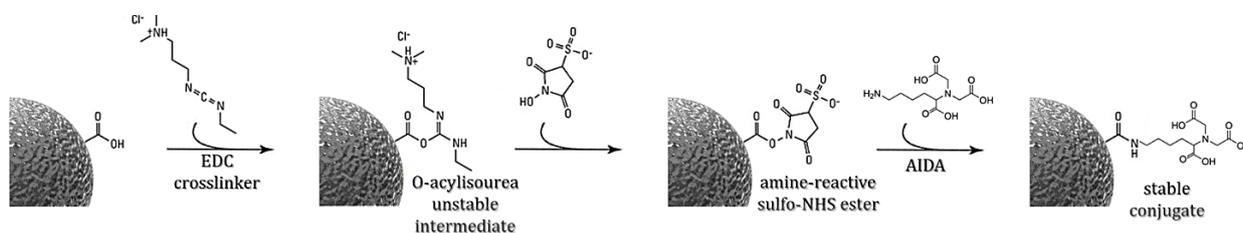


**Figure IV-4.** Schematic representation of the encapsulation step . **A)** Top-down view of the crystal structure of YbZn<sub>16</sub>(quinHA)<sub>16</sub> MC obtained by X-ray diffractometry on a single crystal, done at the Department of Chemistry of the University of Michigan by Tu N. Nguyen and Jeff W. Kampf (colour code: green, Yb; yellow, Zn; red, O; blue, N; grey, C; hydrogen atoms and solvents have been omitted for clarity). **B)** Encapsulation of YbZn<sub>16</sub>(quinHA)<sub>16</sub> complex inside PS/COOH NPs followed by subsequent colour change to orange.

A facile and rapid one-step swelling procedure for the encapsulation of YbZn<sub>16</sub>(quinoHA)<sub>16</sub> inside PS/COOH NPs was used (**Figure IV-4. B**). The procedure includes a strong agitation of an aqueous suspension of the NPs with a solution of YbZn<sub>16</sub>(quinoHA)<sub>16</sub> in DMF, followed by centrifugation and washing steps. This strategy was adapted from the previously reported experimental procedures developed for the encapsulation of hydrophobic organic dyes and/or Ln<sup>3+</sup> inside polymer particles.<sup>17-20</sup> After encapsulation of YbZn<sub>16</sub>(quinoHA)<sub>16</sub>, the apparent change of colour of NPs to yellow could be observed reflecting the success of this step (**Figure IV-4. B**). The resulted YbZn<sub>16</sub>(quinoHA)<sub>16</sub>@PS/COOH NPs were stable in water, maintained colloidal properties and no apparent aggregation or sedimentation was observed even after storage for several months at 4°C.

## 1.2. Surface functionalization

The encapsulation step provided an elegant modification of the interior of NPs while leaving surface COOH groups free for further functionalization. In order to immobilize Gal $\beta$ -ase that possess 6xHis-tag the surface, NPs had to be functionalized with nitrilotriacetic acid (NTA), that is a known chelating agent for Ni<sup>2+</sup>, Zn<sup>2+</sup>, Cu<sup>2+</sup>, Co<sup>2+</sup>, Ca<sup>2+</sup> and Fe<sup>3+</sup>. A two-step procedure of EDC/sulfo-NHS activation of the carboxylated NPs surface was followed by the reaction with (S)-N-(5-amino-1-carboxypentyl)iminodiacetic acid (AIDA) to give PS/NTA NPs (**Figure IV-5.**). The latter were further treated with Ni<sup>2+</sup> or Zn<sup>2+</sup> chlorides to give PS/NTA-Ni<sup>2+</sup> or PS/NTA-Zn<sup>2+</sup>, respectively. The entire surface modification procedure was first established and optimized for PS/COOH NPs and then applied for YbZn<sub>16</sub>(quinoHA)<sub>16</sub>@PS/COOH.



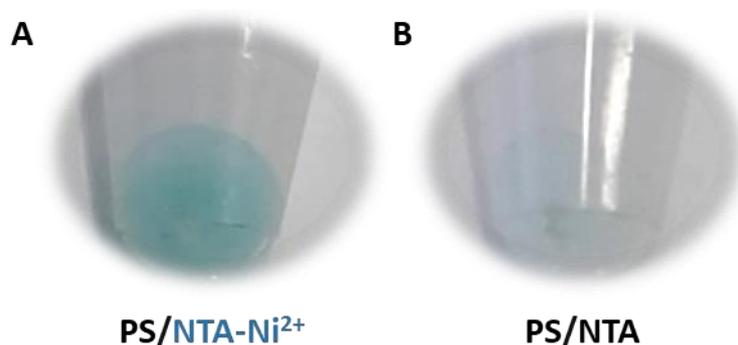
**Figure IV-5.** Schematic representation of the synthetic route used for the preparation of PS/NTA NPs: activation of NPs carboxylic groups with EDC/sulfo-NHS to create the amide bond with AIDA.

The specific PS/NTA-Ni<sup>2+</sup> surface functionalization was primarily chosen based on immobilized metal affinity chromatography (IMAC) technique used for polyhistidine-

tagged enzyme purification which takes advantage of the strong binding interaction between histidine moieties of the enzyme and Ni<sup>2+</sup>. Therefore, NTA-Ni<sup>2+</sup> surface functionalization was exploited as already efficiently proven system for 6xHis-tag Gal $\beta$ -ase immobilization on identical purification system column.

However, it is known that transition metal ions like Cu<sup>2+</sup>, Co<sup>2+</sup> and Ni<sup>2+</sup> can quench Ln<sup>3+</sup> luminescence.<sup>21</sup> Therefore, NPs have been also treated with Zn<sup>2+</sup> chloride to form (YbZn<sub>16</sub>(quinoHA)<sub>16</sub>@)PS/NTA-Zn<sup>2+</sup>. The obtained (YbZn<sub>16</sub>(quinoHA)<sub>16</sub>@)PS/NTA-Zn<sup>2+</sup> NPs were not stable in water and did not maintain colloidal properties. Therefore, (YbZn<sub>16</sub>(quinoHA)<sub>16</sub>@)PS/NTA-Zn<sup>2+</sup> NPs were not suitable for Gal $\beta$ -ase bioconjugation.

NTA is a tetradentate chelating ligand that binds Ni<sup>2+</sup> ions by four coordination groups and forms NTA-Ni<sup>2+</sup> complex with characteristic blue-green colour. An incubation of PS/NTA NPs suspension with the solution of Ni<sup>2+</sup> nitrate resulted in the formation of PS/NTA-Ni<sup>2+</sup> NPs, pellets of which appeared blue-green in colour indicating the successful chelation (**Figure IV-6. A**). In order to demonstrate that such colour change of NPs was induced by the surface coordination of Ni<sup>2+</sup>, PS/NTA-Ni<sup>2+</sup> were treated with ethylenediaminetetraacetic acid (EDTA), a strong hexadentate chelating agent. This step caused pellet to turn back to white due to the formation of more stable EDTA-Ni<sup>2+</sup> chelate (**Figure IV-6. B**). This experiment proved the presence of NTA-Ni<sup>2+</sup> moieties<sup>22</sup> on the surface of NPs which were to facilitate bioconjugation of Gal $\beta$ -ase in the next experimental step.



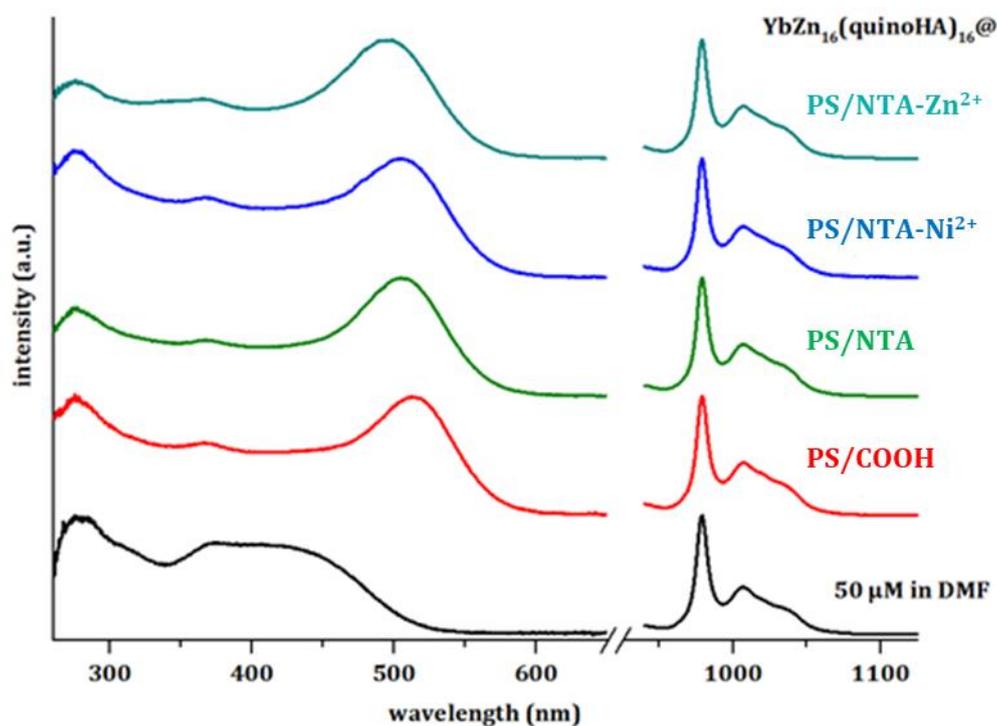
**Figure IV-6.** Photographs of the pellets of PS/NTA-Ni<sup>2+</sup> NPs (**A**) before and (**B**) after incubation with EDTA. The change of colour indicates the release of Ni<sup>2+</sup> from the surface.

In the case of YbZn<sub>16</sub>(quinoHA)<sub>16</sub>@PS/NTA, the observation of such colour changes upon complexation with Ni<sup>2+</sup> was impossible because of the initial yellow colour of the NPs.

Therefore, the presence of  $\text{Ni}^{2+}$  in the supernatant after the treatment of  $\text{YbZn}_{16}(\text{quinoHA})_{16}@PS/\text{NTA}-\text{Ni}^{2+}$  with EDTA was confirmed by ICP measurements (Chapter VI, Experimental part 6.5.5.).

### 1.3. Photophysical properties of $\text{YbZn}_{16}(\text{quinoHA})_{16}$ and corresponding NPs

Photophysical properties (excitation and emission spectra,  $\text{Yb}^{3+}$ -centered quantum yields under ligands excitation ( $Q_{\text{Yb}}^L$ ) and luminescence lifetimes ( $\tau_{\text{obs}}$ ) of aqueous suspension of  $\text{YbZn}_{16}(\text{quinoHA})_{16}@PS/\text{COOH}/\text{NTA}/\text{NTA}-\text{M}^{2+}$  NPs after each functionalization step were acquired and compared with those measured for the initial solution of  $\text{YbZn}_{16}(\text{quinoHA})_{16}$  in DMF. This work allowed to monitor the effects of the encapsulation of  $\text{YbZn}_{16}(\text{quinoHA})_{16}$  inside the PS/COOH NPs, as well as surface functionalization, i.e. the presence of NTA functional groups alone or chelated with  $\text{Ni}^{2+}$  and  $\text{Zn}^{2+}$  ions. First, excitation spectra were acquired upon monitoring the main transition of  $\text{Yb}^{3+}$  at 980 nm (Figure IV-7. left). They exhibit broad ligands-centered bands that extend to 530-570 nm and correlate with the absorption spectra of  $\text{YbZn}_{16}(\text{quinoHA})_{16}$  (Figure IV-2.). Upon excitation at 420 or 510 nm of solution of  $\text{YbZn}_{16}(\text{quinoHA})_{16}$  in DMF and aqueous suspensions of NPs exhibit characteristic emission of  $\text{Yb}^{3+}$  in the range 950-1100 nm with maxima at 980 nm due to  ${}^2\text{F}_{5/2} \rightarrow {}^2\text{F}_{7/2}$  transition (Figure IV-7. right).



**Figure IV-7.** Corrected and normalized (left) excitation ( $\lambda_{\text{em}} = 980 \text{ nm}$ ) and (right) emission spectra of  $\text{YbZn}_{16}(\text{quinoHA})_{16}$  in DMF solution ( $50 \mu\text{M}$ ,  $\lambda_{\text{ex}} = 420 \text{ nm}$ ) or for

aqueous suspensions of YbZn<sub>16</sub>(quinoHA)<sub>16</sub>@PS NPs (6 mg/mL,  $\lambda_{\text{ex}} = 510$  nm), room temperature.

Quantitative photophysical parameters, including  $\tau_{\text{obs}}$  and  $Q_{\text{Yb}}^{\text{L}}$  were also acquired and are summarized in **Table IV-1**.

**Table IV-1.** Luminescence lifetimes ( $\tau_{\text{obs}}$ ) and absolute quantum yields ( $Q_{\text{Yb}}^{\text{L}}$ ) of YbZn<sub>16</sub>(quinoHA)<sub>16</sub> in DMF solution (50  $\mu\text{M}$ ) or for aqueous suspensions of YbZn<sub>16</sub>(quinoHA)<sub>16</sub>@PS NPs (6 mg/mL).<sup>a</sup>

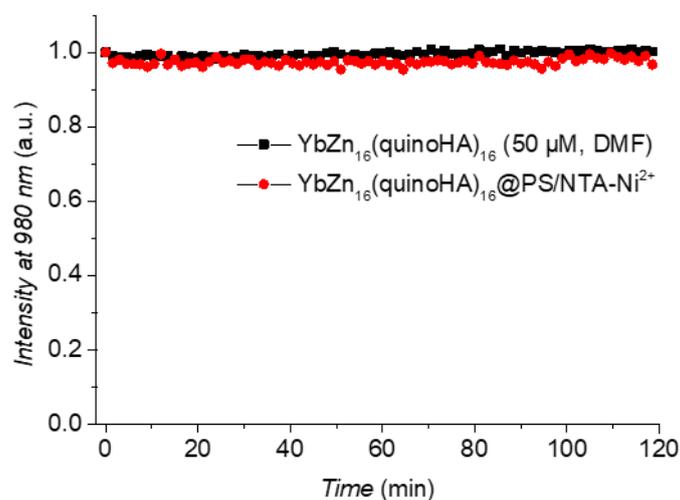
Conditions	$\tau_{\text{obs}}$ ( $\mu\text{s}$ ) <sup>b</sup>	$Q_{\text{Yb}}^{\text{L}}$ (%) <sup>c</sup>
YbZn <sub>16</sub> in DMF	49.2(4)	0.57(1)
YbZn <sub>16</sub> @PS/COOH	24.6(2) : 94(2)% & 9(1) : 6(2)%	0.13(1)
YbZn <sub>16</sub> @PS/NTA	25.3(1)	0.15(1)
YbZn <sub>16</sub> @PS/NTA-Ni <sup>2+</sup>	9.9(1) : 77(1)% & 3.3(2) : 23(1)%	0.041(1)
YbZn <sub>16</sub> @PS/NTA-Zn <sup>2+</sup>	25.2(1)	0.14(1)

<sup>a</sup>Data at room temperature. Standard deviation ( $2\sigma$ ) values between parentheses; estimated relative errors:  $\tau_{\text{obs}}$ ,  $\pm 2\%$ ;  $Q_{\text{Yb}}^{\text{L}}$ ,  $\pm 10\%$ . <sup>b</sup>Under excitation at 355 nm; contributions of each individual lifetime value are provided after the colon. <sup>c</sup>Under excitation at 420 nm for DMF solution and 500 nm for NPs suspensions.

An aqueous suspension of YbZn<sub>16</sub>(quinoHA)<sub>16</sub>@PS/COOH exhibit 2-time shorter luminescence lifetimes and 4.4-times smaller value of  $Q_{\text{Yb}}^{\text{L}}$  compared to the solution of this MC in DMF. Such behaviour can be explained by the presence of high-energy O-H vibrations overtones of which are known to quench Ln<sup>3+</sup> luminescence particularly in the NIR range.<sup>21</sup> So, even so that PS NPs are hydrophobic, they are not providing a complete protection against O-H vibrations. It should be noted also that luminescence decay for YbZn<sub>16</sub>(quinoHA)<sub>16</sub>@PS/COOH could be best fitted by a bi-exponential function, indicating a more significant impact of O-H vibrations on Yb<sup>3+</sup> for YbZn<sub>16</sub>(quinoHA)<sub>16</sub> molecules located closer to the surface of the NP, than for these in the interior. This impact, however, becomes negligible upon further functionalization of the NPs surface and single-exponential decays were recorded for YbZn<sub>16</sub>(quinoHA)<sub>16</sub>@PS/NTA. This functionalization step didn't affect  $Q_{\text{Yb}}^{\text{L}}$  that remained the same within experimental error. Further, complexation of NTA groups on the surface with Zn<sup>2+</sup> does not change quantitative parameters. However, in the case of Ni<sup>2+</sup>, a more than 2.5-times shortening of the  $\tau_{\text{obs}}$  and 3.6-times decrease of  $Q_{\text{Yb}}^{\text{L}}$  are observed. Luminescence decay of

$\text{YbZn}_{16}(\text{quinoHA})_{16}@PS/NTA\text{-Ni}^{2+}$  is biexponential highlighting a distance dependence of  $\text{Ni}^{2+}$  quenching and smaller impact of this phenomenon on the emission of  $\text{YbZn}_{16}(\text{quinoHA})_{16}$  located in the interior of the NPs. Nevertheless, despite this quenching effect, characteristic  $\text{Yb}^{3+}$  emission in the NIR range could be clearly observed for  $\text{YbZn}_{16}(\text{quinoHA})_{16}@PS/NTA\text{-Ni}^{2+}$  (**Figure IV-7.**). Because of the higher affinity of His-tag towards  $\text{Ni}^{2+}$  compared to  $\text{Zn}^{2+}$ , further studies were performed with  $\text{YbZn}_{16}(\text{quinoHA})_{16}@PS/NTA\text{-Ni}^{2+}$ .

The photostability of the imaging probe is an important characteristic that reflects the possibility to perform long-term or repeated experiments, as well as these that require high power excitation sources. To estimate this photostability, a solution of  $\text{YbZn}_{16}(\text{quinoHA})_{16}$  in DMF or an aqueous suspension of  $\text{YbZn}_{16}(\text{quinoHA})_{16}@PS/NTA\text{-Ni}^{2+}$  was placed in capillaries, inserted into an integration sphere and continuously illuminated at 405 nm (maximum of absorption, **Figure IV-2.**) while  $\text{Yb}^{3+}$ -centered emission in the NIR at 980 nm was monitored (**Figure IV-8.**). The results demonstrated that the signal remained constant after 120 min of illumination for both solution of  $\text{YbZn}_{16}(\text{quinoHA})_{16}$  and NPs.



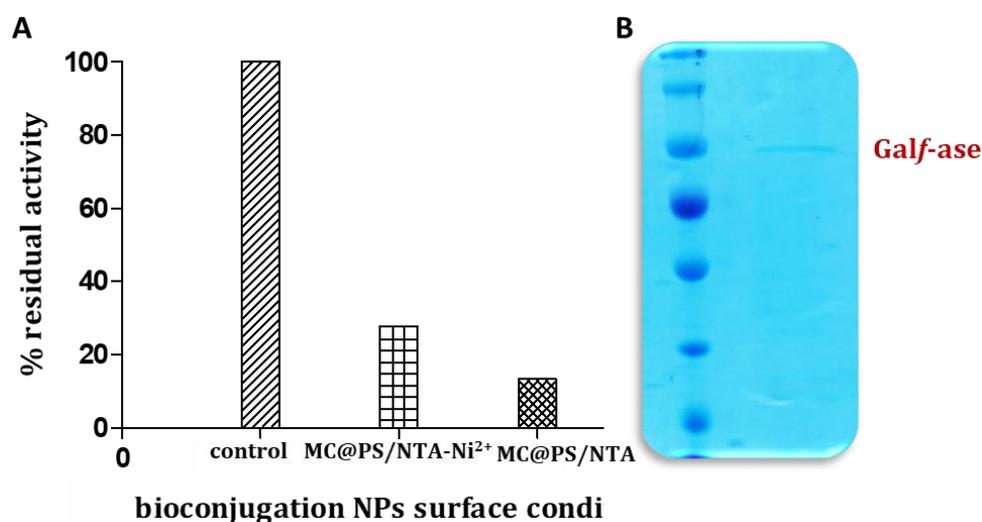
**Figure IV-8.** Changes in emission intensity at 980 nm for a solution of  $\text{YbZn}_{16}(\text{quinoHA})_{16}$  in DMF (50  $\mu\text{M}$ ) or for an aqueous suspension of  $\text{YbZn}_{16}(\text{quinoHA})_{16}@PS/NTA\text{-Ni}^{2+}$  NPs (6 mg/mL) under continuous illumination at 405 nm.

## 2. Bioconjugation of luminescent nanoparticle scaffold

To ensure specificity of constructed NPs scaffold towards *Galf*, surface had to be coated with *Galf*-ase. The bioconjugation of  $(\text{YbZn}_{16}(\text{quinoHA})_{16})@\text{PS}/\text{NTA-Ni}^{2+}$  was done in a one-pot simple and fast procedure that combined the incubation of NPs with 6xHis-tag *Galf*-ase enzyme under agitation to assure optimal interaction. It was assumed that histidine moieties of 6xHis-tag *Galf*-ase can chelate the  $\text{Ni}^{2+}$  immobilized on the surface of the NPs. The same principle was used before and during the affinity chromatography purification of the enzyme. This bioconjugation approach was a preferable choice since it employs mild conditions that could retain enzyme conformation and activity. In principle, enzyme could be covalently attached to the NPs surface, but such binding in the most cases affected the activity of the enzyme.<sup>23,24</sup> Since our primary concern was to preserve the *Galf*-ase function, 6xHis-tag was used to serve for both the initial purification and the subsequent immobilization on the surface of NPs.<sup>25</sup>

To simplify the experimental procedure and the assessment of the enzyme activity, wild-type *Galf*-ase was used for the initial trial rather than E464Q mutant variant which exhibits attenuated hydrolytic activity. Briefly, upon bioconjugation, NPs were submitted to three subsequent washing procedures, meaning pelleting by centrifugation and resuspension in buffered solution by pipetting, in order to remove excess or non-specifically bound enzyme. After a final washing step, NPs were resuspended in buffered solution (205  $\mu\text{L}$ ), to ensure an optimal pH of 4.5, and were used for *Galf*-ase activity assessment by previously described colorimetric assay. To access the specificity of enzyme binding,  $(\text{YbZn}_{16}(\text{quinoHA})_{16})@\text{PS}/\text{NTA}$  NPs lacking  $\text{Ni}^{2+}$  were used as control. Percentage of *Galf*-ase activity originating from  $(\text{YbZn}_{16}(\text{quinoHA})_{16})@\text{PS}/\text{NTA}$  and  $(\text{YbZn}_{16}(\text{quinoHA})_{16})@\text{PS}/\text{NTA-Ni}^{2+}$  NPs was determined and normalized to an initial activity measured for free, non-conjugated enzyme (**Figure IV-9. A**). The specific activity of free *Galf*-ase, tested with *pNP*- $\beta$ -D-*Galf* synthetic substrate, was 0,189 U/g, thus allowing to determine the remaining quantity of enzyme in each set of surface functionalized NPs from activity percentage since for all tested bioconjugation conditions the equal amount of 60  $\mu\text{g}$  of *Galf*-ase was used per 2 mg of NPs. These results suggest that approximately 16  $\mu\text{g}$  of *Galf*-ase remained immobilized on  $\text{NTA-Ni}^{2+}$  surface functionalized NPs, providing 27% of its free enzyme activity under the same experimental conditions, while the 13% of activity corresponds to quantity of 8  $\mu\text{g}$  of *Galf*-

ase that remained non-specifically immobilized on only NTA surface functionalized NPs. In addition, by using an imidazole, histidine side chain structural analogue, bound Galf-ase was released from the NTA-Ni<sup>2+</sup> and the obtained elution supernatant was further analysed by SDS-PAGE. As shown (**Figure IV-9. B**), the observed band at 90 kDa provided the evidence of an efficient conjugation of Galf-ase on the surface of (YbZn<sub>16</sub>(quinoHA)<sub>16</sub>)@PS/NTA-Ni<sup>2+</sup> NPs through the histidine tag strategy. On the other hand, the amount of non-specifically bound Galf-ase on the surface of (YbZn<sub>16</sub>(quinoHA)<sub>16</sub>)@PS/NTA was below the detection level by SDS-PAGE analysis.



**Figure IV-9.** Investigation of the activity of the bioconjugated Galf-ase. **A)** The specific *p*NP- $\beta$ -D-Galf activity of Galf-ase obtained after separate bioconjugation with NTA-Ni<sup>2+</sup> and NTA surface functionalized NPs compared to free enzyme as a control. **B)** The SDS-PAGE analysis of imidazole released Galf-ase from NTA-Ni<sup>2+</sup> surface functionalized NPs upon activity assay.

This bioconjugation procedure allowed to investigate whether the enzyme can retain its activity after being submitted to slightly basic pH conditions with pH 8 during the reaction. Even though these basic conditions do not favour Galf-ase catalytic activity, as was demonstrated in the optimum pH activity assay (**Chapter II, Figure II-12. A**), lowering the pH to 4.5 was avoided since it would lead to a loss of chelated Ni<sup>2+</sup> ions due to the protonation of the chelating carboxylic groups of NTA and would equally protonate the imidazole nitrogen of the histidine residue and therefore disrupt the coordination bond between the 6xHis-tag and the transition metal (**Figure IV-10.**)<sup>26</sup>



### III. Conclusion

The main objective of this chapter was to create a lanthanide(III)-based luminescent probe able to carry an engineered enzyme, glycoside hydrolase that, through inactivation of its catalytic activity, can become a specific optical imaging tool for the detection of its substrate on *in vitro* model. In order to construct such probe, a facile protocol had to be developed. Therefore, polystyrene nanoparticles were chosen since they can encapsulate hydrophobic lanthanide(III)-based complex and possess surface groups to allow further functionalization and attachment of enzyme.

Ytterbium(III)/zinc(II) metallacrown with encapsulated sandwich structure, YbZn<sub>16</sub>(quinoHA)<sub>16</sub>, was chosen and used for encapsulation in PS/COOH NPs. Carboxylic groups on the surface of NPs were then used for further functionalization to obtain YbZn<sub>16</sub>(quinoHA)<sub>16</sub>@PS/NTA-Ni<sup>2+</sup>. Photophysical properties of aqueous suspensions of YbZn<sub>16</sub>(quinoHA)<sub>16</sub>@PS/COOH(/NTA/NTA-Ni<sup>2+</sup>) NPs with different surface groups were investigated in detail and compared to these of the initial YbZn<sub>16</sub>(quinoHA)<sub>16</sub> in DMF solution. The NTA-Ni<sup>2+</sup> surface functionalized NPs were then bioconjugated with Gal $\alpha$ -ase using mild conditions which allowed 6xHis-tag orientated immobilization and conservation of enzyme activity.

The development of luminescent probe for NIR imaging of this kind brought together two concepts, the use of lanthanide(III)-based compounds, as potential optical imaging agents, and neolectins, enzymes that specifically bind to substrate but do not degrade it and can be used as the equivalent of an epitope-specific antibody.

### Résumé

L'objectif principal de ce chapitre était de créer une sonde luminescente à base de lanthanide(III) capable de transporter une enzyme modifiée, la glycoside hydrolase qui, par inactivation de son activité catalytique, peut devenir un outil d'imagerie optique spécifique pour la détection de son substrat sur modèle *in vitro*. Afin de construire une telle sonde, un protocole simple a dû être développé. Par conséquent, les nanoparticules de polystyrène ont été choisies parce qu'elles peuvent encapsuler un complexe hydrophobe à base de lanthanide(III) et possèdent des groupes de surface qui permettent une fonctionnalisation et une fixation plus poussées de l'enzyme.

L'Ytterbium(III)/zinc(II) métallacrown possédait une structure sandwich encapsulée,  $\text{YbZn}_{16}(\text{quinoHA})_{16}$ , a été choisi et utilisé pour l'encapsulation dans les NP PS/COOH. Des groupes carboxyles à la surface des nanoparticules ont ensuite été utilisés pour poursuivre la fonctionnalisation afin d'obtenir  $\text{YbZn}_{16}(\text{quinoHA})_{16}@\text{PS}/\text{NTA}-\text{Ni}^{2+}$ . Les propriétés photophysiques des suspensions aqueuses de  $\text{YbZn}_{16}(\text{quinoHA})_{16}@\text{PS}/\text{COOH}(/ \text{NTA}/\text{NTA}-\text{Ni}^{2+})$  NPs avec différents groupes de surface ont été étudiées en détail et comparées à celles du  $\text{YbZn}_{16}(\text{quinoHA})_{16}$  initial en solution de DMF. Les nanoparticules fonctionnalisées en surface NTA- $\text{Ni}^{2+}$  ont ensuite été bioconjuguées avec *Galf*-ase en utilisant des conditions de conjugaison douces qui ont permis une immobilisation de 6xHis-tag orientée et la conservation de l'activité enzymatique.

La mise au point d'une sonde luminescente pour l'imagerie NIR de ce type a réuni deux concepts: l'utilisation de composés à base de lanthanide(III) comme agents d'imagerie optique et de néolectines, enzymes qui se lient spécifiquement au substrat mais ne le dégradent pas et qui peuvent être utilisées comme équivalent d'un anticorps épitope spécifique.

## Bibliography

1. Benkoulouche, M., Fauré, R., Remaud-Siméon, M., Moulis, C. & André, I. Harnessing glycoenzyme engineering for synthesis of bioactive oligosaccharides. *Interface Focus* **9**, 20180069 (2019).
2. *Lectins: analytical technologies*. (Elsevier, 2007).
3. Arnaud, J., Audfray, A. & Imberty, A. ChemInform Abstract: Binding Sugars: From Natural Lectins to Synthetic Receptors and Engineered Neolactins. *ChemInform* **44**, no-no (2013).
4. Hirabayashi, J. & Arai, R. Lectin engineering: the possible and the actual. *Interface Focus* **9**, 20180068 (2019).
5. The Identification of Aspartic Acid Residue 52 As Being Critical to Lysozyme Activity.
6. Mega, T. & Hase, S. Conversion of egg-white lysozyme to a lectin-like protein with agglutinating activity analogous to wheat germ agglutinin. *Biochim. Biophys. Acta BBA - Gen. Subj.* **1200**, 331–333 (1994).
7. Lienemann, M., Boer, H., Paananen, A., Cottaz, S. & Koivula, A. Toward understanding of carbohydrate binding and substrate specificity of a glycosyl hydrolase 18 family (GH-18) chitinase from *Trichoderma harzianum*. *Glycobiology* **19**, 1116–1126 (2009).
8. Aalto, J., Pelkonen, S., Kalimo, H. & Finne, J. Mutant bacteriophage with non-catalytic endosialidase binds to both bacterial and eukaryotic polysialic acid and can be used as probe for its detection. 8.
9. Almendros, M. *et al.* Exploring the synthetic potency of the first furanothioglycoligase through original remote activation. *Org. Biomol. Chem.* **9**, 8371 (2011).
10. Tsuji, S. *et al.* Human Intelectin Is a Novel Soluble Lectin That Recognizes Galactofuranose in Carbohydrate Chains of Bacterial Cell Wall. *J. Biol. Chem.* **276**, 23456–23463 (2001).
11. Wesener, D. A. *et al.* Recognition of Microbial Glycans by Human Intelectin. *Nat. Struct. Mol. Biol.* **22**, 603–610 (2015).
12. Trivedi, E. R. *et al.* Highly Emitting Near-Infrared Lanthanide “Encapsulated Sandwich” Metallacrown Complexes with Excitation Shifted Toward Lower Energy. *J. Am. Chem. Soc.* **136**, 1526–1534 (2014).

13. Martinić, I., Eliseeva, S. V., Nguyen, T. N., Pecoraro, V. L. & Petoud, S. Near-Infrared Optical Imaging of Necrotic Cells by Photostable Lanthanide-Based Metallacrowns. *J. Am. Chem. Soc.* **139**, 8388–8391 (2017).
14. Martinić, I. *et al.* Near-infrared luminescent metallacrowns for combined in vitro cell fixation and counter staining. *Chem. Sci.* **8**, 6042–6050 (2017).
15. Lutter, J. C., Lopez Bermudez, B. A., Nguyen, T. N., Kampf, J. W. & Pecoraro, V. L. Functionalization of luminescent lanthanide-gallium metallacrowns using copper-catalyzed alkyne-azide cycloaddition and thiol-maleimide Michael addition. *J. Inorg. Biochem.* **192**, 119–125 (2019).
16. Chow, C. Y. *et al.* Ga<sup>3+</sup>/Ln<sup>3+</sup> Metallacrowns: A Promising Family of Highly Luminescent Lanthanide Complexes That Covers Visible and Near-Infrared Domains. *J. Am. Chem. Soc.* **138**, 5100–5109 (2016).
17. Behnke, T. *et al.* Encapsulation of Hydrophobic Dyes in Polystyrene Micro- and Nanoparticles via Swelling Procedures. *J. Fluoresc.* **21**, 937–944 (2011).
18. Behnke, T., Würth, C., Laux, E.-M., Hoffmann, K. & Resch-Genger, U. Simple strategies towards bright polymer particles via one-step staining procedures. *Dyes Pigments* **94**, 247–257 (2012).
19. Laux, E.-M., Behnke, T., Hoffmann, K. & Resch-Genger, U. Keeping particles brilliant – simple methods for the determination of the dye content of fluorophore-loaded polymeric particles. *Anal. Methods* **4**, 1759–1768 (2012).
20. Martinić, I. *et al.* One Approach for Two: Toward the Creation of Near-Infrared Imaging Agents and Rapid Screening of Lanthanide(III) Ion Sensitizers Using Polystyrene Nanobeads. *ACS Appl. Bio Mater.* **2**, 1667–1675 (2019).
21. Reedijk, J. & Poepelmeier, K. R. *Comprehensive Inorganic Chemistry II (Second Edition): From Elements to Applications.* (Elsevier Ltd, 2013).
22. Hennig, A. *et al.* Simple Colorimetric Method for Quantification of Surface Carboxy Groups on Polymer Particles. *Anal. Chem.* **83**, 4970–4974 (2011).
23. Wang, W., Wang, D. I. C. & Li, Z. Facile fabrication of recyclable and active nanobiocatalyst: purification and immobilization of enzyme in one pot with Ni-NTA functionalized magnetic nanoparticle. *Chem. Commun.* **47**, 8115 (2011).
24. Salimi, K., Usta, D. D., Koçer, İ., Çelik, E. & Tuncel, A. Highly selective magnetic affinity purification of histidine-tagged proteins by Ni<sup>2+</sup> carrying monodisperse composite microspheres. *RSC Adv.* **7**, 8718–8726 (2017).

25. Liu, Y.-C. C. *et al.* Specific and reversible immobilization of histidine-tagged proteins on functionalized silicon nanowires. *Nanotechnology* **21**, 245105 (2010).
26. Bornhorst, J. A. & Falke, J. J. [16] Purification of Proteins Using Polyhistidine Affinity Tags. *Methods Enzymol.* **326**, 245–254 (2000).

## **CHAPTER V**

### **Conclusion & perspectives**

## V. Conclusion & perspectives

Interest for galactofuranose (*Gal**f*) and *Gal**f* processing enzymes are main or exclusive object of research interest in a few scientific groups and laboratories around the world. Carbohydrate chemistry is challenge for any organic chemist and no easier is the task of synthesis of less stable furanoside sugars, *Gal**f* as such. Some of them are focused on organochemical aspects of synthesis of *Gal**f* derivatives, while others combine chemoenzymatic approaches. The latter is no less challenging, since the triad of crucial enzymes involved in the biosynthesis and the metabolism of the *Gal**f*-containing molecules, UDP-galactopyranose mutase (UGM), galactofuranosyltransferase (*Gal**f*T) and galactofuranosidase (*Gal**f*-ase), is not completely accessible and characterized for chemoenzymatic or any other usage. Only a few of listed enzymes have been identified and cloned thus available to scientific community.

The Enzymology and Glycobiology research group, is unique in scientific community to be in possession of cloned gene sequences encoding for two UGMs, from *Escherichia coli* and *Lachnoanaerobaculum saburreum*, four putative *Gal**f*Ts from *Leishmania major*, LPG1, LPG1G, LPG1L and LPG1R respectively, and the only reported *Gal**f*-ase from *Streptomyces spp.* The work and results conducted on these enzymes have been acknowledged in recent publications.

To date, only a few furanosidases, originating from arabinofuranosidases, have been used in bioengineering and chemoenzymatic approaches, it was of great interest to explore and evaluate the potential of only *Gal**f*-ase in an original way. Therefore, having access to this valuable enzyme, it was natural to evaluate it from several aspects. As it was already indicated in the thesis objectives, there are three directions thesis was initiated to explore: **groundwork characterization, bioengineering potential and target-specific imaging.**

Primarily subcloning of *Gal**f*-ase gene insert was undertaken, followed by overexpression in prokaryotic host and purification in order to provide us with recombinant single 6xHis-tagged *Gal**f*-ase. This tag has a minor influence on protein structure, due to its 1 kDa size, and served for double purpose, for selective purification and as a linkage to imaging probe scaffold. Substrate specificity was determined through hydrolysis of more than twenty different *p*NP sugars among which only *p*NP- $\beta$ -D-*Gal**f* was selected as exclusive substrate.

Optimal activity conditions demonstrated good stability as well as good catalytic efficiency. This was of outmost importance in order to use *Galf*-ase in chemo-enzymatic synthesis. Altogether obtained results were encouraging to proceed with second part, the bioengineering of neolectins.

Engineering of neolectins from glycosidases by inactivation of hydrolytic activity was demonstrated and reported in the literature. Considering this approach was also employed to obtain first furanothioglycoligase, it was as well our choice of preference. Site-directed mutagenesis method was used to specifically mutate catalytic acid/base residue with catalytically inactive one, to generate four *Galf*-ase mutants. Following successful overexpression and purification three out of four generated mutants, the attenuation of hydrolytic activity was demonstrated through Michaelis-Menten assay. The best mutant variant, E464Q, in aspect of reasonable yield and kinetics, was selected for furanothioglycoside synthesis trials. The represent mutant was tested for thioglycoligase potency in a series of reactions with *p*NP- $\beta$ -D-*Galf* as donor and several thiophenol and 2-mercaptothiazoline derivatives as acceptors. The reactions yielded eight different thioglycoside products which were characterized by HPLC and HRMS-ESI. This strategy proofed the aim to have at once the mutant retaining glycosidase that can efficiently perform the synthesis of *S*-galactofuranosides and act as neolectin due to the low hydrolytic activity.

The final objective of the thesis, target-specific imaging of *Galf* glycan structures on the surface of certain microorganisms, was approached in several steps.

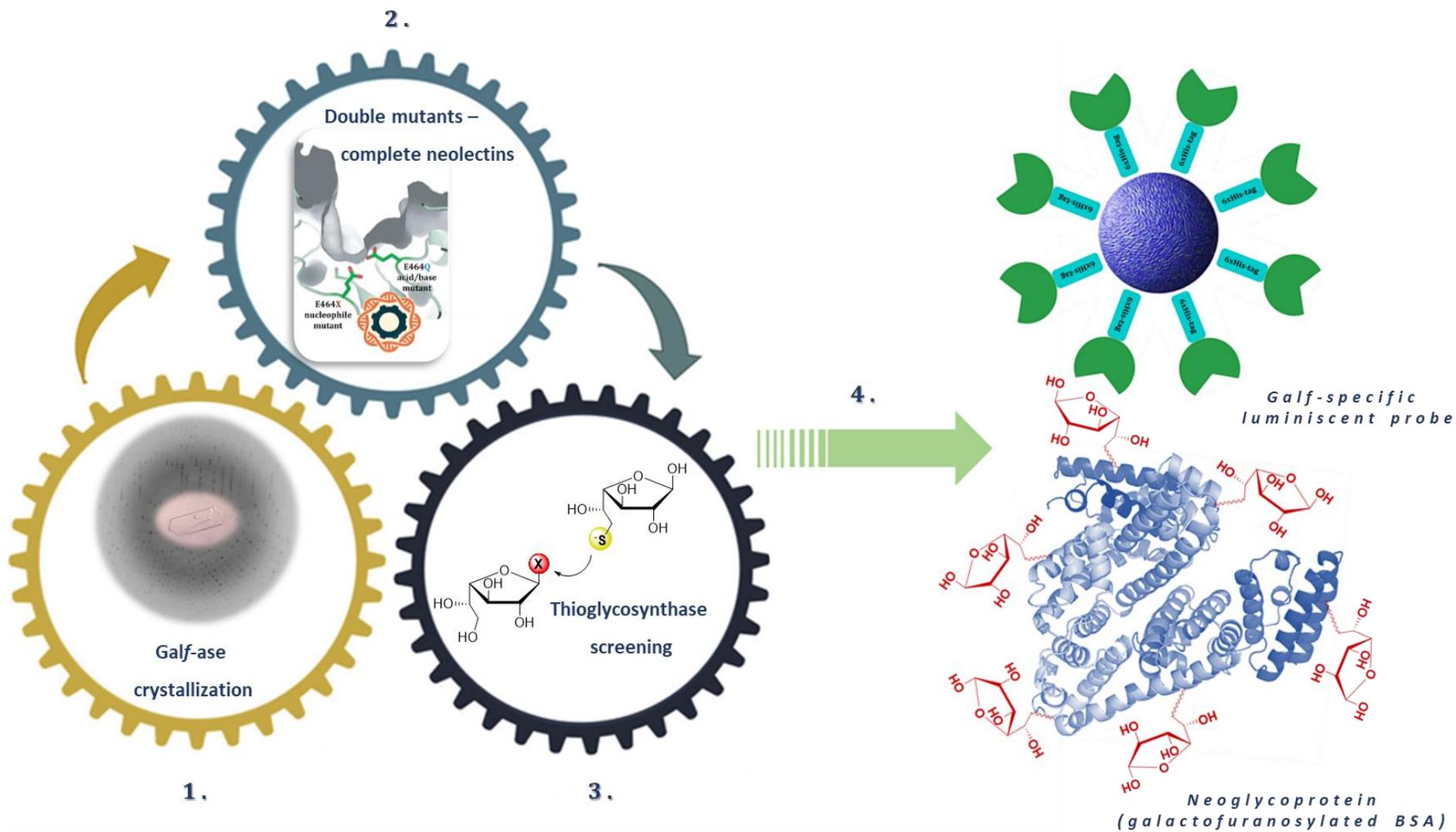
Polystyrene NPs were the first choice for building a NIR luminescent scaffold for neolectin bioconjugation since they are able to encapsulate lanthanide(III) complexes while carboxylated surface was appropriate for further functionalization. The simple and rapid swelling procedure allowed an encapsulation of lanthanide(III) metallacrowns inside NPs followed by a two-step EDC/sulfo-NHS covalent coupling of NTA groups and their further complexation with Ni<sup>2+</sup>. The entire surface functionalization strategy was proposed to specifically immobilize functionally intact neolectins *via* NTA-Ni<sup>2+</sup> and 6xHis-tag interaction. The wild-type *Galf*-ase was used at first to approve the bioconjugation protocol as well as the maintained enzymatic activity. After incubation of NTA-Ni<sup>2+</sup> functionalized NPs with *Galf*-ase and several washing steps, the activity of bound *Galf*-ase

was assessed by previously described *p*NP based enzymatic activity assay. It was proved that the *Galf*-ase on the surface of NPs retained its enzymatic activity upon bioconjugation and will allow us to use such NPs as imaging probes in the future.

This work was preparatory for finalisation of the main objective in a form of finding a sustainable protocol to deliver it. To validate this concept, we focus on finding a simple and repeatable experimental procedure, which will provide us with results sufficient to be a good indicator towards complete concept validation.

In the scope of the work and results presented in this thesis, several short-term and long-term perspectives should be considered for the future work (**Figure V-1.**).

- i) To improve the overexpression of *Galf*-ase, in order to obtain greater yield of pure enzyme, sufficient to conduct crystallization trials.*
- ii) To generate double catalytic acid/base and nucleophile residue mutants to completely attenuate *Galf*-ase hydrolytic activity and create complete neoelectins. The double mutants can be simultaneously screened as thioglycosynthases.*
- iii) To use in vitro optical imaging for testing galactofuranose-specific luminescent probe accessibility towards *Galf* substrate. Neoglycoproteins surface-labelled with *Galf* would serve for the assessment of both imaging probe and *Galf*-ase double-mutated neoelectins.*



**Figure V-1.** A graphical abstract of the main short-term and long-term perspectives for the future work.

## V. Conclusion & perspectives

L'intérêt pour le galactofuranose (Gal $f$ ) et les enzymes de transformation de ce dernier est l'objet principal ou exclusif d'intérêt de recherche dans quelques groupes scientifiques et laboratoires dans le monde. La chimie des glucides est un défi pour tout chimiste organique et il n'est pas plus facile de synthétiser des sucres furanosides moins stables, tels que le Gal $f$ . Certains d'entre eux se concentrent sur les aspects organochimiques de la synthèse des dérivés de Gal $f$ , tandis que d'autres combinent des approches chimioenzymatiques. Ce dernier point n'est pas moins difficile, car la triade d'enzymes cruciales impliquées dans la biosynthèse et le métabolisme des molécules contenant Gal $f$ , UDP-galactopyranose mutase (UGM), galactofuranosyltransférase (Gal $f$ T) et galactofuranosidase (Gal $f$ -ase), n'est pas complètement accessible et caractérisé pour un usage chimioenzymatique ou tout autre. Seules quelques-unes des enzymes répertoriées ont été identifiées et clonées, ce qui permet à la communauté scientifique de les utiliser.

Le groupe de recherche en Enzymologie et Glycobiologie est le seul de la communauté scientifique à posséder des séquences de gènes clonés codant pour deux UGM: *Escherichia coli* et *Lachnoanaerobaculum saburreum*, quatre Gal $f$ T de *Leishmania major*, LPG1, LPG1G, LPG1G et LPG1R respectivement et la seule Gal $f$ -ase de *Streptomyces spp.* Les travaux et résultats réalisés sur ces enzymes ont été publiés récemment.

A ce jour, seules quelques furanosidases, issues d'arabinofuranosidases, ont été utilisées en bioingénierie et en synthèse chimioenzymatique, il était donc très intéressant d'explorer et d'évaluer le potentiel de Gal $f$ -ase de manière originale. Par conséquent, ayant accès à cette précieuse enzyme, il était naturel de l'évaluer sous plusieurs aspects. Comme il a déjà été indiqué dans les objectifs de la thèse, il y a trois directions à explorer: **la caractérisation, le potentiel de bioingénierie et l'imagerie spécifique des cibles.**

Le sous-clonage de l'insert du gène Gal $f$ -ase a été entrepris, suivi d'une surexpression chez l'hôte procaryote et d'une purification afin de nous fournir une seule Gal $f$ -ase recombinante 6xHis-tag. Ce tag a une influence mineure sur la structure des protéines, en raison de sa taille de 1 kDa, et servait à double usage, pour la purification sélective et comme liaison à l'échafaudage d'une sonde d'imagerie. La spécificité du substrat a été déterminée par hydrolyse de plus de vingt sucres pNP différents, dont seul le pNP- $\beta$ -D-

*Galf* a été choisi comme substrat exclusif. Les conditions optimales d'activité ont démontré une bonne stabilité ainsi qu'une bonne efficacité catalytique. C'était d'une importance capitale pour l'utilisation de *Galf*-ase dans la synthèse chimioenzymatique. Dans l'ensemble, les résultats obtenus ont été encourageants et cela nous a permis de continuer sur la deuxième partie, la bioingénierie des néolectines.

L'ingénierie des néolectines à partir des glycosidases par inactivation de l'activité hydrolytique a été démontrée et rapportée dans la littérature. Étant donné que cette approche a également été utilisée pour obtenir la première furanothioglycoligase, c'était aussi notre choix de préférence. La méthode de mutagenèse dirigée a été utilisée pour muter spécifiquement le résidu acide /base catalytique avec un résidu catalytiquement inactif, afin de générer quatre mutants de la *Galf*-ase. Après une surexpression et une purification réussies de trois mutants générés sur quatre, l'atténuation de l'activité hydrolytique a été démontrée par un test Michaelis-Menten. La meilleure variante mutante, E464Q, en ce qui concerne le rendement et la cinétique, a été sélectionnée pour les essais de synthèse des furanothioglycosides. L'activité de la thioglycoligase du mutant représenté a été testée dans une série de réactions avec le *pNP*- $\beta$ -D-*Galf* comme donneur et plusieurs dérivés de thiophénol et de 2-mercaptothiazoline comme accepteurs. Les réactions ont donné huit produits thioglycosidiques différents qui ont été caractérisés par HPLC et HRMS-ESI. Cette stratégie a prouvé l'objectif d'avoir à la fois la glycosidase mutante de rétention qui peut efficacement effectuer la synthèse des *S*-galactofuranosides et agir comme néolectine en raison de la faible activité hydrolytique.

L'objectif final de la thèse, l'imagerie spécifique des structures de *Galf* glycan à la surface de certains micro-organismes, a été abordé en plusieurs étapes.

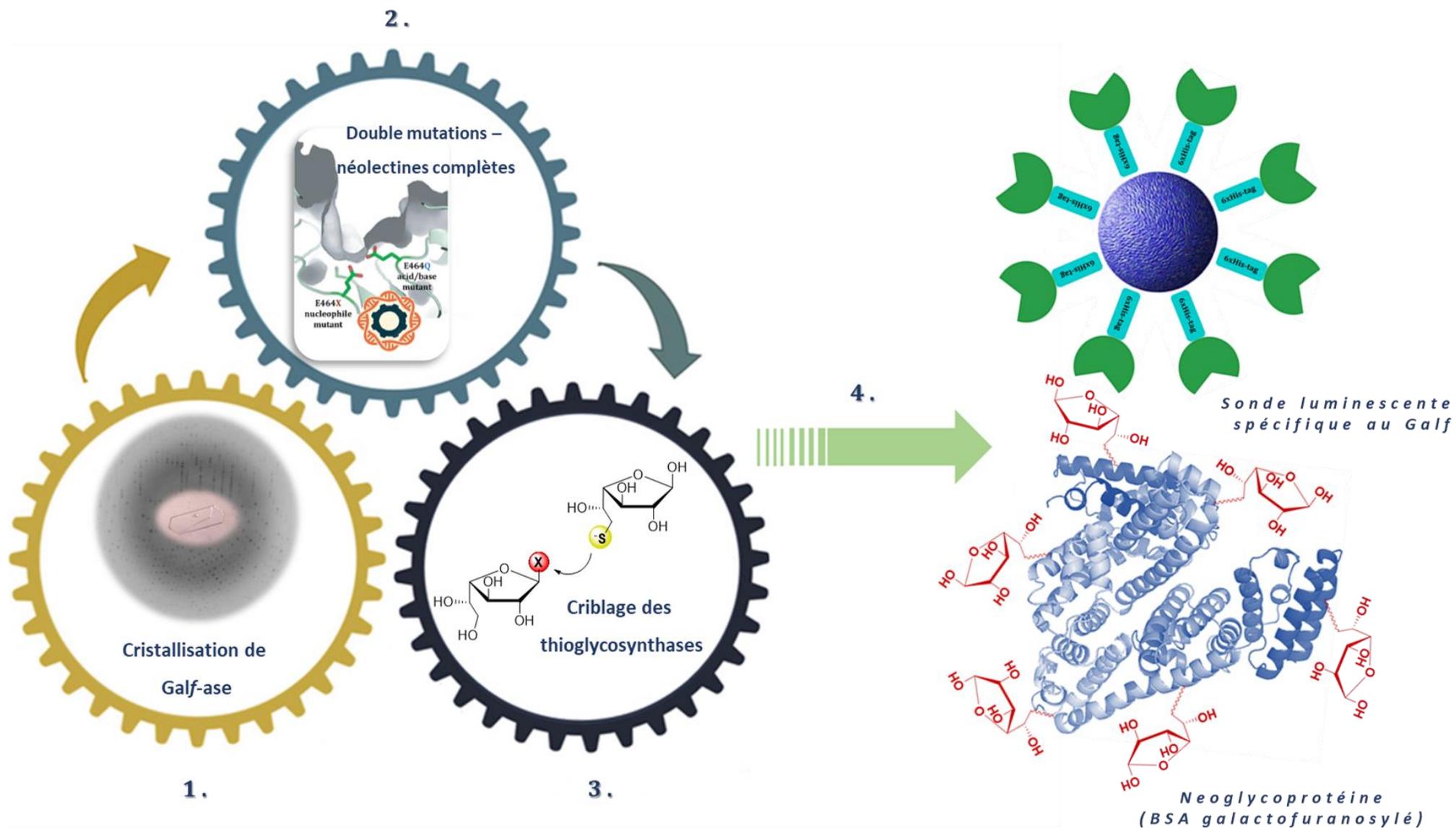
Les nanoparticules de polystyrène disponibles dans le commerce étaient le premier choix pour la construction d'un échafaudage luminescent pour la bioconjugaison de la néolectine puisque leurs cavités étaient faciles à encapsuler avec des complexes de lanthanides et que la surface carboxylée était appropriée pour une facturation ultérieure. La méthode simple et rapide d'encapsulation mécanique des complexes de lanthanides a été suivie d'un couplage covalent EDC/sulfo-NHS en deux étapes de NTA et d'un chargement avec  $\text{Ni}^{2+}$ . La stratégie de fonctionnalisation de la surface entière a été proposée pour immobiliser spécifiquement les néolectines fonctionnellement intactes

par interaction NTA-Ni<sup>2+</sup> et 6xHis-tag. La Galf-ase de type sauvage a été utilisée de première main pour étudier le protocole de bioconjugaison ainsi que l'activité enzymatique maintenue. Après incubation de NP fonctionnalisées à l'acide nitrilotriacétique et de NP chargées de Ni<sup>2+</sup> avec Galf-ase et plusieurs étapes de lavage, l'activité de la Galf-ase restante a été évaluée par un test d'activité enzymatique à base de pNP décrit précédemment. Ceci a montré que la Galf-ase conservait son activité enzymatique lors de la bioconjugaison et nous a permis de l'utiliser comme sonde d'imagerie.

Ces travaux préparaient la finalisation de l'objectif principal sous la forme d'un protocole durable pour l'atteindre. Pour valider ce concept, nous nous sommes concentrés sur la recherche d'une procédure expérimentale simple et reproductible, qui nous fournira des résultats suffisants pour être un bon indicateur vers une validation complète du concept.

Dans le cadre des travaux et des résultats présentés dans cette thèse, plusieurs perspectives à court et à long terme devraient être envisagées pour les travaux futurs (**Figure V-1**).

- i) Améliorer la surexpression de Galf-ase, afin d'obtenir un meilleur rendement en enzyme pure, suffisant pour mener des essais de cristallisation.*
- ii) Générer des mutants d'acide catalytique/base double et des mutants de résidus nucléophiles pour atténuer complètement l'activité hydrolytique de Galf-ase et créer des néoélectines complètes. Les doubles mutants peuvent être simultanément criblés comme thioglycosynthases.*
- iii) Utiliser un substrat d'imagerie in vitro pour tester l'accessibilité des sondes luminescentes spécifiques au galactofuranose vers le substrat Galf. Les néoglycoprotéines marquées en surface avec Galf serviraient à l'évaluation de la sonde d'imagerie et des néoélectines à double mutation Galf-ase.*



**Figure V-1.** Un résumé graphique des principales perspectives à court et à long terme pour les travaux futurs.

## **CHAPTER VI**

### **Experimental part**

<b>6.1. Materials and methods</b> .....	<b>117</b>
<b>6.2. Molecular biology</b> .....	<b>117</b>
6.2.1. <i>In silico</i> sequence analysis.....	117
6.2.2. Polymerase chain reaction (PCR).....	117
6.2.3. DNA analysis and purification.....	121
6.2.4. Bacterial strains and plasmids.....	122
<b>6.3. Biochemistry</b> .....	<b>123</b>
6.3.1. Protein overexpression.....	123
6.3.2. Protein analysis.....	123
6.3.3. Protein purification.....	125
6.3.4. Enzymatic activity assays.....	126
6.3.5. Furanthioglycosylation.....	130
<b>6.4. Coordination chemistry</b> .....	<b>131</b>
6.4.1. Synthesis of Yb <sup>III</sup> /Zn <sup>II</sup> metallacrowns: LnZn <sub>16</sub> (quinoHA) <sub>16</sub> .....	131
<b>6.5. Nanotechnology</b> .....	<b>132</b>
6.5.1. Encapsulation of Yb <sup>III</sup> /Zn <sup>II</sup> metallacrowns (LnZn <sub>16</sub> (quinoHA) <sub>16</sub> ) in polystyrene nanoparticles: LnX@PS/COOH.....	132
6.5.2. Surface modification LnX@PS/COOH to PS/NTA and LnX@PS/NTA.....	132
6.5.3. Synthesis of PS/NTA-Ni <sup>2+</sup> and LnX@PS/NTA-Ni <sup>2+</sup> .....	133
6.5.4. Bioconjugation of 6xHis-tagged recombinant GalF-ase to PS/NTA-Ni <sup>2+</sup> and LnX@PS/NTA-Ni <sup>2+</sup> .....	133
6.5.5. Characterization of PS/NTA-Ni <sup>2+</sup> & LnX@PS/NTA-Ni <sup>2+</sup> nanoparticles.....	134

## 6.1. Materials and methods

Laboratory chemicals used for preparation of compounds or as the starting compounds, reagents and solvents were analytical grade purity and commercially available. The ultrapure water, purified by Veolia or Milli-Q® water purification system, was used for all solutions and buffer preparation and those used for bacterial culture growth were sterilised by autoclaving prior to use. All instruments used for manipulating and maintaining bacterial cultures were done according to aseptic laboratory techniques. Insert gene sequences, wild type and GalF-ase mutants, obtained by site-directed mutagenesis or subcloning were confirmed by custom deoxyribonucleic acid (DNA) sequencing (Eurofins Genomics) before applying in further experimental steps.

## 6.2. Molecular biology

### 6.2.1. *In silico* sequence analysis

GalF-ase nucleotide and amino acid sequences were retrieved from National Center for Biotechnology Information (NCBI) GenBank®<sup>3</sup> database under accession number LC073694. Amino acid sequence alignment search was conducted using a protein query BLAST® (Basic Local Alignment Search Tool) (NCBI)<sup>4</sup> in Protein Data Bank (PDB) database. The deduced protein amino acid sequences, deposited under PDB accession codes 4YPJ (*B. circulans*, 26% identities), 3CMG (*B. fragilis*, 24% identities) and 3GM8 (*B. vulgatus*, 25% identities), were aligned and analysed with CLC Sequence Viewer 7.7 software.

### 6.2.2. Polymerase chain reaction (PCR)

PCR was used to amplify whole plasmid constructs DNA or specific regions of template DNA for gene sequence confirmation and site-directed mutagenesis using custom oligonucleotide primers (Eurofins Genomics, Ebersberg, Germany).

---

<sup>3</sup> <http://www.ncbi.nlm.nih.gov/genbank/>

<sup>4</sup> <http://blast.ncbi.nlm.nih.gov/Blast.cgi>

**Gene amplification**

Wild type *Galf*-ase insert gene from pET-50b(+) plasmid construct was amplified with PCR using *Thermo Scientific™ DreamTaq Green* PCR Master Mix and custom primers (**Table VI-1.**). Reaction was performed in thermocycler (Esco Swift MiniPro Thermal Cycler) at conditions (**Table VI-2.**) according to supplier's instructions.

**Table VI-1.** Primers used for wild type *Galf*-ase gene amplification.

Primer	Primer length/bp	Primer orientation	Primer sequence (5'→3')
T7	25	Forward	CGAAATTAATACGACTCACTATAGG
T7term	15	Reverse	TAGAGGCCCCAAGGG

**Table VI-2.** Cycling parameters for wild type *Galf*-ase gene amplification.

Step	Cycle element	Temperature/°C	Time/min:sec	Cycles
1	Initial denaturation	95	3:00	1
2	Denaturation	95	0:30	40
3	Amplification Annealing	45.6	0:30	
4	Extension	72	2:00	
5	Final extension	72	10:00	1

**Sequencing**

Plasmid construct pET-28a(+) complete double strand gene sequence, containing wild type or mutant *Galf*-ase insert gene, was confirmed using custom oligonucleotide primers (**Table VI-3.**) and was done commercially by Eurofins Genomics.

**Table VI-3.** Primers used for plasmid construct pET-28a(+) gene sequencing.

Primer	Primer length/bp	Primer orientation	Primer sequence (5'→3')
T7	25	Forward	CGAAATTAATACGACTCACTATAGG
<i>Galf</i> -ase MS	20	Forward	GTGCCGTTTCGTCGCAGAGTC
<i>Galf</i> -ase <i>intra</i>	20	Forward	CTCTTGAAGTCCTCTTTCAG
T7term	15	Reverse	TAGAGGCCCCAAGGG

**Site-directed mutagenesis**• **Substitution**

Several Galf-ase catalytic acid/base residue mutants were obtained by site-directed mutagenesis method. Reaction was performed using Agilent Technologies QuikChange Lightning Site-Directed Mutagenesis Kit. Mutagenic primers (**Table VI-4.**) were designed with Agilent Technologies QuikChange Primer Design tool and reaction was performed in thermocycler with pET-28a(+) plasmid construct, containing wild type Galf-ase insert gene, as the template DNA. Reaction conditions (**Table VI-5.**) and the residual parent plasmid DNA digestion with *Dpn* I restriction enzyme was performed according to supplier's instructions while the heat-shock transformation of cloning competent cells was performed according to previously described protocol.

**Table VI-4.** Primers used for site-directed mutagenesis.

Gene mutation	Primer length/bp	Primer orientation	Primer sequence (5'→3') <sup>a</sup>
E464Q	25	Forward	GTCAACCAGAACC <u>CAG</u> GGCTGGGGCC
		Reverse	GGCCCCAGCC <u>CTG</u> GTCTGGTTGAC
E464A	27	Forward	CTGGCCCCAGCC <u>CGC</u> GTCTGGTTGAC
		Reverse	GTCAACCAGAACC <u>GCG</u> GGCTGGGGCCAG
E464C	35	Forward	CGTACTGGCCCCAGCC <u>GCA</u> GTCTGGTTGACCCAC
		Reverse	GTGGGTCAACCAGAAC <u>TGC</u> GGCTGGGGCCAGTACG
E464S	35	Forward	CGTACTGGCCCCAGCC <u>GCT</u> GTCTGGTTGACCCAC
		Reverse	GTGGGTCAACCAGAAC <u>AGC</u> GGCTGGGGCCAGTACG

<sup>a</sup> Underlined sequences indicate the substituted codons

**Table VI-5.** Cycling parameters for site-directed mutagenesis.

Step	Cycle element	Temperature/°C	Time/min:sec	Cycles
1	Initial denaturation	95	2:00	1
2	Denaturation	95	0:20	18
3	Amplification Annealing	60	0:10	
4	Extension	68	4:50	
5	Final extension	68	5:00	1

- **Deletion**

Deletion mutagenesis reaction of EcoRI restriction site was performed using Agilent Technologies QuikChange Lightning Site-Directed Mutagenesis Kit. Mutagenic primers (**Table VI-6.**) were designed with Agilent Technologies QuikChange Primer Design tool and reaction was performed in thermocycler with pET-28a(+) plasmid construct, containing wild type or mutant *Galf*-ase insert gene, as the template DNA. Reaction conditions (**Table VI-7.**) and the residual parent plasmid DNA digestion with *Dpn* I restriction enzyme was performed according to supplier's instructions while the heat-shock transformation of cloning competent cells was performed according to previously described protocol.

**Table VI-6.** Primers used for deletion mutagenesis.

Primer	Primer length/bp	Primer orientation	Primer sequence (5'→3')
<i>Galf</i> -EcoRI $\Delta$	32	Forward	AATGGGTCGCGGATCCAATTCTATGCGTCGAG
<i>Galf</i> -EcoRI $\Delta$	32	Reverse	CTCGACGCATAGAATTGGATCCGCGACCCATT

**Table VI-7.** Cycling parameters for deletion mutagenesis reaction.

Step	Cycle element	Temperature/°C	Time/min:sec	Cycles
1	Initial denaturation	95	2:00	1
2	Denaturation	95	0:20	
3	Amplification	60	0:10	18
4	Extension	68	4:20	
5	Final extension	68	5:00	1

### **Subcloning**

Wild type (E464E) *Galf*-ase gene was subcloned from pET-50b(+) plasmid construct to pET-28a(+) vector in several steps including insert gene preparation, vector preparation, ligation reaction and competent bacteria transformation. Plasmid construct pET-50b(+) was digested with EcoRI and HindIII restriction enzymes according to the supplier's instructions (Thermo Scientific™ FastDigest™ restriction enzymes). The *Galf*-ase gene fragment was separated by agarose-gel electrophoresis from restriction enzyme double digestion reaction mixture, extracted from the low-melting point agarose-gel (1%) and

purified according to supplier's instructions (Thermo Scientific GeneJET Gel Extraction and DNA Cleanup Micro Kit). Vector pET-28a(+) was simultaneously double digested with EcoRI and HindIII restriction enzymes and dephosphorylated with alkaline phosphatase according to supplier's instructions (Thermo Scientific™ FastDigest™ restriction enzymes and FastAP™ Thermosensitive Alkaline Phosphatase). Linearized and dephosphorylated pET-28a(+) vector was purified from reaction mixture according to supplier's instructions (Thermo Scientific™ GeneJET PCR Purification Kit). Ligation reaction was performed with Thermo Scientific Rapid DNA Ligation Kit in thermocycler (Esco Swift MiniPro Thermal Cycler) at 22°C for 10 min and used for cloning strain transformation according to previously described protocol.

### **6.2.3. DNA analysis and purification**

#### ***Plasmid DNA isolation***

Plasmid DNA was prepared from single positive colony inoculated in lysogeny broth (LB) (5 mL) and supplemented with corresponding antibiotic (30 µg/mL kanamycin). After overnight incubation at 37 °C in shaker incubator (250 rpm) plasmid DNA was isolated from bacterial culture and purified according to supplier's instructions (Thermo Scientific™ GeneJET Miniprep Kit).

#### ***Restriction enzyme analysis***

Presence of insert genes, wild type and mutant Gal $\beta$ -ase, in plasmid constructs, pET-50b(+) and pET-28a(+) respectively, was confirmed by restriction enzyme double digestion with appropriate restriction enzymes according to the supplier's instructions (Thermo Scientific™ FastDigest™ restriction enzymes).

#### ***DNA quantification and purity***

Microplate (Thermo Scientific™ µDrop™) was used for DNA sample preparation and spectrophotometer, (Thermo Scientific™ Multiskan™ GO) for photometric measurement (260 nm). Spectral data, quantification and qualification were analysed (Thermo Scientific™ SkanIt Software) and represented as concentration (ng/µL).

### ***Agarose-gel electrophoresis***

Plasmid DNA, PCR products and restriction enzyme digests were separated by agarose-gel electrophoresis. Agarose (1% w/v) was dissolved by heating in 1xTAE buffer (40 mM Tris, 20 mM Acetic acid, 0.1 mM EDTA), DNA gel stain was added (2  $\mu$ L Invitrogen SYBR<sup>®</sup> Safe) and poured into a mould containing the requisite number of wells. DNA samples were diluted (5:1) with loading dye (70% H<sub>2</sub>O, 30% Glycerol, 0.25% Bromophenol blue) and loaded into the wells of the 1xTAE buffer submerged, solid agarose gel. All samples sizes were compared to Thermo Scientific GeneRuler<sup>™</sup> DNA ladder mix depending on the estimated size of the products. Electrophoresis was conducted in the Embi Tec RunOne<sup>™</sup> electrophoresis system at constant voltage (100 V) for approximately 20 min. DNA was visualized under UV-transilluminator (254 nm) and photographic records of gels were taken.

#### **6.2.4. Bacterial strains and plasmids**

Two strains of *Escherichia coli* competent cells were used. *E. coli* DH5 $\alpha$  (Novagen<sup>®</sup>) was used as cloning strain and Rosetta<sup>™</sup> (DE3) (Novagen<sup>®</sup>) as expression strain. Plasmid construct pET-50b(+), encoding wild type GalF-ase gene from *Streptomyces spp.*, was kindly provided by Prof. Kaoru Takegawa, Department of Bioscience and Biotechnology, Faculty of Agriculture, Kyushu University, Japan. Plasmid vector pET-28a(+) (Novagen<sup>®</sup>) was used as destination vector for GalF-ase wild type insert gene subcloning.

#### ***Competent bacteria transformation***

Competent bacterial cells were removed from -80°C, thawed on ice and plasmid DNA (50-100 ng) or ligation product (5  $\mu$ L) was added to aliquot (50  $\mu$ L) of competent cells and let on ice for 20 min, followed by heat-shock at water bath at 42°C for 90 seconds and incubation on ice for 2 min. The cells were let to recover in LB (450  $\mu$ L) and incubated in shaking incubator (250 rpm) at 37°C for an hour. Afterwards aliquot (100  $\mu$ L) of transformed cells was plated on selective, corresponding antibiotic (34  $\mu$ g/mL chloramphenicol and/or 30  $\mu$ g/mL kanamycin) containing LB agar Petri plates and incubated overnight at 37°C.

### 6.3. Biochemistry

#### 6.3.1. Protein overexpression

##### *Protein production*

One single positive Rosetta expression strain colony, previously transformed with plasmid DNA or ligation product according to previously described protocol, was inoculated in LB broth (5 mL) supplemented with corresponding antibiotics (34 µg/mL chloramphenicol and 30 µg/mL kanamycin) and shaker incubated (250 rpm) overnight at 37°C. That was used for preculture preparation (30 mL LB broth, corresponding antibiotics, overnight incubation at 37 °C, shaking 250 rpm). This preculture (10 mL) was inoculated into LB broth (1 L) containing corresponding antibiotics and shaker incubated (250 rpm) at 37°C. Cells were grown to mid-exponential phase (OD<sub>600</sub>: 0.6), shortly cooled on ice, protein expression was induced by adding β-D-thiogalactopyranoside (IPTG) (100 µL; 1 M) and left on shaker incubator (250 rpm) at 15°C for 16 hours. Cells were harvested by centrifugation (4255 g, 30 min, 4 °C), supernatant discharged and the cell pellets were resuspended in lysis buffer (10 mL; 100 mM NaCl, 50 mM Tris/HCl pH 8). Cell suspension was centrifuged (6000 g, 10 min, 4 °C), supernatant discharged and cell pellet stored at -20 °C.

#### 6.3.2. Protein analysis

##### *Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE electrophoresis)*

Protein preparations were separated by SDS-PAGE electrophoresis using stacking gel (4% w/v) and separating gel (8% w/v) in 1x running buffer (0.192 M Glycine, 0.025 M Tris, 0.1% SDS). The samples were diluted (4:1) in Laemmli sample buffer (50% Glycerol, 10% β-mercaptoethanol, 10% SDS, 0.05% Bromophenol blue, 313 mM Tris/HCl pH 6.8), boiled at 95°C for 5 min, centrifuged (10 000 g, 5 min, room temperature) and loaded into the wells. Electrophoresis was conducted using Bio-Rad Mini-PROTEAN® Electrophoresis System at constant voltage (250 V) and current (25 mA). All samples sizes were compared to Thermo Scientific PageRuler™ Prestained Protein Ladder. Protein were fixed and visualized on the gel by incubating for approximately 30 min in a staining solution (20% Ethanol, 10% Acetic acid, 1 mg/mL Coomassie Brilliant Blue G250). Gel was destained by heating in water until bands appeared visible. Scan records of the gels were taken.

***In-gel digestion***

Protein preparations were separated by SDS-PAGE electrophoresis, fixed and visualized according to previously described protocol. Upon electrophoresis, protein bands of interest were excised from the gel with a clean scalpel and in-gel digestion procedure was performed according to the Thermo Scientific In-Gel Tryptic Digestion Kit instructions. An aliquot of the protein digest was further analysed by peptide mass fingerprinting, using matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF/TOF) instrument, performed by The Facility for Mass Spectrometry and Proteomics at the Center for Molecular Biophysics (CBM) platform (FR2708). The obtained data were submitted to database search by Mascot software (Matrix Science Inc.).

***Mass spectrometry***

Previously prepared aliquot (200  $\mu$ L; 4 mg/mL, 20 mM Tris) of pure protein is used to determine the molecular weight using matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS). The analysis was performed by The Facility for Mass Spectrometry and Proteomics at the Center for Molecular Biophysics (CBM) platform (FR2708).

***Protein concentration***

Protein concentration was determined using colorimetric Bio-Rad protein assay based on the Bradford dye-binding method. A standard calibration curve samples were made with a series of six bovine serum albumin (BSA) (1 mg/mL) dilutions (0x, 2x, 2.5x, 3.3x, 5x, 10x, 20x). Five protein dilution samples (0x, 2x, 4x, 10x, 20x, 40x) and non-diluted sample were prepared. Standard solutions and protein samples (10  $\mu$ L) were transferred in 96-well microtiter plate and Bio-Rad protein reagent was added (200  $\mu$ L; 1x). Reaction mixture was briefly stirred, incubated for 5 min at room temperature and absorbance measured (595 nm) was using spectrophotometer (Thermo Scientific™ Multiskan™ GO). Protein concentration (mg/ $\mu$ L) was calculated from standard linear range curve data plot obtained from standard solutions and samples absorbance measurement results.

### 6.3.3. Protein purification

#### *Cell lysis*

Cell pellets were resuspended in solution containing lysis buffer (1:10 v/v; 100 mM NaCl, 50 mM Tris/HCl pH 8, 1 mM phenylmethanesulfonyl fluoride (PMSF), 5% glycerol, 0.1% Triton X-100, 1 mg/L lysozyme). Suspension was incubated on magnetic stirrer for 20 min at 4°C, transferred into tubes, frozen in liquid nitrogen for 5 min at -196°C and thawed in water bath at 37°C. This freeze and thaw procedure were repeated three times. Sonication ([2 min on/2 min pause]x3, 50% cycle) was performed in metal beaker on ice, repeated three times. Lysate was centrifuged (30 000 g, 20 min, 4°C), supernatant collected, filtered (0.45 µm pore filter) and purified by affinity chromatography.

#### *Affinity chromatography*

Protein samples were purified manually or automatically, using an ÄKTA Prime Plus FPLC unit (GE Healthcare Life Sciences) imidazole elution gradient mode, on Thermo Scientific HisPur™ Ni-NTA Chromatography Cartridge (1 mL). Chromatography cartridge was first equilibrated with buffer A (10 mL; 50 mM Tris, 200 mM NaCl, 10 mM Imidazole), followed by loading protein sample and washing with buffer A (20 mL). Bound proteins were eluted in imidazole gradient buffers B, C, D, E (10 mL; 50 mM Tris, 200 mM NaCl, 25, 50, 100, 500 mM Imidazole) and fractions were collected (2 mL). Chromatography cartridge was washed with water (10 mL), ethanol (10 mL; 20%) and stored at 4°C.

#### *Size exclusion chromatography*

Proteins were analysed by size exclusion chromatography using a Superdex™ 200 10/300 GL (GE Healthcare Life Sciences) size exclusion column affixed to an ÄKTA Prime Plus FPLC unit (GE Healthcare Life Sciences). Protein sample was concentrated by ultrafiltration (30 000 MWCO, Sartorius Vivaspin® sample concentrator) and loaded (500 µL) onto the column equilibrated with buffer (50 mM Tris, 200 mM NaCl, pH 8.0) and eluted at a rate of 0.3 ml/min. Fractions (200 µL) were collected and UV-absorption (280 nm) monitored.

### 6.3.4. Enzymatic activity assays

#### *pNP-linked monosaccharide substrates hydrolysis activity screening*

The enzyme-catalysed hydrolysis assay of twenty-one *pNP*-linked monosaccharide substrates (*pNP*- $\alpha$ -L-Ara, *pNP*- $\beta$ -L-Ara, *pNP*- $\alpha$ -L-Araf, *pNP*- $\alpha$ -L-Fuc, *pNP*- $\beta$ -L-Fuc, *pNP*- $\beta$ -D-Fuc, *pNP*- $\alpha$ -D-Gal, *pNP*- $\beta$ -D-Gal, *pNP*- $\beta$ -D-Galf, *pNP*- $\beta$ -D-GalNAc, *pNP*- $\alpha$ -D-Glc, *pNP*- $\beta$ -D-Glc, *pNP*- $\alpha$ -D-GlcNAc, *pNP*- $\beta$ -D-GlcNAc, *pNP*- $\beta$ -D-GlcA, *pNP*- $\alpha$ -L-Rha, *pNP*- $\beta$ -D-Ribf, *pNP*- $\alpha$ -D-Man, *pNP*- $\beta$ -D-Man, *pNP*- $\alpha$ -D-Xyl, *pNP*- $\beta$ -D-Xyl) (**Appendix 7.1.2. Figure VII-2.**) was performed in 96-well microtiter plate in a reaction mixture (50  $\mu$ L) containing purified enzyme (5  $\mu$ L; 0.053 mg/mL), *pNP* monosaccharide substrates (5  $\mu$ L; 10 mM), buffer (8  $\mu$ L; 0.1 M Citric acid/0.2 M Na<sub>2</sub>HPO<sub>4</sub> pH 4.5) and water (32  $\mu$ L). Non-enzyme samples for each *pNP* monosaccharide substrate were taken as controls for the spontaneous hydrolysis of the substrate and performed simultaneously with enzyme-containing reaction mixtures. After 20 min of incubation at 30°C, reaction was terminated (150  $\mu$ L; 1 M Na<sub>2</sub>CO<sub>3</sub>), absorbance (405 nm;  $\epsilon_{405} = 19\,500\text{ M}^{-1}\text{cm}^{-1}$ ) of released *pNP* was measured using spectrophotometer (Thermo Scientific™ Multiskan™ GO). All the reactions were assayed in triplicate and absorbance values were corrected for the spontaneous hydrolysis of the substrate.

#### *Validation of Linearity assay*

To determine the reaction velocity and to evaluate enzyme activity within the linear range, reaction mixture (1 mL) containing purified enzyme (20  $\mu$ L; 0.13 mg/mL), *pNP*- $\beta$ -D-Galf substrate (100  $\mu$ L; 10 mM), buffer (100  $\mu$ L; 0.1 M Citric acid/0.2 M Na<sub>2</sub>HPO<sub>4</sub> pH 5) and water (780  $\mu$ L) was prepared. Reaction was incubated at 37°C (water bath), and after 10, 20, 30, 45, 60 min aliquots of reaction (150  $\mu$ L) were taken, reaction terminated (75  $\mu$ L; 1 M Na<sub>2</sub>CO<sub>3</sub>), aliquoted (200  $\mu$ L) and absorbance (405 nm;  $\epsilon_{405} = 19\,500\text{ M}^{-1}\text{cm}^{-1}$ ) of released *pNP* was measured using spectrophotometer (Thermo Scientific™ Multiskan™ GO). All the reactions were assayed in triplicate and absorbance values were corrected for the spontaneous hydrolysis of the substrate.

***Effect of pH on enzyme activity***

The effect of pH on enzyme activity was tested by varying reaction mixture pH for 0.5 pH units from 2 to 9.5 using different pH adjusted buffer systems (0.1 M Citric acid/HCl pH 2; 0.1 M Citric acid/0.2 M Na<sub>2</sub>HPO<sub>4</sub> pH 2.5 – 6.5; 0.1 M Tris/HCl pH 7 – 8; 0.1 M Tris/NaOH, pH 8.5; 0.1 M Glycine/NaOH pH 9 – 9.5). Reaction (50 µL), containing purified enzyme (5 µL; 0.05 mg/mL), pNP-β-D-Galf substrate (5 µL; 10 mM), buffer (8 µL; pH 2 – 9.5) and water (32 µL), was started by addition of the enzyme and after 20 min of incubation at 30°C (water bath), reaction was terminated (150 µL; 1 M Na<sub>2</sub>CO<sub>3</sub>). Reaction was performed in 96-well microtiter plate and absorbance (405 nm;  $\epsilon_{405} = 19\,500\text{ M}^{-1}\text{cm}^{-1}$ ) of released pNP was measured using spectrophotometer (Thermo Scientific™ Multiskan™ GO). All the reactions were assayed separately in each buffer system, in triplicate and absorbance values were corrected for the spontaneous hydrolysis of the substrate.

***Effect of temperature on enzyme activity***

The effect of temperature on enzyme activity was tested by incubating reaction mixture at various temperatures ranging from 10 to 80°C. The reactions (50 µL), containing purified enzyme (5 µL; 0.05 mg/mL), pNP-β-D-Galf substrate (5 µL; 10 mM), buffer (8 µL; 0.1 M Citric acid/0.2 M Na<sub>2</sub>HPO<sub>4</sub> pH 4.5) and water (32 µL), were performed in PCR reaction tubes and incubated 20 min in thermocycler (Esco Swift MiniPro Thermal Cycler). Reaction was terminated (100 µL; 1 M Na<sub>2</sub>CO<sub>3</sub>) and aliquots of reaction mixtures (200 µL) were taken, transferred in 96-well microtiter plate and absorbance (405 nm;  $\epsilon_{405} = 19\,500\text{ M}^{-1}\text{cm}^{-1}$ ) of released pNP was measured using spectrophotometer (Thermo Scientific™ Multiskan™ GO). All the reactions were assayed separately at different temperatures, in triplicate and absorbance values were corrected for the spontaneous hydrolysis of the substrate.

***Effect of freeze-thaw***

Three samples, containing previously purified enzyme (50 µL; 0.235 mg/mL) different concentrations of glycerol (0 %, 10%, 20%) and water, if required up to solution (100 µL), were prepared. The samples were precooled on ice (30 min), frozen at -20°C and after four days were removed from -20°C, thawed on ice (30 min) and enzymatic activity of sample aliquot (10 µL) was assayed in reaction mixture (200 µL) containing pNP-β-D-Galf

substrate (20  $\mu\text{L}$ ; 10 mM), buffer (35  $\mu\text{L}$ ; 0.1 M Citric acid/0.2 M  $\text{Na}_2\text{HPO}_4$  pH 4.5) and water (135  $\mu\text{L}$ ) at 37°C (water bath), during 10 min, according to previously described enzymatic activity protocol. This predefined freeze and thaw procedure and activity assessment were repeated in four cycles within four days interval. Activity recovery was expressed as a percentage of the activity prior to freezing. All the reactions were assayed separately in duplicate and absorbance values were corrected for the spontaneous hydrolysis of the substrate.

### ***Characterisation of kinetic parameters – Michaelis-Menten kinetics***

For the enzyme kinetic characterisation corresponding volume of an enzyme (2  $\mu\text{L}$ ; 0.13 mg/mL), a *pNP- $\beta$ -D-Galf* substrate in different concentration ranges (0.01 mM, 0.02 mM, 0.03 mM, 0.05 mM, 0.08 mM, 0.1 mM, 0.2 mM, 0.3 mM, 0.5 mM, 0.8 mM, 1 mM, 2 mM, 3 mM, 5 mM), buffer (20  $\mu\text{L}$ ; 0.1 M Citric acid/0.2 M  $\text{Na}_2\text{HPO}_4$  pH 4.5) and water, if required up to reaction mixture (200  $\mu\text{L}$ ), was added. The reaction was started by addition of the enzyme and after 20 min of incubation at 37°C (water bath), reaction was terminated (100  $\mu\text{L}$ ; 1 M  $\text{Na}_2\text{CO}_3$ ). Aliquots of reaction mixtures (200  $\mu\text{L}$ ) were taken, transferred in 96-well microtiter plate and absorbance (405 nm;  $\epsilon_{405} = 19\,500\text{ M}^{-1}\text{cm}^{-1}$ ) of released *pNP* was measured using spectrophotometer (Thermo Scientific™ Multiskan™ GO). All the reactions were assayed in triplicate and absorbance values were corrected for the spontaneous hydrolysis of the substrate. The kinetic parameters ( $K_M$ ,  $V_m$ ,  $k_{\text{cat}}$ ) were calculated using GraphPad Prism 5 software (GraphPad Software, San Diego, CA, USA).

### ***Inhibitor activity screening assay***

The effect of inhibitors **1** – **4** (15  $\mu\text{L}$ ; 10 mM) on enzyme activity was tested in reaction mixture (150  $\mu\text{L}$ ) containing purified enzyme (11.2  $\mu\text{L}$ ; 0.067 mg/mL), *pNP- $\beta$ -D-Galf* substrate (3.75  $\mu\text{L}$ ; 10 mM), buffer (25  $\mu\text{L}$ ; 0.1 M Citric acid/0.2 M  $\text{Na}_2\text{HPO}_4$  pH 4.5) and water (95.5  $\mu\text{L}$ ). After 20 min of incubation at 37°C (water bath), reaction was terminated (100  $\mu\text{L}$ ; 1 M  $\text{Na}_2\text{CO}_3$ ). The inhibitors **1** and **4** were dissolved in 10% dimethyl sulfoxide (DMSO) and the final concentration of DMSO in the reaction medium was 1%. Aliquots of reaction mixtures (200  $\mu\text{L}$ ) were taken, transferred in 96-well microtiter plate and absorbance (405 nm;  $\epsilon_{405} = 19\,500\text{ M}^{-1}\text{cm}^{-1}$ ) of released *pNP* was measured using spectrophotometer (Thermo Scientific™ Multiskan™ GO). Control enzyme activity

reactions, without inhibitors, and those containing DMSO at concentration corresponding to reaction's, were carried out. All the reactions were assayed in triplicate and absorbance values were corrected for the spontaneous hydrolysis of the substrate and tested inhibitors.

#### ***The half-maximal inhibitory concentration assay – IC<sub>50</sub> assay***

The IC<sub>50</sub> assay was performed in the presence of different concentration ranges of inhibitors **1** – **4** (0.1 mM, 0.2 mM, 0.3 mM, 0.5 mM, 0.8 mM, 1 mM, 2 mM, 3 mM, 5 mM, 7 mM), *p*NP-β-D-Galf substrate (3.75 μL; 10 mM), buffer (20 μL; Citric acid/Na<sub>2</sub>HPO<sub>4</sub>, pH 4.5) and water. The reaction (150 μL) was started by addition of the enzyme (11.2 μL; 0.067 mg/mL), and after 20 min of incubation at 37°C (water bath), reaction was terminated (100 μL; 1 M Na<sub>2</sub>CO<sub>3</sub>). Aliquots of reaction mixtures (200 μL) were taken, transferred in 96-well microtiter plate and absorbance of released *p*NP was measured (405 nm; ε<sub>405</sub> = 19 500 M<sup>-1</sup>cm<sup>-1</sup>) using spectrophotometer (Thermo Scientific™ Multiskan™ GO). Control enzyme activity reactions, without inhibitors, and those containing DMSO at concentration corresponding to reaction's (1%), were carried out. All the reactions were assayed in triplicate and absorbance values were corrected for the spontaneous hydrolysis of the substrate. The IC<sub>50</sub> value for each inhibitor was calculated using GraphPad Prism 5 software (GraphPad Software, San Diego, CA, USA).

#### ***Competitive inhibition assay***

The type of inhibition **2** was tested in the presence of different concentration ranges of *p*NP-β-D-Galf substrate (0.05 mM, 0.1 mM, 0.25 mM, 0.5 mM, 1 mM), different concentrations of **2** (0.1 mM, 0.2 mM, 0.4 mM, 1 mM), buffer (25 μL; 0.1 M Citric acid/0.2 M Na<sub>2</sub>HPO<sub>4</sub> pH 4.5) and enzyme (11.2 μL; 0.067 mg/mL). The final concentrations of **2** were altered in four separate reactions (150 μL) which were terminated (100 μL; 1 M Na<sub>2</sub>CO<sub>3</sub>) after 20 min of incubation at 37°C (water bath). Aliquots of reaction mixtures (200 μL) were taken, transferred in 96-well microtiter plate and absorbance (405 nm; ε<sub>405</sub> = 19 500 M<sup>-1</sup>cm<sup>-1</sup>) of released *p*NP was measured using spectrophotometer (Thermo Scientific™ Multiskan™ GO). Control enzyme activity reactions, without **2**, were carried out. All the reactions were assayed in duplicate and absorbance values were corrected for the spontaneous hydrolysis of the substrate. The type of the inhibition and *K<sub>i</sub>* value for **2**

were calculated using GraphPad Prism 5 software (GraphPad Software, San Diego, CA, USA).

### 6.3.5. Furanothioglycosylation

Chemoenzymatic synthesis of furanothioglycosides was performed in the preparative-scale (2 mL) in the presence of *p*NP- $\beta$ -D-Galf glycosyl donor (250  $\mu$ L; 10 mM, 1 eq.), thiophenol and 2-mercaptothiazoline derivatives and 7-mercapto-4-methylcoumarine as acceptors (250  $\mu$ L; methanol, 0.1 M, 10 eq.), 1,4-dithiothreitol (DTT) (250  $\mu$ L; 0.1 M, 10 eq.), buffer (400  $\mu$ L; citric acid/Na<sub>2</sub>HPO<sub>4</sub>, pH 4.5), water (783  $\mu$ L) and E464Q Galf-ase (67  $\mu$ L, 2.7 mg/mL). Reactions were incubated at 37°C (water bath) for 14 h. Transglycosylation products were analyzed by reverse-phase high performance liquid chromatography (HPLC) (Agilent Technologies 1220 Infinity II LC) performed on a ZORBAX Eclipse XDB-C18 column (4.6x50 mm, 3.5  $\mu$ m, Agilent Technologies, USA). The samples and standards were separated with linear gradient elution in the mobile phase delivered at a rate of 1 mL/min with solvents A (0.1% formic acid (v/v) in water) and B (0.1% formic acid (v/v) in acetonitrile). The column effluent was monitored at 250 nm. The products were purified by reversed phase chromatography (Thermo Scientific™ HyperSep™ C18 Cartridges) in appropriate solvent system for each product (**Appendix 7.1.3. Figure VII-3. – Figure VII-10.**), fractions evaporated under reduced pressure (Büchi® Rotavapor® R II) and submitted to structural analysis. The high-resolution electrospray ionization mass spectrometry (HRMS-ESI) (Bruker maXis II™) was performed by the Research Federation at the Institute of Organic and Analytical Chemistry (ICOA) platform (FR2708).

## 6.4. Coordination chemistry

### 6.4.1. Synthesis of Yb<sup>III</sup>/Zn<sup>II</sup> metallacrowns: YbZn<sub>16</sub>(quinoHA)<sub>16</sub>

*Synthesis was done by Dr. Tu N. Nguyen (Department of Chemistry, University of Michigan, USA) following the procedures described below.*

#### **Synthesis of quinoxalinehydroxamic acid (H<sub>2</sub>quinoHA)**

In a three-neck flask, 2-quinoxalinecarboxylic acid (3.1 g, 18.0 mmol) was added into a solution of N-methylmorpholine (2.4 mL, 21.8 mmol) in dichloromethane (50.0 mL). The mixture was stirred under N<sub>2</sub> for 15 minutes and cooled down to 0 °C. Ethylchloroformate (2.0 mL, 21.0 mmol) was added dropwise while stirring. The mixture was stirred for 1 hour and filtered to obtain a brown solution. Meanwhile, a fresh hydroxylamine solution was prepared by reacting hydroxylamine hydrochloride (1.9 g, 27.0 mmol) with potassium hydroxide 85% (1.8 g, 27.0 mmol) in methanol (16.0 mL) at 0°C. The solution was stirred for 10 minutes, filtered, and the filtrate obtained was added dropwise into the brown solution while stirring at 0°C. The reaction mixture was stirred for 2 hours, filtered, and the filtrate was left for slow evaporation to induce the precipitation of a white solid. The solid was recrystallized in hot water, collected by filtration, washed with a copious amount of water, and dried under vacuum to yield the pure powder of H<sub>2</sub>quinoHA. Yield ~ 26% (0.90 g), ESI-MS, calc. for [M - H<sup>+</sup>], C<sub>9</sub>H<sub>6</sub>N<sub>3</sub>O<sub>2</sub>, 188.05; found 188.0. <sup>1</sup>HNMR (400 MHz, DMSO-*d*<sub>6</sub>) δ ppm 11.78 (s, 1 H), 9.40 (s, 1 H), 9.37 (s, 1 H), 8.20–8.15 (m, 2H), 8.00–7.95 (m, 2 H). <sup>13</sup>C NMR (400 MHz, DMSO-*d*<sub>6</sub>) δ ppm 161.0, 144.7, 143.6, 142.8, 139.8, 131.7, 131.2, 129.4, 129.1. Elem. Anal., calc. (found) for (C<sub>9</sub>H<sub>7</sub>N<sub>3</sub>O<sub>2</sub>), C: 57.14 (57.42), H: 3.73 (3.85), N: 22.21 (22.06).

[Zn<sub>16</sub>Yb(quinoHA)<sub>16</sub>(py)<sub>8</sub>](OTf)<sub>3</sub> (YbZn<sub>16</sub>(quinoHA)<sub>16</sub>). H<sub>2</sub>quinoHA (0.060 g, 0.32 mmol) was added into a solution containing 1.0 mL pyridine, 5.0 mL DMF, and 5.0 mL H<sub>2</sub>O. The solution was stirred for 5 minutes to completely dissolve the solid. Zinc triflate (0.12 g, 0.32 mmol) and ytterbium triflate (0.024 g, 0.040 mmol) were added and the orange solution was stirred at 80°C for 2 hours. The solution was then cooled down to room temperature, and filtered. The filtrate was layered with H<sub>2</sub>O to give red-brown crystals after three days. The crystals were collected by filtration, washed with H<sub>2</sub>O, and dried in air. Yield ~ 23% (25 mg). ESI-MS (MeOH), calc. for [M]<sup>3+</sup>, Zn<sub>16</sub>YbC<sub>144</sub>H<sub>80</sub>N<sub>48</sub>O<sub>32</sub>, 1403.46;

found 1403.4. Elem. Anal., calc. (found) for  $(\text{Zn}_{16}\text{YbC}_{187}\text{H}_{120}\text{N}_{56}\text{O}_{41}\text{F}_9\text{S}_3)(\text{H}_2\text{O})_{12}$ , C: 40.76 (40.90), H: 2.63 (2.39), N: 14.24 (14.26).

## 6.5. Nanotechnology

### 6.5.1. Encapsulation of $\text{Yb}^{\text{III}}/\text{Zn}^{\text{II}}$ metallacrowns ( $\text{YbZn}_{16}(\text{quinoHA})_{16}$ ) in polystyrene nanoparticles: $\text{YbZn}_{16}(\text{quinoHA})_{16}@\text{PS}/\text{COOH}$

To an aqueous suspension (600  $\mu\text{L}$ ; 10 mg/mL) of polystyrene, surface carboxylated nanoparticles (PS/COOH) (Micromer®, Micromod Partikeltechnologie GmbH), a solution of lanthanide complex,  $\text{Yb}^{\text{III}}/\text{Zn}^{\text{II}}$  metallacrown  $\text{YbZn}_{16}(\text{quinoHA})_{16}$  (5mM, 100  $\mu\text{L}$ ) in N, N-dimethylformamide (DMF), was added to a final 6:1 volume ratio. After 20 min of strong agitation, water (350  $\mu\text{L}$ ) was added to the suspension and agitation was continued for another 20 min at room temperature. Afterwards, sample was centrifuged (15200 g, 30 min; Fisher Scientific™), supernatant removed and NPs pellet resuspended in 2-morpholinoethanesulfonic acid (MES) buffer (50 mM, 1 mL, pH 6) by combining vortex mixing and an ultrasonic bath (30 min) until suspension appeared homogenous. This washing procedure was repeated twice and the resulted sample ( $\text{YbZn}_{16}(\text{quinoHA})_{16}@\text{PS}/\text{COOH}$ ) was resuspended in MES (50 mM, pH 6) to a final concentration of 6 mg/mL and stored at 4°C until further use.

### 6.5.2. Surface modification of PS/COOH to PS/NTA

To a MES buffered suspension (1 mL; 6 mg/mL) of PS/COOH and  $\text{YbZn}_{16}(\text{quinoHA})_{16}@\text{PS}/\text{COOH}$ , water solutions of 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC) (50 mM; 100  $\mu\text{L}$ ) and N-hydroxysulfosuccinimide (sulfo-NHS) (50 mM; 100  $\mu\text{L}$ ), prepared immediately before use, were added and NPs suspension was incubated by strong agitation at room temperature for 30 min. Afterwards, sample was centrifuged (15200 g, 30 min), supernatant removed and NPs pellet resuspended in MES buffer (50 mM, 1 mL, pH 6) by combining vortex mixing and an ultrasonic bath (30 min) until suspension appeared homogenous. This washing procedure was repeated twice. The sample was finally resuspended in MES (500  $\mu\text{L}$ ) and aqueous solution of  $\text{N}\alpha$ ,  $\text{N}\alpha$ -bis(carboxymethyl)-L-lysine (10 mM; 500  $\mu\text{L}$ ) was added. After 2h 30 min of medium speed agitation at room temperature, sample was centrifuged (15200 g, 30 min), supernatant removed and NPs pellet resuspended in Tris buffer (50 mM, 1 mL, pH 8) by

combining vortex mixing and an ultrasonic bath (30 min) until suspension appeared homogenous. This washing procedure was repeated twice and the sample (PS/NTA or YbZn<sub>16</sub>(quinoHA)<sub>16</sub>@PS/NTA) was finally resuspended in Tris buffer (20 mM, pH = 8) to a concentration of 6 mg/mL and stored at 4°C until further use.

### 6.5.3. Synthesis of PS/NTA-Ni<sup>2+</sup> and YbZn<sub>16</sub>(quinoHA)<sub>16</sub>@PS/NTA-Ni<sup>2+</sup>

The Tris-buffered suspension of PS/NTA or YbZn<sub>16</sub>(quinoHA)<sub>16</sub>@PS/NTA NPs (900 µL; 10 mg/mL) was incubated at medium speed agitation at room temperature for 45 min with aqueous solution of NiCl<sub>2</sub>·6H<sub>2</sub>O (50 mM; 100 µL). Afterwards, samples were centrifuged (15200 g, 30 min), supernatant removed and NPs pellet resuspended in Tris buffer (20 mM, 1 mL, pH 8) by combining vortex mixing and an ultrasonic bath (30 min) until suspension appeared homogenous. This washing procedure was repeated twice and the sample PS/NTA-Ni<sup>2+</sup> or YbZn<sub>16</sub>(quinoHA)<sub>16</sub>@PS/NTA-Ni<sup>2+</sup> was finally resuspended in Tris buffer (20 mM, pH = 8) to a concentration of 6 mg/mL and stored at 4°C until further use.

### 6.5.4. Bioconjugation of 6xHis-tagged recombinant GalF-ase to PS/NTA-Ni<sup>2+</sup> and YbZn<sub>16</sub>(quinoHA)<sub>16</sub>@PS/NTA-Ni<sup>2+</sup>

The Tris buffered suspension of PS/NTA-Ni<sup>2+</sup> or YbZn<sub>16</sub>(quinoHA)<sub>16</sub>@PS/NTA-Ni<sup>2+</sup> NPs (350 µL; 6 mg/mL) was incubated by medium agitation (20 Hz, Fisher Scientific TopMix FB15024 Vortexer) at room temperature for 1 h with solution of GalF-ase (100 µL; 0.4 mg/mL) that corresponds to 35:1 mass ratio. Afterwards, samples were centrifuged (15 200 g, 30 min; Fisher Scientific™), supernatant removed and kept, while the NPs pellet resuspended in Tris buffer (20 mM, 165 µL, pH 8) by pipetting up and down until suspension appeared homogenous. This washing procedure was repeated three times and the sample (PS/NTA-Ni<sup>2+</sup>-6xHisGalF-ase or YbZn<sub>16</sub>(quinoHA)<sub>16</sub>@PS/NTA-Ni<sup>2+</sup>-6xHisGalF-ase) was finally resuspended in Tris buffer (20 mM, 165 µL, pH 8) and, together with the supernatants from all the washing procedures, was kept and used for enzymatic activity assay.

***Bioconjugation assessment of PS/NTA-Ni<sup>2+</sup>-6xHisGalF-ase or YbZn<sub>16</sub>(quinoHA)<sub>16</sub>@PS/NTA-Ni<sup>2+</sup>-6xHisGalF-ase by enzymatic activity assay***

Supernatants (165  $\mu\text{L}$ ), after initial bioconjugation incubation reaction, washing procedures and final suspension system sample (PS/NTA-Ni<sup>2+</sup>-6xHisGal $\beta$ -ase or YbZn<sub>16</sub>(quinoHA)<sub>16</sub>@PS/NTA-Ni<sup>2+</sup>-6xHisGal $\beta$ -ase, 165  $\mu\text{L}$ ) were used in enzymatic activity assay in the presence of *p*NP- $\beta$ -D-Gal $\beta$  substrate (10  $\mu\text{L}$ ; 10 mM) and buffer (30  $\mu\text{L}$ ; 0.1 M Citric acid/0.2 M Na<sub>2</sub>HPO<sub>4</sub> pH 4.5). The reaction mixture (205  $\mu\text{L}$ ) was incubated at 37°C (water bath) for 5 min, terminated by an addition of 100  $\mu\text{L}$  of 1 M Na<sub>2</sub>CO<sub>3</sub> solution and centrifuged (15200 g, 30 min, 10°C; Fisher Scientific Sigma™) if necessary. Aliquots of reaction mixtures (200  $\mu\text{L}$ ) were taken, transferred in 96-well microtiter plate and absorbance (405 nm;  $\epsilon_{405} = 19\,500\text{ M}^{-1}\text{cm}^{-1}$ ) of released *p*NP was measured using spectrophotometer (Thermo Scientific™ Multiskan™ GO). Control enzyme activity reactions, with and without PS/NTA-Ni<sup>2+</sup> or YbZn<sub>16</sub>(quinoHA)<sub>16</sub>@PS/NTA-Ni<sup>2+</sup>, were carried out. All the reactions were assayed in triplicate and absorbance values were corrected for the spontaneous hydrolysis of the substrate and tested NP systems.

#### 6.5.5. Characterization of PS/NTA-Ni<sup>2+</sup> & YbZn<sub>16</sub>(quinoHA)<sub>16</sub>@PS/NTA-Ni<sup>2+</sup> nanoparticles

##### *Dynamic Light Scattering (DLS)*

Hydrodynamic diameters of NPs were determined by DLS using a Zetasizer Nano ZS particle size analyser (Malvern Instruments, Westborough, MA, USA). The stock suspensions of NPs were diluted in water to a 0.06 mg/mL concentration, 500  $\mu\text{L}$  were loaded in appropriate cuvettes and at least three measurements were made at room temperature. The mean hydrodynamic diameter of the particles was calculated using the Malvern software package.

##### *Inductively coupled plasma mass spectrometry (ICP-MS)*

The Tris buffered suspension of PS/NTA-Ni<sup>2+</sup> or YbZn<sub>16</sub>(quinoHA)<sub>16</sub>@PS/NTA-Ni<sup>2+</sup> (20 mM, 1.5 mL; 10 mg/mL) was centrifuged (15 200 g, 30 min, Fisher Scientific™), supernatant removed and pellet resuspended in EDTA aqueous solution (50 mM, 1 mL) by combining vortex mixing and an ultrasonic bath (5 min) until suspension appeared homogenous. Afterwards, sample was centrifuged (15200 g, 30 min), supernatant (1 mL) removed and analysed for Ni<sup>2+</sup> content by inductively coupled plasma mass spectrometry (ICP-MS). From the YbZn<sub>16</sub>(quinoHA)<sub>16</sub>@PS/NTA-Ni<sup>2+</sup> sample (1.5 mL, 10 mg/mL; 150 mg) obtained Ni<sup>2+</sup> concentration was 82 mg/mL.

***Photophysical measurements***

Luminescence data were collected on freshly prepared suspensions of NPs in corresponding buffers (200  $\mu\text{g}/\text{mL}$ ) placed in 2.4 mm i.d. quartz capillaries. Steady-state emission and excitation spectra were measured on a custom-designed Horiba Scientific Fluorolog 3 spectrofluorimeter equipped with either a visible photomultiplier tube (PMT) (220-850 nm, R928P; Hamamatsu) or a NIR PMT (950-1650 nm, H10330-75; Hamamatsu) upon excitation with a continuous Xenon lamp. All excitation and emission spectra were corrected for the instrumental functions. Luminescence lifetimes ( $\tau_{obs}$ ) were determined under excitation at 355 nm provided by a Nd:YAG laser (YG 980; Quantel). Signals were detected in the NIR ranges using an iHR320 monochromator (Horiba Scientific) equipped with a Hamamatsu H10330-75 PMT. The output signal from the detector was fed into a 500 MHz bandpass digital oscilloscope (TDS 754C; Tektronix), transferred to a PC for data processing with the program Origin 8<sup>®</sup>. Luminescence lifetimes are averages of at least three independent measurements. Yb<sup>3+</sup>-centered quantum yields under ligands excitation ( $Q_{Yb}^L$ ) were determined with the Fluorolog 3 spectrofluorimeter based on an absolute method with the use of an integration sphere (Model G8, GMP SA, Renens, Switzerland). Each sample was measured several times under comparable experimental conditions, varying the position of samples. Estimated experimental error for quantum yield determination is ~10 %.

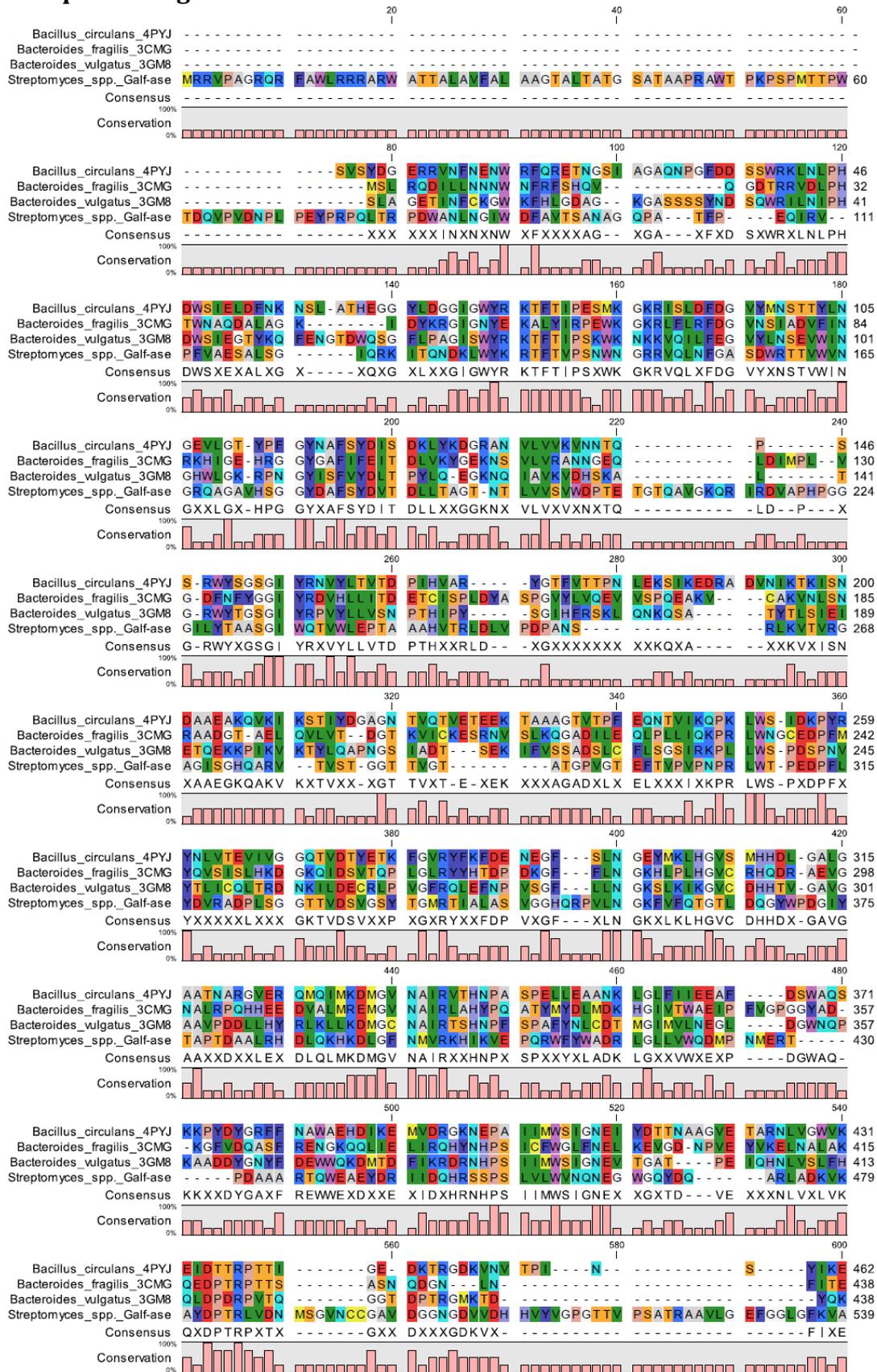
## **CHAPTER VII**

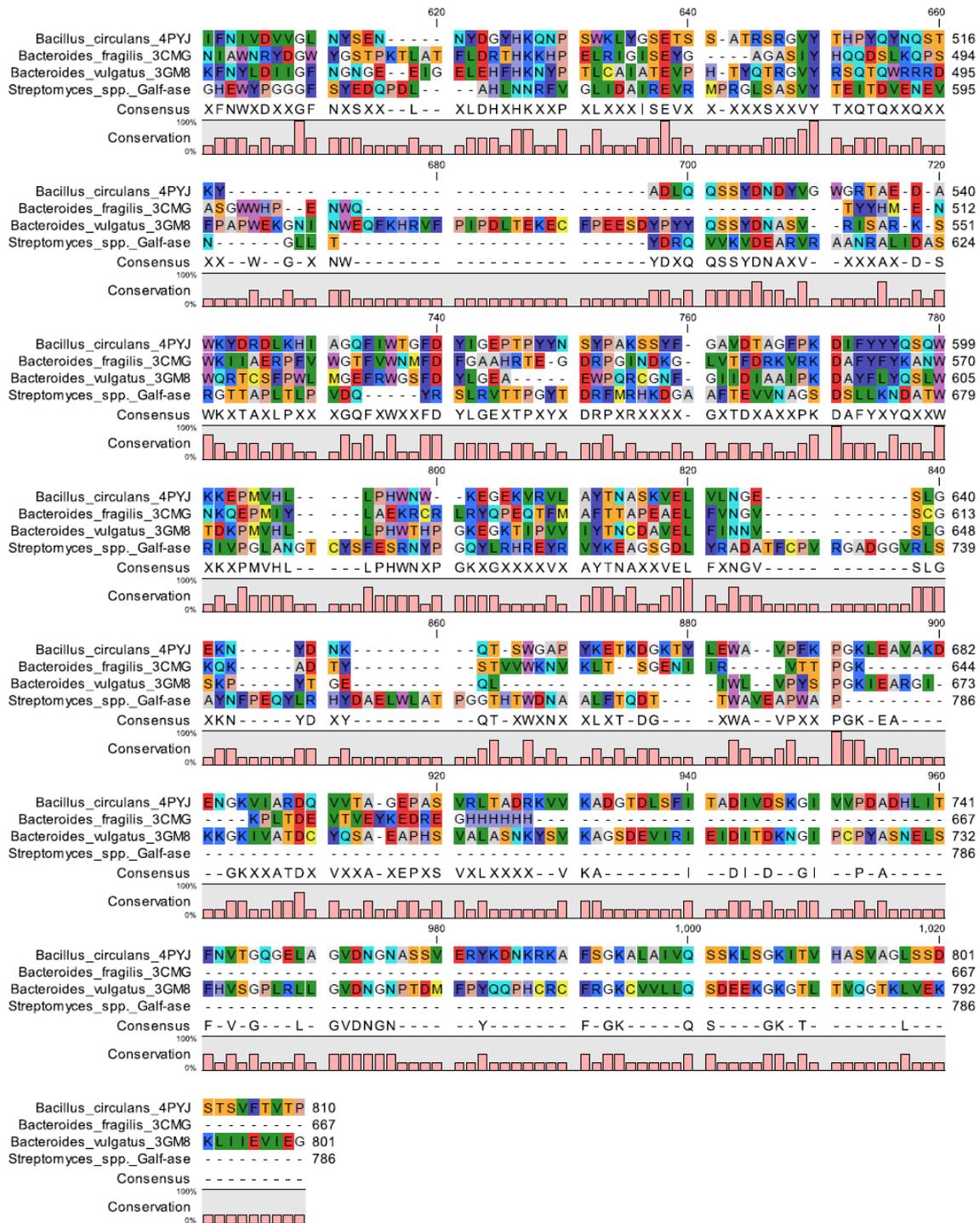
### **Appendix**

<b>7.1. Additional figures.....</b>	<b>136</b>
7.1.1. Sequence alignment.....	136
7.1.2. Structural formulas of <i>p</i> NP-linked substrates .....	138
7.1.3. The HPLC chromatograms and ESI-HRMS spectras of thioglycosylation reactions by E464Q.....	139
7.1.5. Short communication published in the journal Carbohydrate research .....	147

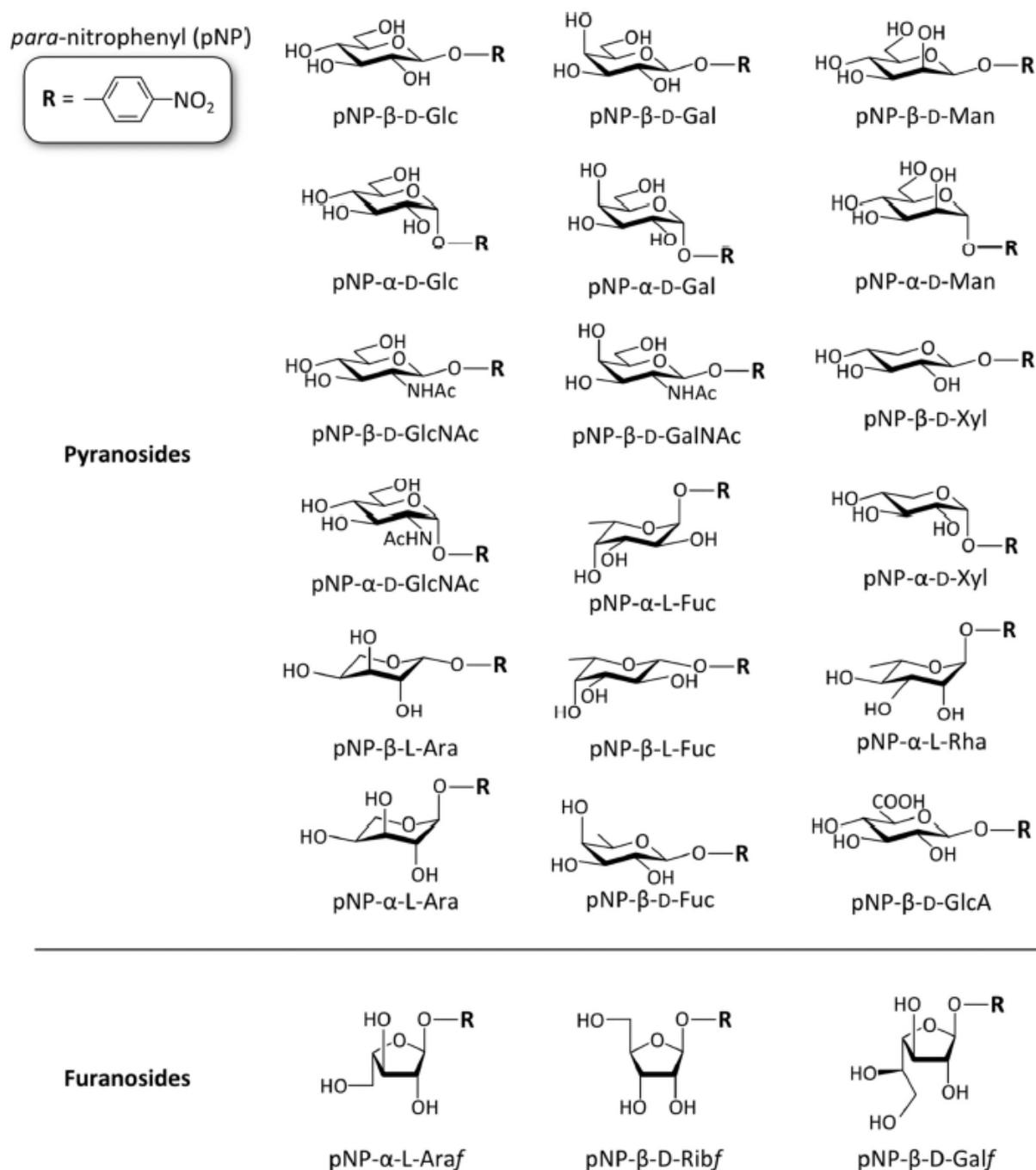
## 7.1. Additional figures

## 7.1.1. Sequence alignment



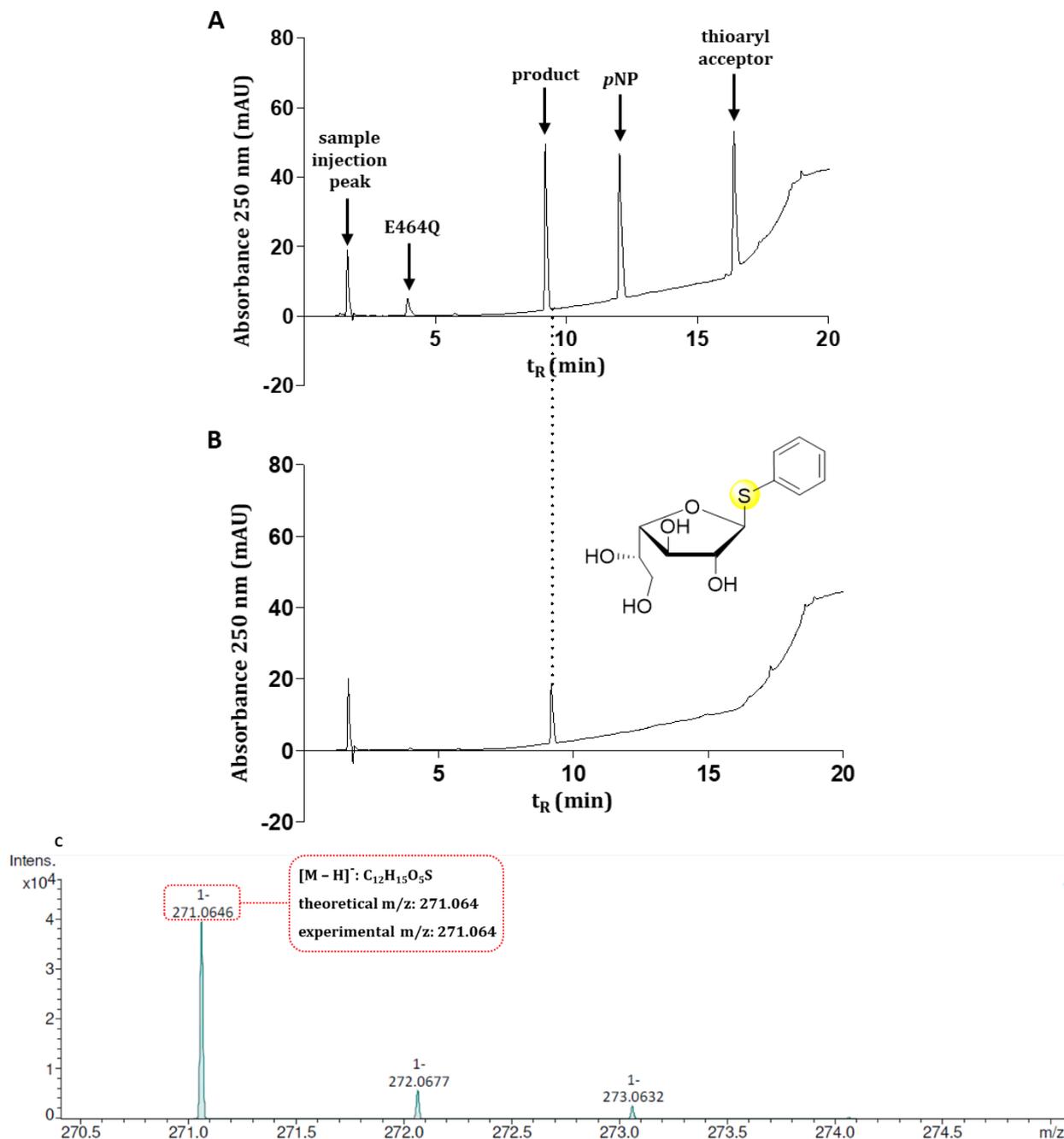


**Figure VII-1.** Alignment of complete ORF Galf-ase (*Streptomyces* spp.) amino acid sequence and of its homologs.

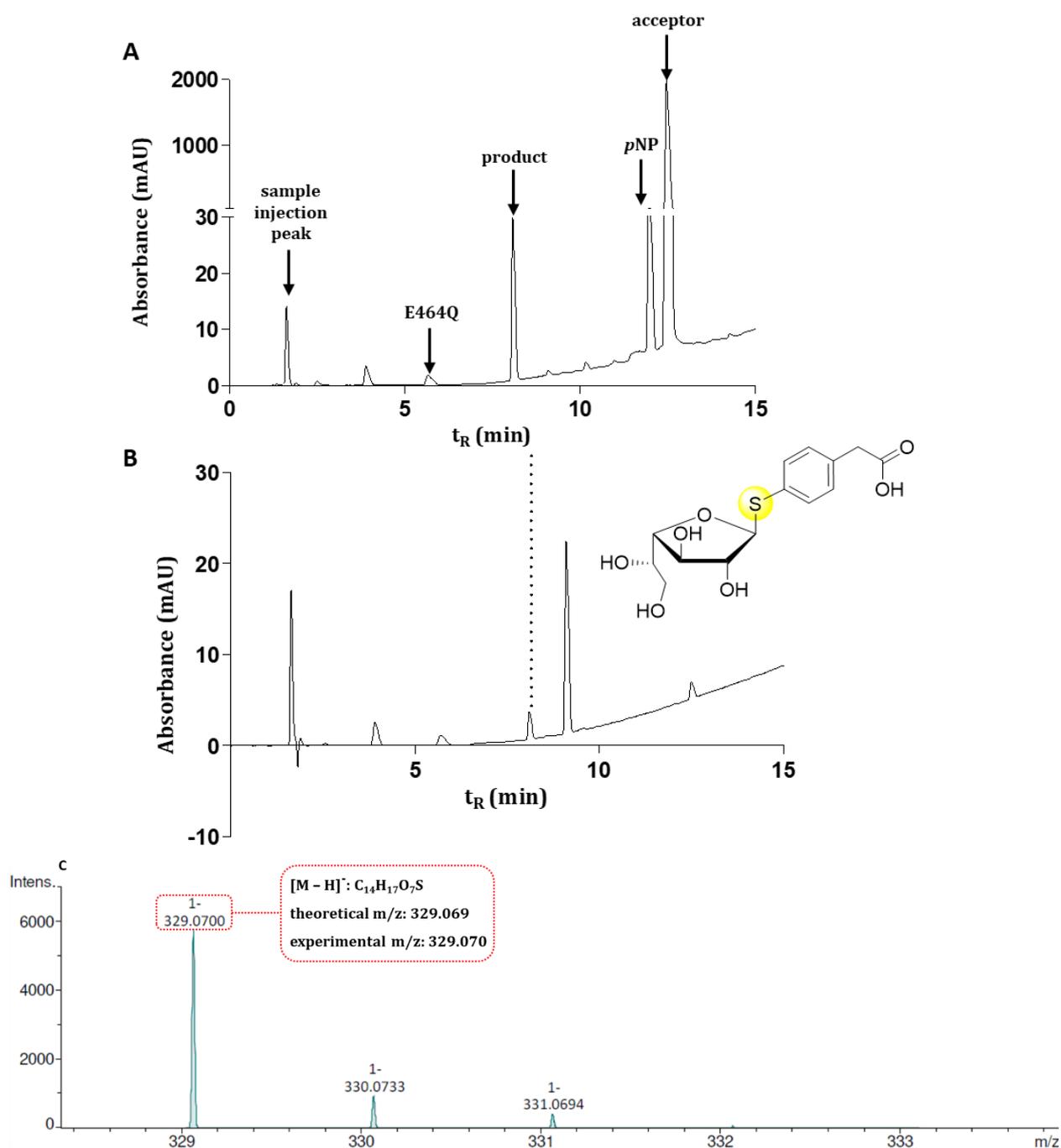
7.1.2. Structural formulas of *p*NP-linked substrates

**Figure VII-2.** Structural formulas of 21 commercially available *p*NP-pyranosyl and -furanosyl substrates used for *Galf*-ase substrate specificity screening.

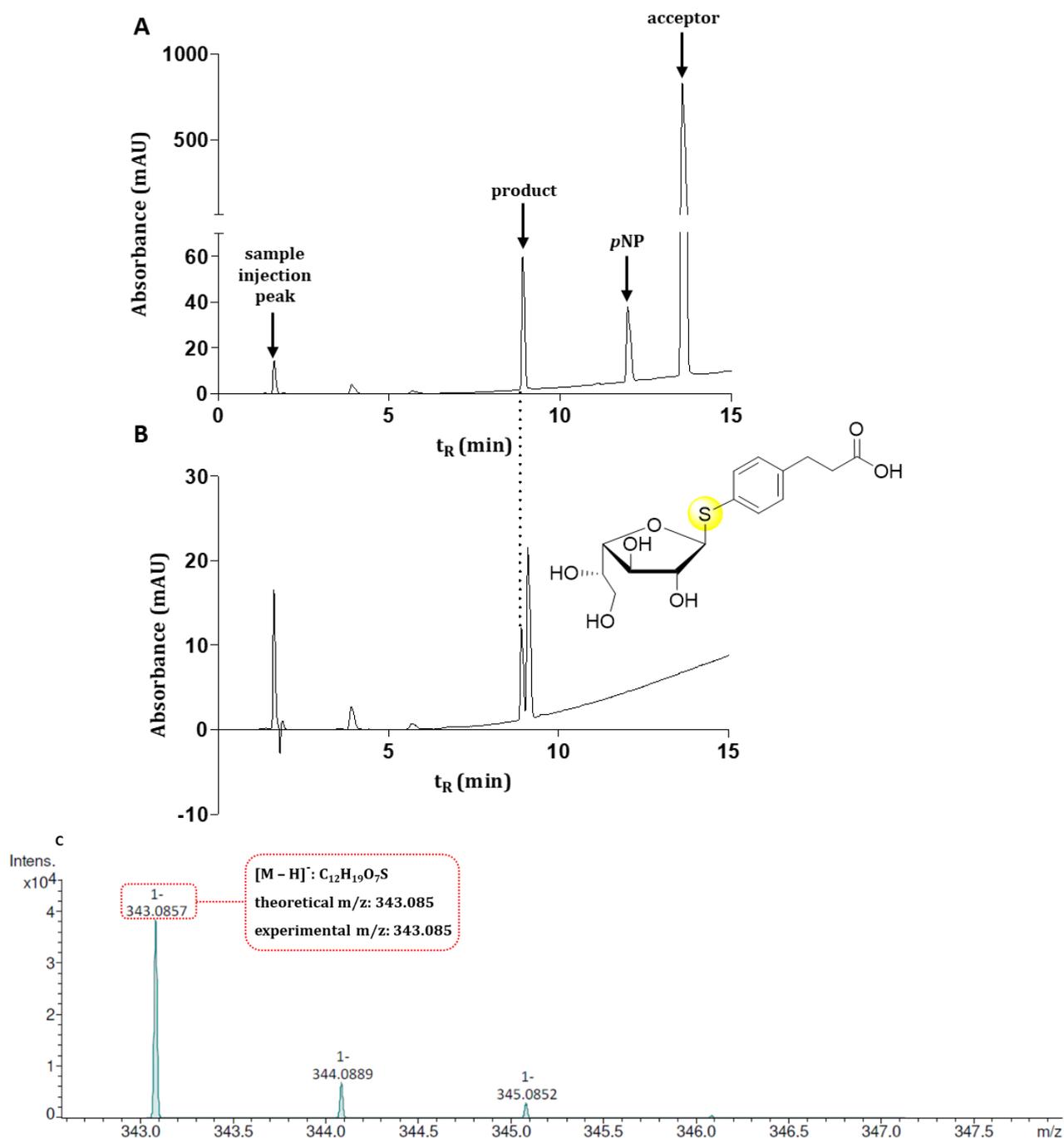
### 7.1.3. The HPLC chromatograms and ESI-HRMS spectras of thioglycosylation reactions by E464Q



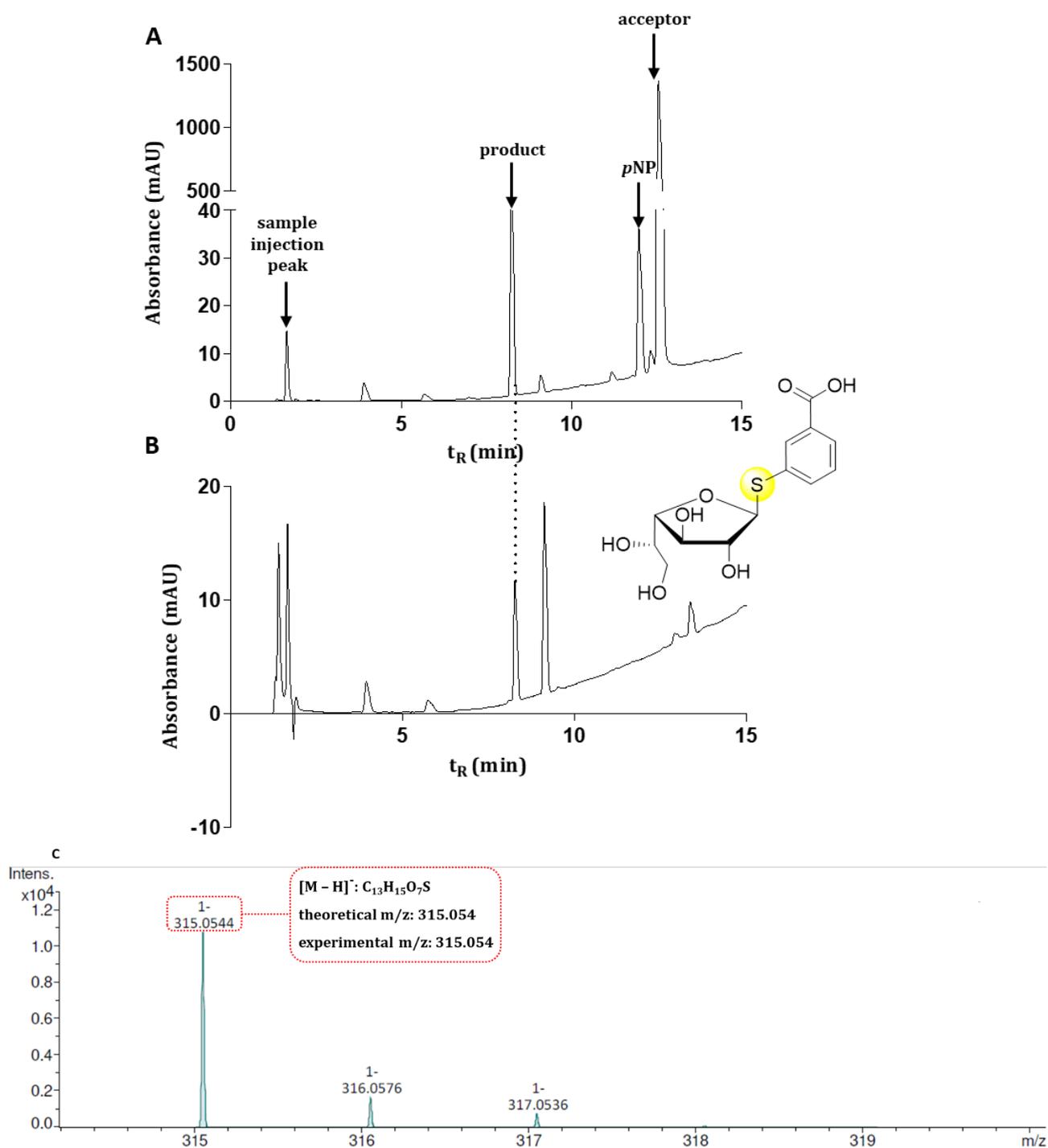
**Figure VII-3.** A) Chromatogram of crude thioglycosylation reaction with thioaryl acceptor. B) Chromatogram of purified product fraction from thioglycosylation reaction. C) The measured HRMS-ESI spectra of thioarylgalactofuranosyl ion in negative mode. The analysed sample is the fraction purified on C18 chromatography resin (eluent solvent mixture: A:B=6:4; A=0.1% formic acid (v/v) in water, B=0.1% formic acid (v/v) in acetonitrile) and corresponds to the peak ( $t_R=9.19$  min) detected on HPLC (Figure VII-3. B).



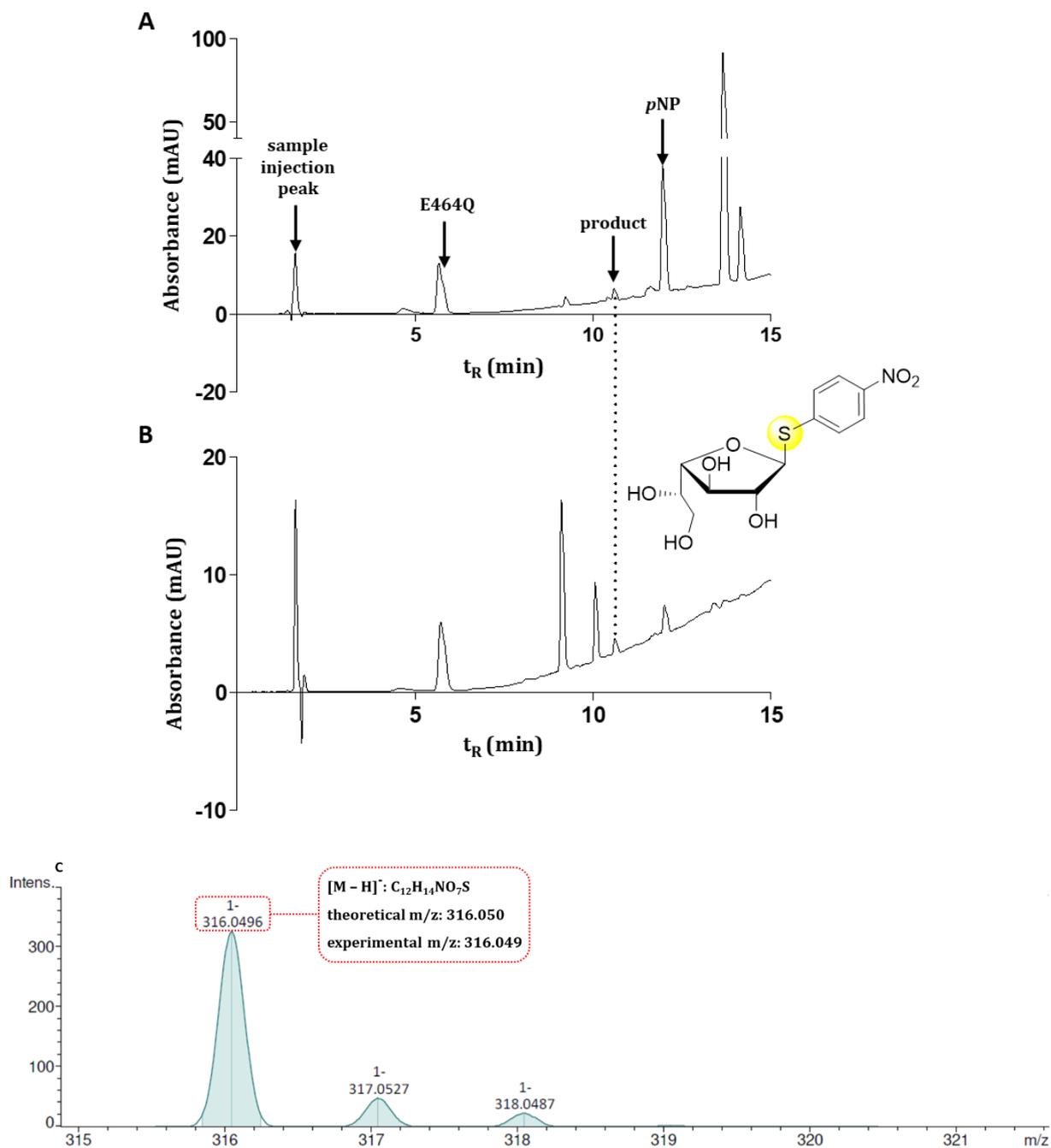
**Figure VII-4.** A) Chromatogram of crude thioglycosylation reaction with mercaptophenylacetic acid as acceptor. B) Chromatogram of purified product fraction from thioglycosylation reaction. C) The measured HRMS-ESI spectra of thioaryl galactofuranosyl ion in negative mode. The analysed sample is the fraction purified on C18 chromatography resin (eluent solvent mixture: A:B=7:3) and corresponds to the peak ( $t_R=8.08$  min) detected on HPLC (Figure VII-4. B).



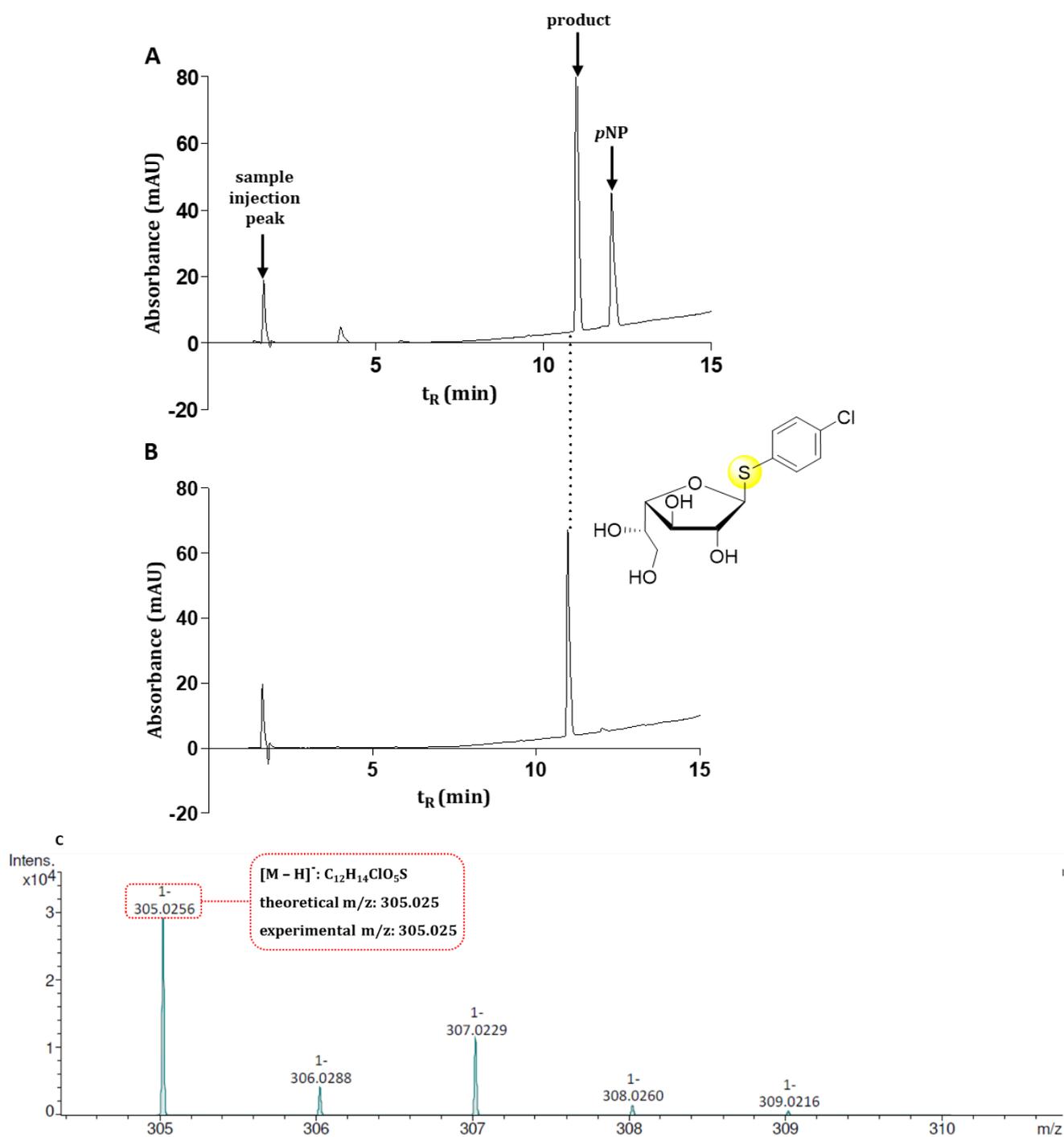
**Figure VII-5.** A) Chromatogram of crude thioglycosylation reaction with 4-mercaptohydrocinnamic acid as acceptor. B) Chromatogram of purified product fraction from thioglycosylation reaction. C) The measured HRMS-ESI spectra of thioaryl galactofuranosyl ion in negative mode. The analysed sample is the fraction purified on C18 chromatography resin (eluent solvent mixture: **A**:**B**=7:3) and corresponds to the peak ( $t_R$ =8.9 min) detected on HPLC (**Figure VII-5. B**).



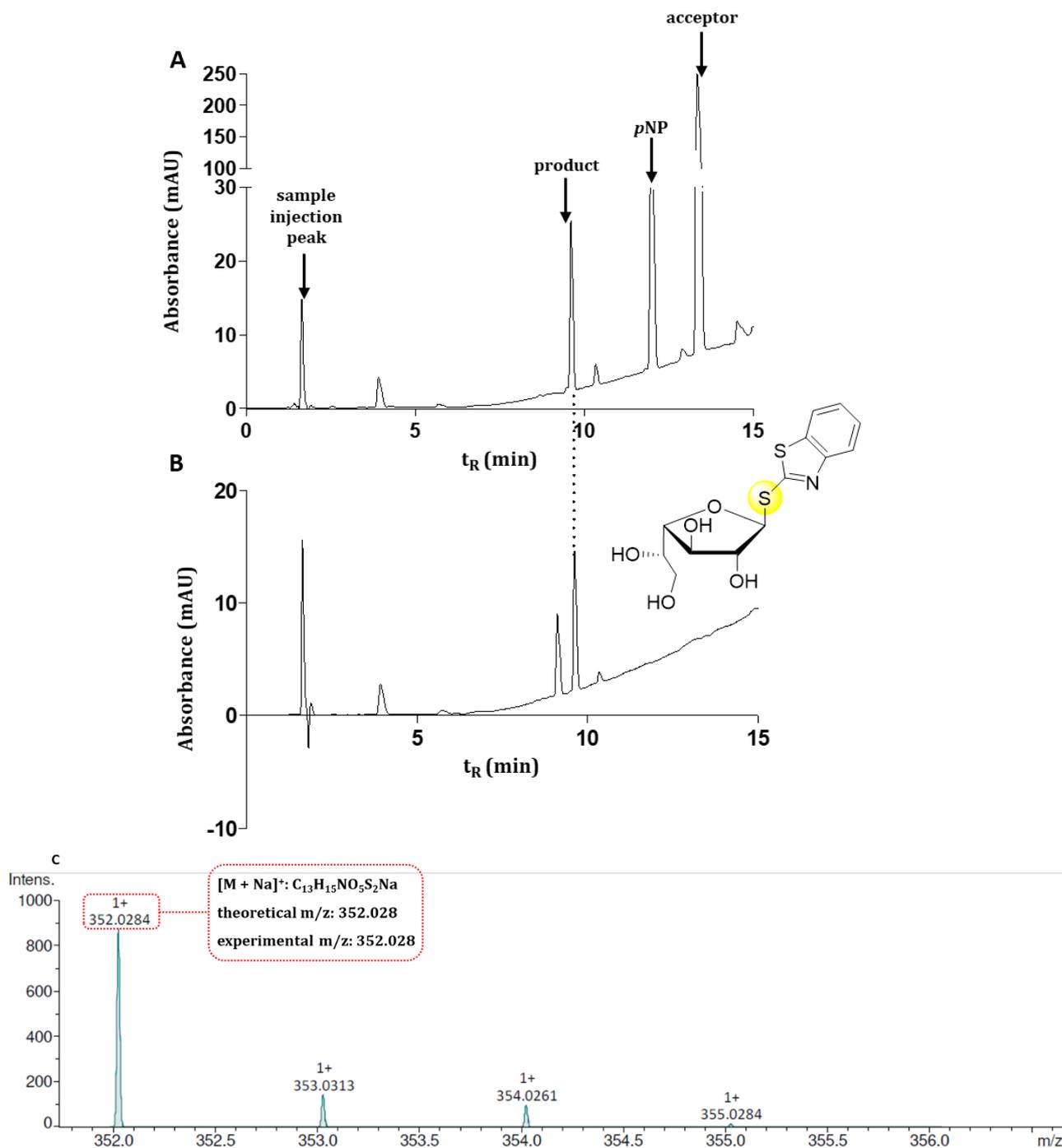
**Figure VII-6.** A) Chromatogram of crude thioglycosylation reaction with 3-mercaptopropionic acid as acceptor. B) Chromatogram of purified product fraction from thioglycosylation reaction. C) The measured HRMS-ESI spectra of thioarylgalactofuranosyl ion in negative mode. The analysed sample is the fraction purified on C18 chromatography resin (eluent solvent mixture: A:B=7:3) and corresponds to the peak ( $t_R=8.2$  min) detected on HPLC (Figure VII-6. B).



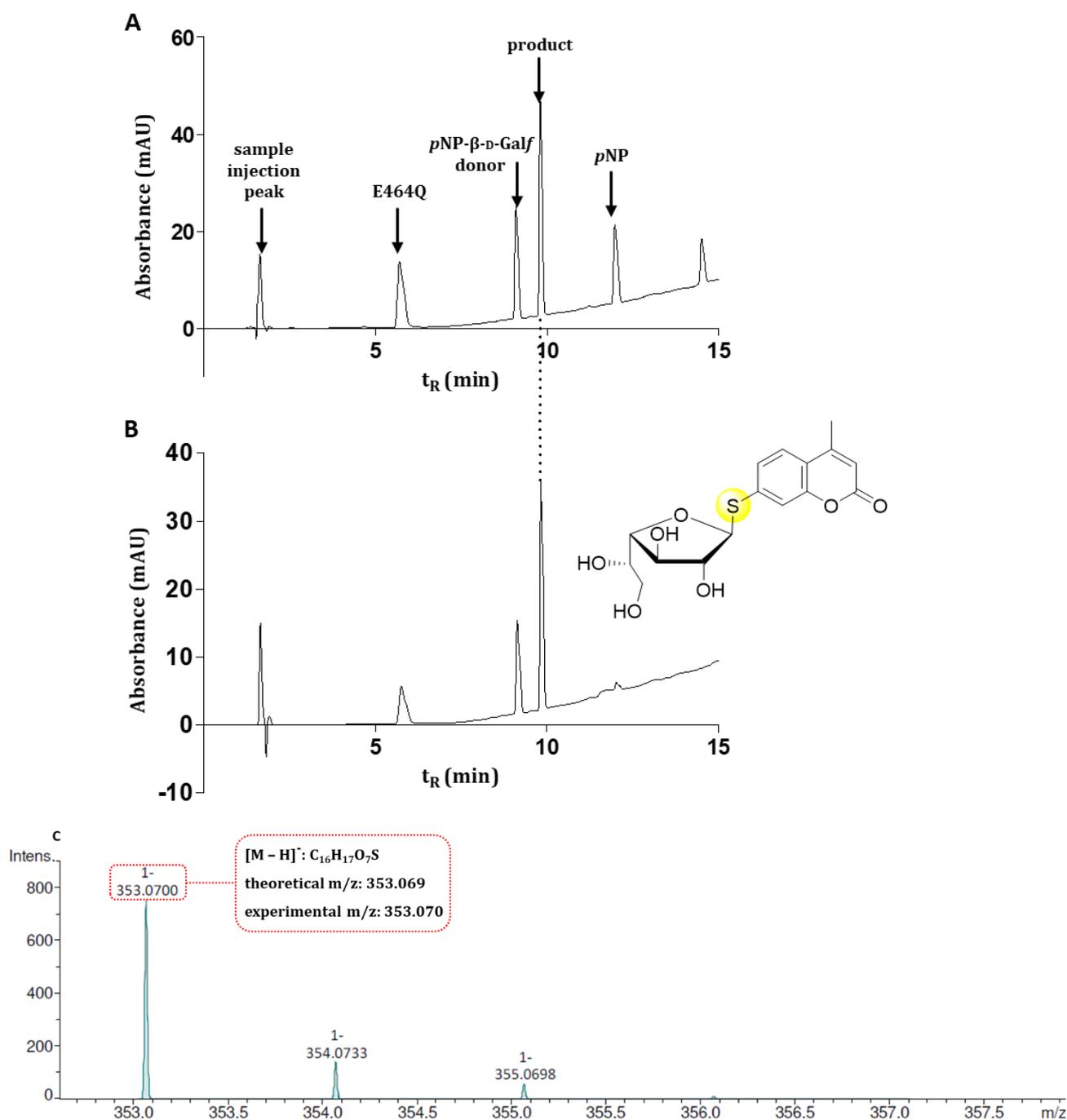
**Figure VII-7.** A) Chromatogram of crude thioglycosylation reaction with 4-nitrothiophenol as acceptor. B) Chromatogram of purified product fraction from thioglycosylation reaction. C) The measured HRMS-ESI spectra of thioaryl galactofuranosyl ion in negative mode. The analysed sample is the fraction purified on C18 chromatography resin (eluent solvent mixture: A:B=6:4) and corresponds to the peak ( $t_R$ =8.9 min) detected on HPLC (Figure VII-7. B).



**Figure VII-8.** A) Chromatogram of crude thioglycosylation reaction with 4-chlorothiophenol as acceptor. B) Chromatogram of purified product fraction from thioglycosylation reaction. C) The measured HRMS-ESI spectra of thioaryl galactofuranosyl ion in negative mode. The analysed sample is the fraction purified on C18 chromatography resin (eluent solvent mixture: A:B=6:4) and corresponds to the peak ( $t_R=10.9$  min) detected on HPLC (Figure VII-8.).



**Figure VII-9.** A) Chromatogram of crude thioglycosylation reaction with 2-mercaptobenzothiazole as acceptor. B) Chromatogram of purified product fraction from thioglycosylation reaction. C) The measured HRMS-ESI spectra of thioaryl galactofuranosyl ion in negative mode. The analysed sample is the fraction purified on C18 chromatography resin (eluent solvent mixture: A:B=6:4) and corresponds to the peak ( $t_R=9.6$  min) detected on HPLC (Figure VII-9.).



**Figure VII-10.** A) Chromatogram of crude thioglycosylation reaction with 7-mercapto-4-methylcoumarin as acceptor. B) Chromatogram of purified product fraction from thioglycosylation reaction. C) The measured HRMS-ESI spectra of thioaryl galactofuranosyl ion in negative mode. The analysed sample is the fraction purified on C18 chromatography resin (eluent solvent mixture: A:B=6:4) and corresponds to the peak ( $t_R$ =9.8 min) detected on HPLC (Figure VII-10.).

## 7.1.5. Short communication published in the journal Carbohydrate research

Carbohydrate Research 480 (2019) 35–41



Contents lists available at ScienceDirect

Carbohydrate Research

Journal homepage: [www.elsevier.com/locate/carres](http://www.elsevier.com/locate/carres)Galactofuranosidase from JHA 19 *Streptomyces* sp.: subcloning and biochemical characterizationMateja Seničar<sup>a,b</sup>, Laurent Legentil<sup>c</sup>, Vincent Ferrières<sup>c</sup>, Svetlana V. Eliseeva<sup>b</sup>, Stéphane Petoud<sup>b</sup>, Kaoru Takegawa<sup>d</sup>, Pierre Lafite<sup>a</sup>, Richard Daniellou<sup>a,\*</sup><sup>a</sup> ICOA UMR CRNS 7311, Université d'Orléans, Rue de Chartres, BP 6759, 45067, Orléans Cedex 2, France<sup>b</sup> Centre de Biophysique Moléculaire, CNRS, UPR 4301, 45071, Orléans, France<sup>c</sup> Univ Rennes, Ecole Nationale Supérieure de Chimie de Rennes, CNRS, ISCR, UMR 6226, F-35000, Rennes, France<sup>d</sup> Department of Bioscience and Biotechnology, Faculty of Agriculture, Kyushu University, Moto-oka 744, Nishi-ku, Fukuoka, Japan

## ARTICLE INFO

## Keywords:

$\beta$ -D-galactofuranosidase  
Subcloning  
Glycoside hydrolase  
Thio- $\beta$ -D-galactofuranoside inhibition

## ABSTRACT

Despite the crucial role of the rare galactofuranose (Gal<sub>f</sub>) in many pathogenic micro-organisms and our increased knowledge of its metabolism, there is still a lack of recombinant and efficient galactofuranoside hydrolase available for chemo-enzymatic synthetic purposes of specific galactofuranosyl-conjugates. Subcloning of the Gal<sub>f</sub>-ase from JHA 19 *Streptomyces* sp. and its further overexpression lead us to the production of this enzyme with a yield of 0.5 mg/L of culture. It exhibits substrate specificity exclusively towards pNP  $\beta$ -D-Gal<sub>f</sub>, giving a  $K_M$  value of 250  $\mu$ M, and the highest enzymatic efficiency ever observed of 14 mM<sup>-1</sup> s<sup>-1</sup>. It proved to be stable to temperature up to 60 °C and to at least 4 freeze-thaw cycles. Thus, Gal<sub>f</sub>-ase demonstrated to be an efficient and stable biocatalyst with greatly improved specificity toward the galactofuranosyl entity, thus paving the way to the further development of transglycosylation and thioligation reactions.

## 1. Introduction

D-Galactose (D-Gal) is extensively distributed in nature as constituent of oligosaccharides and glycoconjugates [1]. However, under its furanose form, galactofuranose (Gal<sub>f</sub>) is totally absent from mammals, but is present in many pathogenic microorganisms [2,3]. Indeed, there are three main and specific enzymes involved in the biosynthesis and the metabolism of the Gal<sub>f</sub>-containing molecules [4]: UDP-galactopyranose mutase (UGM) [5], galactofuranosyltransferase (Gal<sub>f</sub>T) [6] and galactofuranosidase (Gal<sub>f</sub>-ase) [7]. The glycobiology of Gal<sub>f</sub> and enzymes involved in its metabolism have thus become attractive targets for the development of antimicrobial agents [8].

Although the metabolism of  $\beta$ -D-galactofuranosyl conjugates has been extensively studied, mostly in infectious microorganisms such as *Mycobacteria*, *Trypanosoma*, *Leishmania* and *Aspergillus*, there are only few reports related to these enzymes, especially Gal<sub>f</sub>-ase. Over the past forty years, Gal<sub>f</sub>-ase has been identified as responsible for the degradation of the D-Gal<sub>f</sub> containing glycoconjugates [8]. The first Gal<sub>f</sub>-ase identified was an extracellular exo- $\beta$ -D-galactofuranosidase (exo- $\beta$ -D-Gal<sub>f</sub>-ase) isolated and partially purified from *Penicillium fellutanum* (ex-type of *Penicillium charlesii*) culture filtrates [7]. Although described back in 1977 and later, several exo- and endo-Gal<sub>f</sub>-ases were purified

from the culture supernatants and cell lysates of filamentous fungi [7–9], bacteria [10] and protozoa [11], but the genes encoding these enzymes have not been identified and expressed earlier, nor their amino acid sequence determined. Therefore, reported chemo-enzymatic procedures in literature all used cloned  $\alpha$ -L-arabinofuranosidases as biocatalytic tools, based on the structure similarity between pNP  $\beta$ -D-Gal<sub>f</sub> 1 and its analogue pNP  $\alpha$ -L-Araf 2 (Fig. 1.) [12–14]. However, given the poor catalytic efficiency of these biocatalysts toward pNP  $\beta$ -D-Gal<sub>f</sub> 1, there is still a crucial need for the discovery of improved biocatalysts that will allow the community to easily access to furanosyl conjugates (and analogues) owing potential biological properties [3].

Only recently, in 2015, based on the draft genome sequence analysis of soil Gram-positive bacteria *Streptomyces* sp., strain JHA19, an open reading frame that encoded Gal<sub>f</sub> specific enzyme was identified and cloned, and the enzyme characterized [15,16]. Another gene coding for a putative Gal<sub>f</sub>-ase was also found later in the genome of *Streptomyces* sp., strain JHA26. [16,17] However this first cloned Gal<sub>f</sub>-ase protein was co-expressed with a large fusion NusA [18] (N-utilization substance A) tag as solubility enhancer and two His-tags for affinity purification purpose, thus leading to a large protein of around 150 kDa (Fig. 2A) [15]. We thus rapidly thought that these multiple tags might hamper i) the overexpression of the Gal<sub>f</sub>-ase leading to low yield, and ii) the

\* Corresponding author.

E-mail address: [richard.daniellou@univ-orleans.fr](mailto:richard.daniellou@univ-orleans.fr) (R. Daniellou).<https://doi.org/10.1016/j.carres.2019.05.011>

Received 23 April 2019; Received in revised form 16 May 2019; Accepted 22 May 2019

Available online 28 May 2019

0008-6215/ © 2019 Elsevier Ltd. All rights reserved.

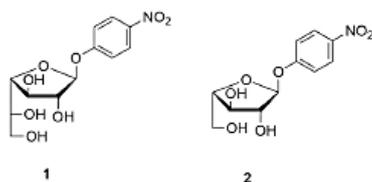


Fig. 1. Chemical structures of pNP  $\beta$ -D-Galf 1 and its analogue pNP  $\alpha$ -L-Araf 2.

physico-chemical and enzymatic properties of this original biocatalyst. In addition, the obtained low yield of protein was incompatible with the development of this biocatalyst.

In view of the previous research considering *Streptomyces* Galf-ase investigations [15], we wish to report herein the optimization of the overexpression and the purification to homogeneity of recombinant ORF1110  $\beta$ -D-galactofuranosidase including only one His-tag (Fig. 2B), with detailed and improved biochemical and kinetic properties, as well as preliminary inhibition studies of this enzyme. This Galf-ase will thus constitute the basis for the further development of biocatalyzed incorporation of Galf moieties in more complex glycoconjugates.

## 2. Materials and methods

### 2.1. Biological and chemical reagents

Restriction enzymes and T4 DNA ligase were purchased from Thermo Scientific™, *E. coli* Rosetta™ (DE3) and plasmid vector pET-28a(+) from Novagen™. All *p*-nitrophenyl monosaccharides (pNP sugars) were purchased from Carbosynth (Compton, UK). Galactofuranosides 3–6 were chemically synthesized as previously described [19,20]. All other laboratory chemicals used as the starting compounds, reagents and solvents were analytical grade purity and commercially available, unless otherwise specified.

### 2.2. Subcloning, overexpression, and protein purification

Plasmid construct pET-50b(+), encoding wild type Galf-ase gene was digested with EcoRI and HindIII restriction enzymes and the resulting 2.5 kb fragment containing Galf-ase gene was agarose-gel purified and ligated into expression vector pET-28a(+) (Novagen™). Ligation was verified by restriction enzyme analysis, and the gene sequence integrity was confirmed by DNA sequencing performed by Eurofins Genomics. *E. coli* Rosetta™ (DE3) (Novagen™) expression strain was transformed with Galf-ase encoding pET-28a(+) plasmid construct by heat shock method and cultured at 37 °C overnight on LB agar plates supplemented with chloramphenicol (34  $\mu$ g/mL) and kanamycin (30  $\mu$ g/mL). One single positive colony was used to inoculate fresh lysogeny broth (LB) (10 mL) supplemented with the same antibiotics and cultured overnight with agitation at 37 °C. The preculture was then inoculated into LB broth (1 L) containing corresponding antibiotics and shaker incubated (250 rpm) at 37 °C. Cells were grown to mid-exponential phase ( $OD_{600}$ : 0.6), shortly cooled on ice and protein expression was induced by the addition of  $\beta$ -D-thiogalactopyranoside (IPTG) (100  $\mu$ L; 1 M) and left on shaker incubator (250 rpm) for further 16 h at 15 °C. Cells were harvested by centrifugation (4255 g, 30 min, 4 °C) and the cell pellets were resuspended in lysis buffer solution (1:10 v/v; 100 mM NaCl, 50 mM Tris/HCl pH 8, 1 mM phenylmethanesulfonyl fluoride (PMSF), 5% glycerol, 0.1% Triton X-100, lysozyme 1 mg/L). Suspension was incubated by stirring for 20 min at 4 °C, lysed by three freeze-thaw cycles and subsequently sonicated ([2 min on/2 min pause]  $\times$  3, 50% cycle) on ice. Lysate was centrifuged (30 000 g, 20 min, 4 °C), supernatant filtered (0.45  $\mu$ m pore filter) and the recombinant protein was then purified from the clarified lysates using pre-equilibrated (10 mL; 50 mM Tris, 200 mM NaCl, 10 mM Imidazole)

Thermo Scientific HisPur™ Ni-NTA Chromatography Cartridge (1 mL). The bound protein was eluted by an imidazole gradient (10–500 mM) and an aliquot of eluted fractions were analysed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) on 8% separating gel according to Laemmli's method [21], and protein bands were visualized by staining with Coomassie Brilliant Blue G250. Fractions containing pure protein were collected, concentrated by ultrafiltration (30 000 MWCO, Sartorius Vivaspin® sample concentrator) and protein quantity was determined using colorimetric Bio-Rad protein assay based on the Bradford dye-binding method, with a bovine serum albumin (BSA) as a standard.

### 2.3. Substrate specificity through hydrolysis of pNP sugars

The hydrolysis of 21 pNP-linked monosaccharide substrates (pNP  $\alpha$  and  $\beta$ -D-Glc, -Man, -Gal, -GlcNAc, -Xyl, -L-Fuc, -L-Ara, pNP  $\alpha$ -L-Rha, pNP  $\beta$ -D-GalNAc and pNP  $\beta$ -D-GlcA, and 3 furanoses namely pNP  $\beta$ -D-Galf, pNP  $\alpha$ -L-Araf and pNP  $\beta$ -D-Ribf) was assayed in reaction mixture (50  $\mu$ L) containing purified enzyme (5  $\mu$ L; 0.053 mg/mL), pNP monosaccharide substrates (5  $\mu$ L; 10 mM), buffer (8  $\mu$ L; 0.1 M Citric acid/0.2 M Na<sub>2</sub>HPO<sub>4</sub> pH 4.5) and water (32  $\mu$ L). Residual spontaneous hydrolysis of the substrate was determined on sample containing H<sub>2</sub>O instead of enzyme. For paranitrophenol (pNP) containing substrates, after 20 min of incubation at 30 °C, 150  $\mu$ L of sodium carbonate 1 M were added, and produced pNP ( $\epsilon_{405} = 19\,500\text{ M}^{-1}\text{ cm}^{-1}$ ) was quantified by absorbance measurement at 405 nm (Thermo Scientific™ Multiskan™ GO). All kinetics parameters were calculated by fitting of saturation curves (as mean of triplicate measurements) with standard Michaelis-Menten equation, using Prism 6 (GraphPad). One unit (U) of enzyme activity was defined as the amount of enzyme required to liberate 1  $\mu$ mol of pNP per min.

### 2.4. Effects of pH and temperature

The optimum pH was determined by incubating the enzyme in various pH adjusted buffers (0.1 M Citric acid/HCl pH 2; 0.1 M Citric acid/0.2 M Na<sub>2</sub>HPO<sub>4</sub> pH 2.5–6.5; 0.1 M Tris/HCl pH 7–8; 0.1 M Tris/NaOH, pH 8.5; 0.1 M Glycine/NaOH pH 9–9.5). The assays were performed separately in each buffer system containing purified enzyme (5  $\mu$ L; 0.05 mg/mL), pNP  $\beta$ -D-Galf substrate (5  $\mu$ L; 10 mM), buffer (8  $\mu$ L; pH 2–9.5) and water (32  $\mu$ L). After 20 min of incubation at 30 °C (water bath), reaction was terminated (150  $\mu$ L; 1 M Na<sub>2</sub>CO<sub>3</sub>) and absorbance (405 nm) of released pNP was measured using spectrophotometer (Thermo Scientific™ Multiskan™ GO). All the reactions were assayed in triplicate and absorbance values were corrected for the spontaneous hydrolysis of the substrate.

The effect of temperature on the enzyme activity was investigated at temperatures ranging from 10 to 80 °C. The reaction mixtures (50  $\mu$ L), containing purified enzyme (5  $\mu$ L; 0.05 mg/mL), pNP  $\beta$ -D-Galf substrate (5  $\mu$ L; 10 mM), buffer (8  $\mu$ L; 0.1 M Citric acid/0.2 M Na<sub>2</sub>HPO<sub>4</sub> pH 4.5) and water (32  $\mu$ L), were incubated 20 min in thermocycler (Esco Swift MiniPro Thermal Cycler). Afterwards, reactions were stopped (150  $\mu$ L; 1 M Na<sub>2</sub>CO<sub>3</sub>). Then, the absorbance (405 nm) of released pNP was measured (Thermo Scientific™ Multiskan™ GO) and values were corrected for the spontaneous hydrolysis of the substrate. All the reactions were assayed in triplicate and absorbance values were corrected for the spontaneous hydrolysis of the substrate.

### 2.5. Effect of freeze-thaw cycles

Three samples (100  $\mu$ L), containing previously purified enzyme (50  $\mu$ L; 0.235 mg/mL) different concentrations of glycerol (0%, 10%, 20%) and water, were prepared. The samples were precooled on ice (30 min), frozen at –20 °C and were removed from –20 °C after four days, thawed on ice (30 min) and the enzymatic activity of a sample aliquot (10  $\mu$ L) was assayed in reaction mixture (200  $\mu$ L) containing

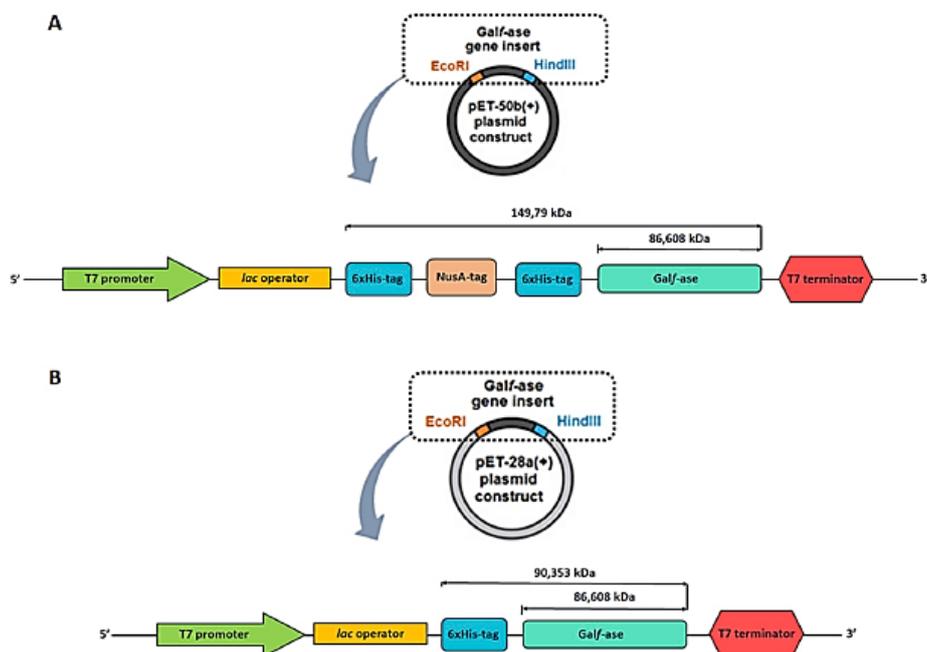


Fig. 2. Schematic representation of pET plasmid constructs cloning/expression regions bearing Galf-ase gene insert. (A) pET-50b(+) plasmid construct ORF coding for N-terminal Nus protein and double 6xHis-tagged Galf-ase. (B) pET-28a(+) plasmid construct ORF coding for N-terminal single peptide 6xHis-tagged Galf-ase.

pNP  $\beta$ -D-Galf substrate (20  $\mu$ L; 10 mM), buffer (35  $\mu$ L; 0.1 M citric acid/0.2 M  $\text{Na}_2\text{HPO}_4$  pH 4.5) and water (135  $\mu$ L) at 37  $^\circ\text{C}$  (water bath), during 10 min, according to previously described enzymatic activity protocol. This predefined freeze and thaw procedure and activity assessment were repeated in four cycles within four days interval. Activity recovery was expressed as a percentage of the activity prior to freezing. All the reactions were assayed separately in duplicate and absorbance values were corrected for the spontaneous hydrolysis of the substrate.

## 2.6. Kinetic studies

Kinetic parameters were determined with pNP  $\beta$ -D-Galf as a substrate in different concentration ranges (0.01 mM–5 mM). The reaction (200  $\mu$ L) containing buffer (20  $\mu$ L; 0.1 M citric acid/0.2 M  $\text{Na}_2\text{HPO}_4$  pH 4.5) and water, if required, was started by addition of the enzyme (2  $\mu$ L; 0.13 mg/mL) and after 20 min of incubation at 37  $^\circ\text{C}$  (water bath), reaction was terminated (100  $\mu$ L; 1 M  $\text{Na}_2\text{CO}_3$ ) and absorbance (405 nm) of released pNP was measured. All the reactions were assayed in triplicate and absorbance values were corrected for the spontaneous hydrolysis of the substrate. The kinetic parameters ( $K_M$ ,  $V_m$ ,  $k_{cat}$ ) were calculated using GraphPad Prism 5 software (GraphPad Software, San Diego, CA, USA).

## 2.7. Inhibition studies

$\text{IC}_{50}$  assay was performed in the presence of different concentration ranges of inhibitors 3–6 (0.1 mM–7 mM), pNP  $\beta$ -D-Galf substrate (3.75  $\mu$ L; 10 mM), buffer (20  $\mu$ L; citric acid/ $\text{Na}_2\text{HPO}_4$ , pH 4.5) and water. The reaction (150  $\mu$ L) was started by addition of the enzyme (11.2  $\mu$ L; 0.067 mg/mL), and after 20 min of incubation at 37  $^\circ\text{C}$  (water bath), reaction was terminated (100  $\mu$ L; 1 M  $\text{Na}_2\text{CO}_3$ ) and absorbance of released pNP was measured (405 nm). Control enzyme activity reactions were carried out and absorbance values were corrected for the spontaneous hydrolysis of the substrate. The type of inhibition for 4 was tested in the presence of different concentration ranges of pNP  $\beta$ -D-Galf

substrate (0.05 mM–1 mM) and with different concentrations of 4 (0.1 mM–1 mM) in the reaction conditions identical to the one previously described. The  $\text{IC}_{50}$  value for each inhibitor, the type of the inhibition and  $K_I$  value for 4 were calculated using GraphPad Prism 5 software (GraphPad Software, San Diego, CA, USA).  $K_I$  value for compounds 3, 5 and 6 was corroborated by the following Cheng & Prusoff equation:  $K_I = \text{IC}_{50}/(1 + [S]/K_M)$  [22].

## 3. Results and discussion

### 3.1. Subcloning, expression and purification of recombinant Galf-ase

Plasmid construct pET-50b(+), containing the gene insert coding for  $\beta$ -D-galactofuranosidase (Galf-ase), was double digested with EcoRI and HindIII restriction enzymes and the resulting 2.4 kb Galf-ase gene fragment was excised and purified from agarose-gel. The gene fragment was ligated to EcoRI and HindIII linearized pET-28a(+) expression vector. Newly obtained plasmid construct was checked for integrity by restriction enzyme analysis and the gene sequence was confirmed by DNA sequencing performed by Eurofins Genomics. Plasmid construct was transformed into *E. coli* Rosetta™ (DE3) (Novagen®) expression system. Galf-ase was expressed as soluble recombinant protein bearing only N-terminal peptide hexahistidine fusion tag (6xHis-tag). 6xHis-tag facilitated selective recombinant protein purification from the bacterial cellular environment by immobilized metal-affinity chromatography (IMAC) and, due to its small size, was less likely to interfere with protein activity and structure [23]. The 6xHis-tagged Galf-ase was expressed and purified by IMAC with a  $\text{Ni}^{2+}$  immobilized on a nitrilotriacetic acid matrix (Ni-NTA) column with a yield of 0.5 mg/L of bacterial culture [24,25]. The purified enzyme displayed an intense band on the SDS-PAGE gel at an apparent molecular weight (90.3 kDa), consistent with that theoretically predicted for the Galf-ase fusion protein (Fig. 3.). This purified Galf-ase was further used for biochemical characterization.

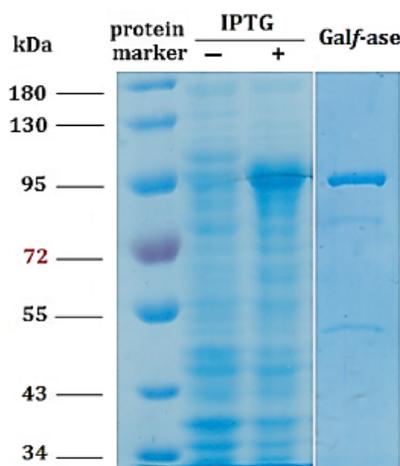


Fig. 3. SDS-PAGE analysis of Galf-ase after IPTG induction and Ni-NTA chromatography purification. Lane 1: protein marker. Lanes 2 & 3: cell's extract before and after IPTG's induction. Lane 4: purified Galf-ase.

### 3.2. Substrate specificity

A Simple one-step colorimetric assay which commonly employs synthetic pNP sugars as artificial substrates was used for the assessment of Galf-ase activity. The assay directly correlates the release of formed pNP with the hydrolytic activity of tested enzyme and is quantified spectrophotometrically at 405 nm. The hydrolytic activity and substrate specificity of Galf-ase was screened against 21 commercially available pNP pyranosyl and furanosyl substrates at 1 mM (Fig. S1.). The only hydrolysed substrate was pNP  $\beta$ -D-Galf and the Galf-ase exhibited no hydrolytic activity towards any other of the 20 tested pNP sugars. It is also noteworthy that, in these experimental conditions, no activity was observed with the pNP  $\alpha$ -L-Araf analogue nor the pNP  $\beta$ -D-Ribf. This result is in total agreement with our previous observation [15].

### 3.3. Temperature and pH properties

The biochemical properties of purified Galf-ase were investigated with pNP  $\beta$ -D-Galf as a substrate in different buffer systems covering the pH intervals between pH 2 and 9.5 and temperature intervals between 10 and 80 °C. The optimum pH was comprised in a narrow range between 3.5 and 5, significantly decreasing above pH 5. The peak of maximum activity is present at pH 4.5 (Fig. 4. Left). Once the optimum

pH conditions were established, the influence of different temperatures on activity was also measured. For 20 min reactions, the optimal temperature was at 60 °C with a clear peak of activity (Fig. 4. Right). Temperatures above the optimum value resulted in dramatic loss of activity. However, due to practical constraints, all the following activity assays were assessed at the physiological temperature of 37 °C. Still, this current Galf-ase biocatalyst exhibits different and improved physico-chemical properties toward the previously reported one, with a maximum pH shift from 5.5 to 4.5 and a maximum temperature at 60 °C, thus demonstrating the drawback of large fusion tags [15].

### 3.4. Freeze-thaw stability

Galf-ase stability towards multiple freeze and thaw cycles was assayed over a different storage conditions, without and in the presence of glycerol (10% and 20%), and throughout 16 days. Percent of activity was normalized to an initial activity measured for each storage condition sample. Galf-ase exhibited stability across all storage conditions tested and maintained minimally 70% activity after three cycles (Fig. S2.). After four cycles, Galf-ase demonstrated low activity (40%) over all storage conditions and exhibited higher sensitivity to inactivation due to multiple freezing. No cryoprotectant was included in this assay to test the stability of crude Galf-ase sample which showed to be resistant to inactivation to multiple freeze-thaw cycles and retaining 80% activity after three cycles. The addition of glycerol did not significantly improve or conserve enzyme stability when compared to one observed in the absence of a cryoprotectant. Glycerol is a common cryoprotectant which protects proteins from inactivation during freezing and thawing by inhibiting the formation of ice crystals upon freezing [26]. Also, it is worth mentioning that the purified enzyme remains fairly active over a minimum of three weeks when stored at 4 °C without presence of any additives such are cryoprotectants, protease inhibitors, reducing agents, metal chelators or antimicrobial agents. All together these experiments demonstrate the good stability of the Galf-ase under storage condition, which will be highly convenient in order to use it during chemo-enzymatic synthesis.

### 3.5. Kinetic analysis

To determine catalytic parameters, Galf-ase activity was further investigated with pNP  $\beta$ -D-Galf as a substrate at optimum pH and at appropriate temperature. The Galf-ase exhibited typical Michaelis-Menten kinetics, with  $K_M$  value of 250  $\mu$ M and with  $k_{cat}$  value of 3.5  $s^{-1}$  giving a  $k_{cat}/K_M$  of 14  $mM^{-1}s^{-1}$  (Fig. 5 and Table 1). Investigation of kinetic parameters was also extended to pNP  $\alpha$ -L-Araf, despite displaying no hydrolysis when screened for substrate specificity. The pNP  $\alpha$ -L-Araf did not appear to show saturation kinetics with concentrations

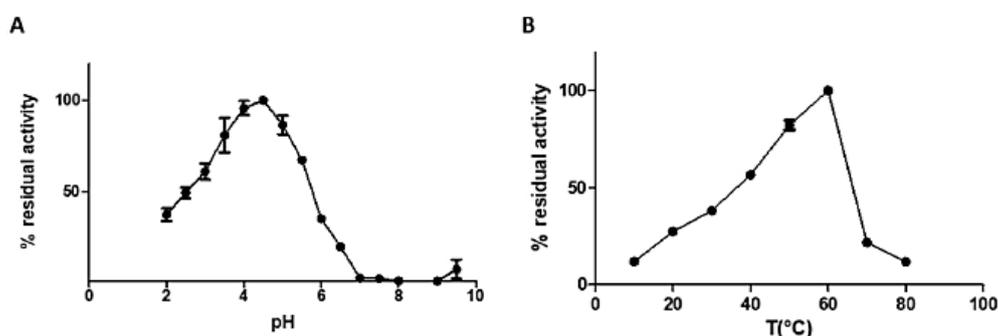


Fig. 4. A: Influence of pH on Galf-ase hydrolytic activity towards pNP  $\beta$ -D-Galf 1. B: Influence of temperature on Galf-ase hydrolytic activity towards pNP  $\beta$ -D-Galf 1. Error bars correspond to the SD of 3 independent measurements.

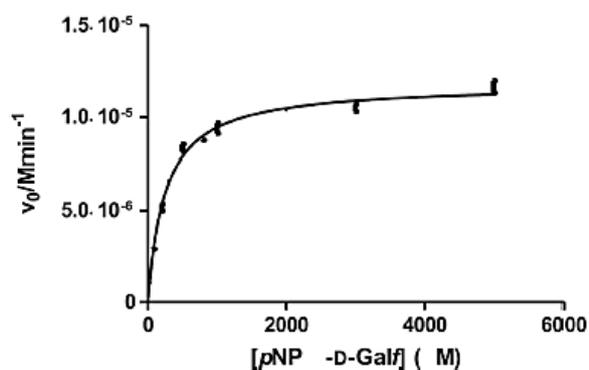


Fig. 5. Michaelis-Menten plot of pNP  $\beta$ -D-Galf 1 hydrolysis reaction catalysed by Galf-ase. Mean values and SD error bars were calculated from three independent experiments and from a single protein preparation. Error bars correspond to the SD of 3 independent measurements.

as high as 15 mM (Fig. S3.). The assay did not yield interpretable kinetic parameters since the hydrolysis rate was too low to give accurate values. These observations emphasized that Galf-ase has a really poor Araf-ase activity as it badly recognised the  $\alpha$ -L-arabinofuranose moiety, thus demonstrating the crucial role of the C-6 hydroxymethyl group in the specific recognition by the Galf-ase. In addition, when compared to the recombinant  $\alpha$ -L-Araf hydrolases AbfD3 and Araf51 owing dual specificity for  $\alpha$ -L-Araf and  $\beta$ -D-Galf (Table 1, lines 1 & 2), the specificity of Galf-ase is incredibly increased toward the hexofuranosyl moiety by a factor of about 100-fold. When compared to the only two reported cloned Galf-ases (line 3), from *Streptomyces sp.* strains JHA19 and JHA26, the values are not in the same order of magnitude. For Galf-ase JHA19 and JHA26,  $K_M$  value ranges from 4.4 to 6.8 mM respectively [15], yielding at least a 18-fold lower substrate affinity compared to Galf-ase's  $K_M$  which is in the  $\mu$ M level (line 4). Finally, the Galf-ase parameters described herein are the best reported to date and fully compatible with the development of the specific and efficient biocatalyzed incorporation of Galf entities into more complex glycoconjugates.

### 3.6. Inhibition studies

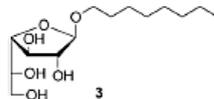
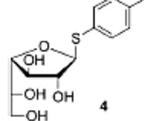
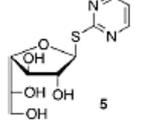
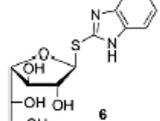
The inhibitory properties of several compounds, i.e. alkyl, thioheteroaryl and thioimidoyl  $\beta$ -D-galactofuranosides 3–6, were investigated by incubating the enzyme, Galf-ase, in the presence of both the substrate pNP  $\beta$ -D-Galf and compounds 3–6 (Fig. S4.). These compounds were tested as inhibitors because (i) the glycone part, that is  $\beta$ -D-Galf, is recognised by the active site, (ii) the aglycone part is mimicking a carbohydrate moiety and (iii) to probe the stability of the thioglycosidic bond [20]. It has been reported previously that alkyl, aryl and heteroaryl 1-thio- $\beta$ -D-galactofuranoside derivatives are good inhibitors of exo  $\beta$ -D-galactofuranosidase, isolated from *P. fellutanum* culture media, and D-galactono-1,4-lactone has been evaluated as a reference inhibitor with an  $IC_{50}$  value of 0.02 mM [28]. In our hands, compounds 3, 5 and

Table 1  
 $K_M$  for synthetic pNP furanoses 1 and 2 toward two different arabinofuranosidases and two Galf-ase.

Enzyme	$K_M$ for 1 (mM)	$k_{cat}$ ( $s^{-1}$ )	$k_{cat}/K_M$ $s^{-1} \cdot mM^{-1}$	$K_M$ for 2 (mM)	$k_{cat}$ ( $s^{-1}$ )	$k_{cat}/K_M$ $s^{-1} \cdot mM^{-1}$
AbfD <sub>3</sub> [12]	> 50	15	0.13	3.5	425	122
Araf51 [13,27]	53	12	0.23	0.25	103	412
JHA 19 & 26 [15]	4.4 to 6.8	Only $V_M$	-	-	-	-
Galf-ase	0.25	3.5	14	nd	nd	nd

nd: not determinable.

Table 2  
Inhibition studies with galactofuranosides 3–6: determination of their respective  $IC_{50}$  &  $K_I$ .

Compound	$IC_{50}$ (mM)	$K_I$ (mM)
	6	3.0
	0.8	0.4
	5	2.5
	3	1.5

6 showed weak inhibitory activity with  $IC_{50}$  values of 6, 5 and 3 mM respectively, while Galf-ase was more sensitive to the inhibition by thiogalactofuranoside 4 (Table 2.). This latter showed the best inhibition activity, with  $IC_{50}$  value of 0.8 mM. The analysis of the Lineweaver-Burk plot (Fig. 6.) indicated that the compound 4 is a competitive inhibitor of the Galf-ase with  $K_I$  value of 0.4 mM. Nevertheless, compared to the only reference L-arabino-1,4-lactone inhibitor previously probed on Galf-ase [15], 4 exhibited a 100 times increased inhibition, that can be mostly attributed to its structural analogy to the artificial substrate.

### 4. Conclusion

In this work, we provided a complete biochemical and kinetic characterization of a subcloned recombinant Galf-ase since the recent publication of the first ever cloned one. This enzyme was obtained with a reasonable yield of 0.5 mg/L of culture, and optimally active at pH 4.5 and at 60 °C. It is inhibited by substrate's structural thio analogue ( $IC_{50}$  = 0.8 mM), which proved to be a moderate competitive inhibitor ( $K_I$  = 0.4 mM). It exhibits substrate specificity exclusively towards pNP  $\beta$ -D-Galf, giving a  $K_M$  value of 250  $\mu$ M, and the highest enzymatic efficiency ever observed of 14  $mM^{-1} s^{-1}$ . Since to date, only a few furanosidases have been involved in the chemoenzymatic approaches, it would be of great interest to explore and evaluate the potential synthetic ability of this enzyme to act both as wild-type in transglycosylation [29] or after site-directed mutagenesis as an efficient thioligase

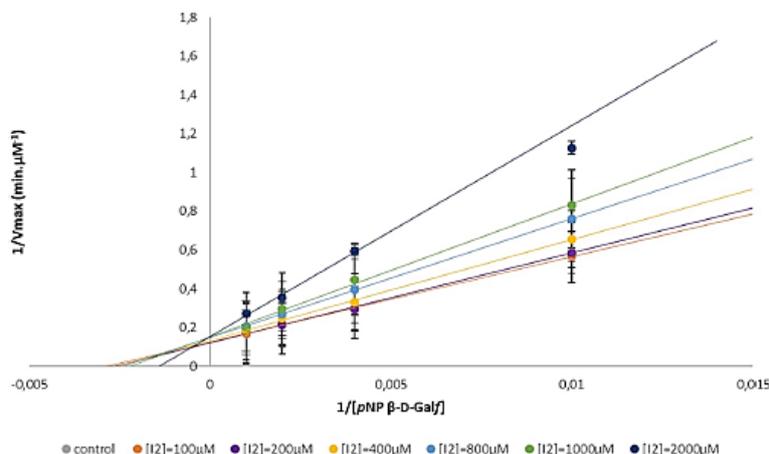


Fig. 6. Lineweaver–Burk plot of competitive inhibition of GalF-ase hydrolytic activity towards pNP  $\beta$ -D-GalF by 4. Mean values and SD error bars were calculated from two independent experiments and from a single protein preparation.

[30] in the efficient preparation of galactofuranosyl conjugates. Works are currently under progress in this direction and results will be reported in due course.

#### Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.carres.2019.05.011>.

#### Funding

This work was supported by the APR IR NeoLect project from the Région Centre-Val de Loire, France. S. V. E. and S. P. acknowledge support from La Ligue Régionale contre le Cancer and from Institut National de la Santé et de la Recherche Médicale (INSERM).

#### References

- [1] A. Varki, Biological roles of oligosaccharides: all of the theories are correct, *Glycobiology* 3 (1993) 97–130, <https://doi.org/10.1093/glycob/3.2.97>.
- [2] M.R. Richards, T.L. Lowary, Chemistry and biology of galactofuranose-containing polysaccharides, *ChemBiochem* 10 (2009) 1920–1938, <https://doi.org/10.1002/cbic.200900208>.
- [3] P. Peltier, R. Euzen, R. Daniellou, C. Nugier-Chauvin, V. Ferrières, Recent knowledge and innovations related to hexofuranosides: structure, synthesis and applications, *Carbohydr. Res.* 343 (2008), <https://doi.org/10.1016/j.carres.2008.02.010>.
- [4] I. Chlubnova, L. Legentil, R. Dureau, A. Pennec, M. Almendros, R. Daniellou, C. Nugier-Chauvin, V. Ferrières, Specific and non-specific enzymes for furanosyl-containing conjugates: biosynthesis, metabolism, and chemo-enzymatic synthesis, *Carbohydr. Res.* 356 (2012), <https://doi.org/10.1016/j.carres.2012.04.002>.
- [5] P.M. Nassau, S.L. Martin, R.E. Brown, A. Weston, D. Monsey, M.R. McNeil, K. Duncan, Galactofuranose biosynthesis in *Escherichia coli* K-12: identification and cloning of UDP-galactopyranose mutase, *J. Bacteriol.* 178 (1996) 1047–1052, <https://doi.org/10.1128/jb.178.4.1047-1052.1996>.
- [6] J. Ati, C. Colas, P. Lafite, R.P. Sweeney, R.B. Zheng, T.L. Lowary, R. Daniellou, The LPG1x family from *Leishmania major* is constituted of rare eukaryotic galactofuranosyltransferases with unprecedented catalytic properties, *Sci. Rep.* 8 (2018) 17566, <https://doi.org/10.1038/s41598-018-35847-w>.
- [7] M. Rietschel-Berst, N.H. Jentoft, P.D. Rick, C. Pletcher, F. Fang, J.E. Gander, Extracellular exo-Galactofuranosidase from *Penicillium charlesii*, *J. Biol. Chem.* 252 (1977) 3219–3226 <http://www.ncbi.nlm.nih.gov/pubmed/863879>.
- [8] C. Marino, C. Gallo-Rodríguez, R.M. de Lederkremer, Galactofuranosyl-containing glycans: occurrence, synthesis and biochemistry, (2012), <https://doi.org/10.1063/1.3243690>.
- [9] H.J. Kamphuis, G.A. De Ruyter, F.M. Rombouts, P. Mischnick, J.H. Van Boom, A.W. Van Bruggen-Van Der Lugt, New structural features of the antigenic extracellular polysaccharides of *Penicillium* and *Aspergillus* species revealed with exo- $\beta$ -D-galactofuranosidase, *J. Bacteriol.* 174 (2016) 6096–6102, <https://doi.org/10.1128/jb.174.19.6096-6102.1992>.
- [10] N. Ramli, M. Fujinaga, M. Tabuchi, K. Takegawa, S. Iwahara, Isolation and characterization of a novel endo- $\beta$ -galactofuranosidase from *Bacillus* sp, *Biosci. Biotechnol. Biochem.* 59 (1995) 1856–1860, <https://doi.org/10.1271/bbb.59.1856>.
- [11] L.C. Miletti, K. Mariño, C. Marino, W. Colli, M.J. Manso Alves, R.M. De Lederkremer, Evidence for exo  $\beta$ -D-galactofuranosidase in *Trypanosoma cruzi*, *Mol. Biochem. Parasitol.* 127 (2003) 85–88, [https://doi.org/10.1016/S0166-6851\(02\)00307-9](https://doi.org/10.1016/S0166-6851(02)00307-9).
- [12] R. Euzen, G. Lopez, C. Nugier-Chauvin, V. Ferrières, D. Plusquellec, C. Rémond, M. O'Donohue, A chemoenzymatic approach for the synthesis of unnatural disaccharides containing D-Galacto- or D-Fucofuranosides, *Eur. J. Org. Chem.* 2005 (2005) 4860–4869, <https://doi.org/10.1002/ejoc.200500525>.
- [13] I. Chlubnová, D. Filipp, V. Spiwok, H. Dvořáková, R. Daniellou, C. Nugier-Chauvin, B. Králová, V. Ferrières, Enzymatic synthesis of oligo-D-galactofuranosides and L-arabinofuranosides: from molecular dynamics to immunological assays, *Org. Biomol. Chem.* 8 (2010), <https://doi.org/10.1039/b926988f>.
- [14] I. Chlubnová, B. Králová, H. Dvořáková, P. Hošek, V. Spiwok, D. Filipp, C. Nugier-Chauvin, R. Daniellou, V. Ferrières, The versatile enzyme AraF51 allowed efficient synthesis of rare pathogen-related  $\beta$ -D-galactofuranosyl-pyranoside disaccharides, *Org. Biomol. Chem.* 12 (2014), <https://doi.org/10.1039/c3ob42519c>.
- [15] E. Matsunaga, Y. Higuchi, K. Mori, N. Yairo, T. Oka, S. Shinozuka, K. Tashiro, M. Izumi, S. Kuhara, K. Takegawa, Identification and characterization of a novel galactofuranose-specific  $\beta$ -D-galactofuranosidase from *Streptomyces* species, *PLoS One* 10 (2015) 1–16, <https://doi.org/10.1371/journal.pone.0137230>.
- [16] E. Matsunaga, Y. Higuchi, K. Mori, K. Tashiro, S. Kuhara, K. Takegawa, Draft genome sequence of *Streptomyces* sp. JHA19, a strain that possesses  $\beta$ -D-galactofuranosidase activity, *Genome Announc.* 3 (2015) 3–4, <https://doi.org/10.1128/genomeA.01171-15>.
- [17] E. Matsunaga, Y. Higuchi, K. Mori, N. Yairo, S. Toyota, T. Oka, K. Tashiro, K. Takegawa, Characterization of a PA14 domain-containing galactofuranose-specific  $\beta$ -D-galactofuranosidase from *Streptomyces* sp, *Biosci. Biotechnol. Biochem.* 81 (2017) 1314–1319, <https://doi.org/10.1080/09168451.2017.1300518>.
- [18] G.D. Davis, C. Elisee, D.M. Newham, R.G. Harrison, New fusion protein systems designed to give soluble expression in *Escherichia coli*, *Biotechnol. Bioeng.* 65 (1999) 382–388 <http://www.ncbi.nlm.nih.gov/pubmed/10506413> (Accessed 12 April 2019).
- [19] L. Legentil, J.-L. Audic, R. Daniellou, C. Nugier-Chauvin, V. Ferrières, Studies of a furanoside as antimycobacterial agent loaded into a biodegradable PBAT/sodium caseinate support, *Carbohydr. Res.* 346 (2011), <https://doi.org/10.1016/j.carres.2011.04.025>.
- [20] G. Lopez, R. Daniellou, M. O'Donohue, V. Ferrières, C. Nugier-Chauvin, Thioimidoyl furanosides as first inhibitors of the  $\alpha$ -L-arabinofuranosidase AbfD3, *Bioorg. Med. Chem. Lett* 17 (2007) 434–438, <https://doi.org/10.1016/j.bmcl.2006.10.032>.
- [21] U.K. Laemmli, Cleavage of structural proteins during the assembly of the head of bacteriophage T4, *Nature* 227 (1970) 680–685 <http://www.ncbi.nlm.nih.gov/pubmed/5432063> (Accessed 16 May 2019).
- [22] C. Yung-Chi, W.H. Prusoff, Relationship between the inhibition constant (KI) and the concentration of inhibitor which causes 50 per cent inhibition (I50) of an enzymatic reaction, *Biochem. Pharmacol.* 22 (1973) 3099–3108, [https://doi.org/10.1016/0006-2952\(73\)90196-2](https://doi.org/10.1016/0006-2952(73)90196-2).
- [23] G.L. Rosano, E.A. Ceccarelli, Recombinant protein expression in *Escherichia coli*: advances and challenges, *Front. Microbiol.* 5 (2014) 1–17, <https://doi.org/10.3389/fmicb.2014.00172>.
- [24] J.J. Winzerling, P. Berna, J. Porath, How to use immobilized metal ion affinity chromatography, *Methods* 4 (1992) 4–13, [https://doi.org/10.1016/1046-2023\(92\)90052-A](https://doi.org/10.1016/1046-2023(92)90052-A).
- [25] V. Gaberc-porekar, V. Menart, Perspectives of immobilized-metal affinity chromatography, *J. Biochem. Biophys. Methods* 49 (2001) 335–360.
- [26] N.J. Alves, K.B. Turner, L.L. Medintz, S.A. Walper, Protecting enzymatic function

- through directed packaging into bacterial outer membrane vesicles, *Sci. Rep.* 6 (2016), <https://doi.org/10.1038/srep24866>.
- [27] E.J. Taylor, N.L. Smith, J.P. Turkenburg, S. D'souza, H.J. Gilbert, G.J. Davies, Structural insight into the ligand specificity of a thermostable family 51 arabinofuranosidase, Araf 51, from *Clostridium thermocellum*, *Biochem. J.* 395 (2006) 31–37, <https://doi.org/10.1042/BJ20051780>.
- [28] E. Repetto, C. Marino, O. Varela, Synthesis of the (1→6)-linked thiodisaccharide of galactofuranose: inhibitory activity against a  $\beta$ -galactofuranosidase, *Bioorg. Med. Chem.* 21 (2013) 3327–3333, <https://doi.org/10.1016/j.bmc.2013.02.057>.
- [29] A. Penneç, R. Daniellou, P. Loyer, C. Nugier-Chauvin, V. Ferrières, Araf51 with improved transglycosylation activities: one engineered biocatalyst for one specific acceptor, *Carbohydr. Res.* 402 (2015) 50–55, <https://doi.org/10.1016/j.carres.2014.10.031>.
- [30] M. Almendros, D. Danalev, M. Franois-Heude, P. Loyer, L. Legentil, C. Nugier-Chauvin, R. Daniellou, V. Ferrières, Exploring the synthetic potency of the first furanothioglycosylase through original remote activation, *Org. Biomol. Chem.* 9 (2011) 8371–8378, <https://doi.org/10.1039/c1ob06227a>.

41

The results obtained during the three years of this PhD thesis were valorised in one publication and the work carried out was presented at several national and international conferences in the form of oral communication and posters:

### **Publication:**

**M. Seničar**, L. Legentil, V. Ferrières, S. V. Eliseeva, S. Petoud, K. Takegawa, P. Lafite, R. Daniellou, Galactofuranosidase from JHA19 *Streptomyces sp.*: subcloning and biochemical characterization. *Carbohydr. Res.*, **480** (2019) 35–41.

### **Oral communication:**

J. Ati, **M. Seničar**, R. B. Zheng, S. V. Eliseeva, P. Lafite, S. Petoud, T. L. Lowary, R. Daniellou, *The Maker, The Breaker & The Shapeshifter*, 27<sup>èmes</sup> Journées du Groupe Français des Glycosciences (GFG), May 21–25, 2018, Nouan-le-Fuzelier (45), France.

### **Posters:**

**M. Seničar**, P. Lafite, S. V. Eliseeva, S. Petoud, R. Daniellou, *Design of furanoside-specific neolectins: bioengineering, imaging and diagnosis*, 1<sup>èmes</sup> Journées du Groupement de Recherche «Agents d'Imagerie Moléculaire» (GDR AIM), June 4–5, 2018, Orléans (45), France.

**M. Seničar**, P. Lafite, S. V. Eliseeva, S. Petoud, R. Daniellou, *Design of furanoside-specific neolectins: bioengineering, imaging and diagnosis*, 31<sup>e</sup> Colloque Biotechnocentre, October 11–12, 2018, Domaine de Seillac (41), France. – **Best poster award** –

**M. Seničar**, P. Lafite, S. V. Eliseeva, S. Petoud, R. Daniellou, *Galactofuranosidase – thioglycosylase potency for thiogalactofuranoside synthesis*, 14<sup>th</sup> International Symposium on Biocatalysis and Biotransformations (BioTrans 2019), July 7–11, 2019, Groningen, The Netherlands.

Mateja SENICAR

## Mise au point de NéoLectines spécifiques des furanosides: bioingénierie, imagerie et diagnostic

### Résumé:

Le ciblage spécifique des infections virales et bactériennes revêt une importance majeure pour le diagnostic précoce de nombreuses maladies. Le galactofuranose (Gal $f$ ) est absent chez l'homme mais se trouve en tant que fragment de glycoconjugués dans un grand nombre d'agents pathogènes pour les Hommes (*Aspergillus*, *Leishmania*, *Trypanosoma*, *Mycobacterium*), offrant ainsi la possibilité de cibler ce sucre dans des applications biotechnologiques utilisant des enzymes impliquées dans son métabolisme (UDP-galactose), mutase (UGM), galactofuranosyl transférase (Gal $f$ T), galactofuranosidase (Gal $f$ -ase). Récemment, la première enzyme spécifique du Gal $f$  de *Streptomyces spp.* et son gène codant ont été rapportés. Ce travail de thèse se concentre sur Gal $f$ -ase et se compose de trois parties principales. La première partie concerne la caractérisation de cette enzyme, y compris le sous-clonage, la surexpression, la purification et la cinétique ( $k_{cat}$ ,  $K_M$ ,  $k_{cat}/K_M$ ), la détermination de l'activité optimale (pH et température) et de la stabilité au stockage. La deuxième partie consiste à appliquer une technologie d'ingénierie enzymatique (mutagenèse dirigée par site) pour générer des néolectines (protéines de reconnaissance des sucres) à partir du type sauvage. Plusieurs enzymes Gal $f$ -ase mutantes ont été surexprimées, purifiées et caractérisées cinétiquement ( $k_{cat}$ ,  $K_M$ ,  $k_{cat}/K_M$ ) pour leur potentiel d'utilisation en tant que néolectines et thioglycoligases pour la synthèse de furanothioglycosides. La troisième et dernière partie porte sur les procédures à suivre pour créer des nanoparticules luminescentes destinées à contenir des complexes de lanthanides et à lier les néolectines de Gal $f$ -ase à sa surface. Les conjugués luminescents résultants seront testés en tant que sondes dans l'imagerie optique proche infrarouge de Gal $f$  située à la surface de microorganismes.

**Mots clés:** Galactofuranosidase, Furanothioglycoligase, Neolectins, Lanthanides

## Design of galactofuranosyl-specific NeoLectins: bioengineering, imaging and diagnosis

### Summary:

Specific targeting of viral and bacterial infections is of major importance for early diagnosis of many diseases. Galactofuranose (Gal $f$ ) is absent in humans but is found as glycoconjugate moiety in a large number of human pathogens (*Aspergillus*, *Leishmania*, *Trypanosoma*, *Mycobacterium*) thus offering possibilities to target this carbohydrate in biotechnological applications using enzymes involved in its metabolism (UDP-galactopyranose mutase (UGM), Galactofuranosyl transferase (Gal $f$ T), Galactofuranosidase (Gal $f$ -ase). Recently, the first Gal $f$ -specific enzyme from *Streptomyces spp.* and gene encoding it has been reported. This thesis work focuses on Gal $f$ -ase and consists of three main parts. First part is groundwork characterization of this enzyme including subcloning, overexpression, purification and kinetic characterization ( $k_{cat}$ ,  $K_M$ ,  $k_{cat}/K_M$ ). Also, determination of optimum activity (pH and temperature) and storage stability conditions. Second part consists of applying enzyme engineering technology (site-directed mutagenesis) to generate neolectins (carbohydrate-binding proteins) from the wild-type. Several mutant Gal $f$ -ase enzymes have been overexpressed, purified and kinetically characterized ( $k_{cat}$ ,  $K_M$ ,  $k_{cat}/K_M$ ) on their potential to be used as neolectins and thioglycoligases for furanothioglycoside synthesis. Finally, third part focuses on procedures in order to create luminescent nanoparticle scaffold to bear lanthanide complexes and to bind Gal $f$ -ase neolectins onto its surface. The resulting luminescent conjugates will be tested as probes in near-infrared optical imaging of Gal $f$  located on the surface of microorganisms.

**Keywords:** Galactofuranosidase, Furanothioglycoligase, Neolectins, Lanthanides



Université d'Orléans, Institut de Chimie Organique et Analytique,  
UMR 7311, CNRS, Rue de Chartres BP6759, 45100 Orléans, France

Centre de Biophysique Moléculaire, UPR 4301, CNRS,  
Rue Charles Sadron CS 8005, 45071 Orléans, France

