



Master Thesis

Production, Isolation, and Characterisation of Carotenoids from Engineered *Rhodothermus marinus*

Nina Diederich

KBTM01

Master's Thesis (30 credits ECTS)

May 2025

Division of Biotechnology and Applied Microbiology

Department of Process and Life Science Engineering

LTH Faculty of Engineering

Lund University

Supervisor: Roya R. Sardari, PhD

Examiner: Prof. Patrick Adlercreutz

Acknowledgements

I want to thank my supervisor Roya for her constant support, for always helping me out and answering every question without judgement, no matter what time of the day. I also want to thank my examiner Patrick Adlercreutz for his valuable guidance throughout the thesis process. A heartfelt thank you to my partner, Lauri, for always being there to support and motivate me. Thank you for listening to all my complaints, big and small, and for always reminding me that things aren't so bad after all. Thank you to everyone in the division who shared a friendly chat, a word of advice or helped me, and for all the great cakes that were consumed on Friday afternoons in the Fika Club. Last but definitely not least, I have to thank my friends, Anna, Emilie, Laura, Lovro, and Renée, that made long days in the lab much more bearable and that I made the most special memories with during those two years of this masters program. Lots of love to you guys!

Popular science summary

Microorganisms that thrive in extreme environments, like hot springs, are called extremophiles. The extremophile, *Rhodothermus marinus*, is known to survive high temperatures and harsh conditions, making it a promising candidate for industrial use. This thesis explores how genetically modified strains can be used to produce carotenoids —natural pigments with potential health and industrial benefits.

The experimental work involved cultivating an engineered strain of *R. marinus*, extracting the carotenoids, and analysing them using advanced laboratory techniques. Additionally, the study explored the enzymatic modification of said carotenoids, aiming to enhance their biosurfactant properties. These modified carotenoids could be used as natural ingredients in products like food, cosmetics, or environmentally friendly cleaning agents. This research helps pave the way for more sustainable production of valuable natural compounds using extreme, adaptable bacteria.

Abstract

Extremophilic bacteria, like *Rhodothermus marinus*, hold promising potential to be used as non-model cell factories for the production of various metabolites, such as carotenoids or exopolysaccharides, due to their inherent adaptability to extreme environmental conditions. This study explores the production, isolation, and characterisation of carotenoids from an engineered *R. marinus* strain termed 'TK21'. The anticipated compound produced by said strain is glycosylated lycopene, which is thought to exhibit biosurfactant properties. Liquid Chromatography Mass Spectrometry (LC-MS) and High-Performance Liquid Chromatography (HPLC) were used to verify the presence of these metabolites and assess structural variations. Additionally, enzymatic transglycosylation reactions using cyclodextrin glucanotransferase (CGtase) were conducted using the model acceptor substrates piceid and crocin. Results indicate that glycosylated lycopene derivatives were produced, though ionisation challenges in MS detection required methodological optimisation. Enzymatic glycosylation reactions successfully modified piceid and crocin derivatives, suggesting the potential for biosurfactant development.

Abbreviations

bp - basepair

CGtase - cyclodextrin glucanotransferase

collision induced dissociation - CID

DI-water - distilled water

EPS - Exopolysccharide

LCMS - Liquid Chromatography Mass Spectrometry

mAU - milli-absorbance units

nm - nanometer

ODC₆₀₀ - optical density at wavelength 600 nm

rpm - rounds per minute

TIC - total ion count

Contents

Acknowledgements	i
Popular science summary	ii
Abstract	iii
Abbreviations	iv
1 Introduction	1
2 Theoretical Background	2
2.1 <i>Rhodothermus marinus</i>	2
2.1.1 Carotenoids from <i>Rhodothermus marinus</i>	2
2.2 Carotenoid biosynthetic pathway elucidation	3
2.3 Creating mutants variants from native <i>Rhodothermus marinus</i> PRI 493	5
2.4 Transglycosylation reaction using CGtase	6
2.5 Aim	8
3 Methods and materials	9
3.1 Chemicals and reagents	9
3.2 Strains	9
3.3 Cultivation of <i>Rhodothermus marinus</i> TK21	9
3.3.1 Media preparation	9
3.3.2 Inoculum	11
3.3.3 Fermentation process	11
3.4 Extraction of carotenoids	12
3.5 Analytical methods	12
3.5.1 HPLC	12
3.5.2 LC-MS	13
3.5.3 Method optimisation	14
3.6 Model enzymatic reaction using piceid and crocin	14

4	Results and discussion	15
4.1	Cultivation and extraction of carotenoids from <i>Rhodothermus marinus</i> TK21	15
4.2	Characterisation of carotenoid extracts from <i>R. marinus</i> TK21	16
4.2.1	Optimisation of MS method	22
4.3	Transglycosylation model reactions using acceptor substrates piceid and crocin . . .	28
4.3.1	Enzymatic synthesis of crocin glycosides	28
4.3.2	Enzymatic synthesis of piceid glycosides	31
5	Conclusion	34
6	Outlook	35
	References	36
A	Appendix	39
A.1	Characterisation of carotenoid extracts from <i>R. marinus</i> TK21	39
A.2	Transglycosylation model reactions using acceptor substrates piceid and crocin . . .	39
A.2.1	Crocin reaction	39
A.2.2	Piceid reaction	39

1 Introduction

Surface-active molecules produced by microorganisms, termed biosurfactants, are gaining growing interest as environmentally friendly alternatives to petrochemically produced surfactants. They play a key role in the growing bioeconomy, providing sustainable and biodegradable solutions that address the pressing challenges of climate change and the need for more sustainable industrial and life science applications.

In addition to environmental advantages, biosurfactants exhibit many desirable properties when compared to petrochemically derived surfactants, like low toxicity, structural diversity, and biodegradability, inevitably also posing a lower risk to human health (Gayathiri et al. (2022), Romero Vega and Gallo Stampino (2025)). Due to these favorable characteristics, biosurfactants are gaining interest in applications in the cosmetics and food industry and waste treatment, where they can be used as emulsifiers, detergents, or preserving agents (Gayathiri et al. (2022)). Furthermore, their potential is being explored in bioremediation, and even soil decontamination, where they increase the surface area of hydrophobic pollutants, such as oil or heavy metals (Ron and Rosenberg (2002), Karmakar et al. (2025)). Although the production costs of biosurfactants currently remain higher than those of conventional surfactants, their market share is growing continuously, with an increase from USD 19 billion in 2023 to an estimated USD 26 billion by 2032 (Marquez et al. (2025)). This reflects the growing interest in sustainable and environmentally friendly alternatives within the surfactant industry.

Of the great variety of microorganisms that produce surfactants, extremophiles constitute a large untapped potential, as they naturally have a high tolerance to extreme conditions, among others temperature, salinity, pH, and pressure, which often are a part of industrial applications. They are because of that excellent candidates for being subjected to genetic modifications in order to use those microorganisms as cell factories for the production of surfactants, but also chemicals, enzymes, and other compounds (Atif et al., 2024).

2 Theoretical Background

2.1 *Rhodothermus marinus*

Rhodothermus marinus is a Gram-negative, aerobic bacterium belonging to the family of Rhodothermaceae. It exhibits thermohalophilic characteristics, meaning it thrives in high-temperature and saline habitats, and was first isolated from marine hot springs in the Isafjardardjup Bay in NW Iceland. Optimal growth of the bacterium has been found to be at 65 °C, pH 7, and 2% (w/v) NaCl (Bjornsdottir et al., 2006), (?).

Like other extremophiles, *R. marinus* has attracted particular interest for a number of enzymes and metabolites that have been found to be produced by the strain that are suitable for industrial applications due to their compatibility with extreme conditions. Because of its heterotrophic nature, *R. marinus* produces thermostable enzymes, in particular glycoside hydrolases, which are capable of degrading carbohydrates found naturally in its native environment for energy production (Ron, 2020),(Bjornsdottir et al., 2006), (Kristjánssdóttir et al., 2022). Additionally, *R. marinus* is a natural producer of exopolysaccharides (EPS). During growth in carbohydrate-rich medium, *R. marinus* cells have been found to form a distinct exopolysaccharide capsule, which is presumably produced by the cell as a protectant and more so in the presence of a high amount of carbon source (Sardari et al., 2017). The third metabolite of interest produced by *R. marinus* are carotenoids, which give the strain their characteristic red colouring.

2.1.1 Carotenoids from *Rhodothermus marinus*

Most research revolving around *R. marinus* has been focused on the DSM 4252 type-strain, of which the complete genome sequence is known (Nolan et al., 2009). Extensive research into the carotenoid composition of this strain and of the strains DSM 4253 and strain PRI 493 (also known as DSM 16675) found that salinixanthin is the most common native carotenoid produced by these strains, which has also previously been found in its closest genetic relative, *Salinibacter ruber* (Ron et al. (2018), Lutnaes et al. (2004)). Salinixanthin is a C₄₀ carotenoid acyl glycoside, which belongs to the xanthophyll class of carotenoids (Figure 1). It has the characteristic C₄₀ polyene skeleton with alternating single and double bonds, contributing to its conjugated system responsible for light absorption. It also possesses a glycosyl end group that is acylated in the C-6 position of the glucose molecule (Britton (1995), Lutnaes et al. (2002)).

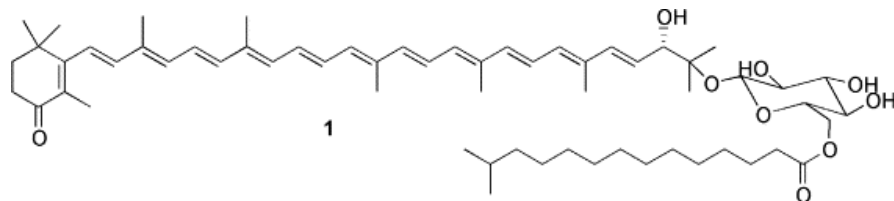


Figure 1: Salinixanthin structure found in *Rhodothermus marinus*. (Source Lutnaes et al. (2004))

The carotenoid acyl glucoside produced by *Rhodothermus marinus* is presumed to be produced by the microorganism to stabilise its membrane and enhance its adaptability to high temperatures (Lutnaes et al. (2002), Lutnaes et al. (2004)). Salinixanthin, the predominant carotenoid in *R. marinus* was found in both hydroxylated and non-hydroxylated forms across all three strains mentioned above. Additionally, a novel carotenoid glycoside ester was identified, exhibiting a structural similarity to salinixanthin but lacking the keto group on the β -ionone ring. In both salinixanthin and the novel carotenoid glycoside ester, variations in fatty acid chain lengths (ranging from C_{11} - C_{17}) were observed in the acyl group (Figure 2) (Ron et al., 2018). Bjornsdottir et al. (Bjornsdottir et al., 2006) reported that the proportion of the difference in chain lengths is dependent on the medium composition, suggesting an environmental impact on carotenoid biosynthesis.

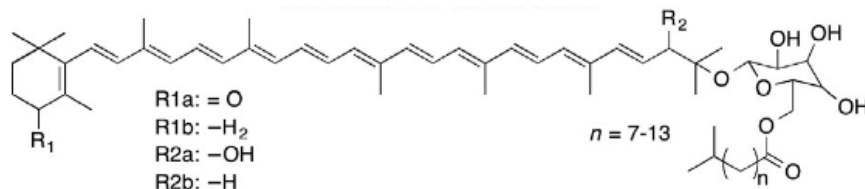


Figure 2: Proposed structure of novel carotenoid found in *Rhodothermus marinus* DSM 4252, DSM 4253 and PRI 493. (Source Ron et al. (2018))

2.2 Carotenoid biosynthetic pathway elucidation

The biosynthetic pathway in Figure 3 was suggested by Ron (2020) after gained insights from the work of Ron et al. (2018) and Kristjansdottir et al. (2020). The pathway begins with the transformation of two geranylgeranyl-pyrophosphates into 1'-hydroxy phytoene, catalysed by phytoene synthase (CrtB) and 1',2'-hydratase (CruF). 1'-hydroxy phytoene is then converted to 1'-hydroxy lycopene by phytoene desaturase (CrtI). From lycopene, salinixanthin is formed, catalysed by glycosyltrans-

ferase (CruC), acyltransferase (CruD), lycopene β -cyclase (CrtL-b/CrtY), and β -carotene 4-ketolase (Kristjansdottir et al. (2020), Ron (2020)). The elucidation of the biosynthetic carotenoid pathway was the first stepping stone towards effectively harnessing *R. marinus* as a microbial cell factory for carotenoid production. The detailed understanding of its biosynthetic pathway and the genetic engineering tools discovered by Bjornsdottir et al. (2007) and Bjornsdottir et al. (2011) were essential to precisely modify and optimise the organism's ability to synthesise and accumulate carotenoids for industrial applications.

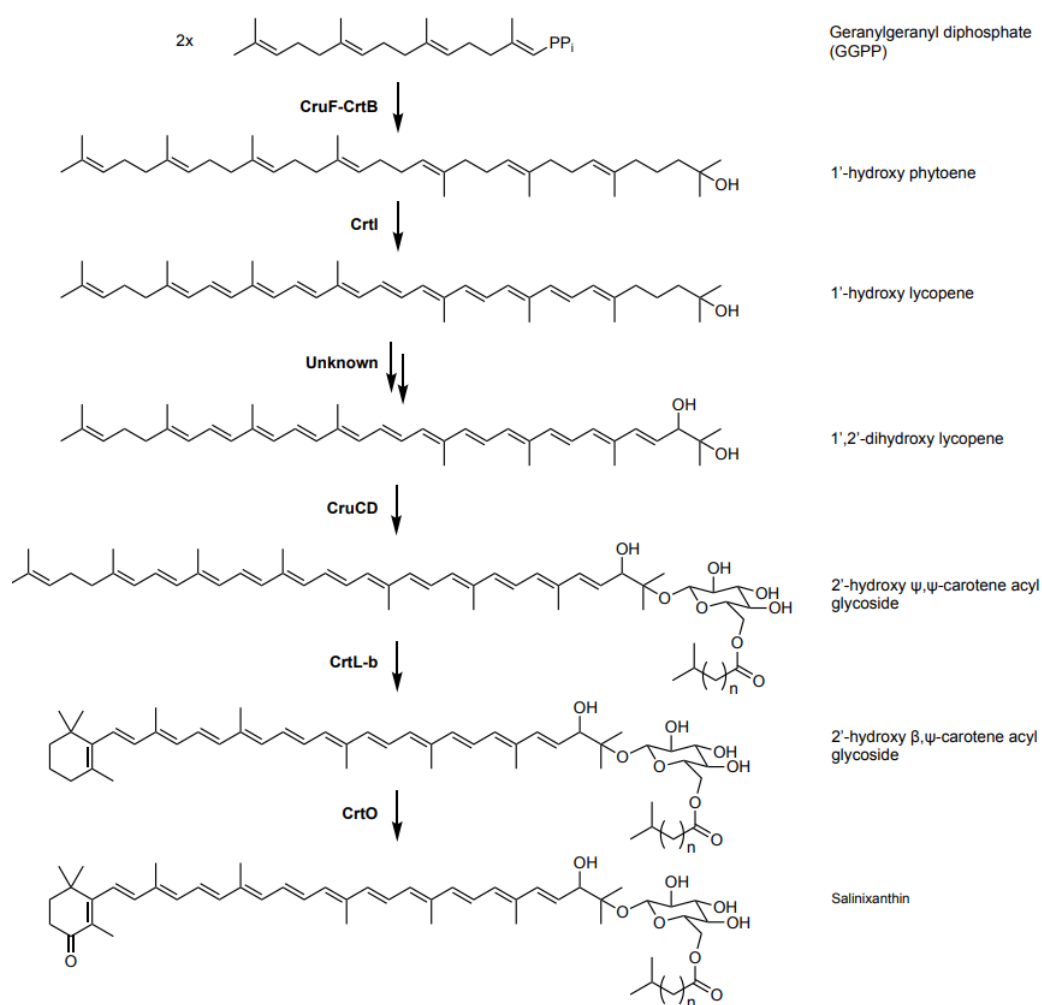


Figure 3: Proposed biosynthetic carotenoid pathway in *Rhodothermus marinus*. (Source Ron (2020))

2.3 Creating mutants variants from native *Rhodothermus marinus* PRI 493

In the pursuit of creating *R. marinus* mutants capable of lycopene production, the strain SB-62 was developed by Bjornsdottir et al. (2011), which originates from the wild type strain PRI 493. The strain was engineered through targeted deletions of the genes *trpB* and *purA*, providing optimal conditions for genetic modifications since those deletions could be used as selection markers (Bjornsdottir et al., 2011). From this strain, Kristjansdottir et al. (2020) engineered the *R. marinus* mutant TK-1, which produced modified lycopene (Figure 4) following the targeted deletion of a 5890 bp region that harboured the *CrtY* and *CrtO* genes. Due to the deletion of those two genes, the resulting compound did not have the β -ionone ring that was part of the compound found in the novel carotenoid found in the *R. marinus* wild types. The resulting compound was successfully extracted, purified, and characterised, prompting further efforts to develop more mutant strains.

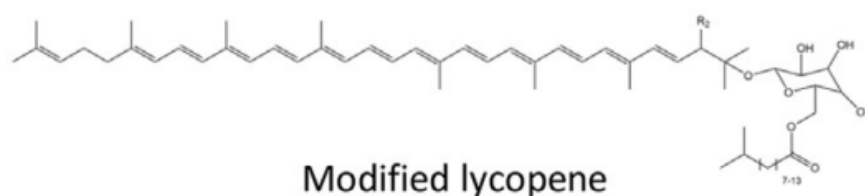


Figure 4: Compound produced by *R. marinus* mutant strain TK-1. Source Kristjansdottir et al. (2020)

Based on TK-1, The mutant strain TK-21 was engineered by knocking out the (*CruD*) gene, encoding for acyltransferase. This modification was expected to modify the lycopene structure further, leading to a glycosylated variant without acylation (Kristjansdottir, unpublished work). The proposed structure can be seen in Figure 5. In the framework of a project at Lund University and in collaboration with other Universities, called SECRETed, plans were made to enzymatically modify this compound because of its amphiphilic nature and the possibility to establish it as a biosurfactant.

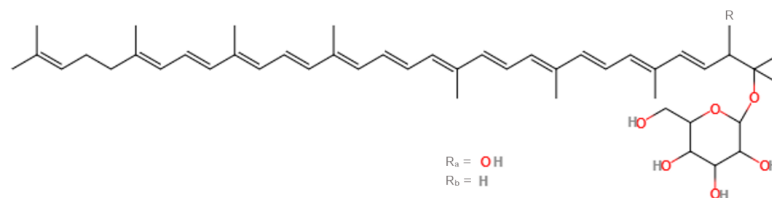


Figure 5: Proposed compound produced by *R. marinus* mutant strain TK-21

2.4 Transglycosylation reaction using CGtase

Cyclodextrin glucanotransferase (CGtase) are enzymes of the α -amylase superfamily and belong to the group of transferases. They catalyse transglycosylation reactions and also catalyse the cleavage of glycosidic bonds between carbohydrates or between a carbohydrate and a non-carbohydrate, forming cyclic products in the process (Qi and Zimmermann, 2005). An enzyme of this group was planned to be used to transform the TK21 extract from mono-glycosylated (substrate acceptor) to oligo-glycosylated, utilising carbohydrate as substrate donor, as seen in Figure 6.

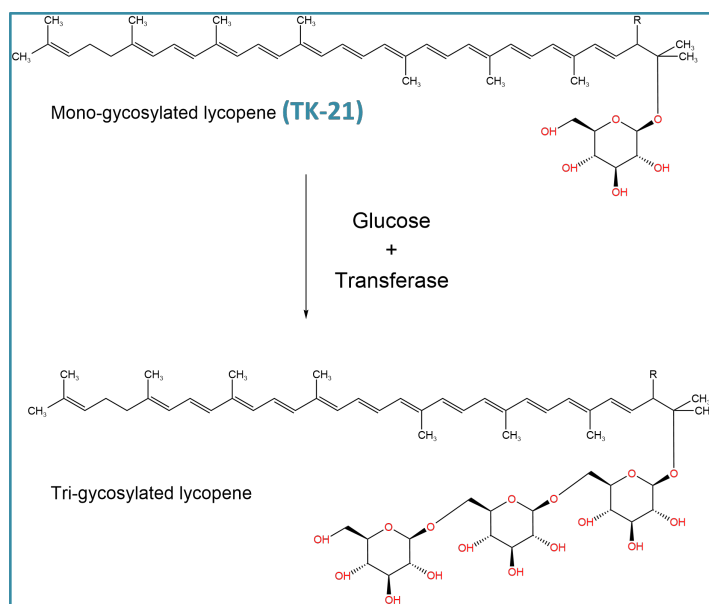
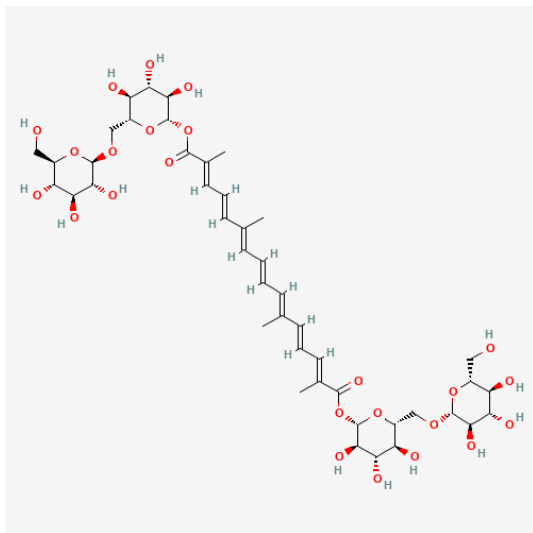
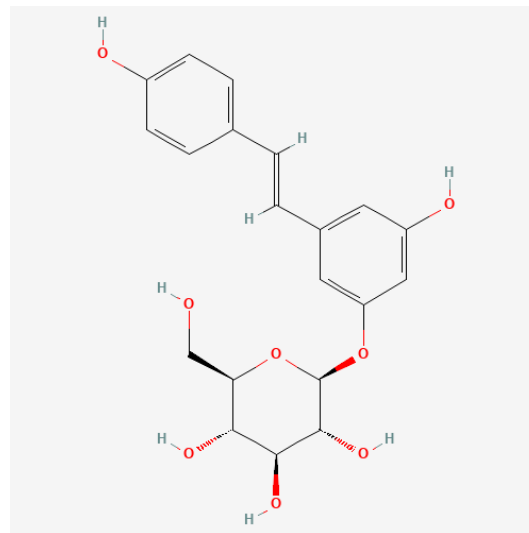


Figure 6: Glycosylation reaction of *R. marinus* TK21. Source SECRETed project

To confirm the efficiency of transglycosylation using *R. marinus* TK21 extracts, a model reaction was designed with two different acceptor substrates structurally similar to the TK21-derived substrate. Crocin, a diester, was selected due to its polyene backbone, which is glycosylated at both polyene chain ends, making it a relevant structural analog (Figure 7a). Additionally, crocin is a carotenoid that naturally exhibits biosurfactant properties, making it an ideal candidate for being used in a model reaction (Naess et al., 2006). Crocin has an absorption maximum at 464 nm (PubChem 5281233).



(a) Chemical structure of crocin-1. Source PubChem 5281233



(b) Chemical structure of piceid. Source PubChem 5281718

Figure 7: Acceptor substrates used in model reactions

2.5 Aim

The aim of this work was to exploit mutant variants of *Rhodothermus marinus* type-strain PRI 493 as cell-factories/biorefineries to produce carotenoid glucosides. Said metabolites were aimed to be characterised using Liquid Chromatography Mass Spectrometry, High-Pressure Liquid Chromatography, Critical Micellar Concentration and their Emulsification Activity and Stability. After characterisation, carotenoid extracts were planned to be subjected to enzymatic alteration, elongating their sugar-end using a CGtase. The success of the enzymatic alteration was aimed to be tested using a model reaction with piceid.

3 Methods and materials

3.1 Chemicals and reagents

All chemicals used were of analytical grade. Solvents used for mobile phases of Liquid Chromatography Mass Spectrometry (LC-MS) were of LC-MS grade. Crocin, piceid and lycopene were sourced from Sigma-Aldrich with a purity of 98.94 %, >95 %, and analytical grade, respectively.

3.2 Strains

The *Rhodothermus marinus* mutant TK-21 was provided by Kristjansdottir (University of Iceland, Reykjavík).

3.3 Cultivation of *Rhodothermus marinus* TK21

3.3.1 Media preparation

Modified thermus 162 medium (M162) with the addition of 2% (w/v) NaCl was used for the fermentation of *Rhodothermus marinus* TK21. The medium recipe was modified after the recipe 'Medium 630' from the German Collection of Microorganisms and Cell Cultures GmbH (DSMZ).

For the preparation of 1 L medium of M162 + 2% (w/v) NaCl the below ingredients were weighed in, solved in 800 mL of DI-water and then autoclaved. Defined minimal medium and phosphate buffer were autoclaved separately and added aseptically after autoclaving. All liquids were autoclaved at 121 °C, 1 bar for 20 minutes.

Component	Amount
Yeast extract	2.5 g
Tryptone	2.5g
NaCl	20 g
Defined mineral medium (see below)	100 mL
Phosphate buffer (see below)	100 mL
DI-water	ad 1L

Defined mineral medium (1L, pH 7.2)

Component	Amount
DI-water	ad 1L
Nitriloacetic acid	1.32 g
CaSO ₄ x2H ₂ O	0.4 g
MgCl ₂ x 6H ₂ O	2.0 g
Trace element solution (see below)	5 mL
0.01 M Fe citrate	5 mL

Phosphate buffer (pH 7.2)

Component	Amount
KH ₂ PO ₄	27.22 g
Na ₂ HPO ₄	28.39 g
DI-water	1L

Trace element solution

Component	Amount
Nitrilotriacetic acid	12.8 g
FeCl ₂ x 4H ₂ O	1.0 g
MnCl ₂ x 4H ₂ O	0.5 g
CoCl ₂ x 4H ₂ O	0.3 g
CuCl ₂ x 2H ₂ O	50.0 mg
Na ₂ MoO ₄ x 2H ₂ O	50.0 mg
H ₃ BO ₃	20.0 mg
NiCl ₂ x 6H ₂ O	20.0 mg
DI-water	1L

3.3.2 Inoculum

Cryocultures of *Rhodothermus marinus* strains were used for inoculation (cultures contained 25 % (v/v) glycerol). Cultures were removed from the -80 °C freezer and thawed on ice. One cryoculture was used to inoculate one Erlenmeyer shake flask (approximately 1.8 mL of culture).

3.3.3 Fermentation process

Rhodothermus marinus was cultivated in a stepwise scale-up fermentation consisting of three stages. During fermentation, Erlenmeyer flasks were filled with 10 % of M162 + 2% (w/v) NaCl of the total volume of the flask. Flasks were cultivated in a shaking incubator (Ecotron, Infors HT) at 65 °C at 200 rpm. As seen in figure 8, for the first fermentation step, 250 mL Erlenmeyer flasks were inoculated as described in section 3.3.2 and incubated in the shaking incubator for 18 hours or until $OD_{600} > 2.5$. The second and third fermentation steps were both inoculated with 10 % (v/v). The second fermentation stage was cultivated under the same conditions as stated above for 12 hours, or until $OD_{600} > 2.5$, and the last fermentation step was cultivated under the same conditions for 14-19 hours.

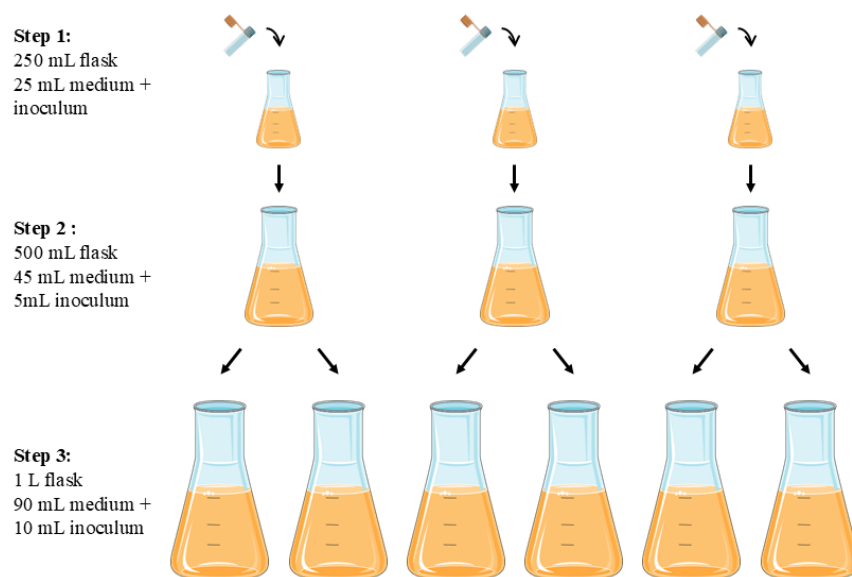


Figure 8: Schematic of stepwise fermentation of *Rhodothermus marinus*. Images adapted from Smart Servier Medical Art (smart.servier.com)

After the third fermentation step was completed, all cells were pooled and centrifuged at 3893xg

(Sorvall LYNX 4000, ThermoFisher Scientific), 4 °C for 25 minutes. The medium was decanted and the cells were washed with DI-water before being centrifuged again under the same conditions. The DI-water was decanted to receive a washed cell pellet.

3.4 Extraction of carotenoids

The extraction method for carotenoids from *Rhodothermus marinus* was conducted using a modified version of the extraction method used by Lutnaes et al. (2004). Carotenoids were extracted from the cell pellet using organic solvent extraction. For the first extraction step, the cell pellet was dissolved in approximately 300 mL of 100 % acetone. The cells were subsequently centrifuged as described in section 3.3.3, and the supernatant was decanted and saved in a Schott flask. The second extraction step was done by dissolving the cell pellet in a small amount of a 80:20 (v/v) mixture of methanol (100%) and acetone (100%). The cells were subsequently centrifuged again, and the solvent was decanted and added to the Schott flask containing the acetone extract. This step was repeated until no further extraction of carotenoids could be observed by colouration of the solvent.

The extract was filtered using vacuum filtration with a Whatman filter to remove any suspended cell particles that were accidentally decanted. The extract was dried using a rotary evaporator (Rotavapor R-300, BÜCHI Labortechnik AG), using the solvent library setting for acetone. During evaporation, the room light and the water bath were both turned off to protect the carotenoids. The extract was resuspended in 3 mL of 100 % acetone and filtered using a 0.2 µm PTFE syringe filter. The concentrate was completely dried under nitrogen gas and stored at -80 °C until further use.

3.5 Analytical methods

For analysis, the dried *R. marinus* extracts were dissolved in 100 % acetone and filtrated using a 0.2 µm PTFE syringe filter. A lycopene standard was used in HPLC and LC-MS runs. For that standard, a 0.3 g/L solution of lycopene in 100 % acetone was prepared and filtrated using a 0.2 µm PTFE syringe filter.

3.5.1 HPLC

HPLC analysis was performed using a Ultra High Performance Liquid Chromatography (UltiMate 3000, Thermo Scientific) system. A diode array detector (DAD) with wavelengths set to 261, 327,

440 and 467 nm was used for detection. Chromatographic separation was achieved using an Acclaim Vanquish C18 column (150 mm x 2.1 mm, 2.2 μ m, Thermo Scientific). The mobile phase consisted of (A) mili-Q water + 0.1 % formic acid and (B) 100 % acetonitrile in a gradient elution programmed as follows: (B) from 5 to 30 % until 5 minutes, from 30 to 50 % until 15 minutes, from 50 to 95 % until 25 minutes, at 95 % until 30 minutes, and down to 5 % until 32 minutes. The run was 35 minutes long, and the flow was set at 0.3 mL/min.

3.5.2 LC-MS

LC-MS analysis was performed using a Ultra High Performance Liquid Chromatography (UltiMate 3000 RS, Thermo Scientific) system, equipped with a UV-Vis detector and coupled to a dual-cell linear ion trap mass spectrometer (Velos Pro, Thermo Scientific). UV-Vis acquire channel wavelengths were set to 250, 450, and 460 nm.

The chromatographic separation was achieved by injection of 20 μ L of sample volume using an ACE 5 C8 column (50 mm x 3.0 mm, ACE HPLC columns, Santec Lab). The mobile phase consisted of (A) mili-Q water + 0.1 % (v/v) formic acid and (B) 100 % acetonitrile and samples were measured in a gradient elution program as seen in Figure 9.

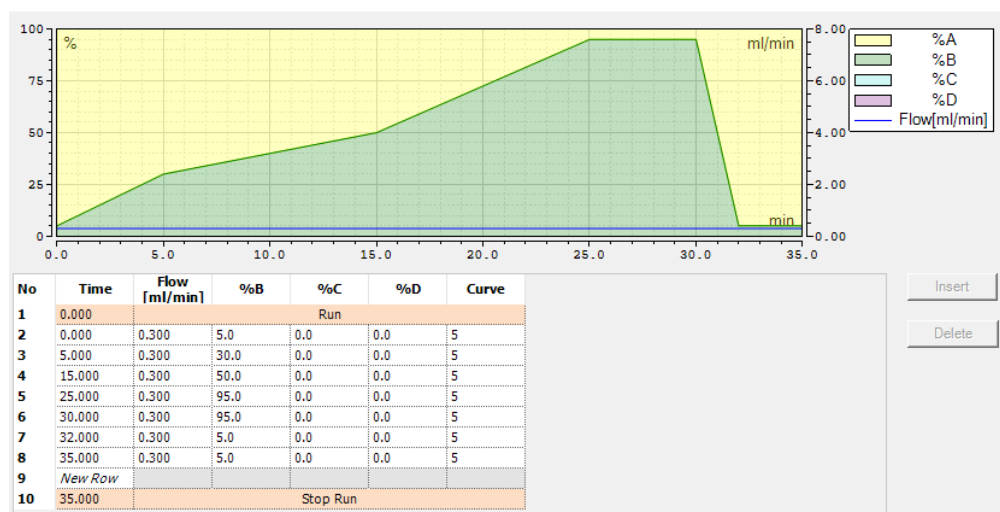


Figure 9: Flow gradient used in instrument method

The mass detector was operated under the following conditions: capillary temperature of 275 °C, source temperature of 250 °C, sheath gas flow of 15 arbitrary units, auxiliary gas flow of 10 arbitrary units, and a source voltage of 3.00 kV. The MS spectra were collected in positive mode. MS² was

performed using collision induced dissociation (CID) with a collision energy of 30.00 V for some of the MS runs.

3.5.3 Method optimisation

In efforts to optimise the MS method, different mobile phase modifications were made. For the first modification, 0.1 % (v/v) formic acid was added to mobile phase (B). For the second modification, 10 mM ammonium formate was added to mobile phase (A).

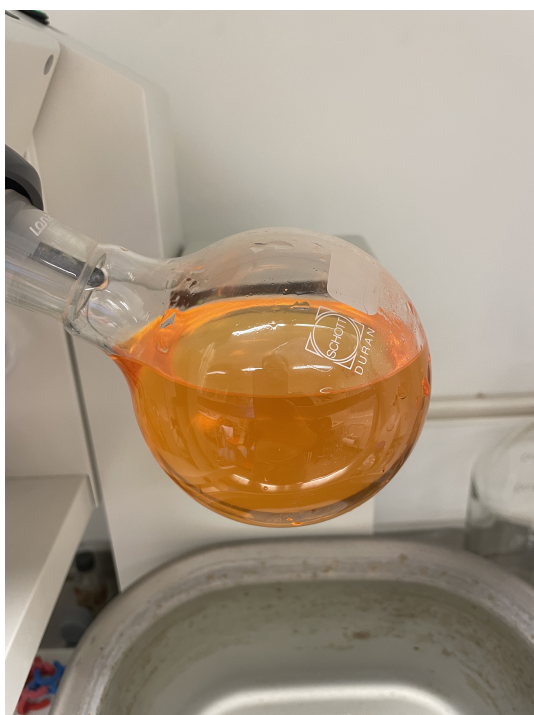
3.6 Model enzymatic reaction using piceid and crocin

Enzymatic reactions were carried out using a modified protocol from Mathew et al. (2012). The reaction mixture consisted of the acceptor substrate crocin 0.1 % (w) in MiliQ-water or piceid 0.1 % (w) in 100 % MeOH, the donor maltodextrin 5 % (w) in 0.02 M citrate phosphate buffer, pH 6.0, methanol (5% v/v), CaCl_2 5 mM and 10 μL enzyme (Toruzyne, Novonesis) in a total volume of 3 mL. The reaction was conducted using a thermoshaker (Hettich Lab Technology) at 700 rpm at 65 °C. For the piceid reaction, samples were taken at 0, 1, 2, 3, and 4 hours. For the crocin reaction, samples were taken at 20, 40, 60 minutes, and 4 hours. Sampling volume was 100 μL each. A crocin standard of 10 g/L was prepared in mili-Q water. A piceid standard of 0.8 g/L was prepared in ethyl acetate. The reactions were stopped immediately after each sampling step by adding 100 μL 0.1 M HCl. To neutralise the solution for HPLC application, 300 μL of 50 mM citrate phosphate buffer pH 6 were added.

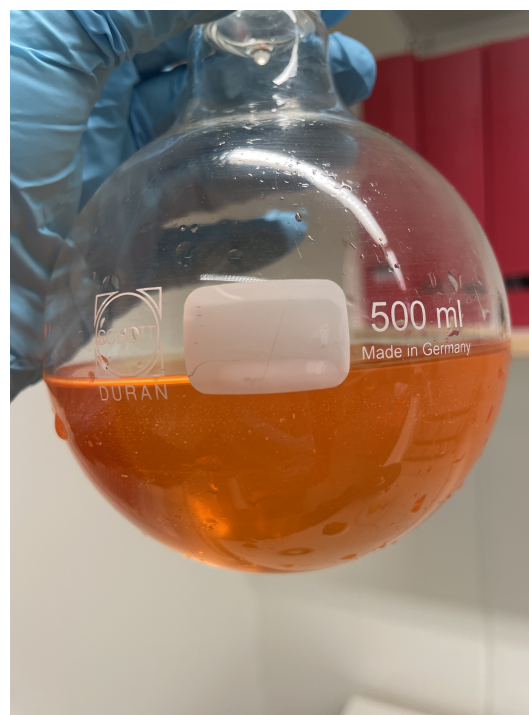
4 Results and discussion

4.1 Cultivation and extraction of carotenoids from *Rhodothermus marinus* TK21

R. marinus TK21 was successfully cultivated using the described protocol. After extraction and filtration, the carotenoids exhibited a vibrant light orange colour, as seen in Figure 10a. During the process, the colour changed towards a more dark, saturated orange, and the forming of small precipitation could be observed (Figure 10b). This effect became more and more visible the more concentrated the extract became.



(a) Carotenoids extracted from *R. marinus* TK21 before concentrating



(b) Carotenoids extracted from *R. marinus* TK21 during concentrating

Figure 10: Process of concentrating the carotenoid extracts

This phenomenon became particularly evident when the samples were completely dried under nitrogen gas, as seen in Figure 11. During solvent evaporation, a dark red precipitate gradually formed, increasing in size as the solvent evaporated. Carotenoids have previously been reported to form aggregates, especially in hydrated polar solvents and in high carotenoid concentrations due to their hydrophobicity. This self-assembly is driven by reversible hydrogen bonding, dipole interactions, van der Waals

forces and hydrophobic effects (Köhn et al., 2008). However, in this specific case, the solvent used was not hydrated, suggesting that the aggregation occurred due to increasing concentration of carotenoids as the solvent evaporated. It was thus of importance to resolve the dried carotenoid extracts in a sufficient amount of solvent and to filtrate the solution before LC-MS analysis, preventing solid particles from entering the sensitive instrumentation.



Figure 11: Precipitate forming in extracts of *Rh. marinus* TK21 during the process of drying

4.2 Characterisation of carotenoid extracts from *R. marinus* TK21

To verify that the *R. marinus* mutant TK21 successfully produced glycosylated lycopene, as anticipated based on its genetic modifications, extracts were analysed using LC-MS. The LC-MS analysis of the first extract can be seen in Figure 12. Because the maximum absorbance wavelength of the glycosylated lycopene is not expected to change significantly, the wavelength of interest here was the same as for lycopene, which is 460 nm (Popescu et al., 2022). The same motivation led to the ion source being chosen to operate in positive mode since PubChem data (PubChem 446925) stated the use of positive mode to detect lycopene. Additionally, Kristjansdottir et al. (2020) reported to have detected the modified lycopene structure (Figure 4) from TK-1 using MS operating in positive mode. The peaks, with retention times between 26.72 and 28.73, that were detected at 460 nm, exhibit the characteristic 4-peak pattern that was already described by Ron et al. (2018) for carotenoids found

in *R. marinus* strains DSM 4252, DSM 4253 and PRI 493. There, the 4 peaks were identified as the hydroxylated and non-hydroxylated forms of the novel carotenoid acyl glucoside (Figure 2) and of the hydroxylated and non-hydroxylated forms of salinixanthin (Figure 1).

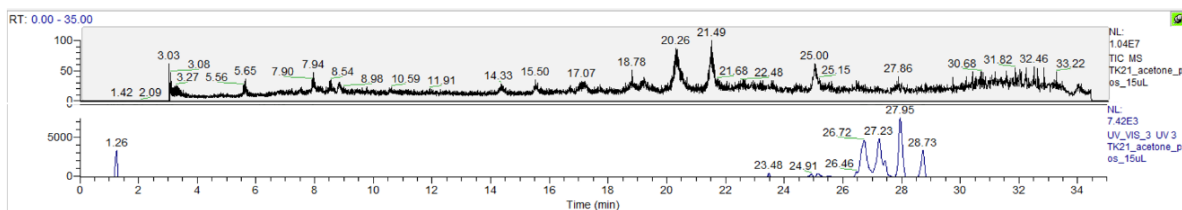


Figure 12: LC-MS data of the first *R. marinus* TK21 extract. Top: total ion count (TIC) chromatogram in (+) mode, bottom: Chromatogram at $\lambda = 460$ nm

Analysis of the MS (+) spectrum of peaks 1 and 2, with retention of times 26.72 and 27.23 minutes, respectively, reveals a very low signal intensity. This suggests inefficient ionisation under the chosen MS settings, and also explains the large number of other ions, which are probably noise and visible due to a very low signal to noise ratio.

In peak 1, the most abundant ion at m/z 730.48 corresponds to the mass of the hydroxylated derivative of the proposed TK21 structure, suggesting that the mutant successfully produces the proposed compounds. The second most abundant ion in that peak at m/z 694.51 has a mass difference of 35.97 Da and could not be identified with certainty. In peak 2, the most abundant ion at m/z 714.48 corresponds to the mass of the non-hydroxylated derivative of the proposed TK21 structure, emphasizing that the mutant does produce the proposed compounds. The second most abundant ion in that peak 2 at m/z 570.46 is not possible to identify with certainty, but it could be that this is a doubly hydroxylated lycopene. Lycopene has a mass of m/z 536.44 and the mass difference of 34.02 Da might suggest a lycopene variant produced by this strain that is hydroxylated in two positions.

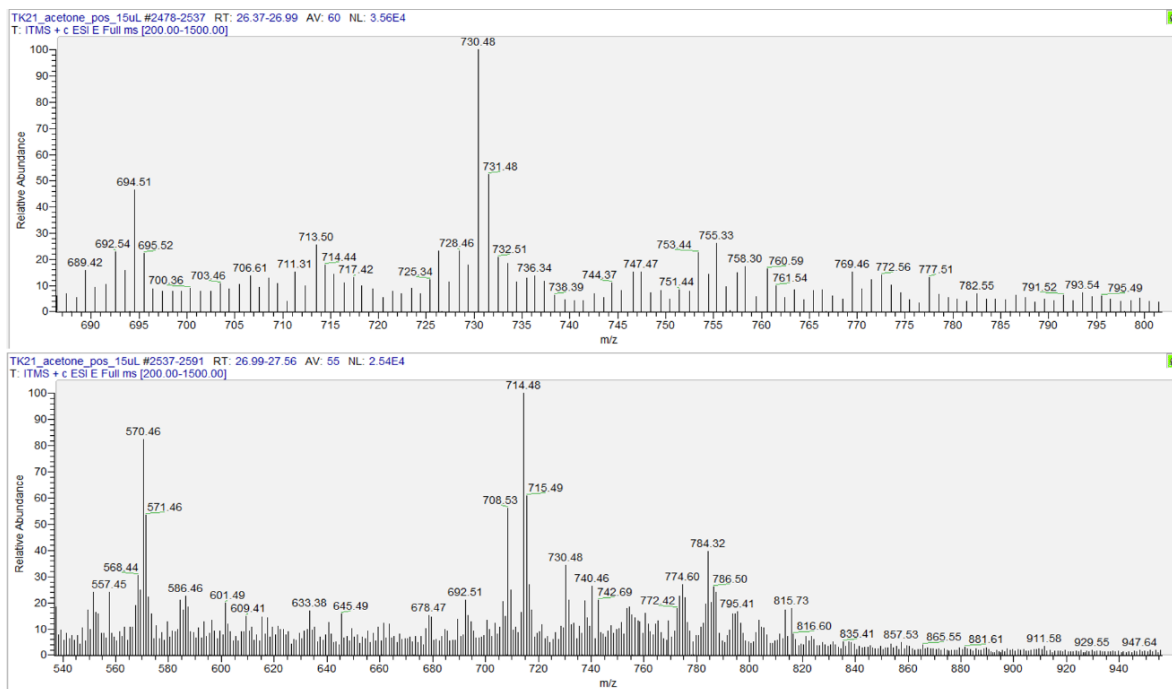


Figure 13: LC-MS data of the first *R. marinus* TK21 extract. Top: MS (+) spectrum of peak 1 ($R_t = 26.72$ min), bottom: MS (+) spectrum of peak 2 ($R_t = 27.23$ min)

Analysis of the MS (+) spectrum of peaks 3 and 4, observed at retention times 27.95 and 28.73 minutes, respectively, also revealed low signal intensity. In peak 3, the most abundant ion at m/z 859.49 may correspond to a salinixanthin derivate, as this peak has previously been identified as salinixanthin by Ron et al. (2018). Further MS² fragmentation analysis in future studies could help confirm its structural identity. In peak 4, many ions with high abundance were detected. However, the most abundant and second most abundant ions at m/z 469.34 and m/z 752.57 remain unknown. In this peak, Ron et al. (2018) previously reported to have found the mass of hydroxylated salinixanthin at m/z 940.6415, suggesting that this peak could represent a derivative of hydroxylated salinixanthin.

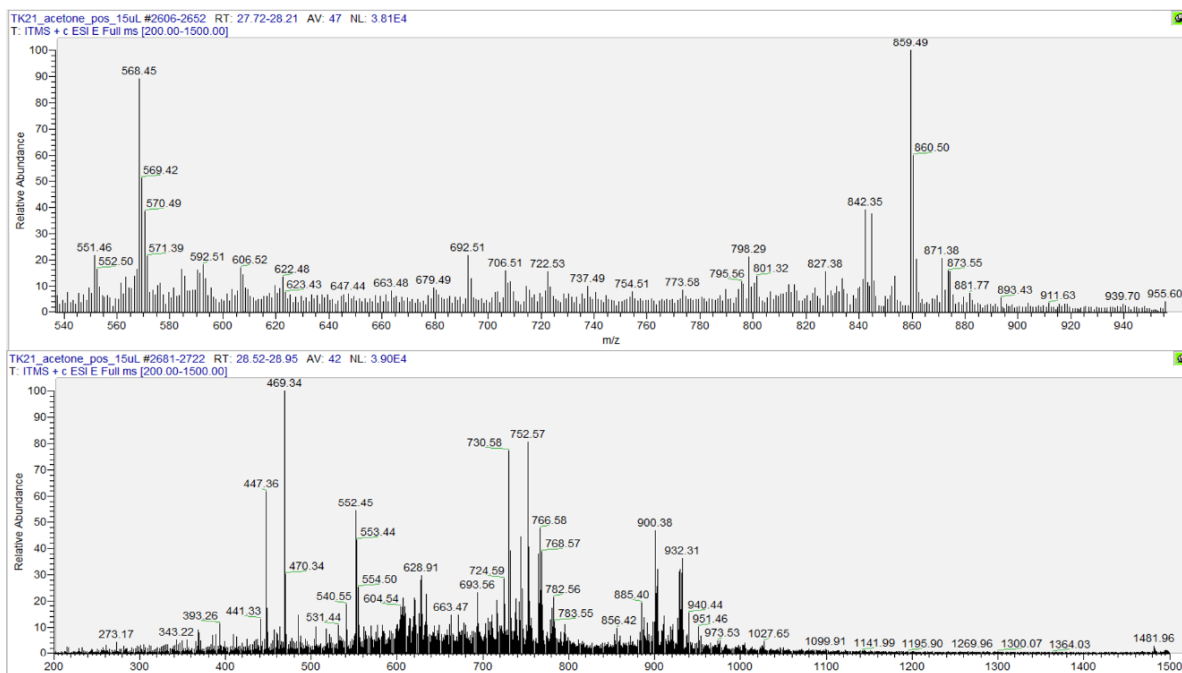


Figure 14: LC-MS data of the first *R. marinus* TK21 extract. Top: MS (+) spectrum of peak 3 ($R_t = 27.95$ min), bottom: MS (+) spectrum of peak 4 ($R_t = 28.73$ min)

To confirm the findings from the initial LC-MS run, additional extracts were analysed using the same conditions (Figure 15). The chromatogram at 460 nm once again displayed the characteristic 4-peak pattern between R_t 26.56 and 28.76 minutes, consistent with observations from the previous extract. Additionally, a fifth peak was visible at 460 nm at a R_t of 30.23 minutes. This peak was identified to be a lycopene standard contamination since the retention time and the mass matched those previously found in LC-MS analyses on lycopene as standard (Appendix A.1, Figure 32). There, lycopene was found to have a retention time of 30.23 and a mass at m/z 536.45, consistent with values in literature (PubChem 446925). This contamination likely originated from the standard sample, which was routinely run before every TK21 analysis. Several column washing steps were conducted in an attempt to remove residual lycopene on the column and prevent contamination of future. However, despite extensive cleaning efforts, the contamination persisted across multiple LC-MS analyses, suggesting that the cleaning of the column was not entirely successful.

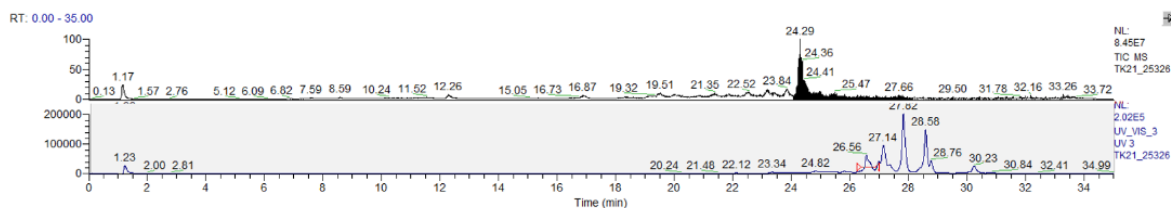


Figure 15: LC-MS data of the second *R. marinus* TK21 extract. Top: total ion count (TIC) chromatogram in (+) mode, bottom: Chromatogram at $\lambda = 460$ nm

Analysis of the MS (+) spectrum of peaks 1 and 2, with retention of times 26.56 and 27.14 minutes, respectively, revealed a higher signal intensity, indicating that higher ionisation of the compounds was achieved in this run. However, the previously detected masses corresponding to glycosylated lycopene (hydroxylated and non-hydroxylated) were not observed in these peaks. Given the similar retention compared to the previous data, similar results were expected. The most abundant ion found in peak 1 was at m/z 706.53, which is a mass difference of 23.95 Da to the anticipated m/z 730.48. This could either suggest a loss of two hydroxy groups from the expected compound mass and indicate degradation or structural modification of the compound or the expected compound has never existed in this extract. To confirm or rule out either of these hypotheses, MS² fragmentation analysis would be necessary. A similar situation was observed for peak 2, where the most abundant ion at m/z 368.43 could not be identified, while the second most abundant ion at m/z 706.53 matched the dominant ion in peak 1.

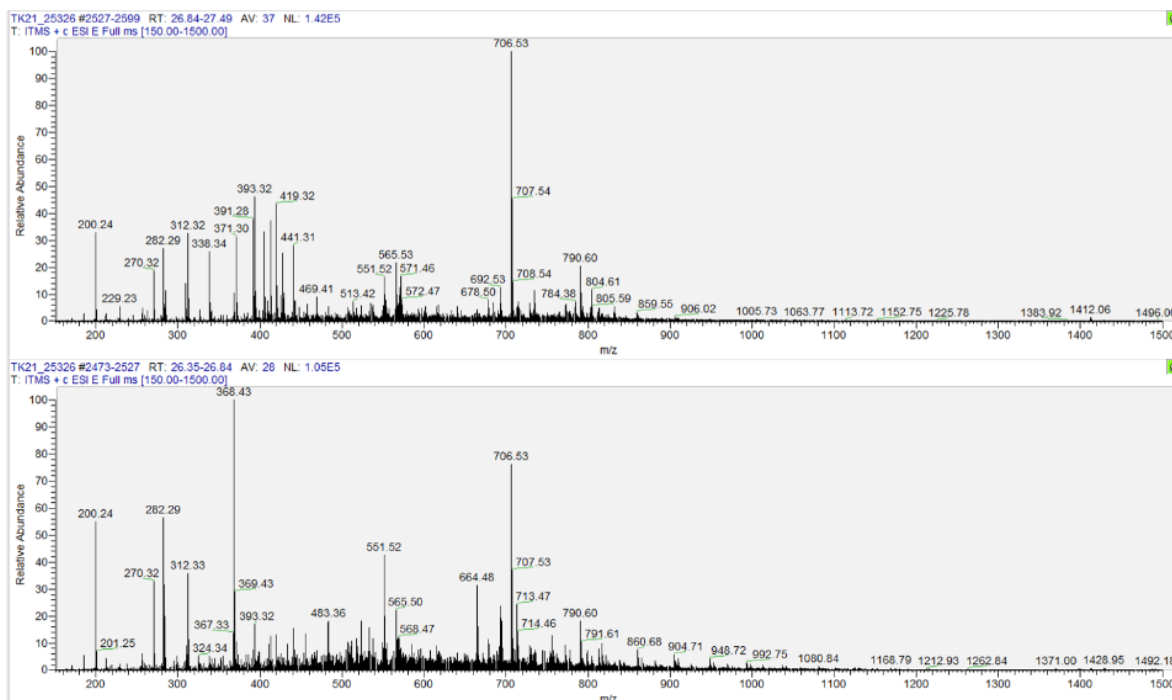


Figure 16: LC-MS data of the second *R. marinus* TK21 extract. Top: MS (+) spectrum of peak 2 ($R_t = 26.56$ min), bottom: MS (+) spectrum of peak 2 ($R_t = 27.14$ min)

Analysis of the MS (+) spectrum of peak 3 (Figure 17, top) with a retention time of 27.62, revealed that the most abundant mass in that peak was at m/z 419.31 and the second most abundant mass was at m/z 859.5, which both could not be identified. In future studies, further analysis for example MS² fragmentation analysis or structural analysis using NMR should be considered to help elucidate the structure or identity of those compounds. Analysis of the MS (+) spectrum of peak 4 (Figure 17, bottom) at a retention time of 29.76 minutes showed that the most abundant mass in that peak was at m/z 447.35, and the second most abundant mass was found to be at m/z 678.51. Both masses could not be identified, but the second mass would later be found in an optimisation run in that same peak, indicating that this mass is a reoccurring presence in extracts from this *R. marinus* mutant. It would thus be interesting to elucidate the structure of this mass to gain further knowledge on the compounds that are produced by this strain. Again, this could be done by MS² fragmentation analysis or structural analysis like NMR in future studies.

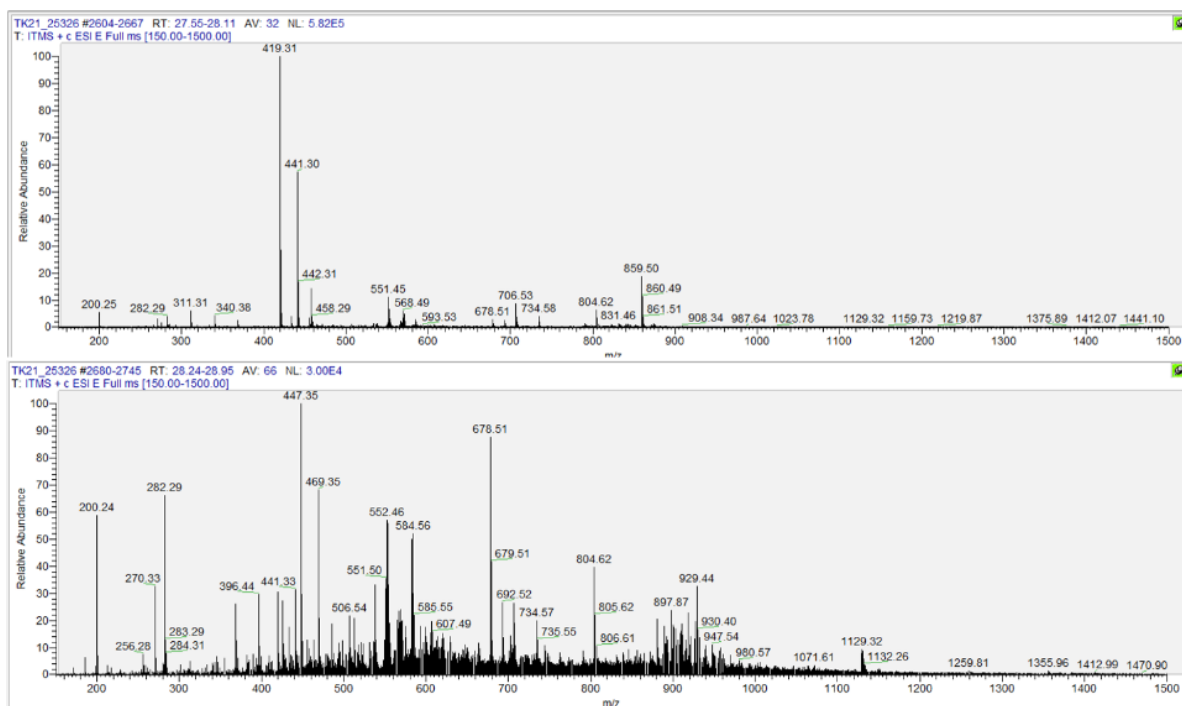


Figure 17: LC-MS data of the second *R. marinus* TK21 extract. Top: MS (+) spectrum of peak 3 ($R_t = 27.62$ min), bottom: MS (+) spectrum of peak 4 ($R_t = 28.76$ min)

4.2.1 Optimisation of MS method

Throughout all LC-MS analyses it was obvious, that the Total Ion Count (TIC) chromatograms never displayed the four characteristic peaks that were identified with the UV detector. This suggests that the compounds responsible for those peaks in at 460 nm either did not ionise or exhibited poor ionisation under the selected MS settings. As a result, efforts were made to optimise the MS method to improve ionisation efficiency. Since the existing mode of operation was successfully used to detect structurally similar compounds from other *R. marinus* mutants, the mode of operation was not changed. Instead two common optimisation approaches were taken, where a) 0.1 % formic acid was added to the mobile phase that did not contain formic acid before (mobile phase B) and b) ammonium formate was added to mobile phase A to help enhance the ionisation efficiency.

Figure 18 shows the optimisation approach a), where 0.1 % formic acid was added to mobile phase B. It can be seen, that the TIC still did not align with the peaks, exhibited at 460 nm. This suggested that the method was not sufficient for optimising ionisation efficiency. Additionally, the lycopene contamination at R_t 30.16 was also seen in this LC-MS run.

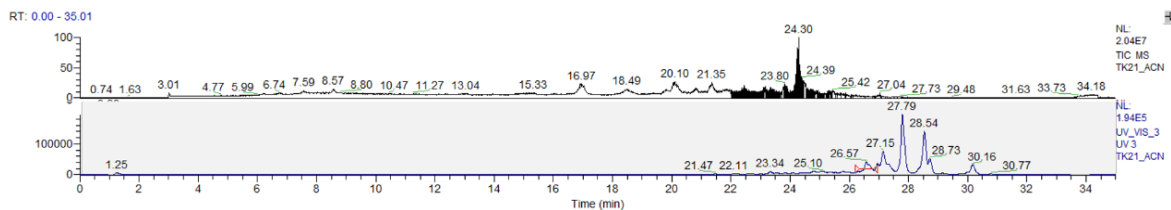


Figure 18: LC-MS data showing optimisation approach a) with added formic acid. Top: total ion count (TIC) chromatogram in (+) mode, bottom: Chromatogram at $\lambda = 460$ nm

Optimisation approach b), where 10 mM ammonium formate was added to mobile phase A, did seem to improve ionisation a lot. As seen in Figure 19, the TIC did show the distinct 4 peak pattern that was detected at 460 nm, indicating that the compounds that eluted in these peaks did actually ionise with the addition of ammonium formate. Unfortunately, in this run, the lycopene contamination was seen again at R_t 30.14 and did also ionise pretty well, which was not wanted.

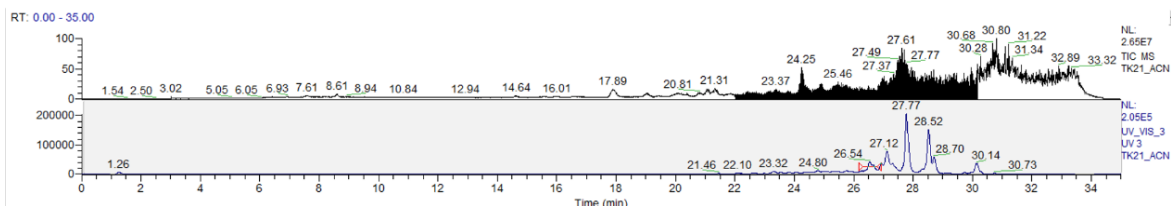


Figure 19: LC-MS data showing optimisation approach b) with 10 mM ammonium formate added to mobile phase A. Top: total ion count (TIC) chromatogram in (+) mode, bottom: Chromatogram at $\lambda = 460$ nm

The MS data for peaks 1 and 2 did not show the expected masses for glycosylated lycopene, so MS^2 was used to subject the most abundant ions within these peaks to collision induced dissociation (CID) fragmentation for structural analysis.

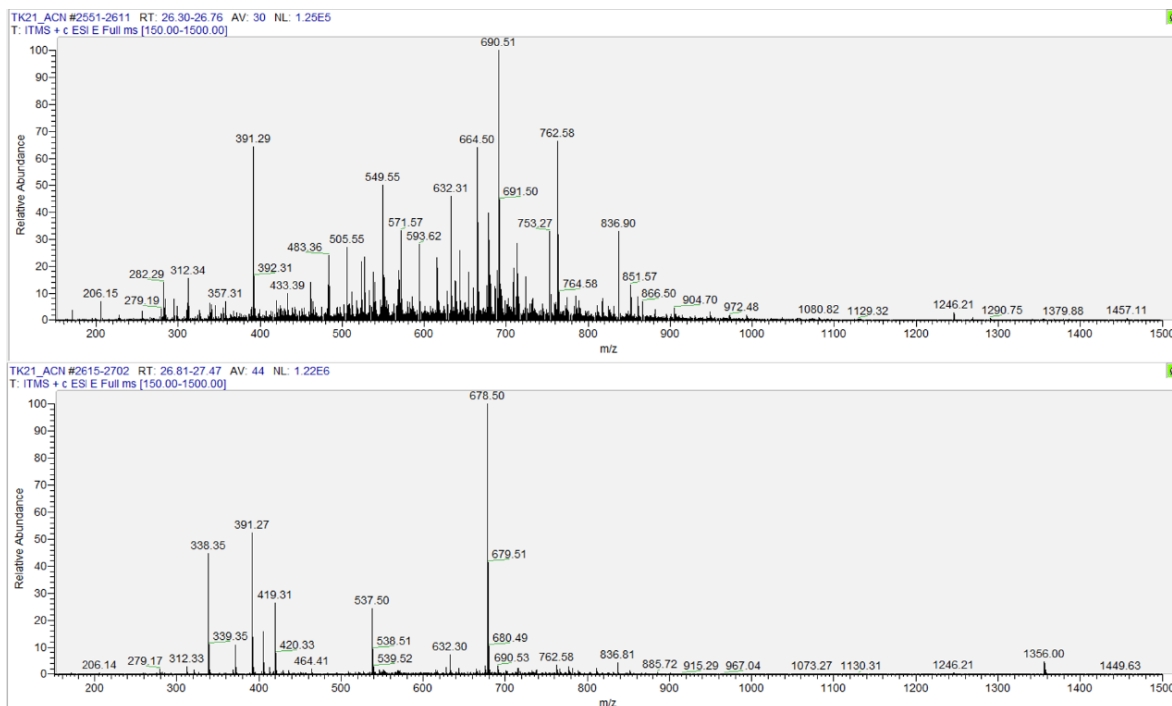
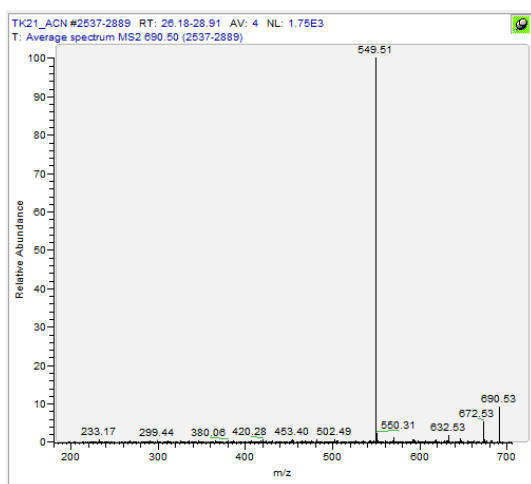
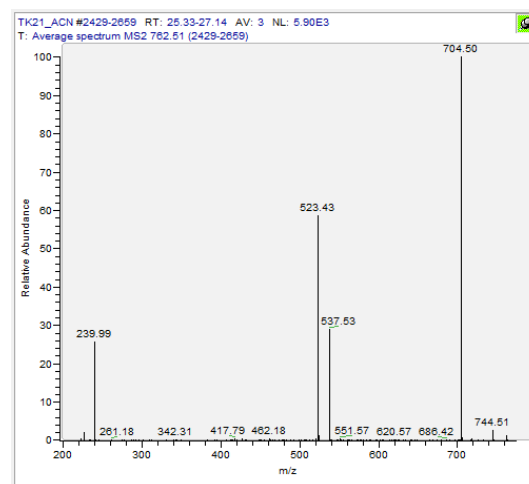


Figure 20: LC-MS data showing optimisation approach b) with 10 mM ammonium formate added to mobile phase A. Top: MS (+) spectrum of peak 1 ($R_t = 26.54$ min), bottom: MS (+) spectrum of peak 2 ($R_t = 27.12$ min)

For peak 1, the ions that were selected as precursor ions for CID were m/z 690.51 and m/z 762.58. Looking at the CID of precursor ion m/z 690.51, resulting fragments include m/z 672.53 from dehydration, m/z 672.53, which is a mass difference of 58 Da, and could not be identified, and m/z 549.51 which is a mass difference of 141.02 Da, and could not be identified. The CID of precursor ion m/z 762.58 resulted in fragments of m/z 744.51 from dehydration, m/z 704.5, which is a mass difference of 58.01 Da and could not be identified and m/z of 537.53, which is a mass difference of 225.05 Da, which could point to the loss of a specific molecular moiety. This is also the mass of lycopene, which suggests whatever structure the precursor ion has, it fragments into lycopene and strongly suggests a carotenoid structure of the kind that was suspected to be produced by *R. marinus* TK21.



(a) MS² spectrum of the precursor ion m/z 690.51



(b) MS² spectrum of the precursor ion m/z 762.58

Figure 21: MS² spectra of precursor ions selected for peak 1

In peak 2, m/z 678.5 was selected as precursor ion for CID. The resulting fragment ion is at m/z 537.51, which is the mass of lycopene and also suggests that the structure of the precursor ion m/z 678.5 in peak 2 must include lycopene in its structure, which would point to the glycosylated lycopene.

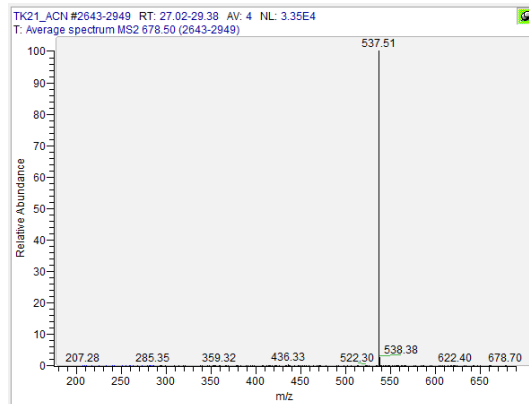


Figure 22: MS² spectra of precursor ion m/z 678.5 selected for peak 2

As for peaks 1 and 2, the MS data for peaks 3 and 4 did not show the expected masses for glycosylated lycopene, so MS² was used to subject the most abundant ions within these peaks to collision induced dissociation (CID) fragmentation for structural analysis (23). For peak 3 and 4, the same precursor ion m/z 678.5 was chosen for CID as for peak 2, leading to the same conclusion.

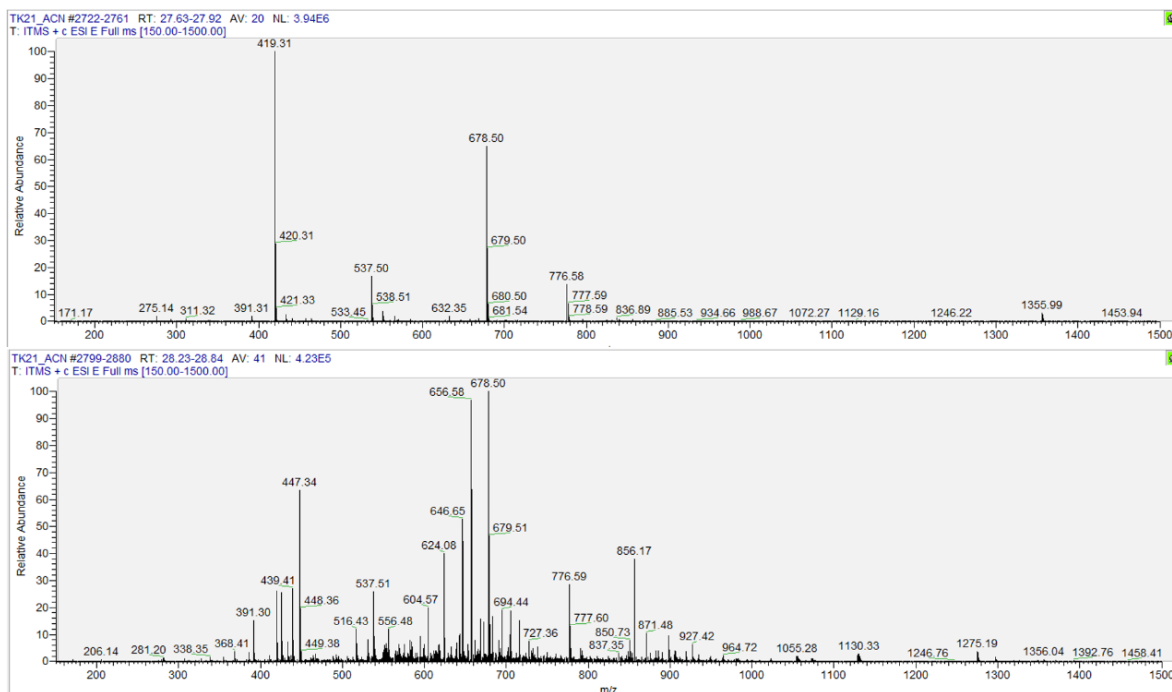


Figure 23: LC-MS data showing optimisation approach b) with 10 mM ammonium formate added to mobile phase A. Top: MS (+) spectrum of peak 3 ($R_t = 27.77$ min), bottom: MS (+) spectrum of peak 4 ($R_t = 28.25$ min)

In peak 3, m/z 776.59 was selected as precursor ion for CID. The resulting fragment ions are at m/z 758.60, from dehydration, at m/z 718.56, which is a mass difference of 58.03 Da and could not be identified. This precursor ion also fragments into m/z 537.49, which is the mass of lycopene, and also suggests that the structure of the precursor ion m/z 776.59 in peak 3 must include lycopene in its structure. This suggests that this precursor ion could also be related to glycosylated lycopene (Figure 24).

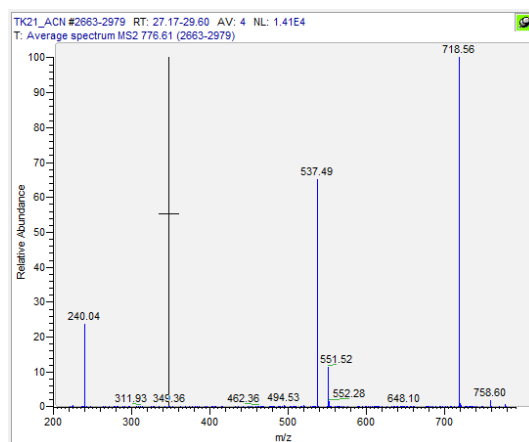
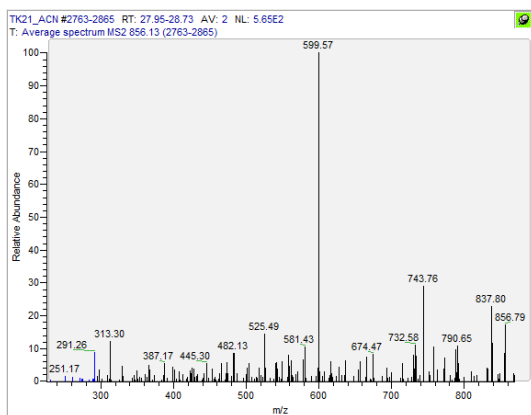
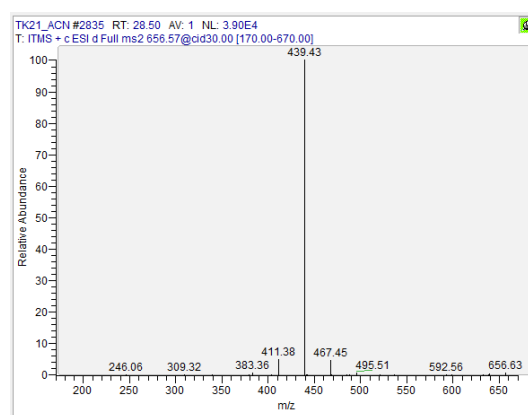


Figure 24: MS² spectra of precursor ion m/z 776.59 selected for peak 3

In peak 4, m/z 856.17 was selected as precursor ion for CID. There are many resulting fragment ions in the MS² spectrum of this precursor ion, with the most abundant one at m/z 599.57. The mass difference from the precursor ion to the fragment mass is 256.6 Da, which could stem from a fatty acid derivative, but the true identity of this precursor ion remains unknown. The precursor ion at m/z at 656.56 from peak 4 was selected as well for CID. The resulting fragment at m/z 439.43 has a mass difference of 217.13 Da, which cannot be identified.



(a) MS² spectrum of precursor ion m/z 856.17



(b) MS² spectrum of the precursor ion m/z 656.56

Figure 25: MS² spectra of precursor ions selected for peak 4

4.3 Transglycosylation model reactions using acceptor substrates piceid and crocin

Model reactions were conducted to validate the enzymatic synthesis of carotenoid glycosides using CGTase, ensuring the methodology was suitable for practical application. Two different acceptor substrates, croci, and piceid, were used in these experiments, as they provide structural similarity in the regions where the enzyme is active.

4.3.1 Enzymatic synthesis of crocin glycosides

To be able to identify crocin peaks in the reaction samples, a crocin standard of 10 g/L was analysed (Figure 26). Crocin was found to have a retention time of 7.87 min and exhibited an absorption maximum between 440 and 468 nm. The chromatogram of the components in the reaction without enzyme and acceptor substrate (negative-control) did not exhibit a peak at that retention time (A.2; Figure 33).

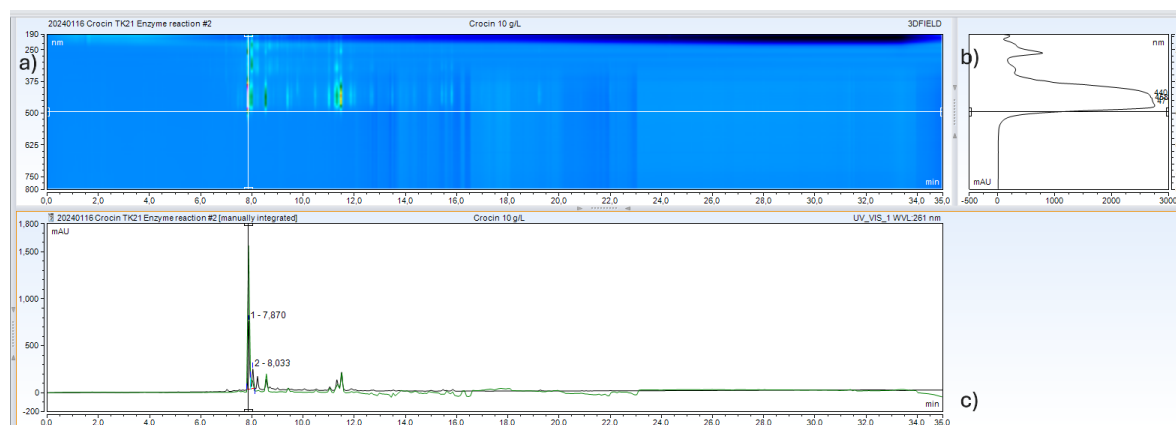


Figure 26: HPLC run of 10 g/L crocin standard. Shown are (a) the 3D field plot (top left window) where the x-axis represents time in minutes, and the y-axis represents wavelength in nm; (b) the spectral plot (top right window) where the x-axis represents wavelength in nm, and the y-axis represents the absorbance in mAU, and (c) the chromatogram where the x-axis represents the time in min, and the y-axis represents the absorbance in mAU

The chromatogram of the 0 h sample exhibited a peak around the same retention time as the crocin standard and also showed an absorption maximum between 440 and 470 nm (Figure 27). This strongly suggests that this peak corresponds to crocin (full chromatogram in Appendix A.2, Figure 34). Notably, even at 0h, the chromatogram exhibits a distinct glycosylation pattern of multiple peaks appearing at earlier retention times, representing the different crocin derivatives with varying sugar chain lengths, as previously described by Mathew et al. (2012). Further analysis of the spectral data

for these peaks supports this, since each shows the characteristic absorbance maximum between 430 and 470 nm (Appendix A.2, Figures 35 to 38.) This unexpected reaction at 0 hours may be due to the unknown concentration of the Toruzyme enzyme solution, as the technical data sheet does not provide specific activity levels. Therefore, the enzyme activity within 10 μL of enzyme solution may have been too high, resulting in rapid catalysis within the short time frame between addition of enzyme, mixing of the reaction solution and sampling.

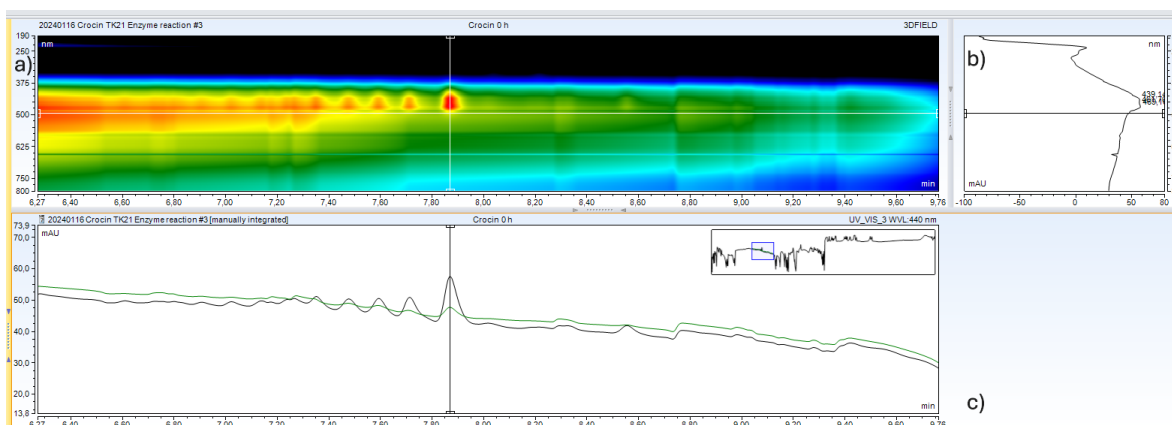


Figure 27: HPLC run of the 0 h sample of crocin. Shown are (a) the 3D field plot (top left window) where the x-axis represents time in minutes and the y-axis represents wavelength in nm; (b) the spectral plot (top right window) where the x-axis represents wavelength in nm and the y-axis represents the absorbance in mAU and (c) the chromatogram where the x-axis represents the time in min and the y-axis represents the absorbance in mAU

The HPLC run of the sample taken after 20 minutes of reaction time the chromatogram showed a major peak at retention time 33 minutes and the overall signal of the run seemed to be very low, with highest peak not reaching 15 mAU. This suggests, that there might have been some sort of contamination in either the sample or column at that point of the reaction (Figure 28). Nevertheless, the chromatogram showed the distinct glycosylation pattern as described earlier, with even more peaks at a retention time lower than than of the crocin standard, even though the spectral data of those peaks does not really show absorption maxima in the expected wavelength range (Appendix A.2; Figures 39 to 41). This also suggests that something did not go quite right with this sample.

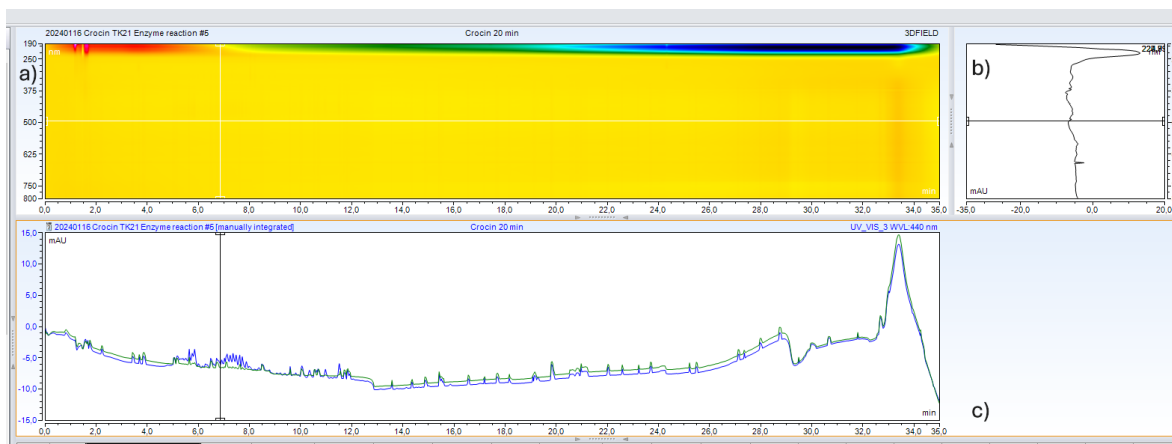


Figure 28: HPLC run of the 20 minute sample of crocin. Shown are (a) the 3D field plot (top left window) where the x-axis represents time in minutes and the y-axis represents wavelength in nm; (b) the spectral plot (top right window) where the x-axis represents wavelength in nm and the y-axis represents the absorbance in mAU and (c) the chromatogram where the x-axis represents the time in min and the y-axis represents the absorbance in mAU

The chromatographic analysis of the 40 minute sample revealed a pattern consistent with the expected glycosylation reaction, with spectral data showing strong absorption between 400 and 500 nm in the range of these peaks (Appendix A.2; Figures 43 to 45). Notably, both this sample and the endpoint sample at 4 hours exhibited low intensity signals, which may be due to a sampling error or potential inconsistencies within the HPLC instrument or method. The 4 hour sample maintained the above described characteristic pattern indicative of glycosylation, with spectral data showing absorption maxima around 440 nm, suggesting that the peaks correspond to crocin-derived structures (Appendix A.2; Figures 46 to 49). To definitely confirm the identity of the compounds eluting earlier than the crocin standard, further structural analysis would be necessary, using adequate techniques, such as Nuclear Magnetic Resonance (NMR) spectroscopy. Nonetheless, the present data indicates that the enzymatic transglycosylation of a glycosylated carotenoid using a CGtase could be successful.

4.3.2 Enzymatic synthesis of piceid glycosides

The piceid standard (0.8 g/L) (Figure 29) was analysed using LC-MS, revealing a UV-Vis spectrum with a retention time of 8.13 minutes. The corresponding mass data confirmed the peak's identity as piceid, since piceid has a monoisotopic mass of 390.13 (Pubchem 5281718) and the second most abundant ion in that peak has a mass at m/z 391.13. Given that the LC-MS and HPLC methods were identical, it can be assumed that piceid elutes at approximately 8 minutes under the given conditions.

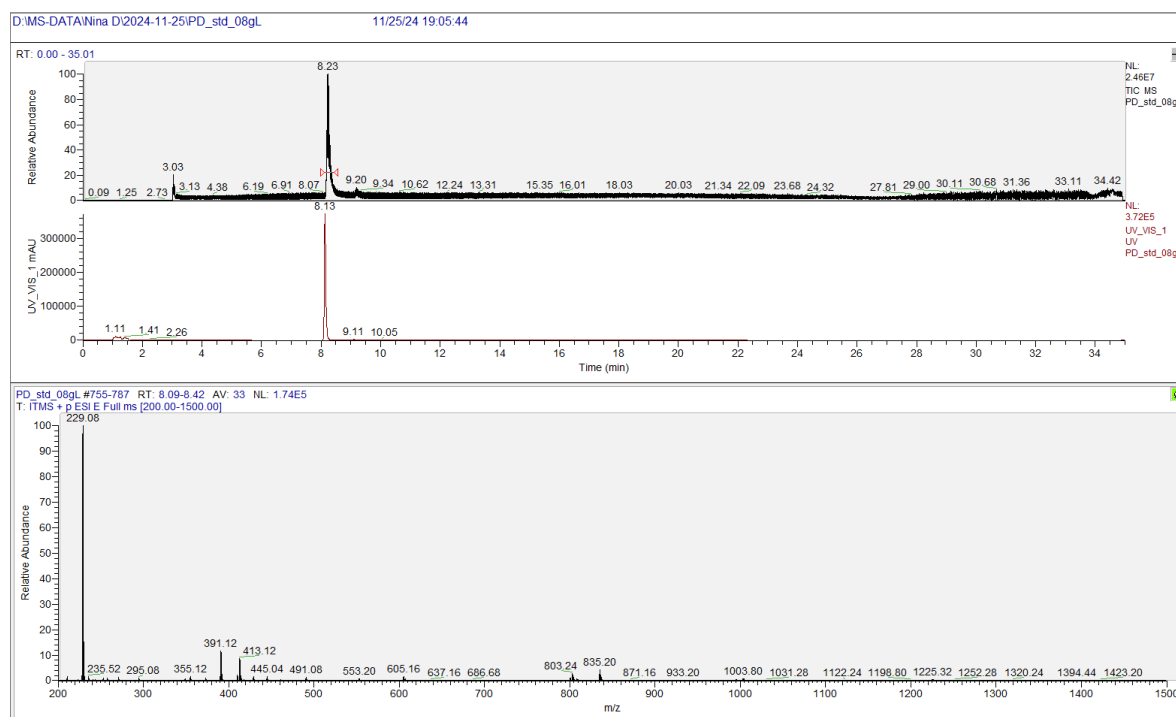


Figure 29: LC-MS data of the piceid standard, from top to bottom: TIC chromatogram, UV-Vis chromatogram, and MS (+) spectrum

The HPLC analysis of the 0h sample did not reveal a distinctive peak at the expected 8 minute retention time (Appendix A.2; Figures 50 and 51). This suggests that something went wrong during sampling or maybe even injection of the sample, especially since both peaks between 1 and 3 minutes exhibit an absorption maximum of around 309 nm, as would be expected for piceid. The sample taken after 1 h displayed the characteristic glycosylation pattern of multiple peaks appearing at earlier retention times than the standard at 8.23 minutes. These peaks likely represent the different piceid derivatives with varying sugar chain lengths, as previously described by Mathew et al. (2012). The spectral data indicates an absorption maximum between 310 and 320 nm, confirming that the peak selected

Figure 30 corresponds to a piceid derivative. The earlier eluting peaks exhibited similar absorption characteristics, supporting the hypothesis that also these peaks are glycosylated piceid derivatives (Appendix A.2; Figures 52 to 55).

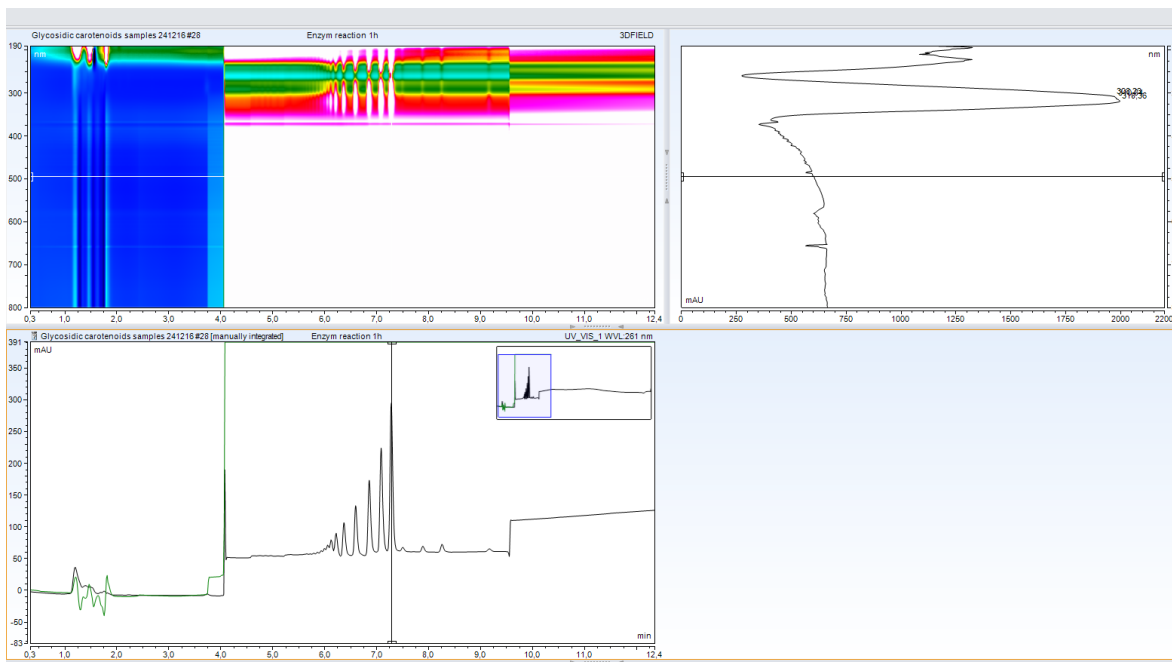


Figure 30: HPLC run of the 1 h sample of the piceid transglycosylation reaction. Shown are (a) the 3D field plot (top left window) where the x-axis represents time in minutes and the y-axis represents wavelength in nm; (b) the spectral plot (top right window) where the x-axis represents wavelength in nm, and the y-axis represents the absorbance in mAU, and (c) the chromatogram where the x-axis represents the time in min and the y-axis represents the absorbance in mAU

A comparison of peak intensities between the sample after 1 hour and the sample after 2 hours reveals minimal change in peak intensities (from approximately 290 mAU to approximately 300 mAU). This suggests that the reaction may have already been completed after 1 hour, similar to the observations in the crocin transglycosylation trial. This might be due to the fact that the concentration of the Toruzyme enzyme solution was unknown, and the enzyme activity within 10 μ L of enzyme solution might have been too high (Figure 31). The HPLC data of the 4 hour endpoint sample did not show an increase in peak intensity but rather a slight decrease, further supporting the hypothesis that the reaction had already been finished after 1 hour of reaction time (Appendix A.2; Figure 56).

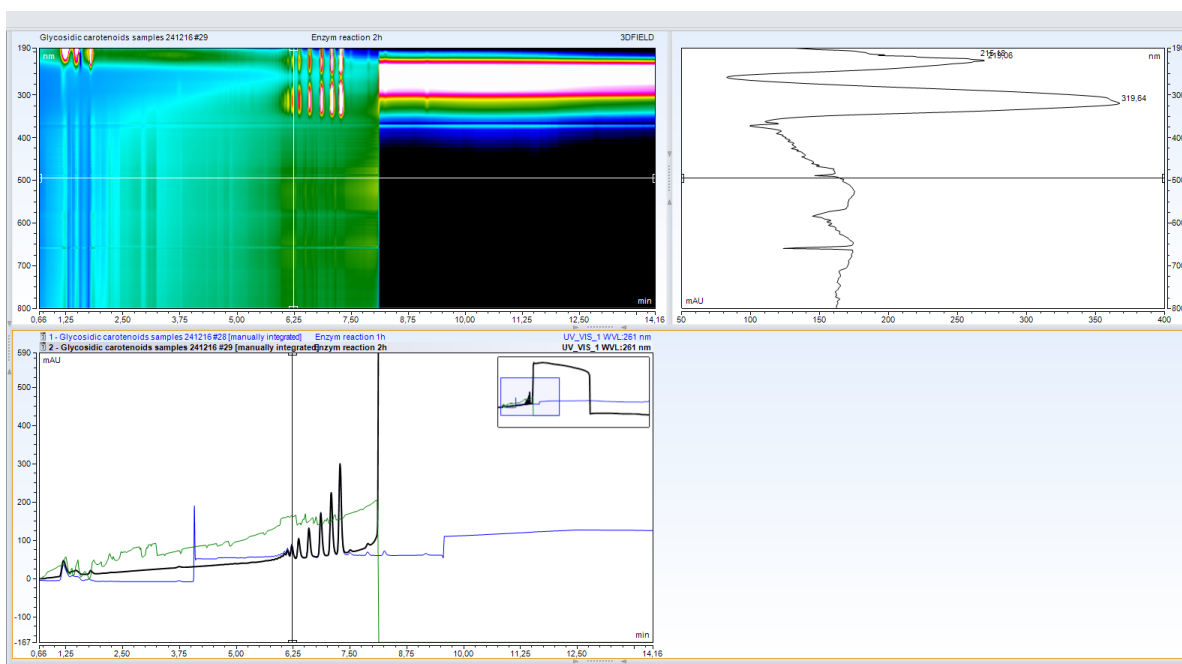


Figure 31: HPLC runs of the 1 h sample and of the 2 h sample of the piceid transglycosylation reaction. Shown are (a) the 3D field plot (top left window) where the x-axis represents time in minutes and the y-axis represents wavelength in nm; (b) the spectral plot (top right window) where the x-axis represents wavelength in nm and the y-axis represents the absorbance in mAU and (c) the chromatogram where the x-axis represents the time in min and the y-axis represents the absorbance in mAU

5 Conclusion

This study demonstrated the successful cultivation of *R. marinus* TK21 and the extraction of carotenoids from this engineered strain, which was designed to produce glycosylated lycopene derivatives. Initial challenges with compound ionisation in mass spectrometry were overcome by method optimisation, enabling a more detailed analysis of the TK21 extracts. The analytical data strongly indicates that the structure of the extracted compound features a lycopene backbone, as suggested by MS fragmentation patterns. However, the exact structural identity of this compound remains unresolved and will require further analysis, such as NMR or advanced MS techniques, to be fully characterised. The enzymatic transglycosylation reaction showed that model compounds piceid and crocin were successfully modified, underscoring the concept of carotenoid transglycosylation reactions for enhanced biosurfactant properties. Overall, this study sets the groundwork for understanding and further optimisation of using *R. marinus* TK21 as a cell factory for carotenoid production and highlights its potential for biotechnological applications.

6 Outlook

While this study provided promising evidence for the production of lycopene-based carotenoids by *R. marinus* TK21, further research is required to elucidate the structure of its extracts fully. Further research should prioritise the development of a purification method capable of separating the carotenoids from impurities like cell debris or phospholipids that could obstruct carotenoid analysis. A purified carotenoid extract would also be more suitable for more sensitive analytical methods like NMR spectroscopy to elucidate the carotenoid structure further.

In addition, it may be beneficial to re-evaluate the collected MS data (using the optimised method) for the ions corresponding to intermediate carotenoids in the biosynthetic pathway. This might help to confirm whether *R. marinus* TK21 produces glycosylated lycopene or an unexpected variant.

To build on the findings of this work, more cultivations and subsequent extractions should be conducted to use these extracts for analysis using the optimised MS method. Moreover, the successful enzymatic transglycosylation of the model compounds encourages to conduct the enzymatic modifications on the TK21-derived carotenoids, especially if purification of said extracts proves to be successful. Future work could explore reaction conditions that enable efficient and selective glycosylation of glycosylated lycopene, as well as assess their surfactant properties in functional assays, like critical micellar concentration assays or emulsification assays.

References

- Atif, F., Maqsood, N., Ali, W., Ali, W., and Irfan, M. (2024). Extremophiles and their enzymatic diversity and biotechnological potential. *Systems Microbiology and Biomanufacturing*, 4(3):833–849.
- Bjornsdottir, S. H., Blondal, T., Hreggvidsson, G. O., Eggertsson, G., Petursdottir, S., Hjorleifsdottir, S., Thorbjarnardottir, S. H., and Kristjansson, J. K. (2006). Rhodothermus marinus: physiology and molecular biology. *Extremophiles*, 10:1–16.
- Bjornsdottir, S. H., Fridjonsson, O. H., Hreggvidsson, G. O., and Eggertsson, G. (2011). Generation of targeted deletions in the genome of rhodothermus marinus. *Applied and environmental microbiology*, 77(15):5505–5512.
- Bjornsdottir, S. H., Fridjonsson, O. H., Kristjansson, J. K., and Eggertsson, G. (2007). Cloning and expression of heterologous genes in rhodothermus marinus. *Extremophiles*, 11:283–293.
- Britton, G. (1995). Structure and properties of carotenoids in relation to function. *The FASEB Journal*, 9(15):1551–1558.
- Gayathiri, E., Prakash, P., Karmegam, N., Varjani, S., Awasthi, M. K., and Ravindran, B. (2022). Biosurfactants: potential and eco-friendly material for sustainable agriculture and environmental safety—a review. *Agronomy*, 12(3):662.
- Karmakar, K., Sarkar, R., Pal, A., Rahaman, S. M., Acharjee, A., and Saha, B. (2025). Recent advances and emerging trends in biosurfactants: A concise review. *Journal of Solution Chemistry*, pages 1–28.
- Köhn, S., Kolbe, H., Korger, M., Köpsel, C., Mayer, B., Auweter, H., Lüddecke, E., Bettermann, H., and Martin, H.-D. (2008). Aggregation and interface behaviour of carotenoids. *Carotenoids: Volume 4: Natural Functions*, pages 53–98.
- Kristjánsdóttir, Þ. et al. (2022). *Development of non-model bacteria as cell factories*. PhD thesis, University of Iceland.
- Kristjansdottir, T., Ron, E. Y., Molins-Delgado, D., Fridjonsson, O. H., Turner, C., Bjornsdottir, S. H., Gudmundsson, S., van Niel, E. W., Karlsson, E. N., and Hreggvidsson, G. O. (2020). Engineering

- the carotenoid biosynthetic pathway in *rhodothermus marinus* for lycopene production. *Metabolic engineering communications*, 11:e00140.
- Lutnaes, B. F., Oren, A., and Liaaen-Jensen, S. (2002). New c40-carotenoid acyl glycoside as principal carotenoid in *salinibacter ruber*, an extremely halophilic eubacterium. *Journal of natural products*, 65(9):1340–1343.
- Lutnaes, B. F., Strand, Å., Pétursdóttir, S. K., and Liaaen-Jensen, S. (2004). Carotenoids of thermophilic bacteria—*rhodothermus marinus* from submarine icelandic hot springs. *Biochemical Systematics and Ecology*, 32(5):455–468.
- Marquez, R., Ortiz, M. S., Barrios, N., Vera, R. E., Patiño-Agudelo, Á. J., Vivas, K. A., Salas, M., Zambrano, F., and Theiner, E. (2025). Surfactants produced from carbohydrate derivatives: Part 2. a review on the value chain, synthesis, and the potential role of artificial intelligence within the biorefinery concept. *Journal of Surfactants and Detergents*, 28(1):25–76.
- Mathew, S., Hedström, M., and Adlercreutz, P. (2012). Enzymatic synthesis of piceid glycosides by cyclodextrin glucanotransferase. *Process Biochemistry*, 47(3):528–532.
- Naess, S. N., Elgsaeter, A., Foss, B. J., Li, B., Sliwka, H.-R., Partali, V., Melø, T. B., and Naqvi, K. R. (2006). Hydrophilic carotenoids: Surface properties and aggregation of crocin as a biosurfactant. *Helvetica chimica acta*, 89(1):45–53.
- Nolan, M., Tindall, B. J., Pomrenke, H., Lapidus, A., Copeland, A., Glavina Del Rio, T., Lucas, S., Chen, F., Tice, H., Cheng, J.-F., et al. (2009). Complete genome sequence of *rhodothermus marinus* type strain (r-10 t). *Standards in genomic sciences*, 1:283–291.
- Popescu, M., Iancu, P., Pleșu, V., Cristina, C., and Todasca, M. (2022). Different spectrophotometric methods for simultaneous quantification of lycopene and β -carotene from a binary mixture. *lwt* 160: 113238.
- Qi, Q. and Zimmermann, W. (2005). Cyclodextrin glucanotransferase: from gene to applications. *Applied Microbiology and Biotechnology*, 66:475–485.
- Romero Vega, G. and Gallo Stampino, P. (2025). Bio-based surfactants and biosurfactants: An overview and main characteristics. *Molecules*, 30(4):863.

- Ron, E. (2020). *Rhodothermus marinus: a cell factory: Exploring products and cultivation technology*. PhD thesis, Lund University.
- Ron, E. Y., Plaza, M., Kristjansdottir, T., Sardari, R. R., Bjornsdottir, S. H., Gudmundsson, S., Hreggvidsson, G. O., Turner, C., van Niel, E. W., and Nordberg-Karlsson, E. (2018). Characterization of carotenoids in *rhodothermus marinus*. *Microbiologyopen*, 7(1):e00536.
- Ron, E. Z. and Rosenberg, E. (2002). Biosurfactants and oil bioremediation. *Current opinion in biotechnology*, 13(3):249–252.
- Sardari, R. R., Kulcinskaja, E., Ron, E. Y., Björnsdóttir, S., Friðjónsson, Ó. H., Hreggviðsson, G. Ó., and Karlsson, E. N. (2017). Evaluation of the production of exopolysaccharides by two strains of the thermophilic bacterium *rhodothermus marinus*. *Carbohydrate polymers*, 156:1–8.

A Appendix

A.1 Characterisation of carotenoid extracts from *R. marinus* TK21

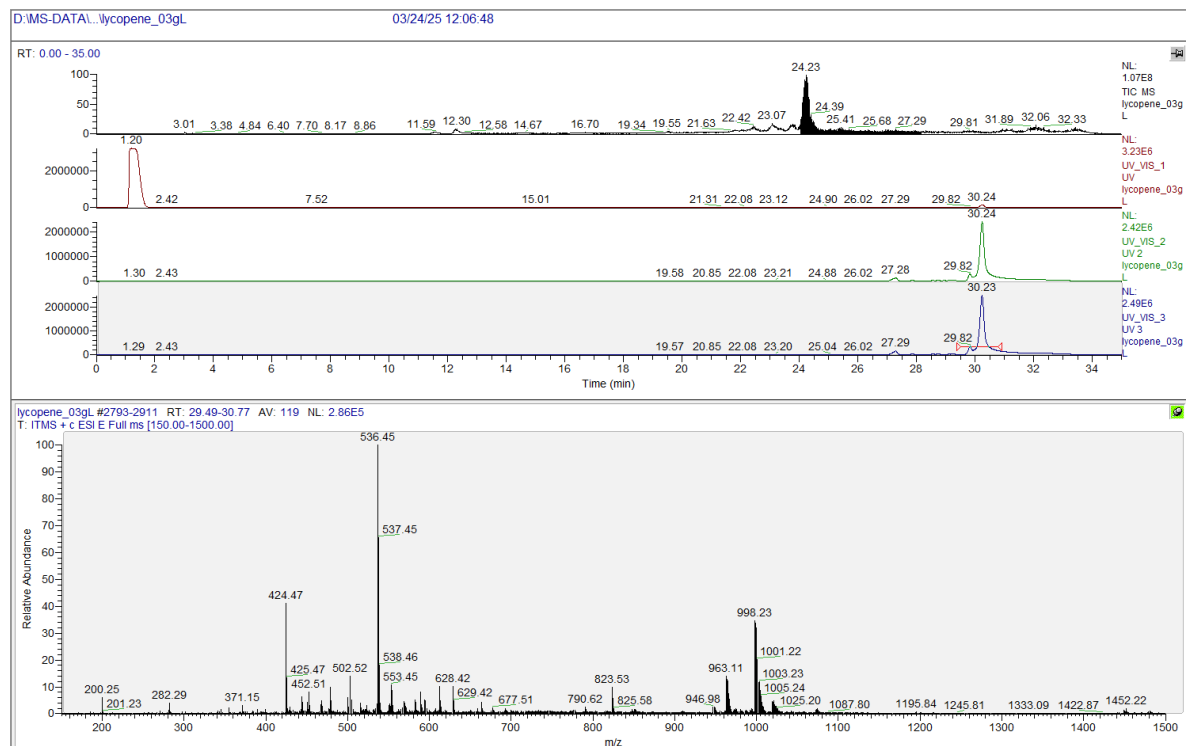


Figure 32: LC-MS data of the lycopene standard. From top to bottom: total ion count (TIC) chromatogram in (+) mode, UV1 chromatogram at $\lambda = 250$ nm, UV2 chromatogram at $\lambda = 450$ nm and UV3 chromatogram at $\lambda = 460$ nm

A.2 Transglycosylation model reactions using acceptor substrates piceid and crocin

A.2.1 Crocin reaction

A.2.2 Piceid reaction

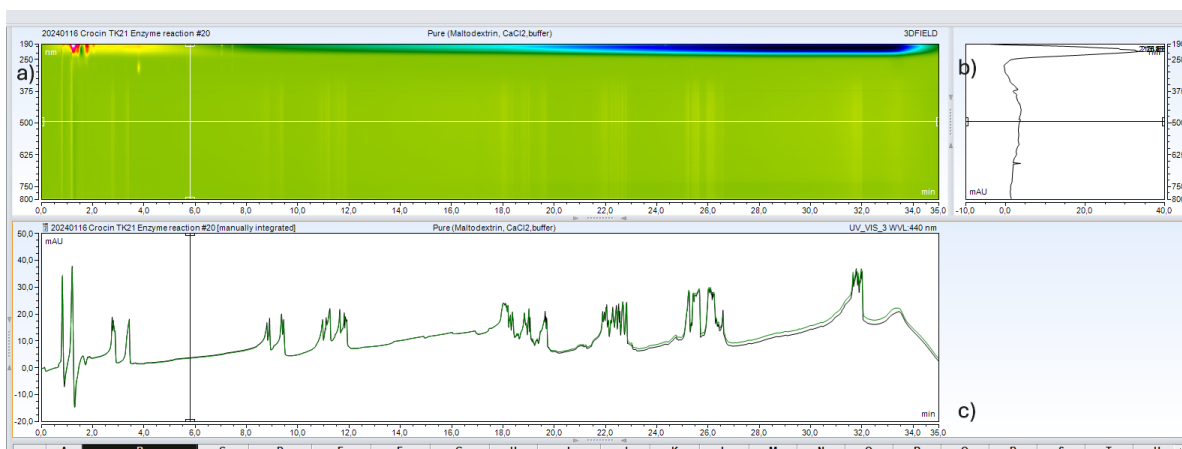


Figure 33: HPLC run of all reaction components except for the enzyme and the acceptor substrate. Shown are (a) the 3D field plot (top left window) where the x-axis represents time in minutes and the y-axis represents wavelength in nm; (b) the spectral plot (top right window) where the x-axis represents wavelength in nm and the y-axis represents the absorbance in mAU and (c) the chromatogram where the x-axis represents the time in min and the y-axis represents the absorbance in mAU

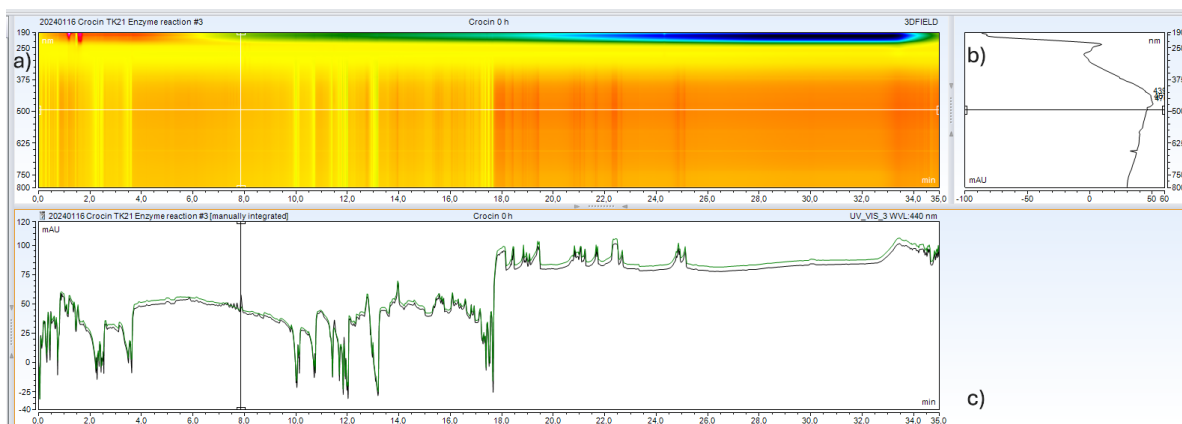


Figure 34: HPLC run of 0 hour sample of crocin transglycosylation reaction. Shown are (a) the 3D field plot (top left window) where the x-axis represents time in minutes and the y-axis represents wavelength in nm; (b) the spectral plot (top right window) where the x-axis represents wavelength in nm and the y-axis represents the absorbance in mAU and (c) the chromatogram where the x-axis represents the time in min and the y-axis represents the absorbance in mAU

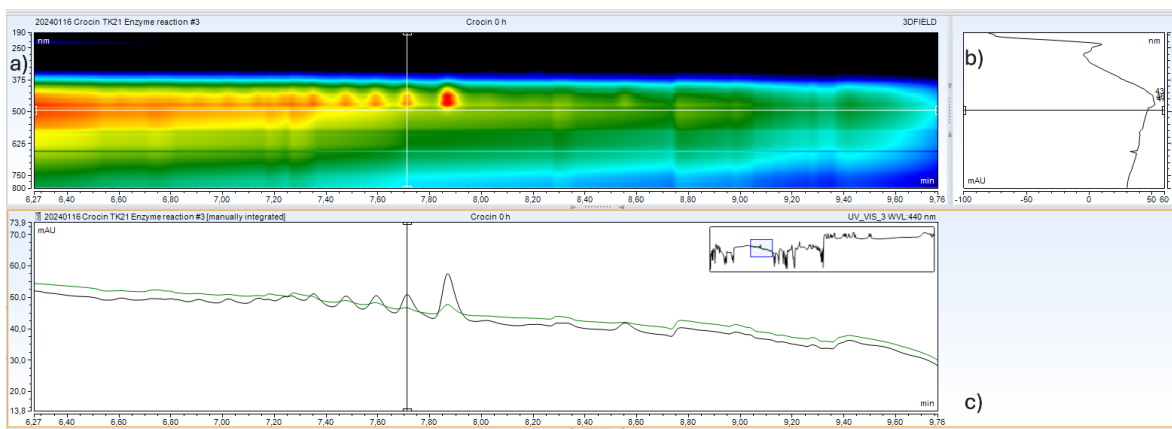


Figure 35: HPLC run of 0 hour sample of crocin transglycosylation reaction, focus on peak 2. Shown are (a) the 3D field plot (top left window) where the x-axis represents time in minutes and the y-axis represents wavelength in nm; (b) the spectral plot (top right window) where the x-axis represents wavelength in nm and the y-axis represents the absorbance in mAU and (c) the chromatogram where the x-axis represents the time in min and the y-axis represents the absorbance in mAU

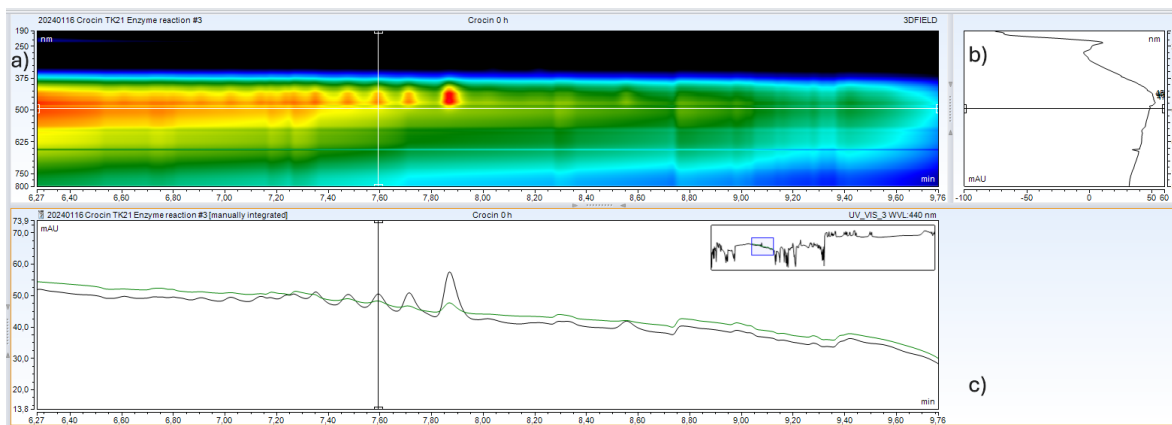


Figure 36: HPLC run of 0 hour sample of crocin transglycosylation reaction, focus on peak 3. Shown are (a) the 3D field plot (top left window) where the x-axis represents time in minutes and the y-axis represents wavelength in nm; (b) the spectral plot (top right window) where the x-axis represents wavelength in nm and the y-axis represents the absorbance in mAU and (c) the chromatogram where the x-axis represents the time in min and the y-axis represents the absorbance in mAU

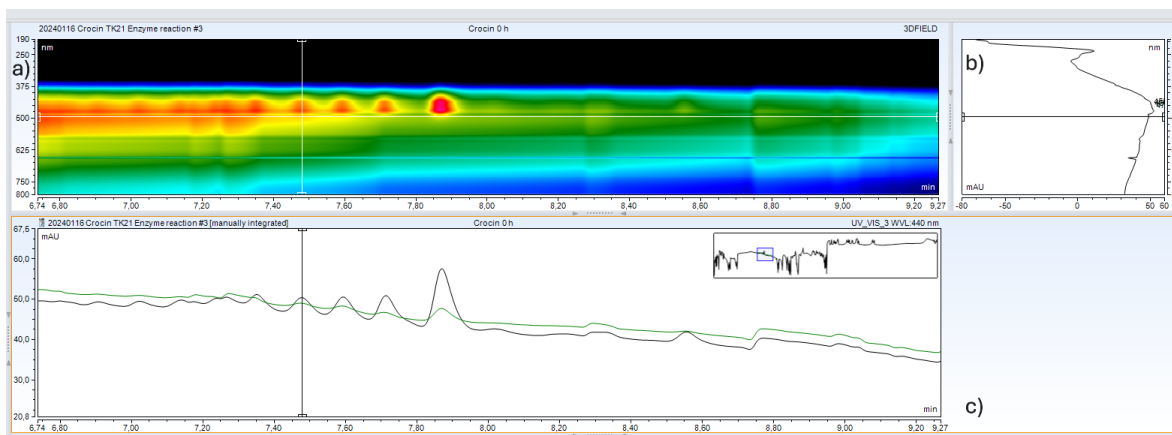


Figure 37: HPLC run of 0 hour sample of crocin transglycosylation reaction, focus on peak 4. Shown are (a) the 3D field plot (top left window) where the x-axis represents time in minutes and the y-axis represents wavelength in nm; (b) the spectral plot (top right window) where the x-axis represents wavelength in nm and the y-axis represents the absorbance in mAU and (c) the chromatogram where the x-axis represents the time in min and the y-axis represents the absorbance in mAU

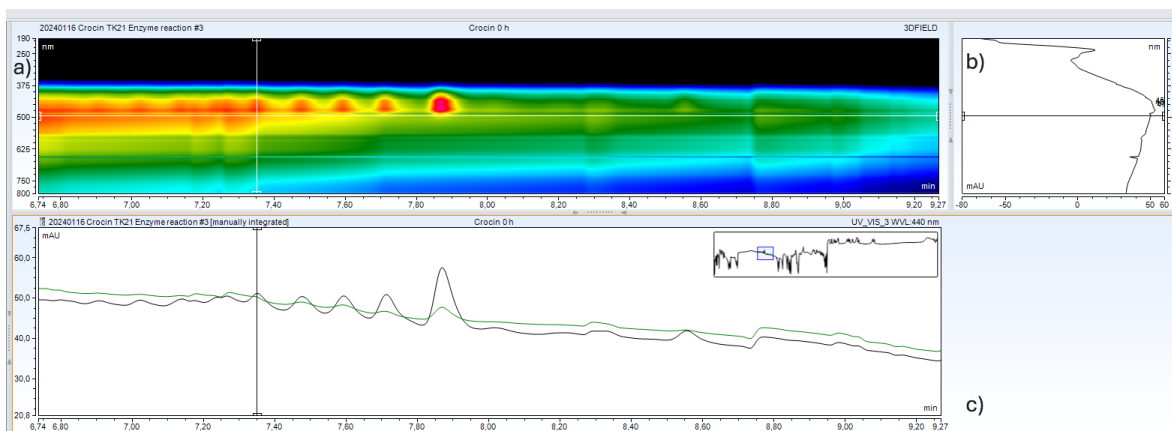


Figure 38: HPLC run of 0 hour sample of crocin transglycosylation reaction, focus on peak 5. Shown are (a) the 3D field plot (top left window) where the x-axis represents time in minutes and the y-axis represents wavelength in nm; (b) the spectral plot (top right window) where the x-axis represents wavelength in nm and the y-axis represents the absorbance in mAU and (c) the chromatogram where the x-axis represents the time in min and the y-axis represents the absorbance in mAU

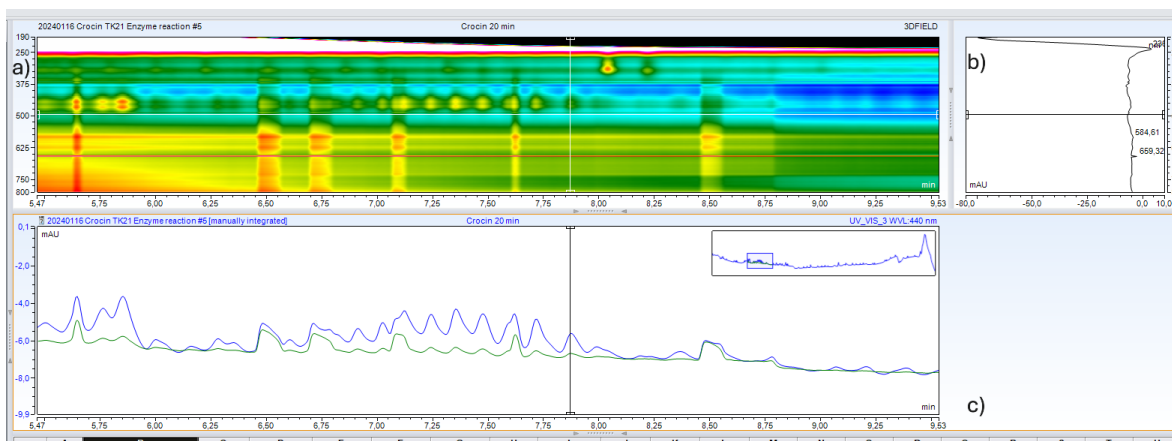


Figure 39: HPLC run of 20 minute sample of crocin transglycosylation reaction, focus on peak 1. Shown are (a) the 3D field plot (top left window) where the x-axis represents time in minutes and the y-axis represents wavelength in nm; (b) the spectral plot (top right window) where the x-axis represents wavelength in nm and the y-axis represents the absorbance in mAU and (c) the chromatogram where the x-axis represents the time in min and the y-axis represents the absorbance in mAU

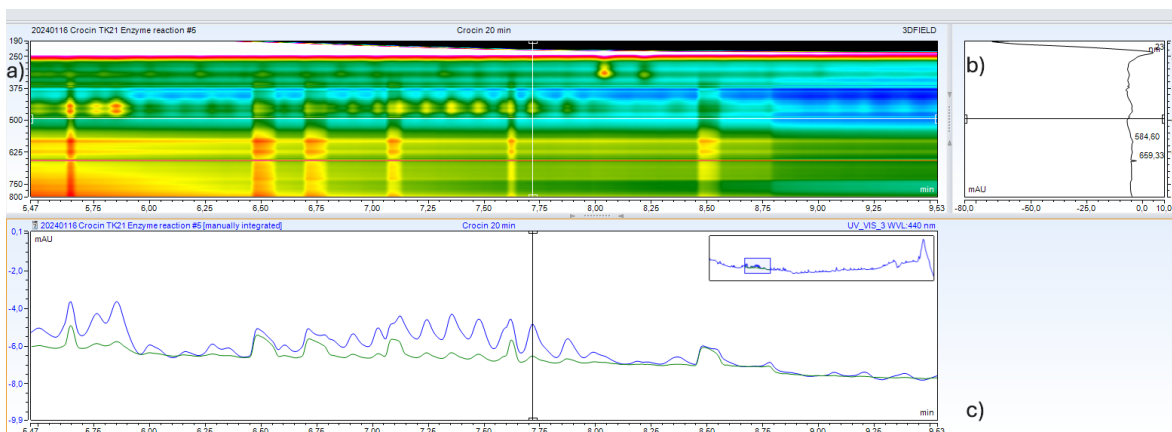


Figure 40: HPLC run of 20 minute sample of crocin transglycosylation reaction, focus on peak 2. Shown are (a) the 3D field plot (top left window) where the x-axis represents time in minutes and the y-axis represents wavelength in nm; (b) the spectral plot (top right window) where the x-axis represents wavelength in nm and the y-axis represents the absorbance in mAU and (c) the chromatogram where the x-axis represents the time in min and the y-axis represents the absorbance in mAU

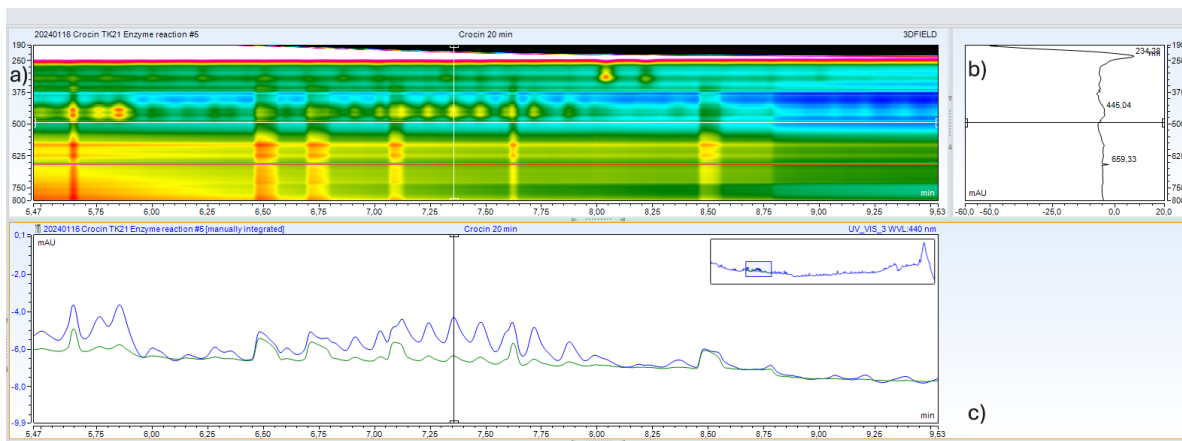


Figure 41: HPLC run of 20 minute sample of crocin transglycosylation reaction, focus on peak 3. Shown are (a) the 3D field plot (top left window) where the x-axis represents time in minutes and the y-axis represents wavelength in nm; (b) the spectral plot (top right window) where the x-axis represents wavelength in nm and the y-axis represents the absorbance in mAU and (c) the chromatogram where the x-axis represents the time in min and the y-axis represents the absorbance in mAU

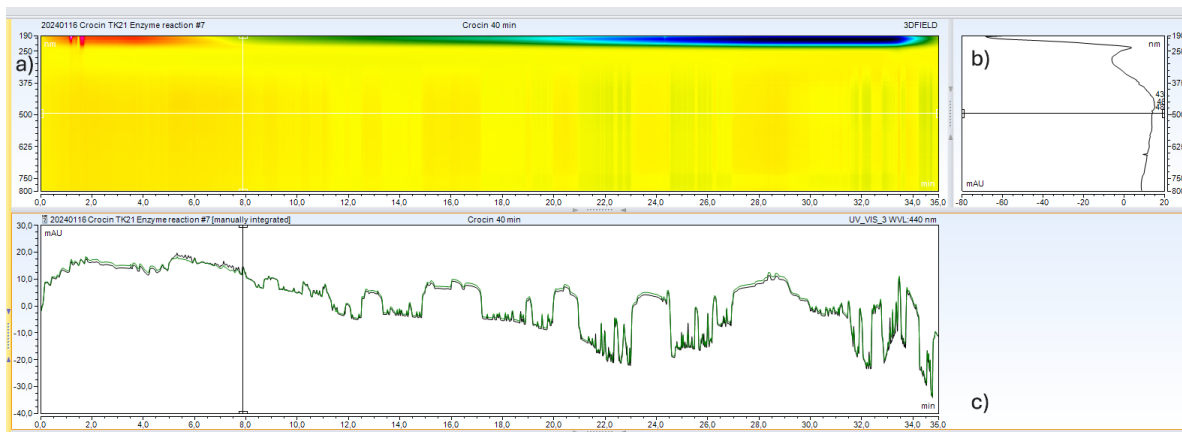


Figure 42: HPLC run of 40 minute sample of crocin transglycosylation reaction. Shown are (a) the 3D field plot (top left window) where the x-axis represents time in minutes and the y-axis represents wavelength in nm; (b) the spectral plot (top right window) where the x-axis represents wavelength in nm and the y-axis represents the absorbance in mAU and (c) the chromatogram where the x-axis represents the time in min and the y-axis represents the absorbance in mAU

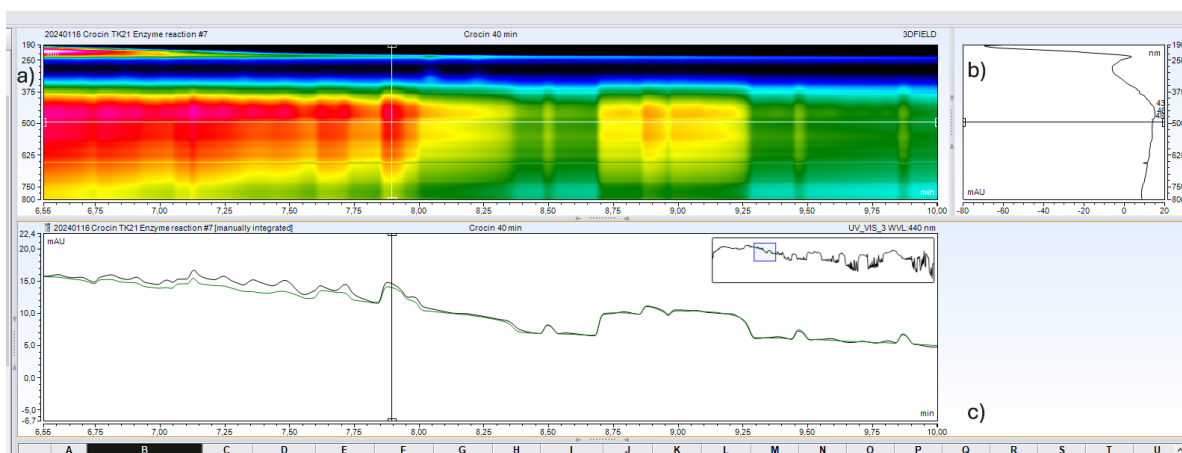


Figure 43: HPLC run of 40 minute sample of crocin transglycosylation reaction, focus on peak 1. Shown are (a) the 3D field plot (top left window) where the x-axis represents time in minutes and the y-axis represents wavelength in nm; (b) the spectral plot (top right window) where the x-axis represents wavelength in nm and the y-axis represents the absorbance in mAU and (c) the chromatogram where the x-axis represents the time in min and the y-axis represents the absorbance in mAU

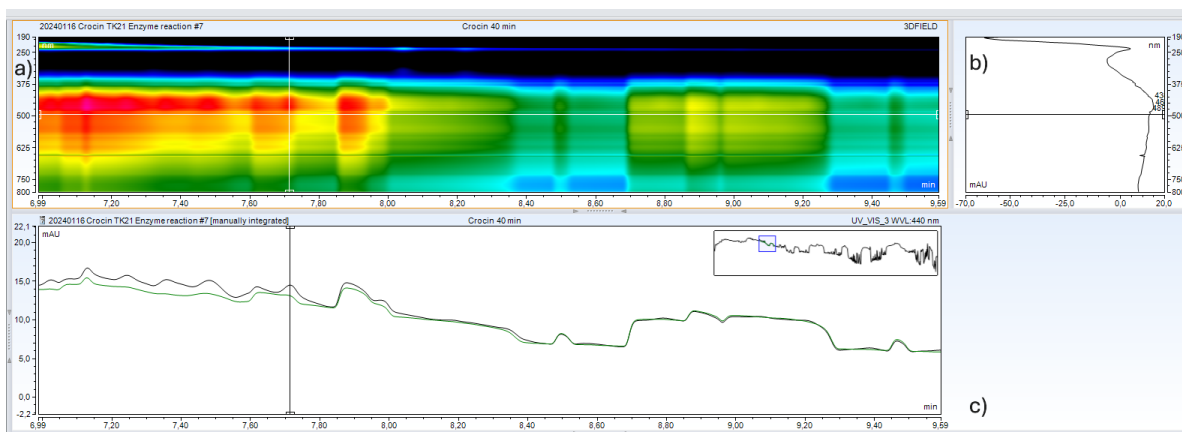


Figure 44: HPLC run of 40 minute sample of crocin transglycosylation reaction, focus on peak 2. Shown are (a) the 3D field plot (top left window) where the x-axis represents time in minutes and the y-axis represents wavelength in nm; (b) the spectral plot (top right window) where the x-axis represents wavelength in nm and the y-axis represents the absorbance in mAU and (c) the chromatogram where the x-axis represents the time in min and the y-axis represents the absorbance in mAU

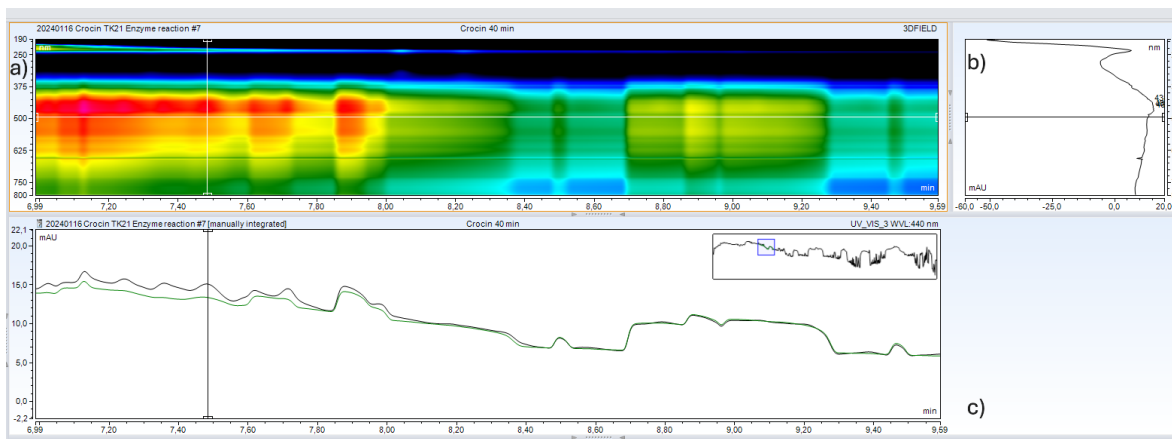


Figure 45: HPLC run of 40 minute sample of crocin transglycosylation reaction, focus on peak 3. Shown are (a) the 3D field plot (top left window) where the x-axis represents time in minutes and the y-axis represents wavelength in nm; (b) the spectral plot (top right window) where the x-axis represents wavelength in nm and the y-axis represents the absorbance in mAU and (c) the chromatogram where the x-axis represents the time in min and the y-axis represents the absorbance in mAU

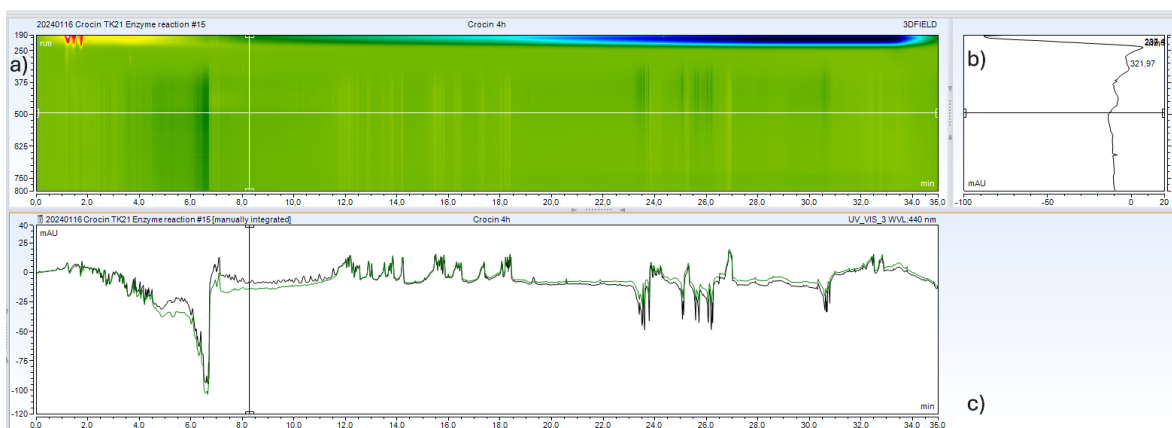


Figure 46: HPLC run of 4 hour sample of crocin transglycosylation reaction. Shown are (a) the 3D field plot (top left window) where the x-axis represents time in minutes and the y-axis represents wavelength in nm; (b) the spectral plot (top right window) where the x-axis represents wavelength in nm and the y-axis represents the absorbance in mAU and (c) the chromatogram where the x-axis represents the time in min and the y-axis represents the absorbance in mAU

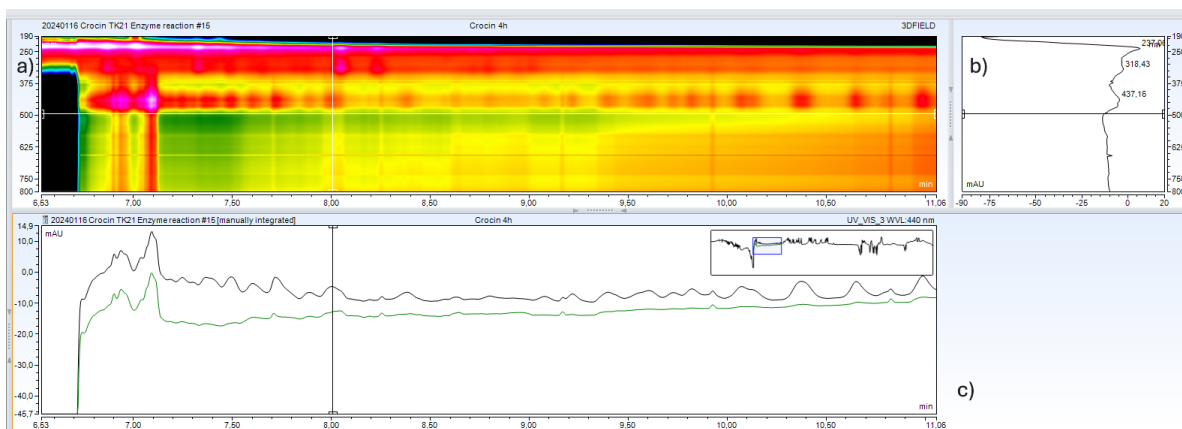


Figure 47: HPLC run of 4 hour sample of crocin transglycosylation reaction, focus on peak 1. Shown are (a) the 3D field plot (top left window) where the x-axis represents time in minutes and the y-axis represents wavelength in nm; (b) the spectral plot (top right window) where the x-axis represents wavelength in nm and the y-axis represents the absorbance in mAU and (c) the chromatogram where the x-axis represents the time in min and the y-axis represents the absorbance in mAU

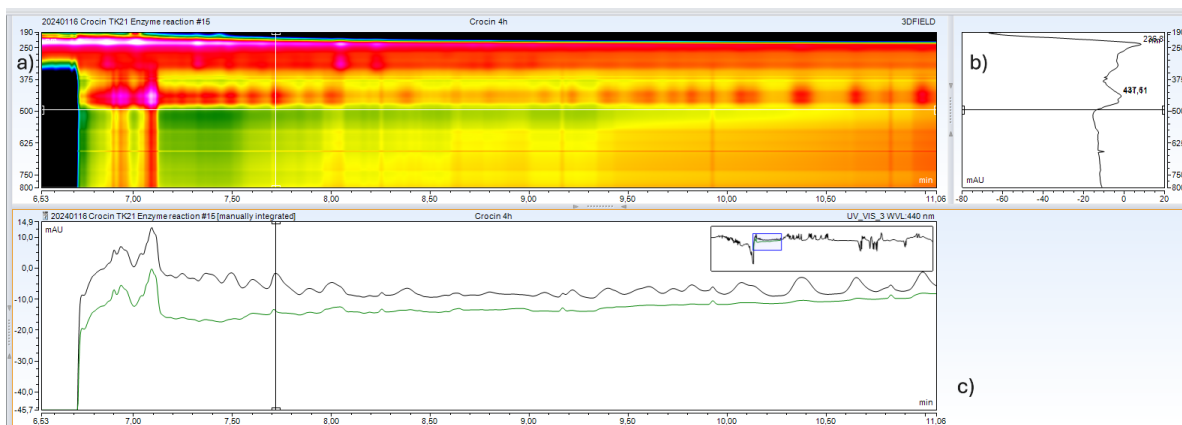


Figure 48: HPLC run of 4 hour sample of crocin transglycosylation reaction, focus on peak 1. Shown are (a) the 3D field plot (top left window) where the x-axis represents time in minutes and the y-axis represents wavelength in nm; (b) the spectral plot (top right window) where the x-axis represents wavelength in nm and the y-axis represents the absorbance in mAU and (c) the chromatogram where the x-axis represents the time in min and the y-axis represents the absorbance in mAU

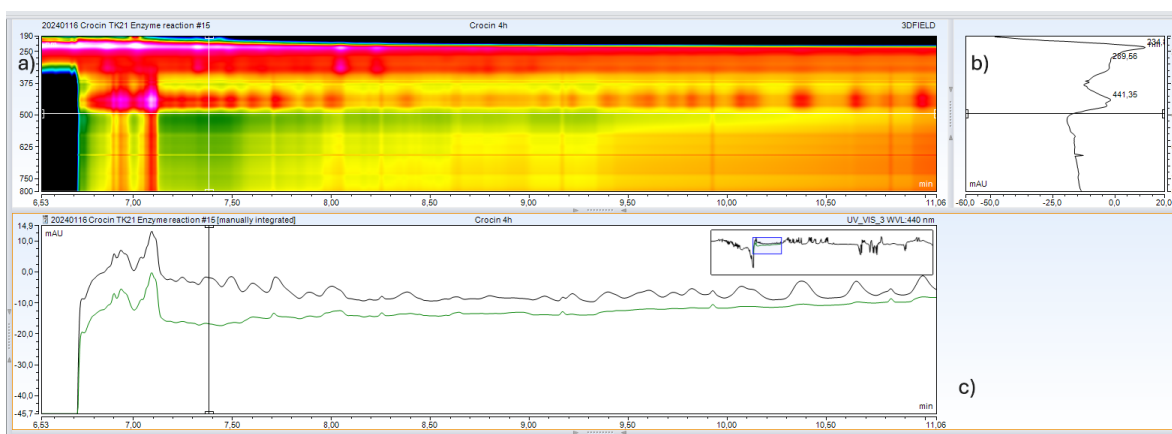


Figure 49: HPLC run of 4 hour sample of crocin transglycosylation reaction, focus on peak 4. Shown are (a) the 3D field plot (top left window) where the x-axis represents time in minutes and the y-axis represents wavelength in nm; (b) the spectral plot (top right window) where the x-axis represents wavelength in nm and the y-axis represents the absorbance in mAU and (c) the chromatogram where the x-axis represents the time in min and the y-axis represents the absorbance in mAU

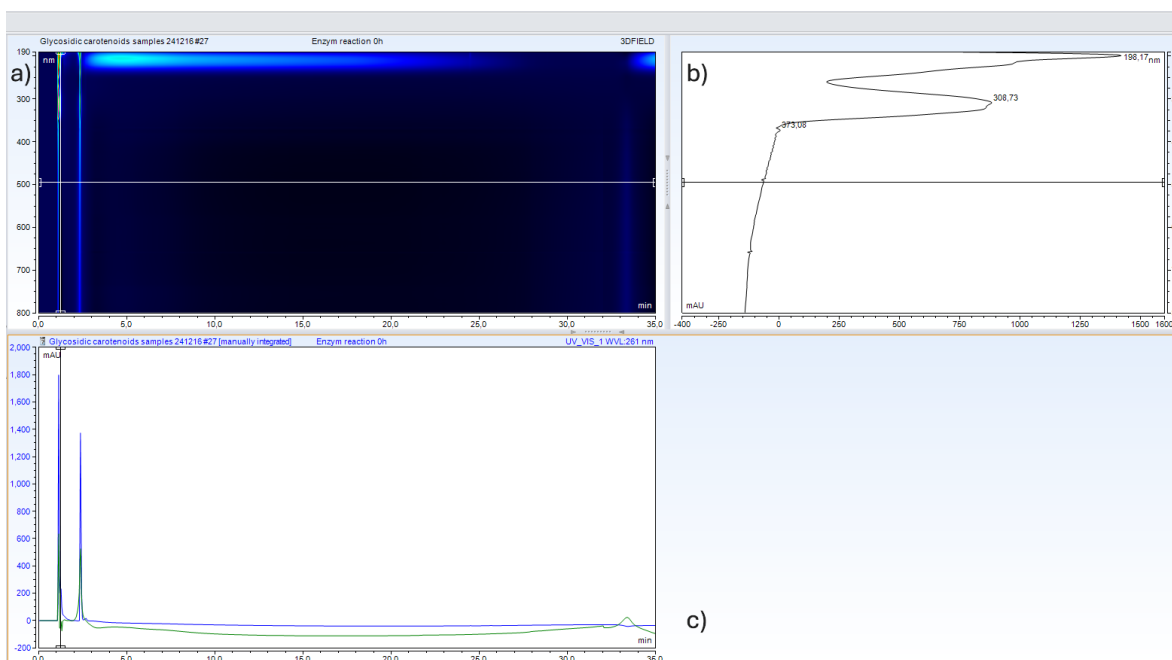


Figure 50: HPLC run of the 0h piceid transglycosylation reaction. Shown are (a) the 3D field plot (top left window) where the x-axis represents time in minutes and the y-axis represents wavelength in nm; (b) the spectral plot (top right window) where the x-axis represents wavelength in nm and the y-axis represents the absorbance in mAU and (c) the chromatogram where the x-axis represents the time in min and the y-axis represents the absorbance in mAU

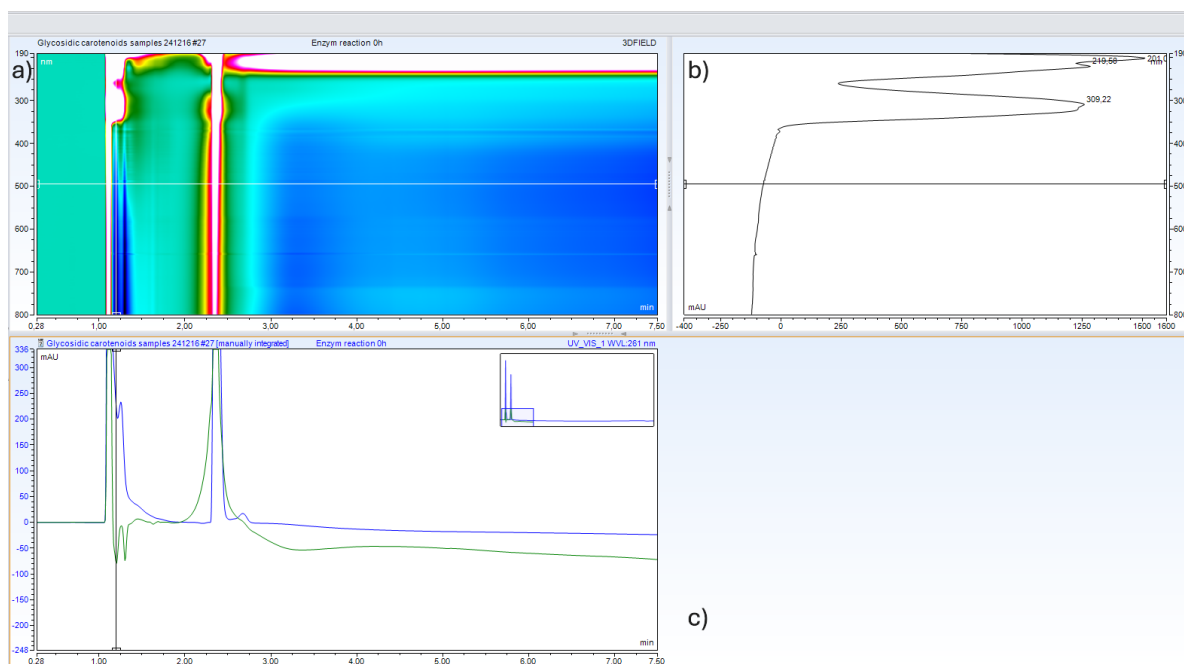


Figure 51: Peak in the in the 0h sample of the piceid transglycosylation reaction. Shown are (a) the 3D field plot (top left window) where the x-axis represents time in minutes and the y-axis represents wavelength in nm; (b) the spectral plot (top right window) where the x-axis represents wavelength in nm and the y-axis represents the absorbance in mAU and (c) the chromatogram where the x-axis represents the time in min and the y-axis represents the absorbance in mAU

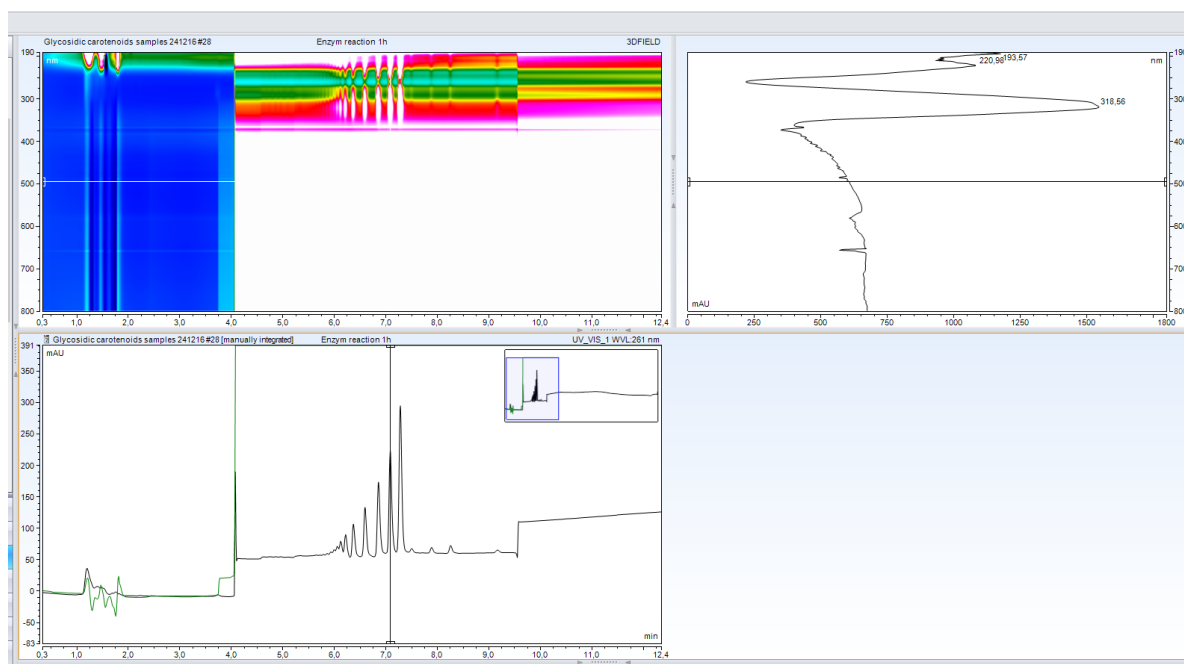


Figure 52: Peak 2 in the in the 1h sample of the piceid transglycosylation reaction. Shown are (a) the 3D field plot (top left window) where the x-axis represents time in minutes and the y-axis represents wavelength in nm; (b) the spectral plot (top right window) where the x-axis represents wavelength in nm and the y-axis represents the absorbance in mAU and (c) the chromatogram where the x-axis represents the time in min and the y-axis represents the absorbance in mAU

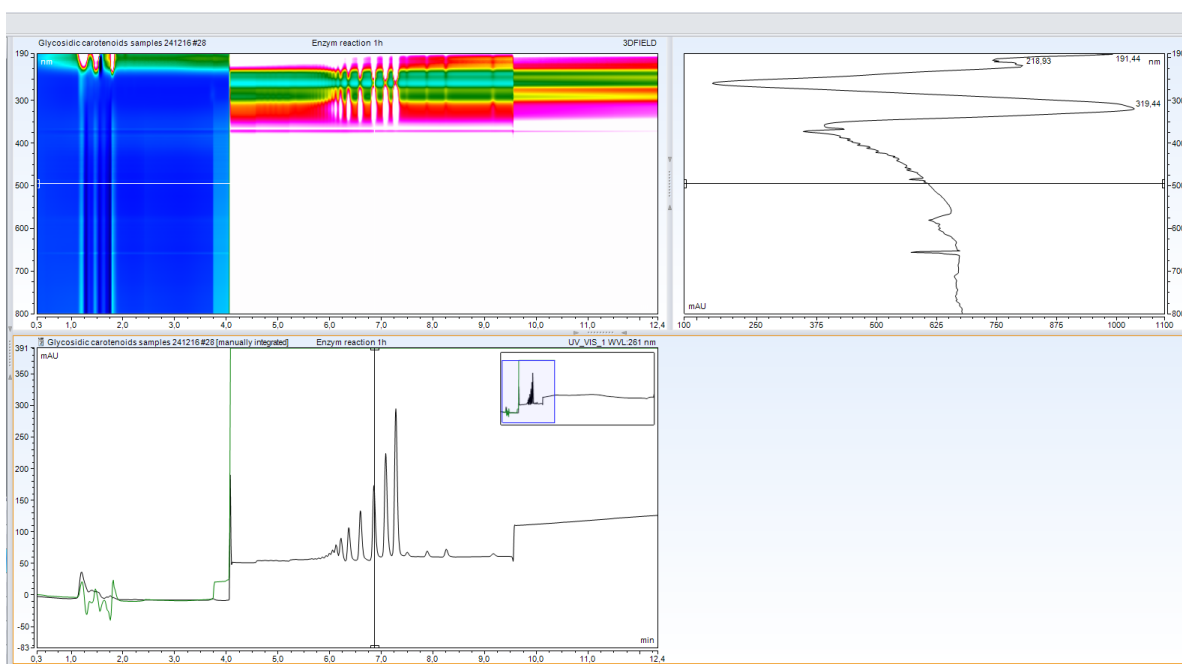


Figure 53: Peak 3 in the in the 1h sample of the piceid transglycosylation reaction. Shown are (a) the 3D field plot (top left window) where the x-axis represents time in minutes and the y-axis represents wavelength in nm; (b) the spectral plot (top right window) where the x-axis represents wavelength in nm and the y-axis represents the absorbance in mAU and (c) the chromatogram where the x-axis represents the time in min and the y-axis represents the absorbance in mAU

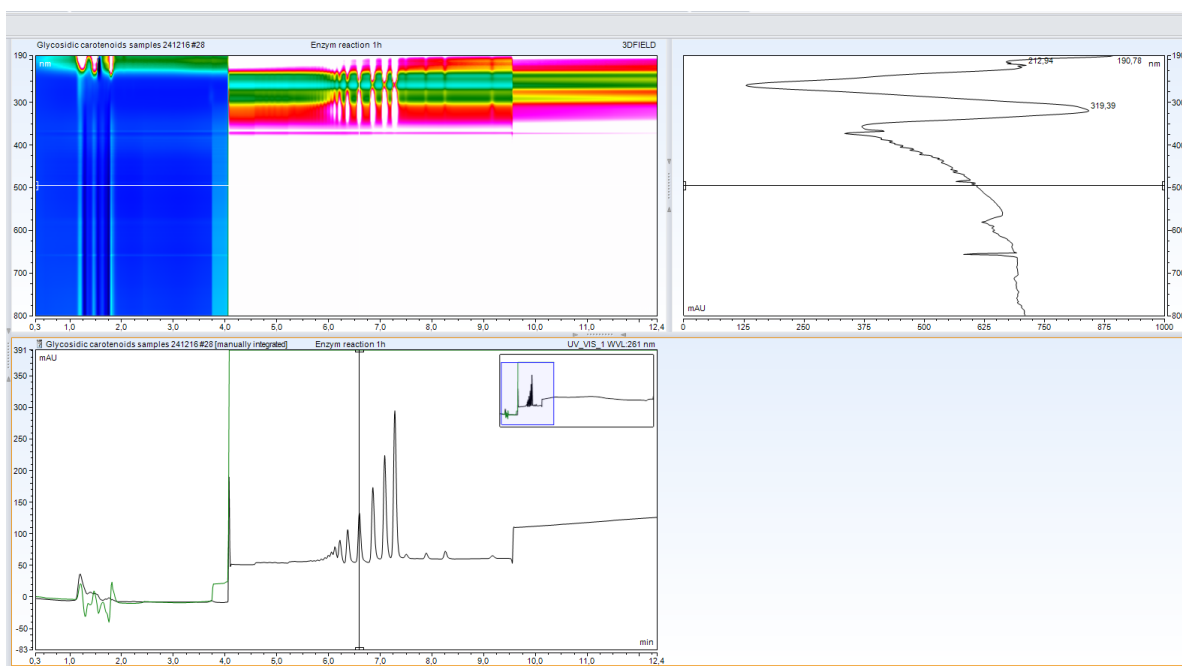


Figure 54: Peak 4 in the in the 1h sample of the piceid transglycosylation reaction. Shown are (a) the 3D field plot (top left window) where the x-axis represents time in minutes and the y-axis represents wavelength in nm; (b) the spectral plot (top right window) where the x-axis represents wavelength in nm and the y-axis represents the absorbance in mAU and (c) the chromatogram where the x-axis represents the time in min and the y-axis represents the absorbance in mAU

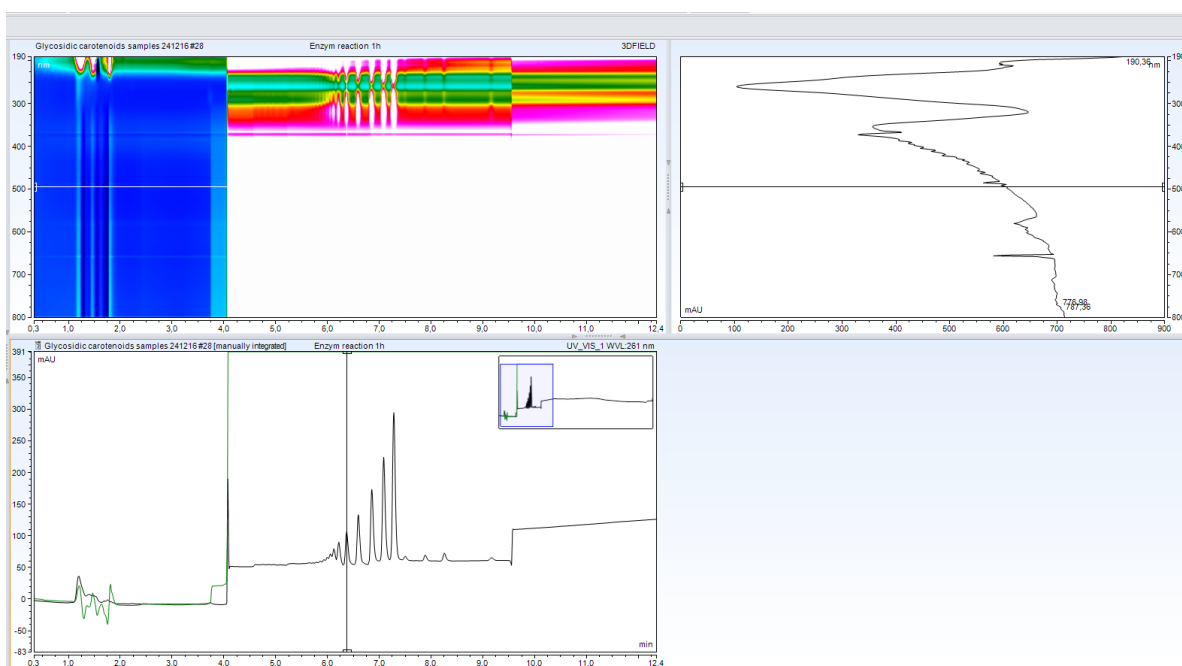


Figure 55: Peak 5 in the in the 1h sample of the piceid transglycosylation reaction. Shown are (a) the 3D field plot (top left window) where the x-axis represents time in minutes and the y-axis represents wavelength in nm; (b) the spectral plot (top right window) where the x-axis represents wavelength in nm and the y-axis represents the absorbance in mAU and (c) the chromatogram where the x-axis represents the time in min and the y-axis represents the absorbance in mAU

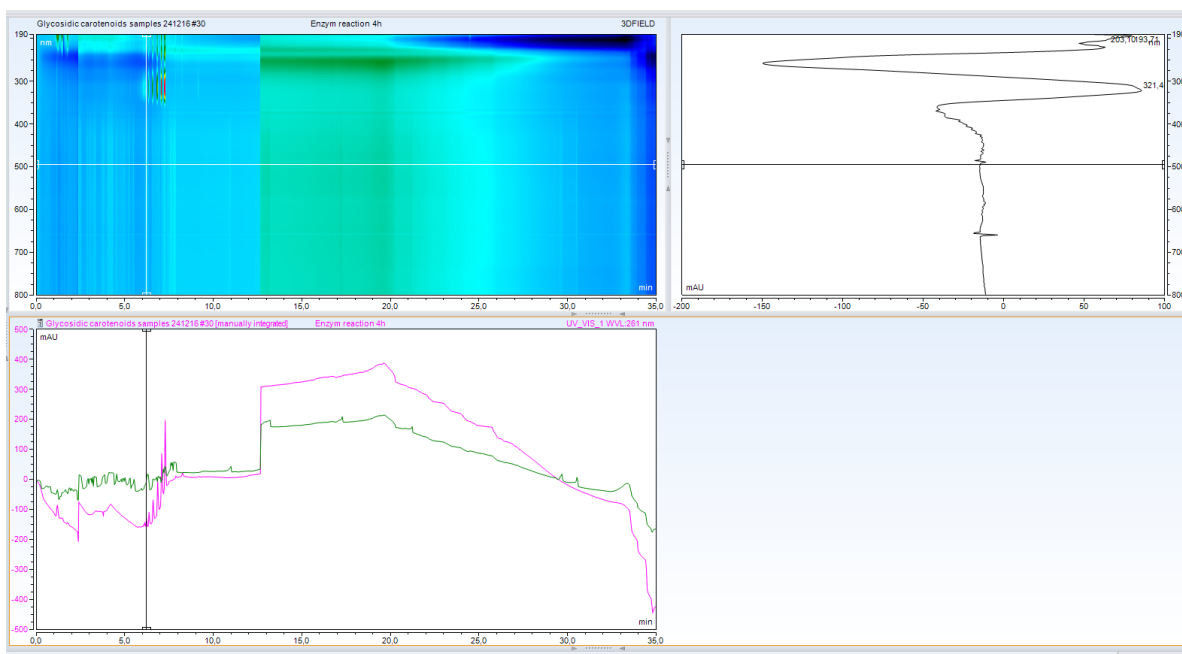


Figure S6: 4h sample of the piceid transglycosylation reaction. Shown are (a) the 3D field plot (top left window) where the x-axis represents time in minutes and the y-axis represents wavelength in nm; (b) the spectral plot (top right window) where the x-axis represents wavelength in nm and the y-axis represents the absorbance in mAU and (c) the chromatogram where the x-axis represents the time in min and the y-axis represents the absorbance in mAU