ENZYME BIOPROSPECTING OF MICROBIAL GLYCOSIDASES AND DIOXYGENASES FOR BIOCATALYSIS.

Francesca Mensitieri

Dottorato in Biotecnologie – 31° ciclo

Università di Napoli Federico II

Dottorato in Biotecnologie – 31° ciclo

Università di Napoli Federico II

ENZYME BIOPROSPECTING OF MICROBIAL GLYCOSIDASES AND DIOXYGENASES FOR BIOCATALYSIS.

Francesca Mensitieri

Dottorando: Francesca Mensitieri

Relatore: Prof. Alberto Di Donato

Coordinatore: Prof. Giovanni Sannia

Ai miei genitori

INDEX

RIASSUNTO pag. 5

ABSTRACT pag. 13

ABBREVIATIONS pag. 15

CHAPTER 1 - Introduction pag. 17

1.1 Role of microbial enzymes in biotransformations and green chemistry. pag. 19

1.2 Ring-cleavage extradiol dioxygenases and their biotechnological applications. pag. 22

1.3 Glycosyl hydrolases, classification and biotechnological applications. pag. 28

1.3.1 The α-L-rhamnosidases. pag. 33

1.3.2 The β-D-galacyosidases. pag. 36

Aims of the thesis pag. 39

CHAPTER 2 – *Novosphingobium* sp. PP1Y as a source of extradiol ring-cleavage dioxygenases of biotechnological relevance. pag. 41

2.1 Recombinant expression and purification of four ERCDs from *N*. sp. PP1Y. pag. 43

2.2 Biochemical characterization. pag. 47

2.2.1. Iron quantification. pag. 47

2.2.2. *Homology modeling*, specific activity and kinetic parameters on catecholic substrates. pag. 48

2.2.3. Bioconversion and kinetic parameters using catechol extrogens as substrates. pag. 53

CHAPTER 3 – RHA-P, biotechnological relevance of a novel α-L-rhamnosidase from *Novosphingobium* sp. PP1Y. pag. 59

3.1 RHA-P subcloning, recombinant expression and purification. pag. 61

3.2 RHA-P *Homology modeling*. pag. 66

3.3 Evidences for calcium presence in RHA-P. pag. 69

3.4 RHA-P substrate specificity. pag. 70

3.5 Kinetic characterization of RHA-P activity on natural flavonoids. pag. 73

3.6 Whole cell bioconversion of naringin using RHA-P. pag. 75

CHAPTER 4 – Identification of novel glycosyl hydrolases from Bacteroides, active on *Mycobacterium tuberculosis* cell wall polysaccharides. pag. 79

4.1 Bacteroides strains as source of novel GHs active towards *Mycobacterium tuberculosis* cell wall. pag. 81

4.2 Cloning, recombinant expression and purification of proteins from *B. finegoldii* DSM 17565. pag. 86

4.3 Activity screening of BF proteins. pag. 87
4.4 Preliminary biochemical characterization of BF04084. pag. 98
4.5 Growth screening of Bacteroides strains on arabinan from *M. smegmatis* pag. 102
CHAPTER 5 – Concluding remarks. pag. 107
REFERENCES pag. 117
APPENDICES pag. 129
APPENDIX I – Materials and Methods. pag. 131
I.1 Generals. pag. 131
I.2 Cloning procedures. pag. 132
I.2.1. Construction of the pET22b(+)/*rha-his* expression vector. pag. 132
I.2.2. Cloning of genes from PULs 41 and 49 of *B. finegoldii* in pET-28a(+). pag. 132
I.3 Recombinant expression procedures. pag. 133
I.3.1. Recombinant expression of the 4 ERCDs from *N*. sp. PP1Y. pag. 133
I.3.2. Recombinant expression of protein RHA-P from *N*. sp. PP1Y. pag. 133
I.3.3. Recombinant expression of 9 proteins from PULs 41 and 49 of *B. finegoldii* pag. 134
I.4 Purification procedures. pag. 134
I.4.1. Purification of recombinant PP26077, PP00124 and PP00193 ERCDs from *N*. sp. PP1Y. pag. 134
I.4.2. Purification of recombinant PP28735 ERCD from *N*. sp. PP1Y. pag. 135
I.4.3. Purification of recombinant RHA-P protein from *N*. sp. PP1Y. pag. 135
I.4.4. Purification of the 9 recombinant proteins from PULs 41 and 49 of *B. finegoldii* pag. 136
I.4.5. Purification by size exclusion chromatography of galactobiose obtained from Galactan after the reaction with protein BF03696 pag. 137
I.4.6. Purification of arabinan obtained from arabinogalactan hydrolysis using proteins BF03696 and BF04084. pag. 137
I.5 Enzyme activity assays. pag. 138
I.5.1. Activity assays of the recombinant ERCDs using 4 catecholic substrates pag. 138
I.5.2. Bioconversion of 2-hydroxy extradiol and 4-hydroxy-extradiol using the recombinant ERCDs. pag. 139
I.5.3. Steady state kinetics of the recombinant ERCDs. pag. 140
I.5.4. Activity assays of the recombinant RHA-P on *p*NPR pag. 140
I.5.5. Steady state kinetics of the recombinant RHA-P on *p*NPR. pag. 140
I.5.6. Steady state kinetics of the recombinant RHA-P on natural flavonoids. pag. 140
I.5.7. Activity assays of RHA-P on synthetic and natural substrates. pag. 141

I.5.8. Activity assays of the recombinant proteins from *B. finegoldii* on *p*NP-sugars pag. 141

I.5.9. Steady state kinetics of BF04084 from *B. finegoldii* on *p*NP-β-D-gal*f* pag. 142

I.5.10. Kinetic parameter determination of protein BF04084 from *B. finegoldii* on the galactobiose obtained from Galactan. pag. 142

I.5.11. Activity assays of recombinant proteins from *B. finegoldii* on natural substrates. pag. 143

I.5.12. Activity assays of proteins BF03696 and BF04084 from *B. finegoldii* on synthetic substrates pag. 143

I.6 Ferene S assay. pag. 144

I.7 Enzymatic conversion of 4-OHE and 2-OHE using ERCD. pag. 144

I.8 HPLC analysis of the reaction products of 4-OHE and 2-OHE enzymatic conversion pag. 144

I.9 Preparation of apo RHA-P and recovery of the enzymatic activity. pag. 145

I.10 Whole-cell enzymatic conversion of naringin using RHA-P. pag. 145

I.11 HPLC analysis of the naringin enzymatic conversion. pag. 146

I.12 Growth of Bacteroides strains. pag. 146

I.13 RT-PCR and q-PCR of PULs 41 and 49 in *B. finegoldii.* pag. 147

I.14 General methods. pag. 147

APPENDIX II – Scientific pubblications. pag. 149

APPENDIX III – Communications. pag. 151

APPENDIX IV – Experiences in foreign laboratories. pag. 155

APPENDIX V – Published papers. pag. 157

RIASSUNTO

BASE SCIENTIFICA DEL PROGETTO

Biocatalisi e Green chemistry.

Le bioconversioni sono processi che impiegano sia enzimi purificati che microrganismi per la trasformazione di molecole organiche in composti ad alto valore aggiunto, utilizzabili in diversi settori industriali, dall’industria chimica alla farmaceutica, dalla tessile a quella agroindustriale.

Numerosi sono i vantaggi dell’utilizzo delle tecnologie di biocatalisi, applicate in alternativa e/o sequenzialmente alle metodiche di sintesi chimica che hanno spesso costi elevati, specificità di reazione di gran lunga inferiore a quella ottenibile con gli enzimi, notevole impatto ambientale e difficoltà legate alle condizioni di pressione, temperatura e pH utilizzate. Gli enzimi, d’altro canto, agiscono in condizioni blande di temperatura e di pH, e vantano caratteristiche uniche di enantio- e regioselettività che sono difficilmente ottenibili mediante le sintesi chimiche.

L’industria dei biocatalizzatori gode quindi di elevate potenzialità, soprattutto per gli avanzamenti nelle tecniche di sequenziamento genomico, manipolazione del DNA ed ingegneria proteica.

Nell’ambito della ricerca di nuovi biocatalizzatori, i microrganismi sono le fonti preferite, grazie alla facilità della loro manipolazione in laboratorio e all’enorme variabilità di adattamenti mostrati dai microrganismi isolati in nicchie ecologiche “peculiari”. Tali ambienti conferiscono spesso ai microorganismi autoctoni la capacità di crescere in condizioni chimico-fisiche estreme e di metabolizzare composti inquinanti e difficilmente degradabili.

Negli ultimi anni, la richiesta di metodiche di bioconversione enzimatica è aumentata notevolmente, anche in concomitanza di una crescente sensibilizzazione riguardo le problematiche relative all’inquinamento ambientale. Il termine “green chemistry” è stato coniato, riferendosi alla progettazione e sviluppo di processi produttivi, materiali e tecnologie aventi un minore impatto sull’ambiente rispetto a quelli tradizionali. Un’ ampia varietà di sostanze chimiche sono attualmente prodotte con l’ausilio di enzimi in processi dell’industria agroalimentare e farmaceutica. In questo contesto, l’‘’enzyme bioprospecting’’ è un approccio scientifico che prevede l’identificazione di nuovi enzimi da sfruttare in ambito industriale attingendo da habitat caratterizzati da peculiari condizioni chimico-fisiche.

Diossigenasi extradioliche - applicazione biotecnologica.

Numerosi batteri sono in grado di utilizzare composti aromatici come unica fonte di carbonio ed energia. Lo step iniziale del metabolismo aerobico degli idrocarburi policiclici aromatici è operato da mono- o diossigenasi che catalizzano l’addizione di due gruppi ossidrilici sull’anello aromatico, portando alla formazione di un cis-diidrodiolo. Il taglio dell'anello ossidrilato è generalmente effettuato da diossigenasi definite come “ring cleavage dioxygenases”. Le diossigenasi possono essere intradioliche (IRCDs) o extradioliche (ERCDs) che agiscono tagliando gli anelli aromatici diossidrilati in posizioni diverse e ottenendo prodotti con diversi gruppi funzionali.

Le RCDs sono biocatalizzatori particolarmente interessanti da un punto dio vista biotecnologico poiché sono in grado di modificare un'ampia varietà di composti in diverse posizioni e con elevata regio- e stereospecificità. L'utilizzo di questi enzimi, e in particolare delle ERCDs, ha già trovato applicazioni nell'industria tessile, alimentare, farmaceutica, oltre che nelle bonifiche ambientali.

Tuttavia, l'utilizzo di questi enzimi è tutt'oggi limitato da diversi fattori, come l'instabilità in ambienti extracellulari e la dipendenza della loro attività dalla rigenerazione di cofattori e dallo stato di ossidazione di cofattori metallici al sito attivo. In molti casi, per questi enzimi la possibilità di utilizzare sistemi di bioconversione in cellule integre o con lisati batterici può essere molto vantaggioso, eliminando anche i costi ed i limiti dei processi di purificazione. In definitiva, lo sviluppo di sistemi efficienti di utilizzo delle diossigenasi batteriche per processi di bioconversione può portare a numerosi vantaggi nella produzione industriale di alcoli, dioli, aldeidi, acidi carbossilici e loro derivati. In generale si può ipotizzare l'utilizzo di queste attività enzimatiche per la modifica di diversi tipi di molecole eterocicliche con diversi gruppi sostituenti in posizioni specifiche, che possono essere successivamente utilizzati come sintoni o "building blocks" nell'industria chimica e farmaceutica.

Glicosil idrolasi – applicazione biotecnologica.

Le glicosil idrolasi, definite più comunemente glicosidasi, sono enzimi responsabili della scissione dei legami glicosidici dei carboidrati e sono coinvolte in una grande varietà di processi biologici. Infatti i carboidrati sono molecole di primaria importanza in tutti gli organismi viventi, e sono coinvolti in svariati processi biologici come ad esempio nelle vie di segnalazione cellulare, oltre ad avere funzioni strutturali e di riserva energetica. Nell'ambito delle biotecnologie, le glicosil idrolasi sono enzimi molto utilizzati, soprattutto nell'industria alimentare, per la produzione di cibi e bevande, nell'industria tessile e della carta e in processi di bioconversione di biomassa vegetale per la produzione di bioetanolo. Inoltre, è stato implementato negli ultimi anni anche l'utilizzo di questi enzimi per lo sviluppo di farmaci ed agenti terapeutici. Infine, questi enzimi risultano essere di grande interesse anche per lo studio e la caratterizzazione strutturale di molecole come carboidrati complessi. Tra le glicosidasi, le α-L-ramnosidasi sono attualmente utilizzate come biocatalizzatori nell'industria alimentare. In particolare, esse vengono utilizzate per esaltare gli aromi dei vini e per dolcificare i succhi di frutta, attraverso l'idrolisi di residui di ramnosio da glicosidi terpenici e flavonoidi naturali. In particolare, la de-ramnosilazione di flavonoidi naturali, oltre a migliorare le proprietà organolettiche dei cibi, aumenta l'assorbimento e biodisponibilità di queste molecole che hanno numerose proprietà benefiche antiossidanti e antiinfiammatorie.

Nonostante le applicazioni descritte, ad oggi sono poche le α-L-ramnosidasi batteriche caratterizzate, e solo un esiguo numero di strutture tridimensionali è depositato nellle banche dati. Naturalmente, l'utilizzo di una proteina in un particolare processo biotecnologico dipende dalla conoscenza che noi abbiamo delle sue caratteristiche e delle sue proprietà. In particolare, l'identificazione di nuove glicosil idrolasi sarebbe certamente di interesse nell'acquisizione di nuovi biocatalizzatori con diverse caratteristiche di attività e specificità.

In questo contesto, la selezione naturale ha conferito ad alcuni microrganismi l'acquisizione di un gran numero di geni codificanti per glicosidasi in grado di degradare e metabolizzare un ampio numero di polisaccaridi complessi. Tra di essi i microorganismi della flora intestinale sono quelli maggiormente specializzati per la degradazione di carboidrati complessi, e rappresentano quindi *reservoir* ideali per la ricerca di nuovi biocatalizzatori ad attività glicosidasica. Insieme ad essi, alcune classi di Sfingomonadali contengono molti geni codificanti per glicosil idrolasi probabilmente deputati anch'essi alla produzione e alla modifica di carboidrati complessi extracellulari.

Obiettivi del progetto di ricerca.

Sulla base delle considerazioni finora descritte, lo scopo di questo progetto di dottorato è stato l'identificazione e caratterizzazione enzimatica di quattro ERCDs e di una α-L-ramnosidasi dal ceppo *Novosphingobium* sp. PP1Y, e l'identificazione in ceppi di Bacteroides di nuovi enzimi ad attività glicosil idrolasica in grado di degradare l'arabinogalattano presente nella parete cellulare di *Mycobacterium tuberculosis*.

RISULTATI

Isolamento e caratterizzazione di ossigenasi ERCDs da *Novosphingobium* sp. PP1Y.

Il ceppo *Novosphingobium* sp. PP1Y (*N*. sp. PP1Y), precedentemente isolato e caratterizzato nel laboratorio in cui è stato svolto questo progetto di dottorato, è stato utilizzato come fonte per l'identificazione di nuovi enzimi ad attività diossigenasica. Il microorganismo *N*. sp. PP1Y ha evidenziato infatti alcune peculiari caratteristiche metaboliche come la capacità di crescere utilizzando un ampio numero di idrocarburi aromatici policiclici e miscele di gasolio e benzina, come unica fonte di carbonio ed energia. L'analisi del suo genoma annotato e sequenziato ha consentito di identificare un elevato numero di geni codificanti per enzimi ad attività ossigenasica appartenenti a diverse famiglie. Tra essi, analisi bioinformatiche di allineamento hanno consentito di identificare quattro enzimi ERCDs (PP28735, PP26077, PP00124 e PP00193) con elevata identità di sequenza con altri enzimi descritti in letteratura e di cui era disponibile la struttura tridimensionale. L'analisi di *homology modeling* e di *docking* molecolare delle corrispondenti sequenze amminoacidiche ha consentito di ottenere il modello della struttura tridimensionale del sito attivo e di effettuare analisi di *docking* nel sito attivo di diversi substrati di interesse per la produzione di estrogeni ed ormoni steroidei. L'analisi bioinformatica ha confermato che questi enzimi sono candidati promettenti per il riconoscimento ed il taglio, in specifiche posizioni, di un'ampia gamma di composti mono- e policiclici aromatici. Di conseguenza, questi enzimi sono stati espressi in maniera ricombinante in cellule di *E. coli* BL21(DE3), ed i protocolli di espressione ricombinante e di purificazione sono stati ottimizzati. Successivamente, l'attività dei quattro enzimi è stata valutata definendone i parametri cinetici su quattro substrati: catecolo (CAT), 3-metilcatecolo (3MC), 4-metilcatecolo (4-MC) e 2,3-DHBP (2,3-diirossibifenile). Il taglio di questi composti da parte delle ERCDs porta alla formazione di semialdeidi cis-idrossi-muconiche, la cui concentrazione può essere determinata spettrofotometricamente. I dati hanno consentito di confermare le indicazioni ottenute dall'analisi di *docking* molecolare, mostrando che PP28735 e PP26077 hanno l'affinità più bassa per i catecoli monoaromatici e solo una debole attività per il 2,3-DHBP, rispetto alle altre due ERCDs. Ciò è in linea con l'analisi di *homology modeling* che aveva mostrato per questi enzimi la presenza di siti attivi più ampi, probabilmente in grado di riconoscere con maggiore affinità substrati policiclici aromatici con più di due anelli. Al contrario, PP00193 ha mostrato una maggiore efficienza catalitica sui substrati testati, confermando di possedere una attività catalitica simile a quella di una catecolo diossigenasi. Infine, PP00124 ha mostrato attività comparabile su tutti i substrati testati, con efficienza migliore per il 2,3-DHBP. Complessivamente, questa analisi ha confermato che le quattro ERCDs sembrano avere differente specificità e che potrebbe essere utilizzata per un loro diverso utilizzo in processi di bioconversione di composti aromatici di interesse biotecnologico.

Infine, sono stati effettuati esperimenti per valutare la bioconversione di catecoli estrogeni quali il 4-idrossiestradiolo e il 2-idrossiestradiolo (4-OHE e 2-OHE), i quali possono essere utilizzati come precursori nella sintesi di diverse famiglie di ormoni steroidei. I risultati hanno evidenziato che solo PP28735 e PP00124 sono in grado di catalizzare efficientemente il taglio dei substrati, con valori di k_{cat}/K_M più elevati rispetto a quelli ottenuti per i catecoli monoaromatici. L' estratto cellulare di *E. coli* contente la proteina PP00124 è stato poi utilizzato in reazioni di bioconversioni preparative che hanno evidenziato la conversione completa del 4-OHE e quella parziale del 2-OHE, confermando così l'effettivo potenziale di utilizzo di questa diossigenasi per la bioconversione dei catecoli estrogeni.

Caratterizzazione e potenziale biotecnologico della α-L-ramnosidasi RHA-P da *Novosphingobium* sp. PP1Y.

Nella seconda parte del progetto, è stata effettuata la caratterizzazione biochimica di RHA-P, una α-L-ramnosidasi isolata dal microorganismo *N.* sp. PP1Y. Precedenti analisi avevano evidenziato che la ramnosidasi RHA-P è una glicosidasi della famiglia GH106 con potenziale biotecnologico per la bioconversione di flavonoidi naturali. In questo lavoro di tesi, è descritta inizialmente l'ottimizzazione del processo di espressione e purificazione ricombinante della proteina. Considerando la bassa resa di purificazione precedentemente ottenuta per la proteina, si è deciso di clonare e di esprimere in maniera ricombinante la proteina fusa con un His-tag al C-terminale. Inoltre, è stata ottimizzata la sua purificazione a partire dalle frazioni periplasmatiche delle colture indotte di *E. coli* BL21(DE3). Infatti, evidenze precedenti avevano mostrato la presenza in RHA-P di un peptide segnale di 23 amminoacidi, successivamente tagliato in seguito a modifiche post-traduzionali e probabilmente coinvolto nel "sorting" della proteina allo spazio periplasmico. Tale ipotesi è stata confermata durante il mio progetto di dottorato da esperimenti di espressione analitica che hanno mostrato la presenza di un'attività specifica della proteina molto più elevata nella frazione periplasmatica, rispetto a quella citoplasmatica. Seguendo il nuovo protocollo di purificazione, è stato possibile ottenere una resa maggiore di proteina, con un maggiore grado di purezza, in un unico step di purificazione attraverso una cromatografia di affinità.

Successivamente, al fine di indagare la topologia del sito attivo di RHA-P ed il suo meccanismo catalitico, è stata effettuata un'analisi di *homology modeling* della proteina, utilizzando come stampo la struttura tridimensionale della proteina BT0986 da *Bacteroides thetaiotaomicron*, l'unica proteina della famiglia GH106 attualmente presente nelle banche dati. L'analisi ha evidenziato che i putativi residui catalitici e quelli coinvolti nel legame del ramnosio al sito attivo, risultano essere conservati in RHA-P. Al contrario, variazioni in altre posizioni del sito attivo hanno suggerito una differente specificità di substrato delle due proteine, come dimostrato dalla maggiore efficienza di BT0986 per l'idrolisi di residui oligosaccaridici rispetto a molecole con sostituenti arilici, che sembrano essere riconosciute preferenzialmente da RHA-P. Inoltre, risultano essere conservati in RHA-P i residui di legame con uno ione calcio al sito attivo, indicando anche in RHA-P l'importanza di questo ione per il meccanismo catalitico. Ciò è supportato anche dall'inattivazione di RHA-P ottenuta in seguito al trattamento dell'enzima con un chelante degli ioni metallici, e la successiva riattivazione in presenza di cationi divalenti, con una efficienza maggiore in presenza di calcio. In definitiva, l'analisi del modello e gli esperimenti appena descritti, suggeriscono per RHA-P un meccanismo catalitico complessivamente simile a quello di BT0986, ma con una differente specificità di substrato.

Successivamente, sono state determinate le costanti cinetiche per la conversione di flavonoidi naturali quali, naringina, rutina, esperidina e quercitrina. I risultati hanno mostrato che RHA-P è in grado di idrolizzare efficientemente tutti i flavonoidi testati, essendo in grado di scindere sia legami glicosidici α-1,2 sia α-1,6. Inoltre, residui disaccaridici legati al flavonoide sembrano essere legati più efficientemente dai residui del sito attivo, ed è stata osservata una preferenza da parte dell'enzima per l'idrolisi di residui di ramnosio legati in posizione C3 dell'anello C dei flavonoidi. Complessivamente, i risultati hanno mostrato una buona efficienza di RHA-P verso questi substrati, confermando il potenziale biotecnologico dell'enzima per l'utilizzo in bioconversioni dei flavonoidi. A questo proposito è stato condotto anche un esperimento di bioconversione della naringina utilizzando cellule integre ricombinanti di *E. coli* BL21(DE3) che esprimono RHA-P. I risultati, ottenuti mediante analisi HPLC, hanno evidenziato che le cellule sono in grado di catalizzare efficientemente l'idrolisi completa della naringina, con rilascio del prodotto de-ramnosilato nel mezzo di coltura. Questo risultato supporta la potenzialità dell'utilizzo di RHA-P in processi di bioconversione dei flavonoidi naturali anche in cellule integre, confermando che le cellule integre rappresentano una ottima strategia nelle bioconversioni industriali che consente di risparmiare sui costi di purificazione del catalizzatore e consente inoltre una più facile purificazione dei prodotti di interesse dal mezzo di crescita.

Identificazione di glicosil idrolasi attive sull'arabinogalattano della parete cellulare di *Mycobacterium tuberculosis*.

L'ultima parte del progetto si è focalizzata sull'identificazione e caratterizzazione di nuovi enzimi ad attività glicosidasica in grado di degradare l'arabinogalattano (AG) presente nella parete cellulare di *Mycobacterium tuberculosis* e composto da residui di galattofuranosio e arabinofuranosio, al fine di valutarne l'utilizzo in terapie enzimatiche contro tale patogeno. *Bacteroides finegoldii*, uno dei batteri del microbiota intestinale, era stato precedentemente caratterizzato per la sua capacità di degradare il polisaccaride complesso dell'AG. Nove glicosidasi sono state identificate nel genoma di *B. finegoldii* come potenziali candidati per l'idrolisi di questo polisaccaride. I geni corrispondenti sono stati clonati con un His-tag al C-terminale, espressi in maniera ricombinante in cellule di *E. coli* TUNER DE3 e le proteine iperespresse sono state purificate mediante cromatografia di affinità. Le proteine purificate sono state utilizzate poi per lo screening di attività su un range di substrati sintetici (pNP-α e β con diversi zuccheri legati) e sui polisaccaridi AG e galattano della parete di *M. tuberculosis*. I risultati hanno evidenziato che cinque delle proteine identificate sono attive esclusivamente su residui di galattofuranosio e sono caratterizzate da una elevata specificità su questi substrati.

Successivamente è stata definita la specificità di substrato per le proteine BF03696 e BF04084, che, utilizzate in combinazione, sono in grado di degradare completamente la porzione polisaccaridica del galattano. Saggi effettuati sui substrati sintetici β(1,5)Gal*f*-ANS e β(1,6)Gal*f*-ANS hanno evidenziato che entrambi gli enzimi agiscono come eso-glicosidasi, e che BF03696 è in grado di scindere esclusivamente legami glicosidici β-1,5, mentre BF04084 è in grado di idrolizzare sia legami glicosidici β-1,5 che β-1,6. Entrambi gli enzimi risultano essere candidati promettenti per la degradazione del galattano presente nella parete di *M. tuberculosis*.

Inoltre, la proteina BF04084 sembra essere membro di una nuova famiglia di glicosidasi, non riconducibile a nessuna delle famiglie GH finora identificate. Al fine di caratterizzare in modo più approfondito tale proteina, è stata effettuata una caratterizzazione cinetica di BF04084 sia sul substrato sintetico *p*NP-β-D-Gal*f* sia sul disaccharide gal*f*-β(1,6)-gal*f* ottenuto dal galattano. Complessivamente, questo enzima si è dimostrato un candidato interessante per la degradazione dell'AG dei micobatteri, mostrando valori di k_{cat}/K_M comparabili a quelli di altri enzimi descritti in letteratura, ed mostrando un optimum di pH a 7.3, un valore differente rispetto agli altri enzimi ad attività galattofuranosidasica già caratterizzati.

Infine, la porzione dell'arabinano di AG è stata purificata ed utilizzata come unica fonte di energia in uno screening per la crescita di altri microorganismi della microflora intestinale. Tra questi, il ceppo *D. gadai* è risultato in grado di crescere utilizzando l'arabinano come fonte di energia, e rappresenta quindi una promettente fonte di enzimi ad attività arabinofuranosidasica. In questo ceppo, la degradazione dell'arabinano di *M. tuberculosis* è stata verificata tramite TLC.

ABSTRACT.

Main aim of this PhD project is the bioprospecting of different microbial enzymatic activities, to evaluate their biotechnological potential in different bioconversion processes. The characterization of four extradiol ring cleavage dioxygenases (ERCDs) and of a α-L-rhamnosidase isolated from *Novosphingobium* sp. PP1Y (*N*. sp. PP1Y) is described. In the last part of the project, gut microorganisms have been used for the identification and characterization of novel glycosyl hydrolases able to degrade the arabinogalactan polysaccharide (AG) of *M. tuberculosis* cell wall.

In this work, the marine microorganism *N*. sp. PP1Y and the gut microorganism *Bacteroides finegoldii* were used as source for oxygenases and glycosyl hydrolases.

More in detail, the optimization of recombinant expression and purification of four novel ERCDs from strain PP1Y allowed to obtain in good yields the corresponding proteins and to carry out their characterization. The activity screening using different catecholic substrates confirmed that these enzymes are able to catalyze the ring cleavage of a variety of mono- and polycyclic aromatic hydrocarbons with different sizes and conformation. In addition, the bioconversion of catechol estrogens, which could be used as precursors in the production of different steroid families and hormones, was evaluated. The results described in this thesis confirmed that these enzymes can be foreseen as a valuable tool for the modification of complex hydroxylated heterocyclic aromatic compounds, which are a starting point in the production pipeline of many pharmacologically active molecules, such as steroid-like molecules. Noteworthy, site-specific cleavage and subsequent modification of aromatic substrates, obtained by enzymatic biocatalysts, is of great advantage for industrial applications when compared to the complex mixture of products that are released instead during chemical modification procedures.

In the second part of the PhD project, the biochemical characterization of RHA-P, a bacterial α-L-rhamnosidase isolated from the microorganism *Novosphingobium* sp. PP1Y, was performed. The active site topology and substrate specificity of RHA-P were investigated by *homology modeling*. The enzyme, whose recombinant expression and purification was optimized, resulted to be appealing from a biotechnological point of view for the bioconversion and de-rhamnosylation of natural flavonoids. The biotechnological use of either wild-type or mutant rhamnosidases is currently a need for food and beverages industry to improve the organoleptic properties of processed vegetal products. Moreover, the possibility of using efficient whole cells biocatalysts for expressing RHA-P has been described. This is particularly interesting because whole cells biocatalysts have numerous advantages in industrial bioconversion processes, allowing costs and process steps reduction.

Finally, the bioprospecting of novel GHs from microorganisms belonging to the human gut microbiome (HGM), able to degrade the AG of *M. tuberculosis* cell wall, have been carried out. The isolation and characterization of different galactofuranosidases is described, which are able to completely degrade the mycobacteria galactan moiety of AG. Moreover, evidences for the presence of arabinofuranosidases were found for another HGM bacterium, which may be used as biocatalysts for the complete degradation of mycobacteria AG polysaccharide.

ABBREVIATIONS

2-OHE = 2-hydroxy extradiol
2,3-DHBP = 2,3-dihydroxy biphenile
3-MC = 3-methyl catechol
4-MC = 4-methyl catechol
α-RHAs = α-L-rhamnosidases
AG = arabinogalactan
Amp = ampicillin
ABS = absorbance
BSA = bovine serum albumin
CAT = catechol
DMF = dimethyl formammide
DMSO = dimethyl sulfoxide
DTT = dithiothreitol
ERCDs = extradiol ring cleavage dioxygenases
GHs = glycosyl hydrolases
GTs = glycosyl transferases
HGM = human gut microbiome
HPAEC-PAD = high performance anionic exclusion chromatography with pulsed amperometric detector
HPLC = high performance liquid chromatography
IPTG = isopropyl β-D-1-thiogalactopyranose
IRCDs = intradiol ring-cleavage dioxygenases
Kan = kanamycin
LB = Luria Bertani medium
MOPS = morpholinopropansulphonic acid
OD$_{600}$ = optical density
ORF = open reading frame
PAHs = polycyclic aromatic hydrocarbons
pNP-β-D-gal*f* = *para*-nitrophenyl galactofuranoside
*p*NPR = *para*-nitrophenyl rhamnopyranoside
PULs = polysaccharide utilization loci
qPCR = quantitative PCR
RCDs = ring-cleavage dioxygenases
RNA-seq = RNA sequencing
RT = room temperature
SDS = sodium dodecyl sulphate
TLC = thin layer chromatography
TYG = tryptone yeast extract glucose medium

CHAPTER 1

INTRODUCTION

CHAPTER 1

Introduction

1.1 Role of microbial enzymes in biotransformations and green chemistry.

Enzyme bioprospecting is a basic research activity devoted to the search for novel biocatalysts [1]. It can be defined as the systematic search for and the development of new enzymatic activities, microorganisms and other valuable bioactive compounds from the environment. The benefits of enzyme bioprospecting include the discovery and study of novel biocatalytsts able to convert an unexpected variety of products, thus widening the potentials of bioconversion approaches in many different fields [2].

In recent years, the existing and recognized applicability of environmental microbiology techniques, has been greatly strengthened by the development of the molecular enzyme technology and metagenomics [3]. Taking advantage of methodological advances in high-throughput DNA sequencing, more than 20,000 bacterial and fungal genomes have been sequenced, and the data have become available in public databases [4]. The metagenomic approach is a powerful tool that allows to discover new enzymes in microbial communities using DNA sequencing without the technically challenging need to culture them as individual species [5, 6]. The availability of complete genome sequences in databases allows to identify *in silico* a target gene with a promising catalytic activity. The gene, thanks to recombinant DNA technology, can be synthesized, cloned into a host microorganism, and produced on an industrial scale. Moreover, genetic engineering techniques can be used to further increase the catalytic properties of the new enzymes, thus allowing to fine-tune their activity towards the desired substrates.

In the search for new enzymes valuable for biocatalytic processes, microorganisms are favored sources, due to their availability, fast growth rate and ease of manipulation. Moreover, special characteristics of microbial enzymes such as thermophilic, acidophilic or alkalophilic, make them the preferred choice for their use in bioprocesses [7]. In fact, these microorganisms have optimal activity at high salinity or high temperatures and their enzymes can be used in many harsh industrial processes where the activity of many mesophilic enzymes would be inhibited. [8]. Moreover, in some cases, bacterial strains bear the metabolic pathways for the degradation of toxic chemical compounds (phenolic compounds, nitriles, amines etc.) via either degradation or conversion [9].

In conclusion, natural selection has endowed microorganisms, especially those living in peculiar extreme environments, with a catalytic machinery able to degrade an extremely wide range of natural substrates and to function in extreme physic-chemical conditions. This great biochemical diversity along with the ease of manipulation and the advances in DNA sequencing techniques, make bacteria a very powerful tool for the isolation of new enzymatic activities.

In recent years, the demand for suitable enzymatic biocatalysts in the form of whole cell catalysts or cell-free systems, with a high applicability potential in industry, is increasing. In the past decades the uprising concerns on environmental pollution, boosted the interest towards renewable energies and waste prevention in many fields. The design of sustainable productive processes became an issue to be addressed and considered. The term “Green chemistry” was created, which includes every process that efficiently uses raw materials in the manufacture of chemical products, generating a minimal amount of waste and avoiding the use of toxic and hazardous reagents. [10]. In this framework, enzymes are the biocatalysts that best meet the requirements of a sustainable production process. The use of enzymes as biocatalysts, in fact, can have significant benefits compared to conventional chemical technologies, allowing high reaction rates and selectivity, improved product purity, and a significant decrease in chemical waste generation by using eco-friendly mild reaction conditions.

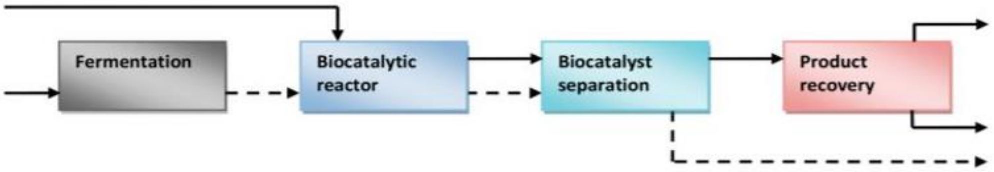
A wide variety of chemical substances is already produced in industrial processes using enzymes [11, 12]. Examples encompass the food industry (*e.g.*, bread, cheese, beer, and vinegar), fine chemicals (*e.g.*, amino acids, vitamins), agricultural (growth hormones), and pharmaceuticals (insulin). Enzymes are also used in environmental bioremediation processes or for analytical and diagnostic purposes [7]. More in detail, the majority of currently used industrial enzymes have hydrolytic activity, being used for the degradation of various natural substances. Proteases are used in detergent fabrications, amylases in food and beverages industry, xylanases, ligninases and cellulases are employed in the paper and pulp industry, textile and biofuel production [13 – 16].

In addition, the application of enzymes is uprising also as therapeutic agents. Examples of therapeutic enzymes are found in the cancer research. For example, recent studies have proved that arginine-degrading enzyme (PEGylated arginine deaminase) can inhibit human melanoma and hepatocellular carcinomas [17]. Moreover, application of enzymes in an antibody-directed enzyme prodrug therapy (ADEPT) have been described. In these cases, a monoclonal antibody carries an enzyme specific to cancer cells where the enzyme activates a prodrug in a site-specific mode. This approach is being used for the discovery and development of cancer therapeutics based on tumor-targeted enzymes that activate prodrugs. In conclusion, applications of bioconversion strategies using microbial enzymes are numerous and increasing rapidly over conventional methods.

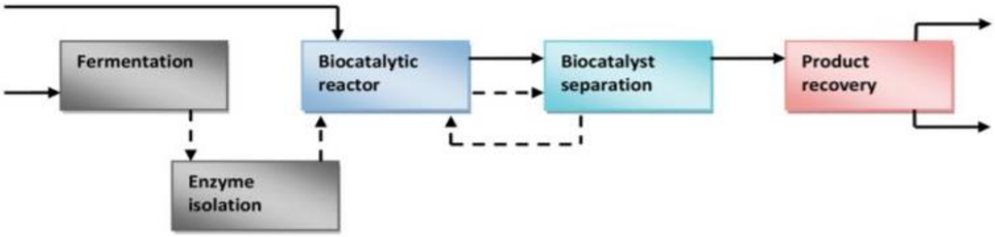
Biocatalytic processes can be carried out by both whole cells and free enzymes (Figure 1). The use of whole-cell biocatalysts even though economically convenient, might lead to cross-reactivity with cellular metabolites. Likewise, cofactors might not be available at the correct rate, and substrates and products can be metabolized through competing cellular reactions. Finally, problems like diffusional limitations of the substrate into the cells, and products recovery have to be taken into account [18]. Alternatively, isolated enzyme biocatalysis can be used. The use of free enzymes overcomes the diffusional problem, even though the added costs for enzyme purification could be excessive. In these cases, the use of non-purified proteins could also be evaluated, provided that the cellular extract does not contain proteases or competing activities [19]. However, conditions for the use of a single enzyme catalyst usually include the need to easily separate the enzyme from the product. In many cases, the best way to achieve a good product purification and catalyst recovery is to use immobilized enzymes. Here, the enzyme is covalently linked usually to a porous support, to facilitate its removal and subsequent recycle. In some cases, immobilized enzymes could also show an improvement of stability over time.

In recent years, the ultimate goal in green chemistry processes is to integrate several catalytic steps into a single procedure without the need to isolate single reaction intermediates [20]. Differently from what happens in traditional organic synthesis, these processes mimic the efficient orchestration of enzymatic steps in the living cells metabolic pathways, thus avoiding the purification of intermediates at each step, which minimizes costs and waste.

a) Whole-cells biocatalysis


Fermentation
Biocatalytic reactor
Biocatalyst separation
Product recovery

b) Isolated enzyme biocatalysis


Fermentation
Biocatalytic reactor
Biocatalyst separation
Product recovery
Enzyme isolation

c) Immobilized enzyme biocatalysis

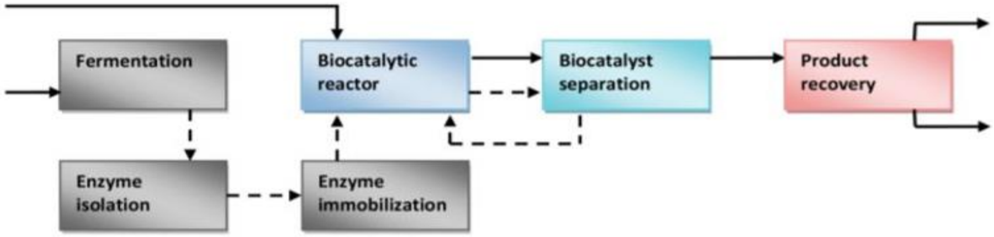

Fermentation
Biocatalytic reactor
Biocatalyst separation
Product recovery
Enzyme isolation
Enzyme immobilization

Figure 1: Flowsheets of the three classes of biocatalysis. Solid lines indicate substrate-/product-rich streams. Dotted lines indicate biocatalyst streams [4].

Furthermore, coupling of enzymatic reactions can be used to drive equilibria toward product, thus avoiding the need for excess reagents. Indeed, multienzyme biocatalysis has come to represent the most effective use of biocatalysis. An example is the four-enzyme cascade synthesis of non-natural carbohydrates starting from glycerol and aldehyde molecules [21]. An alternative to this *in vitro* approach is also to use metabolic engineering to rewire microbial cellular metabolism to enhance the production of native metabolites or to endow cells with the ability to produce new products using the existing metabolic pathways. [22].

All these tools are being used more and more in Green chemistry procedures in order to obtain a more sustainable technology with significant economic and environmental benefits. However, a primary challenge for the future is still to modify laboratory-scale chemistry to make it suitable for industrial implementation, ensuring that the enzymes have sufficient activity and stability under the required conditions. Instruments to achieve this goal start from the basic characterization of the enzymes and then include more and more the use of protein engineering and computational tools to design an effective bioconversion process with a high productivity.

This PhD project was focused on the identification from different microbial sources and characterization of novel glycosyl hydrolases and oxigenase enzymes, with applicability potential for diverse biotechnological applications.

1.2 Ring-cleavage extradiol dioxygenases and their biotechnological applications.

Due to the spreading of chemical pollution in the environment, several bacterial strains have developed the ability to use pollutants as carbon and energy sources [23]. These strains, studied for their ecological role and for their possible use in bioremediation processes, are also a source of very interesting metabolic activities and unusual biocatalysts, which can be used to synthesize pharmacologically active molecules and compounds of industrial interest. In fact, the need to activate and degrade xenobiotic compounds has been the driving force in these microorganisms to develop enzymes able to catalyze chemical reactions for the degradation of peculiar compounds. From this point of view, bacterial strains that are able to use aromatic compounds as the sole source of carbon and energy, are particularly interesting [24].

Marine habitat is one of the natural locations of interest for enzyme bioprospecting activity. The enormous pool of biodiversity found in marine ecosystems is considered in fact a profitable natural reservoir for acquiring an inventory of useful biocatalysts characterized by well-known habitat-related features such as salt tolerance, hyperthermostability, barophilicity and cold adaptivity. In line with this subject, a novel microorganism named *Novosphingobium* sp. PP1Y (*N*. sp. PP1Y), member of the Sphingomonadales order, has been isolated from the surface waters of a small dock bay in the harbor of Pozzuoli, characterized by severe pollution of the water by aromatic hydrocarbons, and has been microbiologically characterized in the laboratory in which my PhD project has been carried out [25]. *N*. sp. PP1Y shows peculiar characteristics when compared to other members of the Sphingomonadales order. It appears not only to be able to grow using, as the sole source of carbon and energy, a wide range of hydrocarbons and mono-and polycyclic aromatic substrates (pyrene, naphthalene and phenanthrene), but also to have evolved an effective adaptation to the growth on complex mixtures of aromatic molecules dissolved in non-polar phases such as diesel and gasoline. Overall, all these phenotypic traits suggest the capacity to catalyze the biodegradation of complex polycyclic aromatic hydrocarbons (PAHs) as part of a specific adaptation strategy of the bacterium to the harsh environment from which it was isolated.

The catabolic pathways that allow bacteria to degrade aromatic compounds involve the action of oxygenase enzymes. These latter catalyze the introduction of one (monooxygenases) or two (dioxygenases) oxygen atoms into an organic molecule. They typically use molecular oxygen (O_2) as the oxygen donor and most of them require reduction counterparts, usually NADH or NADPH. Most oxygenases are cofactor-dependent enzymes requiring metal-ions (e.g., copper or iron), metal-ion complexes (e.g., heme, Fe-S cluster) or coenzymes (e.g., flavin, pterin, pyrroloquinoline quinone) [26 - 29]. The physiological roles of oxygenases are extremely diverse in bacteria but can be divided in two broad categories: (1) biosynthesis of secondary metabolites, hormones and signalling molecules and (2) biodegradation of toxic compounds, such as complex aromatic hydrocarbons, used as primary carbon source.

The aromatic compounds degradation pathway can be generally divided in two steps, indicated as upper and lower pathways [30]. The reactions of the upper pathways involve the activation of the aromatic ring by addition of one or two hydroxyl groups on two adjacent carbon atoms followed by the oxidative cleavage of the ring [31 - 33]. The hydroxylation is catalyzed by monooxygenases that act one after the other, or dioxygenases performing two simultaneous hydroxylation reactions on the aromatic ring [34]. These enzymes are usually enzymatic complexes that use cofactors and carry out redox reactions using oxygen as a co-reagent and often NADH as an electron donor. In the case of polycyclic compounds, more frequently a "ring-hydroxylating dioxygenase" (RHD) coupled to a dihydrodiol dehydrogenase (DDH) operates [35]. Ring activation in the upper pathway is followed by ring cleavage, a reaction catalyzed by specialized dioxygenases. These latter are divided in intra- and extradiol dioxygenases, indicated as IRCD and ERCD respectively, which trigger the so-called lower pathway [33 - 36], leading to the conversion of dihydroxylated aromatic molecules into intermediates of the tricarboxylic acid cycle (TCA), which are readily metabolized by any type of organism. IRCD enzymes cleave the bond between the two hydroxyl groups generating a dicarboxylic compound (a derivative of cis-muconic acid), whereas ERCD enzymes cleave one of the bonds adjacent to the diol, thus generating an alpha-hydroxy acid containing an aldehydic or ketonic group (a derivative of cis-muconic semialdehyde).

Extradiol ring-cleavage dioxygenases, both acting on single and multi-ring aromatic compounds [37, 38], catalyze the incorporation of O_2 oxygen into catechol derivatives, resulting in ring-opening, and muconic semialdehyde compounds.

These dioxygenases are often enzymes with one or two subunits and work using a Fe (II) ion as a cofactor that coordinates an electron transfer from the catechol substrate to an O_2 molecule, in order to cleave the aromatic ring forming a reduced product. The Fe (II) ion is stabilized in the active site by interaction with histidine, glutamate or aspartate residues that help deprotonating the catecholic -OH group thus triggering the reaction; these amino acids play also an important role in the correct positioning of the substrates at the active site, and thus in the enzyme specificity and enantio-selectivity. (Figure 2) [39].

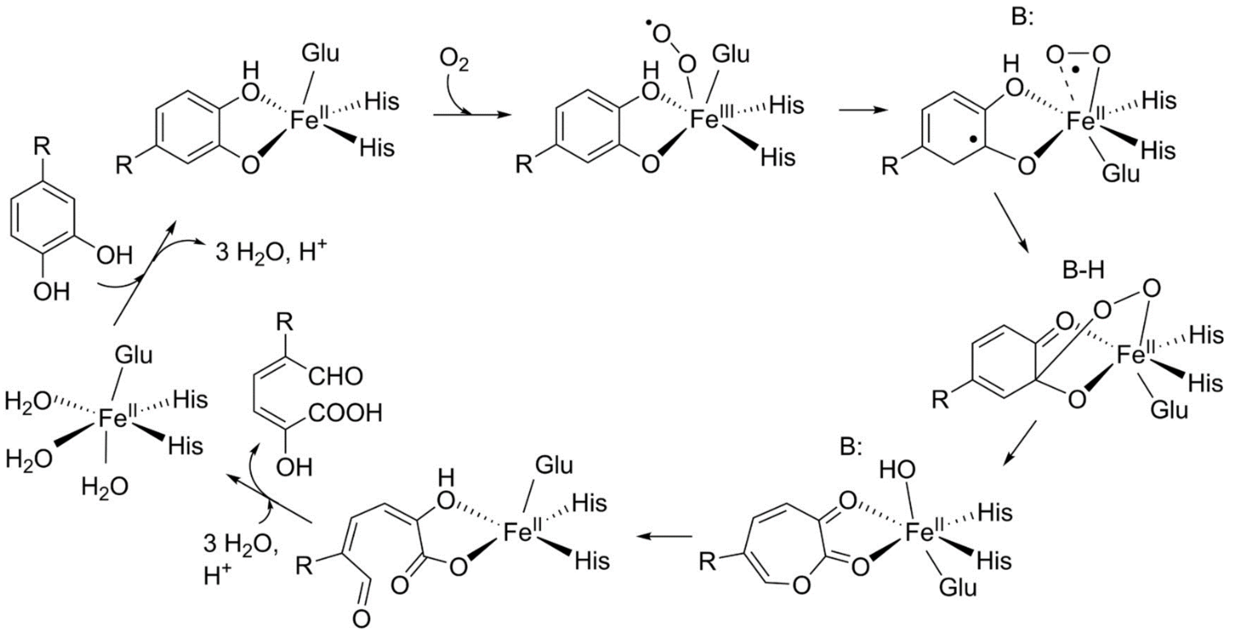


Figure 2: Catalytic mechanism of an extradiol ring-cleaving dioxygenase (ERCD). [40].

Oxygenases are attractive biocatalysts as they carry out the regio-, stereo- and chemio-selective introduction of oxygen in a wide range of organic molecules, in contrast to the poor selectivity of the traditional organic chemical synthesis, using strong reducing agents. [26 - 28]. Therefore, oxygenases hold a great potential for new applications in many fields, such as: textile, food, environment (biodegradation and bioremediation), biosensors, organic synthesis (chiral and asymmetric synthesis, pharmaceuticals), and biofuels [41 - 44].

However, the practical use of oxygenases compared with other industrial enzymes such as hydrolases, lyases and isomerases, is still limited by many factors [26, 27, 44, 45]. Many oxygenases have poor stability in non-native environments as most of them are multicomponent complexes. In addition, in many cases their function depends on the presence of cofactors that need to be regenerated or on the oxidation state of a metal cofactor that tends to be unstable in an outside cell environment [27, 44, 45]. In most cases, whole cell systems for oxygenase-catalyzed biotransformation were used to bypass the need for co-factor regeneration and the industrial processes that involve an oxygenase were carried out as batch or fed-batch fermentations [46]. Nevertheless, if suitable oxygenase biocatalysts are developed in more efficient processes, much shorter routes can replace expensive multistep chemical procedures to synthesize alcohols, diols, aldehydes, sulfoxides, carboxylic acids or their derivatives.

In recent years, protein engineering has been used to improve the catalytic rate, regioselectivity and thermostability of dioxygenases for biosynthesis and bioremediation purposes [44, 47, 48].

Some examples encompass the site directed mutagenesis and molecular modeling of o-xylene dioxygenase (AkbA) from *Rhodococcus sp*. strain DK17 that was used to improve its hydroxylation ability of biphenyl [49]. A site-specific mutagenesis based on homology modeling was used to extend the substrate range of aniline dioxygenase (AtdA) from *Acinetobacter sp*. strain YAA. The best variant showed the most significant enhancement in its bioremediation activity towards aromatic amines [50].

The regiospecificity of a biphenyl dioxygenase (BPDO) from *Burkholderia xenovorans* LB400, which catalyzes the first step in biphenyl biodegradation, was successfully modified to accept 2,2'-dichlorobiphenyl [51]. Moreover, BPDO oxidation of the 7-hydroxyflavone and 5,7-dihydroxyflavone flavonoids to their diol forms was enhanced by mutagenesis.

Catechols and chlorinated catechols are examples of carcinogenic pollutants [52]. Catechol dioxygenases and chlorocatechol dioxygenases play an important role in the detoxification of these hazardous molecules. A site specific mutagenesis based on homology modeling was successfully applied on catechol 1,2-dioxygenase (IsoB) from *A. radioresistens* S13 to fine-tune the catalytic properties and improve the activity on catechol derivatives [52].

Also, enzyme stability was improved in some cases, such as the thermostability increase of catechol-2,3-dioxygenase (C23O) from *Pseudomonas sp*. CGMCC2953 by disulfide bonds insertion [53].

In general, natural or engineered RCDs can be used as catalysts for the oxidative cleavage of polycyclic compounds with high regioselectivity, thus reducing or eliminating side reactions, which make difficult the use of traditional chemical methods. In this context, oxygenases are having an increasing role in the biosynthesis of many chiral compounds, by inserting oxygen at specific positions in complex molecules, and also in the synthesis of rare chemicals [54]. A promising example of oxygenase bioprocess described in literature is the production of a wide variety of terpenes in *E. coli* or other hosts, which in combination with specific oxygenases could allow the efficient production of many currently rare and difficult to isolate oxygenated terpenes. A regiospecific limonene hydroxylase was expressed in *E. coli* and *Saccharomyces* [52] and the metabolic engineering of *E. coli* to improve the production of intermediates such as lycopene was carried out [55].

In conclusion, as the cleavage products of the RCDs possess terminal reactive groups (carboxylic, aldehydic and hydroxyl groups), these enzymes could be used to prepare several types of heterocyclic functionalized molecules, valuable to be used as “synthons” for chemical synthesis design and in diverse biosynthetic pathways.

Finally, the combination of metabolic engineering with efficient and stable oxygenase expression systems will enable the fine-tuning of bioremediation processes from hazardous aromatic pollutants or the production of many oxygenated hydrocarbons, such as terpenes and polyketides, starting from cheap carbon sources. In general, any ortho-hydroxyl-containing aromatic compound, either mono- or polycyclic, prepared either enzymatically using RHDs or chemically, is structurally similar to the physiological substrates of RCDs and can be converted to a specific functionalized molecule with an high enantio-specificity.

The design of an efficient synthetic approach aimed at obtaining pharmacologically active compounds is often hampered by the complexity of these molecules and the presence of multiple different functional groups in precise positions. The formation of many aspecific side products and enantiomers arise in the chemical procedures. In this framework, the use of specific biocatalysts as the RCDs is of great interest to enhance the efficiency of such biosynthetic processes and broaden the range of the obtainable pharmacologically active molecules. Among the compounds of interest for pharmaceutical applications stand the steroids and steroid like molecules (estrogens, 17β-extradiol, estron, catecholestrogens) [56 - 59]. These polycyclic compounds are hormones of primary importance in many physiological processes such as in the neurological or inflammatory response modulation and tissue growth regulation, and they are usually employed in different hormonal therapies [60 - 62].

The catecholestrogens and extradiols are structurally similar to the physiological substrates of many RCDs enzymes (such as dihydroxynaphtalene and dihydroxyphenantrene). Thus, these enzymes in their native or mutagenized form are optimal candidates for the implement in steroid hormones biosynthetic pathways acting with the required regio- and enantio-specificity.

In this framework, the study of *N*. sp. PP1Y as a source of RCDs activities was one of the objectives of this PhD project. In fact, strain PP1Y genome has been sequenced and completely annotated and is comprised of a single 3.9-Mb circular chromosome and of three plasmids, one megaplasmid (Mpl; 1.16 Mb), one large plasmid (Lpl; 0.19 Mb), and one small plasmid (Spl; 0.05 Mb) (Figure 3) [63].

The analysis of *N*. sp. PP1Y genome revealed the presence of hypothetical *orf*s coding for more than 30 RHDs, two membrane bound aromatic monooxygenases, 7 ERCDs and several IRCDs, which constitute an impressive reservoir of oxigenase activities that enable PP1Y to degrade a wide range of mono-, bi-, tri-, tetra- and penta-cyclic aromatic compounds (Figure 3.3) [25].

Among these activities, four ERCDs have been successfully cloned and recombinantly expressed in active form in *E.coli* in the laboratory in which this PhD project has been carried out (Cafaro V. et al, unpublished data), and represent a diverse subset of ring-cleavage dioxygenases to investigate for biosynthetic purposes.

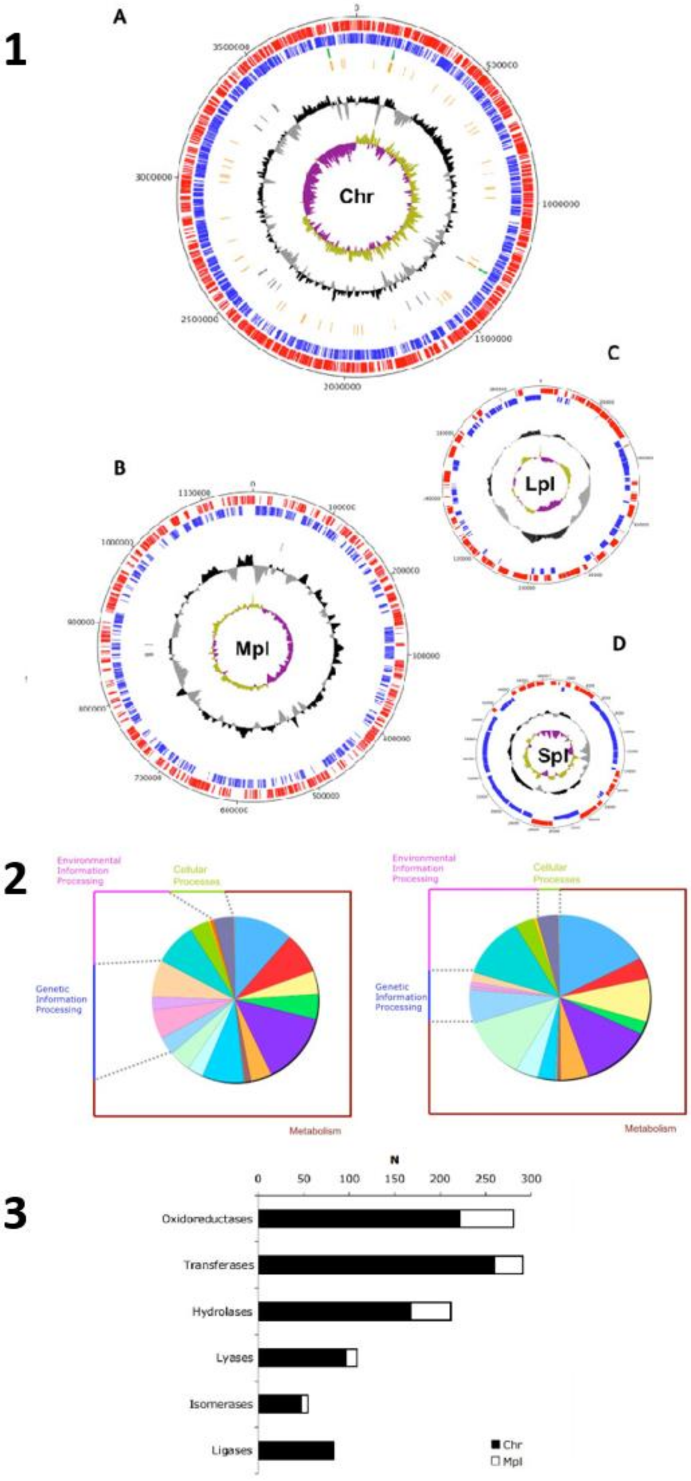


Figure 3: **1**: The main genomic features of PP1Y Chromosome (A) and its plasmids (B, C and D) are shown. For each of them, from outside to the center, there are genes on the forward and on reverse strands (red and blue), GC content (black) and GC skew (green and violet). **2**: Open reading frame annotation according to Kegg database and pathways. **3**: Genes and proteins distribution of Chr and Mpl genes with respect to enzyme classes. N = number of genes. [63].

1.3 Glycosyl hydrolases, classification and biotechnological applications.

Simple and complex carbohydrates are ubiquitous in nature where they play a multitude of functions, ranging from simple sources of energy to molecular recognition scaffolds critical to the interactions/communication among a wide array of biomolecules, cells, tissues and organisms [64]. The synthesis and degradation of carbohydrates in the form of di-, oligo-, and polysaccharides, is crucial in all living organisms.

Two main classes of biocatalysts are involved in the modification of carbohydrates in nature and are involved in a wide spectrum of biological processes: Glycosyl Hydrolases (GHs) and Glycosyl Transferases (GTs), which are responsible for the hydrolysis and the synthesis of the sugars, respectively.

GTs catalyze the transfer of sugar moieties from activated donor molecules to a nucleotide molecule acceptor, forming glycosidic bonds. GHs (also named Glycosidases) are a widespread group of enzymes that hydrolyse the glycosidic linkage between two or more carbohydrates or between a carbohydrate and a non-carbohydrate moiety, leading to the formation of a sugar hemiacetal or hemiketal and the corresponding free aglycon.

The GHs activities can be classified in different ways, based on the substrate sugar specificity, on the stereospecificity of the anomeric carbon in the substrate, distinguishing them in α- and β- glycosidases, or on their activity over long glycosidic chains on which they act as exo- or endo-glycosidases. Finally, the classification in *inverting* and *retaining* is based on the two possible catalytic mechanisms introduced by Koshland in 1953 [65].

The *inverting* or *retaining* catalytic mechanism of GHs is referred to the anomeric configuration of the glycosidic substrate as compared to the product. In fact, the enzymatic cleavage of the glycosidic bond can lead to the formation of a sugar hemiacetal with the opposite (*inverting*) or the same (*retaining*) anomeric configuration of the substrate. More in detail, *inverting* GHs operate with a one-step, single-displacement mechanism, with the assistance of a general acid and a general base group at the active site, where they are typically located 6-11 Å apart. Instead, *retaining* enzymes follow a two-step mechanism. The carboxyl group in the enzymatic active site functions as a general acid/base catalyst, and the carboxylate functions as the nucleophile of the reaction. In this case the catalytic residues are typically located 5.5 Å apart (Figure 4) [66].

Today, an important tool for the GH study and classification is the CAZy database [67]. Online since 1998 and constantly updated, this database is dedicated to the classification and analysis of genomic, structural and biochemical information on: Glycosyl Hydrolases (GHs), Glycosyl Transferases (GTs), Polysaccharide Lyases (PLs) and Carbohydrate Esterases (CEs) generally referred to as Carbohydrate-Active Enzymes (CAZymes). These enzymes are classified in a total of 307 families based on their aminoacidic sequence and some conserved 3D-structure modules. The CAZy classification database is particularly useful because within the families the enzyme fold, catalytic apparatus and mechanism are largely conserved, thus providing useful insights into substrate specificity and structural data of still uncharacterized enzymes.

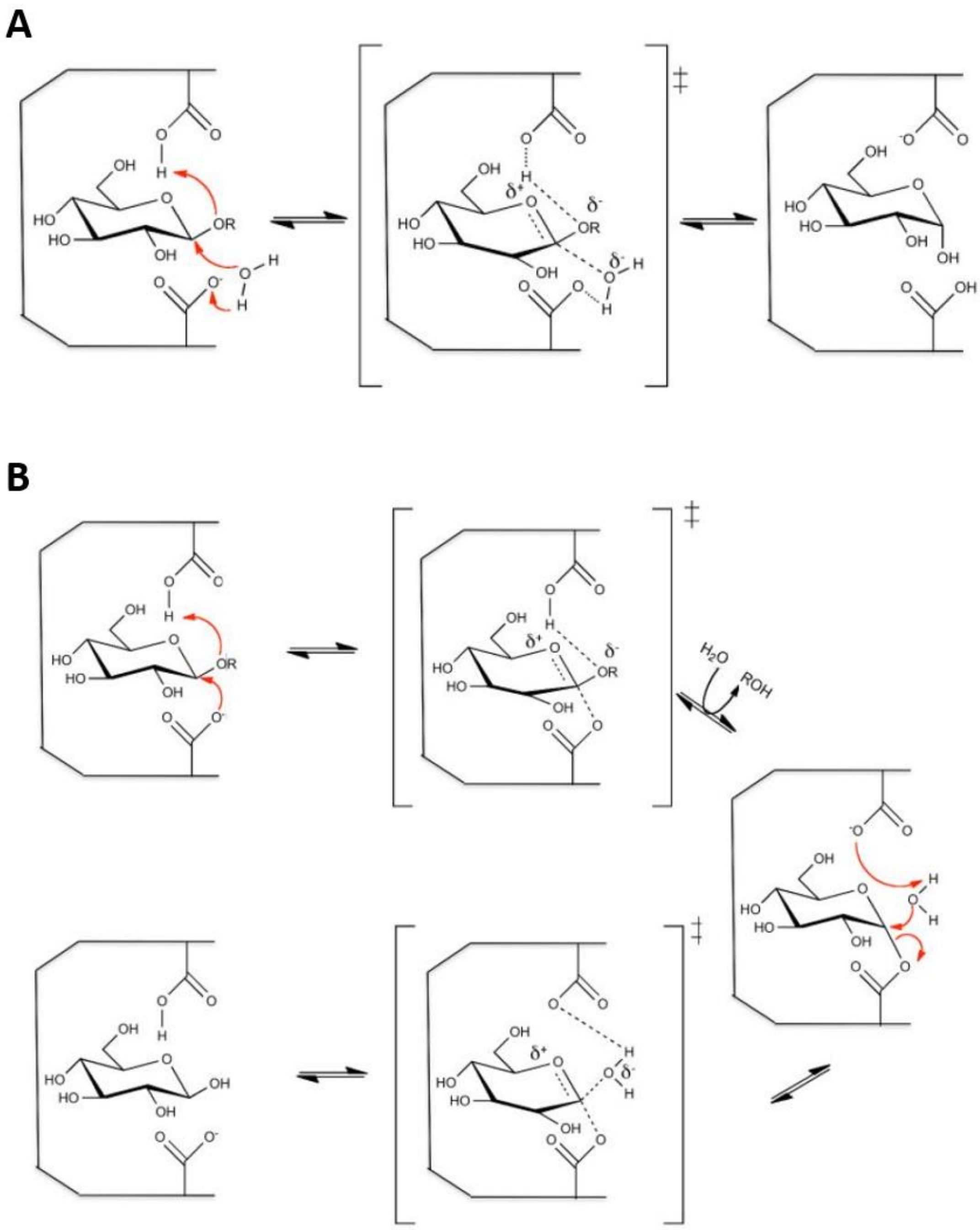

A
B
‡
OH
HO
OR
O
H
δ+
δ-
H2O
ROH

Figure 4: **A**: Catalytic mechanism of an *inverting* β-glycoside hydrolase **B**: Catalytic mechanism of a *retaining* β-glycoside hydrolase [66].

Nevertheless, the broad variety of catalytic activities and substrate specificity among GHs is constantly increasing with the identification and characterization of new enzymes from novel families. This boosted the interest in the identification and characterization of new glycosidases with novel activities for a wide range of substrates. In addition, the study of the structure/function relationship and reaction mechanism of GHs allow to further understand their role in vivo and to develop new methods for oligosaccharide characterization, synthesis and modification.

Also in the case of GHs biocatalysts, natural selection allowed some bacteria living in peculiar environments to evolve a great abundance of enzymes able to degrade a wide range of complex polysaccharides. Among them, the most representative polysaccharide specialized degraders among the microbial communities are the ones of the human gut microbiome (HGM). These bacteria are primary targets for the enzyme bioprospecting of novel GHs activities and for all their features and applications they deserve a special mention.

It is well known that mammalian organisms are not able to degrade the complex structural polysaccharides present in plant and animal material, and they rely on symbiotic gut microorganisms that possess the catabolic machinery to fulfil this task. Three main bacterial phyla dominate the human gut microbiome (HGM), the Bacteroidetes, the Firmicutes and Actinobacteria. The Bacteroidetes, largely represented from the genus Bacteroides, account for the major capability to degrade an extensive array of plant, animal and host complex carbohydrates, while the Firmicutes appear more specialized on plant storage polysaccharides. In general, a tight specialization towards different substrates is obtained among different microorganisms, in order to allow the co-viability of different bacterial species and a cross-feeding in the microbial community [68]. Nonetheless, many dominant gut bacteria show remarkable nutritional flexibility. As an example, *Bacteroides thetaiotaomicron* encodes a huge repertoire of carbohydrate degrading activities and has the ability to switch between diet- and host-derived carbohydrates [69]. The fine-tuning of this unique glycan degradation machinery in Bacteroides is due to the presence of co-transcribed gene clusters, called “polysaccharide utilization loci”, or PULs. Each PUL drives the degradation of a specific glycan and when the substrate glycan is present, the recognition of a specific oligosaccharide or monosaccharide leads in the bacterium to the upregulation of the corresponding PUL, often up to 1,000 fold.

PULs typically encode an array of surface enzymes, glycan binding proteins and Sus-like system transporter proteins devoted to the internalization of smaller oligosaccharides in the periplasm, where additional glycosyl hydrolases (GHs) and polysaccharide lyases (PLs) convert them into simple units [68 - 70]. The main operators of these degradation systems are indeed the GHs enzymes.

Recently, in addition to plant, animal and host glycans, complex carbohydrates produced by microbes have been shown to be a source of nutrients for *Bacteroides thetaiotaomicron*. In fact, the upregulation of specific PULs allowed this microorganism to degrade the α-mannan, an outer layer of the fungal and yeast cell wall in species such as *Saccharomyces cerevisiae* and *Candida albicans* [71]. Moreover, the degradation of the extracellular polysaccharide of different Lactobacilli, and of the 1,6-β-glucan component of many edible fungi cell walls, was observed [72, 73].

These and other recent works on Bacteroides strains not only provided the basic knowledge to unravel the metabolic degradation pathways for the degradation of complex polysaccharides, but also paved the way to the identification of new biotechnologically interesting enzymes.

In this framework, the identification of GH activities able to degrade bacterial arabinogalactan (AG) polysaccharide would be of interest, since this complex polysaccharide is a major component of the cell wall in some pathogenic bacteria, such as *Mycobacterium tuberculosis*. This polysaccharide has a peculiar structure bearing arabinose and galactose in their furanosic ring conformation, which is rarely found in nature. In general, the peculiar structure of mycobacteria cell wall is the reason for the development of specific antituberculosis drugs aimed at blocking cell wall assembly at different stages [74]. Among currently used therapeutics, molecules such as EMB (ethambutol) act by inhibiting the biosynthesis of AG [75]. Nevertheless, in recent years a multidrug resistance towards these compounds was observed, which boosted the research for novel and more effective therapeutic agents [76]. Among novel targets recently identified, a particular attention has been focused on the biosynthetic pathway of the AG domain [77]. Therefore, characterizing enzymes able to specific hydrolyze the linkages in the AG backbone would be of interest from a biotechnological point of view, and also for evaluating their use as therapeutic agents against the pathogen.

In this context, HGM bacteria and Bacteroides strains are optimal candidates to be exploited as source of novel arabinofuranosidase and galactofuranosidase activities.

Another group of microorganisms often specialized for the production and metabolism of complex extracellular polysaccharides is the Sphingomonadales order. In these bacteria the production of exopolysaccharides, known in some cases as sphingans, seems to be part of a specific adaptation strategy aimed at the production of a biofilm that functions as a barrier to protect the bacteria from environmental pollutants and to prevent the nutrient diffusion. [78].

In this framework, the study of *N*. sp. PP1Y genome revealed that, again, this strain is a source of new interesting GHs activities for bioprospecting. More in depth, a great number of genes encoding for both GHs (53 orfs) and GTs (57 orfs) were identified in this microorganism (Figure 3.3) [63]. In particular, the number of GHs is remarkable when compared, respectively, to the 16 GHs and the 32 GHs of the closely related strains of *Sphingomonas wittichii* and *Sphingobium chlorophenolicum* [79, 80]. The 53 GHs of strain PP1Y are distributed among 27 different families. The most represented families are GH3, GH13 and GH23 with 9, 5 and 9 members respectively. Among these, there are at least 8 genes encoding also for α-L-rhamnosidases.

These features make Bacteroides and *Novosphingobium* sp. PP1Y strains optimal candidates for the isolation and characterization of novel GHs activities, appealed for biotechnological applications. GHs are indeed the most important classes of biocatalysts currently used in biotechnological applications. Their applicability as biocatalysts is increased by their general stability in outside cell environment, and the absence of renewable cofactors needed for catalysis. Classical examples include baking processes (amylases), detergent formulations for the removal of glycan spots (cellulases and amylases), fabrics treatment and fading of the denim color (amylases and xylanases) [81], pulp and paper bleaching (xylanases) [82]; production of High Fructose Corn Syrup for soft drinks, where amylases and glucoamylases are employed to convert dextrins to glucose in a process called saccharification [83]. Pectinases are used in fruit juice processing [84], whereas cellulases, glucosidases, and xylanases are also implemented in the conversion of lignocellulosic plant biomass into bioethanol [85]. Bioethanol, unlike petroleum, is a form of renewable energy that can be produced from agricultural feedstocks. It can be obtained from hydrolysis, by amylases, of carbohydrates source as potato, manioc, several cereals and food byproducts, to obtain simple fermentable sugars (first generation bioethanol). Differently, cellulosic-ethanol (second generation bioethanol) is a type of bioethanol produced from lignocellulose (composed mainly of cellulose, hemicellulose and lignin), a structural material that comprises much of the mass of plants, by treatment with cellulase, xylanase and hemicellulase enzymes.

More recently, together with these applications, the specific action of glycosidases working either as hydrolases or in the glycosynthetic mode in the fine chemicals industry has been described [86]. The use of glycosyl hydrolases and glycosyl transferases has been proposed for the stereospecific glycosylation of a wide range of substrates. In these reactions, high substrate concentrations are used for the reverse reaction, thus promoting the formation of glycosidic linkages [87]. In this case, the glycosyl donor can be a monosaccharide, an oligosaccharide or an activated glycoside. This specific GHs implement is of great interest for the production of functional foods and drugs, whose characteristics might be specifically improved by either the addition or removal of specific glycan moieties [88]. The contribution of the attached sugars for many natural products has been studied thoroughly and has been shown to involve important drug properties, including pharmacokinetics, pharmacodynamics, solubility, mechanism and potency.

The GHs enzymes found broad applications also in biomedical technologies, both for direct application as pharmaceuticals and for the production of carbohydrate-containing drugs and pharmaceutically active molecules. Some examples include the production of universal donor blood by α-N-acetylgalactosidase and α-galactosidase hydrolyzing the A and B antigens respectively, to form the common H structure found in the O group [89]. In another case, a α-L-arabinofuranosidase from *Thermobacillus xylanilyticus* was used to synthesize galactofuranose-based glycoconjugates, which are interesting target molecules for drug design since galactofuranose is totally absent in mammals [90]. Synthetic glycopolymers and glycoproteins [91, 92] have been also used as carriers of covalently conjugated drugs, bearing carbohydrate ligands that provide specificity of delivery. In some cases, the carbohydrate ligand specificity is combined with the action of a GH biocatalyst through the construction of novel enzymes and prodrugs conjugated systems. Bipartite systems using biocatalysts for site-selective drug delivery were developed with significant success. The antibody-directed enzyme prodrug therapy (ADEPT) [93] uses mAb-GH enzyme conjugates to release a prodrug at a site determined by the mAb-antigen interaction. In this way, only at the target site the GH hydrolyze the specific glycosidic linkage in the prodrug, thus activating the drug action. A similar developed strategy is the Lectin-directed enzyme-activated prodrug therapy (LEAPT), designed to exploit endogenous carbohydrate-lectin binding by combining it with biocatalysis in a bipartite system [94].

Moreover, glycosidases have recently attracted the attention of many pharmaceutical industries since they are involved in many biological processes such as cell-cell or cell-virus recognition, immune responses, cell growth, and viral and parasitic infections [95].

The use of GHs as therapeutic agents is described also in the case of some genetic disorders in which the presence of a mutated inactive enzyme in the patient leads to a severe disease. In some cases, these disorders are treated with enzyme replacement therapies (ERT), based on intravenous administration of recombinant functional enzyme to restore its function in the patient. An example is the Pompe disease, in which a deficiency in a functional lysosomal α-glucosidase, leads to intra-lysosomal glycogen storage, causing extensive and progressive damages to cardiac and skeletal muscles. [96]. In these patients, administration of recombinant α-glucosidase has shown to stabilize the disease course.

Moreover, GHs substrate specificity could also be used as a tool to investigate the structure of glycosidic linkages in many complex polysaccharides in nature, whose fine chemical structure is still to be described in detail.

In conclusion, GHs are particularly appealing as potential biotechnological tools for their stability, efficiency and wide specificity.

1.3.1. The α-L-rhamnosidases.

α-L-rhamnosidases (α-RHAs) are a subset of GHs that have gained much attention in recent years. These enzymes catalyze the hydrolysis of terminal residues of α-L-rhamnose from a large number of natural compounds [97].

L- Rhamnose is widely distributed in plants as component of flavonoid glycosides, terpenyl glycosides, pigments, signaling molecules, and in cell walls as a component of complex heteropolysaccharides, such as rhamnogalacturonan and arabinogalactan-proteins [98]. In bacteria, L-rhamnose appears to be included in membrane rhamnolipids [99] and complex polysaccharides and glycoproteins, such as gellan gum, rhamnogalacturonan and arabinogalactan-proteins [100]. Bacterial α-RHAs are involved in the construction and metabolism of these molecules.

Among α-RHAs substrates, natural flavonoids are gaining much interest among the food and nutraceutic industry. Natural flavonoids are polyphenolic compounds generally characterized by a three-ring structure, which consists of two phenyl rings (A and B) and a heterocyclic ring (C). These molecules are naturally produced in plants in glycosylated forms, showing the presence of either a rutinoside (6-α-L-rhamnosyl-β-D-glucoside) or a neohesperidoside (2-α-L-rhamnosyl-β-D-glucoside) disaccharidic unit bound in different positions (Figure 5). Flavonoids are produced in plants as secondary metabolites, some of them, such as naringin, hesperidin, rutin and quercitrin are found in grapefruit juices, lemons, sweet oranges and many other vegetables.

These molecules have gained increasing recognition for their health-related qualities, such as their potential antioxidant, antitumor and anti-inflammatory properties [101].

In humans, endogenous β-glycosidases and α-L-arabinosidases in the small intestine are responsible for removing the glucose (or possibly arabinose or xylose) moiety from flavonoids to allow for their effective absorption. These enzymes, however, are not able to cleave terminal rhamnose units, thereby limiting the bioavailability of rhamnosylated flavonoids that are converted to more bioactive forms by the colon microflora [102]. Therefore, enzymatic rhamnose removal from potentially bioactive flavonoids may be the key for improving their intestinal absorption and thus their beneficial properties for human health [103].

Due to their ability to hydrolyze rhamnose from natural flavonoids, α-RHAs are used in several biotechnological applications. Some examples include the hydrolysis of naringin to improve beverage quality by debittering grapefruit and citrus juices [104], and the removal of hesperidin crystals from orange-derived preparations [105]. Other applications of α-RHAs are gaining popularity in the oenological industry, where these enzymes are used to hydrolyze terpenyl glycosides to enhance aroma in wine, grape juices and derived beverages [106]. Moreover, an α-RHA has been implemented for the synthesis of rhamnose-containing chemicals by reverse hydrolysis, suggesting a yet unexplored potential of this enzyme class in the chemical and pharmaceutical industry [107].

Despite their potential as industrial biocatalysts, to date only a limited number of bacterial rhamnosidases has been fully characterized. The commercial preparations of α–L-rhamnosidases, naringinases and hesperidinases available and currently used in oenology, are all isolated from fungal sources such as *Aspergillus niger* and *Penicillium decumbens* [108, 109]. One of the main differences between fungal and bacterial α-RHAs is their different optimal pH values, with the fungal enzymes showing more acidic pH optima when compared to the bacterial counterparts. This characteristic encourages the use of the bacterial "version" of these enzymes for specific biotechnological processes involving flavonoids as substrates because the solubility of these molecules strongly increases at higher pH values [110].

Few attempts have been made so far to engineer bacterial α-RHAs, to unravel the molecular details underlying their catalytic mechanism, to modify their substrate specificity or to optimize their catalytic efficiency towards different substrates. A major obstacle is the limited number of α-RHAs crystal structures that are currently available among the different families of GH enzymes, which include the GH28, GH78, and GH106 families according to the CAZy database. To the best of our knowledge, only five crystal structures of bacterial α-L-rhamnosidases are currently available, four of which belong to the GH78 family, which show an $(\alpha/\alpha)_6$ 3D-structure, and include the putative α-L-rhamnosidase BT1001 from *Bacteroides thetaiotaomicron* VPI-5482 [111], the α-L-rhamnosidase B (BsRhaB) from *Bacillus* sp. GL1 [112], the α-L-rhamnosidase (SaRha78A) from *Streptomyces avermitilis* [113], and the α-L-rhamnosidase (KoRha) from *Klebsiella oxytoca* [114].

Among the members of the GH106 family, which groups 319 different sequences, a single 3D-structure has been reported, the BT0986 from *Bacteroides thetaiotaomicron* that shows an $(\beta/\alpha)_8$ barrel and catalyzes the hydrolysis of an α-L-rhamnopyranoside bound to the C2 position of an arabinofuranoside (L-Rha*p*-α-1,2-L-Ara*f*). In this same family of glycosidases only two enzymes have been characterized so far and the reaction mechanism and the catalytic residues have been inferred from the 3D-structure of BT0986 [115].

Therefore, it is evident that bacterial α-RHAs represent a yet unexplored reservoir of potential biocatalysts for which more functional and structural data are required.

In this regard, preliminary data have shown that *N*. sp. PP1Y crude extract displayed the ability to hydrolyze some natural rhamnosylated flavonoids, such as naringin, diosmin, rutin, hesperidin, neohesperidin dihydrochalcone and quercitrin. A single α-RHA activity, from now on defined as RHA-P, was isolated from the crude extract, recombinantly expressed in *E. coli* BL21(DE3) cells, purified and a partial biochemical characterization of the enzyme was obtained in the laboratory in which this project was carried on [116]. RHA-P was characterized as an *inverting* monomeric glycosidase of ca. 120 kDa belonging to the GH106 family valuable to be further studied for biotechnological implements.

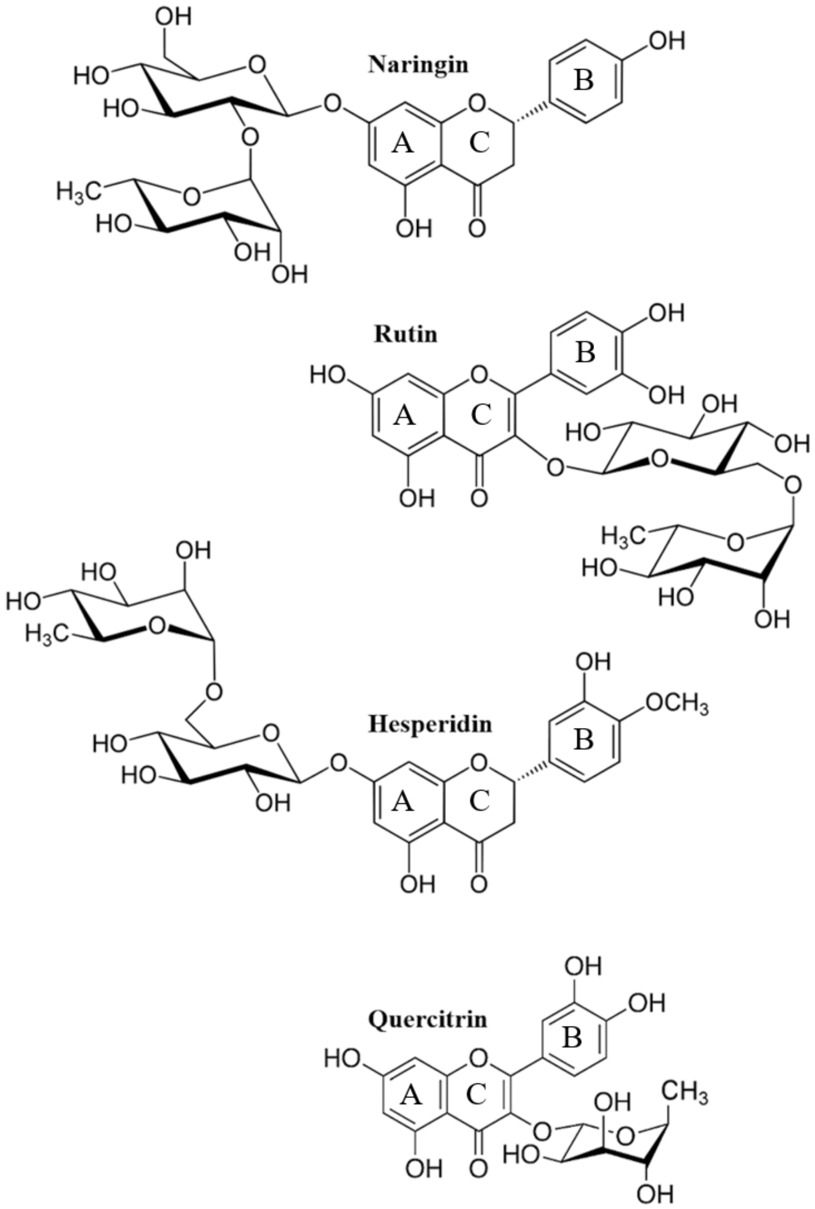

Naringin
Rutin
Hesperidin
Quercitrin
A
B
C

Figure 5: Chemical structure of rhamnosylated natural flavonoids.

1.3.2. The β-D-galactosidases.

β-D-galactosidases are enzymes that hydrolyze D-galactose residues from polymers, oligosaccharides or secondary metabolites. The most common natural substrate of these enzymes is lactose.

β-D-galactosidases are widespread in nature and found in microbes, plants and animal tissues. In plants, these enzymes are involved in several biological processes including plant growth, fruit ripening and cell wall metabolism.

Fungal β-D-galactosidases generally have acidic pH-optima in the range of 2.5–5.4, thus they are most effective for the hydrolysis of lactose present in acidic products such as whey. Moreover, fungal β-D-galactosidases are often thermostable enzymes and the major industrial enzymes currently used are obtained from *Aspergillus* sp. and *Kluyveromyces* sp. [117]. The yeast *Kluyveromyces lactis* is also one of the available commercially sources of β-D-galactosidase [118]. In general, lactose fermenting yeasts produce intracellular β-D-galactosidase in great amounts [119]. This enzyme, due to its stable hydrolytic activity, is currently used to produce lactose-free milk products. Yeast lactases are most active in buffers at pH 6.0–7.0 [120]. Finally, also β-D-galactosidases from bacterial sources have been described and are widely used for lactose hydrolysis due to the ease of fermentation of bacteria, high activity and good stability of the enzymes [121]. β-D-galactosidases are obtained with high yields from various bacterial strains like *Lactobacillus* and *Bifidobacterium* species, and implemented in the food industry. These strains are preferred sources for β-D-galactosidases, since they are effective probiotics agents themselves and are GRAS organisms (“generally regarded as safe”).

Generally, β-D-galactosidases work in a relatively broad pH range: enzymes from fungi act between pH 2.5–5.4, yeast and bacterial enzymes act between pH 6.0–7.0. Most of these enzymes are classified in the CAZy database in glycoside hydrolase family 2 (GH2) and family 35 (GH35). The enzymes of GH2 family are predominantly found in microorganisms whereas nearly 70% of GH35 are present in plants.

Industrially, β-D-galactosidases have various applications. They are widely used in food industry to improve sweetness, solubility, flavor and bioavailability of dairy products [122]. These enzymes are used in the hydrolysis of lactose from milk and milk related products to produce lactose free feedstocks. In fact, lactose absorption requires the activity of lactase enzymes and a significant percentage of the world’s adult population develops deficiency of lactase that leads to a disorder known as lactose intolerance [123]. In addition, lactose is hygroscopic and causes crystallization in food products. Hence, use of β-D-galactosidases allows to obtain food with improved texture and organolectic properties.

Moreover, lactose hydrolysis catlyzed by β-D-galactosidases converts whey lactose into a number of suitable products such as sweet syrups, which are used in bakery and confectionary industry.

β-D-galactosidases in yeasts are also used in the production of ethanol, via lactose fermentation using byproducts of cheese production, such as whey, as substrate [124].

Another application of β-D-galactosidases encompass removal of galactose residues from oligosaccharides and polysaccharides derived from plants, such as pectin and xyloglucans, necessary for their degradation [125].

The use of β-D-galactosidases also in a glycosynthetic mode have been described in literature to produce galacto-oligosaccharides (GOS). GOS trisaccharides and tetrasaccharides are produced from β-galactosidase by the transglycosylation activity during the hydrolysis of lactose [126]. GOS are used as ingredients in prebiotics food, aimed at modifying intestinal microflora and promote the growth of useful bacteria in the intestine such as *Lactobacillus* and *Bifidobacterium* species. Therefore, companies now include GOS in their milk- and cereal-based food products [127]. Currently, GOS are used as additives also in cosmetics, low-calories sweeteners, soft drinks, cereals, baby food, and powdered milk [128].

Finally, drugs containing β-D-galactosidases are also available, which are taken before consuming milk products [129]. These drugs contain fungal-derived β-D-galactosidases, usually from *Aspergillus*, that are stable at low pH thus allowing a proper functioning in the stomach.

As briefly outlined above, β-D-galactosidases are widespread in nature and nowadays a great number of these enzymes have been described, characterized from bacterial and fungal sources and exploited in industry. Conversely, only a few examples of β-D-galactofuranosidases and β-D-galactofuranosyl transferases are known. These GH and GT enzymes specifically recognize galactose in its furanosic ring conformation. In fact, this sugar is rarely found in natural polysaccharides; nonetheless, it is an important component of some bacterial polysaccharides, such as the arabinogalactan in *Mycobacteria* cell wall. Only a few examples of β-D-galactofuranosidases have been described from *Mycobacterium tuberculosis, Bacillus sp.,* and from fungi, *Streptomyces sp.,* and *Aspergillus niger* and little is known about their activity and substrate specificity [130 - 132].

Aims of the thesis

Main aims of this PhD project have been the characterization of the biotechnological potential of four ERCDs and of a bacterial α-L-rhamnosidase isolated from *Novosphingobium* sp. PP1Y, and the identification and characterization in Bacteroides of novel enzymes able to degrade the arabinogalactan complex of the *M. tuberculosis* cell wall.

In brief, the following activities have been performed:

1- Recombinant expression and purification of the four ERCDs from *N*. sp. PP1Y and characterization of their activity on selected hydroxylated polycyclic systems and catechol-estrogens in order to define their substrate specificity and value for biotechnological applications.

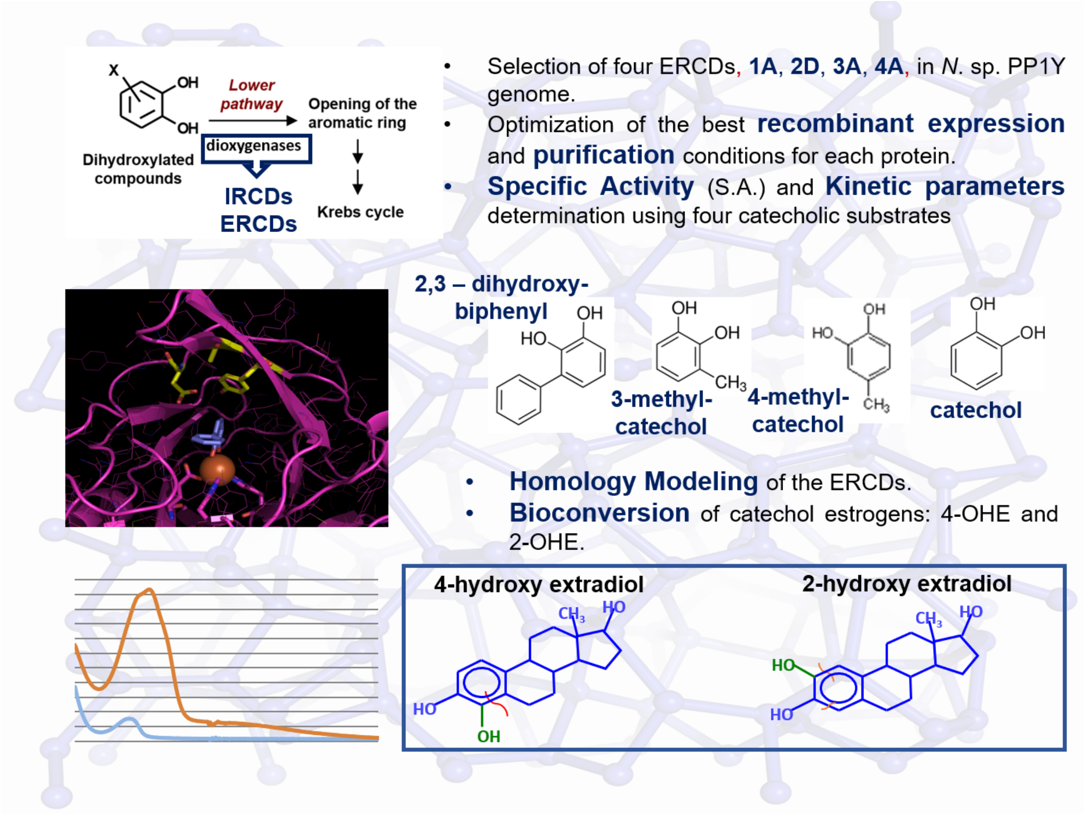
2- Recombinant expression and purification of the RHA-P α-L-rhamnosidase from *N*. sp. PP1Y and characterization of the enzyme potential to catalyze the bioconversion of several flavonoids of pharmaceutical interest. The substrate specificity of RHA-P will be studied in order to evaluate the effective use of this enzyme in bioconversion processes and its potential in analytical chemistry experiments aimed at defining the isomeric configuration, structure and conformation of complex molecules. In addition, the activity of RHA-P will be studied using recombinant whole cells of *E.coli* expressing the enzyme.

3- Identification and characterization of novel GHs enzymes active towards the arabinogalactan cell wall polysaccharide of *M. tuberculosis*, to use as a protection towards *M. tuberculosis* infection. In particular, the project start from the RNA-seq data obtained from Bacteroides species after exposure to the arabinogalactan polysaccharide as a nutrient in the medium. Target GHs were chosen from the identified upregulated PULs, their coding sequence was cloned, the proteins were recombinantly expressed, purified and their enzymatic activity was characterized on different polysaccharidic and arabinogalactan-derived substrates.

The task 3 was developed starting February 2018 through August 2018 at ICaMB (Institute for Cell and Molecular Biosciences), Newcastle University - Newcastle upon Tyne (UK), in the laboratory of Prof. Henry J. Gilbert, under the supervision of Dr. David Bolam.

CHAPTER 2

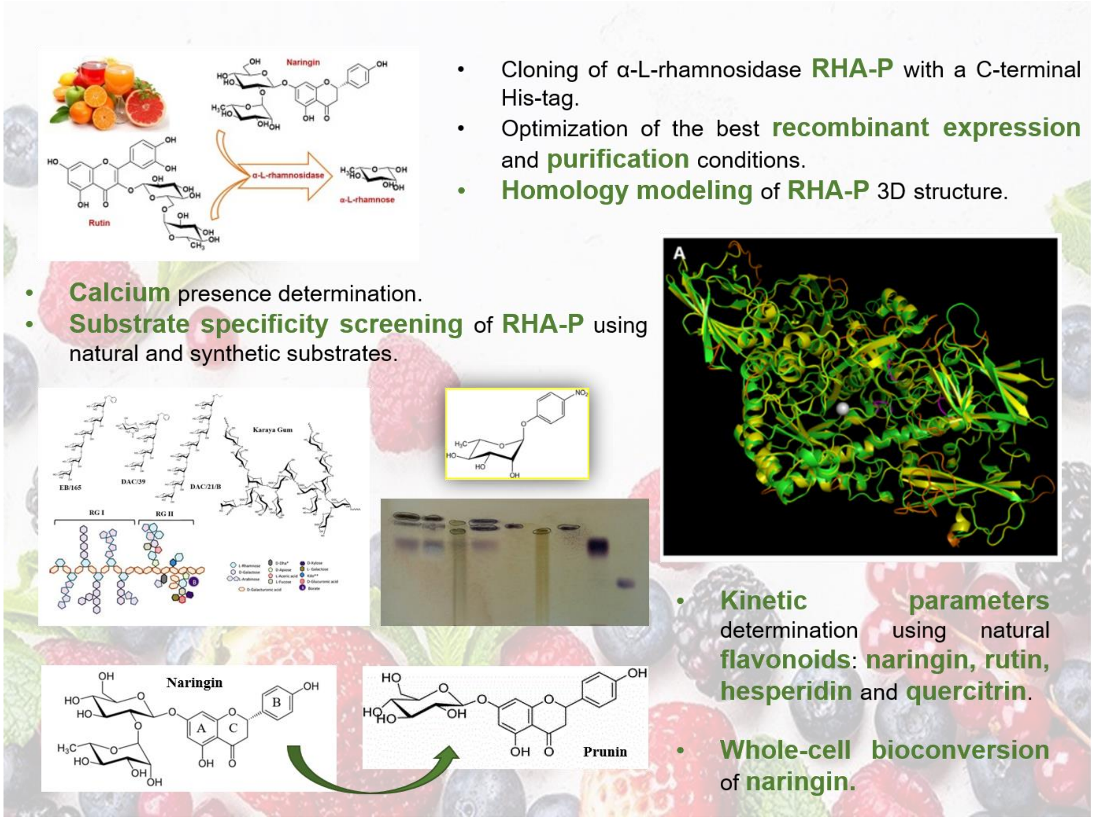
Novosphingobium sp. PP1Y as a source of extradiol ring - cleavage dioxygenases of biotechnological relevance


X
OH
OH
Dihydroxylated compounds
Lower pathway
Opening of the aromatic ring
dioxygenases
IRCDs
ERCDs
Krebs cycle
2,3 – dihydroxy-biphenyl
HO
OH
OH
OH
CH3
3-methyl-catechol
HO
OH
CH3
4-methyl-catechol
OH
OH
catechol
4-hydroxy extradiol
2-hydroxy extradiol
CH3
HO
HO
OH
CH3
HO
HO
HO

- Selection of four ERCDs, **1A**, **2D**, **3A**, **4A**, in *N.* sp. PP1Y genome.
- Optimization of the best **recombinant expression** and **purification** conditions for each protein.
- **Specific Activity** (S.A.) and **Kinetic parameters** determination using four catecholic substrates
- **Homology Modeling** of the ERCDs.
- **Bioconversion** of catechol estrogens: 4-OHE and 2-OHE.

CHAPTER 3

RHA-P, biotechnological relevance of a novel α-L-rhamnosidase from *Novosphingobium* sp. PP1Y.


Naringin
Rutin
α-L-rhamnosidase
α-L-rhamnose
A
Karaya Gum
EB/165
DAC/39
DAC/21/B
RG I
RG II
Naringin
A
B
C
Prunin

- Cloning of α-L-rhamnosidase **RHA-P** with a C-terminal His-tag.
- Optimization of the best **recombinant expression** and **purification** conditions.
- **Homology modeling** of **RHA-P** 3D structure.

- **Calcium** presence determination.
- **Substrate specificity screening** of **RHA-P** using natural and synthetic substrates.

- **Kinetic parameters** determination using natural **flavonoids**: **naringin, rutin, hesperidin** and **quercitrin**.
- **Whole-cell bioconversion** of **naringin.**

CHAPTER 3

RHA-P, biotechnological relevance of a novel α-L-rhamnosidase from *Novosphingobium* sp. PP1Y.

3.1 RHA-P subcloning, recombinant expression and purification.

The crude extract of *Novosphingobium* sp. PP1Y showed the ability to hydrolyze selected natural flavonoidic compounds that have in their chemical structure both α-rhamnose and β-glucose units such as naringin, diosmin, rutin, hesperidin, neohesperidin dihydrochalcone and quercitrin (Figure 14). Eight hypothetical *orf*s coding for rhamnosidases were identified in the *N*. sp. PP1Y genome, which -according to the CAZY database- belong to four glycosyl hydrolase (GH) families known as GH13 (5 *orfs*), GH15 (1 *orf*), GH106 (1 *orf*) and GH-NC (1 *orf*). Among these, structure to function relationships in the GH106 family have not yet been fully clarified, thus hampering the evaluation of their biotechnological implement. To the best of our knowledge, in fact, of the 319 members of the GH106 family, a single 3D-structure has been determined, the BT0986 from *Bacteroides thetaiotaomicron*, which shows a $(\beta/\alpha)_8$-barrel and catalyzes the hydrolysis of a α-L-rhamnopyranoside bound to the C2 position of an arabinofuranoside (L-Rhap-α-1,2-L-Araf) [115]. A single *orf* coding for a α-L-rhamnosidase of the GH106 family was identified in the genome of strain PP1Y, cloned in pET22b(+) vector, recombinantly expressed and partially characterized [116].

This enzyme, named RHA-P, was characterized as an *inverting* monomeric glycosidase of about 120 kDa. A preliminary biochemical characterization using *p*NPR (para-nitrophenyl-α-L-rhamnopyranoside) showed that RHA-P has an optimal activity in the temperature range 35-45°C, between pH 6.0-7.5, and has moderate tolerance to the presence of organic solvents in the reaction mixture [116].

In this research project, further details have been added to the previous biochemical characterization of RHA-P; in addition, the biotechnological potential of this enzyme in bioconversion processes involving the use of natural flavonoids substrates was investigated.

Theoretical parameters of RHA-P were determined by entering the coding sequence in the online software ProtParam and are reported in Table 8.

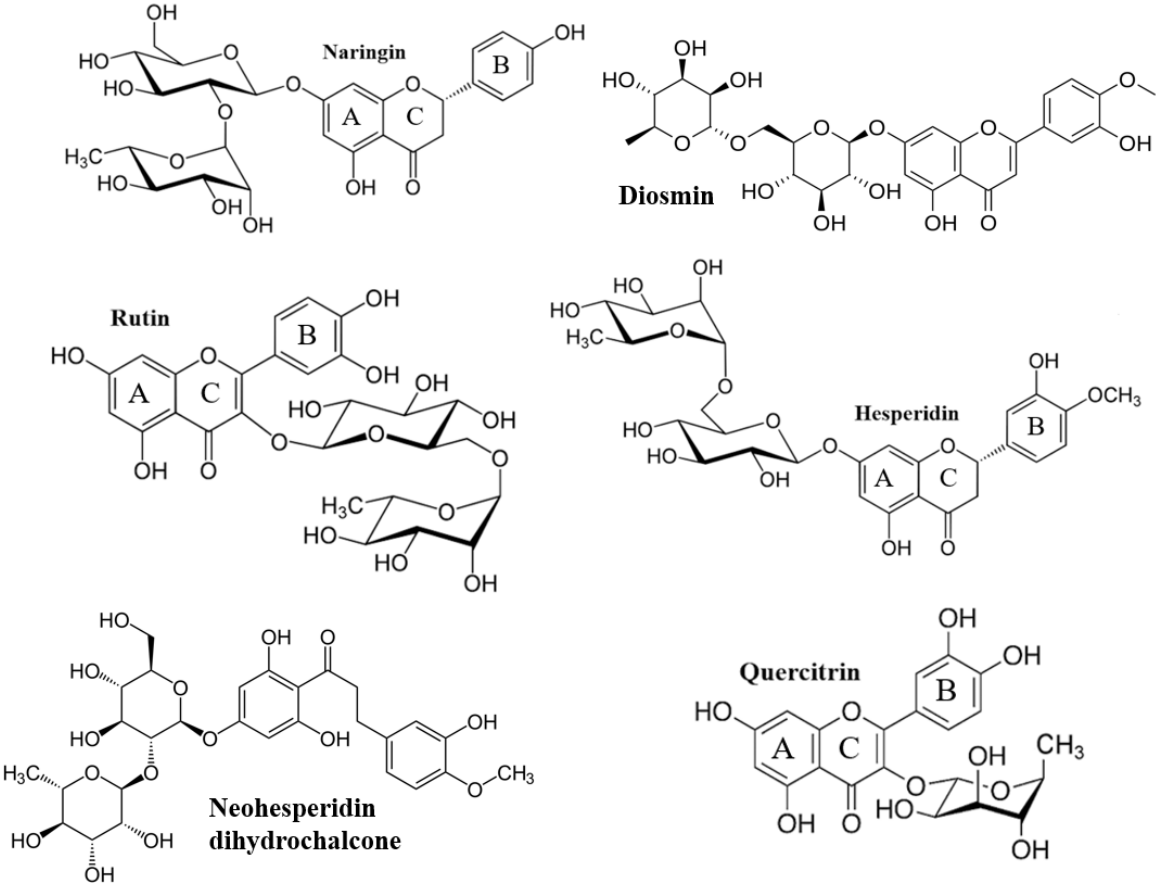

Naringin
Diosmin
Rutin
Hesperidin
Neohesperidin dihydrochalcone
Quercitrin
A
B
C

Figure 14: Chemical structure of rhamnosylated natural flavonoids tested with *N*. sp. PP1Y crude extract.

	Molecular weight (kDa)	**Number of residues**	**Theoretical isoelectric point**	**$\varepsilon_{280\ nm}$ mM^{-1} cm^{-1}**
RHA-P	124.2	1146	5.6	188.2

Table 8: The molecular weight (kDa), number of residues and theoretical isoelectric point of RHA-P from *N.* sp. PP1Y.

Previous work on the untagged recombinant protein RHA-P was hampered by a low yield of purified protein, likely due to the purification strategy used. Therefore, the *orf* coding for RHA-P was subcloned with a C-terminal His-tag (plasmid pET22b(+)/*rha-his*), with the aim of improving the yield and the purity of the recombinant protein. Plasmid pET22b(+)/*rha-his* was obtained as described in paragraph I.2.1. and used to transform *E. coli* BL21(DE3) competent cells. Analytical experiments were performed to determine the best conditions for recombinant expression; this task was done by evaluating in different experimental settings:

- overall expression levels, by SDS-PAGE and densitometric analysis of the soluble and insoluble fractions obtained.
- The specific activity (S.A.) of RHA-P in the soluble fractions, using *p*NPR as substrate

First attempts of expressing the recombinant protein at 37 °C in LB medium, showed the foremost presence of RHA-P in the inclusion bodies. Therefore, analytical expression experiments were set up as outlined in paragraph I.3.2., to optimize the yield of active RHA-P in the soluble fraction of induced cultures.

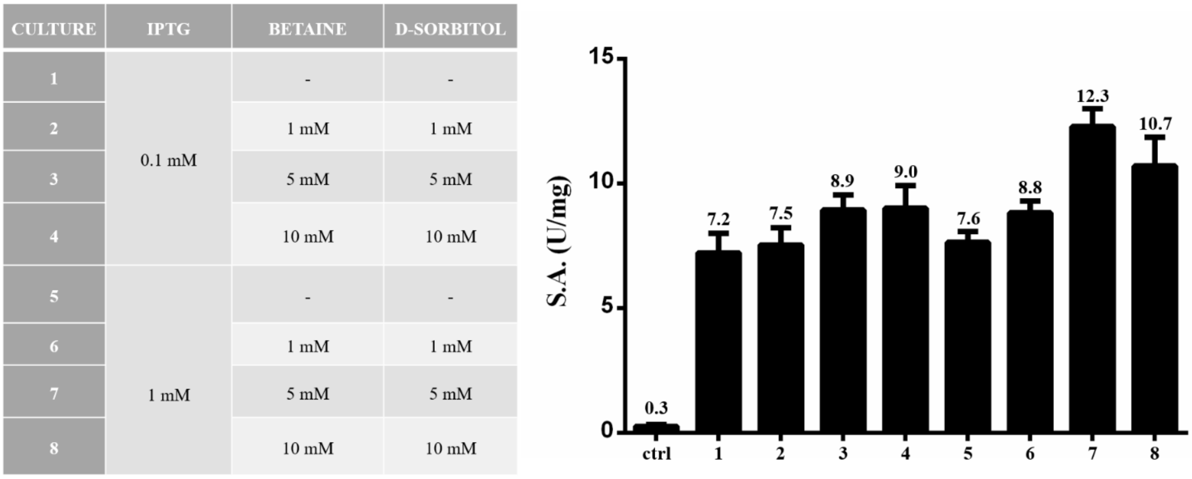
To this purpose, different experimental settings were performed, varying:

- salt and/ or osmolytes (betaine and sorbitol) concentration in LB medium
- induction temperature, (23°C or 37°C)
- concentration of inducer (IPTG) (0.1 or 1 mM)

The best expression condition identified at this stage involved an induction with 1 mM IPTG, at 23°C, using LB supplemented with 5 mM of both betaine and sorbitol (Figure 15). These osmolytes were added to the medium acting as “chemical chaperones” and, along with the induction at 23°C, concurred in lowering the formation of inclusion bodies, as already observed for the recombinant RHA-P previously expressed without His-tag [112].

Once optimized the recombinant expression conditions, a novel purification protocol was set-up, which included several modifications to the previous one [116].

Evidences previously obtained for RHA-P highlighted the presence of a N-terminal signal peptide, which was probably cleaved through a post-translational proteolytic processing. Indeed, the lack - with respect to the translated sequence - of a 23-aminoacids long peptide at the protein N-terminus was retrieved in the mature protein both by LC-MS and N-terminus sequencing analysis [116]. The presence of a similar peptide has also been described for the α-RHA isolated from *Sphingomonas paucimobilis* FP2001, which shares a 48% sequence identity with RHA-P [138]. Moreover, the inspection of the putative signal peptide revealed several peculiarities common to other bacterial signal peptides, such as the presence of a charged region (2 – 5 residues) followed by a hydrophobic stretch of ~ 12 aa and, 1-3 aa upstream of the cleavage site, the occurrence of small and apolar residues. Signal peptides have been also described in α-RHAs isolated from fungal sources, such as the enzymes from *Aspergillus kawachii* and *Aspergillus aculeatus* [109, 139]. In all cases, the efficient cleavage of a 17-20 aminoacids long N-terminal peptide was observed in the recombinant proteins. Although the functional reason for this specific sorting of an α-RHA is still not clear, this might be related to the engagement of this enzymatic activity in the complex molecular machinery involved in the biosynthesis and maintenance of the bacterial cell wall, where L-rhamnose appears to be a component of membrane rhamnolipids and polysaccharides [140]. Therefore, the functional role of this cleavage is possibly related to the sorting of these proteins in the periplasmic space. To confirm the effective presence of RHA-P in the periplasmic fraction of recombinant cells of *E. coli*, samples corresponding to total cell extract, periplasmic and cytoplasmic fractions were obtained as described in paragraph I.4.3., analyzed by SDS-PAGE and the rhamnosidase S.A. was measured in each sample.



CULTURE	IPTG	BETAINE	D-SORBITOL
1	0.1 mM	-	-
2		1 mM	1 mM
3		5 mM	5 mM
4		10 mM	10 mM
5	1 mM	-	-
6		1 mM	1 mM
7		5 mM	5 mM
8		10 mM	10 mM

Figure 15: RHA-P analytical expression experiments. S.A. of RHA-P cell extracts was measured under different experimental conditions (number from 1 to 8 indicate different experimental conditions). [141]

Results showed that only a negligible part of total α-RHA activity present in the crude cell extract was detected in the cytoplasmic fraction, whereas most of the activity could be recovered in the periplasmic fraction (Figure 16 a). This suggested that RHA-P was mainly sorted in the periplasm, which was also evident from the SDS-PAGE analysis (Figure 16 b) [141].

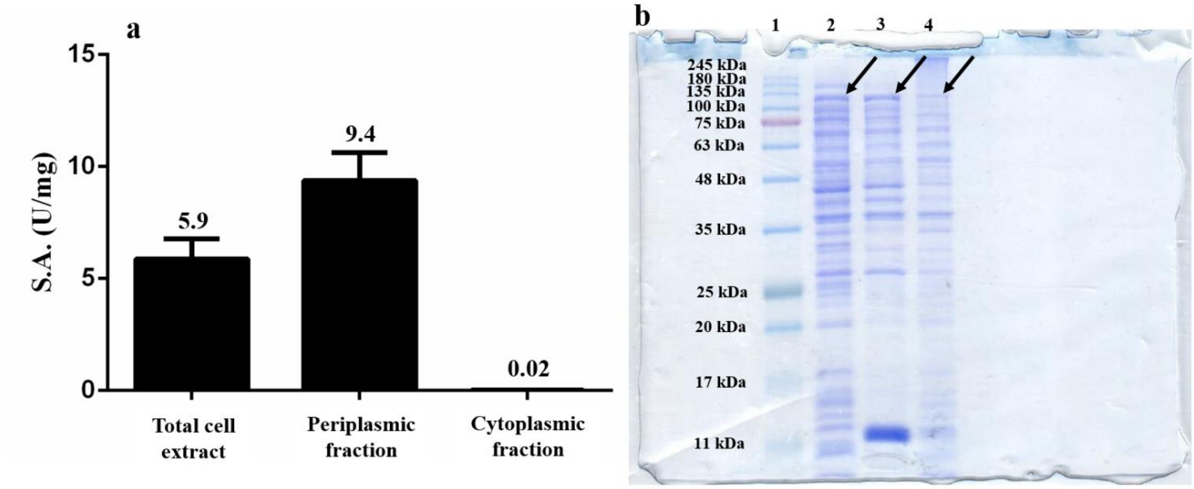


Figure 16: **a**: S.A. of RHA-P in different cellular fractions. **b**: SDS-PAGE analysis of different cellular fractions. Lane 1= MW standard Lane 2= 5 µg of total cell extract. Lane 3= 5 µg of periplasmic fraction. Lane 4 = 5 µg of cytoplasmic fraction.
Black arrows in all lanes indicate a band with the molecular mass expected for RHA-P. [141]

Large-scale recombinant expression of RHA-P was then performed using the optimized condition described above. A large-scale extraction of the periplasm fraction was carried out, and RHA-P purification was performed using a single step affinity chromatography (paragraph I.4.3.). Specific activity and total units of enzyme present in cell lysates and recovered in purified fractions are shown in Figure 17 A. Even though the total yield of RHA-P obtained is not very high (26%), a purification factor of 12.7 is significantly higher compared to that previously obtained with the untagged protein (6.9). In addition, SDS-PAGE analysis of purified RHA-P confirmed the presence of very few contaminants in the sample (Figure 17 B).

At last, to evaluate the potential effect of the His-tag on protein function, S.A. values of purified RHA-P and untagged RHA-P were compared [116]. Data highlighted a negligible decrease of S.A. on *p*NPR, with RHA-P showing a residual 87% rhamnosidase activity compared to the untagged protein. In addition, kinetic analysis confirmed that K_M and k_{cat}/K_M values obtained for RHA-P (156.3 ± 22.3 µM^{-1} and 4.4 sec^{-1} µM^{-1} respectively) were comparable to those obtained with the untagged protein (160.2 ±17.3 µM^{-1} and 4.6 sec^{-1} µM^{-1}).

A

Sample	S.A. (U/mg)	U tot	P.F.	Yield
Periplasm fraction	45.08	4003.60		
RHA-P	571.65	1042.70	12.70	26.00

B

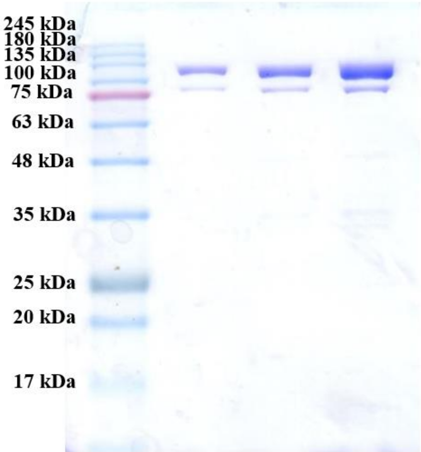


Figure 17: **A**: Purification table of RHA-P. Total enzymatic Units, Specific activity (S.A.), Purification factor (P.F.) and total yield are reported. **B**: SDS-PAGE analysis of RHA-P (1, 3 and 5 µg) purified by affinity chromatography.

3.2 RHA-P *Homology modeling*.

To investigate the active site topology and substrate specificity of RHA-P, a homology model was obtained in collaboration with Dr. Eugenio Notomista at the Department of Biology (University Federico II of Naples), using the structure of rhamnosidase BT0986 from *Bacteroides thetaiotaomicron* (pdb code, 5MWK) as template [141]. α-L-rhamnosidase BT0986 from *B. thetaiotaomicron* hydrolyze cleaves terminal rhamnose residues in oligosaccharidic substrates and has a N-terminal $(\alpha/\beta)_8$-barrel catalytic domain with E593 and E461 residues as potential catalytic base and acid, respectively. Moreover, in the active site, glutamate residues bind calcium, which makes polar interactions with O-2 and O-3 of L-rhamnopyranose (L-Rhap). The importance of a catalytic calcium ion was found within other α-mannosidases *inverting* enzymes, which often use divalent metal ions for substrate binding [142].

More in detail, the BT0986 5MWK structure, which has the E461Q mutation to prevent substrate hydrolysis, was chosen among the available database structures because it shows a branched oligosaccharide bound into the active site. The oligosaccharide has, at the non-reducing end, a rhamnose/arabinose disaccharidic unit with a α-1,2 bond (L-rha-α-1,2-L-ara), which is hydrolyzed by BT0986 [115].

Even though the percent identity with RHA-P is not high (27%) due to several insertions and deletions in RHA-P aminoacidic sequence, the main differences are located on the surface of the protein far from the active site (Figure 18 A). Noteworthy, a deletion and two insertions in RHA-P are in loops contributing to the active site pocket. These three loops are close to each other and are located in the active site pocket corresponding to the arabinose binding sub-site in BT0986.

By contrast, the calcium binding residues and the rhamnose sub-site appear well conserved (Figure 18B and 18C). In BT0986 structure, five residues contribute to Ca^{2+} binding, namely E593, E538, D458, S459 and E561. Of these five residues, only E561, whose mutation into alanine inactivates BT0986 [115], is not conserved in RHA-P that has instead an alanine at that position. It should be kept in mind, however, that the actual role of E561 in BT0986 is still under investigation. In fact, in BT0986 structure without the bound substrate (pdb code, 5MQN), E561 does not bind calcium and in both BT0986 structures E561 forms an ionic couple with K571, which in turn binds to the arabinose residue. Thus, the E561-K571 couple is likely involved in substrate recognition. RHA-P may not have this arabinose binding sub-site, because K571 is replaced with a leucine. Indeed, a loop containing W276, another residue contributing to the arabinose binding site in BT0986, is deleted in RHA-P (Figure 18C), substituted in an equivalent position by a loop bearing an insertion of two residues, D578 and H588 (Figure 18C).

Alll residues contacting rhamnose in the structure of BT0986, are instead conserved in RHA-P or at least show conservative mutations (Figure 18C). The putative acid catalyst, E506, which would correspond to E461 in BT0986, is also maintained (Figure 18B and 18C).

In conclusion, the model obtained suggested that the overall catalytic mechanism of RHA-P might be very similar to that of BT0986, but several details indirectly point to a quite different substrate specificity.

This hypothesis is also supported by the evidence that BT0986 is barely active on *p*NPR and has instead a significant activity on oligosaccharides containing rhamnose, such as rhamnose/arabinose units (α-L-rha-α-1,2-L-ara), for which a k_{cat}/K_M of 23.5 sec^{-1} mM^{-1} has been measured [115]. Conversely, RHA-P is active on aryl-rhamnosides, with a k_{cat}/K_M of 4.4 sec^{-1} mM^{-1} on *p*NPR and has k_{cat}/K_M values ranging from 100 to 500 sec^{-1} mM^{-1} on natural flavonoids. The hypothesis can be advanced that RHA-P active site has evolved to be active on substrates containing aryl- groups and aromatic moieties, rather than poly- and oligosaccharidic compounds.

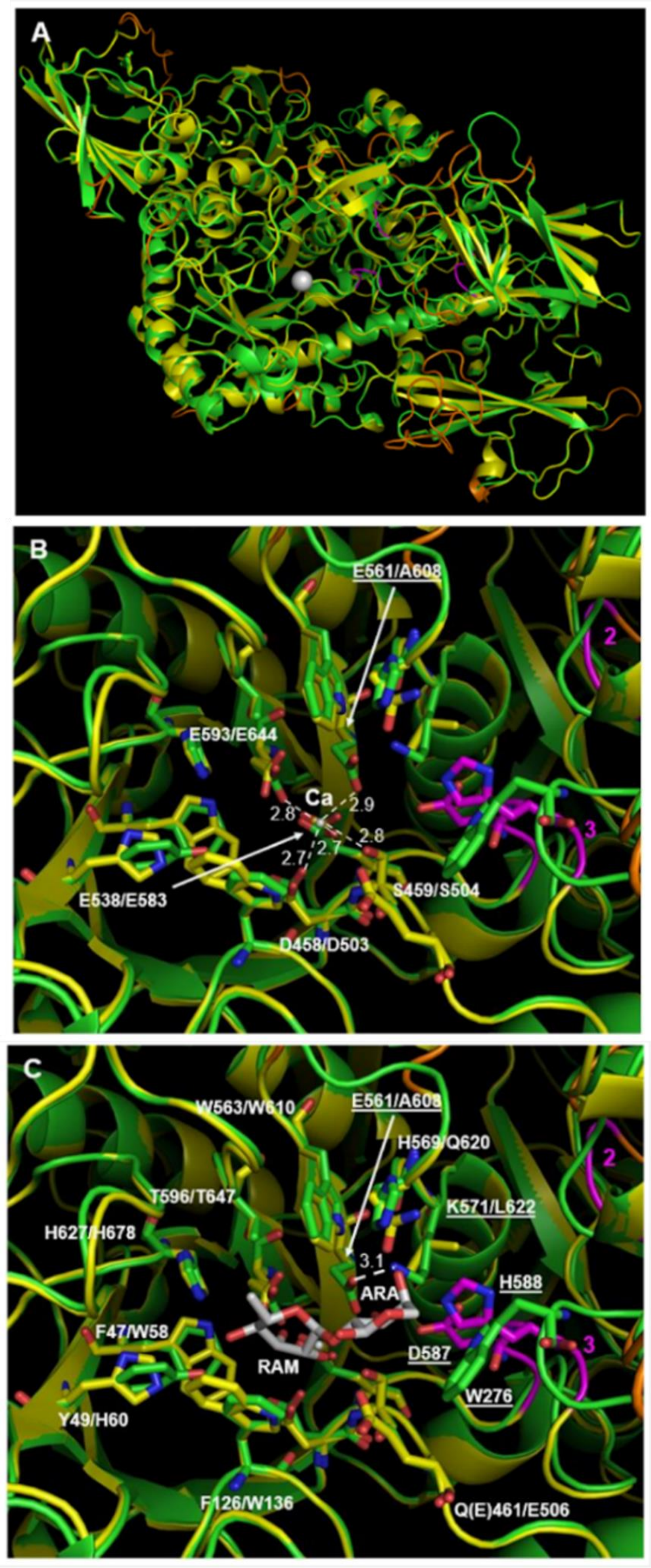

A
B
E561/A608
2
E593/E644
Ca
2.9
2.8
2.8
2.7
2.7
3
E538/E583
S459/S504
D458/D503
C
W563/W610
E561/A608
H569/Q620
2
T596/T647
K571/L622
H627/H678
3.1
H588
ARA
F47/W58
RAM
D587
3
W276
Y49/H60
F126/W136
Q(E)461/E506

Figure 18: Comparison between the structure of BT0986 from *Bacteroides thetaiotaomicron* (pdb code, 5MWK) and the homology model of RHA-P. Panel A: shows the structure and the model as green and yellow cartoons, respectively. In the model of RHA-P, loops with insertions or deletions with respect to BT0986 are in orange. A loop with a deletion (1) and two loops with insertions (2 and 3), close to the active site, are shown in magenta. Magenta arrows show the variation of the three loops with respect to the corresponding loops in BT0986. **Panel B**: shows a closeup of the active site with an orientation approximately similar to that shown in panel A. Residues interacting with calcium in BT0986 (distances are in Angstrom) and corresponding residues in RHA-P, are shown. **Panel C**: shows a close up of the active site with an orientation approximately similar to that shown in panel A. The disaccharide rhamnose (RAM)/arabinose (ARA) and residues contacting the two sugars are shown (RAM: F126/W136, Y49/H60, F47/W58, H627/H678, T596/T647; ARA: H569/Q620, K571/L622, W276 in BT0986, D578 and HY588 in RHA-P; W563/W610, E561/A608 and E461/E506 are at the boundary between rhamnose and arabinose sub-sites).
Significant residues are denoted as sticks colored by atom type (oxygen, red; nitrogen, blue; carbon atoms of the proteins are colored as the cartoons shown in panel A; carbon atoms of the disaccharide in panel C are colored in grey). For each pair of corresponding residues, the residue number of BT0986 is shown first [141].

3.3 Evidences for calcium presence in RHA-P.

As previously outlined, the presence of a calcium ion essential for catalysis has been reported in some GH families, although with different functions. Three crystal structures currently available for GH106 rhamnosidases, among which BT0986 structure used for the homology model, confirmed the presence of this essential cofactor [112, 113, 115]. To investigate the presence of Ca^{2+} also in RHA-P, the enzyme was first incubated in presence of different concentrations of EDTA and then the standard *p*NPR assay was performed, as described in Materials and Methods (Appendix I) [141].
Results presented in Figure 19 A showed a marked reduction of RHA-P specific activity in line with the increase of EDTA concentration, ranging from 0.1 to 1 mM. The enzymatic activity recovery was then evaluated after incubating RHA-P with 2 mM EDTA and then directly adding to the sample either $CaCl_2$ or $MgCl_2$, $MnCl_2$, $ZnCl_2$ at concentrations ranging from 3 to 50 mM (Fig. 19 B). Again, an almost total inactivation of RHA-P after incubation with 2 mM EDTA was observed (4% of residual activity). An almost complete recovery of enzymatic activity was achieved after incubating the EDTA-treated enzyme with all Ca^{2+} concentration used. Noteworthy, also a 56%, 92% and 79% of the initial rhamnosidases specific activity was recovered after incubation with 3 mM, 10 mM and 50 mM of Mg^{2+}, respectively.
As evident from Fig. 19 B, significantly lower percentages of recovery were recorded in presence of increasing concentrations of Mn^{2+}, leading to a maximum recovery of 61% with 50 mM of this metal cation. In contrast, only 3 mM of Zn^{2+} led to a 30% of recovery of enzymatic activity, whereas higher concentrations of this element promptly inactivated RHA-P catalytic activity.
These data altogheter support the evidence of the presence in RHA-P of Ca^{2+} ion essential for catalysis also in RHA-P, therefore suggesting a catalytic mechanism similar to that of BT0986 from *Bacteroides thetaiotaomicron*.

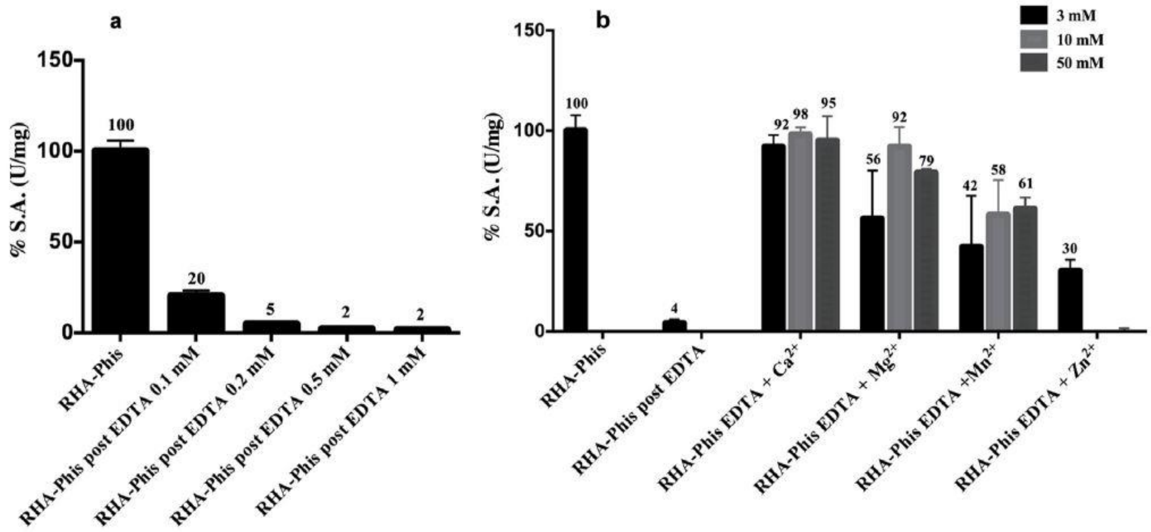
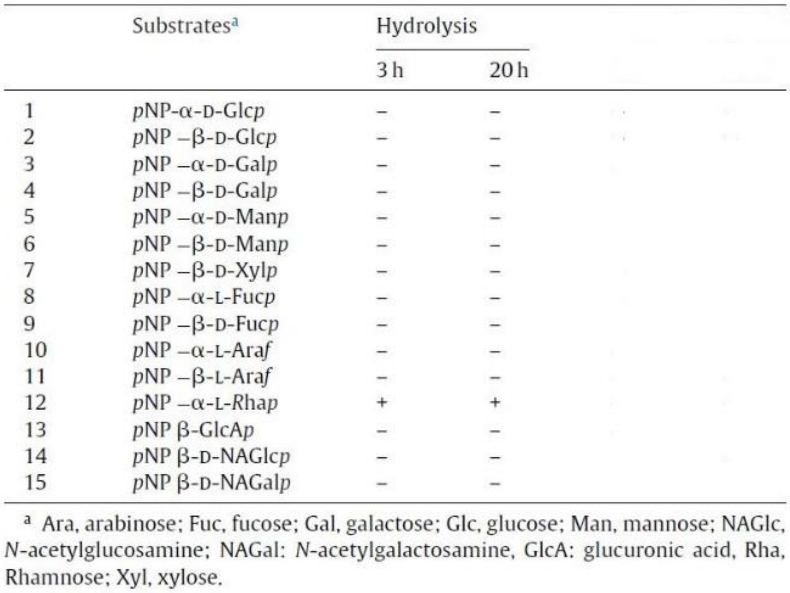


Figure 19: **a**: *p*NPR assays of RHA-P in presence of different EDTA concentrations. **b**: *p*NPR assays of EDTA-treated RHA-P after incubation with different Ca^{2+}, Mg^{2+}, Mn^{2+} and Zn^{2+} concentrations.
Rhamnosidase specific activity is reported as percentage of residual activity compared to control [141].

3.4 RHA-P substrate specificity.

To better define RHA-P substrate specificity, the enzymatic activity was characterized on several *p*NP-α- and *p*NP-β- sugars.
The purified enzyme showed activity only on *p*NPR and not on *p*NP-β-D-Glc, confirming that RHA-P is indeed a pure α-L-rhamnosidase and does not act as a naringinase, a class of glycosyl hydrolase having both α-L-rhamnosidase and β-D-glucosidase activities [106]. No activity on any other *p*NP-α- and *p*NP-β- sugars was detected suggesting RHA-P specificity towards rhamnose residues (Table 9).
RHA-P activity was also tested on some synthetic rhamnose oligosaccharides, named EB/165, DAC/39 and DAC/21/B (Figure 20), and rhamnose-containing polysaccharidic compounds, such as rhamnogalacturonan I and II, karaya gum and galactan from *M. smegmatis* (Figure 20).
At this stage our aim was to evaluate the ability of RHA-P to hydrolyze rhamnose in a range of different glycosidic linkages arrangement with variable chain length and in complex branched polysaccharidic chains. However, TLC analysis showed that RHA-P was unable to catalyze the hydrolysis of these substrates, not even extending incubation time to 24 hr (data not shown).



	Substrates[a]	Hydrolysis	
		3 h	20 h
1	*p*NP-α-D-Glc*p*	–	–
2	*p*NP –β-D-Glc*p*	–	–
3	*p*NP –α-D-Gal*p*	–	–
4	*p*NP –β-D-Gal*p*	–	–
5	*p*NP –α-D-Man*p*	–	–
6	*p*NP –β-D-Man*p*	–	–
7	*p*NP –β-D-Xyl*p*	–	–
8	*p*NP –α-L-Fuc*p*	–	–
9	*p*NP –β-D-Fuc*p*	–	–
10	*p*NP –α-L-Ara*f*	–	–
11	*p*NP –β-L-Ara*f*	–	–
12	*p*NP –α-L-Rha*p*	+	+
13	*p*NP β-GlcA*p*	–	–
14	*p*NP β-D-NAGlc*p*	–	–
15	*p*NP β-D-NAGal*p*	–	–

[a] Ara, arabinose; Fuc, fucose; Gal, galactose; Glc, glucose; Man, mannose; NAGlc, *N*-acetylglucosamine; NAGal: *N*-acetylgalactosamine, GlcA: glucuronic acid, Rha, Rhamnose; Xyl, xylose.

Table 9: Screening of RHA-P reactions with different pNP-α- and pNP-β- sugars as substrates [143].

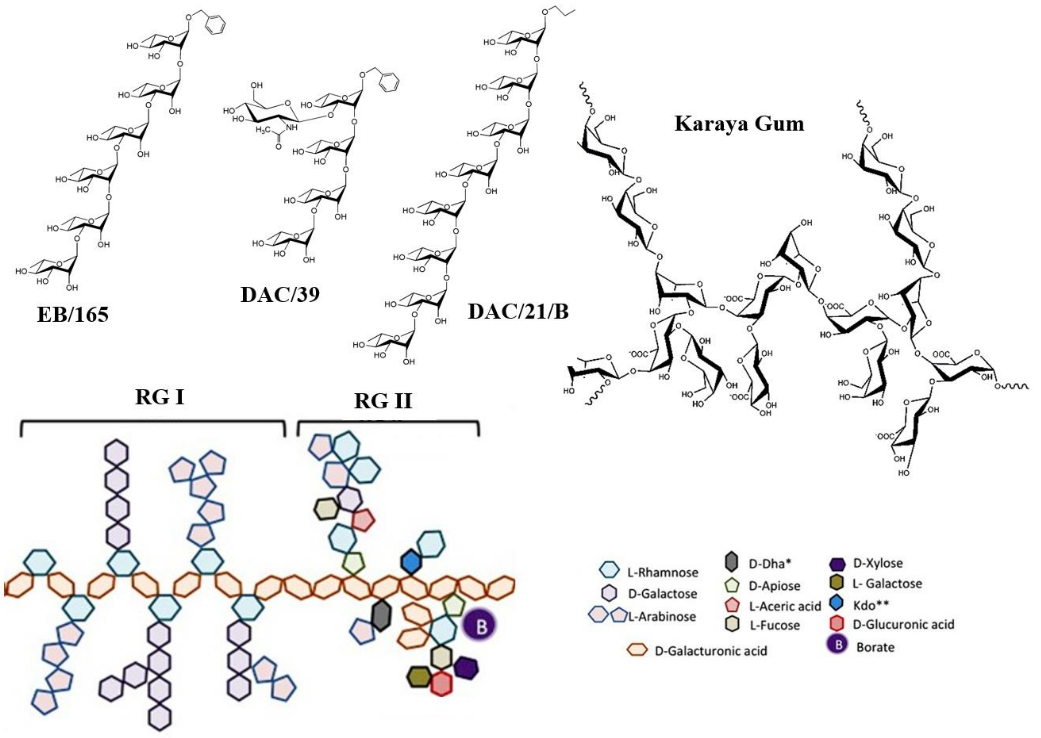

EB/165
DAC/39
DAC/21/B
Karaya Gum
RG I
RG II
L-Rhamnose
D-Galactose
L-Arabinose
D-Galacturonic acid
D-Dha*
D-Apiose
L-Aceric acid
L-Fucose
D-Xylose
L- Galactose
Kdo**
D-Glucuronic acid
B
Borate

Figure 20: Chemical structure of rhamnose containing natural and synthetic substrates used for an activity screening with RHA-P.

Afterwards, the ability of purified RHA-P to hydrolyze natural flavonoids was tested on substrates such as naringin, rutin and neohesperidin dihydrochalcone and time courses were followed by TLC analysis.

Results are reported in Figure 21. After 3 hours of incubation, TLC analysis of reaction mixtures showed a 40-60% hydrolysis of neohesperidin dihydrochalcone and rutin in the corresponding derhamnosylated neohesperidin dihydrochalcone and quercetin-3-β-glucopyranoside (isoquercitrin), respectively. Besides, an almost total conversion of naringin was observed in the corresponding products, rhamnose and prunin. Reactions carried out with naringin in buffer containing 10% DMSO showed an almost complete conversion of naringin in prunin and rhamnose after just 1hour.

The data analyzed by TLC gave preliminary insights into RHA-P substrate specificity, showing the ability of this enzyme to hydrolyze both α-1,2 and α-1,6 linkages. Results supported the indications obtained from the *homology modeling* analysis, which suggested a substrate specificity towards substrates containing aromatic moieties. Conversely, RHA-P was unable to hydrolyze substrates containing oligosaccharic chains (EB/165, DAC/39 and DAC/21/B), neither polysaccharides in which rhamnose is found as a component in the backbone, or in side chains (RGI, RGII and galactan).

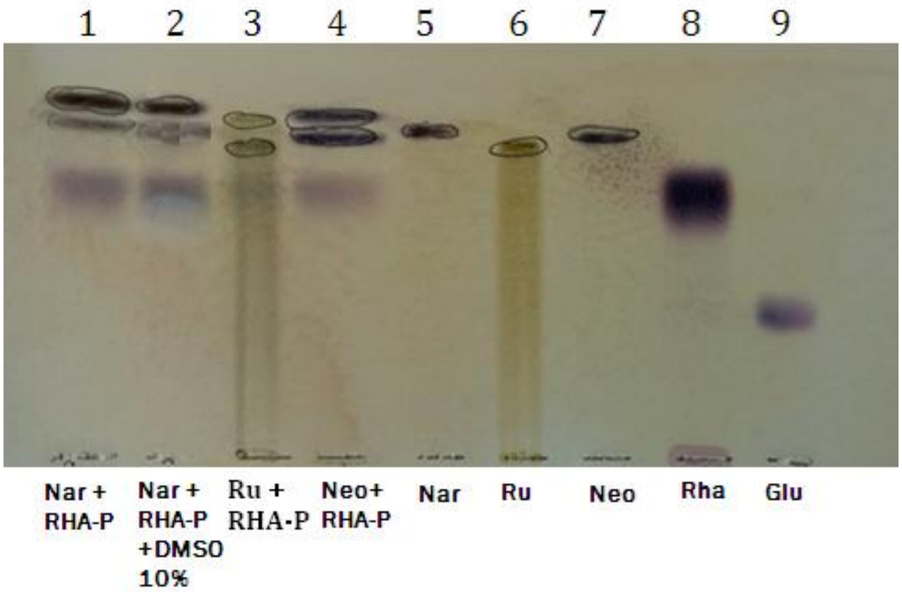


Figure 21: TLC analysis of RHA-P hydrolysis on flavonoids. Lane 1: Reaction mixture of RHA-P containing naringin. Lane 2: Reaction mixture in 10% DMSO of RHA-P containing naringin. Lane 3: Reaction mixture of RHA-P containing rutin. Lane 4: Reaction mixture of RHA-P containing neohesperidin dihydrochalcone. Lane 5: Naringin standard. Lane 6: Rutin standard. Lane 7: Neohesperidin dihydrochalcone standard. Lane 8: Rhamnose standard. Lane 9: Glucose standard.

3.5 Kinetic characterization of RHA-P activity on natural flavonoids.

Experiments were performed to evaluate the kinetic behavior of RHA-P on naringin, rutin, hesperidin and quercitrin (Figure 5) [141].

Reactions were carried out and rhamnose release was determined using high-performance anion-exchange chromatography, with pulsed amperometric detector (HPAEC-PAD), as described in Materials and Methods (Appendix I). Reaction rates (µM min^{-1}), plotted as a function of substrate concentrations (mM), showed typical Michaelis-Menten trends and are reported in Fig. 22. Kinetic constants summarized in Table 10 highlight that RHA-P is active towards all substrates and is able to hydrolyze single rhamnose units, with either α-1,2 or α-1,6 glycosidic linkages. The k_{cat}/K_M values obtained were higher for rutin and naringin (586.28 sec^{-1} mM^{-1} and 359.21 sec^{-1} mM^{-1} respectively), and lower for hesperidin and quercitrin (190.00 sec^{-1} mM^{-1} and 79.83 sec^{-1} mM^{-1} respectively) [141]. The analysis of these values underlined that RHA-P catalytic efficiency was comparable or higher than that of other α-RHAs described in literature [144, 145]. Interestingly, K_M values revealed a higher affinity of RHA-P for these natural flavonoids when compared to homologous enzymes reported in literature [144, 145]. Moreover, RHA-P S.A. values (104.3 ± 4.8 U/mg for naringin, 8.8 ± 0.5 U/mg for rutin, 6.2 ± 0.2 U/mg for hesperidin and 6.9 ± 0.5 U/mg for quercitrin) were higher than the ones measured for the α-RHAs from *Streptomyces avermitilis*, *Clostridium stercorarium* and *Pediococcus acidilactici* [145 - 147].

Specific activities towards naringin and rutin higher from 100 to 300 times were obtained when comparing RHA-P to the α-RHA from *Fusobacterium* K-60 [148], an enzyme considered a good quercitrin-hydrolyzing rhamnosidase. In this specific case, since RHA-P shows an almost equivalent activity for quercitrin, the enzyme from *N. sp.* PP1Y might be considered a valid alternative for the bioconversion of complex biological matrices containing all three flavonoids.

Finally, data suggested that the efficiency of the enzyme could be affected both by the position of the saccharidic linkage and the different nature of the bound aglycone. Overall, results confirmed that RHA-P was able to hydrolyze both α-1,2 and α-1,6 glycosidic linkages [116]. The hydrolysis of a α,1-6 glycosidic linkage bound in position C3 seems to be preferred compared to other types of linkage. Moreover, the presence of a disaccharidic unit seems to improve the catalytic efficiency of the enzyme compared to the rhamnose single unit, present for example in quercitrin. However, the analysis of the k_{cat}/K_M values highlighted some interesting differences in terms of substrate affinity and catalytic efficiency. In particular, rutin and quercitrin share the same flavonoidic aglyconic portion (named quercetin) and both have the saccharidic portion bound at position C3 on the C ring of the flavonoid. However, RHA-P showed a higher catalytic efficiency on rutin. On the other hand, rutin and hesperidin share the same rutinose α-1,6 disaccharidic residue, but bound at different positions of the flavonoidic ring (in position C3 on the C ring for rutin and at position C7 on the A ring for hesperidin). The higher catalytic efficiency of RHA-P towards rutin, might be related to the specific position of the flavonoidic ring to which the saccharidic portion is bound.

Finally, naringin and hesperidin have different glycosidic moieties (neohesperidose α-1,2 and rutinose α-1,6 respectively), bound in the same C7 position on the A ring of the flavonoid. The higher value of k_{cat}/K_M of RHA-P towards the former substrate suggested that the enzyme showed a preferential hydrolysis of the neohesperidose α-1,2 unit.

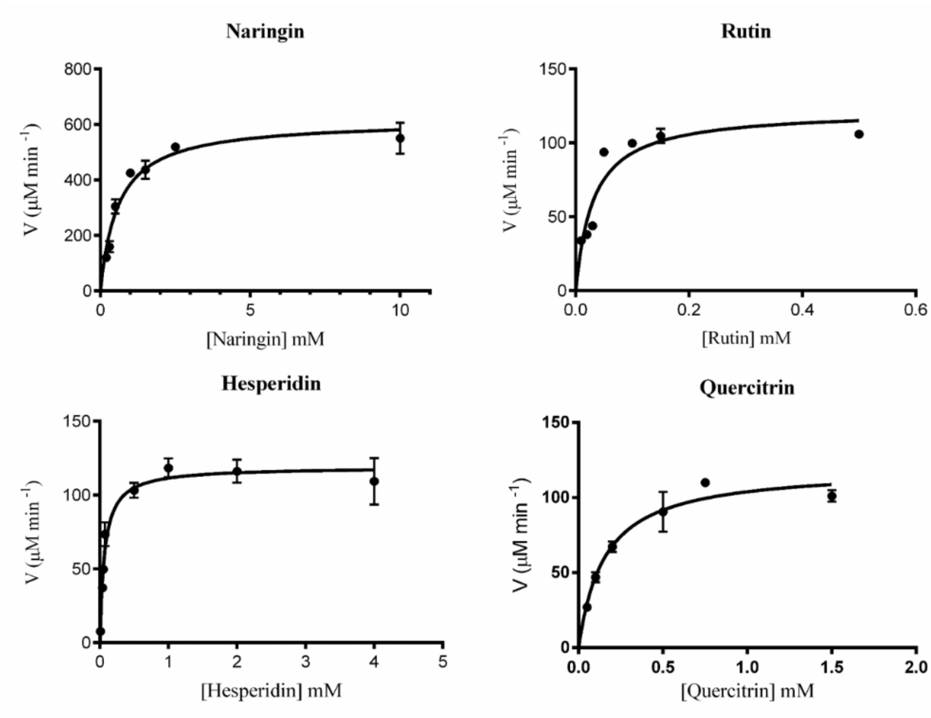


Figure 22: RHA-P kinetic behavior towards natural flavonoids. Reactions rates expressed in µM/min are plotted as a function of flavonoids concentrations [141].

Flavonoid	**Sugar**	**Glycosidic linkage position**	**K_M (mM)**	**k_{cat} (sec^{-1})**	**k_{cat}/K_M (sec^{-1} mM^{-1})**
Naringin	**Neohesperidose (α,1-2)**	**A ring Carbon 7**	**0.59 (±0.09)**	**214 (±10)**	**360 (±70)**
Rutin	**rutinose (α,1-6)**	**C ring Carbon 3**	**0.031 (±0.01)**	**18.2 (±1.2)**	**586 (±172)**
Hesperidin	**rutinose (α,1-6)**	**A ring Carbon 7**	**0.067 (±0.01)**	**12.7 (±0.4)**	**190 (±32)**
Quercitrin	**rhamnose**	**C ring Carbon 3**	**0.15 (±0.03)**	**12.2 (±0.4)**	**80 (±16)**

Table 10: **Kinetic constants obtained for** RHA-P using naringin, rutin, hesperidin and quercitrin as substrates [141].

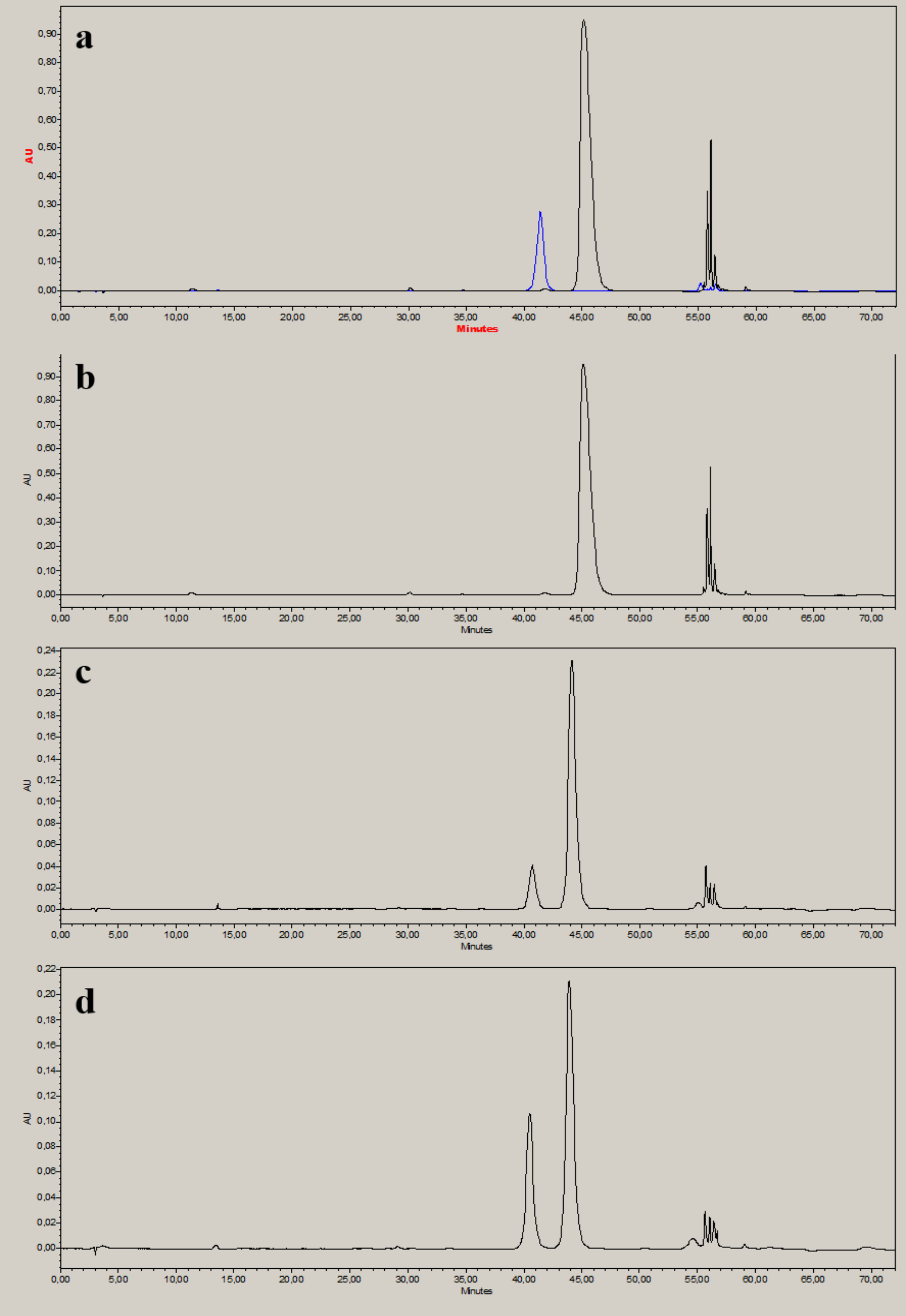
3.6 Whole cell bioconversion of naringin using RHA-P.

Whole cell bioconversion experiments were set up, using naringin as substrate.

After the recombinant expression of RHA-P in *E. coli* BL21(DE3) strain, cells were suspended in minimal media containing 2 mM naringin, as detailed in paragraph I.10, and the hydrolysis of naringin was observed over time using HPLC and monitoring ABS at 300 nm (paragraph I.11).

Time course results, shown in Figure 23, highlight a gradual decrease of naringin chromatographic peak (46 min) and the concomitant appearance of the derhamnosylated product (prunin: 41.5 min). Therefore, results show that whole cells expressing RHA-P are able to catalyze the complete hydrolysis of naringin with concomitant release of the de-rhamnosylated product in the growth medium. The qualitative analysis of the chromatogram highlights that using a 3 OD_{600}/mL concentration of recombinant cells, almost half of naringin is already converted in prunin after 30 minutes.

The same experiment performed with both *E. coli* BL21(DE3) cells and *E. coli* BL21(DE3) cells transformed with an empty pET22b(+) plasmid, used as negative controls, showed only the naringin chromatographic peak, whose area was constant during all time course (data not shown).



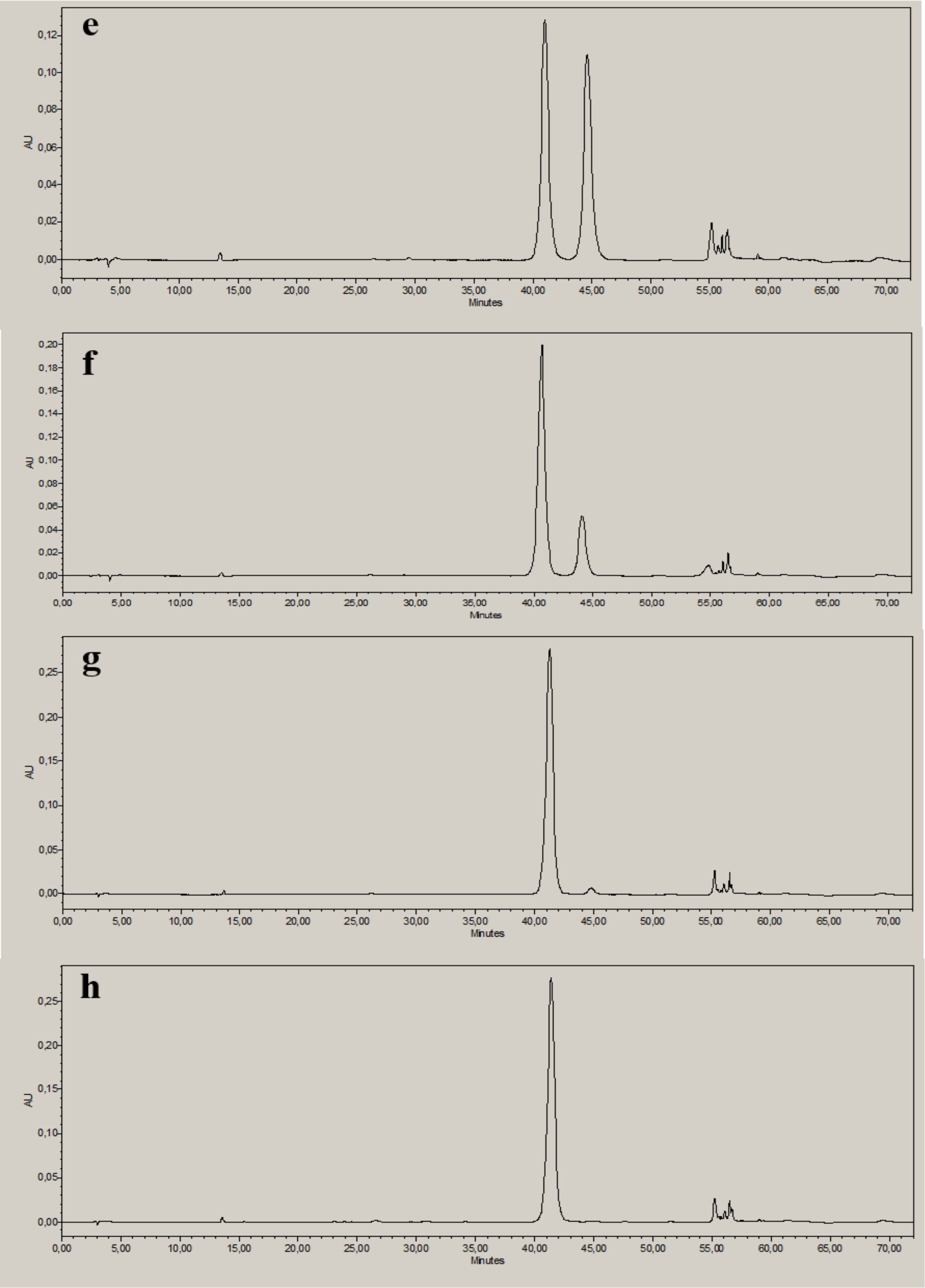
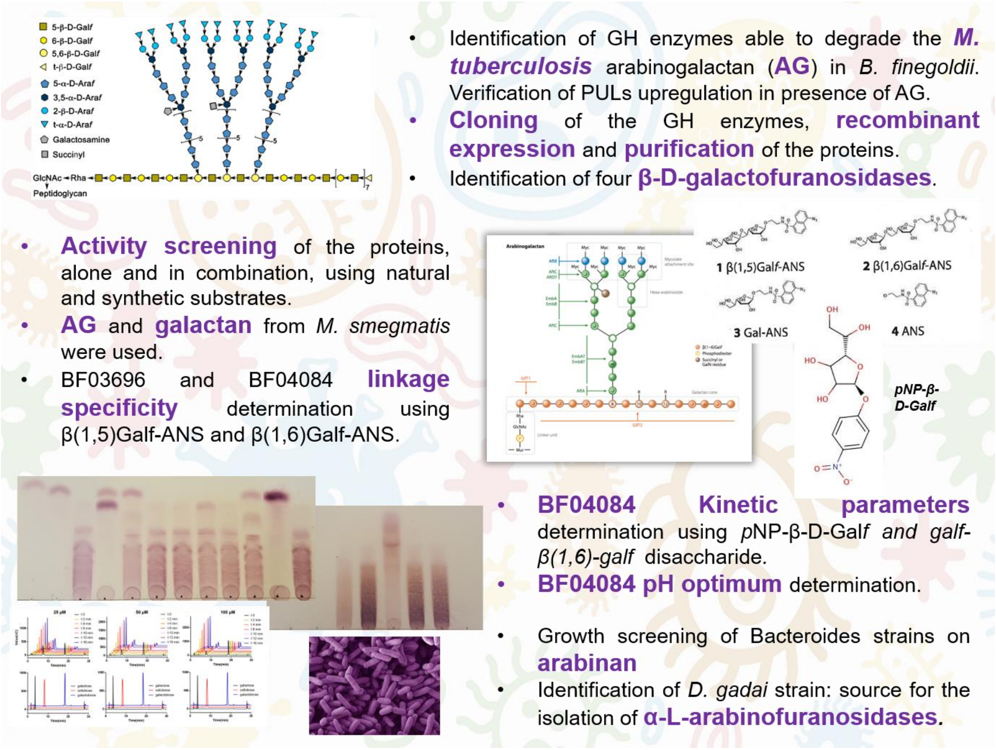


Figure 23: Whole cells time-course HPLC chromatograms at 330 nm. **a**: 2 mM naringin standard (black line) and 0.5 mM prunin (blue line). **b**: t 0 time course. **c**: t 5 minutes time-course. **d**: t 15 minutes time course. **e**: t 30 minutes time course. **f**: time 1 hour time course. **g**: t 1.5 hour time course. **h**: t 2 hours time course.

CHAPTER 4

Identification of novel glycosyl hydrolases from Bacteroides, active on *Mycobacterium tuberculosis* cell wall polysaccharides.


5-β-D-Gal*f*
6-β-D-Gal*f*
5,6-β-D-Gal*f*
t-β-D-Gal*f*
5-α-D-Ara*f*
3,5-α-D-Ara*f*
2-β-D-Ara*f*
t-α-D-Ara*f*
Galactosamine
Succinyl
GlcNAc
Rha
Peptidoglycan
5
7
Arabinogalactan
1 β(1,5)Gal*f*-ANS
2 β(1,6)Gal*f*-ANS
3 Gal-ANS
4 ANS
***p*NP-β-D-Galf**

- Identification of GH enzymes able to degrade the ***M. tuberculosis*** arabinogalactan (**AG**) in *B. finegoldii*. Verification of PULs upregulation in presence of AG.
- **Cloning** of the GH enzymes, **recombinant expression** and **purification** of the proteins.
- Identification of four **β-D-galactofuranosidases**.

- **Activity screening** of the proteins, alone and in combination, using natural and synthetic substrates.
- **AG** and **galactan** from *M. smegmatis* were used.
- BF03696 and BF04084 **linkage specificity** determination using β(1,5)Galf-ANS and β(1,6)Galf-ANS.

- **BF04084 Kinetic parameters** determination using *p*NP-β-D-Gal*f and galf-β(1,6)-galf* disaccharide.
- **BF04084 pH optimum** determination.

- Growth screening of Bacteroides strains on **arabinan**
- Identification of *D. gadai* strain: source for the isolation of **α-L-arabinofuranosidases**.

CHAPTER 5

Concluding remarks.

CHAPTER 5

Concluding Remarks.

The use of enzymes as biocatalysts is an invaluable tool for the biotechnological industry and has numerous advantages with respect to chemical synthesis procedures. The great pool of biodiversity and adaptation found in microbial ecosystems is considered as an invaluable natural reservoir to acquire novel useful biocatalysts.

Main goal of this PhD thesis was the identification and characterization of different microbial enzymatic activities, with the aim of evaluating their biotechnological potential in different bioconversion processes. The characterization of four extradiol ring cleavage dioxygenases (ERCDs) and of a bacterial α-L-rhamnosidase isolated from *Novosphingobium* sp. PP1Y was initially performed. In the last part of the project, gut microorganisms were used for the identification and characterization of novel glycosyl hydrolases able to degrade the arabinogalactan polysaccharide of the *M. tuberculosis* cell wall.

Novosphingobium sp. PP1Y, a member of the order of the Sphingomonadales, is endowed with the ability to grow using a wide array of fuels and hydrocarbons, including gasoil, gasoline and several polycyclic aromatic hydrocarbons and heterocyclic compounds, as unique carbon and energy source. Furthermore, the production of a complex cell matrix has been related [25] to a specific adaptation strategy of this strain as a reaction to the harsh environment from which it is isolated. The analysis of *N*. sp. PP1Y genome revealed several genomic features of great interest for the biotechnological potential of this microorganism. In fact, the presence of particular phenotypic traits is related to a unique abundance in *N*. sp. PP1Y genome of several highly specialized metabolic pathways involved in aromatic hydrocarbon degradation and to the presence of a remarkable number of glycosidase activities, possibly related to the complex array of extracellular carbohydrates produced in different growth conditions.

In the first part of the PhD project, *N*. sp. PP1Y was used for the isolation of novel ERCDs activities. Phylogenetic analysis of RCDs in Sphingomonads showed a great heterogeneity in the distribution of these proteins among the RCD subfamilies [63]. Analysis of the PP1Y genome revealed the presence of thirty-eight ORFs, which were predicted to code for thirty-four different aromatic hydroxylating dioxygenases. This particular set of RCDs might allow strain PP1Y to metabolize complex mixtures of catechols deriving from the simultaneous oxidation of several mono- and polycyclic-aromatic hydrocarbons, which are the preferred growth substrates for this strain [25]. Thus, analysis of the PP1Y genome revealed a broad reservoir of enzymes involved in the PAHs degradation [63]. The coding sequences identified resulted to possess homology with RCDs of closely related strains, *N. aromaticivorans* F199 and *N. pentaromativorans* US6-1. An alignment of these sequences with the ones annotated in related strains, allowed to identify four proteins (PP28735, PP26077, PP00124 and PP00193) with a high overall identity score (ranging between 70 – 95%) with biphenile dioxygenases, active towards the 2,3-DHBP, and (98%) extradiol catechol ring cleavage dioxygenase activities. Therefore, among the ERCDs retrieved in strain PP1Y, our attention was focused on these four enzymes, whose high identity scores with already characterized proteins allowed to perform a *homology modeling* analysis and to obtain the predicted 3D structures. Then, a *molecular docking* study was conducted, showing that these ERCDs could recognize as substrates catechols, dihydroxynaphtalene estrogens and dihydroxyphenantrene estrogens (*Cafaro V. et al, unpublished data*), which can be used as starting point for the synthesis of estrogen-like compounds such as 17-β-estradiol derivatives, hydroxyestradiols, and phytoestrogens analogues. The optimization of the recombinant expression and purification procedures of the four ERCDs from PP1Y, described in chapter 2, allowed obtaining the dioxygenases for their biochemical characterization. The activity screening of the enzymes with different catecholic substrates largely confirmed the evidences highlighted by the *molecular docking* analysis, showing that proteins PP00193 and PP00124, possessing the smaller active sites, are more active towards monoaromatic catechols, with PP00124 preferentially recognizing biphenyl compounds. Kinetic parameters were determined towards catechol, 3-methylcatechol, 4-methylcatechol and 2,3-DHBP. On the basis of our results the hypothesis can be advanced, that proteins PP28735 and PP26077 might have higher affinity for “bigger” substrates, such as polycyclic aromatic hydrocarbons having more than two aromatic rings.

Altogether, these features encourage a further characterization of these ERCDs from PP1Y, since these enzymes seem to be a valuable tool for the modification of complex hydroxylated heterocyclic aromatic compounds, which are a starting point in the production pipeline of many pharmacologically active molecules, such as steroid like molecules. In this framework, the site-specific cleavage and manipulation of aromatic substrates, which can be obtained with enzymatic biocatalysts, is of great advantage when compared to the complex mixture of products released in chemical modification procedures.

Finally, the bioconversion of catechol estrogens, which could be used as precursors in the production of different steroid families and hormones, was evaluated. Analysis of k_{cat}/K_M values highlighted that 4-OHE is a better substrate than monoaromatic catechols for PP28735 and PP00124 enzymes. Results obtained indicated that PP00124 is a potential biocatalyst to be used for the modification of catechol estrogens.

In conclusion, the bioprospecting activity of ERCDs enzymes here described allowed to identify promising biocatalysts to be used in bioconversion of a variety of mono and polycyclic aromatic compounds. The possibility to use these enzymes from cell free extracts was also investigated in this work, highlighting the possibility to avoid expensive and time-consuming enzyme purification procedures. Moreover, the enzymes activity screening pointed out towards a quite different substrate specificity of the ERCDs, compared to other enzymes already described in literature. In the near future, a further characterization and engineering of the catalytic residues in the active site of these enzymes, along with the determination of their crystal structure, will allow to identify ERCDs mutants with different catalytic properties and efficiency towards different substrates of biotechnological interest. Natural or engineered ERCDs might be used as catalysts for oxidative cleavage of catechol-like compounds and to prepare several types of heterocyclic rings, thus generating new families of estradiol-like molecules with high regioselectivity. In fact, the succession of steps of ring-cleavage, and recyclisation of dihydroxylated products is a general strategy for the conversion of terminal aromatic rings of polycyclic compounds into heterocyclic rings. For example, several steroid precursors could be obtained by using this same strategy from phytoestrogens [157]. Moreover, the availability of *N*. sp. PP1Y genome allows the identification, through a bioinformatic approach, of new set of promising enzymes to to use for the bioconversion of different aromatic molecules.

In the second part of the PhD project, the biochemical characterization of RHA-P, a bacterial α-L-rhamnosidase recently isolated from the microorganism *Novosphingobium* sp. PP1Y [116], was performed. This enzyme is an *inverting* GH, belonging to the GH106 family, for which an initial biochemical characterization has already been obtained [116] However, due to previous low purification yields, an optimization of recombinant expression and purification was carried out. A His-tag was added to the C-terminus of the recombinant protein expressed in *E. coli* BL21(DE3). Moreover, the purification of the protein from the periplasmic fractions of the induced cultures was implemented, based on the presence of the recombinant protein in the periplasmic space of induced cells. This is probably due to the presence in the protein of a signal peptide. In fact, RHA-P amino acid sequence, previously verified by MS mapping, showed the lack of the N-terminal peptide, suggesting the presence of a signal peptide presumably cleaved through a post-translational proteolytic processing. A similar evidence has been described for the α-RHA isolated from *S. paucimobilis* FP2001, recombinantly expressed in *E. coli* [138].

The combination of periplasm purification and affinity chromatography based on the C-terminal His-tag, led to an improved purification yield compared to the one obtained with the previous purification protocol. In addition, the analysis of the kinetic constants on *p*NPR suggested that the His-tag did not significantly alter the α-RHA activity of RHA-P on this synthetic substrate.

The *homology modelling* analysis of RHA-P using the structure of rhamnosidase BT0986 from *Bacteroides thetaiotaomicron* (pdb code, 5MWK) as template [115], suggested that the overall catalytic mechanism of RHA-P might be very similar to that of BT0986, but several details indirectly point to a quite different substrate specificity. The putative acid catalyst and all residues contacting rhamnose in the structure of BT0986, are conserved in RHA-P. In contrast, a deletion and two insertions in RHA-P are in the loops contributing to the active site pocket, suggesting the recognition of substrates with structures different from that of rhamnosidase BT0986.

The model obtained, along with the metal depletion and reconstitution experiments, confirmed the importance of the calcium ion stabilization for the catalytic mechanism of RHA-P. As already underlined, the presence of calcium ions as essential cofactors is reported in some GH families, although with different functions [158]. Among the few α-RHAs crystal structures available in literature, two examples of enzymes bearing calcium ions involved in catalysis, besides GH106 BT0986 are available, GH78 BsRhaB from *Bacillus* sp. GL1, and SaRha78A from *Streptomyces avermitilis* [112, 113]. However, differently from what observed in BT0986, in these proteins calcium is bound in a separate domain located near the catalytic domain and contacts the rhamnose residue in the catalytic domain when the substrate is bound in the active site [113]. In particular, Fujimoto Z. and coworkers described for SaRha78A the identification of a novel carbohydrate-binding module (CBM67) highly specific for rhamnose residues [109]. Carbohydrate binding modules (CBM) are defined as non-catalytic domains found in many carbohydrate-active enzymes, which contribute to the enzyme binding specificity towards different carbohydrates and polysaccharides [158]. In RHA-P, and in general in GH106, no separate CBM domains was ever observed, and calcium is probably directly bound in the active site of the enzyme. In BT0986, as well as in RHA-P, calcium seemed rather to play a key catalytic role, similar to the one reported from Zhu Y. et al for the GH92 and GH47 α-mannosidase [142], where calcium was hypothesized to be essential for the *inverting* mechanism of α-mannose residues hydrolysis [142]. The role of the calcium ion in the active site of RHA-P undoubtedly needs further investigation. Nonetheless, data collected so far support its involvement as an essential cofactor at the active site of the enzyme.

Finally, to investigate RHA-P substrate specificity, protein activity was tested on a range of poly- and oligosaccharidic compounds as well as natural flavonoids, highlighting that the enzyme was not active on any of the polysaccharidic compounds. From the homology model analysis, the hypothesis can be advanced that RHA-P active site evolved to be active on substrates containing aryl- groups and aromatic moieties, rather than poly- and oligosaccharidic compounds.

Subsequently, kinetic experiments were performed using natural flavonoids such as naringin, rutin, hesperidin and quercitrin. Results revealed a good enzymatic efficiency for these substrates compared to the one described for other α-RHAs in literature. Moreover, a high specificity of the enzyme towards substrates bearing aromatic moieties was highlighted. Although a slight difference in the substituents of the flavonoidic rings involved should be taken into account in the discussion of these results, several conclusions can be drawn from these experiments. Specifically, data obtained suggested that i) a rhamnose-containing disaccharidic unit is better accommodated in the active site compared to rhamnose alone; ii) when bound to the C7 of the A ring of the flavonoid, the neohesperidose moiety was hydrolyzed more efficiently than rutinose; iii) when rutinose is present as the disaccharidic moiety, the preferential position for its hydrolysis is the C3 of the C ring of the flavonoid.

Finally, in addition to the data presented in this work, previous analyses highlighted also a good thermostability of the enzyme up to 40°C and a moderate tolerance to organic solvents up to 20%, which further encourages the evaluation of RHA-P in bioconversion strategies for the hydrolysis of flavonoidic substrates [116].

In conclusion, RHA-P bioprospecting allowed identifying a novel α-RHA with a valuable kinetic efficiency compared to similar enzymes. Moreover, RHA-P showed a good stability and versatility in various conditions of temperature and pH, which is of utmost importance for biotechnological applications. The production potential and ease of purification of the enzyme was significantly improved in this work to a single-step purification procedure from periplasmic fraction of induced cultures. Moreover, in this work a preliminary structural characterization of the enzyme was obtained. In this context, the possibility to fine-tune RHA-P activity can be foreseen, after the identification of catalytic residues and determination of the active site morphology. The determination of the 3D structure of the enzyme is an important task to fulfil in the near future for better exploring the biotechnological potential of this bacterial rhamnosidase, and to obtain functional mutants.

Whole cell bioconversion experiments were performed, using naringin as substrate and *E. coli* BL21(DE3) cells expressing RHA-P as catalyst. Time course results showed that whole cells expressing RHA-P are able to catalyze the complete hydrolysis of naringin with concomitant release of the de-rhamnosylated product in the growth medium. In this context, the possibility to use efficient whole cells biocatalysts has numerous advantages in industrial bioconversion processes, allowing costs and overall number of processing steps reduction. Further experiments are needed to verify the efficiency of the whole-cell bioconversion system also on complex mixtures and natural products. Moreover, in order to apply flavonoids de-rhamnosylation safely to consumable food, it would be of interest to use GRAS microorganisms as biocatalysts. The heterologous epression of RHA-P into this group of bacteria and subsequent evaluation of whole-cell bioconversion would give an added value for the implement of the enzyme in food industry.

Finally, in the third section of this PhD thesis, results obtained during the visiting period at ICaMB (Institute for Cell and Molecular Biosciences), Newcastle University - Newcastle upon Tyne (UK), are described. The project was focused on the bioprospecting of novel GHs activities, from microorganisms of the human gut microbiome (HGM), able to degrade the arabinogalactan polysaccharide of *M. tuberculosis* cell wall. This research project is intended to identify novel suitable therapeutic agents for the treatment of tuberculosis disease.

HGM bacteria, especially the Bacteroidetes phylum, account for the major capability to degrade an extensive array of plant, animal and host complex carbohydrates. In this project, nine proteins were identified in two PULs from *B. finegoldii* as candidates for the degradation of the arabinogalactan (AG) complex from *M. tuberculosis* cell wall. The activity screening on AG, galactan and a range of different *p*NP-α- and *p*NP-β- sugars, confirmed that the proteins were only active towards galactofuranose residues, with high specificity, not being able to degrade the arabinan moiety of AG. Substrate specificity was further investigated for two of the proteins, BF03696 and BF04084, revealing that BF03696 was only able to hydrolyze β-1,5 glycosidic linkages with high specificity, being also able to act as an exo-GH, hydrolyzing only the terminal galactofuranose residue. Conversely, BF04084, also acting as an exo- GH, was actually able to hydrolyze both β-1,5 and β-1,6 linkages. Overall, these data showed that proteins BF03696 and BF04084 could be promising candidates for the hydrolysis of galactan polysaccharides from *M. tuberculosis* cell wall, being able to completely degrade the polysaccharide when used in combination.

Moreover, the characterization of these enzymes is of interest since not many galactofuranosidases activities have been described in literature so far, and only one crystal structure was described and deposited in PDB. In addition, BF04084 was not listed in any of the known GH families in CAZy database, suggesting that probably the protein possessed a peculiar enzyme fold and active site organization. 3D structure determination will shed light on the catalytic mechanism and substrate specificity of BF04084, thus allowing to confirm if BF04084 is effectively a member of a new GH family.

Kinetic parameters on *p*NP-β-D-Gal*f* were determined, highlighting a higher affinity for this substrate as compared to other enzymes [155,156]. Finally, a k_{cat}/K_M approximately 100 fold higher than the one obtained for *p*NP-β-D-Gal*f* was measured for the gal*f*-β(1,6)-gal*f* disaccharide of galactan, showing a higher efficiency towards oligosaccharidic substrates. A further characterization of the enzyme activity using galactan and other galactan-derived oligosaccharides is required to get more insights in BF04084 kinetic behavior.

These results led to the identification and partial characterization of BF03696 and BF04084 GH enzymes as good candidates for the degradation of galactan backbone of AG. These enzymes exhibited a good stability and a single step purification procedure was set up. The high specificity found for these enzymes towards galactofuranose would be of interest for a future application in the biotechnological modification of these sugar residues.

Finally, in the last part of the work, an experiment was carried out to identify others HGM microorganisms valuable to serve as source of arabinofuranosidase activities active towards the arabinan moiety of the mycobacteria AG. *D. gadai* strain was identified, which was able to hydrolyze the arabinan molecule with arabinose release. Therefore, the presence in this strain of arabinofuranosidases activity was hypothesized. An RNA-sequencing of this strain grown on arabinan would allow identifying novel GH enzymes able to degrade the arabinan moiety of mycobacteria AG.

Future goals of the project will be the identification of arabinofuranosidases activities, and the evaluation of the ability of the galactofuranosidases and arabinofuranosidases enzymes to act in combination for the complete degradation of the AG polysaccharide in vitro. Then, the ability of these GH enzymes to inhibit the growth of *Mycobacterium tuberculosis in vivo* will be evaluated.

In conclusion, in this thesis the bioprospecting of different enzymes from *Novosphingobium* sp. PP1Y and *Bacteroides finegoldii* is described. In the first part of the work, the recombinant expression, purification and preliminary biochemical characterization was obtained for four ERCDs and an α-RHA from *Novosphingobium* sp. PP1Y. These activities represent valuable biocatalysts in industry and are part of still poorly characterized enzymatic families. For all enzymes a substrate specificity study was carried out, and kinetic parameters were obtained. Moreover, the evaluation of the use ERCDs for the bioconversion of extradiolic compounds, and of RHA-P for the transformation of natural flavonoids is described. Results highlighted valuable features of the enzymes in terms of enzymatic efficiency and specificity, compared to other enzymes described in literature. The determination of the 3D structures, identification of catalytic residues and generation of site-specific mutants to fine-tune the enzymes substrate specificity, are among the future goals of this work. A further evaluation of whole-cell bioconversion strategies might be considered for RHA-P.

The last part of the work focused on the identification and characterization of new GH enzymes from the HGM microorganism *Bacteroides finegoldii*, which is able to degrade the AG complex polysaccharide from the *Mycobacterium tuberculosis* cell wall. The work allowed identifying different enzymes with β-D-galactofuranosidase activity, also in this case a class of enzymes still poorly characterized. In particular, two of the identified GHs were able to completely degrade the galactan backbone of the AG polysaccharide, when used in combination. Moreover, another HGM microorganism strain probably possessing α-L-arabinofuranosidase enzymes was identified as a new source of GH enzymes able to hydrolyze also the arabinan backbone.

A preliminary biochemical characterization and substrate specificity study was performed for the two proteins BF03696 and BF04084. In particular, protein BF04084 was an interesting biocatalyst and, according to our data, seemed to belong to a new GH family in the CAZy database.

REFERENCES

1 - Luen-Luen L., McCorkle S. R., Monchy S., Taghavi S., van der Lelie D., *Bioprospecting metagenomes: glycosyl hydrolases for converting biomass*. Biotechnology for Biofuels 2009, 2:10.

2 – Ferrer M., Martínez-Martínez M., Bargiela R., Streit W.R., Golyshina O.V., Golyshin P.N., *Estimating the success of enzyme bioprospecting through metagenomics: current status and future trends*. Microb Biotechnol. 2016, 9(1): 22–34.

3 - Liebl W., Angelov A., Juergensen J., Chow J., Loeschcke A., Drepper T., *Alternative hosts for functional (meta)genome analysis*. Appl Microbiol Biotechnol. 2014, 98: 8099–8109.

4 – Sheldon R.A., Woodley J.M., *Role of Biocatalysis in Sustainable Chemistry*. Chem. Rev. 2018, 118, 801−838.

5 – Kyrpides N.C., Hugenholtz P., Eisen J.A., Woyke T., Göker M., Parker C.T., *Genomic encyclopedia of bacteria and archaea: sequencing a myriad of type strains*. PLoS Biol. 2014, 12: e1001920.

6 - Yarza P., Yilmaz P., Pruesse E., Glöckner F.O., Ludwig W., Schleifer K.H., *Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences*. Nat Rev Microbiol. 2014, 12: 635–645.

7 – Gurung N., Ray S., Bose S., Rai V., *A Broader View: Microbial Enzymes and Their Relevance in Industries, Medicine, and Beyond*. BioMed Research International. 2013, Article ID 329121.

8 – Prakash B., Vidyasagar M., Madhukumar M.S., Muralikrishna G., Sreeramulu K., *Production, purification, and characterization of two extremely halotolerant, thermostable, and alkali-stable α-amylases from Chromohalobacter sp. TVSP 101*. Process Biochemistry. 2009, 44 (2), 210–215.

9 – Singh R., Kumar M., Mittal A., Kumar Mehta P., *Microbial enzymes: industrial progress in 21st century*. Biotech. 2016, 6(2): 174.

10 - Sheldon R.A., *Atom utilisation, E factors and the catalytic solution*. C. R. Acad. Sci. Ser. IIc: Chim. 2000, 3, 541−551.

11 - Straathof A.J.J., Panke S., Schmid A., *The production of fine chemicals by biotransformations*. Current Opinion in Biotechnology. 2002, 13:548-556.

12 - Chen Y., Nielsen J., *Biobased organic acids production by metabolically engineered microorganisms*. Curr. Opin. Biotechnol. 2016, 37, 165−172.

13 – Underkofler L.A., Barton R.R., Rennert S.S., *Production of microbial enzymes and their applgications*. Applied Microbiology. 1957, 6 (3) 212–221.

14 – Deutch C.E., *Characterization of a salt-tolerant extracellular α-amylase from Bacillus dipsosauri*. Letters in Applied Microbiology. 2002, 35 (1) 78–84.

15 – Hube B., Stehr F., Bossenz M., Mazur A., Kretschmar M., Schafer W., *Secreted lipases of Candida albicans: cloning, characterisation and expression analysis of a new gene family with at least ten members*. Archives of Microbiology. 2000, 174 (5) 362–374.

16 – Cupler E.J., Berger K.I., Leshner R.T. *Consensus treatment recommendations for late-onset pompe disease*. Muscle and Nerve. 2012, 45 (3) 319–333.

17 – Ensor C.M., Holtsberg F.W., Bomalaski J.S., Clark M.A., *Pegylated arginine deiminase (ADI-SS PEG20,000 mw) inhibits human melanomas and hepatocellular carcinomas in vitro and in vivo*. Cancer Research. 2002, 62 (19) 5443–5450.

18 – Kell D.B., Swainston N., Pir P., Oliver S.G., *Membrane transporter engineering in industrial biotechnology and whole cell biocatalysis*. Trends Biotechnol. 2015, 33, 237−246.

19 - Bommarius A.S., Paye M.F., *Stabilizing biocatalysts*. Chem. Soc. Rev. 2013, 42, 6534−6565.

20 - Bruggink A., Schoevaart R., Kieboom T., *Concepts of Nature in Organic Synthesis: Cascade Catalysis and Multistep Conversions in Concert*. Org. Process Res. Dev. 2003, 7, 622−640.

21 - Schoevaart R., van Rantwijk F., Sheldon R.A.A., *Four-Step Enzymatic Cascade for the One-Pot Synthesis of Non-natural Carbohydrates from Glycerol*. J. Org. Chem. 2000, 65, 6940−6943.

22 - Nielsen J., Keasling J. *Engineering Cellular Metabolism*. Cell. 2016, 164, 1185−1197.

23 - Balkwill D.L., *et al*., *Taxonomic study of aromatic-degrading bacteria from deep-terrestrial-subsurface sediments and description of Sphingomonas aromaticivorans sp. nov., Sphingomonas subterranea sp. nov., and Sphingomonas stygia sp. nov*. Int J Syst Bacteriol. 1997, 47 (1) 191-201.

24 - Timmis K.N., Pieper D.H., *Bacteria designed for bioremediation*. Trends Biotechnol. 1999, 17, 200-204.

25 - Notomista E., Pennacchio F., Cafaro V., Smaldone G., Izzo V., Troncone L., Varcamonti M., Di Donato A., *The Marine Isolate Novosphingobium sp. PP1Y Shows Specific Adaptation to Use the Aromatic Fraction of Fuels as the Sole Carbon and Energy Source*. MicrobEcol. 2011, 61(3) 582-94.

26 - Urlacher V.B., Schmid R.D., *Recent advances in oxygenase-catalyzed biotransformations*. Curr Opin Chem Biol. 2006, 10, 156–161.

27 - Cirino P.C., Arnold F.H., *Protein engineering of oxygenases for biocatalysis*. Curr Opin Chem Biol. 2002, 6, 130–135.

28 - Torres Pazmiño D.E., Winkler M., Glieder A., Fraaije M.W., *Monooxygenases as biocatalysts: classification, mechanistic aspects and biotechnological applications*. J Biotechnol. 2010, 146, 9–24.

29 - Xu F., *Applications of oxidoreductases: Recent progress*. Industrial Biotechnology. 2005, 1, 38–50.

30 - Arenghi F.L., *et al*., *Organization and regulation of meta cleavage pathway genes for toluene and o-xylene derivative degradation in Pseudomonas stutzeri OX1*. Appl Environ Microbiol. 2001, 67 (7) 3304-8.

31 – Pinyakong O., Habe H., Omori T., *The unique aromatic catabolic genes in sphingomonads degrading polycyclic aromatic hydrocarbons (PAHs)*. J Gen Appl Microbiol. 2003, 49, 1-19.

32 - Harayama S., Rekik M., *Comparison of the nucleotide sequences of the meta-cleavage pathway genes of TOL plasmid pWW0 from Pseudomonas putida with other meta-cleavage genes suggests that both single and multiple nucleotide substitutions contribute to enzyme evolution*. Mol Gen Genet. 1993, 239, 81-89.

33 - Harayama S., Rekik M*., Bacterial aromatic ring-cleavage enzymes are classified into two different gene families*. J Biol Chem. 1989, 264 (7), 15328-15333.

34 - Cafaro V., *et a*l., *Phenol hydroxylase and toluene/o-xylene monooxygenase from Pseudomonas stutzeri OX1: interplay between two enzymes*. Appl Environ Microbiol. 2004, 70 (4), 2211-9.

35 – Jindrová E., Chocová M., Demnerová K., Brenner V., *Bacterial aerobic degradation of benzene, toluene, ethylbenzene and xylene*. Folia Microbiol (Praha). 2002, 47, 83-93.

36 - Viggiani A., *et al*., *The role of the conserved residues His-246, His-199, and Tyr-255 in the catalysis of catechol 2,3-dioxygenase from Pseudomonas stutzeri OX1*. J Biol Chem. 2004, 279 (47) 48630-9.

37 - Lipscomb J.D., Orville A.M., *Mechanistic aspects of dihydroxybenzoate dioxygenases*. Metal Ions Biol Syst. 1992, 28, 243–298.

38 - Vaillancourt F.H., Bolin J.T., Eltis L.D., *The ins and outs of ring-cleaving dioxygenases*. Crit Rev Biochem Mol Biol. 2006, 41, 241–267.

39 – Lipscomb J.D., *Mechanism of extradiol aromatic ring-cleaving dioxygenases*. Curr Opin Struct Biol. 2008, 18 (6), 644–649.

40 – Fetzner S., *Ring-Cleaving Dioxygenases with a Cupin Fold*. Appl. Environ. Microbiol. 2012, 78 (8), 2505–2514.

41 - Parales R.E., Ditty J.L., *Laboratory evolution of catabolic enzymes and pathways*. Curr Opin Biotechnol. 2005, 16, 315–325.

42 - Nolan L., O'Connor K., *Dioxygenase- and monooxygenase-catalysed synthesis of cis -dihydrodiols, catechols, epoxides and other oxygenated products*. Biotechnol Lett. 2008, 30, 1879–1891.

43 - Di Gennaro P., Bargna A., Sello G., *Microbial enzymes for aromatic compound hydroxylation*. Appl Microbiol Biotechnol. 2011, 90, 1817–1827.

44 - Bühler B., Schmid A., *Process implementation aspects for biocatalytic hydrocarbon oxyfunctionalization*. J Biotechnol. 2004, 113, 183–210.

45 - Beilen J.B., Funhoff E.G., *Expanding the alkane oxygenase toolbox: new enzymes and applications*. Curr Opin Biotechnol. 2005, 16, 308–314.

46 - Drora A., Fishmana A., *Engineering Non-Heme Mono- and Dioxygenases for Biocatalysis*. Comput Struct Biotechnol J. 2012, 2 (3) e201209011.

47 - Schmid A., Hofstetter K., Feiten H.J., Hollmann F., Witholt B., *Integrated biocatalytic synthesis on gram scale: the highly enantioselective preparation of chiral oxiranes with styrene monooxygenase*. Adv Synth Catal. 2001, 343, 732–737.

48 - Luetz S., Giver L., Lalonde J., *Engineered enzymes for chemical production*. Biotechnol Bioeng. 2008, 101, 647–653.

49 - Kim D., Choi K., Yoo M., Choi J., Lee C., *et a*l., *Benzylic and aryl hydroxylations of m-xylene by o-xylene dioxygenase from Rhodococcus sp. strain DK17*. Appl Microbiol Biotechnol. 2010, 86, 1841–1847.

50 - Ang E., Obbard J., Zhao H., *Directed evolution of aniline dioxygenase for enhanced bioremediation of aromatic amines*. Appl Microbiol Biotechnol. 2009, 81, 1063–1070.

51 - Vézina J., Barriault D., Sylvestre M., *Family shuffling of soil DNA to change the regiospecificity of Burkholderia xenovorans LB400 biphenyl dioxygenase*. J Bacteriol. 2007, 189, 779–788.

52 - Caglio R., Valetti F., Caposio P., Gribaudo G., Pessione E., *et al*., *Fine-tuning of catalytic properties of catechol 1,2-dioxygenase by active site tailoring*. ChemBioChem. 2009, 10, 1015–1024.

53 - Wei J., Zhou Y., Xu T., Lu B., *Rational design of catechol-2, 3-dioxygenase for improving the enzyme characteristics*. Appl Biochem Biotechnol. 2010, 162, 116–126.

54 - Dong Y., Yan J., Du H., Chen M., Ma T., *et al*., *Engineering of LadA for enhanced hexadecane oxidation using random- and site-directed mutagenesis*. Appl Microbiol Biotechnol. 2012, 94, 1019–1029.

55 - Yoo M., Kim D., Zylstra G.J., Kang B.S., Kim E., *Biphenyl hydroxylation enhanced by an engineered o-xylene dioxygenase from Rhodococcus sp. strain DK17*. Res Microbiol. 2011, 162, 724–728.

56 – Fishman J., Norton B., *Catechol estrogen formation in the central nervous system of the rat*. Endocrinology. 1975, 96 (4), 1054-8.

57 - Lane K.E., Ricci M.J., Ho S.M., *Effect of combined testosterone and estradiol-17 beta treatment on the metabolism of E2 in the prostate and liver of noble rats*. Prostate. 1997, 30 (4) 256-62.

58 - Ball P., Knuppen R., *Formation of 2- and 4-hydroxyestrogens by brain, pituitary, and liver of the human fetus*. J Clin Endocrinol Metab. 1978, 47 (4) 732-7.

59 - Schutze N., Vollmer G., Tiemann I., Geiger M., Knuppen R., *Catecholestrogens are MCF-7 cell estrogen receptor agonists*. J Steroid Biochem Mol Biol. 1993, 46 (6) 781-9.

60 - Hermenegildo C., Oviedo P., Cano J., *Cyclooxygenases regulation by estradiol on endothelium*. Curr Pharm Des. 2006, 12, 205-15.

61 - Bilezikian J.P., *Anabolic therapy for osteoporosis*. Int J Fertil Womens Med. 2005, 50, 53-60.

62 - Oelkers W.H., *Drospirenone in combination with estrogens: for contraception and hormone replacement therapy*. Climacteric. 2005, 3, 19-27.

63 - D'Argenio V., Notomista E., Petrillo M., Cantiello P., Cafaro V., Izzo V., Naso B., Cozzuto L., Durante L., Troncone L., Paolella G., Salvatore F., Di Donato A., *Complete sequencing of Novosphingobium sp. PP1Y reveals a biotechnologically meaningful metabolic pattern*. BMC Genomics. 2014, 15, 384-397.

64 - Gantt R.W., Peltier-Pain P., Cournoyer W.J., Thorson J.S., *Using simple donors to drive the equilibria of glycosyltransferase-catalyzed reactions*. Nature Chemical Biology. 2011, 7, 685–691.

65 - Koshland, D.E., *Stereochemistry and the Mechanism of Enzymatic Reactions*. Biological Reviews. 1953, 28, 416-436.

66 - McCarter J.D., Withers S.G., *Mechanisms of enzymatic glycoside hydrolysis*. Curr Opin Struct Biol. 1994, 4(6), 885-92.

67 - Lombard V., Golaconda Ramulu H., Drula E., Coutinho P.M., Henrissat B., *The carbohydrate-active enzymes database (CAZy) in 2013*, Nucleic Acids Res. 2014, 42, D490–D495.

68 - Flint H.J., Scott K.P., Duncan S.H., Louis P., Forano E., *Microbial degradation of complex carbohydrates in the gut*. Gut Microbes. 2012, 3 (4), 289–306.

69 - Xu J., Bjursell M.K., Himrod J., Deng S., Carmichael L.K., Chiang H.C., *et al*. *A genomic view of the human-Bacteroides thetaiotaomicron symbiosis*. Science. 2003, 299, 2074–6.

70 - Glenwright A.J., Pothula K.R., Bhamidimarri S.P., Chorev D.S., Baslé A., Firbank S.J., Zheng H., Robinson C.V., Winterhalter M., Kleinekathöfer U., Bolam D.N., van den Berg B., *Structural basis for nutrient acquisition by dominant members of the human gut microbiota*. Nature. 2017, 541, 407–411.

71 - Cuskin F., Lowe E.C., Temple M.J., Zhu Y., Cameron E.A., Pudlo N.A., Porter N.T., Urs K., Thompson A.J., Cartmell A., Rogowski A., Hamilton B.S., Chen R., Tolbert T.J., Piens K., *et al*., *Human gut Bacteroidetes can utilize yeast mannan through a selfish mechanism*. Nature. 2015, 517, 165–169.

72 - Lammerts van Bueren A., Saraf A., Martens E.C., Dijkhuizen L., *Differential metabolism of exopolysaccharides from probiotic lactobacilli by the human gut symbiont Bacteroides thetaiotaomicron*. Appl. Environ. Microbiol. 2015, 81, 3973–3983.

73 - Temple M.J., Cuskin F., Baslé A., Hickey N., Speciale G., Williams S.J., Gilbert H.J., Lowe E.C., *A Bacteroidetes locus dedicated to fungal 1,6-β-glucan degradation: unique substrate conformation drives specificity of the key endo-1,6-β-glucanase*. J Biol Chem. 2017, 292 (25), 10639–10650.

74 - 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. 2015, 13 (10), 620–630.

75 - Mikusova´ K., Slayden R.A., Besra G.S., Brennan P.J., *Biogenesis of themycobacterial cellwall and the site of action of ethambutol*. Antimicrob Agents Chemother. 1995, 39, 2484–2489.

76 - Centers for Disease Control and Prevention (CDC), *Emergence of Mycobacterium tuberculosis with extensive resistance to second-line drugs--worldwide, 2000-2004*. MMWR Morb Mortal Wkly Rep. 2006, 55 (11), 301-5.

77 - Ma Y., Stern R.J., Scherman M.S., Vissa V.D., Yan W., Cox Jones V., Zhang F., Franzblau S.G., Lewis W.H., McNeil M.R., *Drug Targeting Mycobacterium tuberculosis Cell Wall Synthesis: Genetics of dTDP-Rhamnose Synthetic Enzymes and Development of a Microtiter Plate-Based Screen for Inhibitors of Conversion of dTDP-Glucose to dTDP-Rhamnose*. Antimicrob Agents Chemother. 2001, 45 (5), 407–1416.

78 - Hui L.W.G., Jiao X., Sun Y., Sun S., Feng Z., Zhou W., Zhu H., *The preparation and characterization of a novel sphingan WL from marine Sphingomonas sp*. Scientific Reports. 2016, 6, 37899.

79 - Miller T.R., *et al*. *Genome sequence of the dioxin-mineralizing bacterium Sphingomonas wittichii RW1*. J Bacteriol. 2010, 192 (22), 6101-2.

80 - Copley S.D., Rokicki J., Turner P., Daligault H., Nolan M., *The Whole Genome Sequence of Sphingobium chlorophenolicum L-1: Insights into the Evolution of the Pentachlorophenol Degradation Pathway*. Genome Biol Evol. 2012, 4 (2), 184-198.

81 - Ito S., Kobayashi T., Ara K., Ozaki K., Kawai S., Hatada Y., *Alkaline detergent enzymes from alkaliphiles: enzymatic properties, genetics, and structures*. Extremophiles. 1998, 2 (3), 185-90.

82 - Kulkarni N., Shendye A., Rao M., *Molecular and biotechnological aspects of xylanases*. FEMS Microbiol Rev. 1999, 23 (4), 411-56.

83 - James J.A., Lee B.H., *Glucoamylases: Microbial sources, industrial applications and molecular biology*. Journal of Food Biochemistry. 1997, 21 1-52.

84 - Kashyap D.R., Vohra P.K., Chopra S., *et al*. *Applications of pectinases in the commercial sector: a review*. Bioresource Technology. 2001, 77, 215-227.

85 - Bayer E.A., Lamed R., Himmel M.E., *The potential of cellulases and cellulosomes for cellulosic waste management*. Curr Opin Biotechnol. 2007, 18 (3), 237-45.

86 - Van Der Borght J., Desmet T., Soetaert W., *Enzymatic production of β-D-glucose-1-phosphate from trehalose*. Biotechnology Journal. 2010, 5 (9), 986.

87 – Trincone A., Giordano A., *Glycosyl hydrolases and glycosyltransferases in the synthesis of oligosaccharides*. Curr. Org. Chem. 2006, 10, 1163-1193.

88 - Cobucci-Ponzano B., Moracci M., *Glycosynthases as tools for the production of glycan analogs of natural products*. Nat Prod Rep. 2012, 29 (6), 697-709.

89 - Liu Q.P., Sulzenbacher G., Yuan H., Bennett E.P., Pietz G., Saunders K., Spence J., Nudelman E., Levery S.B., White T., Neveu J.M., Lane W.S., Bourne Y., Olsson

M.L., Henrissat B., Clausen H., *Bacterial glycosidases for the production of universal red blood cells*. Nat Biotechnol. 2007, 25 (4), 454-64.

90 - Rémond C., Plantier-Royon R., Aubry N., O'Donohue M.J., *An original chemoenzymatic route for the synthesis of β-d-galactofuranosides using an α-l-arabinofuranosidase*. Carbohydr. Res. 2005, 340, 637–644.

91 – Gabius S., Kayser K., Bovin N.V., Yamazaki N., Kojima S., Kaltner H., Gabius, H.-J. *Endogenous lectins and neoglycoconjugates: a sweet approach to tumor diagnosis and targeted drug delivery*. Eur. J. Pharm. Biopharm. 1996, 42, 250-261.

92 Beljaars L., Olinga P., Molema G., de Bleser P., Geerts A., Groothuis G.M.M., Poelstra, K., *Characteristics of the hepatic stellate cell-selective carrier mannose 6-phosphate modified albumin*. Liver. 2001, 21 (5), 320-328.

93 - Bagshawe, K.D., *Antibody-directed enzyme prodrug therapy (Adept)*. J. Controlled Release. 1994, 28, 187-193.

94 - Robinson M.A., Charlton S.T., Garnier P., Wang X., Davis S.S., Perkins A.C., Frier M., Duncan R., Savage T.J., Wyatt D.A., Watson S.A., Davis B.J., *LEAPT: Lectin-directed enzyme-activated prodrug therapy*. Proc Natl Acad Sci U S A. 2004, 101 (40), 14527–14532.

95 – Cerqueira N., Brás N., Ramos M.J. Fernandes P.A., *Glycosidases – A Mechanistic Overview*. Carbohydrates. Intech. 2012.

96 - van der Ploeg A.T., Reuser A.J., *Pompe's disease*. Lancet. 2008, 372, 1342– 1353.

97 - Henrissat B., *A classification of glycosyl hydrolases based on amino-acid sequence similarities*. Biochem. J. 1991, 280, 309–316.

98 - Manzanares P., Vallés S., Ramón D., Orejas M., Industrial Enzymes. (Polaina J, MacCabe AP, eds Springer). *α-L-Rhamnosidases: Old and New Insights*, Industrial Enzymes. Structure, Function and Applications. 2007, 117–140.

99 Nereus W., Gunther I.V., Nun˜ez A., Fett W., Solaiman D.K.Y., *Production of rhamnolipids by Pseudomonas chlororaphis, a nonpathogenic bacterium*. Appl.Environ. Microb. 2005, 71 (5), 2288–2293.

100 - Mutter M., Beldman G., Schols H.A., Voragen A.G.J., *Rhamnogalacturonan α-l-rhamnopyranohydrolase. A novel enzyme specific for the terminalnonreducing rhamnosyl unit in rhamnogalacturonan regions of pectin.* PlantPhysiol. 1994, 106 (1), 241–250.

101 - Panche A.N., Diwan A.D., Chandra S.R., *Flavonoids: an overview*. J Nutr Sci. 2016, 5 e 47.

102 - Bokkenheuser V.D., Shackleton C.H., Winter J., *Hydrolysis of dietary flavonoid glycosides by strains of intestinal Bacteroides from humans*. Biochem J. 1987, 248 (3), 953–956.

103 - Hollman P.C., Bijsman M.N., van Gameren Y., Cnossen E.P., de Vries J.H.,Katan M.B., *The sugar moiety is a major determinant of the absorption of dietary flavonoid glycosides in man*. Free Radic. Res. 1999, 31 (6), 569-573.

104 - González-Barrio R., Trindade L.M., Manzanares P., de Graaff L.H., Tomás-Barberán F.A., Espín J.C., *Production of bioavailable flavonoid glucosides in fruit juices and green tea by use of fungal alpha-L-rhamnosidases*. J. Agric. Food Chem. 2004, 52 (20), 6136-6142.

105 - Terada Y., Kometani T., Nishimura T., Takii H., Okada S., *Prevention of hesperidin crystal formation in canned mandarin orange syrup and clarified orange juice by hesperidin glycosides*. Food Sci Technol. Int. 1995, 1 (1), 29-33.

106 – Manzanares P., Orejas M., Gil V., de Graaff L.H., Visser J.,Ramón D., *Construction of a Genetically Modified Wine Yeast Strain Expressing the Aspergillus aculeatus rhaA Gene, Encoding an α-L-Rhamnosidase of Enological Interest*. Appl. Environ. Microbiol. 2003, 69 (12), 7558–7562.
107 - Xu L., Liu X., Yin Z., Liu Q., Lu L., Xiao M., *Site-directed mutagenesis of α-L-rhamnosidase from Alternaria sp. L1 to enhance synthesis yield of reverse hydrolysis based on rational design*. Appl Microbiol Biotechnol. 2016, 100 (24), 10385-10394.
108 - Li L., Yu Y., Zhang X., Jiang Z., Zhu Y., Xiao A., Ni H., Chen F., *Expression and biochemical characterization of recombinant α-L-rhamnosidase r-Rha1 from Aspergillus niger JMU-TS528*. Int J Biol Macromol. 2016, 85, 391-399.
109 - Koseki T., Mese Y., Nishibori N., Masaki K., Fujii T., Handa T., Yamane Y., Shiono Y., Murayama T., Iefuji H., *Characterization of an alpha-L-rhamnosidase from Aspergillus kawachii and its gene*. Appl. Microbiol. Biotechnol. 2008, 80 (6), 1007-1013.
110 - Puri M., *Updates on naringinase: structural and biotechnological aspects*. Appl. Microbiol. Biot. 2012, 93 (1), 49-60.
111 - Bonanno J.B., Almo S.C., Bresnick A., Chance M.R., Fiser A., Swaminathan S., Jiang J., Studier F.W., Shapiro L., Lima C.D., Gaasterland T.M., Sali A., Bain K., Feil I., Gao X., Lorimer D., Ramos A., Sauder J.M., Wasserman S.R., Emtage S., D'Amico K.L., Burley S.K., *New York-Structural GenomiX Research Consortium (NYSGXRC): a large scale center for the protein structure initiative*. J Struct Funct Genomics. 2005, 6 (2-3), 225-232.
112 - Cui Z., Maruyama Y., Mikami B., Hashimoto W., Murata K., *Crystal structure of glycoside hydrolase family 78 α-L-rhamnosidase from Bacillus sp. GL1*. J. Mol. Biol. 2007, 374 (2), 384–398.
113 - Fujimoto Z., Jackson A., Michikawa M., Maehara T., Momma M., Henrissat B., Gilbert H.J., Kaneko S., *The structure of a Streptomyces avermitilis α-L-rhamnosidase reveals a novel carbohydrate-binding module CBM67 within the six-domain arrangement*. J. Biol. Chem. 2013, 288 (17), 12376-12385.
114 - O'Neill E.C., Stevenson C.E., Paterson M.J., Rejzek M., Chauvin A.L., Lawson D.M., Field R.A., *Crystal structure of a novel two domain GH78 family α-rhamnosidase from Klebsiella oxytoca with rhamnose bound*. Proteins. 2015, 83 (9), 1742-1749.
115 - Ndeh D., Rogowski A., Cartmell A., Luis A.S., Baslé A., Gray J., Venditto I., Briggs J., Zhang X., Labourel X., Terrapon N., Buffetto F., Nepogodiev S., Xiao Y., Field R.A., Zhu Y., O'Neill M.A., Urbanowicz B.R., York W.S., Davies G.J., Abbott D.W., Ralet M.C., Martens E.C., Henrissat B., Gilbert H.J., *Complex pectin metabolism by gut bacteria reveals novel catalytic functions*. Nature. 2017, 544 (7648), 65-70.
116 - De Lise F., Mensitieri F., Tarallo V., Ventimiglia N., Vinciguerra R., Tramice A., Marchetti R., Pizzo E., Notomista E., Cafaro V., Molinaro A., Birolo L., Di Donato A., Izzo V., *RHA-P: isolation, expression and characterization of a bacterial α-L-rhamnosidase from Novosphingobium sp. PP1Y*. J. Mol. Catal. B: Enzym. 2016, 134, 136–147.
117 - Zhou Q.Z.K., Chen X.D., *Effect of temperature and pH on the catalytic activity of the immobilized β galactosidase from Kluyveromyces lactis*. Biochem. Eng. 2001, 9: 33–40.
118 - Tello-Solís S.R., Jiménez-Guzmán J., Sarabia-Leos C., Gómez-Ruíz L., CruzGuerrero A.E., Rodríguez-Serrano G.M., García-Garibay M., *Determination of the secondary structure of Kluyveromyces lactis betagalactosidase by circular dichroism*

and its structure-activity relationship as a function of the pH. J. Agric. Food Chem. 2005, 53: 10200–10204.
119 - Numanoglu Y., Sungur S., *β galactosidase from Kluyveromyces lactis cell disruption and enzyme immobilization using a cellulose-gelatin carrier system*. Proc. Biochem. 2004, 39: 703–709.
120 - Genari A.N., Passos F.V., Passos F.M., *Configuration of a bioreactor for milk lactose hydrolysis*. J. Dairy Sci. 2003, 86: 2783–2789.
121 - Picard C., Fioramonti J., Francois A., Robinson T., Neant F., Matuchansky C., *Review article: bifidobacteria as probiotic agents -- physiological effects and clinical benefits*. Aliment Pharmacol. Ther. 2005, 22: 495–512.
122 - Grosova Z., Rosenberg M., Rebros M., *Perspectives and applications of immobilized β galactosidase in food industry: a review*. Czech J. Food Sci. 2008, 26: 1–14.
123 - Sitanggang A.B., Drews A., Kraume M., *Recent advances on prebiotic lactulose production*. World J Microbiol. Biotechnol. 2006, 32:154.
124 - Guimarães P.M., Teixeira J.A., Domingues L., *Fermentation of lactose to bio-ethanol by yeasts as part of integrated solutions for the valorisation of cheese whey*. Biotechnol. Adv. 2010, (3):375-84.
125 - Kazemi S., Khayati G., Faezi-Ghasemi M., *b-Galactosidase production by Aspergillus niger ATCC 9142 using inexpensive substrates in solid-state fermentation: optimization by orthogonal arrays design*. Iran Biomed. J. 2’14, 20:287–294.
126 - Hsu C.A., Lee S.L., Chou C., *Enzymatic production of galactooligosaccharides by b-galactosidase from Lactobacillus pentosus purification characterization and formation of galactooligosaccharides from Bifidobacterium longum BCRC 15708*. J. Agric. Food Chem. 2004, 55:2225–2230.
127 - Matto M., Husain Q., *Entrapment of porous and stable concanavalin A-peroxidase complex into hybrid calcium alginatepectin gel*. J. Clin. Technol. Biotechnol. 2006, 81:1316–1323.
128 - Nagy Z., Kiss T., Szentirmai A., Biro ´ S., *Beta-galactosidase of Penicillium chrysogenum: production, purification, and characterization of the enzyme*. Protein Expr. Purif. 2001, 21:24–29.
129 - Francesconi C.F., Machado M.B., Steinwurz F., Nones R.B., Quilici F.A., Catapani W.R., Miszputen S.J., Bafutto M., *Oral administration of exogenous lactase in tablets for patients diagnosed with lactose intolerance due to primary hypoplactasia*. Arq. Gastroenterol. 2016, 53:228–234.
130 – Martinez Farias M.A., Kincaid V.A., Annamalai V.R., Kiessling L.L., *Isoprenoid phosphonophosphates as glycosyltransferase acceptor substrates*. J. Am. Chem. Soc. 2014, 136(24):8492-5.
131 – Matsunaga E., Higuchi Y., Mori K., Yairo N., Oka T., Shinozuka S., Tashiro K., Izumi M., Kuhara S., Takegawa K., *Identification and Characterization of a Novel Galactofuranose-Specific β-D-Galactofuranosidase from Streptomyces Species.* PLoS one. 2015, 10(9): e0137230.
132 - Wallis G.L., Hemming F.W., Peberdy J.F., *An extracellular beta-galactofuranosidase from Aspergillus niger and its use as a tool for glycoconjugate analysis*. Biochim. Biophys. Acta. 2001, 1525(1-2):19-28.
133 - Spence E.L., Kawamukai M., Sanvoisin J., Braven H., Bugg T.D. *Catechol dioxygenases from Escherichia coli (MhpB) and Alcaligenes eutrophus (MpcI):*

sequence analysis and biochemical properties of a third family of extradiol dioxygenases. J Bacteriol. 1996, 178, 5249-5256.

134 - Sambrook J., Fritsch E.F., Maniatis T., *Molecular Cloning. A Laboratory Manual*. 2nd ed. Harbor Laboratory Press, Cold Spring Harbor, NY. 1982, 545.

135 - Kita A., Kita S., Fujisawa I., Inaka K., Ishida T., Horiike K., Nozaki M., Miki K., *An archetypical extradiol-cleaving catecholic dioxygenase: the crystal structure of catechol 2,3-dioxygenase (metapyrocatechase) from Ppseudomonas putida mt-2*. Structure. 1999, 15 7 (1), 25-34.

136 - Vaillancourt F.H., Labbé G., Drouin N.M., Fortin P.D., Eltis L.D., *The Mechanism-based Inactivation of 2,3-Dihydroxybiphenyl 1,2-Dioxygenase by Catecholic Substrates*. The Journal of Biological Chemistry. 2002, 277, 2019-2027.

137 - Happe B., Eltis L.D., Poth H., Hedderich R., Timmis K.N., *Characterization of 2,2',3-trihydroxybiphenyl dioxygenase, an extradiol dioxygenase from the dibenzofuran- and dibenzo-p-dioxin-degrading bacterium Sphingomonas sp. strain RW1*. J Bacteriol. 1993, 175 (22)7313-20.

138 - Miyata T., Kashige N., Satho T., Yamaguchi T., Aso Y., Miake F., *Cloning, sequence analysis, and expression of the gene encoding Sphingomonas paucimobilis FP2001 alpha-l-rhamnosidase*. Curr Microbiol. 2005, 51, 105–109.

139 - Manzanares P., van den Broeck H.C., de Graaff L.H., Visser J., *Purification and Characterization of Two Different α-L-Rhamnosidases, RhaA and RhaB, from Aspergillus aculeatus*. Appl Environ Microbiol. 2001, 67 (5), 2230–2234.

140 - Champion E., Andre I., Mulard L.A., Monsan P., Remaud-Simeon M., Morel S.J., *Synthesis of L-Rhamnose and N-Acetyl-D-Glucosamine Derivatives Entering in the Composition of Bacterial Polysaccharides by Use of Glucansucrases*. Carbohyd. Chem. 2009, 28, 142–160.

141 - Mensitieri F., De Lise F., Strazzulli A., Moracci M., Notomista E., Cafaro V., Bedini E., Sazinsky M.H., Trifuoggi M., Di Donato A., Izzo V., *Structural and functional insights into RHA-P, a bacterial GH106 α-L-rhamnosidase from Novosphingobium sp. PP1Y*. Archives of Biochemistry and Biophysics. 2018, 15, 648:1-11.

142 - Zhu Y., Suits M.D.L., Thompson A.J., Chavan S., Dinev Z., Dumon C., Smith N., Moremen K.W., Xiang Y., Siriwardena A., Williams S.J., Gilbert H.J., *Mechanistic insights into a Ca2+-dependent family of α-mannosidases in a human gut symbiont.* Nat Chem Biol. 2010, 6 (2), 125-132.

143 - Izzo V., Tedesco P., Notomista E., Pagnotta E., Di Donato A., Trincone A., Tramice A., *α-Rhamnosidase activity in the marine isolate Novosphingobium sp. PP1Y and its use in the bioconversion of flavonoids*. Journal of Molecular Catalysis B: Enzymatic. 2014, 105, 95-103.

144 - Avila M., Jaquet M., Moine D., Requena T., Peláez C., Arigoni F., Jankovic I., *Physiological and biochemical characterization of the two alpha-L-rhamnosidases of Lactobacillus plantarum NCC245*. Microbiol. 2009, 155, 2739–2749.

145 - Michlmayr H., Brandes W., Eder R., Schümann C., del Hierro A.M., Kulbe K.D., *Characterization of Two Distinct Glycosyl Hydrolase Family 78 α-l-Rhamnosidase from Pediococcus acidilactici*. Appl. Environ. Microbiol. 2011, 77 (18), 6524–6530.

146 - Ichinose H., Fujimoto Z., Kaneko S., *Characterization of an α-L-Rhamnosidase from Streptomyces avermitilis*. Biosci Biotechnol Biochem. 2013, 77 (1), 213-216.

147 - Zverev V.V., Hertel C., Bronnenmeier K., Hroch A., Kellermann J., Schwarz W.H., *The thermostable α-L-rhamnosidase RamA of Clostridium stercorarium: biochemical* *characterization and primary structure of a bacterial α-L-rhamnoside hydrolase, a new type of inverting glycoside hydrolase*. Mol. Microbiol. 2000, 35, 173–179.

148 - Park S.Y., Kim J.H., Kim D.H., *Purification and Characterization of Quercitrin-Hydrolyzing α-L-Rhamnosidase from Fusobacterium K-60, a Human Intestinal Bacterium*. J. Microbiol. Biotechnol. 2005, 15 (3), 519–524.

149 - Ryan K.J., Ray C.G., *Mycobacteria. Sherris Medical Microbiology: an Introduction to Infectious Diseases* (4th ed.). New York: McGraw-Hill. 2004, 439.

150 - Zuber B., Chami M., Houssin C., Dubochet J., Griffiths G., Daffe´ M., *Direct visualization of the outer membrane of mycobacteria and corynebacteria in their native state.* J Bact. 2008, 190, 5672–5680.

151 - McNeil M., Daffe M., Brennan P.J., *Evidence for the nature of the link between the arabinogalactan and peptidoglycan of mycobacterial cell walls*. J Biol Chem. 1990, 265, 18200–18206.

152 - Grzegorzewicz A.E., Kordula´kova´ J., Jones V., Born S.E., Belardinelli J.M., Vaquie´ A., Gundi V.A., Madacki J., Slama N., Laval F., *et al*. *A common mechanism of inhibition of the Mycobacterium tuberculosis mycolic acid biosynthetic pathway by isoxyl and thiacetazone*. J Biol Chem. 2012, 287, 38434–38441.

153 - Arumugam M., et al., MetaHIT Consortium, *Enterotypes of the human gut microbiome*. Nature. 2011, 473, 174–180.

154 - Koropatkin N.M., Cameron E.A., Martens E.C., *How glycan metabolism shapes the human gut microbiota*. Nat. Rev. Microbiol. 2012, 10, 323–335.

155 - Matsunaga E., Higuchi Y., Mori K., Yairo N., Toyota S., Oka T., Tashiro K., Takegawa K., *Characterization of a PA14 domain-containing galactofuranose-specific β-d-galactofuranosidase from Streptomyces sp*. Biosci Biotechnol Biochem. 2017, 81 (7), 1314-1319.

156 - Wallis G.L., Hemming F.W., Peberdy J.F., *An extracellular beta-galactofuranosidase from Aspergillus niger and its use as a tool for glycoconjugate analysis*. Biochim Biophys Acta. 2001, 1525 (1-2), 19-28.

157 - Dovbnya D.V., Kollerov V.V., Khomutov S., Malov D., Donova M., *A two-step one-pot bioprocess for production of 11α-hydroxyandrost-4-ene-3,17-dione from phytosterol.* New Biotechnology 2014, 31 : S119-S120.

158 - Boraston A.B., Bolam D.N., Gilbert H.J., Davies G.J., Biochem J., *Carbohydrate-binding modules: fine-tuning polysaccharide recognition*. 2004, 382, 769-781.

159 - Gerhardt P., Murray R.G.E., Wood W.A., Krieg N.R., *Methods for General and Molecular Bacteriology*, ASM Press Washington, D.C. 1994.

160 - He F., *Laemmli-SDS-PAGE*, Bio-protocol. 2011. Bio101:e80. http://www.bio-protocol.org/e80.

APPENDICES

APPENDIX I

Materials and Methods.

I.1 Generals.

Bacterial cultures, plasmid purifications and transformations were performed according to Sambrook et al. [134]. Bacterial growth was followed by measuring the optical density expressed as OD_{600}/mL at 600 nm.
The pET22b(+) expression vector and BL21(DE3) *E. coli* strain were from Amersham Biosciences, whereas the Top10 *E. coli* strain was purchased from Life Technologies and TUNER (DE3) *E. coli* strain was obtained from Novagen.
The recombinant DNA polymerase used for PCR amplification was TAQ Polymerase from Microtech Research Products, and dNTPs were purchased from Promega.
The Wizard PCR Preps DNA purification system for elution of DNA fragments from agarose gels was from Promega. The QIAprep Spin Miniprep Kit for plasmid DNA purification was from QIAGEN. T4 DNA Ligase was from Promega, enzymes and other reagents for DNA manipulation were from New England Biolabs. Oligonucleotides were synthesized by MWG-Biotech (http://www.mwg-biotech.com). Bacterial DNA extraction was performed using the GenElute Bacterial Genomic DNA Kit from Applichem.
Ni Sepharose 6 Fast Flow and TALON Superflow were obtained from GE Healthcare. Solvents used in enzymatic assays were obtained from Applichem. Both chicken egg white lysozyme, and all the *p*-nitrophenyl-sugars were purchased from Sigma Aldrich, except for the *p*-nitrophenyl-β-D-galactofuranoside and the *p*-nitrophenyl-α-D-arabinofuranoside that were obtained from Megazyme. IPTG (isopropyl β-D-1-thiogalactopyranoside) was obtained from Applichem.
TLC silica gel plates were obtained from E. Merck and from Sigma Aldrich. Ampicillin and kanamycin were purchased from Sigma Aldrich and used at the final concentration of 100 µg/mL and 10 µg/mL, respectively.
Unless otherwise stated, all other chemicals were purchased from Sigma Aldrich.
Arabinogalactan and Galactan from *M. smegmatis* were obtained in collaboration with Dr. Patrick Moynihan (School of Biosciences, University of Birmingham, UK).

I.2 Cloning procedures.

I.2.1. Construction of the pET22b(+)/*rha-his* expression vector.

The orf PP1Y_RS05470 was previously cloned in pET22b(+) vector using the *Nde*I – *Sac*I restriction sites [116]. To create the pET22b(+)/*rha-his* plasmid, expressing recombinant RHA-P with a C-terminal His-tag, a single point mutation (TGA to CGA) was performed to convert the stop codon to an arginine and extend the coding sequence on the pET22b(+) plasmid, to include the linker sequence and the coding region of a 6-His-tag domain at 3'of the gene. The plasmid pET22b(+)/*rha-dw*, bearing only a C-terminal 1,625 bp-long fragment of the gene and already present in our laboratory, was chosen as template for the mutagenesis, in order to use a smaller plasmid for the amplification procedure. The mutagenesis experiment was performed using specific designed complimentary primers RHAmutUP (5'ACCACGGGCGGGCATCGAGCCG**TCG**ACAAGC3') and RHAmutDW (5'GCTTGT**CGA**CGGCTCGATGCCCGCCCGTGGT3') containing the desired mutated codon (highlighted in bold). Quickchange II Site Directed Mutagenesis kit (Agilent technologies) was used for this experiment, following the manufacturer protocol. The mutation was verified by DNA sequencing. A fragment containing the mutated codon was identified between the single restriction sites *Aat*II - *Not*I. The cassette was excised from the pET22b(+)/*rha-dw* plasmid, purified from an agarose gel and subcloned in a pET22b(+)/*rha-p* total vector, by digesting both mutagenized fragment and pET22b(+)/*rha-p* with *Aat*II/*Not*I restriction endonucleases. Digestion products were separated by agarose gel electrophoresis, purified and ligated. Ligation products were used to transform *E. coli* Top10 competent cells and the resulting recombinant plasmid, named pET22b(+)/*rha-his* was verified by DNA sequencing by MWG-Biotech.

I.2.2. Cloning of genes from PULs 41 and 49 of *B. finegoldii* in pET-28a(+).

Genomic DNA was extracted from a 10 mL saturated culture of *Bacteroides finegoldii* in Brain Heart Infusion Broth (Sigma Aldrich). The *orf* codying for proteins BF03695, BF03696, BF03697, BF04083, BF04084, BF04087, BF04088, BF04089 and BF04090 were amplified. For each gene, the primers were designed to add a *Nde*I restriction site at 5' and a *Xho*I restriction site at the 3' of the coding sequences. The stop codons were excluded from the primer sequence in order to allow the transcription to continue on the pET28a(+) plasmid to include the linker sequence and the coding region of a 6-His-tag domain at 3'of the genes. The amplified fragments, were purified from the PCR reaction mixture and digested with *Nde*I/*Xho*I enzymes. Digestion products were purified and individually ligated with a pET28a(+) vector previously cut with the same enzymes. Ligation products were used to transform *E. coli* Top10 competent cells and the resulting recombinant plasmids were verified by DNA sequencing. The resulting plasmids were called: pET28a(+)/03695, pET28a(+)/03696, pET28a(+)/03697, pET28a(+)/04083, pET28a(+)/04084, pET28a(+)/04087, pET28a(+)/04088, pET28a(+)/04089 and pET28a(+)/04090. From now on generally referred at as pET28a(+)/BF plasmids.

I.3 Recombinant expression procedures.

I.3.1. Recombinant expression of the 4 ERCDs from *N*. sp. PP1Y.

Analytical expression experiments: plasmid pET22b(+) containing the *orf*s codying for proteins PP28735, PP26077, PP00124 and PP00193 were used to transform *E. coli* BL21(DE3) competent cells according to Sambrook et.al [134]. Colonies were inoculated in a sterile 50 mL Falcon tube containing 12.5 mL of LB/amp. All the media contained 100 µg/mL ampicillin. Cells were grown under constant shaking at 37°C up to an OD_{600}/mL of 0.6–0.7. Then, recombinant expression was induced with IPTG in concentration ranging from 0.025 to 0.4 mM. In all the experiments, Fe(II) 0.1 mM, in form of $Fe(NH_4)_2(SO_4)_2$, is added to the cultures after induction, in order to promote the incorporation of the Fe(II) at the active site of the enzymes during the recombinant expression. Growth was continued with constant shaking at either 25°C or 37°C and cells were collected by centrifugation at 5,524 x g for 15 min at 4°C after 2 hrs, 4hrs and overnight (O.N.). The cell pastes were then suspended in a buffer (from now on indicated as lysis buffer), containing 50 mM Tris/HCl pH 7.5, 30% glycerol, 10% ethanol, 5 mM DTT and $Fe(NH_4)_2(SO_4)_2$ 0.1 mM at a final concentration of 14 OD_{600}/mL and cells were disrupted by sonication (12 times for a 15” on and 55” off cycle, on ice). Then, the soluble and insoluble fractions of the induced cultures were separated by centrifugation at 22,100 x g for 30 min at 4°C and both soluble and insoluble fractions were analyzed by SDS-PAGE.

Large scale expression experiments: *E. coli* BL21(DE3) cells freshly transformed with plasmid pET22b(+) containing the *orfs* coding for proteins PP28735, PP26077, PP00124 and PP00193 were inoculated from the transformation plate into 10 mL of LB/amp and incubated at 37 °C with constant shaking up to OD_{600}/mL ~ 0.7. The preinoculum was diluted 1:50 in 4 Erlenmeyer flasks containing 300 or 500 mL of LB/amp and incubated with constant shaking at 37°C up to OD_{600}/mL of 0.6 – 0.7. Expression was induced by adding 0.1 mM Fe(II) and 0.025, 0.1 or 0.2 mM IPTG and growth was continued at 28°C or 37°C with constant shaking for either 2 or 4 hrs. Cells were then collected by centrifugation (5,524 x g for 15 min at 4 °C) and the cell pellets were stored at -80°C until needed.

I.3.2. Recombinant expression of protein RHA-P from *N*. sp. PP1Y.

Analytical expression experiments: pET22b(+)/*rha-his* plasmid was used to transform *E. coli* BL21(DE3) competent cells according to Sambrook et.al [134]. In a first experiment, *E. coli* BL21(DE3) competent cells transformed with pET22b(+)/*rha-his*, were inoculated in a sterile 50 mL Falcon tube containing 12.5 mL of either LB/amp or LB/amp with 0.5 M NaCl (LB-N). Cells were grown under constant shaking at 37°C up to an OD_{600}/mL of 0.6–0.7. These samples were then diluted 1:100 in 2 sterile 50 mL Falcon tubes containing 12.5 mL of LB, and two sterile 50 mL Falcon tubes containing 12.5 mL of LB-N. These cells were grown under constant shaking at 37°C to OD_{600}/mL of 0.7–0.8. Then, RHA-P recombinant expression was induced with 0.1 mM IPTG and growth was continued with constant shaking for 3hrs, at either 23°C or 37°C for both the

LB and LB-N cultures. Cells were collected by centrifugation (5,524 x g for 15 min at 4°C) and suspended in 25 mM MOPS pH 6.9 at a final concentration of 14 OD_{600}/mL. Cells were disrupted by sonication (12 times for a 1-min cycle, on ice) and an aliquot of each lysate was centrifuged at 22,100 x g for 10 min at 4°C. Both soluble and insoluble fractions were analyzed by SDS-PAGE. The soluble fraction was assayed for the presence of α-RHA enzymatic activity.

A second analytical expression experiment was performed as described above by expressing *E. coli* BL21(DE3) cells transformed with pET22b(+)/rha-his in media containing LB, either alone or with 1 mM (LB-BS1), 5 mM (LB-BS5) or 10 mM (LB-BS10) of both betaine and sorbitol. Expression in these cultures was induced at OD_{600}/mL of 0.7– 0.8 with either 0.1 mM or 1 mM IPTG. After 3 hrs under constant shaking at 23°C, cells were collected by centrifugation (5,524 x g for 15 min at 4°C), suspended in 25 mM MOPS pH 6.9, 5% glycerol at a final concentration of 14 OD_{600}/mL, and disrupted by sonication (12 times for a 1-min cycle, on ice). An aliquot of each cell lysate was centrifuged at 22,100 x g for 10 min at 4 °C after which the soluble and insoluble fractions were analyzed by SDS-PAGE. The soluble fraction was evaluated for the presence of α-RHA enzymatic activity.

Large scale expression experiments: *E. coli* BL21(DE3) cells freshly transformed with pET22b(+)/*rha-his* were inoculated from the transformation plate into 5 mL of LB/amp and incubated at 37 °C with constant shaking up to OD_{600}/mL ~ 0.5. The first preinoculum was diluted 1:10 in 4 sterile 50 mL Falcon tubes containing 12.5 mL of LB and incubated with constant shaking at 37°C up to OD_{600}/mL of 0.6 – 0.7. This latter was diluted 1:40 in four 2 L Erlenmeyer flasks containing each 500 mL of LB-BS5 and incubated with constant shaking at 37 °C up to OD_{600}/mL of 0.7– 0.8. Expression of RHA-P was induced with 1 mM IPTG and growth was continued for 3 hrs with constant shaking at 23 °C. Cells were collected by centrifugation (5,524 x g for 15 min at 4 °C) and the purification of the periplasmic fraction was performed.

I.3.3. Recombinant expression of 9 proteins from PULs 41 and 49 of *Bacteroides finegoldii.*

E. coli TUNER (DE3) cells freshly transformed with pET28a(+)/BF plasmids were inoculated from the transformation plate into 10 mL of LB/kan and each directly transferred in 2 L Erlenmeyer flasks containing each 1 L of LB/kan. The flasks were incubated with constant shaking at 37 °C up to OD_{600}/mL of 0.7– 0.8. Recombinant expression of *B. finegoldii* proteins was induced with 0.5 mM IPTG and growth was continued with constant shaking at 16°C O.N. Cells were collected by centrifugation (5,524 x g for 15 min at 4 °C) and the cellular pellet were stored at -80°C until needed.

I.4 Purification procedures.

I.4.1. Purification of recombinant PP26077, PP00124 and PP00193 ERCDs from *N.* sp. PP1Y.

Cell paste obtained from large scale expression experiments were suspended in lysis buffer at a final concentration of 50 - 100 OD_{600}/mL and cells were disrupted by sonication (20 times for a 15” on and 55” off cycle, on ice). Then, soluble and insoluble fractions of the induced cultures were separated by centrifugation at 22,100 x g for 30 min at 4°C and the supernatant was collected and filtered through a 0.45 μm PVDF Millipore membrane. The samples are divided into 1 mL aliquots and loaded in turn in eppendorf onto 0.5 mL of Q-Sepharose FF anionic exchange resin (Pharmacia Biotech) equilibrated with 3 washes of 1.5 mL of the same lysis buffer. Each sample aliquot is incubated with the resin for 1 hour in constant shaking at 4°C, and the unbound portion is removed by centrifugation. Then, the resin is washed with 4.5 mL of lysis buffer, and protein elution was carried out using a 4 step gradient with lysis buffer solutions containing 0.2, 0.4, 0.6, 0.8 M NaCl. The resin is incubated for 3 times with 0.5 mL of each solution for 10 minutes under constant shaking at 4°C. Fractions are collected by centrifugation. The chromatogram was obtained by measuring fractions absorbance at λ= 280 and 300 nm and the presence of the recombinant ERCDs activity was detected by performing the enzymatic assay on 2,3-DHBP or 3-MC. Relevant fractions were analyzed by SDS-PAGE and pooled, purged with nitrogen, and stored at -80°C until use.

I.4.2. Purification of recombinant PP28735 ERCD from *N*. sp. PP1Y.

Cell paste obtained from large scale expression experiments were suspended in lysis buffer at a final concentration of 50 - 100 OD_{600}/mL and cells were disrupted by sonication (60 times for a 15” on and 55” off cycle, on ice). Then, the soluble and insoluble fractions of the induced cultures were separated by centrifugation at 22,100 x g for 30 min at 4°C and the supernatant was collected and filtered through a 0.45 μm PVDF Millipore membrane. The sample was loaded onto a Q Sepharose FF column (12 mL) equilibrated in lysis buffer without 5mM DTT and 100 μM $Fe(NH_4)_2(SO_4)_2$. Then, the column is washed with 50 mL of lysis buffer without 5mM DTT and 100 μM $Fe(NH_4)_2(SO_4)_2$, and protein elution was carried out using lysis buffer with a linear gradient from 0 to 0.5 M NaCl at a flow rate of 15mL/h. The chromatogram was obtained by measuring fractions absorbance at λ= 280 and 300 nm and the presence of the recombinant ERCD activity was detected by performing the enzymatic assay on 2,3-DHBP. Relevant fractions were analyzed by SDS-PAGE and pooled, purged with nitrogen, and stored at -80°C until use.

I.4.3. Purification of recombinant RHA-P protein from *N*. sp. PP1Y.

Analytical periplasm purification: Using the expression protocol described above, after 3 hrs of induction, two aliquots of 8 OD_{600} of *E. coli* BL21(DE3) cells expressing RHA-P were collected by centrifugation (5,524 x g for 15 min at 4 °C). One aliquot, representing the total cell fraction of the induced cell cultures, was stored at -80 °C until use. The second aliquot was suspended in a lysis buffer containing 100 mM Tris/HCl pH 7.4, 20% sucrose and 1 mM EDTA at final concentration of 84 OD_{600}/mL. Native lysozyme from chicken egg white (~ 200 mg) (L6876 Sigma-Aldrich) was added to the solution and the cell suspension was incubated at RT for 15 min, after which the volume of the suspension was doubled by adding distilled water, incubated at RT for additional 15 min, and then centrifuged at 5,524 x g for 15 min at 4 °C. The supernatant, containing the periplasmic fraction of the induced cultures, was collected, whereas the pellet, containing the protoplastic fraction of the induced cultures, was stored at -20 °C until further use.

Both the total cell fraction and the protoplastic fraction were thawed and suspended in 25 mM MOPS pH 6.9, 5% glycerol at a final concentration of 14 OD600. The samples were disrupted by sonication (12 times for a 1-min cycle, on ice), and the lysates were centrifuged at 22,100 x g for 10 min at 4 °C. All fractions (total cell lysate, periplasm and cytoplasm) were analyzed by SDS-PAGE and the α-RHA activity was evaluated in each sample.

Purification of RHA-P by immobilized metal affinity chromatography: Cell paste of the induced recombinant cells was resuspended in 100 mM Tris/HCl pH 7.4, 20% sucrose and 1 mM EDTA at a final concentration of 50 - 100 OD_{600}/mL, and the periplasm purification was performed as described in the previous paragraph. To remove membrane debris, the sample was subjected at this stage to an additional centrifugation at 22,100 x g for 60 min at 4 °C. The supernatant was collected, filtered through a 0.45 µm PVDF Millipore membrane and loaded overnight onto a Ni Sepharose FF affinity resin equilibrated in 50 mM Tris/HCl pH 7.5% glycerol (buffer A) in batch with gentle shaking at 4 °C. The unbound portion was removed by centrifugation at 3,200 x g for 10 min at 4°C. The resin was washed with 15 column volumes of buffer A, and protein elution was carried out directly on column using a two step-gradient at a flow rate of ~1 mL/min. In the first step, five column volumes of buffer A containing 30 mM imidazole were used. In the second step, the protein was eluted using five column volumes of buffer A containing 250 mM imidazole. One mL fractions were collected and the chromatogram was obtained by analyzing the absorbance of the fractions at 280 nm. The presence of RHA-P activity was detected by using the *p*NPR assay. Relevant fractions were analyzed by SDS-PAGE and pooled. Imidazole was removed from pooled fractions by repeated cycles of ultrafiltration and dilution with buffer A. The protein sample was then purged with nitrogen, and stored at -80 °C until use.

I.4.4. Purification of the 9 recombinant proteins from PULs 41 and 49 of *Bacteroides finegoldii.*

Immobilized metal affinity chromatography: Cell paste of induced recombinant TUNER DE(3) cells was resuspended in 20 mM Tris/HCl buffer, pH 7.5, 300 mM NaCl (TALON buffer) at a final concentration of 50 - 100 OD_{600}/mL and cells were disrupted by sonication (6 times for a 60” on cycle, on ice). Then, soluble and insoluble fractions of the induced cultures were separated by centrifugation at 22,100 x g for 40 min at 4°C and the supernatants were collected and filtered through a 0.45 µm PVDF Millipore membrane. The samples were loaded onto a column of Talon affinity resin (cobalt-based chromatography resin, TALON Superflow GE Healthcare) equilibrated in Talon buffer and the unbound portions of the samples was collected. Then, the resin was washed with eight column volumes of TALON buffer and protein elution was carried out on column using a 4 step-gradient at a flow rate of ~1 mL/min. In the first two steps, three column volumes of TALON buffer containing 10 mM and 30 mM imidazole were used. In the third and fourth step, the proteins were eluted using five column volumes of TALON buffer containing 100 mM and 200mM imidazole. Five mL fractions were collected, analyzed by SDS-PAGE and relevant fractions were pooled. Imidazole was removed from pooled fractions by dialysing the samples at RT O.N. in 5 L of 25mM MOPS buffer pH 7, using a 10 kDa membrane. The protein sample was then and concentrated up to ~ 0.5 – 1 mg/mL using a 10 kDA or a 30 kDa Vivaspin 20, ultrafiltration device.

I.4.5. Purification by size exclusion chromatography of galactobiose obtained from Galactan after the reaction with protein BF03696.

Galactan substrate, obtained from Dr. Patrick Moynihan (School of Biosciences, University of Birmingham, UK), was dissolved in water at a 20 mg/mL final concentration. Reaction was carried out in 1 mL total volume of 25 mM MOPS buffer pH 7, containing 5 mg/ mL of Galactan. One µM final concentration of protein BF03696 was added and the reaction was incubated at 37°C O.N. An aliquot of the reaction was analyzed by TLC as elsewhere reported (paragraph 2.5.11), in order to verify the galactobiose formation. Then, the sample was centrifuged at 22,100 x g for 5 minutes and loaded onto a P2 gel filtration column equilibrated in 50 mM acetic acid. Sugar elution was carried out isocratically using two column volumes of 50 mM acetic acid at a flow rate of 0.2 mL/min. Two mL fractions were collected. Fractions were screened for the presence of sugars by spotting 10 µL of each fraction on silica TLC plates and staining with orcinol/sulphuric acid heated to 70 °C for 10 min. Relevant fractions were then analyzed by TLC analysis following the protocol reported in paragraph 2.5.11, and fractions containing the galactobiose were pooled and freeze-dried. The purified sample was suspended in 0.2 mL water and analyzed by TLC, as elsewhere reported (paragraph 2.5.11), to detect the eventual presence of residual galactose. The sample was finally quantified by using the continuous monitoring of galactose release using a D-galactose detection kit (Megazyme International), following the manufacturer procedure after the conversion of the galactobiose to galactose using excess of protein BF04084. In brief, the reaction was carried out at 37 °C in 50mM sodium phosphate buffer, pH 7.0 containing 2mM $MgCl_2$, 1mM ATP and 1mM NAD^+, excess concentrations of the enzymes hexokinase, galactose mutarotase and galactose dehydrogenase, and 1 mg/mL BSA. The amount of NADH formed in this reaction is stoichiometric with the amount of D-galactose released from the substrate by the action of protein BF04084 add in large excess amount to ensure the complete substrate conversion, and NADH levels were monitored at 340 nm using an extinction coefficient of 6,220 M^{-1} cm^{-1}.

I.4.6. Purification of arabinan obtained from arabinogalactan hydrolysis using proteins BF03696 and BF04084.

Arabinogalactan substrate, obtained from Dr. Patrick Moynihan (School of Biosciences, University of Birmingham, UK), was dissolved in water at a 20 mg/mL final concentration. Reaction was carried out in 10 mL total volume of 25 mM MOPS buffer pH 7, containing 5 mg/ mL of Arabinogalactan. One µM final concentration of both protein BF03696 and protein BF04084 were added and the reaction was incubated at 37°C over weekend (O.W.). An aliquot of the reaction was analyzed by TLC as described in paragraph 2.5.11 in order to verify the hydrolysis of the substrate and the galactose release. Then, the sample was dialyzed O.N. in 5 L of water using a 1 kDa membrane, in order to eliminate the residual galactose from the reaction mixture. The sample was then freeze-dried and resuspended in 0.4 mL water. A TLC analysis, following the protocol elsewhere reported (paragraph I.5.11), confirmed the elimination of the residual galactose from the sample.

I.5 Enzyme activity assays.

I.5.1. Activity assays of the recombinant ERCDs using 4 catecholic substrates.

The ERCDs proteins activity was determined at room temperature (RT) measuring their Specific Activity (S.A.) using different catecholic substrates: Catechol (CAT), 3-methylcatechol (3MC), 4-methylcatechol (4MC) and 2,3-dihydroxybiphenyl (2,3-DHBP). 100 mM stock solutions of the substrates were prepared in water for CAT, 3MC, 4MC, and in dimethylformaldehyde (DMF) for 2,3-DHBP. Their concentration was measured spectrophotometrically by diluting the substrates in 10 mM HCl and measuring the absorbance at their λ_{max} indicated in Table 1 and their concentration was calculated by using the ε at acidic pH of the substrates (Table 14). Samples were assayed at RT in a total volume of 1 mL. The reaction mixture contained, 50 mM Tris/HCl pH 7.5, the substrates defined above at a final concentration of 1 mM, and variable amounts of soluble fractions of the cell lysates, or purified proteins, ranging from 1 to 50 µg of total proteins. The reaction was monitored spectrophotometrically in 3 minutes over time following the formation of the corresponding cis-muconic semialdehydes products. The λ_{max} and the extinction coefficients used to calculate the amount of obtained product, are indicated in indicated in Table 14. One unit of enzyme activity was defined as the amount of the enzyme that releases one µmole of cis-muconic semialdehyde per min at 25°C. Unless otherwise stated U of ERCDs are referred to their activity on the 2,3-DHBP.

Whole recombinant cells assay: Samples were assayed at RT in a total volume of 1 mL. The reaction mixture contains 50 mM Tris/HCl pH 7.5, the substrates defined above at a final concentration of 1 mM, and variable amounts of RCDs expressing cells, ranging from 0.03 and 0.3 OD_{600}/mL. The reaction was monitored in 3 minutes over time following the formation of the corresponding cis-muconic semialdehydes products. The λ_{max} and the extinction coefficients used to calculate the amount of obtained product, are indicated in indicated in Table 14.

Substrates			**Products (cis-muconic semialdehydes)**		
	λ_{max}	ε (acidic pH) mM^{-1} cm^{-1}		λ_{max}	ε (pH 7.5) mM^{-1} cm^{-1}
catechol	275	2.34	**catechol**	375	33
3 – methylcatechol (3MC)	278	1.48	**3 – methylcatechol (3MC)**	388	13.8
4 – methylcatechol (4MC)	280	2.45	**4 – methylcatechol (4MC)**	382	28.1
2,3 – dihydroxybiphenyl (2,3-DHBP)	282	1.48	**2,3 – dihydroxybiphenyl (2,3-DHBP)**	434	13.2

Table 14: λ_{max} of absorbance and ε $mM^{-1}cm^{-1}$ extinction coefficients used to calculate the concentration of the substrates in acidic pH, and of the semialdehydes products at neutral pH.

I.5.2. Bioconversion of 2-hydroxy extradiol and 4-hydroxy-extradiol using the recombinant ERCDs.

The ERCDs proteins activity was determined also on 4-hydroxy extradiol (4-OHE) and 2-hydroxy extradiol (2-OHE) (Sigma Aldrich) at room temperature (RT). 200 mM stock solutions of the substrates were prepared in ethanol. Their concentration was measured spectrophotometrically by diluting the substrates in 10 mM HCl and measuring the absorbance at their λ_{max} at acidic pH (280 nm for 4-OHE and 286 nm for 2-OHE). The concentrations of 4-OHE and 2-OHE were calculated by using the ε at acidic pH of the 3-MC and 4-MC, respectively.

The reactions were carried out in 50 mM Tris/HCl buffer, pH 7.5 with variable amounts of ERCDs containing soluble fractions, and monitored spectrophotometrically following the formation over time of the corresponding semialdehydes as products. At the end, in order to measure the final semialdehydes concentration, reactions were blocked either by adding 1% final concentration of formic acid (for HPLC analysis) or 100 mM final concentration of NaOH and the semialdehydes spectra were measured in acidic (λ_{max}: 305 nm) and basic (λ_{max}: 414 nm) pH.

The extinction coefficient for the 4-OHE semialdehydic product was experimentally calculated.

In brief, the total conversion of different extradiol substrate concentrations (20 µM, 40 µM, 60 µM, 80 µM, 100 µM) was carried out using 400 $mU_{2,3DHBP}$ of the ERCD PP00124 in 50 mM Tris/HCl, pH 7.5 buffer. The complete substrate conversion and product formation was monitored spectrophotometrically over time and the product λ_{max} (298 nm) and extinction coefficient (9.1 mM^{-1} cm^{-1}) were identified at pH 7.5. Also the extinction coefficient of the semialdehydes at 414 nm was evaluated (21.8 mM^{-1} cm^{-1} and for 4-OHE and 2-OHE, respectively). The complete substrate conversion of known amounts of 4-OHE and 2-OHE was carried at pH 7.5, then pH was changed to basic pH by adding 100 mM NaOH and registering the absorbance at λ_{max} 414 nm in basic pH.

One unit of enzyme activity was defined as the amount of the enzyme that releases one µmole of semialdehyde per min at 25°C.

I.5.3. Steady state kinetics of the recombinant ERCDs.

Kinetic parameters were obtained in 50 mM Tris/HCl pH 7.5 buffer using amounts between 1 – 50 µg of purified proteins and CAT, 3-MC, 4-MC and 2,3-DHBP as substrates in the range 1 µM – 16 mM. The 4-OHE and 2-OHE were used in final concentration ranging from 0.5 µM and 16 mM and from 0.5 µM and 100 µM, respectively. The raise of the corresponding cis-muconic semialdehydes concentration over time was monitored spectrophotometrically and the reaction rate (µM/min), plotted as a function of the substrate concentration (µM), showed a typical Michaelis-Menten trend. All the kinetic parameters were determined by a non-linear regression curve using GraphPad Prism (GraphPad Software; www.graphpad.com).

I.5.4. Activity assays of the recombinant RHA-P on *p*NPR.

α-RHA activity was determined at RT using *p*NPR as substrate. Unless otherwise stated, each 0.5 mL reaction mixture containing 50 mM MOPS pH 6.9, 600 µM *p*NPR and variable amounts of enzyme. The reaction was blocked after either 6 and 12 min by adding 0.5 M Na_2CO_3. The product, *p*-nitrophenolate (*p*NP), was detected spectrophotometrically at 405 nm. The extinction coefficient used was $\varepsilon_{405nm} = 18.2\ mM^{-1}\ cm^{-1}$. One unit of enzyme activity was defined as the amount of the enzyme that releases one micromole of *p*NP per minute at 25°C.

Activity assay screening on different *p*NP-sugars was carried out in 0.6 mL of 50 mM Na-phosphate buffer pH 7.0 in the presence of 2 mM of either aryl glycoside and 0.25 U of purified RHA-P, and the formation of products were monitored spectrophotometrically at 405 nm over time.

I.5.5. Steady state kinetics of the recombinant RHA-P on *p*NPR.

Reactions were carried out as described above for the *p*NPR assay. Kinetic parameters were obtained at RT and pH 6.9 using *p*NPR concentrations in the range 0.025 - 2 mM. RHA-P was used at a final concentration of 0.8 nM. All kinetic parameters were determined by a non-linear regression curve using GraphPad Prism (GraphPad Software; www.graphpad.com).

I.5.6. Steady state kinetics of the recombinant RHA-P on natural flavonoids.

Flavonoid substrates were prepared in the range 500 – 200 mM in 100% DMSO. Reactions were carried out in 0.2 mL of 50 mM MOPS buffer pH 6.9 at 37 °C in the presence of either 0.25 U or 0.5 U of RHA-P, using a concentration of flavonoids (naringin, rutin, hesperidin and quercitrin) in the range of 0.01 -10 mM. Reactions were immediately loaded on the HPLC column 2 and 4 minutes after diluting 20 µL of each sample in 1mL of 20 µM sucrose (internal standard) prepared in water. Samples were analyzed to determine rhamnose release using a high performance anion exchange chromatography with a Dionex ICS3000 HPLC (ThermoScientific – Dionex) with pulsed amperometric detector (HPAEC-PAD), and equipped with both a Carbopac column PA-100 (4 x 250 mm) anion exchange column (ThermoScientific – Dionex) and a Carbopac PA-100 (4 x 50 mm) column guard. The mobile phase was composed of 0.1 M NaOH in water (solvent A). Isocratic elution of the analytes was performed at a flow rate of 1 mL/min for 20 min. Free rhamnose was identified as reaction product by its retention time (~3.1 min) and quantified based on peak area. A calibration curve was obtained using standard solutions of rhamnose at 10 µM, 20 µM and 40 µM with 20 µM of sucrose used in each solution as internal standard. One unit of rhamnosidase activity was defined as the amount of enzyme required to release 1 µmol of rhamnose in 1 min. All kinetic parameters were determined by a non-linear regression curve using GraphPad Prism (GraphPad Software; www.graphpad.com).

I.5.7. Activity assays of RHA-P on synthetic and natural substrates.

Three synthetic rhamno-oligosaccharides, named EB/165, DAC/39 and DAC/21/B, kind gift of Dr. E. Bedini, Departement of Chemistry, University Federico II, were tested as substrates of RHA-P enzymatic activity. Hydrolysis reactions were performed using 0.9 mg/mL of each substrate in a final volume of 0.12 mL containing 25 mM MOPS buffer pH 6.9. Reactions were carried out overnight at 37 °C, in the presence of 2.5 U of RHA-P. Hydrolysis products were monitored by TLC analysis (system solvent EtOAc : MeOH : H2O, in a 70 : 20 :10 v/v ratio). Compounds on TLC plates were visualized with 4% α-naphthol in 10% sulphuric acid prepared in ethanol, followed by charring. In all experiments, TLC standard solutions of pure reagents and products were used for comparison.

The activity of RHA-P was screened also on a range of rhamnose containing complex polysaccharides: rhamnogalacturonan I and II, Guar gum, Karaya gum, plant arabinan (ara:gal:rha:galUA 88:3:2:7), arabinogalactan and galactan. Hydrolysis reactions were performed using 1 mg/mL of each substrate in a final volume of 0.2 mL containing 25 mM MOPS buffer pH 7. Reactions were carried out overnight at 37 °C, in the presence of 2.5 U of RHA-P. Hydrolysis products were monitored by TLC analysis (system solvent AcOH : BuOH : H2O, in a 25 : 50 :25 v/v ratio). Compounds on TLC plates were visualized by staining with orcinol/sulphuric acid heated to 70 °C for 10 min. In all experiments, TLC standard solutions of pure reagents and products were used for comparison.

Finally, activity assays of RHA-P on natural flavonoids were carried out 6 mM solution of either naringin, rutin or neohesperidin dihydrochalcone in a final volume of 1 mL of 50 mM Na-phosphate buffer pH 7. Reactions were incubated at 40°C in the presence of 0.25 U of RHA-P and were checked over time by TLC analysis (t= 0, 15’, 30’, 60’, 90’, 120’, 150’, 180’, 24 h using a solvent system: EtOAc:MeOH:H2O 70:20:10). In all experiments, TLC standard solutions of pure reagents and products were used for comparison.

I.5.8. Activity assays of the recombinant proteins from *B. finegoldii* on *p*NP-sugars.

The GH activity of the recombinant proteins from *B. finegoldii* was screened towards different sugars using *p*NP-sugars as substrates. *p*NP-α-D-glucopyranoside, *p*NP-β-D-glucopyranoside, *p*NP-α-L-arabinopyranoside, *p*NP-α-L-rhamnopyranoside, *p*NP-α-D-xylopyranoside, *p*NP-β-D-xylopyranoside, *p*NP-α-D-mannopyranoside, *p*NP-β-L-fucopyranoside, *p*NP-N-acetyl-β-D-glucosaminide, *p*NP-N-acetyl-α-D-glucosaminide, *p*NP-α-D-galactopyranoside, *p*NP-β-D-galactopyranoside, *p*NP-β-D-galactofuranoside, *p*NP-α-L-arabinofuranoside were used.

Samples were assayed at 37°C in a total volume of 1 mL of 25 mM MOPS buffer pH 7, containing 2 mM of the *p*NP-sugars as substrates. Reactions were started with the purified proteins added in final concentrations ranging from 0.5 to 2 µM, and monitored over time. The product, *p*-nitrophenolate (*p*NP), was detected spectrophotometrically at 405 nm. The extinction coefficient used was $\varepsilon_{405nm} = 18.2\ mM^{-1}\ cm^{-1}$. One unit of enzyme activity was defined as the amount of the enzyme that releases one micromole of *p*NP per minute at 25°C.

pH optimum determination of protein BF04084: pH optimum for BF04084 activity was determined in the pH range 4 – 9. The activity assay was performed as described above, at 37°C using 5, 10 or 20 nM of BF04084 purified protein and *p*NP-β-D-galactofuranoside (*p*NP-β-D-gal*f*) as substrate at a final concentration of 1.5 mM. The buffer systems used are: 25 mM sodium acetate (pH 4 – 5.6), 25 mM MES buffer (pH 5.6 – 6.6), 25 mM Hepes buffer (pH 6.6 – 7.6) and 25 mM Tris/HCl (pH 7.6 – 9).

I.5.9. Steady state kinetics of BF04084 from *B. finegoldii* on *p*NP-β-D-gal*f*.

Reactions were carried out as described above for the *p*NP-β-D-gal*f* assay. Kinetic parameters were obtained at 37°C in 25 mM MOPS pH 7.2, using *p*NP-β-D-gal*f* concentrations in the range 0.05 - 6 mM. BF04084 was used at a final concentration of 0.1 µM. All kinetic parameters were determined by a non-linear regression curve using GraphPad Prism (GraphPad Software; www.graphpad.com).

I.5.10. Kinetic parameter determination of protein BF04084 from *B. finegoldii* on the galactobiose obtained from Galactan

The galactobiose oligosaccharide purified from Galactan, as described in paragraph 2.4.5., was used as substrate. A time course experiment was performed using three substrate concentrations. Reactions were carried out in 0.5 mL of 25 mM MOPS buffer pH 7 at 37 °C in the presence of 20 nM of BF04084, and using substrate concentrations ranging from 12.5 to 100 µM. 500 µM cellobiose was added in all reactions as an internal standard. 50µL aliquots of the reaction were collected after: t0, 2 min, 4 min, 8 min, 10 min, 12 min, 16 min and boiled at 100°C for 10 min to stop the reaction. Samples were analyzed using high performance anion exchange chromatography to determine galactose release and substrate depletion over time using a Dionex ICS3000 HPLC (ThermoScientific – Dionex) with a pulsed amperometric detector (HPAEC-PAD), and equipped with both a Carbopac column PA-1 anion exchange column (ThermoScientific – Dionex). Samples were loaded on the column and sugars were eluted isocratically at a flow rate of 0.25 mL/min for 20 min. Mobile phase was composed of 0.1 M NaOH in water (solvent A). Substrate depletion was observed over time and the galactobiose peak areas were measured for each sample and corrected using the area of the cellobiose internal standard. The k_{cat}/K_M were obtained by monitoring the substrate depletion rate at a range of substrate concentrations significantly below the K_M such that a linear plot of V_0 as a function of [S] was obtained. Thus, the substrate depletion rate obtained allowed to calculate k_{cat}/K_M. The equation $\ln(k_{cat}/K_M) = (S_0/S_t)/[E]$ was used, in which [E] and [S] are enzyme and substrate concentration, respectively.

I.5.11. Activity assays of recombinant proteins from *B. finegoldii* on natural substrates.

The activity of the recombinant proteins from *B. finegoldii* was evaluated on the arabinogalactan and galactan polysaccharides derived from *Mycobacterium smegmatis* cell wall and obtained in collaboration with Dr. Patrick Moynihan (School of Biosciences, University of Birmingham, UK). Hydrolysis reactions were performed using 1 mg/mL of either arabinogalactan or galactan in a final volume of 0.1 mL containing 25 mM MOPS buffer pH 7. Reactions were carried out overnight at 37 °C, in the presence of 1 µM of purified recombinant proteins either alone or in combination. Hydrolysis products were monitored by TLC and HPLC analysis.

TLC analysis: 6-10 µL of the reaction mixtures were spotted on a TLC silica plate and then run in a TLC solvent system of butanol/acetic acid/water buffer (2:1:1). The plates were dried and carbohydrates visualized by orcinol/sulphuric acid heated to 70 °C for 10 min. In all experiments, TLC standard solutions of pure reagents and products were used for comparison.

HPAEC-PAD analysis: The samples were analyzed using high performance anion exchange chromatography to determine galactose release and substrate depletion, using a Dionex ICS3000 HPLC (ThermoScientific – Dionex) with a pulsed amperometric detector (HPAEC-PAD), and equipped with both a Carbopac column PA-1 anion exchange column (ThermoScientific – Dionex). The samples were diluted 1:10 and loaded on the column and sugars were eluted isocratically at a flow rate of 0.25 mL/min for 20 min. The mobile phase was composed of 0.1 M NaOH in water (solvent A).

I.5.12. Activity assays of proteins BF03696 and BF04084 from *B. finegoldii* on synthetic substrates.

The activity of the recombinant proteins BF03696 and BF04084 was evaluated on 2 synthetic substrates: β(1,5)Gal*f*-ANS and β(1,6)Gal*f*-ANS obtained in collaboration with Dr. Patrick Moynihan (School of Biosciences, University of Birmingham, UK). Hydrolysis reactions were performed using 1 mM of either β(1,5)Gal*f*-ANS or β(1,6)Gal*f*-ANS in a final volume of 0.5 mL containing 25 mM MOPS buffer pH 7. Reactions were carried out at 37 °C, in the presence of 1.5 µM of purified recombinant proteins. A time course experiment was carried out and 50µL aliquots of the reactions were collected after

1min, 5min, 10min, 20min, 30min, 1hr, 2hr and O.N. The aliquots were boiled for 10 min at 100°C to stop the reactions, and then analyzed both at the epi-UV light and by TLC analysis following the protocol described in the previous paragraph.

I.6 Ferene S assay.

The Ferene S assay was used in order to verify the presence of the iron cofactor in the purified proteins, and its oxidation state. Ferene S, (3-(2-pyridyl)-5,6-difurylsulfonic acid-1,2,4-triazine sodium salt), is a complex heterocyclic compound containing 2 furyl groups substituted with sulfonate residues. This compound acts as a chelating agent of iron ions, turning into a product absorbing in the visible region, with λ_{max} at 593 nm, after the conjugation with Fe(II). A calibration curve was obtained performing the assay in presence of known amounts of $Fe(NH_4)_2(SO_4)_2$ and measuring the corresponding Abs at 593nm. Then, the assay was performed on the purified proteins. DTT and $Fe(NH_4)_2(SO_4)_2$ were removed from the buffer with a chromatographic step using a PD-10 column (Pharmacia). Then, aliquots of 450 µL of the pooled fractions were incubated for 10 minutes at RT in buffer containing 0.5% HCl. Then, TCA (trichloroacetic acid) was added to a final percentage of 2.6% and precipitated proteins were removed by centrifugation. Then, Ferene S was added to a final concentration of 1.6 mM in 3% sodium acetate, and the absorbance at 593 nm was measured. The concentration of total Fe(II) amount was obtained by determining the total concentration of the visible product using ε_{593nm}= 34.32 $mM^{-1}cm^{-1}$.
For each sample, a second aliquot was used for the determination of total iron concentration (Fe(II) + Fe(III)). In this case, a 6 mM final concentration of Vitamin C was added. This latter causes the complete reduction of Fe(III) to Fe(II), allowing the total amount of iron in solution to react with Ferene S. Again, the absorbance at 593 nm was measured, and the concentration of total iron amount was obtained by determining the total concentration of the visible product. The Total amount of Fe(III) was finally measured by subtracting the Fe(II) concentration from the total concentration obtained in this second experiment.

I.7 Enzymatic conversion of 4-OHE and 2-OHE using ERCD.

The enzymatic conversion of 4-OHE and 2-OHE was performed using the enzymes PP28735 and PP00124. Reactions were carried out at RT in 1 mL of 50 mM Tris/HCl, pH 7.5 containing either 100 µM of 4-OHE or 200 – 400 µM of 2-OHE. 260 mU/mL of unpurified ERCDs enzymes in soluble fractions were added to the sample to start the reaction. The product formation (λ_{max} 285 and 298 nm) was followed spectrophotometrically over 15 minutes, between 230 and 500 nm.

I.8 HPLC analysis of the reaction products of 4-OHE and 2-OHE enzymatic conversion.

Reactions with 4-OHE and 2-OHE as substrates were performed as outlined in the previous paragraph. At the end, reactions were stopped by the addition of 1% final concentration of formic acid, the semialdehydes spectra were measured in acidic pH and the samples were centrifuged at 22,100 x g for 10 minutes and analyzed by HPLC.

An aliquot of the reaction mixture was collected before the addition of the enzyme, and 1% final concentration of formic acid was added. The sample was centrifuged at 22,100 x g for 10 minutes and analyzed by HPLC as a negative control.

Samples were analyzed using a Waters 1525 binary pump HPLC with a photodiode array detector (Waters 2996), and equipped with a C18 ultrasphere column (Beckman Coulter 4.6 mm x 25 cm). 0.5 mL of the sample were injected. A linear gradient elution was performed at a flow rate of 1 mL/min, using a mobile phase composed of: 0.1% formic acid in water (solvent A) and 0.1% formic acid in methanol (solvent B).

The 4-OHE elution was obtained using the following gradient:

1) 0 min: 10% solvent B to 3 min: 10% solvent B.
2) 3 min: 10% solvent B to 23 min: 75% solvent B.
3) 23 min: 75% solvent B to 25 min: 98% solvent B.
4) 25 min: 98% solvent B to 30 min: 98% solvent B.

The retention times measured for the 4-OHE and its products were of 23.9 min (λ_{max}= 280,3 nm) and of 22.5 min (λ_{max}= 305,3 nm), respectively.

The 2-OHE elution was performed using the following gradient:

1) 0 min: 20% solvent B to 3 min: 20% solvent B.
2) 3 min: 20% solvent B to 28 min: 80% solvent B.
3) 28 min: 80% solvent B to 30 min: 98% solvent B.
4) 30 min: 98% solvent B to 40 min: 98% solvent B.

The retention times measured for the 2-OHE and its products were of 25.2 min (λ_{max}= 287,5 nm) and of 22.6 min (λ_{max}= 305,3 nm), respectively.

I.9 Preparation of apo RHA-P and recovery of the enzymatic activity.

Five aliquots of 1 µM RHA-P were incubated in 50 mM Tris/HCl buffer pH 7 containing 2 mM EDTA, in a 2,000 fold excess compared to protein concentration. Samples were incubated for 30 min in a thermocycler, using a temperature gradient of 1 °C/min increase from 25 °C to 38 °C, 30 sec at 38 °C and 1 °C/min decrease from 38 °C back to 25 °C. Samples were then centrifuged at 22,100 x g for 10 min at 4 °C, and the supernatants were collected in clean tubes.

In order to recover the enzymatic activity, 3 mM, 10 mM and 50 mM of either $CaCl_2$, or $MgCl_2$, or $MnCl_2$, or $ZnCl_2$ were separately added to aliquots of EDTA-treated RHA-P. Samples were incubated for 5 minutes at RT; afterwards, the same temperature gradient described above for the depletion was used. The *p*NPR assay was performed after each step of incubation, and the specific activity was evaluated.

I.10 Whole-cell enzymatic conversion of naringin using RHA-P.

E. coli BL21(DE3) cells freshly transformed with pET22b(+)/*rha-his* were inoculated from the transformation plate into 6 mL of LB/amp and incubated at 37 °C with constant shaking up to OD_{600}/mL ~ 0.5. The first preinoculum was diluted 1:10 in 2 sterile 50 mL Falcon tubes containing 15 mL of LB/amp and incubated with constant shaking at 37°C

up to OD_{600}/mL of 0.6 – 0.7. Expression of RHA-P was induced with 1 mM IPTG and growth was continued for 3 hrs with constant shaking at 23 °C. Then, cells were collected by centrifugation (5,524 x g for 15 min at 4 °C) and gently suspended in M9 minimal media containing 0.4% glucose, at a final concentration of 1.5 OD_{600}/mL. Naringin was added to the medium at a 2 mM final concentration, and the cell suspensions were incubated at 23 °C in constant shaking. A time course experiment was performed to evaluate the conversion of naringin over time. Aliquots of the samples were collected at different time intervals, (T0, T15min, T30min, T1h, T1h30min, T2h, T2h30min, TON) and centrifuged at 5,524 x g for 5 min at 4 °C. The supernatant was collected for each time and reactions were stopped by the addition of 0.1% final concentration of formic acid and analyzed by HPLC.
Aliquots of *E. coli* BL21(DE3) and *E. coli* BL21(DE3) transformed with an empty pET22b(+), were used as negative controls following the same experimental procedures.

I.11 HPLC analysis of the naringin enzymatic conversion.

The supernatant of *E. coli* BL21(DE3) expressing RHA-P M9 medium with 2 mM naringin was collected at each time point and analyzed by HPLC to verify the presence of naringin and prunin. The samples were diluted 1:20 in 0.1% formic acid and injected in the chromatographic system.
Samples were analyzed using a Waters 1525 binary pump HPLC with a photodiode array detector (Waters 2996), and equipped with a C18 Symmetry column (4.6 x 75 mm, particle size 3.5 µm and pore size 10 nm). 0.2 mL of the diluted sample were injected. A linear gradient elution was performed at a flow rate of 0.3 mL/min, using a mobile phase composed of: 0.1% formic acid in water (solvent A) and 0.1% formic acid in methanol (solvent B).
The naringin and prunin elution was performed using the following gradient:

1) 0 min: 10% solvent B to 5 min: 10% solvent B.
2) 5 min: 10% solvent B to 10 min: 33% solvent B.
3) 10 min: 33% solvent B to 50 min: 37% solvent B.
4) 50 min: 37% solvent B to 55 min: 95% solvent B.
5) 55 min: 95% solvent B to 60 min: 95% solvent B.

The retention times measured for the naringin and prunin were of 46 min (λ_{max}= 330 nm) and of 41.5 min (λ_{max}= 330 nm), respectively.

I.12 Growth of Bacteroides strains.

Aliquots of Bacteroides strains glycerol stocks were inoculated in 30 mL glass tubes containing 5 mL of either sterile tryptone yeast extract glucose (TYG) medium , or sterile Brain Heart Infusion broth (BHI – Sigma Aldrich), or sterile Clostridial nutrient media (Sigma Aldrich). Porcine hematin (1.2 mg/mL Sigma Aldrich) was added after sterilization to all the media. The cells were incubated O.N. at 37°C into an anaerobic cabinet (Whitley A35 Workstation; Don Whitley, UK). When growths with specific substrates are performed, a sterile minimal media is used [141], supplemented with the desired substrate as sole energy and carbon source. Growths are incubated at 37°C into an anaerobic cabinet (Whitley A35 Workstation; Don Whitley, UK).

I.13 RT-PCR and q-PCR of PULs 41 and 49 in *B. finegoldii*.

Comparison of the levels of the PULs 41 and 49 of *B. finegoldii* transcript expression was performed by qPCR. *B. finegoldii* was inoculated in 10 mL of Brain Heart Infusion broth (BHI) O.N. into an anaerobic cabinet at 37°C. Then, 0.2 mL of the overnight growth were inoculated in:

- 3 tubes containing Minimal media supplemented with 5 mg/mL glucose
- 3 tubes containing Minimal media supplemented with 5 mg/mL arabinogalactan from *M. smegmatis*.

Negative controls were performed without the addition of *B. finegoldii*.

Triplicate bacterial cultures were harvested at mid-log phase (0.8 OD_{600}/mL) and placed in RNAprotect (Qiagen), then stored at -80°C overnight, before purification with RNeasy kit (Qiagen). RNA purity was assessed spectrophotometrically, and 1 µg of RNA was used immediately for reverse transcription (QuantiTect Reverse Transcription kit, Qiagen). Quantitative RT-PCR was performed in a 96-well plate on a LightCycler 480 System (Roche) with FastStart Essential DNA Green Master (Roche) using gene specific primers for the 9 genes of interest in PULs 41 and 49 of *B. finegoldii* genome. Reactions were carried out in 10 mL, consisting of 5 mL SYBR Green mix, 20 total ng of cDNA and 1 µM or 0.125 µM (16 S rRNA) primer mix. Reaction conditions used were: 95 °C 600 sec, followed by 45 cycles of 95 °C for 10 sec, 55 °C for 10 sec, 72 °C for 10 sec. Cq values were calculated using LightCycler 480 SW 1.5. Data were normalized to 16 S rRNA transcript levels, and change in expression level calculated as fold-change compared with minimal media, glucose cultures.

I.14 General methods.

Protein concentration was measured with the Bio-Rad Protein System [159] using bovine serum albumin (BSA) as a standard. Polyacrylamide gel electrophoresis was carried out using standard techniques [160]. SDS–PAGE 15% Tris–glycine gels were run under denaturing conditions and proteins were stained with Coomassie brilliant blue G-250. “Wide range” (200–6.5 kDa) molecular weight standard was from Sigma (ColorBurst™ Electrophoresis Marker).

Chemical parameters determination of the proteins was conducted using the online software ProtParam (https://web.expasy.org/protparam/).

APPENDIX II

Scientific pubblications.

1. De Lise F., **Mensitieri F.**, Tarallo V., Ventimiglia N., Vinciguerra R., Tramice A., Marchetti R., Pizzo E., Notomista E., Cafaro V., Molinaro A., Birolo L., Di Donato A. and Izzo V. *RHA-P: isolation, expression and characterization of a novel α-L-rhamnosidase from Novosphingobium sp. PP1Y*. Journal of Molecular Catalysis B: Enzymatic 134 (2016) 136–147.

2. Pingeon M., Charlier B., De Lise F., **Mensitieri F**., Dal Piaz F., Izzo V. *Novel Drug Targets for the Treatment of Cardiac Diseases*. Current Pharmacogenomics and Personalized Medicine, 2017, 15, 1-13.

3. **Mensitieri F.**, De Lise F., Strazzulli A., Moracci M., Notomista E., Cafaro V., Bedini E., Sazinsky M.H., Trifuoggi M., Di Donato A., Izzo V. *Structural and functional insights into RHA-P, a bacterial GH106 α-L-rhamnosidase from Novosphingobium sp. PP1Y*. Archives of Biochemistry and Biophysics, 2018, 15, 648:1-11.

4. Petruk G., Roxo M., De Lise F., **Mensitieri F.**, Notomista E., Wink M., Izzo V., Monti DM. *The marine Gram-negative bacterium Novosphingobium sp. PP1Y as a potential source of novel metabolites with antioxidant activity*. Accepted in Biotechnology Letters, 2019. *doi.org/10.1007/s10529-018-02636-4.*

APPENDIX III

Communications.

1. Manzo V., Donadio G., Notomista E., Russomanno G., Sarcinelli C., De Lise F., **Mensitieri F.**, Ventimiglia N., Di Cristo C., Pizzo E., Pezzella A., Di Donato A., Izzo V. *Microbial Oxygenase Activities for the Biosynthesis of Novel Aromatic Antioxidant Compounds*. Giornate della Facoltà di Farmacia e Medicina a Salerno. Maggio 2014, Università degli Studi di Salerno. Translational Medicine @ UniSa (2014) - ISSN 2239-9747 2014, Special Issue (1): 9.

2. Di Donato A., De Lise F., **Mensitieri F.**, Donadio G., Tramice A., Trincone A., Cafaro V., Notomista E., Izzo V. *From Bioremediation to Biocatalysis: biotecnologically relevant enzymes from microorganisms adapted to polluted environments*. Japan-Italy Symposium on New Trends in Science and Engineering of Enzyme and Microbiology for Sustainable Society, Nara, Japan (2014). Oral communication.

3. De Lise F., **Mensitieri F.**, Rusciano G., Molinaro A., dal Piaz F., Di Cosmo A., Sasso A., Di Donato A., Izzo V. *Isolation and purification of extracellular nanostructures from N.sp.PP1Y: a novel example of Outer Membrane Vesicles?* “ISEV 2016”. Rotterdam, The Netherlands. May 2016.

4. **Mensitieri F.**, De Lise F., Strazzulli A., Del Sorbo M., Tramice A., Molinaro A., Birolo L., Moracci M., Di Donato A., Izzo V. *RHA-P: insight into a Novel Bacterial α-L-Rhamnosidase from Novosphingobium sp. PP1Y*. “CBM12”. Wien. April 2017.

5. **Mensitieri F.**, De Lise F., Cafaro V., Lumacone M., Strazzulli A., Notomista E., Moracci M., Di Donato A., Izzo V. *RHA-P: a Novel α-L-Rhamnosidase of Biotechnological Relevance from Novosphingobium sp. PP1Y*. “SIB 2017”. Caserta, Italy. September 2017.

6. De Lise F., **Mensitieri F.**, Castaldi S., Rusciano G., Dal Piaz F., Sasso A., Zarrelli A., Di Donato A., Izzo V. *Purification and Characterization of Extracellular*

Nanostructures from N. sp. PP1Y: a Novel Example of Outer Membrane Vesicles. “SIB 2017” Caserta, Italy. September 2017. Oral Communication.

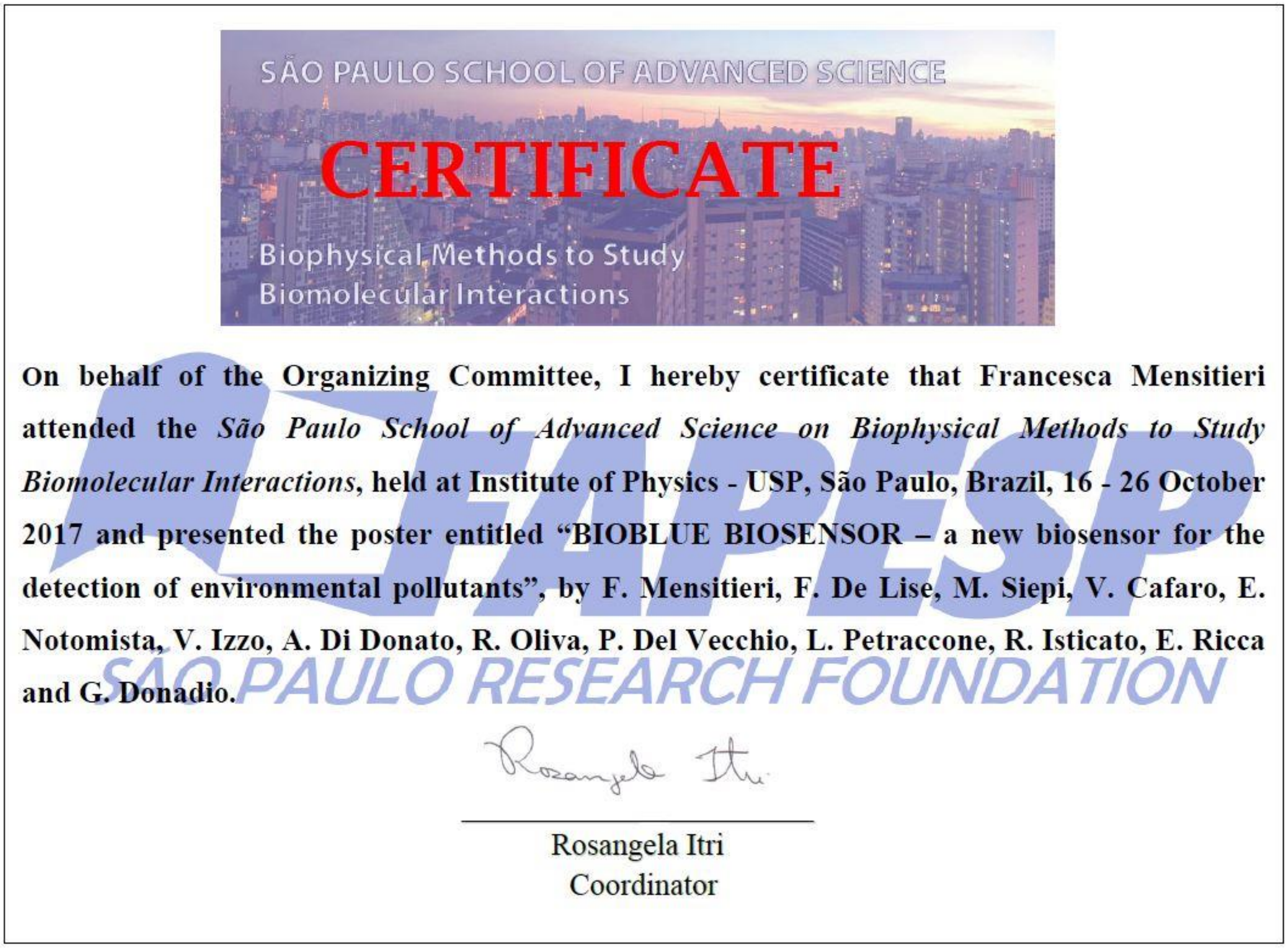
7. **Mensitieri F.**, De Lise F., Siepi M., Cafaro V., Notomista E., Izzo V., Di Donato A., Oliva R., Del Vecchio P., Petraccone L., Isticato R., Ricca E., Donadio G. *BIOBLUE BIOSENSOR – a new biosensor for the detection of environmental pollutants*. “Sao Paulo School of Advanced Science on Biophysical Methods to Study Biomolecular Interactions” Sao Paulo, Brasil. 16-27 October 2017.

8. De Lise F., **Mensitieri F.**, Izzo V., Notomista E., Dal Piaz F., Sasso A., Rusciano G., Zarrelli A., Di Donato A. *OMVs from Novosphingobium sp. PP1Y, a marine microorganism adapted to hydrocarbons polluted environment*. “Extremophiles 2018” Ischia, Italy. September 2018.

APPENDIX IV

Abstract scholarship winner for participation in the Workshop.

From 16th to 27th October 2017, participation to the "Sao Paulo School of Advanced Science - Biophysical Methods to Study Biomolecular Interactions" - ESPCA. Institute of Physics – University of Sao Paulo - Sao Paulo, Brasil.



SÃO PAULO SCHOOL OF ADVANCED SCIENCE

CERTIFICATE

Biophysical Methods to Study
Biomolecular Interactions

On behalf of the Organizing Committee, I hereby certificate that Francesca Mensitieri attended the *São Paulo School of Advanced Science on Biophysical Methods to Study Biomolecular Interactions*, held at Institute of Physics - USP, São Paulo, Brazil, 16 - 26 October 2017 and presented the poster entitled "BIOBLUE BIOSENSOR – a new biosensor for the detection of environmental pollutants", by F. Mensitieri, F. De Lise, M. Siepi, V. Cafaro, E. Notomista, V. Izzo, A. Di Donato, R. Oliva, P. Del Vecchio, L. Petraccone, R. Isticato, E. Ricca and G. Donadio.

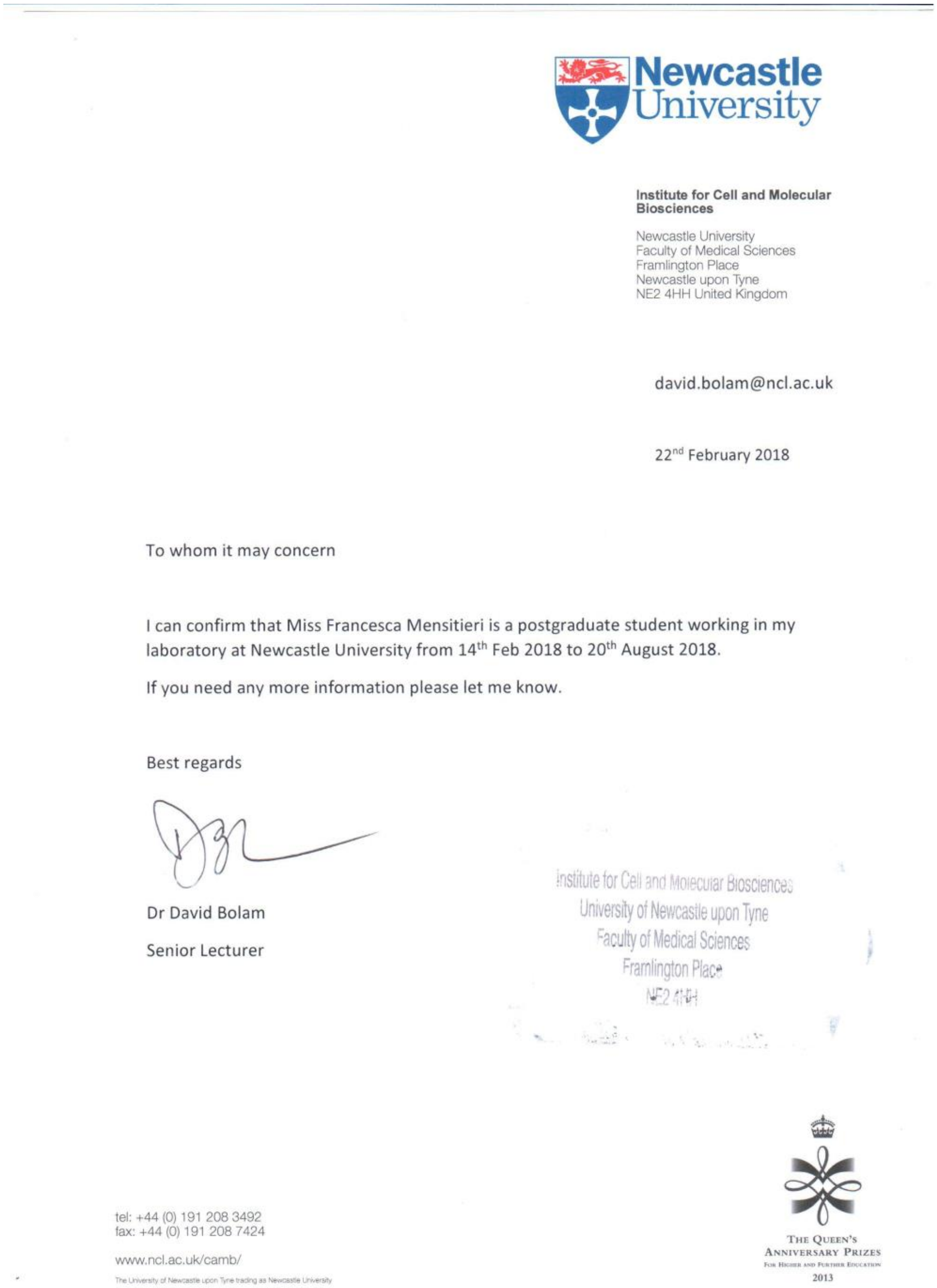
Rosangela Itri

Rosangela Itri
Coordinator

FAPESP
SÃO PAULO RESEARCH FOUNDATION

Experience in foreign laboratories

From February 14th to August 20th 2018, the research activity of Dr. Mensitieri has been carried out at the ICaMB (Institute for Cell and Molecular Biosciences), Newcastle University - Newcastle upon Tyne (UK), in the laboratory of Prof. Henry J. Gilbert, under the supervision of Dr. David Bolam.



Newcastle University

Institute for Cell and Molecular Biosciences

Newcastle University
Faculty of Medical Sciences
Framlington Place
Newcastle upon Tyne
NE2 4HH United Kingdom

david.bolam@ncl.ac.uk

22nd February 2018

To whom it may concern

I can confirm that Miss Francesca Mensitieri is a postgraduate student working in my laboratory at Newcastle University from 14th Feb 2018 to 20th August 2018.

If you need any more information please let me know.

Best regards

Dr David Bolam

Senior Lecturer

Institute for Cell and Molecular Biosciences
University of Newcastle upon Tyne
Faculty of Medical Sciences
Framlington Place
NE2 4HH

tel: +44 (0) 191 208 3492
fax: +44 (0) 191 208 7424

www.ncl.ac.uk/camb/

The University of Newcastle upon Tyne trading as Newcastle University

The Queen's Anniversary Prizes
For Higher and Further Education
2013

APPENDIX V

Published papers.

Journal of Molecular Catalysis B: Enzymatic 134 (2016) 136–147

Contents lists available at ScienceDirect

Journal of Molecular Catalysis B: Enzymatic

journal homepage: www.elsevier.com/locate/molcatb

RHA-P: Isolation, expression and characterization of a bacterial α-L-rhamnosidase from *Novosphingobium* sp. PP1Y

Federica De Lise[a,1], Francesca Mensitieri[a,1], Vincenzo Tarallo[b], Nicola Ventimiglia[a], Roberto Vinciguerra[b], Annabella Tramice[c], Roberta Marchetti[b], Elio Pizzo[a], Eugenio Notomista[a], Valeria Cafaro[a], Antonio Molinaro[b], Leila Birolo[b], Alberto Di Donato[a], Viviana Izzo[d,*]

[a] *Dipartimento di Biologia, Università Federico II di Napoli, Via Cinthia I, 80126 Napoli, Italy*
[b] *Dipartimento di Scienze Chimiche, Università Federico II di Napoli, via Cinthia I, 80126 Napoli, Italy*
[c] *Istituto di Chimica Biomolecolare, Consiglio Nazionale delle Ricerche, Via Campi Flegrei 34, 80072 Pozzuoli (NA), Italy*
[d] *Dipartimento di Medicina, Chirurgia e Odontoiatria "Scuola Medica Salernitana", Università degli Studi di Salerno, via S. Allende, 84081 Baronissi (SA), Italy*

ARTICLE INFO

Article history:
Received 29 July 2016
Received in revised form
23 September 2016
Accepted 3 October 2016
Available online 3 October 2016

Keywords:
α-L-Rhamnosidases
Flavonoids
Novosphingobium sp. PP1Y
Glycosyl hydrolases
Biotransformation

ABSTRACT

α-L-Rhamnosidases (α-RHAs) are a group of glycosyl hydrolases of biotechnological potential in industrial processes, which catalyze the hydrolysis of α-L-rhamnose terminal residues from several natural compounds. A novel α–RHA activity was identified in the crude extract of *Novosphingobium* sp. PP1Y, a marine bacterium able to grow on a wide range of aromatic polycyclic compounds. In this work, this α-RHA activity was isolated from the native microorganism and the corresponding *orf* was identified in the completely sequenced and annotated genome of strain PP1Y. The coding gene was expressed in *Escherichia coli*, strain BL21(DE3), and the recombinant protein, rRHA-P, was purified and characterized as an inverting monomeric glycosidase of ca. 120 kDa belonging to the GH106 family. A biochemical characterization of this enzyme using *p*NPR as substrate was performed, which showed that rRHA-P had a moderate tolerance to organic solvents, a significant thermal stability up to 45 °C and a catalytic efficiency, at pH 6.9, significantly higher than other bacterial α-RHAs described in literature. Moreover, rRHA-P was able to hydrolyze natural glycosylated flavonoids (naringin, rutin, neohesperidin dihydrochalcone) containing α-L-rhamnose bound to β-D-glucose with either α-1,2 or α-1,6 glycosidic linkages. Data presented in this manuscript strongly support the potential use of RHA-P as a biocatalyst for diverse biotechnological applications.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

α-L-Rhamnosidases (α-RHAs) are a group of glycosyl hydrolases (GHs) that catalyze the hydrolysis of terminal residues of α-L-rhamnose from a large number of natural compounds [1]. L-Rhamnose is widely distributed in plants as component of flavonoid glycosides, terpenyl glycosides, pigments, signaling molecules, and in cell walls as a component of complex heteropolysaccharides, such as rhamnogalacturonan and arabinogalactan-proteins [2–7]. In bacteria, L-rhamnose appears to be included in membrane rhamnolipids [8,9] and polysaccharides [10]. According to the similarities among their amino acidic sequences, α-RHAs are grouped in the CAZy (carbohydrate-active enzymes) database (www.cazy.org) into four different families: GH28, GH78, GH106, and NC (non-classified).

In the last decade, α-RHAs have attracted a great deal of attention due to their potential application as biocatalysts in a variety of industrial processes and in particular in the food industry [1]. Several dietary products are rich in glycosylated flavonoids that show the presence of either a rutinoside (6-α-L-rhamnosyl-β-D-glucose)

Abbreviations: α-RHAs, α-L-rhamnosidases; GHs, glycosyl hydrolases; GTs, glycosyl transferases; PPMM, potassium phosphate minimal medium; *p*NPR, *p*-nitrophenyl-α-L-rhamnopyranoside; MOPS, 3-(*N*-morpholino) propanesulfonic acid; *p*NP, *p*-nitrophenolate; BSA, bovine serum albumin; LB, Luria Bertani medium; LB-N, Luria Bertani medium containing a final concentration of 0.5 M NaCl; LB-BS, Luria Bertani medium supplemented with 1 mM of both betaine and sorbitol; LB-NBS, Luria Bertani medium containing a final concentration of 0.5 M NaCl and 1 mM of both betaine and sorbitol; IPTG, Isopropyl β-D-1-thiogalactopyranoside.

* Corresponding author.
E-mail address: vizzo@unisa.it (V. Izzo).

[1] These authors equally contributed to the work.

http://dx.doi.org/10.1016/j.molcatb.2016.10.002
1381-1177/© 2016 Elsevier B.V. All rights reserved.

or a neohesperidoside (2-α-L-rhamnosyl-β-D-glucose) disaccharidic unit. In particular, naringin, hesperidin and rutin, flavanone glycosides found in grapefruit juices, lemons, sweet oranges and vegetables, have gained increasing recognition for their potential antioxidant, antitumor and anti-inflammatory properties [11–14].

The ability to hydrolyze glycosylated flavonoids has been used to mitigate the bitterness of citrus juices, which is primarily caused by naringin. Rhamnose removal from naringin allows softening the bitter taste of citrus juice [15,16]. Moreover, the corresponding de-rhamnosylated compound, prunin, is endowed with antimicrobial properties [14], and shows an improved intestinal assimilation when compared to naringin. Other applications of α-RHAs are gaining popularity in the oenological industry, where these enzymes are used to hydrolyze terpenyl glycosides and enhance aroma in wine, grape juices and derived beverages [17–19].

Application of α-RHAs to improve flavonoids bioavailability has also been recently described [20]. In humans, flavonoids absorption occurs primarily in the small intestine where the attached glucose (or possibly arabinose or xylose) is removed by endogenous β-glucosidases [21–23]. Terminal rhamnose is not a suitable substrate for human β-glucosidases. Therefore, unabsorbed rhamnosylated flavonoids arrive unmodified in the colon, where they are hydrolyzed by α-rhamnosidase activities expressed by the local microflora [24]. To improve intestinal absorption of rhamnosylated flavonoids, and thus their bioavailability in humans, a removal of the terminal rhamnose group catalyzed by α-RHAs would be indeed beneficial [25–27].

The absence of human α-RHAs has been the key to the development of a novel targeted drug delivery strategy, indicated as LEAPT (Lectin-directed enzyme activated prodrug therapy) [28,29], a bipartite system based on the internalization of an engineered α-RHA bearing a glycosidic moiety that is recognized by specific lectins present on the surface of different eukaryotic cell lines. In the LEAPT system, the intake of a rhamnosylated prodrug, which cannot be processed by mammalian enzymes, allows a site-selective action of the drug in cells where the engineered α-RHA has been prelocalized.

To date, microbial α-RHAs have been mainly purified from fungal strains such as *Penicillium* and *Aspergillus* [30–32]; only one example of α-RHA isolated from a viral source has been described [33], and it is noteworthy that only a limited number of bacterial rhamnosidases has been fully characterized [34–39]. One of the main differences between fungal and bacterial α-RHAs is their different optimal pH values, with the fungal enzymes showing more acidic pH optima when compared to the bacterial counterparts, for which neutral and alkaline values have generally been described. This characteristic suggests diverse applications for fungal and bacterial enzymes, making bacterial α-RHAs suitable in biotechnological processes requiring good activity in more basic solutions such as, for example, the production of L-rhamnose from the hydrolysis of naringin or hesperidin, flavonoids whose solubility strongly increases at higher pH values [40]. In addition, bacterial rhamnosidases could be ideal candidates for dietary supplements having activity across the entire gastrointestinal (GI) tract, and more specifically in the small intestine where flavonoids absorption should mostly occur to enhance the beneficial effect of these molecules on human health.

In order to elucidate the real biotechnological potential of bacterial α-RHAs, further investigation is undoubtedly needed. Few details on the catalytic mechanism of bacterial α-RHAs are available, and most importantly, to the best of our knowledge, no attempt to improve the catalytic efficiency or modify substrate specificity of these enzymes by mutagenesis has been performed yet. This is in part consequential to the fact that only a very limited number of crystal structures of α-RHAs are currently available, such as the α-L-rhamnosidase B (BsRhaB) from *Bacillus* sp. GL1 [41], and the α-L-rhamnosidase from *Streptomyces avermitilis* [42]. Therefore, it is evident that bacterial α-RHAs represent a yet unexplored reservoir of potential biocatalysts for which more functional and structural data are required.

A member of the order of the Sphingomonadales, recently isolated and microbiologically characterized, *Novosphingobium* sp. PP1Y [43,44], appears to be a valuable source for the isolation of α-RHA activities. Sphingomonadales are a group of Gram-negative α-proteobacteria whose genomes show the presence of a great abundance of both glycosyl hydrolases (GHs) and glycosyltransferases (GTs). These activities are probably involved in the biosynthesis of complex extracellular polysaccharides and microbial biofilms [45]. The interest for *Novosphingobium* sp. PP1Y carbohydrate-active enzymes has grown as the sequencing and annotation of the whole genome, recently completed, allowed the identification of a great number of genes encoding for both GHs (53 *orfs*) and GTs (57 *orfs*) [46].

Recently, a α-RHA activity in *Novosphingobium* sp. PP1Y crude extract was described, which showed an alkaline pH optimum and a moderate tolerance to organic solvents [47]. *Novosphingobium* sp. PP1Y crude extract expressing this enzymatic activity was used for the bioconversion of naringin, rutin and hesperidin. Based on these preliminary results, a more detailed biochemical characterization of the α-RHA activity was essential.

In this work, the isolation, recombinant expression and partial characterization of a α-RHA from *Novosphingobium* sp. PP1Y is reported. This enzyme, named RHA-P, belongs to the GH106 family [48], a subgroup for which, according to our knowledge, no crystal structure is available yet. As evident from our analyses, RHA-P is a promising candidate for several biotechnological applications.

2. Materials and methods

2.1. Generals

General molecular biology techniques were performed according to Sambrook et al. [49]. Bacterial growth was followed by measuring the optical density at 600 nm (OD_{600}). pET22b(+) expression vector and *E. coli* strain BL21(DE3) were from Amersham Biosciences; *E. coli* strain Top10 was purchased from Life Technologies. *N.* sp. PP1Y was isolated from polluted seawater in the harbor of Pozzuoli (Naples, Italy) as previously described [43].

The thermostable recombinant DNA polymerase used for PCR amplification was TAQ Polymerase from Microtech Research Products. dNTPs, T4 DNA ligase, and the Wizard PCR Preps DNA purification system for elution of DNA fragments from agarose gels were purchased from Promega. The QIAprep Spin Miniprep Kit for plasmid DNA purification was from QIAGEN. Enzymes and other reagents for DNA manipulation were from New England Biolabs. Oligonucleotides synthesis and DNA sequencing were performed by MWG-Biotech. N-terminus of rRHA-P was sequenced by Proteome Factory AG. The presence and location of a potential signal peptide cleavage site on the amino acidic sequence of RHA-P was analyzed using SignalP 4.1 server (http://www.cbs.dtu.dk/services/SignalP).

Q-Sepharose Fast Flow and *p*-nitrophenyl-α-L-rhamnopyranoside (*p*NPR) were from Sigma Aldrich; Sephacryl S200 High Resolution was purchased from Amersham Biosciences. IPTG (isopropyl β-D-1-thiogalactopyranoside) was obtained from Applichem.

Solvents used in enzymatic assays were either from Applichem (DMSO) or from Romil (acetone). TLC silica gel plates were from E. Merck (Darmstadt, Germany).

2.2. *Growth of* Novosphingobium *sp. PP1Y cells*

N. sp. PP1Y cells were grown in Potassium Phosphate Minimal Medium (PPMM) at 30 °C for 28 h under orbital shaking at 220 rpm. A pre-inoculum in LB was prepared by transferring 50 μL from a glycerol stock stored at −80 °C to a 50 mL Falcon tube containing 12.5 mL of sterile LB. The pre-inoculum was allowed to grow at 30 °C O/N under orbital shaking and then used to inoculate 1 L of PPMM at an initial cell concentration of 0.01- 0.02 OD_{600}. 0.3 mM naringine was added to the medium and used as an inducer of the α-RHA activity, as previously described [47].

2.3. *Cloning of* orfPP1Y_RS05470 *and construction of the* pET22b(+)/rha-p *expression vector*

Genomic DNA was extracted from a 50 mL saturated culture of *N.* sp. PP1Y as described elsewhere [44]. *Orf*PP1Y_RS05470 coding for the α-RHA activity was amplified in two contiguous fragments, owing to the considerable length of the *orf* (3441 bp). The first fragment, named *rha-up* (1816 bp), was amplified using an internal downstream primer, RHA-Intdw (5′-AGGCGGCCATGGGAATGT-3′), which included an internal *Nco*I site already present in *orf*PP1Y_RS05470, and an upstream primer, RHA-up (5′-GGGAATTCCATATGCCGCGCCTTTCGCT-3′), designed to add a *Nde*I restriction site at 5′ of *orf*PP1Y_RS05470. The second half of the gene, named *rha-dw* (1625 bp), was amplified using the upstream primer RHA-Intup (5′-ACATTCCCATGGCCGCCT-3′), complementary to RHA-Intdw, and the downstream primer RHA-dw (5′-AAAACCGAGCTCTCAATGCCCGCCCGTG-3′) that was intended to incorporate a *Sac*I restriction site downstream of the amplified *orf.*

The amplified fragments, *rha-up* and *rha-dw,* were purified from agarose gel, digested, respectively, with *Nde*I/*Nco*I and *Nco*I/*Sac*I, and individually cloned in pET22b(+) vector previously cut with the same enzymes.

Ligated vectors were used to transform *E. coli*, strain Top10, competent cells. The resulting recombinant plasmids, named pET22b(+)/*rha-up* and pET22b(+)/*rha-dw*, were verified by DNA sequencing. Next, the construction of complete *rha-p* gene in pET22b(+) was performed. First, both pET22b(+)/*rha-dw* and pET22b(+)/*rha-up* were digested with *Nco*I/*Sac*I restriction endonucleases to obtain, respectively, fragment *rha-dw* and linearized pET22b(+)/*rha-up*.

Digestion products were purified from agarose gel electrophoresis, eluted and ligated. Ligation products were used to transform *E. coli* Top10 competent cells and the resulting recombinant plasmid, named pET22b(+)/*rha-p* was verified by DNA sequencing.

2.4. *α-L-Rhamnosidase recombinant expression*

Protein expression was carried out in *E. coli* BL21(DE3) strain transformed with pET22b(+)/*rha-p* plasmid.

All the media described in this paragraph contained 100 μg/mL of ampicillin.

2.4.1. *Analytical expression*

E. coli BL21(DE3) competent cells transformed with plasmid pET22b (+)/*rha-p* were inoculated in a sterile 50 mL Falcon tube containing 12.5 mL of either LB [50] or LB containing a final concentration of 0.5 M NaCl (LB-N). Cells were grown under constant shaking at 37 °C up to 0.6–0.7 OD_{600}. This preinoculum was diluted 1:100 in 12.5 mL of either one of the four following media: LB, LB-N, LB supplemented with 1 mM of both betaine and sorbitol (LB-BS), or LB containing a final concentration of 0.5 M NaCl and 1 mM of both betaine and sorbitol (LB-NBS). Cells were grown under constant shaking at 37 °C up to 0.7–0.8 OD_{600}. RHA-P recombinant expression was induced with 0.1 mM IPTG at either 23 °C or 37 °C; growth was continued in constant shaking for 3 h. Cells were collected by centrifugation (5,524 × *g* for 15 min at 4 °C) and suspended in 25 mM MOPS pH 6.9 at a final concentration of 14 OD_{600}. Cells were disrupted by sonication (12 times for a 1-min cycle, on ice) and an aliquot of each lysate was centrifuged at 22,100 × *g* for 10 min at 4 °C. Both soluble and insoluble fractions were analyzed by SDS-PAGE. The soluble fraction was assayed for the presence of α-RHA enzymatic activity.

2.4.2. *Large scale expression*

Fresh transformed cells were inoculated into 10 mL of LB-N and incubated in constant shaking at 37 °C O/N. The preinoculum was diluted 1:100 in four 2 L Erlenmeyer flasks containing each 500 mL of LB-NBS and incubated in constant shaking at 37 °C up to 0.7–0.8 OD_{600}.

Expression of the recombinant protein, named rRHA-P, was induced with 0.1 mM IPTG and growth was continued for 3 h at 23 °C. Cells were collected by centrifugation (5,524 × *g* for 15 min at 4 °C) and stored at −80 °C until needed.

2.5. *Native and recombinant α-L-rhamnosidase purification*

Both native and recombinant RHA-P were purified following three chromatographic steps. Cell paste was suspended in 25 mM MOPS pH 6.9, 5% glycerol (buffer A), at a final concentration of 100 OD_{600} and cells were disrupted by sonication (10 times for a 1-min cycle, on ice). Cell debris were removed by centrifugation at 22,100 × *g* for 60 min at 4 °C and the supernatant was collected and filtered through a 0.45 μm PVDF Millipore membrane.

Afterwards, cell extract was loaded onto a Q Sepharose FF column (30 mL) equilibrated in buffer A. The column was washed with 50 mL of buffer A, after which bound proteins were eluted by using a 300 mL linear gradient of buffer A from 0 to 0.4 M NaCl at a flow rate of 15 mL/h. The chromatogram was obtained by analyzing fractions absorbance at λ = 280 nm and the presence of the recombinant α-RHA activity was detected using the *p*NPR assay. Relevant fractions were analyzed by SDS-PAGE, pooled and concentrated at a final volume of ~ 0.5 mL using a 30 kDa Amicon ultra membrane, Millipore. The sample was then loaded onto a Sephacryl HR S200 equilibrated with buffer A containing 0.2 M NaCl (buffer B).

Proteins were eluted from the gel filtration column with 250 mL of buffer B at a flow rate of 12 mL/h. Fractions were collected, analyzed and screened for the presence of α-RHA activity as previously described. At this stage, NaCl was removed from pooled fractions by repeated cycles of ultrafiltration and dilution with buffer A. The sample was then loaded on a Q Sepharose FF column (30 mL) equilibrated in buffer A. The column was washed with 50 mL of buffer A, after which bound proteins were eluted with 300 mL of a linear gradient of buffer A from 0 to 0.25 M NaCl at a flow rate of 13 mL/h. Fractions were collected, analyzed by SDS-PAGE and screened for the presence of α-RHA activity. Relevant fractions were pooled, concentrated, purged with nitrogen, and stored at −80 °C until use.

2.6. *Analytical gel filtration*

Analytical gel-filtration experiments were carried out as follows: 100 μL of a 40 pM protein sample was loaded on a Superdex 200 HR 10/300 column previously equilibrated in buffer B, installed on an AKTA^{TM}FPLC^{TM} (GE Healthcare Life Science). Samples were eluted isocratically at RT at a flow rate of 0.5 mL/min. Protein elution was monitored at λ = 280 nm. A molecular weight calibration was performed in the same buffer with the following proteins of known molecular weight: β-amylase (200 kDa), glyceraldehyde-3-phosphate dehydrogenase (143 kDa), carbonic anhydrase (29 kDa).

2.7. Enzyme activity assays

α-RHA activity was determined using *p*NPR as substrate (*p*NPR assay). Otherwise stated, all activity assays were performed at RT. The reaction mixture contained, in a final volume of 0.5 mL of 50 mM MOPS pH 6.9, *p*NPR at a final concentration of 600 μM and variable amounts of the sample tested. The reaction was blocked after 10 and 20 min by adding 0.5 M Na_2CO_3; the product, *p*-nitrophenolate (*p*NP), was detected spectrophotometrically at λ = 405 nm. The extinction coefficient used was $\varepsilon_{405} = 18.2\ mM^{-1}\ cm^{-1}$. One unit of enzyme activity was defined as the amount of the enzyme that releases one micromole of *p*NP per min.

2.7.1. Kinetic parameters determination

Kinetic parameters were obtained at pH 6.9 using a *p*NPR concentration in the range 0.025–1 mM. All kinetic parameters were determined by a non-linear regression curve using GraphPad Prism (GraphPad Software; www.graphpad.com).

2.7.2. pH optimum

pH optimum for α-RHA activity was determined in the range 4.7–8.8. Enzyme assays were performed as described above, using the following buffers: 50 mM potassium acetate (pH 4.7–5.7), 50 mM MOPS/NaOH (pH 5.7–7.7) and 50 mM Tris/HCl (pH 7.7–8.8).

2.7.3. Temperature optimum and stability

Optimum temperature was evaluated by performing the standard *p*NPR assay and incubating the reaction mixture at different temperatures, in the range 25–55 °C.

The thermal stability of the enzyme was determined by incubating the enzyme for one and 3 h at 30, 40, 50, 60 °C and measuring, after each incubation, the residual specific activity.

2.7.4. Organic solvents tolerance

The tolerance of the enzymatic activity to the presence of organic solvents in the reaction mixture, such as DMSO, acetone or ethanol, was evaluated by performing the standard *p*NPR assay in 50 mM MOPS pH 6.9 to which either 10% or 50% of solvent was added.

2.8. Substrate specificity

Reactions were carried out in 0.6 mL of 50 mM Na-phosphate buffer pH 7.0 under magnetic stirring at 40 °C in the presence of 20 mM of either aryl glycoside and 0.25 U of rRHA-P. Reactions were monitored over time (0–24 h) by TLC analysis (system solvent: EtOAc:MeOH:H_2O 70:20:10). Compounds on TLC plates were visualized under UV light or charring with α-naphthol reagent.

Hydrolysis reactions of maltose, pullulan, starch, amylopectin, sucrose, raffinose, lactose, xylan from birchwood, xylan from oat spelt, hyaluronic acid, α − cellulose, cellobiose, chitosan, β-Glucan from barley, laminarin, curdlan, fucoidan from *Fucus versiculosus*, rhamnogalacturonan, rutinose were performed using 2.5 mg of each substrate, which was suspended in 0.5 mL of 50 mM Na-phosphate buffer pH 7.0. The reaction was carried out at 40 °C under magnetic stirring, in the presence of 0.25 U of rRHA-P. Hydrolysis products were monitored by TLC analysis (system solvent HCOOH:HAc:H_2O:2-propanol: EtOAc:1:10:15:5:25).

Flavonoidic compounds such as naringin, rutin, neohesperidin dihydrochalcone, were screened as possible substrates. In this case, a 6 mM solution of each compound in a final volume of 1 mL of 50 mM Na-phosphate buffer pH 7 was incubated at 40 °C in the presence of 0.25 U of rRHA-P. Reactions were checked over time by TLC analysis (t = 0, 15′, 30′, 60′, 90′, 120′, 150′, 180′, 24 h) with the following solvent system: EtOAc:MeOH:H_2O 70:20:10. An additional hydrolysis reaction of naringin was performed in conditions similar to those reported above in the presence of 10% DMSO, and the reaction was monitored over time.

In all experiments, TLC standard solutions of pure reagents and products were used for comparison.

2.9. Mass spectrometry analysis

Identification of native and recombinant RHA-P by mass spectrometry was performed by enzymatic digestion, either in solution or *in situ* after separation by SDS-PAGE.

When digesting in solution, the protein sample was lyophilized and then dissolved in denaturing buffer (300 mM Tris/HCl pH 8.8, 6 M urea, 10 mM EDTA). Cysteines were reduced with 10 mM dithiothreitol in denaturing buffer at 37 °C for 2 h, and then carbamidomethylated with 55 mM iodoacetamide dissolved in the same buffer at RT for 30 min, in the dark. Protein sample was desalted by size exclusion chromatography on a Sephadex G–25 M column (GE Healthcare, Uppsala, Sweden). Fractions containing RHA-P were lyophilized and then dissolved in 10 mM NH_4HCO_3 buffer (pH 8.0). Enzyme digestion was performed using proteases with different specificity such as trypsin, chymotrypsin and endoproteinase GluC (V8-DE) (Sigma), using a protease/RHA-P ratio of 1:50 (w/w) at 37 °C for 16 h.

For the enzymatic digestion *in situ*, Coomassie blue stained bands excised from polyacrylamide gels were destained by several washes with 0.1 M NH_4HCO_3 pH 7.5 and acetonitrile. Cysteines were reduced for 45 min in 100 μL of 0.1 M NH_4HCO_3, 10 mM dithiothreitol, pH 7.5. Carbamidomethylation of thiols was achieved after 30 min in the dark by adding 100 μL of 55 mM iodoacetamide dissolved in the same buffer. Enzymatic digestion was carried out on different bands of RHA-P by using 100 ng of proteases with different specificity (same as above) in 10 mM NH_4HCO_3 buffer, pH 7.5. Samples were incubated for 2 h at 4 °C. Afterwards, the enzyme solution was removed and a fresh protease solution was added; samples were incubated for 18 h at 37 °C. Peptides were extracted by washing the gel spots/bands with acetonitrile and 0.1% formic acid at RT, and were filtered using 0.22 um PVDF filters (Millipore) following the recommended procedures.

Peptide mixtures were analyzed either by matrix-assisted laser-desorption/ionization mass spectrometry (MALDI-MS) or capillary liquid chromatography with tandem mass spectrometry detection (LC–MS/MS).

MALDI-MS experiments were performed on a 4800 Plus MALDI TOF/TOF mass spectrometer (Applied Biosystems, Framingham, MA) equipped with a nitrogen laser (337 nm). The peptide mixture (1 μL) was diluted (1:1, v/v) in acetonitrile/50 mM citrate buffer (70:30 v/v) containing 10 mg/mL of α-cyano-4-hydroxycinnamic acid. Mass calibration was performed using external peptide standards purchased from Applied Biosystems.

Spectra were acquired using a mass (*m*/*z*) range of 300–4000 amu and raw data were analyzed using Data Explorer Software provided by the manufacturer. Experimental mass values were compared with calculated masses derived from an *in silico* digestion of RHA-P obtained with different proteases, using MS-Digest, a proteomics tool from Protein Prospector software (prospector.ucsf.edu/).

Peptide mixtures were analyzed by nanoLC–MS/MS on a CHIP MS 6520 QTOF equipped with a capillary 1200 HPLC system and a chip cube (Agilent Technologies, Palo Alto, Ca). After sample loading, the peptide mixture was first concentrated and then washed at 4 nL/min in a 40 nL enrichment column (Agilent Technologies chip) with 2% acetonitrile + 0.1% formic acid.

The sample was then fractionated on a C18 reverse-phase capillary column at a flow rate of 400 nL/min by using a two-solvents system consisting of 2% acetonitrile + 0.1% formic acid (solvent A) and 95% acetonitrile + 0.1% formic acid (solvent B). Separation was carried out with a 50 min linear gradient from 5 to 60% of solvent B.

Peptide analysis was performed using data-dependent acquisition of one MS scan (mass range from 300 to 2,000 m/z) followed by MS/MS scans of the five most abundant ions in each MS scan.

MS/MS spectra were acquired automatically when the MS signal detected upraised the threshold of 50,000 counts. Double and triple charged ions were isolated and fragmented. The acquired MS/MS spectra were transformed in Mascot Generic format (.mgf) and used to query either a NCBInr 02-2015 (21,322,359,704 amino acid sequences; 10,835,265,410 residues) or a *N.* sp. PP1Y (4635 aminoacid sequences; 1,493,994 residues) sequence database.

A licensed version (2.4.0.) of MASCOT software (www.matrixscience.com) was used with the following search parameters: trypsin, chymotrypsin or GluC endoproteinase as enzyme; 3 or 2 as allowed number of missed cleavages; carbamidomethylation of cysteines as fixed modification; 10 ppm MS tolerance and 0.6 Da MS/MS tolerance; peptide charge from +2 to +3. Oxidation of methionine and formation of pyroglutamic acid from glutamine residues at the N-terminal position of peptides were considered as variable modifications. Ions score is −10*Log(P), where P is the probability that the observed match is a random event. Individual ions score >14 indicate identity or extensive homology ($p < 0.05$). Protein scores derive from ion scores as a non-probabilistic basis for ranking protein hits. Individual ion score threshold provided by MASCOT software, necessary to evaluate the quality of matches in MS/MS data and to calculate protein score, was 52 in NCBInr database and 14 in PP1Y Database. In Table 1 only peptides with ion score ≥ 10 are reported.

2.10. NMR Analysis

All the experiments were performed at 298 K on a Bruker 600 MHz DRX spectrometer equipped with a cryo probe. *p*NPR was dissolved in D_2O to a final concentration of 3 mM in a 5-mm NMR tube. The ligand proton resonances were assigned by a combination of 1D and 2D NMR experiments including COSY, TOCSY and HSQC. ^{1}H NMR spectra were acquired with 32 k data points. Double quantum-filtered phase sensitive COSY and TOCSY experiments were performed by using data sets of 2096 × 256 (t1 x t2) points. Total Correlation Spectroscopy (TOCSY) spectrum was recorded with a spin lock time of 100 ms. Heteronuclear single quantum coherence (HSQC) experiment was measured in the ^{1}H-detected mode via single quantum coherence with proton decoupling in the ^{13}C domain, by using data sets of 2048 × 256 points. Experiments were carried out in the phase-sensitive mode according to the method of States and coworkers [51]. In all homonuclear spectra the data matrix was zero-filled in the F1 dimension to give a matrix of 4096 × 2048 points and resolution was enhanced in both dimensions by a cosine-bell function before Fourier transformation.

After the acquisition of the initial spectra of the ligand alone, rRHA-P was added to a final concentration of ~1 μM (~60 μg) in the NMR tube from a ~5.9 μM solution in 25 mM MOPS containing 5% glycerol. The NMR sample was immediately placed in the spectrometer probe to record experiments on the mixture.

Data acquisition and processing were performed with TOPSPIN 2.1 software.

2.10. Analytical methods

Protein concentration was measured using the Bio-Rad Protein System [52] using bovine serum albumin (BSA) as standard. Polyacrylamide gel electrophoresis was carried out using standard techniques [53]. SDS–PAGE 15% Tris–glycine gels were run under denaturing conditions and proteins were stained with Coomassie brilliant blue G-250. "Wide range" (200–6.5 kDa) molecular weight standard was from Sigma (ColorBurst^{TM} Electrophoresis Marker).

3. Results

3.1. Isolation and identification of a α-RHA from *Novosphingobium sp. PP1Y*

In a previous work [47], a α-RHA activity was detected in the crude extract of *N.* sp. PP1Y grown in minimal medium and in the presence of 0.3 mM naringin. In order to identify and characterize this enzymatic activity, *N.* sp. PP1Y cells were grown under the same experimental conditions described by the authors [47]. Cells were collected by centrifugation after 24 h of growth in PPMM (optical density of ~1 OD_{600}) at 5,524 × *g* for 15 min at 4 °C. α-RHA purification was carried out following the procedures described in this work in Materials and Methods. In all three chromatographic steps used for the purification of *N.* sp. PP1Y crude extract, a unique peak of α-RHA activity was always detected, but no major protein band was ever evident from the SDS-PAGE analysis of the corresponding active fractions (data not shown).

Peak fractions obtained from the first two chromatographic steps were digested with trypsin and analyzed by LC–MS/MS analysis. Raw data were used to search NCBI database (Bacteria as taxonomy restriction) with the MS/MS ion search program of Mascot software. Among the identified proteins, a member of the glycoside hydrolase family (*Novosphingobium* sp. PP1Y:WP_013837086.1) was identified with 10 (20% of sequence coverage) and 32 peptides (46% of sequence coverage) in the first and in the second chromatographic step, respectively (data not shown). The putative MW, deduced from the amino acid sequence, was of 124,225 Da.

Peak fractions obtained from the last chromatographic step (Q-sepharose) showed the presence of a relatively limited number of protein bands (Fig.S1, Supplementary material). Five bands migrating as expected for proteins of MW higher than 100 kDa were selected and excised from the polyacrylamide gel stained with Colloidal Coomassie Blue. Bands were digested *in situ* with trypsin, and the resulting peptides mixture was analyzed by LC–MS/MS. Raw data were used to search NCBI database (Bacteria as taxonomy restriction) with MASCOT software.

One of the protein bands analyzed, indicated in Fig. S1 by a red arrow, was identified (52 peptides, 71% sequence coverage, Table 1) as the protein encoded by the *orf*PP1Y_RS05470; this latter is located in one of the extrachromosomal elements of *N.* sp. PP1Y, the Megaplasmid referred to as Mpl [44]. This protein, from now on indicated as RHA-P, had been previously annotated as a member of the GHs family and is composed by 1,146 aa for a calculated molecular weight of about 124,225 Da.

3.2. Cloning, recombinant expression and purification of recombinant RHA-P

The recombinant plasmid pET22b(+)/*rha-p* (Materials and Methods), in which *orf* PP1Y_RS05470 was cloned, was used to transform *E. coli* BL21(DE3) strain. First attempts of expression of the recombinant protein at 37 °C showed the presence of a protein band with the expected molecular weight of recombinant RHA-P (rRHA-P) almost exclusively in the insoluble fraction of the induced culture (Fig. S2). As evident from Fig. S2 of the Supplementary material, when *E. coli* was instead induced at 23 °C [54], the presence of a protein band with the expected MW for rRHA-P in the soluble fraction (lane 5) appeared to be markedly increased when compared to the soluble fraction obtained after an induction performed

Table 1
Identification of native RHA-P by *in situ* digestion followed by LC–MS/MS analysis.

Protein [Species] NCBInr Accession number	Mass	Protein score	Sequence coverage (%)	n° of peptides	Ion score[a] and peptide sequence
Glycoside hydrolase family protein [*Novosphingobium* sp. PP1Y] gi|334145605|ref| WP_013837086.1|	124,225	2318	71	52	34 R.TGDAIK.A 31 R.VGGFYGK.G 38 R.GVVQNEK.R 28 R.LDNELVR.G 21 R.GPEIEPVK.A 11 R.VLPITGRK.A 19KVGSKNEPR.G 37 R.AVAEAAFDK.V 42 R.GVVQNEKR.Q 48KLSSPNAMALR.A 13KLSSPNAMALR.A+Oxidation (M) 37 R.DITDQVPVLR.L 13 R.ITVVGDITMDR.L 41 R.ITVVGDITMDR.L+Oxidation (M) 32 R.LGRPEDYYVK.L 17KAPIVSVTLYYK.V 24KVEKWFGQIPR.G 55KVGLHILGGLASSR.L 52 R.VKAQAEAELAQK.L 61 R.LDATAIAEAQER.L 35 R.TFVVSAPVQADR.T 28 R.MGWLLPAVTQDK.L 15 R.MGWLLPAVTQDK.L+Oxidation (M) 27 R.KAPIVSVTLYYK.V 52 R.DLAAMAALTPAQVK.A 70 R.DLAA*M*AALTPAQVK.A+Oxidation (M) 74 R.AVVSTLSGEIGGLER.V 32 R.LKPMLEQAFGAWK.A 31 R.LKPMLEQAFGAWK.A+Oxidation (M) 74KGATLAEGQVYSDDPAK.Y 28 R.RAVVSTLSGEIGGLER.V 50KFVLANGLTTIVHTDR.K 18KSAAPVKPLGAAVPAAQGR.I 23 R.NWTAPGINDPDTPALK.V 37 R.MGWLLPAVTQDKLDK.Q 33 R.MGWLLPAVTQDKLDK.Q+Oxidation (M) 47 R.AVAEAAFDKVLADYLR.D 65 R.ALPNQFETNAAVQAAIR.K 25KGWSYGVYSSVTQPTGPR.T 29KWMSRPVYALNVVPGPR.T 26KWMSRPVYALNVVPGPR.T+Oxidation (M) 20 R.TFVVSAPVQADRTGDAIK.A 18KGATLAEGQVYSDDPAKYK.R 36 R.IVLIDRPNSPQSVIMAGR.V 43 R.IVLIDRPNSPQSVIMAGR.V+Oxidation (M) 58KADTEALGLANDVLGGGFLSR.L 47 R.GPEIEPVKAGPVTLPAPLSR.D 55 R.HATIGSMADLDAATLTDVRK.W 56 R.HATIGSMADLDAATLTDVRK.W+Oxidation (M) 21 R.EEKGWSYGVYSSVTQPTGPR.T 46KAIIADVGAFPGQKPITQEELQR.V 70 R.TNYVETVPTGALDLALFMESDR.M 50 R.AIEASANEFLRPGGMTYVVVGDR.K 59 R.AIEASANEFLRPGGMTYVVVGDR.K+Oxidation (M) 65 R.AIEASANEFLRPGGMTYVVVGDRK.I 54 R.QGDNQPYGLFDYAQADGLLPVGHPYR.H+ Gln->pyro-Glu (N-term Q) 71 R.QGDNQPYGLFDYAQADGLLPVGHPYR.H 23 R.LNKDLREEKGWSYGVYSSVTQPTGPR.T 38KWFTEHYGPNNVVLVLSGDIDAATARPK.V 35KIVEPQLEGLGLPIDVQQAPQAAADDQSVE.- 55 R.KIVEPQLEGLGLPIDVQQAPQAAADDQSVE.- 40 R.ALGSTLFGQHPYAQPTDGLGNAASLAALTPAALR.A

[a] In the table only peptides with ion score ≥10 are reported.

at 37 °C (lane 3). Neither the use of different IPTG concentrations nor the variation of the induction time seemed to improve rRHA-P solubility.

To increase the yield of active rRHA-P in the soluble fraction of cell cultures induced at 23 °C, a novel set of analytical expression experiments using either LB or LB-N (Materials and Methods) supplemented with 1 mM of both betaine and sorbitol (two osmolytes), were performed. Data in literature suggest that the addition of these molecules to the growth medium of induced recombinant cells of *E. coli* might increase protein solubility. This effect should be improved by the co-presence in the medium of a high salt concentration (LB-N medium), which might ***i)*** increase the uptake rate in the cytoplasm of betaine and sorbitol from the extracellular

Table 2
Purification Table of rRHA-P.

	Total Units[b]	Specific Activity (Units/mg)[b]	P.F.[a]	Yield (%)
Cell lysate	13,728.8	49.8		
I Q-Sepharose	5,939.6	54.7	1.1	43.3
Gel filtration	1,370.6	66.1	1.3	9.9
II Q-Sepharose	920.9	346.9	6.9	6.7

[a] Purification Factor.
[b] Standard deviations, expressed as percentage errors, were always within 20%.

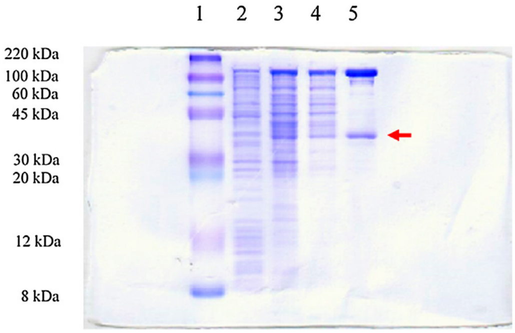


Fig. 1. SDS-PAGE analysis of rRHA-P purification steps. Lane 1: MW standard. Lane 2: cell lysate of *E. coli* strain BL21(DE3). Lane 3: peak fraction after the first Q-sepharose. Lane 4: peak fraction after S200-gel filtration chromatography. Lane 5: purified rRHA-P; the red arrow in this line indicates the contaminant protein found at the end of purification. For each sample an amount of ca. 4 μg of total proteins was loaded. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 3
TLC results of RHA-P-catalyzed reaction using different *p*NP-α- and *p*NP-β- derivatives as substrates.

	Substrates[a]	Hydrolysis		Transglycosylation	
		3 h	20 h	3 h	20 h
1	*p*NP-α-D-Glc*p*	–	–	–	–
2	*p*NP –β-D-Glc*p*	–	–	–	–
3	*p*NP –α-D-Gal*p*	–	–	–	–
4	*p*NP –β-D-Gal*p*	–	–	–	–
5	*p*NP –α-D-Man*p*	–	–	–	–
6	*p*NP –β-D-Man*p*	–	–	–	–
7	*p*NP –β-D-Xyl*p*	–	–	–	–
8	*p*NP –α-L-Fuc*p*	–	–	–	–
9	*p*NP –β-D-Fuc*p*	–	–	–	–
10	*p*NP –α-L-Ara*f*	–	–	–	–
11	*p*NP –β-L-Ara*f*	–	–	–	–
12	*p*NP –α-L-Rha*p*	+	+	–	–
13	*p*NP β-GlcA*p*	–	–	–	–
14	*p*NP β-D-NAGlc*p*	–	–	–	–
15	*p*NP β-D-NAGal*p*	–	–	–	–

[a] Ara, arabinose; Fuc, fucose; Gal, galactose; Glc, glucose; Man, mannose; NAGlc, *N*-acetylglucosamine; NAGal: *N*-acetylgalactosamine, GlcA: glucuronic acid, Rha, Rhamnose; Xyl, xylose.

environment [55], and/or ***ii)*** induce the expression of additional heat-shock proteins assisting in protein folding [56].

For these experiments, IPTG concentration and induction time were respectively 0.1 mM and 3 h. Activity assays showed that the best expression of active rRHA-P was obtained inducing the cultures at 23 °C and using LB-NBS (LB-N to which 1 mM of both betaine and sorbitol were added) as growth medium. In this case, specific activity of cell extract was 46 times higher than what obtained with the soluble fraction of recombinant cells grown in LB and induced at 37 °C.

Large-scale recombinant expression was performed by using the optimized conditions suggested by the analytical experiments.

rRHA-P purification was carried out following three different chromatographic steps. The purification table (Table 2) showed a final purification factor (PF) of 6.9. SDS-PAGE analysis of peak fractions representative of each purification step (Fig. 1) showed in the final purified sample of rRHA-P the additional presence of a major contaminant protein with an approximate MW of ∼35 kDa.

The 35 kDa protein band was excised, digested *in situ* with trypsin and analyzed by LC–MS/MS. The protein, identified by searching the *E. coli* sequence database (Materials and Methods), resulted to be glyceraldehyde 3-phosphate dehydrogenase, as identified by 17 peptides matches for a sequence coverage of 71%.

Purified rRHA-P amino acidic sequence was verified up to a 94% coverage, by MS Mapping analysis carried out by MALDI-TOF and LC–MS/MS. However, it is worth noting that the protein N-terminus, detected by MS analysis, lacks the first 23 amino acids from the expected cloned sequence (Fig. 2 and Fig. S3).

To confirm the lack of the putative N-terminal peptide, purified rRHA-P was subjected to both a MALDI-TOF and LC–MS/MS analysis after an *in situ* digestion with either trypsin or chymotrypsin, and to the N-terminus sequencing. Mass spectrometry analysis allowed the identification of peptides ^{23}ESRDDAAEVAPSTRPEPSLEQAF^{45} (MH+ 2502.17 $^{\text{ppm}}$) and ^{23}ESRDDAAEVAPSTR^{36} (MH^{+} 1503.70) when rRHA-P was digested with chymotrypsin or trypsin, respectively. This evidence was also confirmed by N-terminus sequencing results, which showed that the first 5 amino acids of the recombinant protein are **ESRDD**. Interestingly, also in native RHA-P the same N-terminal peptide, obtained by fragmentation spectrum of the triple charged ion at 834.73 *m*/*z* from the *in situ* digestion of the protein with chymotrypsin (Fig. S4), had been detected in the first place. The scarce amount of native protein, however, did not allow us to confirm the effective lack of the N-terminus peptide by direct amino acidic sequencing.

3.3. *Biochemical characterization of rRHA-P*

3.3.1. *Analytical gel filtration*

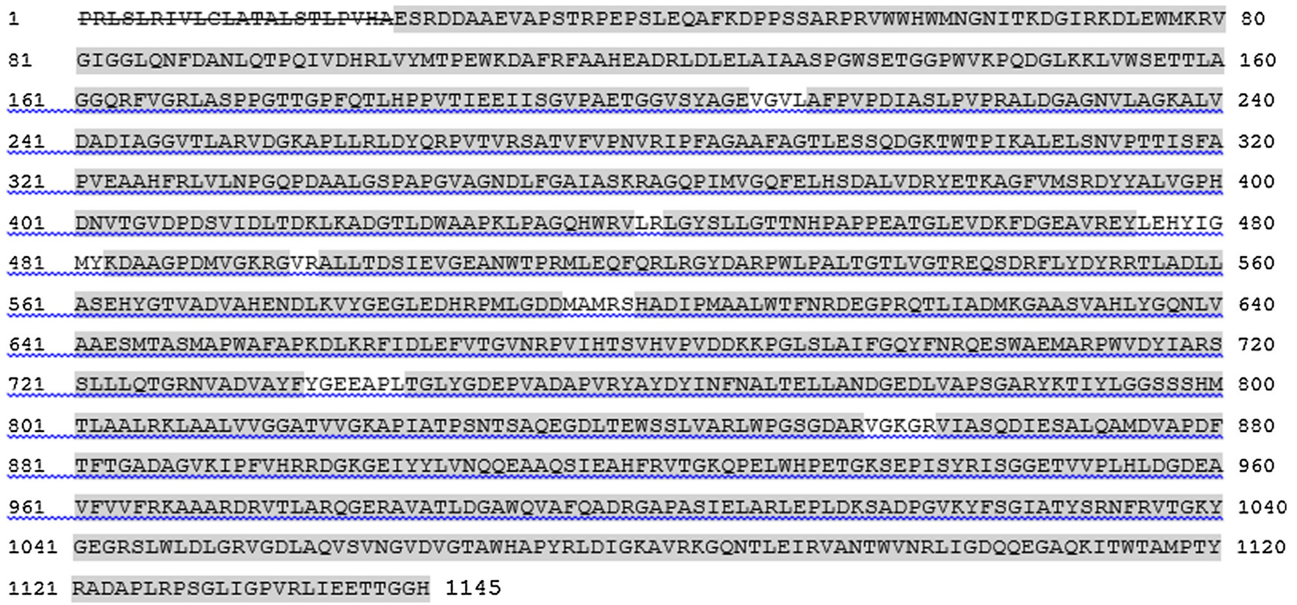
The oligomeric state of rRHA-P was verified by analytical gel filtration using a Superdex 200 analytical column; the protein eluted as a single peak with an apparent molecular weight of 101,500 ± 5,000 Da, which indicates that rRHA-P is a monomer.

3.3.2. *Substrate specificity*

To investigate rRHA-P substrate specificity in hydrolysis and self-condensation processes, several *p*NP-α- and *p*NP-β- derivatives were used as substrates in enzymatic assays performed as described in Materials and Methods. Reactions were followed over time by TLC analysis and data are reported in Table 3. As evident from Table 3, the purified enzyme showed activity only on *p*NPR, confirming that rRHA-P is indeed a α-L-rhamnosidase. Moreover, no activity on *p*NP-β-D-Glc was detected, excluding that rRHA-P might act as a naringinase, a class of glycosyl hydrolase having both α-L-rhamnosidase and β-D-glucosidase activities [40]. Furthermore, oligo- and polysaccharides such as rutinose, maltose, sucrose, lactose, pullulan, starch, amylopectin, raffinose, xylan from birchwood, xylan from oat spelt, hyaluronic acid, α-cellulose, cellobiose, chitosan, β-Glucan from barley, laminarin, curdlan, fucoidan from *Fucus versiculosus*, were also tested as substrates for rRHA-P. TLC analysis of all these reactions, performed as described in Materials and Methods, showed the lack of activity of this enzyme on any of these compounds.

3.3.3. *Mechanism of action*

To understand whether rRHA-P behaves as an inverting or a retaining glycosidase [1], ^{1}H NMR spectroscopy was used to monitor the stereochemical course of *p*NPR cleavage. As shown in Fig. 3, after the addition of the enzyme, a mixture of α and β



1 ~~PRLSLRIVLCLATALSTLPVHA~~ESRDDAAEVAPSTRPEPSLEQAFKDPPSSARPRVWWHWMNGNITKDGIRKDLEWMKRV 80
81 GIGGLQNFDANLQTPQIVDHRLVYMTPEWKDAFRFAAHEADRLDLELAIAASPGWSETGGPWVKPQDGLKKLVWSETTLA 160
161 GGQRFVGRLASPPGTTGPFQTLHPPVTIEEIISGVPAETGGVSYAGEVGVLAFPVPDIASLPVPRALDGAGNVLAGKALV 240
241 DADIAGGVTLARVDGKAPLLRLDYQRPVTVRSATVFVPNVRIPFAGAAFAGTLESSQDGKTWTPIKALELSNVPTTISFA 320
321 PVEAAHFRLVLNPGQPDAALGSPAPGVAGNDLFGAIASKRAGQPIMVGQFELHSDALVDRYETKAGFVMSRDYYALVGPH 400
401 DNVTGVDPDSVIDLTDKLKADGTLDWAAPKLPAGQHWRVLRLGYSLLGTTNHPAPPEATGLEVDKFDGEAVREYLEHYIG 480
481 MYKDAAGPDMVGKRGVRALLTDSIEVGEANWTPRMLEQFQRLRGYDARPWLPALTGTLVGTREQSDRFLYDYRRTLADLL 560
561 ASEHYGTVADVAHENDLKVYGEGLEDHRPMLGDDMAMRSHADIPMAALWTFNRDEGPRQTLIADMKGAASVAHLYGQNLV 640
641 AAESMTASMAPWAFAPKDLKRFIDLEFVTGVNRPVIHTSVHVPVDDKKPGLSLAIFGQYFNRQESWAEMARPWVDYIARS 720
721 SLLLQTGRNVADVAYFYGEEAPLTGLYGDEPVADAPVRYAYDYINFNALTELLANDGEDLVAPSGARYKTIYLGGSSSHM 800
801 TLAALRKLAALVVGGATVVGKAPIATPSNTSAQEGDLTEWSSLVARLWPGSGDARVGKGRVIASQDIESALQAMDVAPDF 880
881 TFTGADAGVKIPFVHRRDGKGEIYYLVNQQEAAQSIEAHFRVTGKQPELWHPETGKSEPISYRISGGETVVPLHLDGDEA 960
961 VFVVFRKAAARDRVTLARQGERAVATLDGAWQVAFQADRGAPASIELARLEPLDKSADPGVKYFSGIATYSRNFRVTGKY 1040
1041 GEGRSLWLDLGRVGDLAQVSVNGVDVGTAWHAPYRLDIGKAVRKGQNTLEIRVANTWVNRLIGDQQEGAQKITWTAMPTY 1120
1121 RADAPLRPSGLIGPVRLIEETTGGH 1145

Fig. 2. rRHA-P sequence coverage by MS analysis. Highlighted in grey are the peptides, identified by MS analysis (MALDI-TOF and/or LC–MSMS), of the enzymatic digestions of rRHA-P carried out either *in situ* or in solution with proteases of different specificity: trypsin, chymotrypsin and endoproteinase GluC (V8-DE). A global sequence coverage of 94% was obtained. Crossed is the N-terminus sequence expected for rRHA-P (reference sequence WP_013837086.1) and not retrieved, either by N-terminus or by MS analysis.

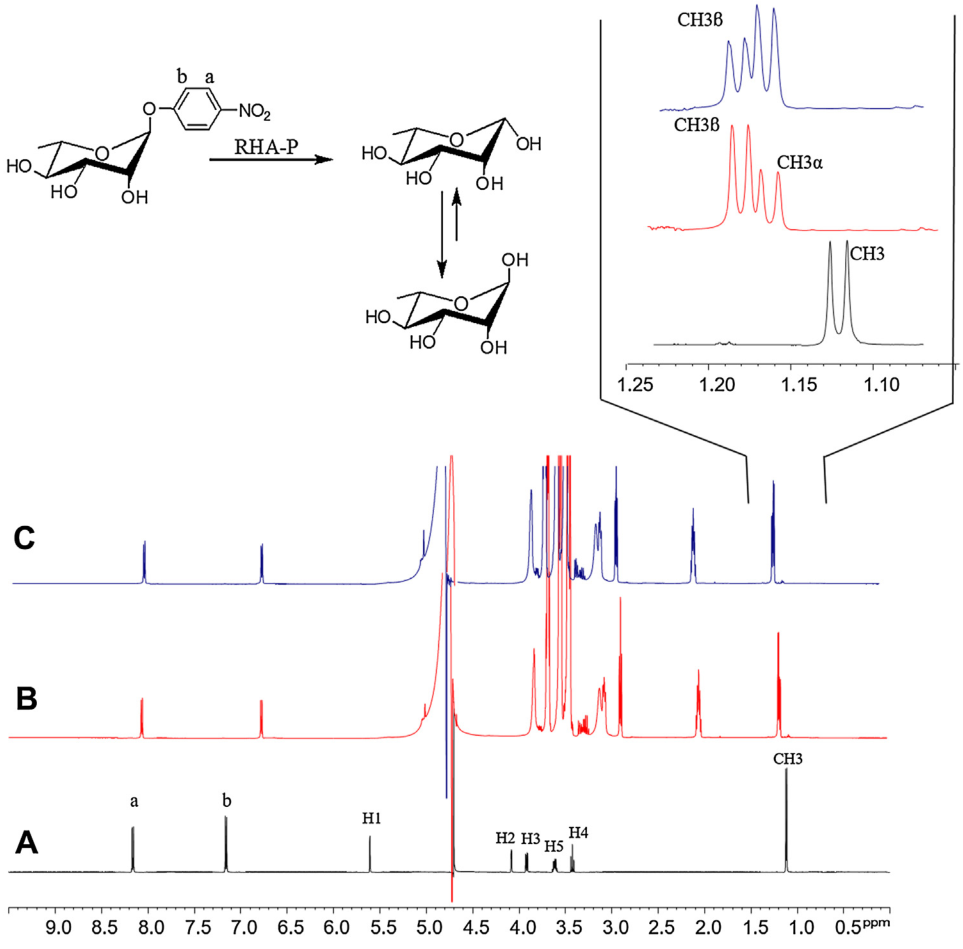


Fig. 3. ^{1}H NMR analysis of pNPR after rRHA-P catalyzed hydrolysis. A. Spectrum of the ligand alone in solution (3 mM). B. Spectrum following the addition of the enzyme to the NMR tube (~60 μg), which showed the appearance of CH_3 signals relative to α- and β-anomers. C. Spectrum obtained after a 30 min reaction; a mixture of α- and β-anomers with a different relative ratio was observed due to emiacetal equilibrium.

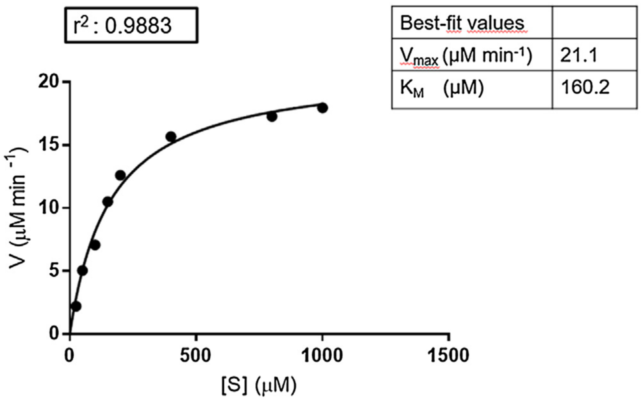


Fig. 4. rRHA-P kinetic behavior. Reaction rate expressed in μM/min is plotted as a function of *p*NPR concentration.

rhamnose anomers was observed in solution, as evident from the presence of two different methyl group resonances. TOCSY spectrum allowed the assignment of both α and β spin systems up to the methyl group resonances located at position 6 of each rhamnose residue. Therefore, the less shielded methyl signal was assigned to β-rhamnose. The analysis of the intensity of the different methyl proton signals showed that the thermodynamically unfavorable β-rhamnose prevailed soon after the addition of rRHA-P (Fig. 3B). Afterwards, as a consequence of the equilibrium established by a reducing monosaccharide in water, the relative ratio of the two peaks changed up to the point in which the α anomeric product was predominant (Fig. 3C). Hence, the data suggested that rRHA-P acts as an inverting enzyme as it cleaves the α-glycoside ligand yielding a β-anomer, which in water solution equilibrates according to the anomeric effect.

3.3.4. *Enzymatic characterization*

All enzymatic assays described in this subsection are detailed in Materials and Methods section and were performed using *p*NPR as substrate. Kinetic constants on *p*NPR were determined using a substrate concentration ranging from 0.025 to 1 mM. The reaction rate (μM/min), plotted as a function of the substrate concentration (μM), showed a typical Michaelis-Menten trend and is reported in Fig. 4. K_M constant was of 160.2 (±17.3) μM, a value significantly lower than most other α-RHAs described in literature [32,57], indicating a higher affinity of rRHA-P for *p*NPR. A V_{max} of 21.1 (±8.0) μM/min, a k_{cat} of 734.4 (±212.9) sec^{-1} and a K_s of 4.6 (±1.7) sec^{-1} $μM^{-1}$ were obtained.

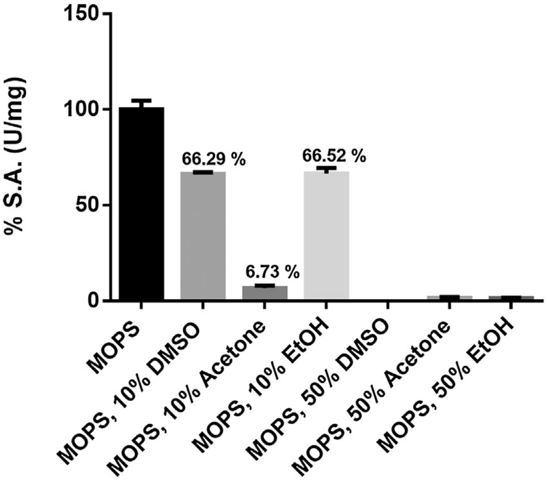


Fig. 6. rRHA-P tolerance to organic solvents. Data are reported as percentage of residual specific activity (S.A.) compared to the control (25 mM MOPS, pH 6.9). MOPS buffer is used at a final concentration of 25 mM and at pH 6.9.

rRHA-P revealed an optimal activity in the range 6.0-7.5 with an optimum at pH 6.9 (Fig. 5A), retaining 47% and 36% of activity at pH 7.9 and 8.8, respectively. In addition, rRHA-P activity was assayed in the temperature range 25 °C–55 °C. Data are presented in Fig. 5B and show an optimal activity temperature at ca. 40.9 °C, with a 44% retaining of its initial value up to 50 °C. To evaluate the thermal stability of rRHA-P in the same range used to determine the optimal temperature, an incubation of the purified protein for either one or 3 h was performed at different temperatures; afterwards, the residual rhamnosidase activity of the enzyme was calculated (Fig. S5). The enzymatic activity appeared to be stable between 25 and 40 °C, independently from the duration of the incubation. At 40 °C, corresponding to the temperature optimum, the activity is stable up to 24 h (data not shown).

In order to establish the tolerance of the enzymatic activity of rRHA-P to the presence of organic solvents, *p*NPR assay in 50 mM MOPS pH 6.9 containing 10% or 50% of either DMSO, acetone or ethanol was performed (Fig. 6). While activity dramatically decreased in all concentration of acetone and in 50% of ethanol and DMSO, a 66% of activity was still recovered in the presence of 10% of either DMSO or ethanol.

3.3.5. *Activity on flavonoids*

The ability of rRHA-P to hydrolyze natural flavonoids was evaluated. Enzymatic assays using naringin, rutin and neohesperidin

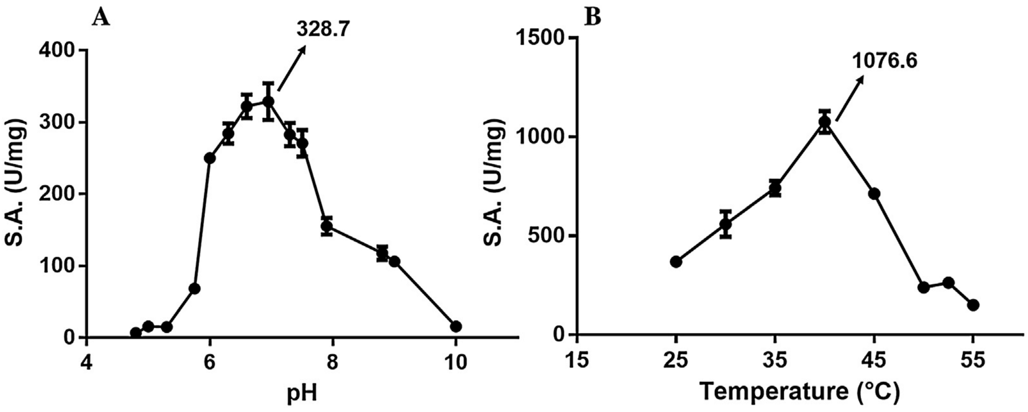


Fig. 5. rRHA-P activity assays. pH (A) and temperature (B) optimum curves. In both graphs specific activity (S.A.) is plotted as a function of pH and temperature, respectively. The maximum value of S.A. obtained is reported in both graphs and indicated by a black arrow.

dihydrochalcone were performed and followed by TLC analysis, as described in Materials and Methods. To this purpose, a 6 mM solution of each compound was incubated with purified rRHA-P in 50 mM Na-phosphate buffer pH 7.0 at 40 °C, and the reaction was monitored by TLC analyses over time. After 3 h of incubation, TLC analysis of the reaction mixtures showed a 40–60% hydrolysis of neohesperidin dihydrochalcone and rutin in the corresponding derhamnosylated neohesperidin dihydrochalcone and quercetin-3-β-glucopyranoside (isoquercitrin). Besides, a total conversion of naringin was observed as the corresponding products, rhamnose and prunin, were the only compounds observed on the TLC plate. Reaction on naringin was also carried out in the presence of 10% DMSO. In this case, in our experimental conditions, the complete conversion of naringin in prunin and rhamnose occurred after just 1h.

4. Discussion

α-RHAs are a group of glycosyl hydrolases (GHs) that have attracted a great deal of attention due to their potential application as biocatalysts in a variety of industrial processes. These enzymes are of particular interest for the biotransformation of several natural compounds used in pharmaceutical and food industry. Enzymatic derhamnosylation catalyzed by α-RHAs can be used, for example, in functional foods and beverages containing molecules with enhanced health-related properties. Some examples encompass the biotransformation of natural steroids, antibiotics, flavonoids, and terpenyl glycosides responsible for wine aromas [2,40].

Bacterial α-RHAs are still poorly characterized, but their biochemical properties might be of key importance for biotransformation processes involving reactions conditions that are unfavorable for fungal rhamnosidases. In this work, we have successfully expressed and partially characterized a novel recombinant α-RHA from *Novosphingobium* sp. PP1Y, named RHA-P. The α-RHA activity has been isolated from the native source and the corresponding protein sequence was verified by MS analysis. The sequence was used for a BLAST search against all protein databases, and resulted to be homologous to GHs present in the genomes of several Sphingomonads; among these, a 48% sequence identity of RHA-P with the α-RHA (rhaM) from *Sphingomonas paucimobilis* FP2001, was retrieved [58]. Noteworthy, an analysis of the amino acidic sequence of RHA-P highlighted the presence of a 23 amino acids long signal peptide [59] located at the N-terminus of the protein.

The new α-RHA identified is encoded by *orf*PP1Y_RS05470, a 3,441 bp gene located in Megaplasmid Mpl of *N.* sp. PP1Y [44], which was amplified and cloned into pET22b (+) plasmid to allow for the recombinant expression in *E. coli* BL21(DE3) strain. First attempts of expression of the recombinant protein, rRHA-P, resulted in the foremost presence of the protein in the insoluble portion of the induced cultures. A lower induction temperature, shifted from 37 °C to 23 °C, along with the use of a high-salt LB formulation containing both betaine and sorbitol, which presumably act as “chemical chaperones”, efficiently concurred in improving the yield of soluble, active rRHA-P [54–56].

rRHA-P was purified following a three-step purification protocol. The relatively poor yield obtained at the end of the purification was not unexpected at this stage, as we chose after each step to avoid collecting rRHA-P active fractions that showed the presence of too many protein contaminants.

Purified rRHA-P sequence was verified by MS mapping. It is interesting to underline that the protein N-terminus expected from the cloned sequence was never detected, either by MS analysis (both native and recombinant RHA-P) or by N-terminus sequencing (only recombinant RHA-P). The lack of this peptide, which had been expected by a preliminary analysis of the amino acidic sequence, confirmed the presence of a signal peptide presumably cleaved through a post-translational proteolytic processing. A similar evidence has been described for the α-RHA isolated from *S. paucimobilis* FP2001 and recombinantly expressed in *E. coli* [58]. Takeshi M. and coworkers reported in this case the presence of a putative signal peptide cleaved between positions A24 and N25 [58]. Similarly, in rRHA-P the cleavage site is located between A23 and E24. The amino acidic sequence of the N-terminus of both α-RHA from strain FP2001 and strain PP1Y revealed several peculiarities owned by bacterial signal peptides, such as the presence of a charged region (2–5 residues) followed by a hydrophobic stretch of ∼ 12 aa, along with the presence of small and apolar residues located 1 and 3 aa upstream of the cleavage site [60]. The functional role of this cleavage, as well as the possible sorting of these processed proteins in the periplasmic space of the native bacteria, needs to be further investigated.

A preliminary structural characterization showed that rRHA-P is a monomeric protein with an approximate molecular weight of 101,500 ± 5,000 Da. Mechanism of action, kinetic parameters and optimum reaction conditions were determined using a synthetic substrate, *p*NPR.

As for the reaction mechanism, GHs are generally grouped in two main classes, inverting and retaining. A typical inverting glycosidase requires the presence of a catalytic acid residue and a catalytic basic residue, whereas a typical retaining glycosidase contains a general acid/base residue and a nucleophile. These different hydrolysis mechanisms lead, respectively, to an inversion or to a retention of the anomeric oxygen configuration in the product compared to the substrate [61]. NMR experiments showed that rRHA-P acts as an inverting GH; further investigation will give additional insights in identifying residues in the active site of this enzyme, which are directly involved in catalysis.

rRHA-P catalytic efficiency, defined by a K_s value of $4.6\ s^{-1}\ \mu M^{-1}$, underlined that the activity of this enzyme resulted to be comparable or even higher than that of other bacterial α-RHAs described in literature. When clearly reported, such as in lactic acid bacteria *Lactobacillus plantarum,* and *Pediococcus acidilactici,* K_s values on *p*NPR range in fact from $0.01\ s^{-1}\ \mu M^{-1}$ to $0.019\ s^{-1}\ \mu M^{-1}$ [38,39]. In addition, K_s of rRHA-P resulted to be approximately 3 times higher than the K_s of $1.74\ s^{-1}\ \mu M^{-1}$ reported for the α-RHA isolated from *S. paucimobilis* FP2001 [57]. The comparison among the K_M values present in literature, ranging from 1.18 mM for *S. paucimobilis* FP 2001 to 16.2 mM for the α-RHA isolated from *P. acidilactici*, showed a higher apparent affinity of rRHA-P for *p*NPR (K_M of 160 μM), which might be in part responsible for the great differences found in K_S values.

To investigate the effective potential of using rRHA-P in biotechnological processes involving the biotransformation of flavonoids, optimal reaction conditions in terms of pH, temperature and presence of organic solvents were evaluated. In particular, rRHA-P, in line with other bacterial α-RHAs [34–38,55], has an optimal activity at pH 6.9, a value that highlights a marked difference with fungal α-RHAs acting at pH values ranging from 4 to 5 [17,31,32]. rRHA-P shows an optimal activity at 40.9 °C and a significant overall thermal stability, retaining ∼66% of the activity up to 45 °C, a temperature range mainly used in most industrial processes. In this respect, also a great majority of other fungal and bacterial enzymes exhibit similar values of temperature stability and optimal activity [17,32,34,37,38,55], whereas only a limited number of α-RHAs is more thermophilic and has an optimum of temperature at ca. 60 °C [31,36,37,39]. rRHA-P has a moderate tolerance to organic solvents, usually employed for the biotransformation of flavonoids that are poorly soluble in water, retaining the 66% of activity on *p*NPR in solutions containing 10% of either DMSO or ethanol. Naringin conversion is even faster in 10% DMSO than in the absence of this solvent, probably owing to the fact that this flavonoid is more soluble in the presence of DMSO. Noteworthy, flavonoids conversion catalyzed by other α-RHA has been described in the presence of 2–5% DMSO [37–39]; higher concentration of DMSO, as the one used for our bioconversion, might further increase the solubility of many flavonoids of commercial interest. These data altogether suggest a higher tolerance of rRHA-P to the presence of solvents compared to α-RHAs described in literature for which a residual 20 − 30% activity in 12% ethanol and a 24% activity in 25% DMSO has been reported [17,39]. It is worth mentioning that rRHA-P residual activity in ethanol-containing buffers might be of additional importance as many biologically relevant rhamnosylated flavonoids are found in wine and possibly citrus derived alcoholic beverages.

As previously underlined, the kinetic data obtained so far for this enzyme undoubtedly encourage its use for the bioconversion of glycosylated natural flavonoids that are more soluble at high temperature (ca. 50 °C) [62], basic pH values, and in the presence of organic solvents. In this context, experiments to verify the hydrolytic properties of rRHA-P on few natural flavonoids containing either α-1,2 or α-1,6 glycosidic linkages between α-L-rhamnose and β-D-glucose were performed. TLC analysis on naringin (α-1,2), rutin (α-1,6) and neohesperidin (α-1,6) showed a total conversion of naringin in the corresponding prunin and rhamnose, while a partial hydrolysis of rutin and neohesperidin was observed, thus confirming the ability of this enzyme to hydrolyze both α-1,2 and α-1,6 linkages. Likewise, the majority of other bacterial and fungal α-RHAs exhibit similar substrate specificity, even though these enzymes often show a slight preference for the hydrolysis of α-1,6 linkages [31,32,37–39].

In conclusion, the results described in this work boost the interest toward a further biochemical characterization of RHA-P activity using for example a wide array of glycosylated flavonoids, to investigate the effective potential of this novel enzyme for various biomass and food-processing applications aimed at recovering high-value added compounds from complex vegetable matrixes.

Acknowledgment

This research was supported by FARB 2014 ORSA 132082. The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.molcatb.2016.10.002.

References

[1] B. Henrissat, A classification of glycosyl hydrolases based on amino-acid sequence similarities, Biochem. J. 280 (1991) 309–316.

[2] P. Manzanares, S. Vallés, D. Ramón, M. Orejas, α-L-Rhamnosidases: old and new insights, industrial enzymes, structure, function and application, in: J. Polaina, A.P. MacCabe (Eds.), Industrial Enzymes, Springer, 2007, pp. 117–140.

[3] M. Mutter, G. Beldman, H.A. Schols, A.G.J. Voragen, Rhamnogalacturonan α-L-rhamnopyranohydrolase. A novel enzyme specific for the terminal nonreducing rhamnosyl unit in rhamnogalacturonan regions of pectin, Plant Physiol. 106 (1) (1994) 241–250.

[4] D.C. Crick, S. Mahapatra, P.J. Brennan, Biosynthesis of the arabinogalactan-peptidoglycan complex of *Mycobacterium Tuberculosis*, Glicobiology 11 (9) (2001) 107R–118R.

[5] E.G. Maxwell, I.J. Colquhoun, H.K. Chau, A.T. Hotchkiss, K.W. Waldron, V.J. Morris, N.J. Belshaw, Rhamnogalacturonan I containing homogalacturonan inhibits colon cancer cell proliferation by decreasing ICAM1 expression, Carbohydr. Polym 132 (2015) 546–553.

[6] E. Nguema-Ona, M. Vicré-Gibouin, M. Gotté, B. Plancot, P. Lerouge, M. Bardor, A. Driouich, Cell wall *O*-glycoproteins and *N*-glycoproteins: aspects of biosynthesis and function, Front. Plant Sci. 5 (2014) 499.

[7] W. Hashimoto, O. Miyake, H. Nankai, K. Murata, Molecular identification of an α-L-rhamnosidase from *Bacillus* sp. strain GL1 as an enzyme involved in complete metabolism of gellan, Arch. Biochem. Biophys. 415 (2) (2003) 235–244.

[8] R. Rahim, U.A. Ochsner, C. Olvera, M. Graninger, P. Messner, J.S. Lam, G. Soberón-Chávez, Cloning and functional characterization of the *Pseudomonas aeruginosa* rhlC gene that encodes rhamnosyltransferase 2, an enzyme responsible for di-rhamnolipid biosynthesis, Mol. Microbiol. 40 (3) (2001) 708–718.

[9] -W. Nereus, I.V. Gunther, A. Nun˜ez, W. Fett, D.K.Y. Solaiman, Production of rhamnolipids by *Pseudomonas chlororaphis*, a nonpathogenic bacterium, Appl. Environ. Microb. 71 (5) (2005) 2288–2293.

[10] E. Champion, I. Andre, L.A. Mulard, P. Monsan, M. Remaud-Simeon, S. Morel, Synthesis of L-rhamnose and *N*-acetyl-D-glucosamine derivatives entering in the composition of bacterial polysaccharides by use of glucansucrases, J. Carbohydr. Chem. 28 (2009) 142–160.

[11] E.Jr. Middleton, T.C. Theoharides, The effects of plant flavonoids on mammalian cells: implications for inflammation, heart disease, and cancer, Pharmacol. Rev. 4 (52) (2000) 673–751.

[12] F.L. Perez-Vizcaino, J. Duarte, Flavonols and cardiovascular disease, Mol. Aspects Med. 31 (6) (2010) 478–494.

[13] P. Fresco, F. Borges, M.P.M. Marques, C. Diniz, The anticancer properties of dietary polyphenols and its relation with apoptosis, Curr. Pharm. Des. 16 (2010) 114–134.

[14] G. Ceíliz, M.C. Audisio, M. Daz, Antimicrobial properties of prunin, a citric flavanone glucoside, and its prunin 6-*O*-lauroyl ester, J. Appl. Microbiol. 109 (4) (2010) 1450–1457.

[15] M. Puri, S.S. Marwaha, R.M. Kothari, J.F. Kennedy, Biochemical basis of bitterness in citrus fruit juices and biotech approaches for debittering, Crit. Rev. Biotechnol. 16 (1996) 145–155.

[16] R. González-Barrio, L.M. Trindade, P. Manzanares, L.H. de Graaff, F.A. Tomás-Barberán, J.C. Espín, Production of bioavailable flavonoid glucosides in fruit juices and green tea by use of fungal alpha-L-rhamnosidases, J. Agric. Food Chem. 52 (20) (2004) 6136–6142.

[17] M.V. Gallego, F. Piñaga, D. Ramón, S. Vallés, Purification and characterization of an α-L-rhamnosidase from *Aspergillus terreus* of interest in winemaking, J. Food Sci. 66 (2) (2001) 204–209.

[18] Z. Gunata, S. Bitteur, J.M. Brillouet, C. Bayonove, R. Cordonnier, Sequential enzymatic hydrolysis of potentially aromatic glycosides from grape, Carbohydr. Res. 184 (1988) 139–149.

[19] P. Manzanares, M. Orejas, J.V. Gil, L.H. de Graaff, J. Visser, D. Ramón, Construction of a genetically modified wine yeast strain expressing the *Aspergillus aculeatus* rhaA gene, encoding an α-L-rhamnosidase of enological interest, Appl. Environ. Microbiol. 69 (12) (2003) 7558–7562.

[20] M.E. Melo Branco de Araújo, Y.E. Moreira Franco, T. Grando Alberto, M. Alves Sobreiro, M.A. Conrado, D. Gonçalves Priolli, A.C.H. Frankland Sawaya, A.L.T.G. Ruiz, J.E. de Carvalho, P. de Oliveira Carvalho, Enzymatic de-glycosylation of rutin improves its antioxidant and antiproliferative activities, Food Chem. 141 (2013) 266–273.

[21] V.D. Bokkenheuser, C.H. Shackleton, J. Winter, Hydrolysis of dietary flavonoid glycosides by strains of intestinal Bacteroides from humans, Biochem. J. 248 (3) (1987) 953–956.

[22] A.J. Day, J.M. Gee, M.S. DuPont, I.T. Johnson, G. Williamson, Absorption of quercetin-3-glucoside and quercetin-4′-glucoside in the rat small intestine: the role of lactase phlorizin hydrolase and the sodium-dependent glucose transporter, Biochem. Pharmacol. 65 (7) (2003) 1199–1206.

[23] S. Wolffram, M. Block, P. Ader, Quercetin-3-glucoside is transported by the glucose carrier SGLT1 across the brush border membrane of rat small intestine, J. Nutr. 132 (4) (2002) 630–635.

[24] A. Scalbert, G. Williamson, Dietary intake and bioavailability of polyphenols, J. Nutr 130 (2000) 2073S–2085S.

[25] P.C. Hollman, J.M. van Trijp, M.N. Buysman, M.S. van der Gaag, M.J. Mengelers, J.H. de Vries, M.B. Katan, Relative bioavailability of the antioxidant flavonoid quercetin from various foods in man, FEBS Lett. 418 (1997) 152–156.

[26] K. Németh, G.W. Plumb, J.G. Berrin, N. Juge, R.J. Hassan, H.Y. Naim, G. Williamson, D.M. Swallow, P.A. Kroon, Deglycosylation by small intestinal epithelial cell β-glucosidases is a critical step in the absorption and metabolism of dietary flavonoid glycosides in humans, Eur. J. Nutr. 42 (2003) 29–42.

[27] P.C. Hollman, M.N. Bijsman, Y. van Gameren, E.P. Cnossen, J.H. de Vries, M.B. Katan, The sugar moiety is a major determinant of the absorption of dietary flavonoid glycosides in man, Free Radic. Res. 31 (6) (1999) 569–573.

[28] P. Garnier, X.T. Wang, M.A. Robinson, S. van Kasteren, A.C. Perkins, M. Frier, A.J. Fairbanks, B.G. Davis, Lectin-directed enzyme activated prodrug therapy (LEAPT): synthesis and evaluation of rhamnose-capped prodrugs, J. Drug Target. 18 (10) (2010) 794–802.

[29] M.A. Robinson, S.T. Charlton, P. Garnier, X.T. Wang, S.S. Davis, A.C. Perkins, M. Frier, R. Duncan, T.J. Savage, D.A. Wyatt, S.A. Watson, B.G. Davis, LEAPT: lectin-directed enzyme-activated prodrug therapy, Proc. Natl. Acad. Sci. 101 (40) (2004) 14527–14532.

[30] N.M. Young, R.A.Z. Johnston, J.C. Richards, Purification of the α-L-rhamnosidase of *Penicillium decumbens* and characteristics of two glycopeptide components, Carbohydr. Res. 191 (1989) 53–62.

[31] T. Koseki, Y. Mese, N. Nishibori, K. Masaki, T. Fujii, T. Handa, Y. Yamane, Y. Shiono, T. Murayama, H. Iefuji, Characterization of an alpha-L-rhamnosidase from *Aspergillus kawachii* and its gene, Appl. Microbiol. Biotechnol. 80 (6) (2008) 1007–1013.

[32] P. Manzanares, H.C. van den Broeck, L.H. de Graaff, J. Visser, Purification and characterization of two different alpha-L-rhamnosidases, RhaA and RhaB, from *Aspergillus aculeatus*, Appl. Environ. Microbiol. 67 (5) (2001) 2230–2234.

[33] J.E. Chua, P.A. Manning, R. Morona, The Shigella flexneri bacteriophage Sf6 tailspike protein (TSP)/endo rhamnosidase is related to the bacteriophage P22 TSP and has a motif common to exo- and endoglucanases and C-5 epimerases, Microbiology 145 (7) (1999) 1649–1659.

[34] W. Hashimoto, H. Nankai, N. Sato, S. Kawai, K. Murata, Characterization of α-L-rhamnosidase of *Bacillus* sp. GL1 responsible for the complete depolymerization of gellan, Arch. Biochem. Biophys. 415 (2) (1999) 56–60.

[35] A.G. Orrillo, P. Ledesma, O.D. Delgado, G. Spagna, J.D. Breccia, Cold-active α-L-rhamnosidase from psychrotolerant bacteria isolated from a sub-Antarctic ecosystem, Enz. Microbiol. Technol. 40 (2) (2007) 236–241.

[36] V.V. Zverev, C. Hertel, K. Bronnenmeier, A. Hroch, J. Kellermann, W.H. Schwarz, The thermostable α-L-rhamnosidase RamA of *Clostridium stercorarium*: biochemical characterization and primary structure of a bacterial α-L-rhamnoside hydrolase a new type of inverting glycoside hydrolase, Mol. Microbiol. 35 (2000) 173–179.

[37] M. Avila, M. Jaquet, D. Moine, T. Requena, C. Peláez, F. Arigoni, I. Jankovic, Physiological and biochemical characterization of the two alpha-L-rhamnosidases of *Lactobacillus plantarum* NCC245, Microbiology 155 (2009) 2739–2749.

[38] J. Beekwilder, D. Marcozzi, S. Vecchi, R. de Vos, P. Janssen, C. Francke, J. van Hylckama Vlieg, R.D. Hall, Characterization of rhamnosidases from *Lactobacillus plantarum* and *Lactobacillus acidophilus*, Appl. Environ. Microbiol 75 (2009) 3447–3454.

[39] H. Michlmayr, W. Brandes, R. Eder, C. Schümann, A.M. del Hierro, K.D. Kulbe, Characterization of two distinct glycosyl hydrolase family 78 α-L-rhamnosidases from *Pediococcus acidilactici*, Appl. Environ. Microbiol. 77 (18) (2011) 6524–6530.

[40] M. Puri, Updates on naringinase: structural and biotechnological aspects, Appl. Microbiol. Biotechnol. 93 (1) (2012) 49–60.

[41] Z. Cui, Y. Maruyama, B. Mikami, W. Hashimoto, K. Murata, Crystal structure of glycoside hydrolase family 78 α-L-rhamnosidase from *Bacillus* sp. GL1, J. Mol. Biol. 374 (2) (2007) 384–398.

[42] Z. Fujimoto, A. Jackson, M. Michikawa, T. Maehara, M. Momma, B. Henrissat, H.J. Gilbert, S. Kaneko, The structure of a *Streptomyces avermitilis* α-L-rhamnosidase reveals a novel carbohydrate-binding module CBM67 within the six-domain arrangement, J. Biol. Chem. 288 (17) (2013) 12376–12385.

[43] E. Notomista, F. Pennacchio, V. Cafaro, G. Smaldone, V. Izzo, L. Troncone, M. Varcamonti, A. Di Donato, The marine isolate *Novosphingobium* sp. PP1Y shows specific adaptation to use the aromatic fraction of fuels as the sole carbon and energy source, Microb. Ecol. 61 (3) (2011) 582–594.

[44] V. D'Argenio, M. Petrillo, P. Cantiello, B. Naso, L. Cozzuto, E. Notomista, G. Paolella, A. Di Donato, F. Salvatore, De novo sequencing and assembly of the whole genome of *Novosphingobium* sp. strain PP1Y, J. Bacteriol. 193 (16) (2011) 4296.

[45] W. Hashimoto, K. Murata, Alpha-L-rhamnosidase of *Sphingomonas* sp. R1 producing an unusual exopolysaccharide of sphingan, Biosci. Biotechnol. Biochem. 62 (6) (1998) 1068–1074.

[46] V. D'Argenio, E. Notomista, M. Petrillo, P. Cantiello, V. Cafaro, V. Izzo, B. Naso, L. Cozzuto, L. Durante, L. Troncone, G. Paolella, F. Salvatore, A. Di Donato, Complete sequencing of *Novosphingobium* sp. PP1Y reveals a biotechnologically meaningful metabolic pattern, BMC Genomics 15 (2014) 384.

[47] V. Izzo, P. Tedesco, E. Notomista, E. Pagnotta, A. Di Donato, A. Trincone, A. Tramice, Rhamnosidase activity in the marine isolate *Novosphingobium* sp. PP1Y and its use in the bioconversion of flavonoids, J. Mol. Catal. B: Enzym. 105 (2014) 95–103.

[48] B.L. Cantarel, P.M. Coutinho, C. Rancurel, T. Bernard, V. Lombard, B. Henrissat, The carbohydrate-Active EnZymes database (CAZy). an expert resource for glycogenomics, Nucl. Acids Res. 37 (2009) D233–D238.

[49] J. Sambrook, E.F. Fritsch, T. Maniatis, Molecular Cloning. A Laboratory Manual, 2nd ed., Cold Spring Harbor Lab. Press, 1982, pp. 545.

[50] P. Gerhardt, R.G.E. Murray, W.A. Wood, N.R. Krieg, Methods for General and Molecular Bacteriology, ASM Press, Washington D.C, 1994.

[51] D.J. States, R.A. Haberkorn, D.J. Ruben, A two-dimensional nuclear overhauser experiment with pure absorption phase in four quadrants, J. Magn. Reson. 48 (1982) 286–292.

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

[53] F. He, Laemmli-SDS-PAGE, Bio-protocol. 2011. Bio101:e80. http://www.bio-protocol.org/e80.

[54] J.R. Blackwell, R. Horgan, A novel strategy for production of a highly expressed recombinant protein in an active form, FEBS Lett. 295 (1–3) (1991) 10–12.

[55] N. Oganesyan, I. Ankudinova, S.H. Kim, R. Kimb, Effect of osmotic stress and heat shock in recombinant protein overexpression and crystallization, Prot. Expr. Purif. 52 (2) (2007) 280–285.

[56] M. Kilstrup, S. Jacobsen, K. Hammer, F.K. Vogensen, Induction of heat shock proteins DnaK, GroEL, and GroES by salt stress in *Lactococcus lactis*, Appl. Environ. Microbiol. 63 (5) (1997) 1826–1837.

[57] F. Miake, T. Satho, H. Takesue, F. Yanagida, N. Kashige, K. Watanabe, Purification and characterization of intracellular alpha-L-rhamnosidase from *Pseudomonas paucimobilis* FP2001, Arch. Microbiol. 173 (1) (2000) 65–70.

[58] M. Takeshi, K. Nobuhiro, S. Tomomitsu, Y. Tadatoshi, A. Yoichi, M. Fumio, Cloning, Sequence analysis, and expression of the gene encoding *Sphingomonas paucimobilis* FP2001 a-L-rhamnosidase, Curr. Microbiol. 51 (2005) 105–109.

[59] T.N. Petersen, S. Brunak, G. von Heijne, H. Nielsen, SignalP 4.0: discriminating signal peptides from transmembrane regions, Nat. Methods 8 (2011) 785–786.

[60] L.M. Gierasch, Signal sequences, Biochemistry 28 (1989) 923–930.

[61] B. Cobucci-Ponzano, M. Moracci, Glycosidases and glycosynthases, in: Biocatalysis in Organic Synthesis of Faber K, Fessner W-D Turner NJ, 2016, pp. 483–506.

[62] T. Sathishkumar, R. Baskar R, S. Shanmugam, P. Rajasekaran, S. Sadasivam, V. Manikandan, Optimization of flavonoids extraction from the leaves of *Tabernaemontana heyneana* Wall using L_{16} Orthogonal design, Naturesci 6 (3) (2008), 1545–0740.

Archives of Biochemistry and Biophysics 648 (2018) 1–11

Contents lists available at ScienceDirect

Archives of Biochemistry and Biophysics

journal homepage: www.elsevier.com/locate/yabbi

Structural and functional insights into RHA-P, a bacterial GH106 α-L-rhamnosidase from *Novosphingobium* sp. PP1Y

Francesca Mensitieri[a], Federica De Lise[a], Andrea Strazzulli[a], Marco Moracci[a,b], Eugenio Notomista[a], Valeria Cafaro[a], Emiliano Bedini[c], Matthew Howard Sazinsky[d], Marco Trifuoggi[c], Alberto Di Donato[a], Viviana Izzo[e,*]

[a] *Department of Biology, University Federico II of Naples, Via Cinthia 26, 80127, Naples, Italy*
[b] *Institute of Biosciences and Bioresources, National Research Council of Italy, Via P. Castellino 111, 80131, Naples, Italy*
[c] *Department of Chemical Sciences, University Federico II of Naples, Via Cinthia 26, 80127, Naples, Italy*
[d] *Department of Chemistry, Pomona College, Sumner Hall, 333 N College Way, Claremont, CA, 91711, United States*
[e] *Department of Medicine, Surgery and Dentistry "Scuola Medica Salernitana", University of Salerno, Via S. Allende 2, 84131, Salerno, Italy*

ARTICLE INFO

Keywords:
Rhamnosidase
Flavonoid
Homology modeling
Site-directed mutagenesis
Novosphingobium sp. PP1Y

ABSTRACT

α-L-Rhamnosidases (α-RHAs, **EC 3.2.1.40**) are glycosyl hydrolases (GHs) hydrolyzing terminal α-L-rhamnose residues from different substrates such as heteropolysaccharides, glycosylated proteins and natural flavonoids. Although the possibility to hydrolyze rhamnose from natural flavonoids has boosted the use of these enzymes in several biotechnological applications over the past decades, to date only few bacterial rhamnosidases have been fully characterized and only one crystal structure of a rhamnosidase of the GH106 family has been described. In our previous work, an α-L-rhamnosidase belonging to this family, named RHA-P, was isolated from the marine microorganism *Novosphingobium* sp. PP1Y. The initial biochemical characterization highlighted the biotechnological potential of RHA-P for bioconversion applications. In this work, further functional and structural characterization of the enzyme is provided. The recombinant protein was obtained fused to a C-terminal His-tag and, starting from the periplasmic fractions of induced recombinant cells of *E. coli* strain BL21(DE3), was purified through a single step purification protocol. Homology modeling of RHA-P in combination with a site directed mutagenesis analysis confirmed the function of residues D503, E506, E644, likely located at the catalytic site of RHA-P. In addition, a kinetic characterization of the enzyme on natural flavonoids such as naringin, rutin, hesperidin and quercitrin was performed. RHA-P showed activity on all flavonoids tested, with a catalytic efficiency comparable or even higher than other bacterial α-RHAs described in literature. The results confirm that RHA-P is able to hydrolyze both α-1,2 and α-1,6 glycosidic linkages, and suggest that the enzyme may locate different polyphenolic aromatic moities in the active site.

1. Introduction

α-L-Rhamnosidases (EC 3.2.1.40) are glycosyl hydrolases (GHs) that cleave terminal α-L-rhamnose from a large number of natural products [1]. The action of α-L-rhamnosidases (α-RHAs) has been reported, among others, on different complex substrates such as heteropolysaccharides or glycosylated proteins containing rhamnose units (gellan gum, rhamnogalacturonan and arabinogalactan-proteins) [2–5]. Among α-RHAs substrates, natural flavonoids are gaining much interest among the food and nutraceutic industry. Natural flavonoids are polyphenolic compounds generally characterized by a three-ring structure, which consists of two phenyl rings (A and B) and a heterocyclic ring (C). These molecules are naturally produced in plants in glycosylated forms, showing the presence of either a rutinoside (6-α-L-rhamnosyl-β-D-glucoside) or a neohesperidoside (2-α-L-rhamnosyl-β-D-glucoside) disaccharidic unit bound in different positions. These molecules are very interesting due to their potential antioxidant, antitumor and anti-inflammatory properties [6,7]. In humans, endogenous β-glycosidases and α-L-arabinosidases in the small intestine are responsible for removing the glucose (or possibly arabinose or xylose) moiety from

---

Abbreviations: α-RHA, α-L-rhamnosidase; ABS, absorbance; BSA, bovin serum albumin; *E. coli*, *Escherichia coli*; GHs, glycosyl hydrolases; GTs, glycosyl transferases; IPTG, isopropyl β-D-1-thiogalactopyranoside; LB, Luria Bertani; *N.* sp. PP1Y, *Novosphingobium* sp. PP1Y; OD, optical density; ORF, open reading frame; *p*NPR, *para*-nitrophenyl-α-L-rhamnopyranoside; RT, room temperature; SDS, sodium dodecyl sulphate; TLC, thin layer chromatography

[*] Corresponding author.

E-mail address: vizzo@unisa.it (V. Izzo).

https://doi.org/10.1016/j.abb.2018.04.013
Received 13 February 2018; Received in revised form 12 April 2018; Accepted 16 April 2018
Available online 18 April 2018
0003-9861/ © 2018 Elsevier Inc. All rights reserved.

flavonoids thus facilitating their absorption. These enzymes, however, are not able to cleave terminal rhamnose units, thereby limiting the bioavailability of rhamnosylated flavonoids that are converted to more bioactive forms by the colon microflora [8,9]. Therefore, enzymatic rhamnose removal from potentially bioactive flavonoids may be the key for improving their intestinal absorption and thus their beneficial properties for human health [10,11].

Due to their ability to hydrolyze rhamnose from natural flavonoids, α-RHAs are used in several biotechnological applications. Some examples include the hydrolysis of naringin to improve beverage quality by debittering grapefruit and citrus juices [12,13], and the removal of hesperidin crystals from orange-derived preparations [14]. Other applications of α-RHAs are gaining popularity in the oenological industry, where these enzymes are used to hydrolyze terpenyl glycosides to enhance aroma in wine, grape juices and derived beverages [15]. Moreover, an α-RHA has been implemented for the synthesis of rhamnose-containing chemicals by reverse hydrolysis, suggesting a yet unexplored potential of this enzymatic class in the chemical and pharmaceutical industry [16]. Despite their potential as industrial biocatalysts, to date only a limited number of bacterial rhamnosidases has been fully characterized [17–21]. The commercial preparations of α–L-rhamnosidases, naringinases and hesperidinases available and currently used in oenology, are all isolated from fungal sources such as *Aspergillus niger* and *Penicillium decumbens* [22–24]. The different operational parameters observed among the bacterial and fungal sources, with bacterial RHAs being more efficient at higher pH and temperature, suggest this class of α-RHAs to be an alternative source of biocatalysts to use for carbohydrates biotransformation at high pH values.

Few attempts have been made so far to engineer bacterial α-RHAs to unravel the molecular details underlying their catalytic mechanism, to modify their substrate specificity or to optimize their catalytic efficiency towards different substrates. A major obstacle is the limited number of α-RHAs crystal structures that are currently available among the different families of GH enzymes, which include the GH28, GH78, and GH106 families according to the CAZy database [25–27]. To the best of our knowledge only five crystal structures of bacterial α-L-rhamnosidases are currently available, four of which belong to the GH78 family, which show an $(\alpha/\alpha)_6$ 3D-structure, and include the putative α-L-rhamnosidase BT1001 from *Bacteroides thetaiotaomicron* VPI-5482 [28], the α-L-rhamnosidase B (BsRhaB) from *Bacillus* sp. GL1 [29], the α-L-rhamnosidase (SaRha78A) from *Streptomyces avermitilis* [30], and the α-L-rhamnosidase (KoRha) from *Klebsiella oxytoca* [31].

Among the members of the GH106 family, which groups 319 different sequences, a single 3D-structure has been reported, the BT0986 from *Bacteroides thetaiotaomicron* that shows a $(\beta/\alpha)_8$ barrel and catalyzes the hydrolysis of an α-L-rhamnopyranoside bound to the C2 position of an arabinofuranoside (L-Rha*p*-α-1,2-L-Ara*f*). In this same family of glycosidases only two enzymes have been characterized so far, and the reaction mechanism and the catalytic residues have been inferred from the 3D-structure of BT0986 [27].

In our previous work, a novel α-L-rhamnosidase was isolated from the marine microorganism *Novosphingobium* sp. PP1Y, a Gram-negative bacterium isolated from a polluted marine environment in the Pozzuoli harbor (Naples, Italy) [32–34]. The α-RHA from *N.* sp. PP1Y was isolated from the native bacterium, expressed in *E. coli* and partially characterized [35]. This enzyme, named RHA-P, was characterized as an *inverting* monomeric glycosidase of *ca.* 120 kDa belonging to the GH106 family. A preliminary biochemical characterization using the synthetic substrate *p*NPR (*para*-nitrophenyl-α-L-rhamnopyranoside) showed that RHA-P has moderate tolerance to organic solvents and optimal activity between pH 6.0–7.5 at 35–45 °C. Moreover, TLC analysis showed that RHA-P is able to hydrolyze rhamnose from natural flavonoids such as naringin, rutin and neohesperidin dihydrochalcone [35].

Here we extend our previous analysis on the functional and structural features of RHA-P, subcloning the enzyme with a C-terminal His-tag, improving the yield and purity of the recombinant protein, and making way for a detailed kinetic characterization of RHA-P activity on natural flavonoids. In addition, an homology model of RHA-P based on BT0986 3D-structure in combination with the characterization of the substrate specificity of both the wild type and mutant enzymes, revealed a new specificity for GH106 enzymes expanding the possible applicative field of this glycosidase family.

2. Materials and methods

2.1. General

Bacterial cultures, plasmid purifications and transformations were performed according to Sambrook et al. [36]. Bacterial growth was followed by measuring the optical density expressed as OD/mL at 600 nm (from now on defined as OD_{600}). When cells were centrifuged and the supernatant discarded, we indirectly referred to the wet weight of the collected cells as “total OD_{600}”, a value obtained by multiplying the volume of cells suspension by the optical density.

The pET22b(+) expression vector and BL21(DE3) *E. coli* strain were from Amersham Biosciences, whereas the Top10 *E. coli* strain was purchased from Life Technologies.

The recombinant DNA polymerase used for PCR amplification was TAQ Polymerase from Microtech Research Products, and dNTPs were purchased from Promega.

The Wizard PCR Preps DNA purification system for elution of DNA fragments from agarose gels was from Promega. The QIAprep Spin Miniprep Kit for plasmid DNA purification was from QIAGEN. T4 DNA Ligase was from Promega, enzymes and other reagents for DNA manipulation were from New England Biolabs. Oligonucleotides were synthesized by MWG-Biotech (http://www.mwg-biotech.com).

Ni Sepharose 6 Fast Flow was obtained from GE Healthcare. Solvents used in enzymatic assays were obtained from Applichem (DMSO). Both chicken egg white lysozyme, and *p*-nitrophenyl-α-L-rhamnopyranoside (*p*NPR) were purchased from Sigma Aldrich, IPTG (isopropyl β-D-1-thiogalactopyranoside) was obtained from Applichem.

TLC silica gel plates were obtained from E. Merck (Darmstadt, Germany). Unless otherwise stated, all other chemicals were purchased from Sigma Aldrich.

2.2. Construction of the pET22b(+)/rha-his expression vector

The *orf* PP1Y_RS05470, coding for the wild-type RHA-P protein, was previously cloned in pET22b(+) vector; the corresponding recombinant plasmid was named pET22b(+)/*rha-p* [35]. To obtain the pET22b(+)/*rha-his* plasmid, which expressed recombinant RHA-P with a C-terminal His-tag (from now on defined as RHA-Phis), our cloning strategy involved a single point mutation (TGA to CGA) to convert the stop codon to an arginine, and the extension of *rha-p* coding sequence on the pET22b(+) plasmid to include a linker sequence and the coding region of a 6-His-tag domain at 3′of the gene. The plasmid pET22b(+)/*rha-dw*, bearing only the C-terminal 1625 bp-long coding fragment of *rha-p* and already available in our laboratory [35], was chosen as template for the PCR procedure that allowed the insertion of the 6-His-tag at the 3′ of the gene. The use of pET22b(+)/*rha-dw* as template was useful to avoid the amplification of the entire *rha-p* gene, which is 3441 bp long. The mutagenesis experiment was performed using specific designed complimentary primers RHAmutUP (5′ACCACGGGCGGGCATCGAGCCG**TCG**ACAAGC3′) and RHAmutDW (5′GCTTGT**CGA**CGGCTCGATGCCCGCCCGTGGT3′) containing the desired mutated codon (highlighted in bold). Quickchange II Site Directed Mutagenesis kit (Agilent technologies) was used for this experiment, following the manufacturer protocol. The mutation was verified by DNA sequencing. A fragment containing the mutated codon was identified between the single restriction sites *Aat*II - *Not*I. The cassette was excised from the pET22b(+)/*rha-dw* plasmid, purified from an agarose gel and subcloned in a pET22b(+)/*rha-p* vector, containing the entire *rha-p* gene, by digesting both mutagenized fragment and pET22b(+)/*rha-p* with *Aat*II/*Not*I restriction endonucleases. Digestion products were separated by agarose gel electrophoresis, purified and ligated. Ligation products were used to transform *E. coli* Top10 competent cells and the resulting recombinant plasmid, named pET22b(+)/*rha-his* was verified by DNA sequencing by MWG-Biotech.

2.3. Mutagenesis of the pET22b(+)/rha-his expression vector and construction of RHA-Phis single mutants

A mutagenesis cassette containing all the residues selected for modification was identified in the pET22b(+)/*rha-his* sequence, between the single restriction sites *Aat*II and *Kpn*I. The mutagenesis cassette was excised from pET22b(+)/*rha-his* vector, purified from agarose gel and ligated into a pGEM-3Z vector digested using the same enzymes. Ligation products were used to transform *E. coli* Top10 competent cells and the resulting recombinant plasmid, named pGEM-3Z/*rha*, was verified by DNA sequencing. The plasmid pGEM-3Z/*rha* was then used in five different mutagenesis experiments in which residues D503, E506, D552, E644, and D763 were singularly mutated to alanine. For D503, D552 and D763 the GAT codon was converted in GTC, for E506 the GAG codon was converted in GCG and for E644 the GAA codon was converted in GCA. Therefore, five pairs of specific mutagenic primers were designed (Table S1). The QuikChange II Site-Directed Mutagenesis kit (Agilent Technologies) was again used on pGEM-3Z/*rha* in five separate reactions, following the manufacturer protocol. Mutagenized fragments in pGEM-3Z/*rha* vector were then individually subcloned into pET22b(+)/*rha-his* vector using *Aat*II/*Kpn*I restriction endonucleases. Mutagenized clones (generally indicated as pET22b(+)/D-EXXXA) were named pET22b(+)/D503A, pET22b(+)/E506A, pET22b(+)/D552A, pET22b(+)/E644A, pET22b(+)/D763A, and all sequences were verified by DNA sequencing.

2.4. Recombinant expression

Protein expression was carried out in *E. coli* BL21(DE3) cells transformed by heat shock with either pET22b(+)/*rha-his* or pET22b(+)/D-EXXXA plasmids. All the media contained 100 μg/mL ampicillin.

2.4.1. Analytical expression

Analytical expression experiments were performed to evaluate the best conditions for RHA-Phis recombinant expression as already used for *wt*-RHA-P. In a first experiment, *E. coli* BL21(DE3) competent cells transformed with pET22b(+)/*rha-his* [37], were inoculated in a sterile 50 mL Falcon tube containing 12.5 mL of either LB or LB with 0.5 M NaCl (LB-N). Cells were grown under constant shaking at 37 °C up to an OD_{600} of 0.6–0.7. These samples were then diluted 1:100 in 2 sterile 50 mL Falcon tubes containing 12.5 mL of LB, and 2 sterile 50 mL Falcon tubes containing 12.5 mL of LB-N. These cells were grown under constant shaking at 37 °C to OD_{600} of 0.7–0.8. Then, RHA-Phis recombinant expression was induced with 0.1 mM IPTG and growth was continued with constant shaking for 3 h, at either 23 °C or 37 °C for both the LB and LB-N cultures. Cells were collected by centrifugation (5524 × *g* for 15 min at 4 °C) and suspended in 25 mM MOPS pH 6.9 to obtain a concentration of 14 OD_{600}. Cells were disrupted by sonication (12 times for a 1-min cycle, on ice) and an aliquot of each lysate was centrifuged at 22100 × *g* for 10 min at 4 °C. Both soluble and insoluble fractions were analyzed by SDS-PAGE. The soluble fraction was assayed for the presence of α-RHA enzymatic activity.

A second analytical expression experiment was performed as described above by expressing *E. coli* BL21(DE3) cells transformed with pET22b(+)/*rha-his* in media containing LB, either alone or with 1 mM (LB-BS1), 5 mM (LB-BS5) or 10 mM (LB-BS10) of both betaine and sorbitol. Expression in these cultures was induced at OD_{600} of 0.7–0.8 with either 0.1 mM or 1 mM IPTG. After 3 h under constant shaking at 23 °C, cells were collected by centrifugation (5524 × *g* for 15 min at 4 °C), suspended in 25 mM MOPS pH 6.9, 5% glycerol at a final concentration of 14 OD_{600}, and disrupted by sonication (12 times for a 1-min cycle, on ice). An aliquot of each cell lysate was centrifuged at 22,100 × *g* for 10 min at 4 °C; afterwards, soluble and insoluble fractions were analyzed by SDS-PAGE. The soluble fraction was evaluated for the presence of α-RHA enzymatic activity.

2.4.2. Large scale expression

E. coli BL21(DE3) cells freshly transformed with pET22b(+)/*rha-his,* pET22b(+)/D503A, pET22b(+)/E506A, pET22b(+)/D552A, pET22b(+)/E644A, or pET22b(+)/D763A, were inoculated from the transformation plate into 5 mL of LB and incubated at 37 °C with constant shaking up to OD_{600} ~ 0.5. The first preinoculum was diluted 1:10 in 4 sterile 50 mL Falcon tubes containing 12.5 mL of LB and incubated at 37 °C up to OD_{600} of 0.6–0.7. This latter was diluted 1:40 in four 2 L Erlenmeyer flasks containing each 500 mL of LB-BS5 and incubated at 37 °C up to OD_{600} of 0.7–0.8. Expression of RHA-Phis or RHA-Phis mutants, was induced with 1 mM IPTG and growth was continued for 3 h with constant shaking at 23 °C. Cells were collected by centrifugation (5524 × *g* for 15 min at 4 °C) and the purification of the periplasmic fraction was performed.

2.5. Analytical periplasm purification

Using the expression protocol described above, after 3 h of induction, 2 aliquots of 8 total OD_{600} of *E. coli* BL21(DE3) cells expressing RHA-Phis were collected by centrifugation (5524 × *g* for 15 min at 4 °C). One aliquot, representing the total cell fraction of the induced cell cultures, was stored at −80 °C until use. The second aliquot was suspended in a lysis buffer containing 100 mM Tris/HCl pH 7.4, 20% sucrose and 1 mM EDTA at final concentration of 84 OD_{600}. Native lysozyme from chicken egg white (~200 mg) (L6876 Sigma-Aldrich) was added to the solution and the cell suspension was incubated at RT for 15 min, after which the volume of the suspension was doubled by adding distilled water, incubated at RT for additional 15 min, and then centrifuged at 5524 × *g* for 15 min at 4 °C. The supernatant, containing the periplasmic fraction of the induced cultures, was collected, whereas the pellet, containing the protoplastic fraction of the induced cultures, was stored at −20 °C until further use.

Both the total cell fraction and the protoplastic fraction were thawed and suspended in 25 mM MOPS pH 6.9, 5% glycerol at a final concentration of 14 total OD_{600}. The samples were disrupted by sonication (12 times for a 1-min cycle, on ice), and the lysates were centrifuged at 22100 × *g* for 10 min at 4 °C. All fractions (total cell lysate, periplasm and cytoplasm) were analyzed by SDS-PAGE and the α-RHA activity was evaluated in each sample.

2.6. RHA-Phis and RHA-Phis mutants purification

The cell paste of the induced recombinant cells was suspended in 100 mM Tris/HCl pH 7.4, 20% sucrose and 1 mM EDTA, and the periplasm purification was performed as described in the previous paragraph. To remove membrane debris, the sample was subjected at this stage to an additional centrifugation at 22100 × *g* for 60 min at 4 °C. The supernatant was collected, filtered through a 0.45 μm PVDF Millipore membrane and incubated overnight with 2.5 mL of Ni Sepharose FF resin equilibrated in 50 mM Tris/HCl pH 7.4, 5% glycerol (buffer A) in batch with gentle shaking at 4 °C. The unbound portion was removed by centrifugation at 3200 × *g* for 10 min at 4 °C. The resin was washed with 15 column volumes of buffer A, and protein elution was carried out in column (1 cm diameter, 17 cm length), using a 2 step-gradient at a flow rate of ~1 mL/min. In the first step, five column volumes of buffer A containing 30 mM imidazole were used. In the second step, the protein was eluted using five column volumes of buffer

A containing 250 mM imidazole. One mL fractions were collected and the chromatogram was obtained by analyzing the absorbance of the fractions at 280 nm. The presence of either RHA-Phis or RHA-Phis mutant activity was detected by using the *p*NPR assay. Relevant fractions were analyzed by SDS-PAGE and pooled. Imidazole was removed from pooled fractions by repeated cycles of ultrafiltration and dilution with buffer A. The protein sample was then purged with nitrogen, and stored at −80 °C until use.

2.7. Enzyme activity assays

α-RHA activity was determined at room temperature using *p*NPR as substrate (*p*NPR assay). Unless otherwise stated, each reaction mixture contained, in a final volume of 0.5 mL, 50 mM MOPS pH 6.9, 600 μM *p*NPR and variable amounts of enzyme. The reaction was blocked after 6 and 12 min by adding 0.5 M Na_2CO_3. The product, *p*-nitrophenolate (*p*NP), was detected spectrophotometrically at 405 nm. The extinction coefficient used was $\varepsilon_{405nm} = 18.2\,mM^{-1}\,cm^{-1}$. One unit of enzyme activity was defined as the amount of the enzyme that releases 1 μmol of *p*NP per minute at 25 °C.

2.7.1. Whole cells pNPR assay

Samples were assayed in a total volume of 0.5 mL. The reaction mixtures contained 50 mM MOPS pH 6.9, *p*NPR at a final concentration of 600 μM, and variable amounts of either RHA-Phis or RHA-Phis mutants expressing cells. Cells were preincubated in M9 minimal medium containing 0.4% glucose, and assayed at a final cell density ranging from 0.05 to 0.2 OD_{600}. Reactions were incubated at RT, and after 6 and 12 min aliquots of 0.2 mL were collected and centrifuged at 5524 × *g* for 3 min in order to remove whole cells from the assay mixture. Supernatants were collected and directly added to 0.5 M Na_2CO_3 in cuvette. The product, *p*-nitrophenolate (*p*NP), was detected spectrophotometrically as described above. Specific Activity was reported as mU/OD_{600}.

2.7.2. Assays on synthetic substrates

Three synthetic rhamno-oligosaccharides (Fig. S1 of the Supplemental Material) were tested as substrates of RHA-P enzymatic activity. Hydrolysis reactions were performed using 0.9 mg/mL of each substrate in a final volume of 0.12 mL containing 50 mM MOPS buffer pH 6.9. Reactions were carried out overnight at 37 °C, in the presence of 2.5 U of RHA-Phis. Hydrolysis products were monitored by TLC analysis (system solvent EtOAc: MeOH: H_2O, in a 70:20:10 v/v ratio). Compounds on TLC plates were visualized with 4% α-naphthol in 10% sulphuric acid prepared in ethanol, followed by charring. In all experiments, TLC standard solutions of pure reagents and products were used for comparison.

2.7.3. Steady state kinetics with pNPR

Reactions were carried out according to the general *p*NPR assay. Kinetic parameters were obtained at pH 6.9 using a *p*NPR concentration in the range 0.025–2 mM. RHA-Phis was used at a final concentration of 0.79 nM, whereas mutants RHA-PhisD552A and RHA-PhisD763A were used at a final concentration of 0.97 nM and 3.38 nM, respectively. All kinetic parameters were determined by a non-linear regression curve using GraphPad Prism (GraphPad Software; www.graphpad.com).

2.7.4. Steady state kinetics with flavonoids

Reactions were carried out in 0.2 mL of 50 mM MOPS buffer pH 6.9 at 37 °C in the presence of either 0.25 U or 0.5 U of RHA-Phis, using a concentration of flavonoids (naringin, rutin, hesperidin and quercitrin) in the range 0.01–10 mM. Reactions were immediately loaded on the HPLC column 2 and 4 min after diluting 20 μL of each sample in 1 mL of 20 μM sucrose (internal standard) prepared in water. Samples were analyzed to determine rhamnose release using a Dionex ICS3000 HPLC (ThermoScientific – Dionex) with pulsed amperometric detector (HPAEC-PAD), and equipped with both a Carbopac column PA-100 (4 × 250 mm) anion exchange column (ThermoScientific – Dionex) and a Carbopac PA-100 (4 × 50 mm) column guard. The mobile phase was composed of 0.1 M NaOH in water (solvent A). Isocratic elution of the analytes was performed at a flow rate of 1 mL/min for 20 min. Free rhamnose was identified as reaction product by its retention time (~3.1 min) and quantified based on peak area. A calibration curve was obtained using standard solutions of rhamnose at 10 μM, 20 μM and 40 μM with 20 μM of sucrose used in each solution as internal standard. One unit of rhamnosidase activity was defined as the amount of enzyme required to release 1 μmol of rhamnose in 1 min. All kinetic parameters were determined by a non-linear regression curve using GraphPad Prism (GraphPad Software; www.graphpad.com).

2.8. Treatment of RHA-Phis with EDTA and recovery of the enzymatic activity

Five aliquots of 1 μM RHA-Phis were separately incubated in 50 mM Tris/HCl pH 7 and 2 mM EDTA; the chelating agent was used in a 2000 fold excess compared to protein concentration. Samples were incubated for 5 min at RT and for 30 min in a thermocycler, using a temperature gradient of 1 °C/min increase from 25 °C to 38 °C, 30 s at 38 °C and 1 °C/min decrease from 38 °C back to 25 °C. Samples were then centrifuged at 22100 × *g* for 10 min at 4 °C, and the supernatants were transferred to clean tubes.

To recover the enzymatic activity, 3 mM, 10 mM and 50 mM of either $CaCl_2$, or $MgCl_2$, or $MnCl_2$, or $ZnCl_2$ were separately added to aliquots of EDTA-treated RHA-Phis. Samples were incubated for 5 min at RT; afterwards, the same temperature gradient used for the depletion described above was used to facilitate the metal insertion in the protein active site. The *p*NPR assay was performed after each step of incubation, and the specific activity was evaluated as described in previous paragraphs.

2.9. Homology modeling

The homology model of RHA-P was prepared using the "DeepView Project Mode" option of the Swiss-Model server [38,39]. Sequences of RHA-P and of the BT0986 rhamnosidase from *Bacteroides thetaiotaomicron* were first aligned using the *lalign* server (https://embnet.vital-it.ch/software/LALIGN_form.html) [40]. Next, an initial raw model was prepared using the DeepView/Swiss-PdbViewer software [39]. The sequence of RHA-P was manually superimposed to the structure of the BT0986 rhamnosidase from *Bacteroides thetaiotaomicron* (pdb code, 5MWK) using the *lalign* alignment as reference, then the alignment was manually adjusted. Finally, the corresponding DeepView project file obtained was uploaded to the Swiss-Model server. BT0986 structures and the RHA-P model were analyzed using PyMol (http://www.pymol.org/).

2.10. Analytical methods

Protein concentration was measured with the Bio-Rad Protein System [37] using bovine serum albumin (BSA) as a standard. Polyacrylamide gel electrophoresis was carried out using standard techniques [41]. SDS–PAGE 15% Tris–glycine gels were run under denaturing conditions and proteins were stained with Coomassie brilliant blue G-250. "Wide range" (200–6.5 kDa) molecular weight standard was from Sigma (ColorBurst™ Electrophoresis Marker).

3. Results

3.1. Cloning, recombinant expression and purification of RHA-Phis

Previous work on the untagged recombinant protein rRHA-P was hampered by a non-optimal purification yield that could not be improved any further [35]. Therefore, we decided to fuse a His-tag to the C-terminus of rRHA-P, thus facilitating the purification of the protein by affinity chromatography.

The plasmid pET22b(+)/*rha-his,* expressing recombinant RHA-P to which a 6-His-tag has been added to the C-terminal (from now on defined as RHA-Phis), was used to transform *E. coli* BL21(DE3) competent cells. Experiments were performed in order to evaluate the best conditions for RHA-Phis recombinant expression. Growth and induction performed at 37 °C in LB medium led to the foremost presence of the recombinant protein in the inclusion bodies. Therefore, we set up analytical expression experiments to evaluate the recombinant expression of RHA-Phis either in the presence or absence of high salt concentrations, and after induction at either 37 or 23 °C.

The SDS-PAGE analysis of the soluble and insoluble fractions of induced cultures showed this time the presence of a protein band with the molecular weight expected for RHA-Phis almost exclusively in the soluble fraction obtained at an induction temperature of 23 °C (data not shown).

Afterwards, the effect of adding to the culture medium different concentrations of IPTG as inducer, and betaine and sorbitol as “chemical chaperones” [35], was evaluated. The data showed that the best expression conditions involved an induction performed with 1 mM IPTG at 23 °C using LB supplemented with 5 mM of both betaine and sorbitol (Fig. S2).

In order to verify the presence of RHA-Phis in the periplasmic fraction, an additional analytical expression experiment was performed. Samples corresponding to the total cell extract, the periplasmic and the cytoplasmic fractions were analyzed by SDS-PAGE and catalytic activity was individually evaluated in each sample. Results showed that only a negligible part of the total α-RHA activity present in the crude cell extract was detected in the cytoplasmic fraction, whereas most of the activity could be retrieved in the periplasmic fraction (Fig. 1A). This suggested that RHA-Phis was mainly sorted in the periplasm of the recombinant *E.coli* strain, which was also evident from the SDS-PAGE analysis (Fig. 1B).

Large-scale recombinant expression of RHA-Phis was performed by using the optimized conditions described above. RHA-Phis purification was carried out using a single step affinity chromatography; a final purification factor (PF) of 12.7 and a total yield of 25.3% were obtained.

To evaluate the potential effect of the His-tag on protein function, specific activities of the purified RHA-Phis and untagged RHA-P were compared [35]. Data highlighted a negligible decrease of the specific activity on the synthetic substrate *p*NPR, with RHA-Phis showing a residual 87% rhamnosidase activity compared to the untagged protein. In addition, kinetic analysis confirmed that K_M and k_{cat}/K_M values found for RHA-Phis ($157 \pm 22\,\mu M^{-1}$ and $4.4\,s^{-1}\mu M^{-1}$ respectively) were comparable to the values obtained for the untagged protein ($160 \pm 17\,\mu M^{-1}$ and $4.6\,s^{-1}\mu M^{-1}$).

3.2. RHA-P alignment and homology modeling

To identify RHA-P catalytic residues, two complementary approaches were used.

First, the RHA-P amino acid sequence was aligned to 130 GH106 family homologs by using the ClustalO multialignment tool (https://www.ebi.ac.uk/Tools/msa/clustalo) (Table 1 and Table S2).

Five highly conserved aspartic or glutamic acid residues, with a > 90% consensus retrieved among all the sequences aligned, were identified: D503, E506, D552, E644, D763.

In addition, a RHA-Phis homology model was created using the structure of rhamnosidase BT0986 from *Bacteroides thetaiotaomicron* (pdb code, 5MWK) as a template, the only available so far for the GH106 family. The 5MWK structure, bearing the E461Q mutation to prevent the hydrolysis of the substrate, was chosen among those available because it shows the presence of a branched oligosaccharide bound into the active site. The oligosaccharide has, at the non-reducing end, a rhamnose/arabinose disaccharidic unit with a α-1,2 bond (L-Rha-α-1,2-L-Ara), which is hydrolyzed by BT0986.

The alignment used for the modeling is shown in Fig. S3. Even though the percent identity is not high (27%) and several insertions and deletions are present, the majority of the differences are located on the surface of the protein far from the active site (Fig. 2A). Only a deletion and two insertions in RHA-Phis are in loops contributing to the active site pocket (Fig. S3). Of note, these three loops are close to each other and are located in the active site pocket corresponding to the arabinose binding sub-site in BT0986. By contrast, the calcium binding residues and the rhamnose sub-site appears well conserved (Fig. S3, Fig. 2B and C). In the BT0986 structure, five residues contribute to the binding of the Ca^{2+} ion, namely E593, E538, D458, S459 and E561. Of these five residues, only E561, whose mutation into alanine inactivates BT0986 [27], is not conserved in RHA-Phis that has an alanine in this position. It should be kept in mind, however, that the actual role of E561 in BT0986 is still under investigation. In the BT0986 structure without the bound substrate (pdb code, 5MQN), E561 does not bind the calcium ion. However, in both BT0986 structures, E561 forms an ionic couple

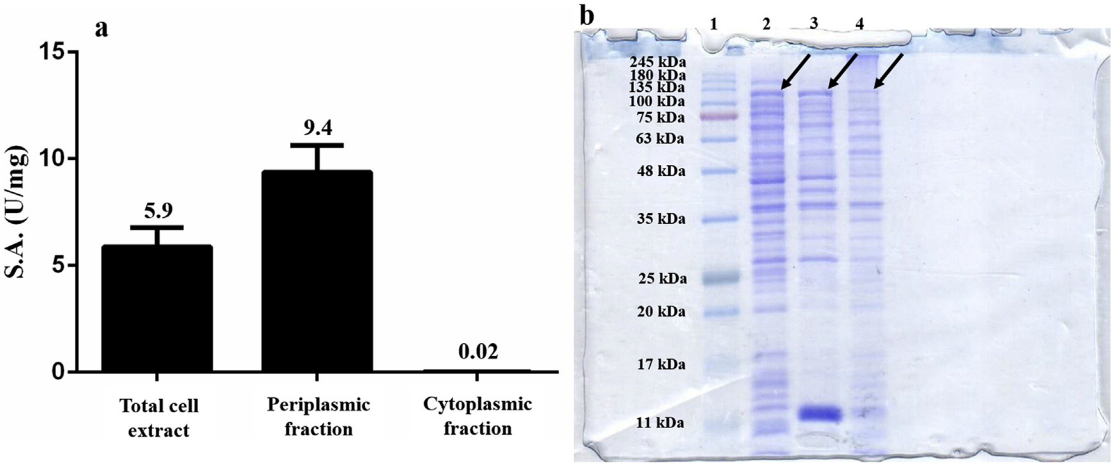


Fig. 1. SDS-PAGE analysis of RHA-Phis expression in the different cellular compartments of induced recombinant cells of strain BL21(DE3). a) Rhamnosidase specific activity (S.A), expressed in U/mg, of *E. coli* BL21(DE3) cells transformed with pET22b(+)/*rha-his* plasmid. S.A. values of total cell extract, periplasmic and cytoplasmic fractions are reported. b) SDS-PAGE analysis of the total cell extract, periplasmic and cytoplasmic fraction of induced *E. coli* BL21(DE3) cells transformed with pET22b(+)/*rha-his* plasmid. Lane 1 = MW standard (numbers reported refer to the MW expressed in kDa). Lane 2 = total cell lysate. Lane 3 = periplasmic fraction. Lane 4 = cytoplasmic fraction. Black arrows in all lanes indicate a band with the molecular mass expected for RHA-Phis.

Table 1
Highly conserved residues. Table showing the highly conserved carboxylic residues in the GH106 family sequences used for the alignment. The corresponding selected residues are indicated both in the RHA-P sequence and in the BT0986 sequence (used for the homology model).

Multialignment posix	Conservation	Quality	Consensus %	Consensus residue	RHA-P residue	BT0986 residue
1106	8	361.92047	88.46	D 88%	D465	D421
1153	10	375.69186	96.15	D 96%	D503	D458
1161	10	375.40494	95.38	E 95%	E506	Q461
1189	2	275.2199	67.69	D 67%	D527	D482
1238	10	334.8154	90.77	D 90%	D552	D507
1296	1	318.931	81.54	D 81%	D603	D556
1301	0	288.4868	79.23	E 79%	A608	E561
1361	3	366.1997	96.15	E 96%	E644	E593
1390	0	304.88068	81.54	D 81%	D665	D614
1477	2	312.0073	80.77	D 80%	D733	D683
1569	4	337.38684	90.00	D 90%	D763	D713

with K571, which in turn binds to the arabinose residue; therefore, the E561-K571 might be involved in substrate recognition. RHA-Phis may not have this arabinose-binding sub-site, because K571 is replaced with a leucine. Moreover, a loop containing W276, another residue contributing to the arabinose binding site in BT0986, is deleted in RHA-Phis (Fig. S3 and Fig. 2C). In an equivalent position in RHA-Phis, a loop bearing an insertion of two residues, D578 and H588, is present (Fig. 2C).

All the residues that contact rhamnose in the structure of BT0986, are also conserved in RHA-Phis or at least show conservative mutations (Fig. 2C). The putative acid catalyst, E506, which would correspond to E461 in BT0986, is also conserved (Fig. 2B and C).

Overall, the model suggests that the catalytic mechanism of RHA-Phis might be very similar to that of BT0986, but several details indirectly point to a different substrate specificity.

This hypothesis is also supported by the evidence that BT0986 is barely active on *p*NPR and has only a significative activity on oligosaccharides containing rhamnose, such as the rhamnose/arabinose units (α-L-Rha-α-1,2-L-Ara), for which it shows a k_{cat}/K_M of $23.5\,s^{-1}\,mM^{-1}$ [27]. On the other hand, RHA-Phis is active on arylrhamnosides, showing a k_{cat}/K_M of $4.4\,s^{-1}\mu M^{-1}$ on *p*NPR, and has k_{cat}/K_M values ranging from 200 to $400\,s^{-1}\,mM^{-1}$ on natural flavonoids. All these data point to a specificity of RHA-P for substrates containing aryl-groups and aromatic moieties, rather than poly- and oligosaccharidic compounds.

3.3. *Expression and characterization of RHA-Phis mutants*

To investigate the role of the carboxylic acid residues identified from the multiple alignment approach, an alanine scanning site directed mutagenesis was carried out. According to the homology model performed using the template BT0986, among these five residues initially identified, aminoacids D503, E506 and E644 were most likely essential for catalysis.

RHA-Phis/D503A, RHA-Phis/E506A, RHA-Phis/D552A, RHA-Phis/E644A and RHA-Phis/D763A single mutants were obtained as already described in this work for RHA-Phis. After optimized conditions were validated also for the expression of RHA-Phis mutants in *E. coli* BL21(DE3) (data not shown), large-scale expression experiments were performed. After 3 h of induction, an initial screening for α-RHA activity was carried out using *E. coli* BL21(DE3) whole cells. No detectable activity was found for mutants RHA-Phis/D503A and RHA-Phis/E644A, and only 0.60% of residual activity was obtained for mutant RHA-Phis/E506A. A marked reduction of α-RHA activity for mutants RHA-Phis/D552A (11.3%) and RHA-Phis/D763A (12.2%) was observed (Fig. 3A).

RHA-Phis mutants were purified following the same protocol used for RHA-Phis, and the specific activity of all mutants was compared to *wt* RHA-Phis. In this case also, RHA-Phis/D503A and RHA-Phis/E644A were found to be inactive; RHA-Phis/E506A had a negligible residual activity (0.05%). Mutants RHA-Phis/D552A and RHA-Phis/D763A showed a residual activity of 47.9% and 15.8%, respectively (Fig. 3B).

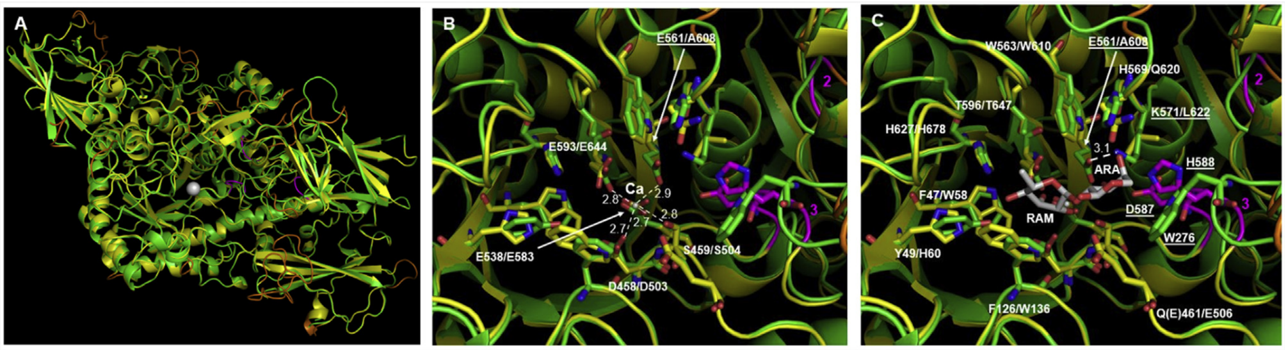


Fig. 2. Comparison between the structure of BT0986 from *Bacteroides thetaiotaomicron* (pdb code, 5MWK) and the homology model of RHA-P. Panel A shows the structure and the model as green and yellow cartoons, respectively. In the model of RHA-P, loops with insertions or deletions with respect to BT0986 are coloured in orange. A loop with a deletion (1) and two loops with insertions (2 and 3), close to the active site, are shown in magenta. Magenta arrows show the variation of the three loops with respect to the corresponding loops in BT0986. Panels B and C show a close up of the active site with an orientation approximately similar to that shown in panel A. Significant residues are shown as sticks coloured by atom type (oxygen, red; nitrogen, blue; carbon atoms of the proteins are coloured as the cartoons shown in panel A; carbon atoms of the disaccharide in panel C are coloured in grey). For each pair of corresponding residues, the residue number of BT0986 is shown first. Significantly different residues in RHA-P and BT0986 are underlined. Panel B shows the residues interacting with the calcium ion in BT0986 (distances are in Angstrom) and the corresponding residues in RHA-P. Panel C shows the disaccharide rhamnose (RAM)/arabinose (ARA) and the residues contacting the two sugars (RAM: F126/W136, Y49/H60, F47/W58, H627/H678, T596/T647; ARA: H569/Q620, K571/L622, W276 in BT0986, D578 and HY588 in rRHA-P; W563/W610, E561/A608 and E461/E506 are at the boundary between the rhamnose and the arabinose sub-sites). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

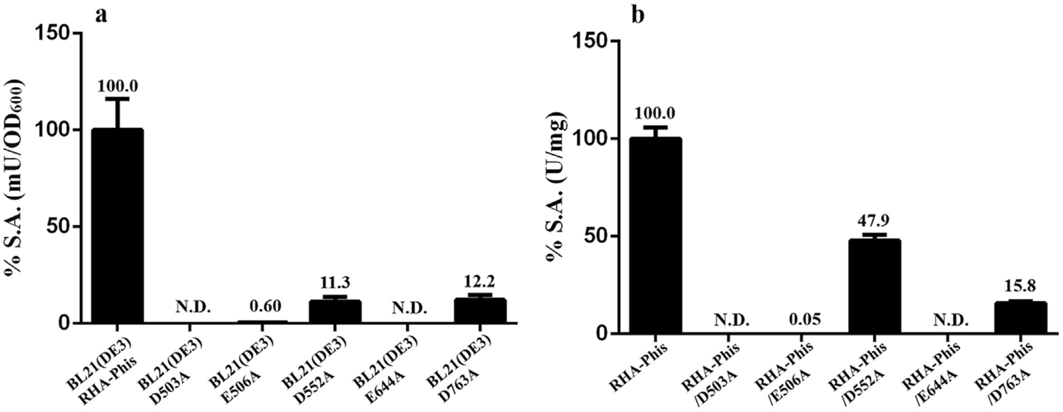


Fig. 3. Residual specific activity of RHA-Phis mutants. a) pNPR assays on whole cells of *E. coli* BL21(DE3) transformed with pET22b(+)/D-EXXXA, compared to the control recombinant cells transformed with pET22b(+)*/rha-his*. Rhamnosidase specific activity is reported as the percentage of residual activity, expressed as mU/OD_{600}, compared to the control. b) pNPR assays of purified RHA-Phis mutants, compared to wt RHA-Phis. Rhamnosidase specific activity is reported as percentage of residual activity, expressed as U/mg, compared to the wt protein.

These data suggested that residues D503, E644 and E506 might have an important role in the catalytic mechanism of the enzyme.

Kinetic parameters on *p*NPR were measured for RHA-Phis/D552A and RHA-Phis/D763A. The reaction rate (μM/min) was plotted as a function of the substrate concentration (μM), showing a typical Michaelis-Menten trend, which is reported in Fig. S4. Table 2 summarizes the kinetic constants measured for RHA-Phis and the active mutants. k_{cat}/K_M values equal to the 65% and 21% of *wt* RHA-Phis k_{cat}/K_M were obtained for RHA-Phis/D552A and RHA-Phis/D763A, respectively. In RHA-Phis/D552A this reduced catalytic efficiency is mainly due to a significative variation in the turnover number, defined by the k_{cat} value, for which a 34% decrease is observed. On the other hand, a remarkable 82% decrease in k_{cat} was retrieved for RHA-Phis/D763A, while the K_M was similar within the experimental error.

3.4. Evidence for calcium presence in RHA-Phis

The presence of a calcium ion essential for catalysis has been reported in some GH families, although with different functions. Three crystal structures currently available for GH106 rhamnosidases confirmed the presence of this essential cofactor [27,29,30].

To verify the presence of Ca^{2+} in RHA-Phis, we first incubated the enzyme in the presence of different concentrations of EDTA and then performed the standard *p*NPR assay. Results presented in Fig. 4A showed a marked reduction of RHA-Phis specific activity in line with the increase of EDTA concentration in the assay, ranging from 0.1 to 1 mM. The recovery of the enzymatic activity was evaluated after incubating the protein with 2 mM EDTA and then directly adding to the sample either $CaCl_2$ or $MgCl_2$, $MnCl_2$, $ZnCl_2$ at concentrations ranging from 3 to 50 mM (Fig. 4B).

We confirmed an almost total inactivation of RHA-Phis after incubation with 2 mM EDTA (4% of residual activity). An almost complete recovery of the enzymatic activity was achieved after incubating the EDTA-treated enzyme with either Ca^{2+} concentration used in the assay. Noteworthy, a 56%, 92% and 79% of the initial rhamnosidases activity was recovered after incubation, respectively, with 3 mM, 10 mM and 50 mM of Mg^{2+}. As evident from Fig. 4B, significantly lower percentages of recovery were obtained in the presence of increasing concentrations of Mn^{2+}, leading to a maximum recovery of 61% in the presence of 50 mM of this metal cation. By contrast, only 3 mM of Zn^{2+} led to a 30% of recovery of the enzymatic activity, while higher concentrations of this element promptly inactivated RHA-Phis catalytic activity.

Table 2
***p*NPR kinetic constants of RHA-Phis and mutants**. RHA-Phis, RHA-Phis/D552A and RHA-Phis/D763A kinetic constants using *p*NPR as substrate.

Protein	KM (μM)	Kcat (sec-1)	kcat/KM (sec-1 μM-1)
Rha-Phis	157 (± 22)	692 (± 33)	4.4 (± 0.9)
Rha-Phis D552A	159 (± 23)	458 (± 21)	2.9 (± 0.5)
Rha-Phis D763A	138 (± 15)	126 (± 6)	0.9 (± 0.2)

3.5. Kinetic characterization of RHA-Phis on natural and synthetic compounds

To gain a deeper insight into the substrate specificity of the enzyme, experiments were performed to evaluate the kinetic behavior of RHA-Phis on different natural flavonoids such as naringin, rutin, hesperidin and quercitrin. The corresponding chemical structures of the flavonoids are reported in Fig. 5 and kinetic plots are presented in Fig. S5. The kinetic constants summarized in Table 3 highlight that RHA-Phis is active towards all the substrates and is able to hydrolyze single rhamnose units, with either α-1,2 or α-1,6 glycosidic linkages. The k_{cat}/K_M values obtained were highest for rutin and naringin ($586 \pm 172\,s^{-1}\,mM^{-1}$ and $360 \pm 70\,s^{-1}\,mM^{-1}$ respectively), and lower for hesperidin and quercitrin ($190 \pm 32\,s^{-1}\,mM^{-1}$ and $80 \pm 16\,s^{-1}\,mM^{-1}$ respectively). These k_{cat}/K_M values are all higher than those observed for the α-RHAs from *Fusobacterium* K-60 and *Pediococcus acidilactici* [42,43]. Conversely, K_M constants for all the tested flavonoidic compounds are ca. 50 times higher than the ones reported for *Lactobacillus plantarum*, and *Pediococcus acidilactici* [19,42].

RHA-Phis specific activities measured on these compounds are of 104.3 ± 4.8 U/mg for naringin, 8.8 ± 0.5 U/mg for rutin, 6.2 ± 0.2 U/mg for hesperidin and 6.9 ± 0.5 U/mg for quercitrin (expressed alternatively as 12,920 ± 595 μmol product min^{-1}/μmol enzyme for naringin, 1091 ± 62 μmol product min^{-1}/μmol enzyme for rutin, 768 ± 25 μmol product min^{-1}/μmol enzyme for hesperidin and 856 ± 62 μmol product min^{-1}/μmol enzyme for quercitrin).

More in detail, RHA-Phis specific activities for naringin, hesperidin and rutin were approximately between 5 and 14 times higher than the ones measured for the α-RHAs from *Streptomyces avermitilis*, *Clostridium stercorarium* and *Pediococcus acidilactici* [5,18,42]. Between 100 and 300 times higher specific activities towards naringin and rutin were obtained when compared to the α-RHA from *Fusobacterium* K-60 [43]. Noteworthy, this enzyme has been recognized as a good quercitrin-hydrolyzing rhamnosidase, and RHA-Phis shows an almost equivalent specific activity for quercitrin [43].

All the substrates tested in this work share the presence of a polyphenolic aromatic portion, together with a mono- or disaccharidic

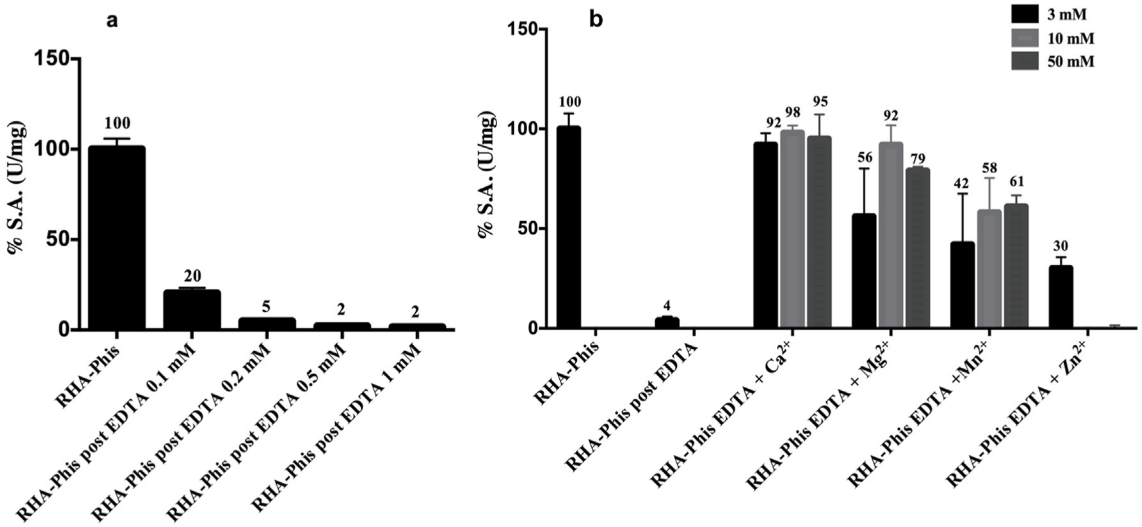


Fig. 4. RHA-Phis specific activity in the presence of increasing concentrations of EDTA, and after incubation with either Ca^{2+}, Mg^{2+}, Mn^{2+} or Zn^{2+} cations. Rhamnosidase specific activity is reported as percentage of residual activity compared to control. a) *p*NPR assays of RHA-Phis in the presence of different EDTA concentrations. b) *p*NPR assays of EDTA-treated RHA-Phis after incubation with different Ca^{2+}, Mg^{2+}, Mn^{2+} and Zn^{2+} concentrations.

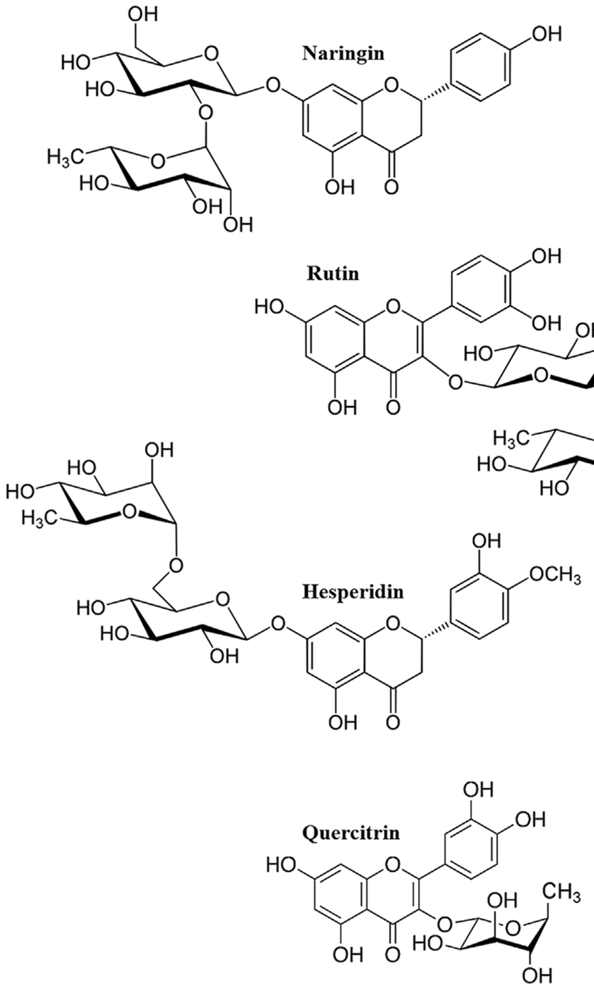


Fig. 5. Chemical structure of the flavonoidic compounds used as substrates in RHA-P catalyzed reactions.

moiety. To further explore RHA-P substrate specificity, we evaluated the activity of the enzyme on synthetic rhamnose oligosaccharides showing variable chain length and glycosidic linkages arrangement (Fig. S1) [44–46]. However, TLC analysis showed that RHA-Phis was unable to catalyze the hydrolysis of these substrates, not even extending the enzymatic assay incubation time to 24 h.

4. Discussion

RHA-P is a bacterial α-L-rhamnosidase recently isolated from the microorganism *Novosphingobium* sp. PP1Y [35]. This enzyme is an inverting glycosidase, belonging to the GH106 family, for which an initial biochemical characterization has highlighted an interesting biotechnological potential that needs to be better defined through a deeper functional and structural characterization of the protein.

Due to previous low purification yields related to the excessive number of purification steps used, a His-tag was added to the C-terminus of the recombinant protein expressed in the strain BL21(DE3) of *E.coli.* In our previous work we showed that "standard" conditions of recombinant expression led to the foremost presence of the untagged rRHA-P in the insoluble fraction of the induced cultures. The use of a lower induction temperature (23 °C instead of 37 °C), and a high-salt (0.5 M NaCl) LB medium, containing both betaine and sorbitol, allowed to markedly increase the yield of active rRHA-P in the soluble fraction of induced cultures [35].

RHA-Phis in standard expression conditions is also mostly expressed as inclusion bodies. Once having optimized the expression conditions, which involved again the use of a lower induction temperature and the addition of both betaine and sorbitol to the growth medium, we verified the presence of the recombinant protein in the periplasmic space of induced cells, as suggested by the identification of a signal peptide at the N-terminus of the protein. The presence of a similar peptide has also been described for the α-RHA isolated from *Sphingomonas paucimobilis* FP2001, which shares a 48% sequence identity with RHA-P [26]. Signal peptides have been described in α-RHAs isolated from fungal sources, such as the enzymes from *Aspergillus kawachii* and *Aspergillus aculeatus* [24,47]. In these latter, a slight difference was found in the sequence features compared to their bacterial counterparts; nonetheless, the efficient cleavage of a 17–20 aminoacids long N-terminal peptide was observed in the recombinant proteins. Although the functional reason for this specific sorting of an α-RHA is still not clear, this might be related to the engagement of this enzymatic activity in the complex molecular machinery involved in the biosynthesis and maintenance of the bacterial cell wall, where L-rhamnose appears to be a component of membrane rhamnolipids and polysaccharides [48,49].

Table 3
Kinetic constants of RHA-Phis on natural flavonoids. RHA-Phis kinetic constants on naringin, rutin, hesperidin and quercitrin are shown.

Flavonoid	Sugar	Glycosidic linkage position	K_M (mM)	k_{cat} (sec^{-1})	k_{cat}/K_M (sec^{-1} mM^{-1})
Naringin	neohesperidose (α,1–2)	A ring Carbon 7	0.59 (± 0.09)	214 (± 10)	360 (± 70)
Rutin	rutinose (α,1–6)	C ring Carbon 3	0.031 (± 0.01)	18.2 (± 1.2)	586 (± 172)
Hesperidin	rutinose (α,1–6)	A ring Carbon 7	0.067 (± 0.01)	12.7 (± 0.4)	190 (± 32)
Quercitrin	rhamnose	C ring Carbon 3	0.15 (± 0.03)	12.2 (± 0.4)	80 (± 16)

The presence of a putative N-terminal signal peptide in RHA-P had been previously suggested by the lack of an expected 22 aminoacid-long peptide, presumably cleaved through a post-translational proteolytic processing, in both the native and the recombinant protein [35].

The use of the periplasmic fraction as starting material and the presence of a unique purification step concurred in improving the purification yield to 25.3%, which is significantly higher than what obtained with the previous purification protocol (6.7%). In addition, the analysis of the kinetic constants evaluated on *p*NPR suggested that the His-tag did not significantly alter the α-RHA activity of RHA-Phis on this synthetic substrate.

To gather more insights into the catalytic mechanism of RHA-Phis for a future fine-tuning of this biotechnologically relevant enzyme, a combined approach of sequence alignment and homology modeling was used. The alignment of RHA-P sequence with 130 homologs of the GH106 family allowed identifying five highly conserved Glu and Asp residues. To confirm the role of these conserved residues in the active site, the homology model of RHA-Phis was built by using the three-dimensional structure α-RHA BT0986 from *Bacteroides thetaiotaomicron* as template. The modeled structure shows that, whereas residues D503, E506 and E644 are conserved also in the template protein BT0986, positions D552 and D763 do not seem to be directly involved in the "architecture" of the active site. To shed light on this aspect, an alanine scanning strategy on all five conserved residues initially retrieved by the sequence alignment was performed.

Disruptive mutations at positions D503, E506 and E644 led to an almost total inactivation of the enzyme, thus suggesting that these aminoacids either represent the catalytic residues in the active site or are directly involved in the binding of a calcium ion, a metal cofactor that has been proven to be essential for catalysis in some GH families.

Both whole recombinant cells and purified proteins enzymatic assays show that mutants at positions D552 and D763 have residual rhamnosidase activity (Fig. 4). The different values of specific activity measured on *p*NPR for each mutant in the two different assays probably account for different expression levels of the mutant proteins in the strain BL21(DE3) of *E. coli*. Of course, this aspect becomes irrelevant when purified proteins are instead assayed for their rhamnosidase activity. Although the whole cells assay at this stage is undoubtedly less informative to evaluate the differential catalytic role of the mutagenized residues, its feasibility shows that we have an invaluable tool to perform in the near future the preliminary screening of a large number of RHA mutants. In addition, from a biotechnological point of view, the low stability of purified enzymes often favors the use of whole cells in the set-up and development of bioconversion processes.

Purified D552 and D763 mutants exhibited a residual activity of 47.9% and 15.7%, respectively; likely, these residues are not directly involved in catalysis. This hypothesis is also supported by the kinetic constants measured on *p*NPR for these two mutants. Indeed, the decrease in catalysis efficiency, expressed by the k_{cat}/K_M, is almost exclusively dependent on the k_{cat} value and not on a difference in the apparent affinity of the enzyme for the substrate. These positions might not be directly involved in the binding or positioning of the substrate in the active site, as confirmed by the RHA-Phis model described in this work, which confirms that these two residues are located in quite distant positions from the active site.

The reduced specific activity observed in these mutants might thus be related to protein structural modifications leading to a soluble but partially inactive protein. More specifically, a possible functional role in the protein folding process can be foreseen for D763, for which a notable decrease in k_{cat} value is observed compared to *wt* RHA-Phis.

Residues D503, E506 and E644, which are conserved also in the template α-RHA BT0986, probably share a fundamental role in the active site of the enzyme. In particular, E506 is homologous to residue E461 in BT0986 (mutated in Q461 in the reference structure 5MWK) and represents the potential acid or base catalyst (highlighted in cyan in Fig. S3). On the other hand, residues D503 and E644 are homologous to residues D458 and E592 in BT0986, and seem to serve as calcium binding residues at the active site (these residues are highlighted in red in Fig. S3). The marked decrease of the activity observed when these residues are mutagenized supports the presence of an essential calcium ion at the active site of RHA-Phis as well, and a similar catalytic mechanism. It is worth noting that all four residues involved in the calcium binding in RHA-Phis model are conserved also in the template, thus indicating the importance of the calcium ion stabilization for the catalytic mechanism of the enzyme.

Although the presence of calcium in RHA-P active site is confirmed by the homology modeling, some additional observations can be made following the metal depletion and reconstitution experiments performed on RHA-Phis. EDTA-treated RHA-Phis recovers its activity in the presence of divalent cations such as Mg^{2+} and Ca^{2+}, but with an apparent higher efficiency when calcium is used, as suggested by the almost total recovery of the activity in the presence of a 3000-fold excess of this metal compared to the 56% obtained with a similar excess of Mg^{2+}. The results obtained with Mg^{2+} indirectly suggest that the two alkaline earth metals might interchange in RHA active site, but that the two metal cations probably have different binding parameters. It is important to underline, however, that a detailed description of the binding kinetic of these two elements was beyond the scope of our work. By contrast, incubation of EDTA-treated RHA-Phis with manganese or zinc led to significantly lower values of recovery of the enzymatic activity. For Mg^{2+} and Zn^{2+}, a relative decrease of the enzymatic activity was observed for the highest concentration used, *i.e.* 50 mM; more in detail, when using zinc a marked protein precipitation was observed in the sample, which may be responsible for the almost complete loss of enzymatic activity.

As already underlined, the presence of calcium ions as essential cofactors is reported in some GH families, although with different functions [50]. Among the few α-RHAs crystal structures available in literature, two examples are provided bearing calcium ions involved in catalysis, besides GH106 BT0986: GH78 BsRhaB from *Bacillus* sp. GL1 and SaRha78A from *Streptomyces avermitilis* [29,30]. However, differently from what observed in BT0986, in these proteins calcium is bound in a separate domain located near the catalytic domain and contacts the rhamnose residue in the catalytic domain when the substrate is bound in the active site [30]. In particular, Fujimoto Z. and coworkers described for SaRha78A the identification of a novel carbohydrate-binding module (CBM67) highly specific for rhamnose residues [30]. Carbohydrate binding modules (CBM), are defined as non-catalytic domains found in many carbohydrate-active enzymes and contribute to the enzyme binding specificity towards different carbohydrates and polysaccharides [50]. In RHA-Phis, and in general GH106, no separate CBM domains have been observed, and calcium is probably directly bound in the active site of the enzyme. In BT0986, as well as in RHA-Phis, calcium seems rather to play a key catalytic role, similar to the one reported from Zhu Y. and coworkers for the GH92 and GH47 α-mannosidase enzymes [51], where calcium is hypothesized to be essential for the inverting mechanism of α-mannose residues hydrolysis [51]. The role of the calcium ion in the active site of RHA-Phis undoubtedly needs further investigation; nonetheless, the data collected so far sustain its involvement as an essential cofactor at the active site of the enzyme.

The model suggests that RHA-Phis and protein BT0986 have similar catalytic mechanisms, and all the residues involved in the calcium binding seem to be conserved in the two proteins. In addition, aminoacids contributing to rhamnose binding in the structure of BT0986 are also conserved in RHA-Phis (residues of interest are highlighted in green in Fig. S3 of the Supplementary Material). Noteworthy, the RHA-Phis model points out also some significant differences with the BT0986 template, which allows predicting a different substrate specificity. In particular, RHA-Phis lacks a loop region that in BT0986 is involved in the arabinose binding sub-site; this loop might be the key factor responsible for the accommodation of different substrates in the active site of these enzymes.

To investigate RHA-Phis substrate specificity on natural substrates, kinetic experiments were performed using flavonoids such as naringin, rutin, hesperidin and quercitrin. Results confirm that RHA-Phis is able to hydrolyze both α-1,2 and α-1,6 glycosidic linkages [35]. However, the analysis of the k_{cat}/K_M values highlights some interesting differences in terms of substrate affinity and catalytic efficiency. In particular, rutin and quercitrin share the same flavonoidic aglyconic portion (named quercetin) and both have the saccharidic portion bound at position C3 on the C ring of the flavonoid. However, RHA-Phis shows a higher catalytic efficiency on rutin. On the other hand, rutin and hesperidin share the same rutinose α-1,6 disaccharidic residue, but bound at different positions of the flavonoidic ring (in position C3 on the C ring for rutin and at position C7 on the A ring for hesperidin). The higher catalytic efficiency of RHA-Phis towards rutin, might be related to the specific position of the flavonoidic ring to which the saccharidic portion is bound.

Finally, naringin and hesperidin have different glycosidic moieties (neohesperidose α-1,2 and rutinose α-1,6 respectively), bound in the same C7 position on the A ring of the flavonoid. The higher value of k_{cat}/K_M of RHA-Phis towards the former substrate suggests that the enzyme shows a preferential hydrolysis of the neohesperidose α-1,2 unit.

Although a slight difference in the substituents of the flavonoidic rings involved should be taken into account in the discussion of these results, several conclusions can be drawn from these experiments. Specifically, the data obtained suggest that *i*) a rhamnose-containing disaccharidic unit is better accommodated in the active site compared to rhamnose alone, *ii*) when bound to the C7 of the A ring of the flavonoid, the neohesperidose moiety is hydrolyzed more efficiently than rutinose, and *iii*) when rutinose is present as the disaccharidic moiety, the preferential position for its hydrolysis is the C3 of the flavonoid C ring.

All the hydrolyzed flavonoids used in this work share the presence of a polyphenolic aromatic portion that can be accommodated in the active site of RHA-Phis. Conversely, the rhamnose oligosaccharides tested (a penta-, hexa- and octasaccharide, Fig. S1 of the Supplementary Material), which lack the polyphenolic portion, were not hydrolyzed by RHA-Phis. It should be noted, however, that all these oligosaccharides are characterized by the same α-1,3 glycosidic linkage at the non-reducing terminal of the chain. The lack of hydrolysis could be due either to the inability of the enzyme to hydrolyze α-1,3 bonds, or to the impossibility to efficiently accommodate these oligosaccharidic chains in the active site.

5. Conclusions

In this work we have gained a substantial amount of structural and functional information on RHA-P, a novel recombinant rhamnosidase of the GH106 family recently isolated from the marine bacterium *N.* sp. PP1Y. The GH106 is a still poorly characterized family of glycosyl hydrolases and, to the best of our knowledge, only the crystal structure of protein BT0986 from *Bacteroides thetaiotaomicron,* has been solved so far. Therefore, in the absence of RHA-P crystal structure, the insights obtained from the combined approach involving a multiple sequences alignment and the homology modeling described in this work allowed to identify critical residues in the active site of RHA-P, which is of fundamental importance for the future fine-tuning of this enzymatic activity for biotechnological applications. Our results have highlighted that the biotechnological use of RHA-P, which has shown a remarkable catalytic efficiency on several flavonoids, encompass the use of either the purified protein or the whole recombinant cells expressing the His-tagged protein. Lastly, the optimization of the recombinant expression and purification protocol will undoubtedly pave the way in the near future to crystallization experiments that are required to shed light on the catalytic and structural features of this unique family of glycosyl hydrolases.

Author contributions

FM and VI conceived the idea and experiments. FM and FDL carried out the experiments. FM wrote the manuscript with support from VI, EN, ADD, MM, MHS. MM, VC, VI and ADD verified the analytical methods. AS, VC contributed in experimental procedures. EB provided substrates and chemical reagents and EN performed the homology modeling analysis. MM, MT, EN, ADD contributed to the interpretation of the results. All authors provided critical feedback and helped shape the research, analysis and manuscript.

Acknowledgements

This research did not receive any specific grant from funding agencies in the public, commercial or not-for-profit sectors.

We are grateful to Prof. Harry J. Gilbert of the Institute for Cell and Molecular Biosciences, Newcastle University (Newcastle upon Tyne, UK), for the useful discussion and the support in the data analysis of RHA-P homology modeling.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.abb.2018.04.013.

References

[1] B. Henrissat, A classification of glycosyl hydrolases based on amino-acid sequence similarities, Biochem. J. 280 (1991) 309–316.

[2] P. Manzanares, S. Vallés, D. Ramón, M. Orejas, Industrial enzymes, in: J. Polaina, A.P. MacCabe (Eds.), α-L-Rhamnosidases: Old and New Insights, Industrial Enzymes, Structure, Function and Applications, Springer, 2007, pp. 117–140.

[3] S. Matsumotoa, H. Yamadab, Y. Kunishigea, S. Takenakac, M. Nakazawaa, M. Uedaa, T. Sakamotoa, Identification of a novel *Penicillium chrysogenum* rhamnogalacturonan rhamnohydrolase and the first report of a rhamnogalacturonan rhamnohydrolase gene, Enzym. Microb. Technol. 98 (2017) 76–85.

[4] W. Hashimoto, O. Miyake, H. Nankai, K. Murata, Molecular identification of an α-L-rhamnosidase from *Bacillus* sp. strain GL1 as an enzyme involved in complete metabolism of gellan, Arch. Biochem. Biophys. 415 (2) (2003) 235–244.

[5] H. Ichinose, Z. Fujimoto, S. Kaneko, Characterization of an α-l-rhamnosidase from Streptomyces avermitilis, Biosci. Biotechnol. Biochem. 77 (1) (2013) 213–216.

[6] A.N. Panche, A.D. Diwan, S.R. Chandra, Flavonoids: an overview, J. Nutr. Sci. 5 (2016) e47.

[7] O. Benavente-García, J. Castillo, Update on uses and properties of citrus flavonoids: new findings in anticancer, cardiovascular, and anti-inflammatory activity, J. Agric. Food Chem. 56 (15) (2008) 6185–6205.

[8] K. Németh, G.W. Plumb, B.G. Berrin, N. Juge, R.J. Hassan, H.Y. Naim, G. Williamson, D.M. Swallow, P.A. Kroon, Deglycosylation by small intestinal epithelial cell β-glucosidases is a critical step in the absorption and metabolism of dietary flavonoid glycosides in humans, Eur. J. Nutr. 42 (2003) 29–42.

[9] V.D. Bokkenheuser, C.H. Shackleton, J. Winter, Hydrolysis of dietary flavonoid glycosides by strains of intestinal Bacteroides from humans, Biochem. J. 248 (3) (1987) 953–956.

[10] P.C. Hollman, M.N. Bijsman, Y. van Gameren, E.P. Cnossen, J.H. de Vries, M.B. Katan, The sugar moiety is a major determinant of the absorption of dietary flavonoid glycosides in man, Free Radic. Res. 31 (6) (1999) 569–573.

[11] A.J. Day, J.M. Gee, M.S. DuPont, I.T. Johnson, G. Williamson, Absorption of quercetin-3-glucoside and quercetin-4′-glucoside in the rat small intestine: the role of lactase phlorizin hydrolase and the sodium-dependent glucose transporter, Biochem. Pharmacol. 65 (7) (2003) 1199–1206.

[12] R. González-Barrio, L.M. Trindade, P. Manzanares, L.H. de Graaff, F.A. Tomás-Barberán, J.C. Espín, Production of bioavailable flavonoid glucosides in fruit juices and green tea by use of fungal alpha-L-rhamnosidases, J. Agric. Food Chem. 52 (20) (2004) 6136–6142.

[13] M. Puri, S.S. Marwaha, R.M. Kothari, J.F. Kennedy, Biochemical basis of bitterness in citrus fruit juices and biotech approaches for debittering, Crit. Rev. Biotechnol. 16 (1996) 145–155.

[14] Y. Terada, T. Kometani, T. Nishimura, H. Takii, S. Okada, Prevention of hesperidin crystal formation in canned Mandarin orange syrup and clarified orange juice by hesperidin glycosides, Food Sci. Technol. Int. 1 (1) (1995) 29–33.

[15] P. Manzanares, M. Orejas, V. Gil, L.H. de Graaff, J. Visser, D. Ramón, Construction of a genetically modified wine yeast strain expressing the *Aspergillus aculeatus* rhaA gene, encoding an α-l-rhamnosidase of enological interest, Appl. Environ. Microbiol. 69 (12) (2003) 7558–7562.

[16] L. Xu, X. Liu, Z. Yin, Q. Liu, L. Lu, M. Xiao, Site-directed mutagenesis of α-L-rhamnosidase from *Alternaria* sp. L1 to enhance synthesis yield of reverse hydrolysis based on rational design, Appl. Microbiol. Biotechnol. 100 (24) (2016) 10385–10394.

[17] W. Hashimoto, H. Nankai, N. Sato, S. Kawai, K. Murata, Characterization of α-L-rhamnosidase of *Bacillus* sp. GL1 responsible for the complete depolymerization of gellan, Arch. Biochem. Biophys. 415 (2) (1999) 56–60.

[18] V.V. Zverev, C. Hertel, K. Bronnenmeier, A. Hroch, J. Kellermann, W.H. Schwarz, The thermostable α-L-rhamnosidase RamA of *Clostridium stercorarium*: biochemical characterization and primary structure of a bacterial α-L-rhamnoside hydrolase, a new type of inverting glycoside hydrolase, Mol. Microbiol. 35 (2000) 173–179.

[19] M. Avila, M. Jaquet, D. Moine, T. Requena, C. Peláez, F. Arigoni, I. Jankovic, Physiological and biochemical characterization of the two alpha-L-rhamnosidases of *Lactobacillus plantarum* NCC245, Microbiology 155 (2009) 2739–2749.

[20] J. Beekwilder, D. Marcozzi, S. Vecchi, R. de Vos, P. Janssen, C. Francke, J. van Hylckama Vlieg, R.D. Hall, Characterization of rhamnosidases from *Lactobacillus plantarum* and Lactobacillus acidophilus, Appl. Environ. Microbiol. 75 (2009) 3447–3454.

[21] F. Miake, T. Satho, H. Takesue, F. Yanagida, N. Kashige, K. Watanabe, Purification and characterization of intracellular alpha-L-rhamnosidase from *Pseudomonas paucimobilis* FP2001, Arch. Microbiol. 173 (1) (2000) 65–70.

[22] L. Li, Y. Yu, X. Zhang, Z. Jiang, Y. Zhu, A. Xiao, H. Ni, F. Chen, Expression and biochemical characterization of recombinant α-L-rhamnosidase r-Rha1 from *Aspergillus Niger* JMU-TS528, Int. J. Biol. Macromol. 85 (2016) 391–399.

[23] N.M. Young, R.A.Z. Johnston, J.C. Richards, Purification of the α-L-rhamnosidase of *Penicillium decumbens* and characteristics of two glycopeptide components, Carbohydr. Res. 191 (1989) 53–62.

[24] T. Koseki, Y. Mese, N. Nishibori, K. Masaki, T. Fujii, T. Handa, Y. Yamane, Y. Shiono, T. Murayama, H. Iefuji, Characterization of an alpha-L-rhamnosidase from *Aspergillus kawachii* and its gene, Appl. Microbiol. Biotechnol. 80 (6) (2008) 1007–1013.

[25] V. Lombard, H. Golaconda Ramulu, E. Drula, P.M. Coutinho, B. Henrissat, The carbohydrate-active enzymes database (CAZy) in 2013, Nucleic Acids Res. 42 (2014) D490–D495.

[26] T. Miyata, N. Kashige, T. Satho, T. Yamaguchi, Y. Aso, F. Miake, Cloning, sequence analysis, and expression of the gene encoding *Sphingomonas paucimobilis* FP2001 alpha-L-rhamnosidase, Curr. Microbiol. 51 (2005) 105–109.

[27] D. Ndeh, A. Rogowski, A. Cartmell, A.S. Luis, A. Baslé, J. Gray, I. Venditto, J. Briggs, X. Zhang, X. Labourel, N. Terrapon, F. Buffetto, S. Nepogodiev, Y. Xiao, R.A. Field, Y. Zhu, M.A. O'Neill, B.R. Urbanowicz, W.S. York, G.J. Davies, D.W. Abbott, M.C. Ralet, E.C. Martens, B. Henrissat, H.J. Gilbert, Complex pectin metabolism by gut bacteria reveals novel catalytic functions, Nature 544 (7648) (2017) 65–70.

[28] J.B. Bonanno, S.C. Almo, A. Bresnick, M.R. Chance, A. Fiser, S. Swaminathan, J. Jiang, F.W. Studier, L. Shapiro, C.D. Lima, T.M. Gaasterland, A. Sali, K. Bain, I. Feil, X. Gao, D. Lorimer, A. Ramos, J.M. Sauder, S.R. Wasserman, S. Emtage, K.L. D'Amico, S.K. Burley, New York-Structural GenomiX Research Consortium (NYSGXRC): a large scale center for the protein structure initiative, J. Struct. Funct. Genom. 6 (2–3) (2005) 225–232.

[29] Z. Cui, Y. Maruyama, B. Mikami, W. Hashimoto, K. Murata, Crystal structure of glycoside hydrolase family 78 α-L-rhamnosidase from *Bacillus* sp. GL1, J. Mol. Biol. 374 (2) (2007) 384–398.

[30] Z. Fujimoto, A. Jackson, M. Michikawa, T. Maehara, M. Momma, B. Henrissat, H.J. Gilbert, S. Kaneko, The structure of a *Streptomyces avermitilis* α-L-rhamnosidase reveals a novel carbohydrate-binding module CBM67 within the six-domain arrangement, J. Biol. Chem. 288 (17) (2013) 12376–12385.

[31] E.C. O'Neill, C.E. Stevenson, M.J. Paterson, M. Rejzek, A.L. Chauvin, D.M. Lawson, R.A. Field, Crystal structure of a novel two domain GH78 family α-rhamnosidase from *Klebsiella oxytoca* with rhamnose bound, Proteins 83 (9) (2015) 1742–1749.

[32] E. Notomista, F. Pennacchio, V. Cafaro, G. Smaldone, V. Izzo, L. Troncone, M. Varcamonti, A. Di Donato, The marine isolate *Novosphingobium* sp. PP1Y shows specific adaptation to use the aromatic fraction of fuels as the sole carbon and energy source, Microb. Ecol. 61 (3) (2011) 582–594.

[33] V. D'Argenio, M. Petrillo, P. Cantiello, B. Naso, L. Cozzuto, E. Notomista, G. Paolella, A. Di Donato, F. Salvatore, De Novo sequencing and assembly of the whole genome of *Novosphingobium* sp. Strain PP1Y, J. Bacteriol. 193 (16) (2011) 4296.

[34] V. D'Argenio, E. Notomista, M. Petrillo, P. Cantiello, V. Cafaro, V. Izzo, B. Naso, L. Cozzuto, L. Durante, L. Troncone, G. Paolella, F. Salvatore, A. Di Donato, Complete sequencing of *Novosphingobium* sp. PP1Y reveals a biotechnologically meaningful metabolic pattern, BMC Genom. 15 (1) (2014) 384.

[35] F. De Lise, F. Mensitieri, V. Tarallo, N. Ventimiglia, R. Vinciguerra, A. Tramice, R. Marchetti, E. Pizzo, E. Notomista, V. Cafaro, A. Molinaro, L. Birolo, A. Di Donato, V. Izzo, RHA-P: isolation, expression and characterization of a bacterial α-L-rhamnosidase from *Novosphingobium* sp. PP1Y, J. Mol. Catal. B Enzym. 134 (A) (2016) 136–147.

[36] J. Sambrook, E.F. Fritsch, T. Maniatis, Molecular Cloning. A Laboratory Manual, second ed., Harbor Laboratory Press, Cold Spring Harbor, NY, 1982, p. 545.

[37] P. Gerhardt, R.G.E. Murray, W.A. Wood, N.R. Krieg, Methods for General and Molecular Bacteriology, ASM Press Washington, D.C, 1994.

[38] K. Arnold, L. Bordoli, J. Kopp, T. Schwede, The SWISS-MODEL workspace: a web-based environment for protein structure homology modeling, Bioinformatics 22 (2) (2006) 195–201.

[39] N. Guex, M.C. Peitsch, T. Schwede, Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: a historical perspective, Electrophoresis 30 (1) (2009) S162–S173.

[40] X. Huang, W. Miller, A time-efficient, linear-space local similarity algorithm, Adv. Appl. Math. 12 (3) (1991) 337–357.

[41] F. He, Laemmli-sds-page, Bio-protocol 101 (2011) e80 http://www.bio-protocol.org/e80.

[42] H. Michlmayr, W. Brandes, R. Eder, C. Schümann, A.M. del Hierro, K.D. Kulbe, Characterization of two distinct glycosyl hydrolase family 78 α-l-Rhamnosidase from Pediococcus acidilactici, Appl. Environ. Microbiol. 77 (18) (2011) 6524–6530.

[43] S.Y. Park, J.H. Kim, D.H. Kim, Purification and characterization of quercitrin-hydrolyzing α-l-rhamnosidase from *Fusobacterium* K-60, a human intestinal bacterium, J. Microbiol. Biotechnol. 15 (3) (2005) 519–524.

[44] E. Bedini, M. Parrilli, C. Unverzagt, Oligomerization of a rhamnanic trisaccharide repeating unit of O-chain polysaccharides from phytopathogenic bacteria, Tetrahedron Lett. 43 (2002) 8879–8882.

[45] E. Bedini, G. Barone, C. Unverzagt, M. Parrilli, Synthesis of the pentasaccharide repeating unit of the major O-antigen component from *Pseudomonas syringae* pv. *ribicola* NVPPB 1010, Carbohydr. Res. 339 (2004) 393–400.

[46] E. Bedini, A. Carabellese, D. Comegna, C. De Castro, M. Parrilli, *Tetrahedron* Synthetic Oligorhamnans Related to the Most Common O-chain Backbone from Phytopathogenic Bacteria vol. 62, (2006), pp. 8474–8483.

[47] P. Manzanares, H.C. van den Broeck, L.H. de Graaff, J. Visser, Purification and characterization of two different α-l-rhamnosidases, RhaA and RhaB, from Aspergillus aculeatus, Appl. Environ. Microbiol. 67 (5) (2001) 2230–2234.

[48] W. Nereus, I.V. Gunther, A. Nuñez, W. Fett, D.K.Y. Solaiman, Production of rhamnolipids by *Pseudomonas chlororaphis*, a Nonpathogenic bacterium, Appl. Environ. Microbiol. 71 (5) (2005) 2288–2293.

[49] E. Champion, I. Andre, L.A. Mulard, P. Monsan, M. Remaud-Simeon, S. Morel, Synthesis of l-rhamnose and N-Acetyl-D-Glucosamine derivatives entering in the composition of bacterial polysaccharides by use of glucansucrases, J. Carbohydr. Chem. 28 (2009) 142–160.

[50] A.B. Boraston, D.N. Bolam, H.J. Gilbert, G.J. Davies, Carbohydrate-binding modules: fine-tuning polysaccharide recognition, Biochem. J. 382 (2004) 769–781.

[51] Y. Zhu, M.D.L. Suits, A.J. Thompson, S. Chavan, Z. Dinev, C. Dumon, N. Smith, K.W. Moremen, Y. Xiang, A. Siriwardena, S.J. Williams, H.J. Gilbert, Mechanistic insights into a Ca^{2+}-dependent family of α-mannosidases in a human gut symbiont, Nat. Chem. Biol. 6 (2) (2010) 125–132.