



SCHOOL of  
GRADUATE STUDIES

EAST TENNESSEE STATE UNIVERSITY

East Tennessee State University  
**Digital Commons @ East  
Tennessee State University**

---

Electronic Theses and Dissertations

Student Works

---

12-2016

# Structural and Functional Analysis of Grapefruit Flavonol-Specific-3-O-GT Mutant P145T

Sangam Kandel Mr

*East Tennessee State University*

Follow this and additional works at: <https://dc.etsu.edu/etd>



Part of the [Biology Commons](#)

---

## Recommended Citation

Kandel, Sangam Mr, "Structural and Functional Analysis of Grapefruit Flavonol-Specific-3-O-GT Mutant P145T" (2016). *Electronic Theses and Dissertations*. Paper 3150. <https://dc.etsu.edu/etd/3150>

This Thesis - Open Access is brought to you for free and open access by the Student Works at Digital Commons @ East Tennessee State University. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of Digital Commons @ East Tennessee State University. For more information, please contact [digilib@etsu.edu](mailto:digilib@etsu.edu).

# Structural and Functional Analysis of Grapefruit Flavonol-Specific-3-*O*-GT Mutant P145T

---

A thesis

presented to

The faculty of the Department of Biological Sciences

East Tennessee State University

In partial fulfillment

of the requirements for the degree

Master of Science in Biology

---

by

Sangam Kandel

December 2016

---

Cecilia A. McIntosh, PhD, Chair

Ranjan Chakraborty, PhD

Shivakumar P. Devaiah, PhD

Keywords: Regiospecificity, Mutation, Homology modeling, Molecular docking

## ABSTRACT

Structural and Functional Analysis of Grapefruit Flavonol-Specific-3-*O*-GT Mutant P145T

by

Sangam Kandel

This research is focused on the study of the effect of mutating proline 145 to threonine on the substrate and regiospecificity of flavonol specific 3-*O*-glucosyltransferase (Cp3GT). While the mutant P145T enzyme did not glucosylate anthocyanidins, it did glucosylate flavanones and flavones in addition to retaining activity with flavonols. HPLC was used for product identification and showed mutant P145T glucosylated naringenin at the 7-OH position forming naringenin-7-*O*-glucoside and flavonols at the 3-OH position. Homology modeling and docking was done to predict the acceptor substrate recognition pattern and models were validated by experimental results. In other related work, a thrombin cleavage site was inserted into wild type Cp3GT and recombinant P145T enzyme between the enzyme and the C-myc tags in order to be able to cleave off tags. This provides the tool needed for future efforts to crystallize these proteins for structural determination.

Copyright 2016 by Sangam Kandel  
All Rights Reserved

## ACKNOWLEDGEMENTS

I would like to express my sincere thanks and gratitude to my research advisor and committee chair Dr. Cecilia McIntosh for granting me an opportunity to work at her lab. This thesis would not have been possible without her guidance, motivation, and support. Her inspiration and motivation helped me improve my knowledge, research skills, and develop keys to be an independent researcher. I would like to express my sincere thanks to Dr. Shivakumar Devaiah who provided me hands-on training on basic molecular biology techniques and his willingness to help me whenever needed. My sincere thanks to Dr. Ranjan Chakraborty for his support and valuable suggestions and recommendations during my research. I am greatly indebted to my committee members for their guidelines and support throughout my MS program.

I would like to thank School of Graduate Studies, ETSU for granting me admission as well as providing me a Graduate School Research Grant to help complete my MS thesis, and the Department of Biological Science for accepting me for MS program. I am thankful to all past and present members of the Dr. McIntosh lab. I am also thankful to all the faculties, staffs as well as my colleagues for their support and encouragement. Lastly, I would like to thank my parents and my sisters for their blessings, support and encouragement.

## TABLE OF CONTENTS

	Page
ABSTRACT.....	2
ACKNOWLEDGEMENTS.....	4
LIST OF TABLES.....	8
LIST OF FIGURES .....	9
Chapter	
1. INTRODUCTION .....	10
Different Classes of Secondary Metabolites .....	11
Nitrogen Containing Metabolites .....	11
Terpenoids.....	12
Phenolics .....	12
Flavonoids .....	13
Flavonoid Biosynthetic Pathway.....	14
Importance of Flavonoids in Plants.....	18
Importance of Flavonoids in Mammals.....	23
Glycosylation.....	26
Plant Secondary Product Glucosyltransferase.....	27
Classification and Structure of GT Enzymes .....	28
Crystal Structures of Family1 GTs.....	31
The Model Plant; <i>Citrus paradisi</i> .....	35
Flavonol Specific 3-O-Glucosyltransferase from Grapefruit.....	36
Hypotheses .....	39
2. STRUCTURAL AND FUNCTIONAL ANALYSIS OF GRAPEFRUIT FLAVONOL-SPECIFIC-3-O-GT MUTANT P145T .....	40
Abstract .....	41
1.Introduction .....	42
2.Materials and methods.....	48
2.1. Reagents.....	48
2.2. Scale up expression of recombinant wild type Cp3GT and mutant P145T.....	49
2.3. Protein extraction and purification.....	50
2.4. Glucosyltrasnferase activity.....	51

2.5. Product identification.....	52
2.6. Homology modeling and docking.....	53
2.7. Insertion of thrombin site.....	54
2.7.1. Primer design .....	54
2.7.2. Site-directed mutagenesis .....	54
2.7.3. Heterologous expression.....	55
3. Results and discussion.....	56
3.1. Acceptor substrate specificity .....	56
3.2. Product identification .....	57
3.3. Enzyme characterization .....	62
3.4. Structural similarities and differences between wild type Cp3GTand P145T. ....	66
3.5. Molecular docking with acceptor substrates and protein-ligand interaction.....	73
3.6. Insertion of thrombin site and heterologous expression.....	81
4. Conclusions .....	87
Acknowledgements.....	88
References .....	89
3. SUMMARY AND DIRECTIONS FOR FUTURE RESEARCH .....	95
REFERENCES .....	100
APPENDICES .....	123
APPENDIX A: RECIPES .....	123
APPENDIX B: METHODS .....	137
SDS-PAGE.....	137
Western blotting .....	137
Insertion of thrombin cleavage site .....	138
Primer design.....	138
Site-directed mutagenesis.....	139
DpnI digestion of parental DNA of the site-direct mutagenesis PCR product .....	140
Transformation into E.coli competent cells .....	140
PCR screening of positive transformants .....	141
Re-streaking of positive transformants.....	142
Miniprep plasmid extraction for sequencing.....	142
Midiprep plasmid extraction for transformation into yeast.....	144

Linearization using SacI for transformation in to yeast .....	146
Phenol-chloroform extraction of linearized plasmid DNA .....	147
Competent cell preparation for transformation of DNA into yeast.....	148
APPENDIX C: SUPPLEMENTARY FIGURES.....	150
VITA.....	152



## LIST OF TABLES

Table	Page
1.1 UGTs with crystal structures. ....	32
2. 1 Activity Screening of Wild Type Cp3GT and Mutant P145T. ....	58
B. 1 Primers for insertion of the thrombin site. ....	138
B. 2 PCR conditions for insertion of thrombin site using site directed mutagenesis. ....	139
B. 3 AOX primer sequence used for colony PCR ....	141
B. 4 Colony PCR reaction conditions to identify the positive transformants. ....	141
B. 5 Miniprep Plasmid Extraction. ....	144
B. 6 Samples sent for sequencing. ....	144
B. 7 Midiprep plasmid extraction and purification. ....	146
B. 8 Phenol chloroform extraction. ....	148

## LIST OF FIGURES

Figure	Page
1. 1 The generic structure of a flavonoid .....	13
1. 2 Structure and examples of some common classes of flavonoids .....	15
1. 3 Flavonoid biosynthetic pathway. ....	16
1. 4 Inverting and retaining GTs. ....	29
1. 5 Structure of GT-A and GT-B folds of GTs.....	30
1. 6 Schematic of a Cp3GT catalyzed reaction.....	37
2. 1 Generic structure of major classes of flavonoids and substitution patterns.....	43
2. 2 Structure of Cp3GT modelled with VvGT. ....	46
2. 3 Multiple sequence alignment of Cp3GT and VvGT using Bionedit. ....	47
2. 4 Reaction product identification of mutant P145T with acceptor quercetin. ....	59
2. 5 Reaction product identification of mutant P145T with acceptor kaempferol.....	60
2. 6 Reaction product identification with acceptor naringenin. ....	61
2. 7 Time course study with different concentration of enzyme with quercetin.....	63
2. 8 Time course study with different concentration of enzyme with kaempferol. ....	64
2. 9 Effect of temperature on enzyme activity with quercetin.....	65
2. 10 Effect of temperature on enzyme activity with kaempferol.....	65
2. 11 Homology modeling of mutant P145T with VvGT.....	67
2. 12 Structural similarities and differences. ....	69
2. 13 Superimposition of VvGT and Cp3GT.....	70
2. 14 Docking of wild type Cp3GT and mutant P145T with naringenin.....	71
2. 15 Docking of wild type Cp3GT and mutant P145T with quercetin.....	71
2. 16 Docking of wild type Cp3GT and mutant P145T with kaempferol.....	72
2. 17 Molecular docking of wild type Cp3GT and mutant P145T with quercetin. ....	75
2. 18 Molecular docking of wild type Cp3GT and mutant P145T with naringenin. ....	77
2. 19 Molecular docking of wild type Cp3GT and mutant P145T with kaempferol. ....	79
2. 20 Colony PCR of positive transformants having thrombin site. ....	81
2. 21 Colony PCR of positive transformants of <i>E.coli</i> . having thrombin cleavage site. ....	82
2. 22 Sequencing of mutant P145T with thrombin site. ....	84
2. 23 Sequencing of wild type Cp3GT with thrombin cleavage site. ....	86
2. 24 Digestion of plasmid having thrombin site with SacI.....	86
C. 1 Multiple sequence alignment of wild type Cp3GT and VvGT using Bionedit.....	150
C. 2 Electrophoresis of site directed mutagenesis PCR product.....	151
C. 3 DpnI digestion of site-directed PCR product to insert thrombin site.....	151

## CHAPTER 1

### INTRODUCTION

Plants produce many compounds that are important for their growth and development. These products are a diverse group of chemical compounds that are produced by plants via various biosynthetic pathways. These compounds have been an important source of pharmaceuticals, insecticides, flavorings, fragrances and food colorants (Kieran et al. 1997, Newman and Gordon 2016 and ref. therein). Primary metabolites have important roles in the growth and development of plants (Ashihara et al. 2006). These include carbohydrates, proteins, lipids, phytosterols, acyl lipids, nucleotides, amino acids, and organic acids (Croteau et al. 2000 and ref. therein; Ashihara et al. 2006). In addition to these primary metabolites, plants synthesize low molecular weight compounds called secondary metabolites that have specific functions in the plants in which they are synthesized. Because humans are totally dependent on plants for food, plant natural products and their effect in human health has been of great interest especially the metabolic pathways and the modification of biosynthesis in order to synthesize novel compounds that can be used for human benefit (McIntosh and Owens 2016 and ref. therein).

Several decades ago, plant secondary metabolites were considered just as waste products but more recent research has determined roles of many of the metabolites. Secondary metabolites have varied chemical structures and many different biological and ecological functions. They are produced from unique biosynthetic pathways from primary metabolites and intermediates. It has been found that secondary metabolites are important for the survival of plants as well as their reproductive fitness (Wink 2003; Tiwari et al. 2016 and ref. therein). Secondary metabolites have a key role in defense against herbivores, pests, and pathogens (Bennet and Wallsgrove 1994 and

ref. therein), act as insect and animal attractants for pollination and seed dispersal (Osbourn and Lanzetti, 2009 and ref. therein), as allelopathic agents (Weir et al. 2004), protection from harmful UV (Brouillard and Dangles; 1993 Li et al. 1993), and nitrogen fixation in root nodules of legumes (Peters et al. 1986; Treuuter 2005, and ref. therein).

Similarly, flavonoids have medicinal value in human health including anti-fungal (Cushnie and Lamb 2005), anti-viral (Hanasaki et al. 1994; Cushnie and Lamb 2005), anti-bacterial (Cushnie and Lamb 2005), anti-cancer (reviewed in Benavente et al. 1997), antiallergic (Hope et al. 1983; Hanasaki et al. 1994; Middleton and Kandaswami 1994) and vasodilatory (Duarte et al. 1993) activities. In addition, flavonoids are found to inhibit lipid peroxidation (Salvayre et al. 1987; Ratty and Das 1998), platelet aggregation (Tzeng et al. 1991), and many other different pharmaceutical properties.

### Different Classes of Secondary Metabolites

Secondary metabolites are classified into different classes based on their chemical structure and composition, metabolic pathway by which they are synthesized, and/or their stability and solubility in different solvents (Owens and McIntosh 2011; Kabera et al. 2014). There are many different classes of secondary products the majority of which include phenolics, nitrogen containing compounds, and terpenoids (Harborne 1991).

#### Nitrogen Containing Metabolites

Alkaloids are the largest group of nitrogen containing compounds (Harborne 1991). These are the organic bases having a nitrogen atom as a part of their structure and usually as a part of their carbon cyclic system (Ziegler and Facchini 2008 and ref. therein; Cushnie et al.

2014 and ref. therein). Common examples of alkaloids include morphine, caffeine, and cocaine. Other classes of nitrogen containing metabolites include non-protein amino acids, cyanogenic glycosides, and glucosinolates. Because the supply of nitrogen may be limited in plants, these classes of secondary metabolites may be found in lower amounts or they may have restricted distribution in plants even in the legumes that fix atmospheric nitrogen (Harborne 1991; Cushnie et al. 2014 and ref. therein).

### Terpenoids

Terpenoids are the most diverse and the largest family of plant secondary products (Harborne 1991). Terpenoids are synthesized from isopentenyl diphosphate and dimethylallyl diphosphate (Eisenreich et al. 2001 and ref. therein). Terpenoids are formed from acetyl coenzyme A via the mevalonic acid pathway (Harborne 1991), and are derived by the fusion of C-5 units based on isopentane skeleton (Langenheim 1994). However, the discovery of a non-mevalonate (non-MVA) pathway changed the conventional concept of biosynthesis of terpenoids (Dubey et al. 2003). The non-mevalonate pathway involves biosynthesis of hemi, mono, sesqui and di-terpenes, as well as the phytol chain of chlorophyll and carotenoids (Dubey et al. 2003). Thermal decomposition of terpenoids produce the alkene gas isoprene and polymerization of these monomers under suitable chemical conditions can produce numerous terpenoid skeletons. Because of this, terpenoids are also called isoprenoids.

### Phenolics

Phenolics are derived from the pentose phosphate, shikimate, and phenylpropanoid pathways (Harborne 1993 and ref. therein; Randhir et al. 2004). These compounds have aromatic structures with one or more hydroxyl groups. The majority are polyphenols in which the

hydroxyl groups are methyl or glycosyl substituted (Harborne 1991). Some common examples of phenolics include flavonoids, phenylpropanoids, xanthenes, coumarins, lignans, phenols, stilbenoids, tannins, and quinones. Over 10,000 flavonoids exhibiting structural diversity have been identified in higher plants (Ferrer et al. 2008; Buer 2010 and ref. therein; Weston and Mathesius 2013 and ref. therein).

### Flavonoids

Flavonoids are polyphenolic low molecular weight compounds, having two aromatic rings, A and B, containing 6- carbon atoms that are linked by a 3-carbon chain that forms a closed ring or the C-ring (heterocyclic ring) containing oxygen (Fig. 1.1) (Harborne 1993 and ref. therein).

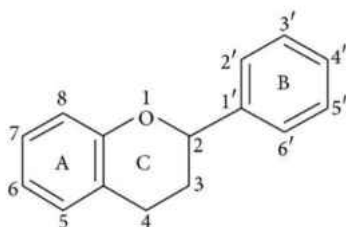


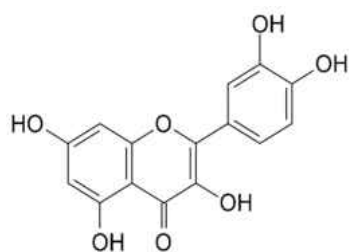
Fig. 1. 1 The generic structure of a flavonoid

The chemical structure and orientation of various substituents determines the biochemical activities of flavonoids and their metabolites (Cody 1987; Harborne 1993). Flavonoids are ubiquitous in plants and have diverse chemical structure and properties (Harborne 1993). The structure of flavonoids varies widely with different substitutions such as hydrogenation, hydroxylation, methylation, malonylation, sulphation, methoxylation, prenylation, and glycosylation (Harborne 1986; Harborne 1988; Dixon and Pasinetti 2010; Owens and McIntosh

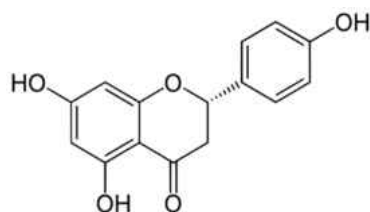
2011). Many flavonoids occur naturally as flavonoid glycosides (Kuhnau 1975; Havsteen 1983; Harborne 1993 and ref. therein; Owens and McIntosh 2009) and carbohydrate substitutions such as D-glucosides, L-rhamnoside, arabinosides, and galactosides (Harborne 1993 and ref. therein). Flavonoids are classified into different sub-groups based on the oxidation of C-15 ring system (Owens and McIntosh 2011 and ref. therein). Common flavonoid classes are: aurone, chalcone, flavanone, flavone, dihydroflavanol, flavonol, isoflavone, leucocyanidin, and anthocyanidin. Flavonoids possess a close structural and chemical interrelationship that reflect a close relationship in the biosynthetic pathways (Fig. 1.2) (Owens and McIntosh 2011 and ref. therein).

### Flavonoid Biosynthetic Pathway

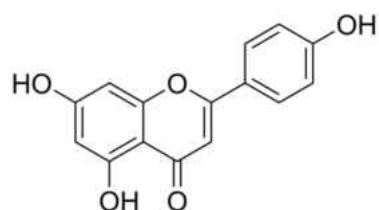
Biosynthesis of flavonoids as well as other numerous plant secondary metabolites starts from the phenylpropanoid pathway that uses phenylalanine synthesized via the shikimate pathway and this has been recently reviewed (Fig 1.3; Owens and McIntosh 2011). Almost all of the enzymes of the phenylpropanoid pathway have been biochemically characterized and some of the proteins have been resolved structurally by crystallization (Yu et al. 2006). Phenylalanine ammonia lyase (PAL; EC 4.3.1.5), which was first characterized from *Herdeum vulgare* L. var. Aravat (barley), converts phenylalaline to trans-cinnamate (Koukol and Conn 1961). PAL catalyzes the elimination of ammonia and the pro-3S hydrogen to generate trans-cinnamate (Owens and McIntosh 2011). Cinnamate 4-hydroxylase (C4H; EC 1.14.13.11) p-hydroxylates trans-cinnamate to produce 4-coumarate. This enzyme requires molecular oxygen and NADPH as electron donor for its activity (Potts et al. 1974; Russell 1971). The 4-coumarate is then activated by 4-coumarate: CoA ligase (4CL; EC 6.2.1.12) to form 4-coumaroyl-CoA through the



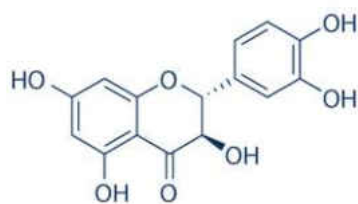
Flavanol (Quercetin)



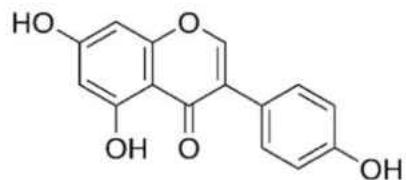
Flavanone (Naringenin)



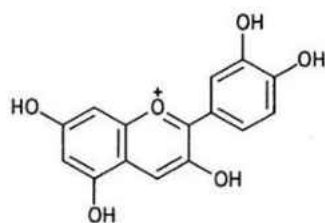
Flavone (Apigenin)



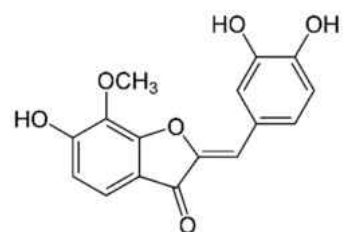
Dihydroflavanol (Dihydroquercetin)



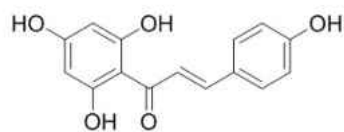
Isoflavone (Genistein)



Anthocyanidin (Cyanidin)



Aurone (Leptosidin)



Chalcone (Naringenin chalcone)

Fig. 1. 2 Structure and examples of some common classes of flavonoids



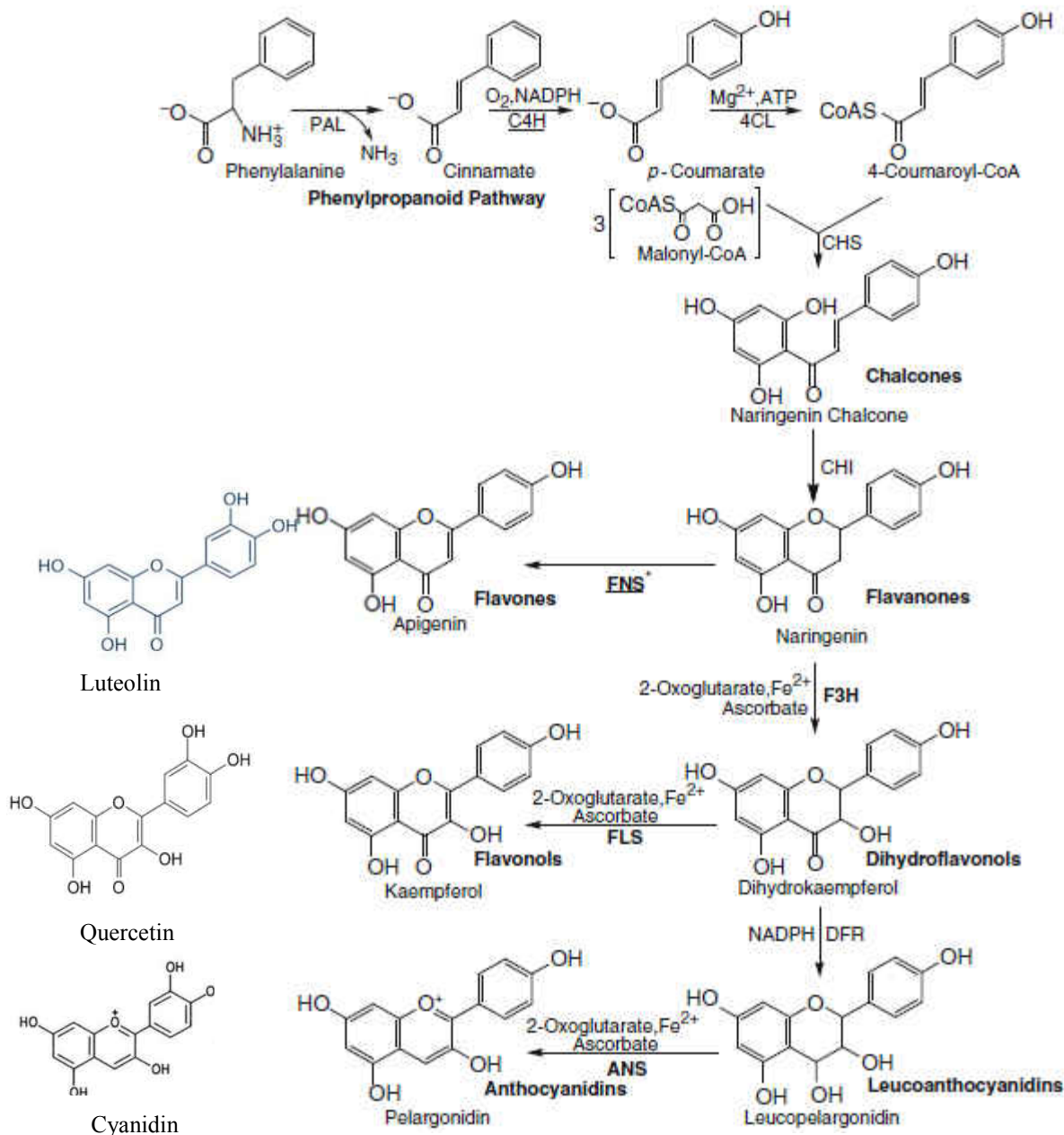


Fig. 1. 3 Flavonoid biosynthetic pathway. PAL, phenylalanine ammonia-lyase; 4CL, 4-coumarate: CoA ligase; CHS, Chalcone synthase; DFR, dihydroflavonol 4-reductase; CHI, chalcone isomerase; FSI/FS2, flavone synthase; IFR, isoflavone reductase; C4H, cinnamate-4-hydroxylase; I2'H, isoflavone 2'-hydroxylase; FNS, Flavone synthase; F3'5'H, flavonoid 3'5' hydroxylase; F3H, flavanone 3-hydroxylase; F3'H, flavonoid 3' hydroxylase; ANS, anthocyanidin synthase. Hydroxytransferase can add further OH groups. (Adapted from Owens and McIntosh 2011).

formation of CoA esters.  $Mg^{2+}$  and ATP are required co-factors for enzyme activity (Ragg et al. 1961; Hahlbrock and Grisebach 1970).

The 4-coumaroyl-CoA is a key intermediate in phenylpropanoid metabolism in higher plants (Schroder et al. 1979; Hahlbrock and Scheel 1989) serving as a branch point substrate for biosynthesis of different metabolites, such as flavonoids, lignins, and various cinnamate esters (Ragg et al. 1961).

The next step in the biosynthesis of flavonoids is the production of chalcone which serves as the first committed step for production of a large number of flavonoids having a common C-15 backbone (Heller and Hahlbrock 1980; Owens and McIntosh 2011 and ref. therein). Chalcone synthase (CHS; EC 2.3.1.74) catalyzes the condensation of three molecules of malonyl-CoA with 4-coumaroyl Co-A to form chalcone (Heller and Forkmann 1988). Stereospecific cyclization of these bicyclic chalcones by chalcone isomerase (CHI; EC 5.5.1.6) forms the tricyclic flavanone which is a pre-requisite for the synthesis of the majority of downstream flavonoid subclasses (Moustafa and Wong 1967; Owens and McIntosh 2011 and ref. therein). Flavanone 3 $\beta$ -hydroxylase (F3H; EC 1.14.11.9) catalyzes the stereospecific 3 $\beta$ -hydroxylation of flavanones to form dihydroflavonols (Britsch et al. 1981; Britsch and Grisebach 1986). By introducing a double bond between C2 and C3 of a flavanone, the enzyme flavone synthase (FNS; EC 1.14.11.22) produces the corresponding flavone. This was first characterized in cell suspension cultures of parsley and was found to be encoded by a 2-oxoglutarate- dependent oxygenase (Sutter et al. 1975; Britsch et al. 1981; Martens et al. 2001). Flavonols are synthesized from dihydroflavonols catalyzed by flavonol synthase (FLS; E.C.1.14.11.23) that is also a 2-oxoglutarate-dependent dioxygenase requiring cofactors  $Fe^{2+}$  and ascorbate (Britsch et al.

1981; Owens and McIntosh 2011 and ref. therein). Similarly, dihydroflavonol 4-reductase (DFR; EC 1.1.1.219) catalyzes the stereospecific conversion of dihydroflavonols to the corresponding leucoanthocyanidins. This requires a cofactor, NADPH (Britsch et al. 1981; Stafford and Lester 1982; Heller et al. 1985). Conversion of leucoanthocyanidins to anthocyanidins is catalyzed by anthocyanidin synthase (ANS, LDOX; EC 1.14.11.19). It is also a 2-oxoglutarate-dependent dioxygenase requiring cofactors,  $\text{Fe}^{2+}$  and ascorbate (Saito et al. 1999).

Hydroxytransferases further add OH groups at different positions and modify the compounds. Flavanone 3'-hydroxylase (F3'H; EC 1.14.13.21) mediates hydroxylation of flavanones, flavones, dihydroflavonols, and flavonols, but is not active with anthocyanidins (Hagmann et al. 1983; Owens and McIntosh 2011 and references therein). Similarly, flavanone 3', 5'-hydroxylase (F3'5' H; EC 1.14.13.88) hydroxylates flavanones and dihydroflavonols having 4'-OH groups at 3' and 5' positions as well as hydroxylates 3', 4'-OH containing flavanones and dihydroflavonols at 5'-OH positions (Stotz and Forkmann 1982).

### Importance of Flavonoids in Plants

Flavonoids have a number of biological and ecological functions, including protection of plants against abiotic and biotic stresses such as UV rays and pathogen attack, signaling during nodulation, male fertility, and auxin transport (Mol et al. 1998; Winkle-Shirley 2002; Thill et al. 2013 and ref. therein). Flavonoids are well known because of the accumulation of red, blue and purple colored anthocyanin pigments that are distributed throughout different plant tissues (Winkle-Shirley 2001). These colored pigments can attract insects and other seed dispersal agents thereby promoting plant reproduction and distribution.

Anthocyanins are suggested to have a role in photo protection of chlorophyll that protects leaves from oxidative damage and promotes the recovery of nutrients during senescence (Field et al. 2001; Diaz et al. 2006). The accumulation of anthocyanins can be an indication of fruit ripening and fruit quality (Medina-Puche et al. 2014). In *Arabidopsis thaliana*, accumulation of anthocyanins is considered as a marker of a plant's response against adverse conditions unfavorable for growth and development (Chalker-Scoot 1999). Anthocyanins are accumulated in plants in response to abiotic and biotic stresses such as low availability of nutrients such as nitrogen and phosphorus (Diaz et al. 2006), pathogen attack (Lorenk-Kukula et al. 2005), drought (Castellarin et al. 2007), UV exposure (Kolesnikov and Zore 1957) and cold temperatures (Christie et al. 1994). Anthocyanins are found to accumulate in uninfected epidermal cells of maize plants surrounding the restricted lesions on leaves of resistant cultivar and it was suggested that these anthocyanins protect the healthy cells from toxic metabolites accumulated during the expression of resistance (Hipskind et al. 1996).

Flavonoids can screen harmful UV radiation and protect plants from pathogen attack (Li et al. 1993; Kolesnikov and Zore 1957; Tanaka et al. 2008). It was shown that flavonol kaempferol and other phenolic compounds played important roles *in vivo* to protect plants from UV-B radiation (Li et al. 1993). The ability of flavonoids to absorb light at 280– 320 nm is considered to be a fundamental property to prevent DNA damage (Stapleton and Walbot 1994). Study of *Arabidopsis* mutant's pigment analyses suggested an increase in production of UV absorbing compounds when exposed to UV radiation (Li et al. 1993; Bieza and Lois 2001). The increase in absorption in the UV region is due to an increase in flavonoid and sinapate accumulation (Bieza and Lois 2001). The expression of chalcone synthase (CHS) was

upregulated in mutants, suggesting an increase in UV blocking ability may be due to change in expression level of CHS genes (Bieza and Lois 2001).

Flavonoids provide resistance to aluminum toxicity in highly acidic soil, pH <5 (Kidd et al. 2001). Aluminum toxicity is a major factor that may affect the productivity of crops in acidic soil (Barcelo and Poschenrieder 2002 and ref. therein). Phenolics and organic acid chelates aluminum and detoxifies its effect (Barcelo and Poschenrieder 2002 and ref. therein). It was shown that in some of the aluminum resistant varieties of maize, plants that are pretreated with silicon exuded high levels of phenolics such as catechin and quercetin that were 15-fold higher than those plants not treated with silicon, suggesting that flavonoids have a significant role in Si induced amelioration of Al toxicity in plants (Kidd et al. 2001). This finding is consistent with the metal binding affinity of many flavonoids (Winkle-Shirley 2002 and ref. therein).

Similarly, flavonoids play an important role in plant-microbe interactions. Many legumes have N<sub>2</sub>-fixing bacteria that helps the plant obtain nitrogen when required (Treueter 2005 and ref. therein). The legume roots exude flavonoid signal molecules to attract bacteria to the root (Peters et al. 1986) and triggers the transcription of bacterial nodulation genes (i.e. *nod*, *nol* or *noe* genes). The protein products are required for infection that establishes the N<sub>2</sub>-fixing symbiosis of bacteria with leguminous plants resulting in the formation of nodules that facilitates the N<sub>2</sub>-fixation (Peters et al. 1986; Hungria and Stacey 1997 and ref. therein; Broughton et al. 2003; Kobayashi et al. 2004). A few examples of flavonoids that were isolated from the leguminous hosts under sterile conditions are quercetin, kaempferol, naringenin, apigenin, daidzein, myricetin, chalcones, eriodictyol, malvidin, genistein, luteolin, and petunidin (Firmin et al. 1986; Peters et al. 1986; Kossak et al. 1987; Hungria et al. 1991a; Hungria et al. 1991b;

Philips et al. 1992; Kobayashi et al. 2004). Productivity of economically important legume crops such as soybeans, peas, lentils, lupines, and peanuts can be enhanced by inducing nodulation using flavonoid compounds. Even in non-leguminous plants such as wheat, flavonoids are found to stimulate plant-microbe interactions. Naringenin was found to stimulate the root colonization of wheat by *Azorhizobium caulinodans* and other diazotrophic bacteria (Webster et al. 1998).

Flavonoids can also act as signaling molecules in regulating the expression of genes critical to phytopathogenesis (Peters and Verma 1990). Phenolic signaling molecules such as flavonoids mediate the transcription of virulence (*vir*) genes required for the transfer of T-DNA to host cells that results in the formation of crown galls. This was first observed in *Agrobacterium tumefaciens* (Stachel et al. 1985). Flavonol glycosides (quercetin 3-rutinosyl-4'-glucoside and kaempferol 3-rutinosyl-4'-glucoside) and a flavanone glucoside (dihydrowogonin 7-glucoside) from the crude extracts of Sweet Cherry leaves were found to trigger the expression of a virulence gene, *syfB*, required for synthesis of the phytotoxin, syringomycin (Mo et al. 1995).

Flavonoids protect plants from ozone damage by scavenging hydroxyl radicals, free oxygen and hydrogen peroxide (Appel 1993; Foy et al. 1995). It was demonstrated that the ozone tolerance in soybean cultivars is associated with the production of kaempferol glycosides (Foy et al. 1995). The most sensitive transgenic lines of soybean to ozone stress were the lines without kaempferol glycosides (Foy et al. 1995).

Flavonoids are found to play an important role in mediating allelopathic plant-plant interaction. Allelopathy is the inhibition of the growth and occurrence of one plant species due to chemicals secreted by another plant species (Rice 1984; Weir et al. 2004; Weston and Mathesius

2013 and ref. therein). The use of allelopathic compounds in agriculture may prevent the deterioration of environment by minimizing the use of herbicides, fungicides, and insecticides (Chou 1999). It was shown that *Centaurea aculosa* (spotted knapweed) exudes the allelochemical (–)-catechin from its roots that inhibits the growth and seed germination of susceptible species such as *Arabidopsis thaliana* (Bais et al. 2003). Similarly, flavonoid aglycones formononetin, medicarpin, and kaempferol as well as glycosides of kaempferol and quercetin have been reported to have potential for allelopathy and autoallelopathic interactions (Weston and Mathesius 2013 and ref. therein).

Flavonoids have major role in mediating plant-insect interactions (Simmonds 2003 and ref. therein). *Polyommatus icarus*, a common blue butterfly withdraws flavonoids from their larval host plants and use them to attract males (Burghardt et al. 2000). Resin on the surface of leaves of the shrub *Mimulus aurantiacus* Cortis contains flavonoids that protects plants by inhibiting the growth of *Euphydryas chalcedona* (Lincoln 1985; Lincoln and Wala 1986). Flavonoids also act as feeding deterrents and several insects are deterred by flavonoids (Harborne 1988 and ref. therein; Thoison et al. 2004). For example, crude water extracts of Japanese larch, *Larix leptolepis*, wood that contain higher amounts of dihydroflavonols (Yasuda et al. 1975) exhibited strong feeding deterrent activities against the subterranean termite, *Coptotermes formosanus* (Chen et al. 2004). Similarly, flavonoids are found to stimulate the oviposition of the monarch butterfly *Danaus plexippus*. Three quercetin galactosides; quercetin-3-rutinosyl, quercetin -3- B-D-glucopyranosyl (1-6) B-D-galactopyranosyl, quercetin - 3-2'',6'', alpha-L-dirhamnopyranosyl-B-D- galactopyranosyl were isolated from *Asclepias currassavica* that stimulated oviposition of *Danaus plexippus* (Baur et al. 1998).

Flavonoids have anti-oxidant properties in plants. Reactive oxygen species, such as hydroxyl radicals (OH), hydrogen peroxide (H<sub>2</sub>O<sub>2</sub>), singlet oxygen (<sup>1</sup>O<sub>2</sub>), and superoxide anions (O<sub>2</sub><sup>-</sup>) are produced and accumulated when plants are exposed to different environmental stresses (Asada 2006; Van Breusegem and Dat 2006). This leads to oxidative damage of different cellular components, such as DNA, lipids, proteins and sugars (Asada 2006; Van Breusegem and Dat 2006). Flavonoids protect plants from these oxidative damages by scavenging free radicals (Agati et al. 2012 and ref. therein).

### Importance of Flavonoids in Mammals

Flavonoids form an important part of the human diet and have many potential health impacts. Flavonoids are abundant in fruits, seeds, vegetables, and beverages, the consumption of which has been shown to be beneficial for human health (Quideau et al. 2011 and ref. therein). In humans, flavonoids are found to have anti-cancer (Benavente et al. 1997 and ref. therein; Birt et al. 2001; Clere et al. 2011), antibacterial (Hanasaki et al. 1994), antiviral (Hanasaki et al. 1994), anti-inflammatory (Chirumbolo, 2010; Gonzalez et al. 2011), antiallergic (Hope et al. 1983; Hanasaki et al. 1994; Middleton and Kandaswami, 1994) and vasodilatory (Duwarte et al. 1993) activities.

Flavonoids are reported to possess strong antioxidant activity and free-radical-scavenging ability because of their capacity to donate electrons or hydrogen atoms (Robak and Gryglewski 1988; Yuting et al. 1990; Hernandez et al. 2009). A hydroxyl group in the C-3 position, presence of a double bond between C-2 and C-3, a carbonyl group in C-4, and polyhydroxylation of the A and B aromatic rings confers ability of flavonoids to act as an anti-oxidant and free radical scavengers (Cook and Samman 1996). Quercetin, hesperetin, naringenin, and rutin were tested in



two different *in vitro* experimental models;  $\text{Fe}^{2+}$  induced linoleate peroxidation ( $\text{Fe}^{2+}$ -ILP) and auto-oxidation of rat cerebral membranes (ARCM) (Saija et al. 1995). The ability of flavonoids to interact with and penetrate the lipid bilayer of model membranes was studied (Saija et al. 1995). The capacity of these flavonoids to modify free-radical-induced membrane lipoperoxidation is related to their ability to interact with and penetrate the lipid bilayers (Saija et al. 1995).

Quercetin, a flavonol, has a number of pharmaceutical implications. It is a strong antioxidant which prevents oxidation of low density lipoproteins *in vitro* (Hollman et al. 1998). It has been suggested that the dietary intake of flavonols and flavones is inversely related to the occurrence of coronary heart disease in humans (Hollman et al. 1998). Quercetin has anti-cancerous properties as well (Boly et al. 2011). It was found to inhibit 16 different protein kinases that are directly involved in controlling mitotic processes (Boly et al. 2011). Similarly, quercetin has been found to play a major role in prevention of neurodegenerative diseases. The neuroprotective effects of quercetin against  $\text{H}_2\text{O}_2$ -induced apoptosis in human neuronal SH-SY5Y cells were investigated (Suematsu et al. 2011). It was found that quercetin induced the expression of anti-apoptotic Bcl-2 gene and repressed pro-apoptotic Bax gene in SH-SY5Y cells (Suematsu et al. 2011).

More recently, the impact of flavonoids on human gut microbiota was studied and it was shown that an increase in the dietary uptake of fruits and vegetables containing flavonoids could potentially inhibit the growth of pathogenic *Clostridia* (Klinder et al. 2016). Flavonoids were shown to mediate the protective effects by interacting with gut microbiota (Klinder et al. 2016). Similarly, Citrus flavonoids such as hesperetin, naringenin, tangeretin, and nobiletin have been

reported as important therapeutics for the treatment of metabolic dysregulation (Mulvihill et al. 2016 and ref. therein). Citrus flavonoids trigger inflammatory response in tissues of kidney, aorta, liver, and the adipose tissue (Mulvihill et al. 2016 and ref. therein). More recently, Citrus flavonoids were found to have an important role in the treatment of obesity, insulin resistance, dyslipidemia, atherosclerosis, and hepatic steatosis (Mulvihill et al. 2016 and ref. therein).

Similarly, the antiviral effect of 10 different flavonoids against feline calicivirus (FCV) and murine norovirus (MNV) was investigated on Crandell-Reese feline kidney (CRFK) cells and RAW 264.7 cells and it was found that the flavonols quercetin, fisetin, and kaempferol inhibited FCV and MNV in a dose dependent manner (Seo et al. 2016). Flavonoids and polyphenols such as catechins, proanthocyanidins, theaflavins, and flavonols were reported to possess anti-viral activity against influenza A and B viruses (Yang et al. 2014).

In a recent study of the structure–activity relationship of phenolic hydroxyls and alcoholic hydroxyls on the reduction of acrylamide formation by flavonoids, it was shown that phenolic hydroxyls contribute to reduction effects of acrylamide (Zhang et al. 2016). As acrylamide is a carcinogen and genotoxin in rodents with high-dose exposure and also a neurotoxin in humans (Zhang et al. 2016), the acrylamide chemoprevention role of flavonoids clearly highlights the benefits of presence of flavonoid compounds in foods we consume.

Flavonols and isoflavonols were found to be cytotoxic against human oral squamous cell carcinoma and salivary gland tumor cell lines (Sakagami et al. 1999). Some dietary flavonoids such as quercetin, naringin, hesperetin and catechin were reported to inhibit the replication and infectivity of certain RNA (Respiratory Syncytial Virus, Par influenza virus, Polio virus) and DNA (Herpes Simplex Virus) viruses (Kaul, 1985). In addition, flavonoids have been found to

inhibit lipid peroxidation (Salvayre et al. 1987; Ratty and Das, 1998), platelet aggregation (Tzeng et al. 1991), capillary permeability and fragility (Torel et al. 1986; Robak et al. 1987). Specific flavonoids act as enzyme inhibitors, hormones or immune modulators and prevent human diseases (reviewed in Lee et al. 2007).

### Glycosylation

Glycosylation is the general term for the reactions involving the transfer of a sugar moiety from sugar donors to acceptor substrates including plant metabolites, phytotoxins and xenobiotics (Bowles et al. 2005). Many sugar donors have a sugar moiety and a UDP group (Owens and McIntosh 2011 and ref. therein; Sharma et al. 2014). Glycosyltransferases (GTs; EC 2.4.1. x) constitute a large family of transferase enzymes that are involved in the biosynthesis of a myriad of plant metabolites such as oligosaccharides, polysaccharides or glycoconjugates (Taniguchi et al. 2002). Glycosylation modifies the structure and hence influences chemical properties and complexity (Wang 2009 and ref. therein). It enhances the solubility and stability through the protection of reactive nucleophile groups, thus facilitating their storage, accumulation, and transport, as well as regulating their bioavailability for other metabolic processes (Jones and Vogt 2001; Bowles et al. 2005; Gachon et al. 2005; Owens and McIntosh, 2009; Bhat et al. 2013). Increased stability and water solubility of glycosylated products assist in the inactivation and detoxification of xenobiotic and harmful compounds, as well as regulation of hormones (Gachon et al. 2005).

Glycosylation of flavonoids is considered to be a key mechanism in metabolic homeostasis of plant cells (Bowles et al. 2005). UDP-glycosyltransferases (UGTs) catalyze the transfer of a glycosyl group such as glucose, galactose, xylose, rhamnose, mannose,

glucosamine, etc. from a nucleotide diphosphate sugar donor (UDP-sugar) to different acceptor substrates (Wang 2009 and ref. therein). UGTs exhibit specificity to sugar donors as well as acceptor molecules. Some are highly specific to acceptor substrates that mediate glycosylation of only one or two types of acceptor molecules (McIntosh et al. 1990; Shao et al. 2005; Owens and McIntosh 2009) whereas some glycosylate a broad range of acceptor substrates (Osmani et al. 2009 and ref. therein).

### Plant Secondary Product Glucosyltransferase

Glucosyltransferases are the soluble proteins that catalyze the transfer of a glucose molecule from a nucleotide-activated sugar donor, UDP-glucose, onto acceptor molecules resulting in the formation of glucosides (Harborne 1993 and ref. therein). Glucosylation is also the final step in the biosynthesis of many plant secondary metabolites (McIntosh et al. 1990 and ref. therein; Jones and Vogt, 2001; Devaiah et al. 2016). Among different families of GTs, the largest is the GT1 enzymes that use UDP-activated sugars as donor in enzymatic reactions (Vogt and Jones, 2000; Bowles et al. 2006; Yonekura-Sakakibara and Hanada, 2011), hence called uridine diphosphate glucosyltransferase (UGTs). One of the characteristic features of this family of GTs is the presence of a conserved motif near the carboxy-terminal (44-amino acid residues), known as the Plant Secondary Product Glucosyltransferase Box (Hughes and Hughes, 1994; Vogt and Jones 2000; Owens and McIntosh 2011 and ref. therein; Devaiah et al. 2016). The residues within this signature motif have roles in forming the donor substrate binding pocket (Hughes and Hughes 1994; Vogt and Jones 2000; Paquette et al. 2003; Offen et al. 2006; Li et al. 2007). In contrast, the acceptor substrate binding pocket is formed within the N-terminal domain (Mackenzie et al. 1997; Hu and Walker, 2002; Shao et al. 2005; Offen et al. 2006; Osmani et al.

2009 and ref. therein). The sequences within the PSPG box are almost 60-80% similar within the related flavonoid GTs, however, the overall sequence similarity is fairly low (Sarkar et al. 2007; Owens and McIntosh 2009).

### Classification and Structure of GT Enzymes

There are two major classification systems for GTs, based on the reactions they catalyze according to the recommendations of the International Union of Biochemistry and Molecular Biology (Webb 1992) and classification based on the similarity of inferred amino acid sequence (Campbell et al. 1997). Based on amino acid similarity, 98 families of GTs have been defined in the CAZy (Carbohydrate-Active Enzymes) database (Cantarel et al. 2009; <http://www.cazy.org/GlycosylTransferases.html>). As of September 2013, the CAZy database reports sequence information on almost 119,910 putative GTs (Lombard et al. 2014) including all nucleotide sugar-dependent enzymes, glycosyltransferases that utilize sugar donors such as lipid diphospho-sugars, nucleoside monophospho-sugars, sugar-1-phosphates, and dolichol-phospho-sugars (Coutinho et al. 2003). These enzymes catalyze the *O*-glycosylation reaction and generates the phospho-containing group that departs from their corresponding activated sugar donors (Coutinho et al. 2003). However, this classification does not include enzymes such as cyclodextrin glucanotransferases (EC 2.4.1.19), dextranase (EC 2.4.1.5), and xyloglucan endotransferases (EC 2.4.1.207) (Coutinho et al. 2003) that are classified as glycosidases (Henrissat 1991; Henrissat and Bairoch 1993; Henrissat and Bairoch 1996; Coutinho et al. 2003).

CAZy enzymes are classified as “inverting” or “retaining” (Fig. 1.4) based on the stereochemistry of the carbon atom of the sugar donor (Lairson et al. 2008 and ref. therein). Inverting GTs catalyze the reaction through a single nucleophilic substitution step. The acceptor

mediates a nucleophilic attack at C1 of the sugar donor and the configuration of the sugar is changed (Liu and Mushegian, 2003). Most of the inverting GTs need a divalent cation, typically  $\text{Mg}^{2+}$  or  $\text{Mn}^{2+}$ . However, there are some metal-independent enzymes in this group as well (Lairson et al. 2008 and ref. therein). In the reactions catalyzed by inverting GTs, there is net inversion of stereochemistry of the sugar donor whereas there is net retention of stereochemistry of the donor substrate in the reactions catalyzed by retaining GTs (Sinnott 1990; Coutinho et al. 2003; Lairson et al. 2008). When the nucleophile reacts with the glycosyl donor it generates a glycosyl-enzyme complex with inversion of stereochemistry (Lairson et al. 2008 and ref. therein). Then a second inversion of stereochemistry occurs after the reaction with the acceptor that generates the product of the reaction with a net retention of stereochemistry (Lairson et al. 2008 and ref. therein).

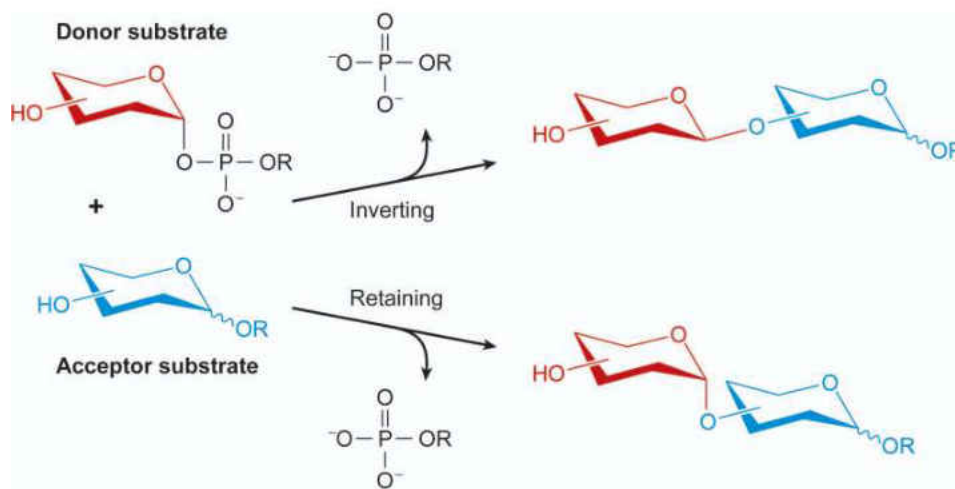


Fig. 1. 4 Inverting and retaining GTs. Catalysis of the GT enzymes via inversion or retention of configuration of the sugar donor (Lairson et al. 2008, Used with permission)

All nucleotide dependent GTs exhibit two structural folds, GT-A and GT-B (Fig. 1.5) (Lairson et al. 2008 and ref. therein). More recently GT-C fold have been identified (McIntosh and Owens 2016 and ref. therein) and these are integral membrane, and inverting N-glucosyltransferases (Liang et al. 2015 and ref therein). The GT-A fold was first observed for SpsA from *Bacillus subtilis* (Lairson et al. 2008 and ref. therein) and GT-B for the phage T4  $\beta$ -glucosyltransferase and glycogen phosphorylase (Bourne and Henrissat 2001).

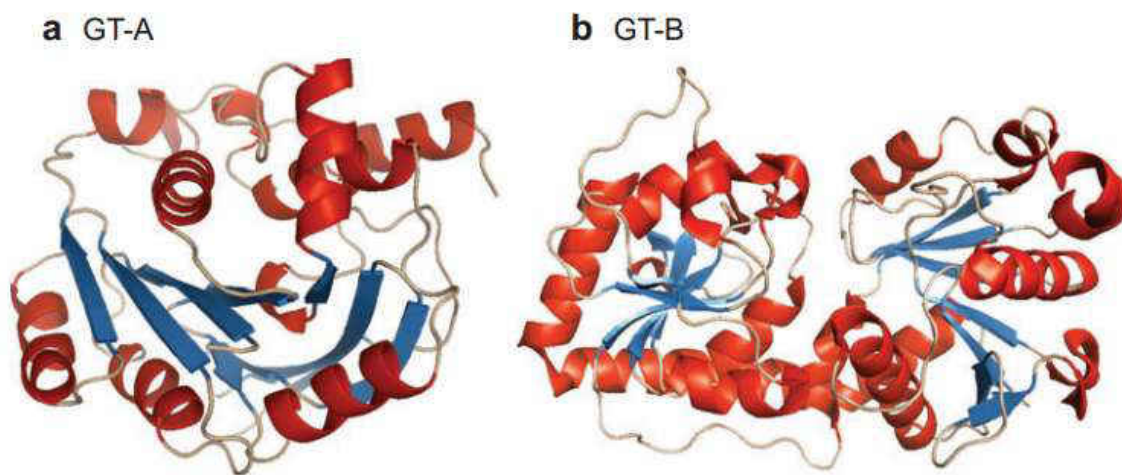


Fig. 1. 5 Structure of GT-A and GT-B folds of GTs. The GT-A represented by SpsA from *Bacillus subtilis* and (b) the GT-B fold by phage T4  $\beta$ -glucosyltransferase (Lairson et al. 2008, Used with permission)

GT-A and GT-B have  $\beta / \alpha / \beta$  Rossmann-like domain (Coutinho et al. 2003; Sharma et al. 2013). As shown in Fig. 1.5, the fold is made up of alternating secondary structure consisting of  $\beta$ -strands and  $\alpha$ -helices (Coutinho et al. 2003; Osmani et al. 2009 and ref. therein). In the GT-A fold, the  $\beta / \alpha / \beta$  secondary structure is associated tightly and abutting each other whereas in GT-B the structure is loosely arranged and faces each other (Coutinho et al. 2003). The GT-A

enzymes require a divalent cation such as  $Mg^{2+}$  or  $Mn^{2+}$  and have a single Rossmann domain (Lairson et al. 2008) whereas the GT-B fold has two Rossmann domains (Coutinho et al. 2003; Sharma et al. 2013). The nucleotide binding site is present on the N-terminal domain and the acceptor binding on the C-terminal domain of the GT-A enzymes whereas this is just opposite in the case of GT-B enzymes i.e. nucleotide binding on the on the C-terminal domain and acceptor binding on the N-terminal domain (Coutinho et al. 2003).

### Crystal Structures of Family1 GTs

Crystal structures of GTs provide insight to the interactions between the enzyme's active site and its acceptor and donor substrates. Crystal structures of Uridine diphosphate glycosyltransferases (UGTs) in complex with donor and acceptor substrates provide a basis for the identification of key residues that are directly involved in the enzyme function and activity (Offen et al. 2006). The amino acid sequence conservation is low in GTs (Campbell et al. 1997; Hu and Walker 2002, Sarkar et al. 2007), even among the enzymes with established similar substrate and regiospecificity (McIntosh and Owens 2016 and ref. therein). However, the tertiary structure remains highly conserved among these GTs (Osmani et al. 2007). The study of the 3-D structure of protein is an important tool to compare the structure and function of GTs.

To date, 6 crystal structures of plant secondary product family1 GTs have been solved (Table 1.1). All of these possess a typical GT-B fold structure, with inverting catalysis which is highly conserved in plant UGTs (Breton et al. 2012). Common to all of these crystallized UGTs, the sugar donors are bound to the C-terminal domain and the acceptors bind to the N-terminal domain (Li et al. 2007). The N- and C-terminal domain of these UGTs contain several strands (ranging from six to ten) of parallel  $\beta$  sheets flanked by  $\alpha$  helices. These two domains face each



other and are loosely associated, and active site is formed in a cleft between these two domains (Osmani et al. 2009 and ref. therein). The major conformational differences between these UGTs are observed in their N-terminal domains, mainly in the linker region that connects the N and C-terminal (Li et al. 2007). This is important because the active site is formed in a cleft in between these two domains. Similarly, in all of these UGTs, two key amino acid residues, His and Asp, are identified as catalytic residues. Histidine acts as Bronsted base for deprotonation of the acceptor molecule (Modolo et al. 2009), and an aspartate helps in stabilizing the structure after deprotonation (Offen et al. 2006). Table 1.1 summarizes the function of GTs whose crystal structures have been solved.

Table 1.1 UGTs with crystal structures.

S.N	GTs	Function	Reference
1	UGT71G1, <i>Medicago truncatula</i> (PDB 2ACV )	Triterpenes and flavonols	Shao et al. 2005
2	VvGT1, <i>Vitis vinifera</i> (PDB 2C1Z)	Anthocyanidins, flavonols	Offen et al. 2006
3	UGT72B1, <i>Arabidopsis thaliana</i>	Dichloroanilines and trichlorophenols	Braizer-Hicks et al. 2007
4	UGT85H2, <i>Medicago truncatula</i>	Flavonol, isoflavone	Li et al. 2007
5	UHT78G1, <i>Medicago truncatula</i>	Flavonols, isoflavones, anthocyanidins	Molodo et al. 2009
6	UGT78K6, <i>Clitoria ternatea</i>	Anthocyanidins	Hiromoto et al. 2013 and 2015

UGT71G1 is a multifunctional triterpene and flavonoid glycosyltransferase from *Medicago truncatula*. Its crystal structure in complex with UDP or UDP-glucose has been solved (Shao et al. 2005). The crystal structure of UGT71G1 in complex with UDP-glucose suggested that His-22 and Asp-121 are the catalytic residues that help in deprotonation of the acceptor

substrate by forming an electron transfer chain within the catalytic base (Shao et al. 2005). UGT71G1 is involved in the glucosylation of sapogenins forming saponins, however it also favors kinetics with quercetin (Shao et al. 2005). The N-terminal domain is made up of seven  $\beta$ -sheets flanked by eight  $\alpha$ -helices on both sides and a small two-stranded  $\beta$  sheet (Shao et al. 2005). The C-terminal domain contains a six-stranded  $\beta$  sheet flanked by eight  $\alpha$  helices (Shao et al. 2005). UDP-binding pocket is formed in a deep cleft between two domains (Shao et al. 2005).

Similarly, VvGT1 from *Vitis vinifera* glucosylates anthocyanidins but has preferred specificity towards flavonols such as quercetin and kaempferol (Offen et al. 2006). The crystal structure of VvGT1 in Michaelis complex with UDP-glucose, with acceptor substrate kaempferol, and in complex with UDP and quercetin has been solved. This UGT is promiscuous with respect to the sugar donor specificity as it can accept wide range of sugar donors such as UDP- glucose, UDP-xylose, UDP-mannose, UDP-galactose and UDP- *N*-acetyl-D-glucosamine. However, the activity with UDP-glucose was higher compared to other sugar donors. Three residues (Asp374, Gln375 and Thr141) were found to interact directly with the hydroxyl groups of the glucose moiety of UDP-glucose (Offen et al. 2006).

UGT72B1 is a bifunctional *O*-glucosyltransferase (OGT) and *N*-glucosyltransferase (NGT) from *Arabidopsis thaliana* that is involved in the *O*-glycosylation of trichlorophenols and *N*-glycosylation of chloroanilines (Braizer-Hicks et al. 2007). The crystal structure of UGT72B1 was solved in a complex with the donor UDP-glucose, UDP and Tris buffer, and in Michaelis complex (for *O*-glycosyltransfer) with trichlorophenol and UDP-2- deoxy-2-fluoro glucose (Braizer-Hicks et al. 2007). This UGT has application in the detoxification of xenobiotics (Braizer-Hicks et al. 2007).

Similarly, UGT85H2 is a (iso) flavonoid glycosyltransferase from *Medicago truncatula* which shows activity towards kaempferol, biochanin A, and isoliquiritigenin (Li et al. 2007). Kinetic studies suggested that this UGT has preference and higher catalytic activity with flavonols than for isoflavones or chalcones (Li et al. 2007).

UGT78G1 from *Medicago truncatula* is a glycosyltransferase that glycosylates kaempferol and myricetin (flavonols), formononetin (isoflavone), and pelargonidin and cyanidin (anthocyanidin) (Molodo et al. 2009). Interestingly, under certain conditions this UGT removes the sugar moiety from the glycosylated products (Molodo et al. 2009). The crystal structure of UGT78G1 was solved in complexes with UDP or with UDP and myricetin. From the crystal structure of UGT78G1, glutamate 192 was identified as a key residue for the glycosidase activity. This enzyme catalyzes the conversion of biochanin A-7-*O*-glucoside, genistein-7-*O*-glucoside, kaempferol-3-*O*-glucoside, and quercetin-3-*O*-glucoside into the corresponding aglycones.

Ct3GT-A and UGT78K6 from *Clitoria ternatea* catalyzes the transfer of glucose from UDP-glucose to anthocyanidins such as delphinidin (Hiromoto et al. 2013). The crystal structure of Ct3GT-A (Hiromoto et al. 2013) and the crystal structure of UGT78K6 and its complex forms with anthocyanidins (delphinidin and petunidin) and with the flavonol kaempferol were determined (Hiromoto et al. 2015). UGT78K6 from *Clitoria ternatea* is anthocyanidin specific and had highest activity with delphinidin as acceptor. However lower relative activity (less than 30%) were detected on pelargonidin and cyanidin and weak glucosylation activities (less than 10%) were reported on different flavonols such as isorhamnetin, quercetin, kaempferol, and myricetin (Hiromoto et al. 2015)

### The Model Plant; *Citrus paradisi*

Citrus varieties are known to produce an array of flavonoid compounds that are present naturally in glycosylated form (Owens and McIntosh 2011 and ref. therein). Almost all the enzymes of the flavonoid biosynthetic pathway are present in Citrus. Citrus species are found to produce high levels of flavonoid glycosides such as flavanone, flavone, and flavonols as well as dihydroflavonol and chalcone glycosides (McIntosh et al. 1990; McIntosh and Mansell 1997; Berhow et al. 1998; Owens and McIntosh 2009; Owens and McIntosh 2011 and ref. therein). The most prevalent glycosylated flavonoids in Citrus are flavanone- and flavone-7-*O*-diglycosides (Owens and McIntosh 2011 and ref. therein). Flavanone-7-, flavone-7-, and flavonol-*O* glucosyltransferases were identified and partially purified from young leaves of grapefruit (McIntosh and Mansell 1990; McIntosh et al. 1990). Glucosyltransferases that mediate C-glucosylation of aglycones were also reported in Citrus, but fairly in low amount (Jay et al. 2006). Approximately 60 different flavonoid compounds containing D-glucose and L-rhamnose have been reported in Citrus juices (Barreca et al. 2012).

Grapefruit is an important cash crop as well an excellent model system to study the biosynthesis and metabolism of glycosylated flavonoids due to the presence of diverse groups of flavonoid glycosides (Owens and McIntosh 2011; Devaiah et al. 2016). A flavanone specific glucosyltransferase catalyzing the 7-*O*- glucosylation of flavanone aglycones in grapefruit seedlings was purified to near homogeneity and was characterized for its biochemical and kinetic properties (McIntosh et al. 1990). This enzyme catalyzes the transfer of sugar in the first reaction that synthesizes flavanone-7-*O*-neohesperidosides or rutinosides (McIntosh and Mansell 1990; McIntosh et al. 1990). The major flavonoid found in *Citrus paradisi* is the bitter compound naringin, the 7- $\beta$ -neohesperidoside of naringenin (4', 5, 7-trihydroxyflavanone) which accounts

for 40 to 70% of the dry weight of small green fruits (Kesterson and Hendrickson 1953; Jourdan et al. 1985). The quantitative distribution of naringin in different tissues and plant parts using radioimmunoassay showed that the young tissues of grapefruit accumulated high levels of naringin (Jourdan et al. 1985). The bitter taste of grapefruit is due the presence of two compounds, naringin and the bitter terpenoid limonin (Mansell et al. 1983). Naringin is the major bitter compound that imparts immediate bitterness to the grapefruit juice (Owens and McIntosh 2011).

#### Flavonol Specific 3-*O*-Glucosyltransferase from Grapefruit

Flavonol specific 3-*O*-Glucosyltransferase from grapefruit is an enzyme that catalyzes the transfer of a glucose molecule from an activated sugar donor, UDP-Glucose, to acceptor substrates (flavonol aglycones) at the 3-OH position (Fig. 1.7). The products of the reaction catalyzed by this enzyme are flavonol glucosides (Owens and McIntosh, 2009). The recombinant enzyme, Cp3GT, has MW of 51.2 kDa and a pI of 6.27. This has 96% amino acid identity with a *Citrus sinensis* homologue for which direct biochemical activity was not tested (Lo Piero et al. 2005). Cp3GT was expressed in *E.coli* and the recombinant enzyme was screened for activity with a number of different flavonoid compounds. The enzyme glucosylated only the flavonol aglycones quercetin, kaempferol, and myricetin. Glucosylation of flavonols occurred at 3-OH position and the products of the reaction were confirmed by HPLC and TLC by using standard compounds as ref. (Owens and McIntosh 2009). The optimum pH and temperature for this enzyme was 7.5 and 40-50°C, respectively. Tris-HCl buffer had pronounced effect on enzyme activity, with reduction of enzyme activity by 51% under optimal pH conditions. The enzyme was inhibited by different divalent metal ions such as Cu<sup>2+</sup>, Fe<sup>2+</sup>, and Zn<sup>2+</sup> as well as UDP, a

product of the reaction. Cysteine, histidine, arginine, tryptophan, and tyrosine residues were identified as important amino acids for enzyme activity by the treatment of enzyme with a variety of amino acid modifying compounds. (Owens and McIntosh 2009).

A large amount of insoluble recombinant protein was contained within the inclusion bodies when expressed in the bacterial expression system and this was completely absent when expressed in yeast (Devaiah et al. 2016). Expression in yeast yielded higher amounts of soluble protein as compared to bacterial expression systems (Devaiah et al. 2016). The substrate preference for flavonols, effect of metals ions,  $K_i$  value for inhibitor UDP, temperature optimum and temperature stability were similar for enzyme from both expression systems (Devaiah et al. 2016). The major difference observed was the increased stability at higher pH which could be due to the nature and placement of tags. In the yeast expression system, the C-myc tag and 6x-Histidin tags are present at the C- terminus whereas the thioredoxin tag and 6x-Histidine tags are present at the N-terminus when a bacterial expression system was used (Owens and McIntosh 2009; Devaiah et al. 2016).

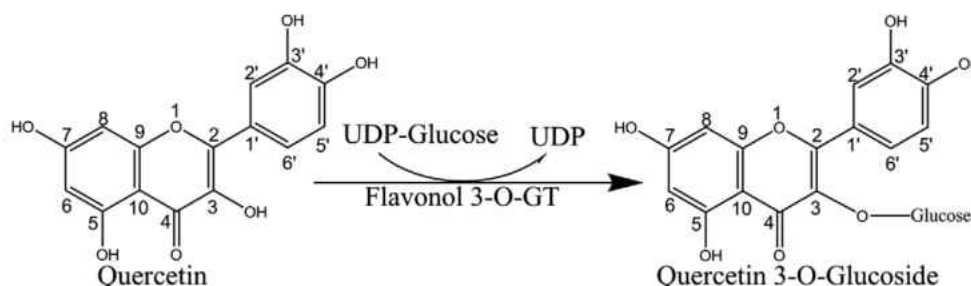


Fig. 1. 6 Schematic of a Cp3GT catalyzed reaction. The Cp3GT catalyzed reaction with quercetin as the acceptor substrate and UDP-glucose as an activated sugar donor. (Owens and McIntosh, 2009).

The grapefruit 3-*O*-glucosyltransferase enzyme glucosylates flavonol aglycones at the 3-OH position (Owens and McIntosh, 2009) forming 3-*O*-glucosides. The red grape enzyme UDP-

glucose; flavonoid 3-*O*-glycosyltransferase from *Vitis vinifera* (VvGT) glucosylates flavonols as well as anthocyanidins at the 3-OH position (Offen et al. 2006). VvGT shares 56% sequence identity and 83% homology with Cp3GT. The crystal structure of VvGT showed that threonine at position 141 of red grape enzyme VvGT is present on the N-terminal domain. The residues within the N-terminal domain are involved in forming the acceptor substrate binding pocket. Further it was found that threonine at this position interacts with the hydroxyl group of UDP-glucose, the sugar donor and hence may be important in enzyme activity (Offen et al. 2006). Through multiple sequence alignment and 3-D modeling, the corresponding amino acid residue was found to be proline at position 145 of Cp3GT. Homology modeling was done to predict the 3-D structure of Cp3GT and compared with the solved crystal structure of VvGT as template. It was observed that threonine 141 was closer to the substrate binding pocket of VvGT and the corresponding residue in Cp3GT and proline 145 was found to be closer to the substrate binding pocket. This modeling analysis gave an insight into prospects of substituting proline with threonine. Another important factor to substitute proline with threonine is the nature of these amino acids. Proline has a cyclic side chain that forces a bend in the structure of protein. This may affect the protein folding and might affect the interaction with other residues within the active site of the protein. Thus, we hypothesized that replacing the amino acid that imparts structural rigidity on the tertiary structure of a protein by a completely different type of amino acid may allow free rotation resulting in different orientation or may expose protein to different interactions resulting in a change in the structure and hence the catalytic properties of the enzyme.

The main objective of the research in our lab is the structure-function study of this particular GT from grapefruit. Site-directed mutagenesis was performed to test hypotheses

regarding the substrate and regiospecificity of the flavonol specific Cp3GT. Mutated enzyme is biochemically characterized and the structure modelled using different computational tools. Experimental results are used to validate the computational models. This research is focused on the characterization of a mutant, P145T, and study of the effects of tags on the recombinant enzyme activity.

### Hypotheses

This research is the continuation of the study on the structural and functional analysis of flavonol specific 3-*O*-glucosyltransferase from grapefruit. This research focused on analyzing the effect of a substituting proline at position 145 of Cp3GT with threonine as well as characterization wild type and mutant P145T in presence and absence of tags. A thrombin cleavage site was inserted in both wild type and mutant P145T for the removal of the C-myc and 6x-His tags.

The two hypotheses of this research are:

**Hypothesis 1:** Mutation of proline in Cp3GT to threonine can alter substrate specificity of Cp3GT.

**Hypothesis 2:** Mutation of proline in Cp3GT to threonine can alter regiospecificity of Cp3GT.



## CHAPTER 2

### STRUCTURAL AND FUNCTIONAL ANALYSIS OF GRAPEFRUIT FLAVONOL-SPECIFIC-3-*O*-GT MUTANT P145T

Sangam Kandel<sup>1</sup>, Shivakumar P. Devaiah<sup>2</sup>, Cecilia A. McIntosh<sup>1, 3</sup>

<sup>1</sup>Department of Biological Sciences, East Tennessee State University, Johnson City, TN, 37614

<sup>2</sup>BioStrategies-LC, State University, Arkansas, 72467

<sup>3</sup>School of Graduate Studies, East Tennessee State University, Johnson City, TN, 37614

Corresponding Author:

Cecilia A. McIntosh

Department of Biological Sciences

East Tennessee State University

Box 70703, Johnson City, TN 37614

Phone: (423)-439-5838

Email: [mcintosc@etsu.edu](mailto:mcintosc@etsu.edu)

## Abstract

The main objective of the research in our lab is the structural and functional analysis of flavonol-specific-3-*O*-GT from grapefruit (Cp3GT). This research has two main objectives: 1) Study of the effect of mutating proline to a threonine on the substrate and regiospecificity of Cp3GT, 2) Enzymatic characterization of wild type and mutant P145T with and without C-myc and 6x-His tags. The mutant P145T altered the substrate and regiospecificity, glucosylating flavanones and flavones along with flavonols, while the wild type glucosylated only flavonols. The most intriguing result was the glucosylation by P145T enzyme of flavanone and flavones. HPLC was used for product identification and showed mutant P145T glucosylated naringenin forming naringenin-7-*O*-glucoside. Homology modeling and molecular docking was performed to identify the acceptor substrate recognition patterns and models were validated by experimental results. Site-directed mutagenesis was performed to insert a thrombin cleavage site into the recombinant wild type Cp3GT and mutant P145T enzymes between the enzyme and the C-myc tags. This was done in preparation for future studies to crystallize the proteins and to test effect of tags on enzyme activity.

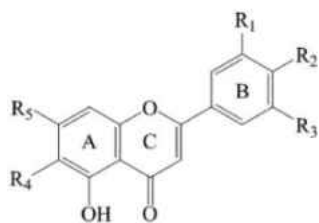
## 1. Introduction

Flavonoids are a major group plant secondary metabolites produced as a final product of the phenylpropanoid pathway [1]. Over 10,000 flavonoids exhibiting structural diversity have been identified in higher plants [2, 3, 4]. Oxidation of the C-15 ring system is the major distinguishing feature for classifying flavonoids into different sub-classes [5, 6] such as flavonol, flavone, flavanone, iso-flavonoid, and anthocyanidin (Fig. 2.1).

Flavonoids have been of interest to scientists for decades because of their important biological and ecological functions in plants. Flavonoids have important roles in photo protection, ability to scavenge reactive oxygen species [7], coloration of plants including flower pigmentation and UV-patterning [8], plant-insect interactions as feeding deterrents and attractants [5], mediates allelopathic plant-plant interactions that contribute to maintaining plant the diversity of plants, as well as enhance the agricultural productivity [4, 9]

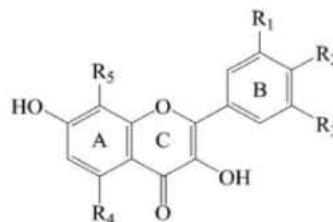
Flavonoid glycosides also have important roles in human health [10]. Flavonoids are shown to possess antioxidant properties, anti-inflammatory, anti-allergenic, anti-carcinogenic, anticancer, and antiviral activities, free radical scavenging capacity, coronary heart disease prevention, and hepatoprotective activities [11].

Most flavonoids are present in plants in glycosylated form. Glycosylation is mediated by a ubiquitous family of enzymes called glycosyltransferases. Glycosylation is an important reaction that modifies the structure and hence influences chemical properties and complexity of flavonoids [12].



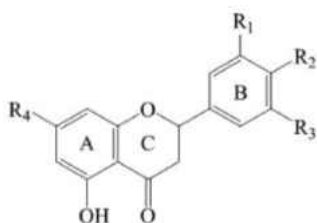
Flavone

Luteolin: R1=R2=R5=OH; R3=R4=H  
 Apigenin: R1=R3=R4=H; R2=R5=OH  
 Diosmetin: R1=R5=OH; R2=OCH3; R3=R4=H  
 Scutallerein: R1=R3=H; R2=R4=OH; R5=O-glu



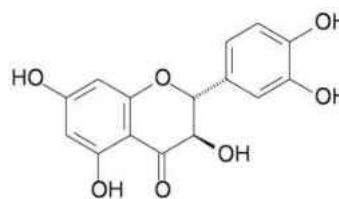
Flavonol

Quercetin: R1=R2=R4=OH; R3=R5=H  
 Kaempferol: R1=R3=R5=H; R2=R4=OH  
 Fisetin: R1=R4=R5=H; R2=R3=OH  
 Gossypetin: R1=R2=R4=R5=OH; R3=H

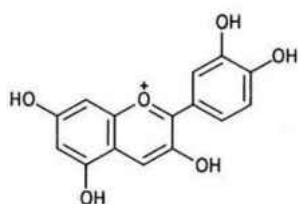


Flavanone

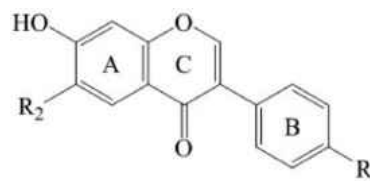
Naringenin: R1=R3=H; R2=R4=OH  
 Hesperetin: R1=H; R2=OCH3; R3=R4=OH  
 Eriodictyol: R1=R2=R4=OH; R3=H  
 Isosakuranetin: R1=R3=H; R2=OCH3; R4=OH



Dihydroflavonol (Dihydroquercetin)



Anthocyanidin (Cyanidin)



Isoflavone

4'-acetoxy-7-hydroxy-6-methoxy isoflavone:  
 R1=CH<sub>2</sub>COOCH<sub>3</sub>; R2= OCH<sub>3</sub>

Fig. 2. 1 Generic structure of major classes of flavonoids and substitution patterns

Glucosylation enhances the solubility and stability through the protection of reactive nucleophile groups, thus facilitating their storage, accumulation, and transport, as well as

regulating their bioavailability for other metabolic processes [6, 13-18]. Glucosyltransferases mediate glucosylation by transferring a glucose molecule from a nucleotide-activated sugar donor such as UDP-glucose onto acceptor molecules [15, 19]. To date, 98 families of GTs have been defined in CAZy (Carbohydrate-Active Enzymes) database (<http://www.cazy.org/GlycosylTransferases.html>) [20] based on the sequence identity score [21, 22]. Out of 98 families, family 1 GTs use UDP-activated sugar as donor in the enzymatic reactions [14]. One of the characteristic features of family 1 GTs is the presence of a 44-amino acid conserved motif at the C-terminal domain, known as the Plant Secondary Product Glycosyltransferase Box [23, 24]. This conserved motif has a role in forming the donor substrate binding pocket, whereas the acceptor substrate binding pocket is formed within the N-terminal domain [25-29]. These are inverting GTs and have GT-B fold topology [29].

Plant UGTs glycosylate a large number of diverse compounds *in vivo* which is evident from the enormous variety of glycosides found in plants [13]. Some UGTs are highly specific and glycosylate one or a few acceptor substrates [6, 30-32], whereas others have broad specificity and glycosylate a wide range of acceptor substrates [33-35]. A 3-*O*-GT from *Citrus paradisi* (grapefruit) is specific to flavonols and glucosylates in the 3-OH position [6]. *Vitis vinifera* 3-*O*-GT (VvGT1) glucosylates anthocyanidin as well as the flavonol kaempferol [28]. These two GTs from different plants have 56% sequence identity and 87% similarity.

Although plant UGTs have low sequence identity, they have highly similar secondary and tertiary structure [29]. The crystal structures of six plant UGTs have been published [27, 28, 35-39]. All of these have conserved tertiary structure. Bioinformatics have been extensively tested for ability to predict the substrate and regiospecificity of GTs based on primary sequences by generating computational models [10, 40, 41]. However, due to insufficient information on

the established biochemical function of GTs, the precise function of GTs cannot be predicted based on primary sequence and homology modeling [10, 42]. Biochemical characterization and study of acceptor substrate specificity *in vitro* remains the most effective way to study the function of plant UGTs [1, 10].

Our current research is focused on the structural and functional analysis of flavonol-specific Cp3GT previously characterized in our lab [6, 43]. Homology modeling, site-directed mutagenesis, analysis of acceptor substrate specificity and kinetic study of wild type and mutant Cp3GT were performed for the structure functional analysis of Cp3GT. Site directed mutagenesis is an effective method to identify the amino acids that are critical for enzyme activity and function. Multiple sequence alignment and protein homology modeling were used to identify the candidate residues for mutation. Both wild type Cp3GT and mutant P145T have been modeled using VvGT as template (Fig. 2.2). Molecular docking and modeling provides potential insight into the active site of the enzyme as well as basis for substrate recognition.

For this study, Cp3GT was modeled with a flavonoid 3-*O*-GT from *Vitis vinifera* (VvGT) that can glucosylate both flavonols and anthocyanidins. We identified a proline residue at position 145 of Cp3GT that corresponded to a threonine residue in VvGT (Fig. 2.3) and designed a Cp3GT-P145T mutant to test the hypothesis that that mutation of proline in Cp3GT to threonine could alter substrate and regiospecificity of Cp3GT.

The gold standard of structural determination is protein crystallization. Out of six different crystal structures of GTs published to date, VvGT glucosylates both anthocyanidin and flavonols [27]. In contrast, wild type Cp3GT has strict specificity towards flavonols which is different from other GTs crystallized to date. None crystal structures of GTs with strict specificity towards flavonols have been published to date. Along with the mutant P145T, there

are several other mutants of Cp3GT designed at our lab that had altered substrate and regiospecificity [44, 45]. Crystallization of wild type Cp3GT as well as some other mutations of our lab possessing altered substrate and regiospecificity could be a significant contribution towards understanding the acceptor substrate recognition pattern of flavonol specific GTs and structure function study. In order to crystalize proteins, the tags should be removed. We used site-directed mutagenesis to insert a thrombin cleavage site into the wild type and recombinant mutant P145T between the enzyme and the C-myc tags in C-terminal domain in order to be able to remove tags.

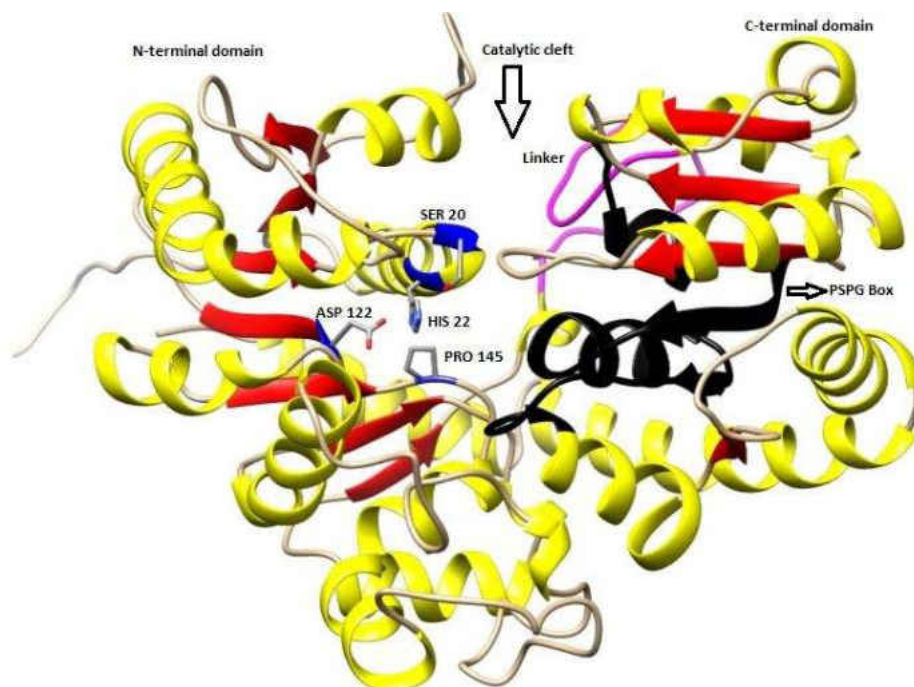


Fig. 2. 2 Structure of Cp3GT modelled with VvGT. Yellow, red, gold, and purple indicate the  $\alpha$ -helices,  $\beta$ -sheets, loops and the flexible linker region that connects N-terminal and C-terminal domain respectively. PSPG box is represented by black. The catalytic residues His22, Asp122, and Ser20 and the residue Pro145 chosen for mutation are represented by blue in the catalytic cleft.

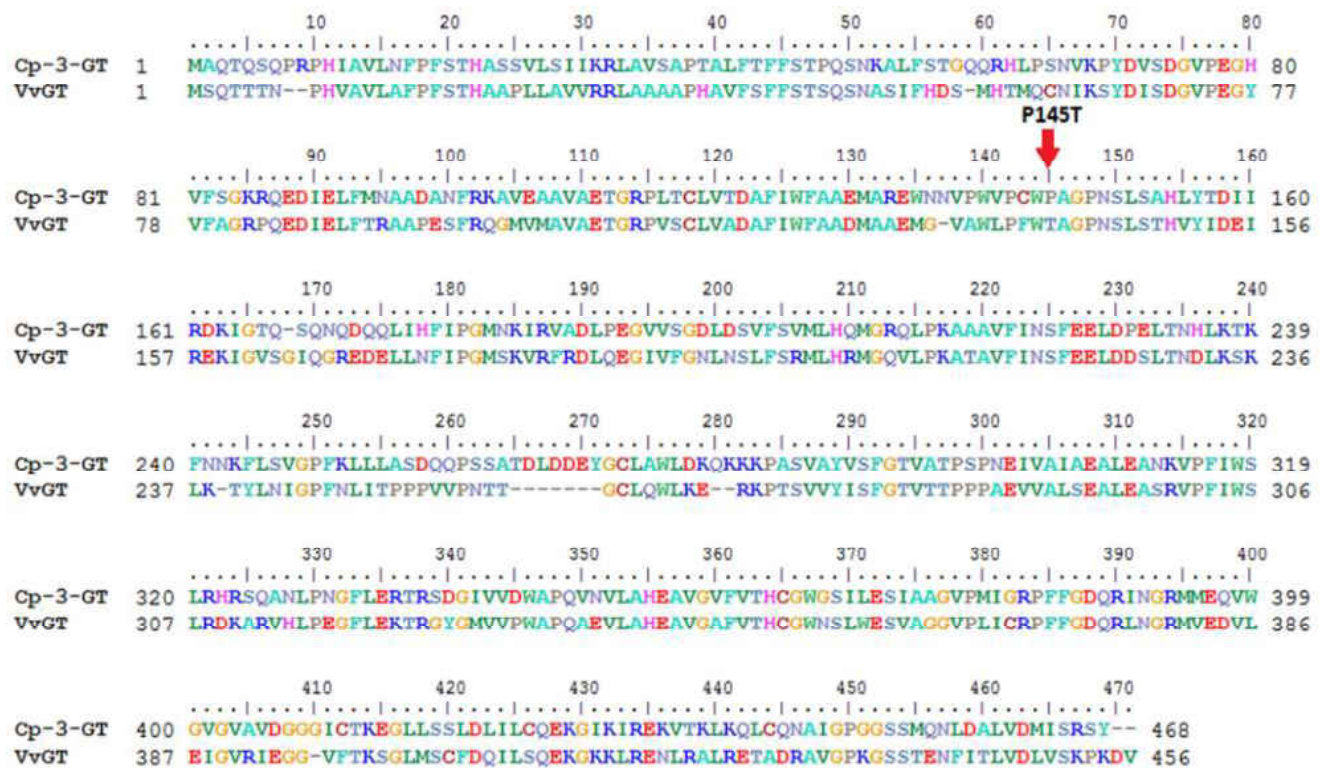


Fig. 2. 3 Multiple sequence alignment of Cp3GT and VvGT using Biomedit. Targeted mutation is shown with red arrow. Proline at position 145 of Cp3GT corresponds to threonine at position 141 of VvGT. Both of these are present within the N-terminal domain which is involved in forming the acceptor substrate binding pocket.

Recombinant proteins are cloned with C-myc and 6x-His tags at the C-terminus for ease in purification and to track proteins during the purification process. Because Cp3GT does not have an internal thrombin cleavage site, the most effective and efficient way to cleave off tags from our recombinant protein was to insert a thrombin site just in between our protein and the tags. This permits testing of tag-effects on the recombinant proteins [6], as well as cleaving off tags in order to crystallize proteins.

The recombinant enzyme Cp3GT from *E.coli* was tested for the effects of tag on enzyme activity and function [6]. The pCD1 vector harbors the thrombin site so the recombinant purified



enzyme which was expressed in *E.coli* using this vector was treated with thrombin enzyme to cleave off the N-terminal thioredoxin and 6x-His tags. The tags were successfully cleaved off after treating with thrombin enzyme [6]. To avoid the problem with inclusion bodies during expression and purification of recombinant protein using *E.coli*, Cp3GT was expressed in yeast, *Pichia pastoris* [43]. Presence or absence of tags had no effect on enzyme activity under similar assay conditions when it was expressed in *E.coli* [6]. The substrate preference for flavonols, effect of metals ions,  $K_i$  value for inhibitor UDP, temperature optimum and temperature stability were similar in both expression system [43]. The major difference observed was the increased stability at higher pH which could be due to the tags. In order to test the effects of tags while using yeast for recombinant expression of this protein, a site-directed mutagenesis approach was used to insert thrombin cleavage site into the recombinant wild type and mutant P145T.

In this study, hypothesis on the effect of mutating proline to threonine on the substrate and regiospecificity of Cp3GT was tested. Site-directed mutagenesis, homology modeling and biochemical testing of recombinant wild type Cp3GT and mutant P145T with different flavonoid substrates were done for structure function analysis of flavonol specific Cp3GT mutant P145T. *In silico* analysis was done to generate computational models and validated by *in vitro* results.

## **2. Materials and methods**

### **2.1. Reagents**

Quercetin, quercetin 3-*O*-glucoside, kaempferol 3-*O*-glucoside, naringenin and dihydroquercetin were purchased from Sigma (St. Louis, MO, USA); kaempferol, naringenin-7-*O*-glucoside, hesperetin, eriodictyol, isosakuranetin, apigenin, luteolin, diosmetin, scutallerein, fisetin, myricetin, gossypetin, cyanidin chloride, and 4'-acetoxy-7-hydroxy-6-methoxyflavonol

were purchased from Indofine (Hillsborough, NJ, USA); Phenylmethylsulfonyl fluoride (PMSF) was purchased from MP Biomedicals (Solon, OH). Acid washed glass beads (pore size-0.5mm), and ethidium bromide (EtBr) were purchased from Sigma. UDP-glucose was purchased from Calbiochem (Gibbstown, NJ, USA); UDP-[U-14C] glucose (specific activity 293mCi/mmol) was from Perkin Elmer, QuikChange Site-Directed Mutagenesis kit was purchased from Agilent Technologies, Inc. (Santa Clara, CA), GeneJET plasmid mini-prep kit was purchased from Thermo scientific, Wizard Plus Midipreps DNA purification System was purchased from Promega, restriction enzymes SacI and DpnI were purchased from Promega (Madison, WI), all other reagents were purchased from Fisher Scientific.

## 2.2. Scale up expression of recombinant wild type Cp3GT and mutant P145T

A 3mL aliquot of YPD media containing 100 µg/mL zeocin was inoculated with glycerol stock of yeast cells containing wild type Cp3GT or mutant P145T stored at -80°C and incubated overnight (16-18hrs) at 30°C at 250 rpm. A 500µL of overnight culture was transferred to a 1L flask containing 250mL of BMGY and incubated at 30°C, 250 rpm to an O.D.<sub>600</sub> of 2-6. The overnight culture was transferred to different 50mL centrifuge tubes and pellets were collected by centrifuging for 5 minutes at 2800 x g at room temperature and washed with 50mL of BMMY by suspending the cell pellets. The cell pellets were washed twice with BMMY media to remove glycerol. The washed cell pellets were resuspended in 250ml of freshly prepared BMMY in 1L baffled flask and incubated for 24hrs at 30°C at 250 rpm. After 24hrs (optimum time from time course induction study and western blot analysis), the overnight culture was transferred to different 50mL centrifuge tubes and cells were harvested by centrifuging for 5 minutes at 2800 x g at room temperature and stored at -80°C.

### 2.3. Protein extraction and purification

The cell pellets stored at -80°C were resuspended in 5-10 mL of breaking buffer (50mM sodium phosphate buffer at pH 7.5, 1mM EDTA, 5% glycerol) containing 5 mM  $\beta$ ME and 1 mM PMSF. Cells were lysed using French press at 1120 psi pressure. The cells were put in ice during entire lysis process. Before using French press, it was first washed with distilled water and rinsed with breaking buffer twice to ensure it was free from other proteins. Lysis was repeated for at least 4 cycles ensuring the yeast cells were broken efficiently. The lysed cells were centrifuged at 13,000 x g for 20 minutes at 4 °C and the supernatant was collected and kept on ice.

PD-10 columns for buffer exchange were equilibrated with 25mL equilibration buffer (50 mM sodium phosphate buffer at pH 7.5, 300 mM sodium chloride) containing 5mM  $\beta$ ME. A 2.5 mL volume of the crude protein was subjected to buffer exchange with PD-10 column and the flow through the PD-10 was discarded. A 3.5mL volume of the equilibration buffer was used to elute the protein from PD-10 column. The PD-10 column was again equilibrated and the remaining crude protein was eluted with 3.5mL of equilibration buffer as described. A 2mL bed volume TALON<sup>R</sup> IMAC cobalt metal affinity resin column was equilibrated with 25mL of equilibration buffer containing 5mM  $\beta$ ME prior to use. The 7.5mL of eluted protein from the PD-10 column was loaded on a 2mL bed volume TALON<sup>R</sup> IMAC cobalt metal affinity resin column at a flow rate of 0.5 mL min<sup>-1</sup> and discarded the flow through. The column was washed with equilibration buffer until the O.D.<sub>280</sub> of the eluate reached almost zero. Equilibration buffer was used as blank to measure the O.D. of the washed fractions. Once all the non-specific proteins were washed off from the column, the bound protein was eluted in 2 mL fractions with elution buffer (50 mM sodium phosphate buffer pH 7.5, 300 mM sodium chloride, 5 mM  $\beta$ ME,

and 150 mM imidazole). The protein concentration of each eluted fraction was estimated using Nanodrop at 280nm. The fractions with higher concentrations of proteins were pooled, concentrated, and desalted using Amicon Centricon 30 MWCO centrifugal filters (Millipore, Billerica, MA, USA) by centrifuging at 2800 x g at 4°C for 10 to 30min until the volume reached 500µL. After the volume reached 500µL, 2mL of assay buffer (50 mM sodium phosphate buffer pH 7.5, 14 mM βME) was added in the Amicon Centricon Centrifugal Filter and centrifuged again to remove the imidazole that was used for elution. The concentrated protein was collected and concentration was measured and kept on ice for further assay and biochemical testing. A 50µL aliquot of each protein fraction from desalting, flow through, washing, and elution were collected for western blot analysis.

#### 2.4. Glucosyltransferase activity

Glucosyltransferase activity was assayed by determining the incorporation of <sup>14</sup>C-glucose into the reaction product [6, 29]. For initial screening, reactions were performed using 0.025 µCi of UDP-[U-<sup>14</sup>C] glucose (specific activity of 263mCi/mmol) diluted to 20,000cpm/10uL in water. Each reaction for activity screening was a 75µL reaction mixture containing 5 µL of flavonoid aglycone (50nmol/5µL) in ethylene glycol monomethyl ether, 10 µL of sugar donor (UDP-<sup>14</sup>C glucose), 50 µL of 50mM phosphate buffer (pH 7.5) containing 14 mM βME, and 10µL of recombinant enzyme (3µg/10µL). The reactions were incubated in a 37°C water bath for 5 minutes. The reactions were terminated by the addition of 15 µL of 6 M HCl. The product of the reaction was then extracted using 250 µL of EtOAc and vigorous mixing, and centrifuging for few seconds. A 150 µL aliquot of the EtOAc extract was placed in 2 mL of CytoScint scintillation cocktail (Thermo Fisher) and incorporation measured by a Beckman LS 6500 scintillation counter.

For other biochemical analysis and kinetics, each standard 75 $\mu$ L reaction contained 5 $\mu$ L of flavonoid aglycone (50nmol/5 $\mu$ L) in ethylene glycol monomethyl ether, 50  $\mu$ L of 50mM phosphate buffer (pH 7.5) containing 14 mM  $\beta$ ME, 10 $\mu$ L UDP-glucose (100nmol/10 $\mu$ L) having 50,000cpm/10 $\mu$ L, and 10 $\mu$ L of 0.5 $\mu$ g/10 $\mu$ L enzyme (linearity obtained with 0.5 $\mu$ g of enzyme), The reactions were incubated at 37 °C for 10 minutes unless otherwise described. The reactions were terminated by the addition of 15  $\mu$ L of 6 M HCl. The products of the reaction was then extracted using 250  $\mu$ L of EtOAc and vigorous mixing, and spinning for few seconds. A 150  $\mu$ L aliquot of the EtOAc extract was placed in 2 mL of CytoScint scintillation cocktail (Thermo Fisher) and incorporation measured by a Beckman LS 6500 scintillation counter. Incorporation of the radioactive glucose in the reaction product was calculated by multiplying cpm measured with 1.67 to obtain total cpm incorporated. Nanomoles of products are determined by cpm incorporated divided by 500.

## 2.5.Product identification

The standard 150 $\mu$ L reaction for product identification contained 10 $\mu$ L of aglycone (100nmols/10 $\mu$ L) in ethylene glycol monomethyl ether, 100  $\mu$ L of 50mM phosphate buffer (pH 7.5) containing 14 mM  $\beta$ ME, 20 $\mu$ L of non-radioactive UDP-glucose (200nmol/20 $\mu$ L), and 20  $\mu$ L of enzyme (6-10 $\mu$ g). The reactions were incubated for 30-60 minutes at 37°C to ensure sufficient product formation for analysis. The reactions were terminated by the addition of 15  $\mu$ L of 6 M HCl. The products were extracted using 500  $\mu$ L of ethyl acetate and dried under a stream of N<sub>2</sub>. The residues were resuspended in HPLC grade methanol (60  $\mu$ L) and 10  $\mu$ L injected for reversed-phase HPLC analysis on a Waters Breeze HPLC system composed of an in-line degasser AF, a binary HPLC pump 1525, and a dual  $\lambda$  absorbance detector 2487 operated by Breeze software version 3.30 (Waters, Milford, MA, USA). Samples were fractionated at room

temperature using a Bridge C18,  $4.6 \times 150$  mm column at a flow rate of  $1.0 \text{ mL min}^{-1}$ .

Fractionation was achieved with HOAc/H<sub>2</sub>O (15:85) as the mobile phase and a linear gradient of HPLC grade CH<sub>3</sub>CN (5–65%) over 25 min as the organic phase. Chromatograms were detected at A<sub>290</sub> and A<sub>365</sub> for analysis of the product of the reaction. Retention time of reaction products were compared with authentic standards.

## 2.6. Homology modeling and docking

Homology modeling was performed using the solved crystal structure of the *Vitis vinifera* flavonoid-3-*O*-glucosyltransferase as template (Offen et al. 2006). UCSF Chimera (<https://www.cgl.ucsf.edu/chimera/>) was used to generate the 3-D structure of protein. EasyModeller 4.0 [46] was used to generate models having different molpdf (molecular probability distribution), DOPE (Discrete Optimized Protein Energy) and GA341 scores. The model with lowest DOPE score was considered as best fit model [46] and was used for further docking analysis. The model was refined by energy minimization for molecular docking by selecting all the residues using Swiss PDB viewer (<http://spdbv.vital-it.ch/>). This generates the energy minimized model to the most stable state. The best model was superimposed with the template to check similarities and differences in the structure among loop as well as  $\alpha$ -helix and  $\beta$ -sheets using Chimera. This energy minimized model was used with PyRx (<http://pyrx.sourceforge.net/>) to dock the protein with various ligands of choice. The protein model generated by Swiss PDB viewer and a reference ligand (different flavonoid aglycones, downloaded from PDB database) was added in the system before running docking program (PyRx). Then the models were analyzed using Autodock 1.5.6 of MGL tools software

(autodock.scripps.edu). When model with lowest DOPE score did not agree with experimental results, additional models were examined.

## 2.7. Insertion of thrombin site

### 2.7.1. Primer design

Thrombin recognizes the consensus sequence leucine-valine-proline-arginine-glycine-serine, cleaving the peptide bond between arginine and glycine. Primers specific for thrombin site were designed (Supplementary Table B.1) using Quickchange Primer Design web tool from Agilent Technologies.

([http://www.genomics.agilent.com/primerDesignProgram.jsp?&\\_requestid=994370](http://www.genomics.agilent.com/primerDesignProgram.jsp?&_requestid=994370)). Site-directed mutagenesis PCR was performed to insert the thrombin site between the wild type Cp3GT and mutant P145T enzymes and the C-myc tags.

### 2.7.2. Site-directed mutagenesis

Site-directed mutagenesis PCR was performed to insert thrombin site in between the wild type Cp3GT and mutant P145T enzyme and the C-myc tags using QuikChange<sup>R</sup> Lightning Site-Directed Mutagenesis Kit from Agilent Technologies. The reaction mixture contained:

#### P145T-thrombin cleavage site

5μL of 10X reaction buffer

1.25μL of 125ng/μL Primer (forward)

1.25μL of 125ng/μL Primer (reverse)

1μL of dNTP mix

1.5μL of Quick solution reagent

1 μL of Quick change DNA polymerase enzyme

39μL of P145T DNA template (45ng)

#### Cp3GT-thrombin cleavage site

5 $\mu$ L of 10X reaction buffer

1.25 $\mu$ L of 125ng/ $\mu$ L Primer (forward)

1.25 $\mu$ L of 125ng/ $\mu$ L Primer (reverse)

1 $\mu$ L of dNTP mix

1.5 $\mu$ L of Quick solution reagent

1  $\mu$ L of Quick change DNA polymerase enzyme

39 $\mu$ L of Cp3GT DNA template (35ng)

PCR was run using the following conditions as mentioned on the QuikChange<sup>R</sup> Lightning Site- Directed Mutagenesis Protocol from Agilent Technologies.

Table 2. 2 PCR conditions for insertion of thrombin site using site directed mutagenesis.

Step	Temperature ( <sup>o</sup> C)	Time(s)	Cycles
Initiation	95	30	1
Denaturation	95	30	16
Annealing	55	60	
Elongation	68	600	
Extension	68	600	1

#### 2.7.3. Heterologous expression

The site directed mutagenesis PCR product was treated with the restriction enzyme DpnI (Agilent Technologies) to digest the parental template DNA (Supplementary Fig. C.3) and the wild type Cp3GT and mutant P145T plasmid with thrombin site was transformed into Top 10 competent *E.coli* cells (Invitrogen) by the heat shock method. The transformants were grown on low-salt LB plates containing 25 $\mu$ g/mL zeocin and incubated at 37 °C. Positive colonies were



identified by colony-PCR using vector specific AOX primers (Fig 2.21) and restreaked on low-salt LB plates containing 25µg/mL zeocin and incubated at 37 °C. Plasmids were extracted using Quantum Prep™ Plasmid Miniprep Kit (Invitrogen) and the presence of the thrombin site in-frame was confirmed by sequencing (Fig. 2.22-2.23). Plasmid were extracted using Quantum Prep™ Plasmid Midiprep Kit (Invitrogen) and linearized using SacI enzyme (Promega) to transform into competent *Pichia pastoris* (Mut+) cells by electroporation. Transformation into yeast is in progress.

### **3. Results and discussion**

#### **3.1. Acceptor substrate specificity**

The purified wild type Cp3GT and mutant P145T activities were screened with 15 different aglycone substrates belonging to six different flavonoid subclasses (Table 2.1). Wild type Cp3GT is flavonol specific and adds sugar at 3-OH only and it glucosylates quercetin to the highest degree (Owens and McIntosh, 2009). In addition to glucosylating flavonols, the mutant P145T glucosylated flavones and flavanones in the screening assay (Table 2.1). This is significant because flavanones and flavones do not contain a 3-OH group for glucosylation. In contrast to the wild type Cp3GT enzyme, the mutant P145T enzyme had 2X higher level of activity with kaempferol as a substrate. In the screening assay, wild type Cp3GT had no activity with flavonoid substrates other than flavonols (Owens and McIntosh 2009). However, mutant P145T had activity with naringenin, hesperetin and scutallerein (Table 2.1).

These results suggest that the mutant P145T substrate specificity was broadened to accept different classes of flavonoids as well as regiospecificity in terms of attaching the sugar at

different hydroxyl positions. Because this is a hypersensitive assay and it might give false positive results, results need confirmation by kinetic assay.

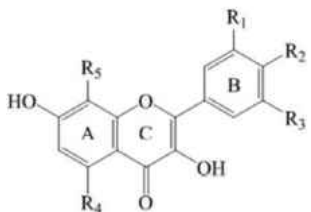
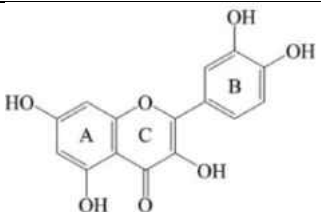
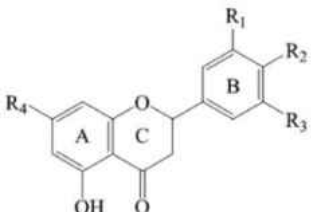
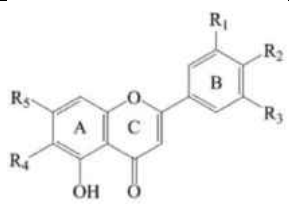
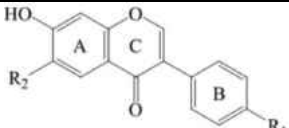
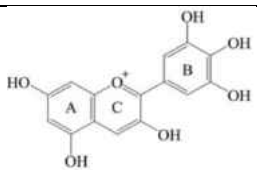
### 3.2. Product identification

High Performance Liquid Chromatography (HPLC) was performed to identify the products of the reaction catalyzed by the mutant P145T as in Owens and McIntosh, 2009. Additionally, product identification verified the activity screening results. Retention times of the reaction products were compared with those of standard compounds to identify the products of the reaction using quercetin, kaempferol and naringenin as acceptor substrates. The products of the reaction were quercetin-3-*O*-glucoside (Fig. 2.4), kaempferol-3-*O*-glucoside (Fig. 2.5) and naringenin-7-*O*-glucoside (Fig. 2.6) with quercetin, kaempferol and naringenin as acceptor substrate, respectively. As has been previously observed [6, 42], naringenin was not glucosylated by wild type Cp3GT.

The mutant P145T glucosylated quercetin at the 3-OH group forming quercetin-3-*O*-glucoside. Similarly, the mutant P145T glucosylated kaempferol at the 3-OH position forming kaempferol-3-*O*-glucoside. Naringenin does not contain a 3-OH group for glucosylation. However, HPLC chromatograms showed that the mutant P145T enzyme glucosylated naringenin at the 7-OH group forming prunin. No glucoside peak was obtained for the reaction catalyzed by wild type Cp3GT using naringenin as acceptor substrate. These results suggested that mutant P145T had broadened substrate and regiospecificity of Cp3GT and also confirmed the initial screening assay. These results indicate that a single point mutation altered the substrate and regiospecificity of Cp3GT. This might be due to the structural change in the protein caused by a single mutation that caused an alteration in the activity and functioning of the mutant protein.

Table 2. 1 Activity Screening of Wild Type Cp3GT and Mutant P145T.

Activity results of each acceptor substrates are expressed as percentage relative activity with respect to glucosylation of quercetin by wild type enzyme.

Flavonoid subclass	Flavonoid substrate	Base Structure	Relative activity	
			WT	P145T
<b>Flavonol</b>	Quercetin: R1=R2=R4=OH; R3=R5=H		100	90
	Kaempferol: R1=R3=R5=H; R2=R4=OH		52.5	119
	Fisetin: RA=R4=R5=H; R2=R3=OH		42	54
	Gossypetin: R1=R2=R4=R5=OH; R3=H		25	42.5
<b>Dihydroflavonol</b>	Dihydroquercetin		*	*
<b>Flavanone</b>	Naringenin: R1=R3=H; R2=R4=OH		9	25
	Hesperetin: R1=H; R2=OCH3; R3=R4=OH		6.5	34
	Eriodictyol: R1=R2=R4=OH; R3=H		2	5
	Isosakuranetin: R1=R3=H; R2=OCH3; R4=OH		8	6
<b>Flavone</b>	Luteolin: R1=R2=R5=OH; R3=R4=H		1	7
	Apigenin: R1=R3=R4=H; R2=R5=OH		1	2.5
	Diosmetin: R1=R5=OH; R2=OCH3; R3=R4=H		5.5	12
	Scutallerein: R1=R3=H; R2=R4=OH; R5=O-glu		11.5	33
<b>Isoflavone</b>	4'-acetoxy-7-hydroxy-6-methoxy isoflavone R1=COOH; R2= OCH3		6.5	10.5
<b>Anthocyanidin</b>	Cyanidin chloride		1.5	4.5

\*Needs confirmation by a regular assay

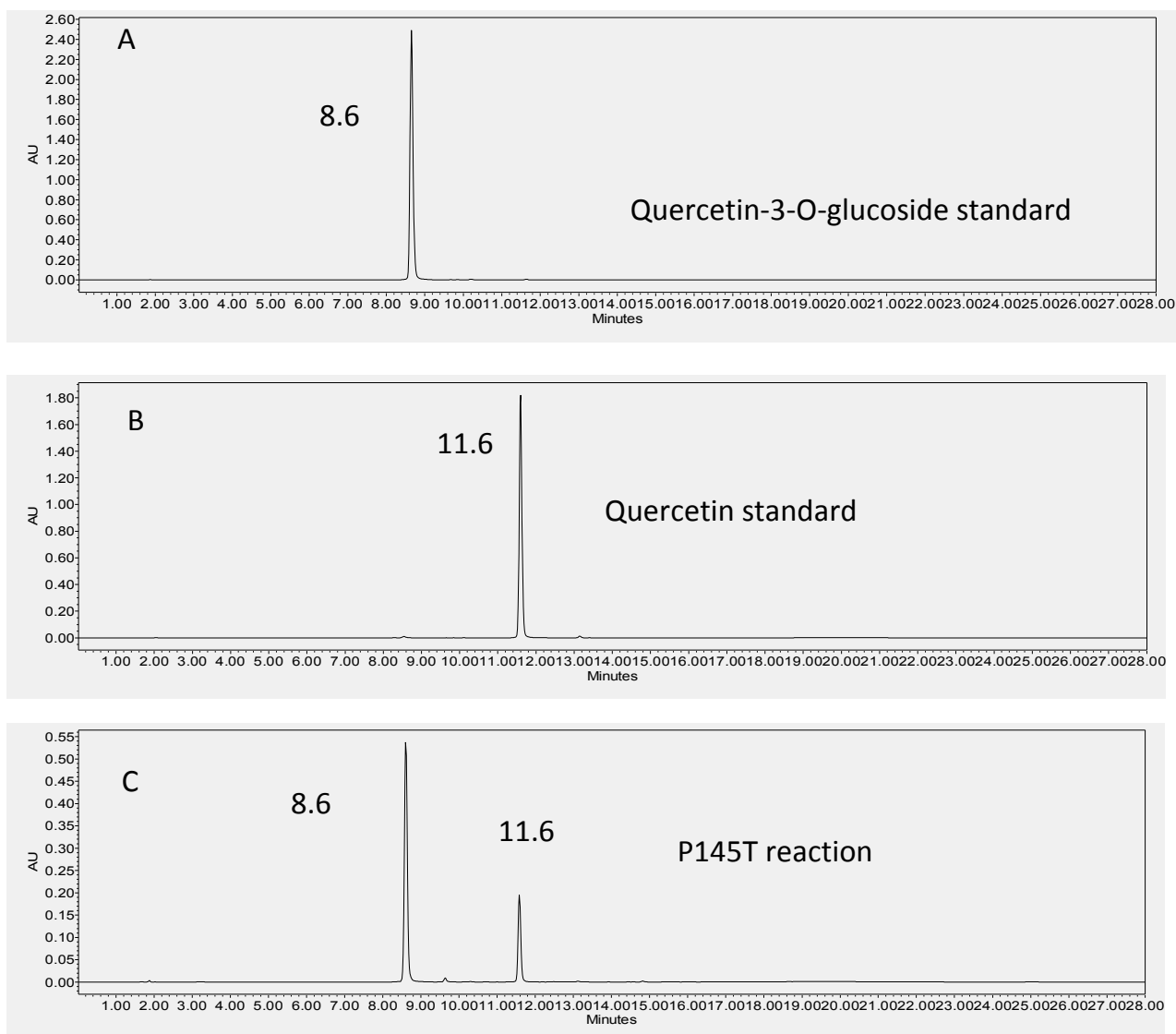


Fig. 2. 4 Reaction product identification of mutant P145T with acceptor quercetin. HPLC chromatograms of A) standard quercetin-3-*O*-glucoside, retention time 8.6 minutes, B) standard quercetin, retention time 11.6 minutes, and C) products of the reaction catalyzed by recombinant mutant P145T. Retention times (minutes) are indicated by the peaks.

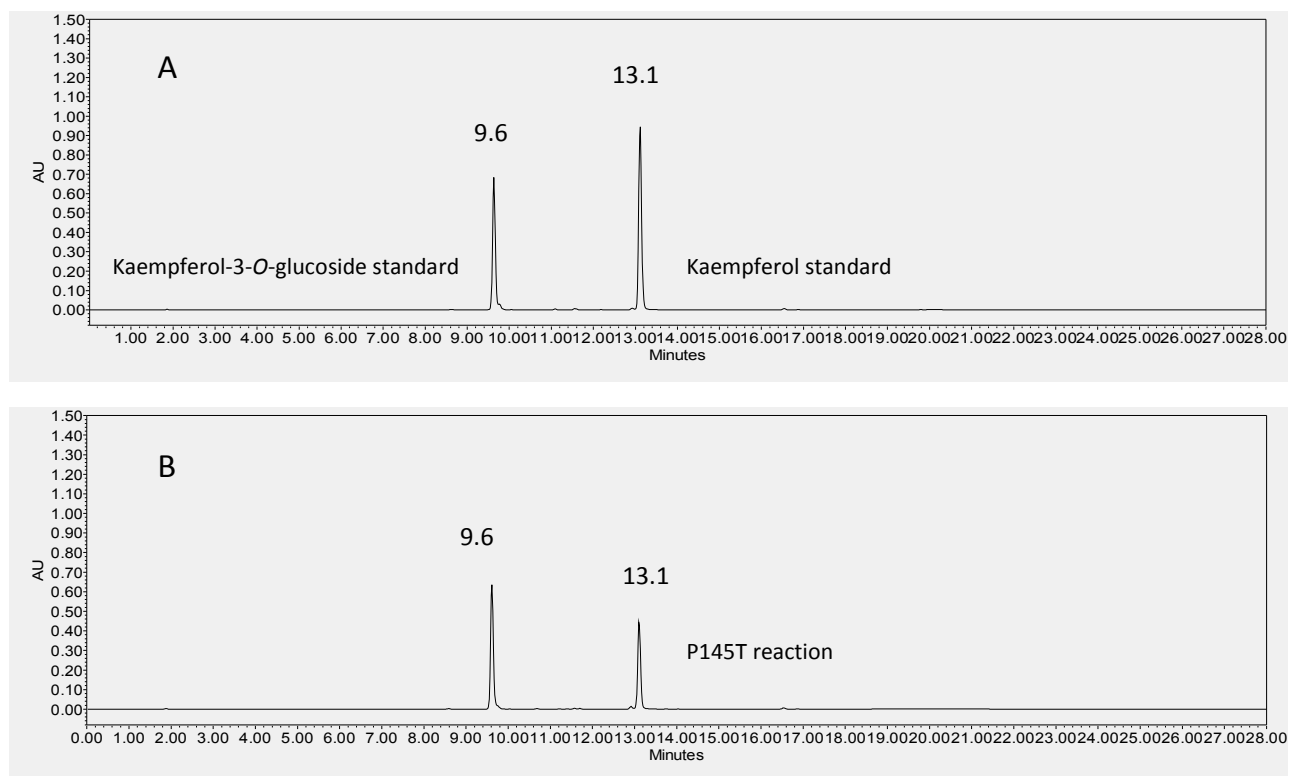


Fig. 2. 5 Reaction product identification of mutant P145T with acceptor kaempferol. HPLC chromatograms of A) Standard kaempferol-3-*O*-glucoside, retention time 9.6 minutes, and standard kaempferol, retention time 13.1 minutes, and B) the products of the reaction catalyzed by recombinant mutant P145T. Retention times (minutes) are indicated by the peaks.

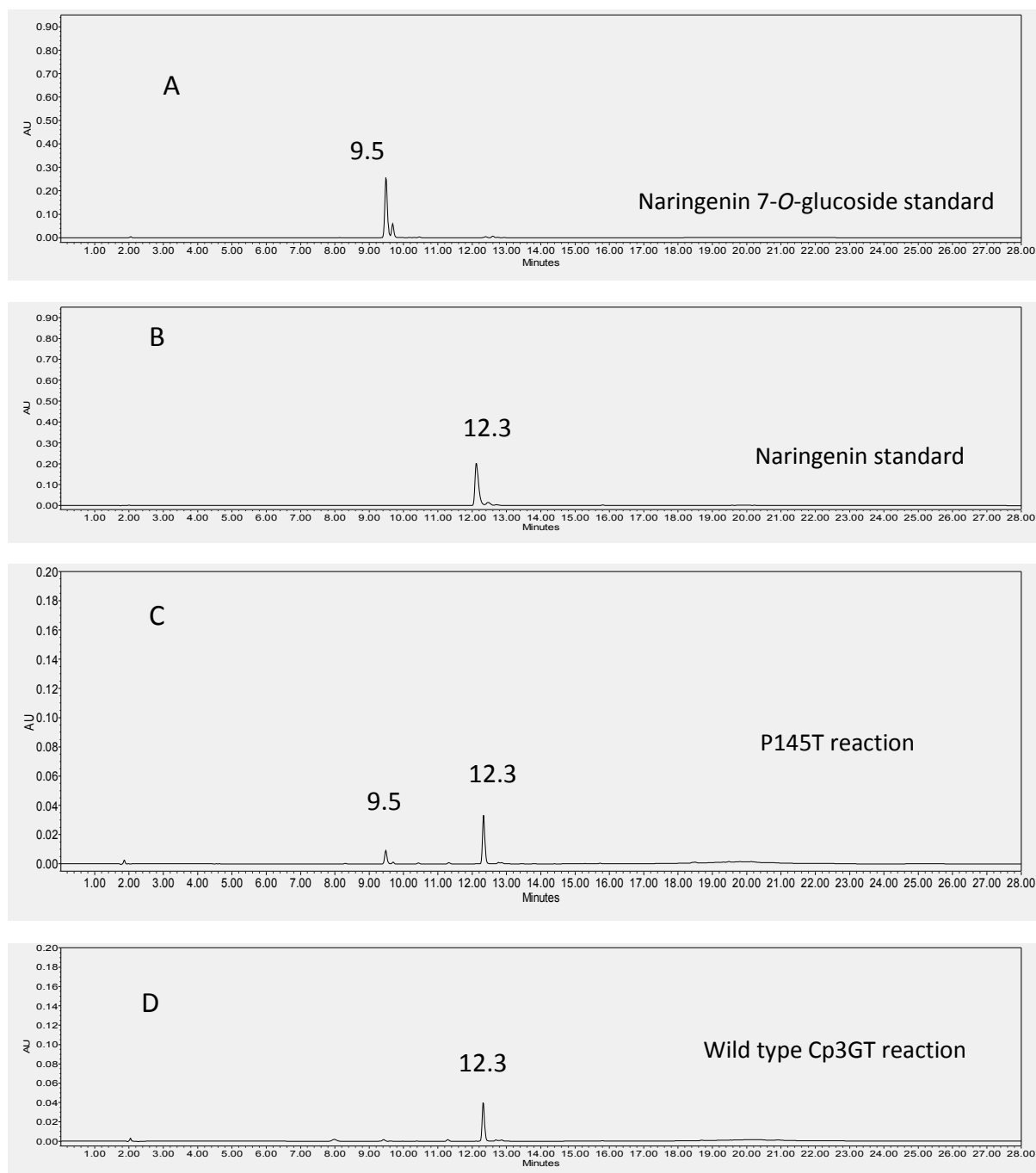


Fig. 2. 6 Reaction product identification with acceptor naringenin. HPLC chromatograms of A) naringenin 7-*O*-glucoside standard, retention time 9.5 minutes, B) naringenin standard, retention time 12.2 minutes, C) reaction products catalyzed by mutant P145T, and D) reaction products catalyzed by wild type Cp3GT enzyme. Retention times (minutes) are indicated by the peaks.

Homology modeling and molecular docking were performed to provide insight into possible structural changes brought about by the P145T mutation.

### 3.3. Enzyme characterization

Wild type and mutant P145T were characterized using quercetin and kaempferol as acceptor substrates. Wild type had higher percentage relative activity with quercetin although glucosylates kaempferol whereas the mutant P145T had higher percentage relative activity with kaempferol although it glucosylates quercetin as well (Table 2.1).

#### 3.3.1. Optimal reaction condition

##### 3.3.1.1. Enzyme concentration

Three different concentrations of enzymes were taken for reaction to test the effect of enzyme concentration on the rate of reaction. The reaction mixture was incubated at 37°C at different time intervals from 0, 10, 20, 30, 60 and 90 minutes. With 0.25 µg of enzyme, the reaction was linear for 15 minutes for both wildtype and mutant enzyme with quercetin. Similarly, the reaction was linear for 15 minutes with 0.5 µg of enzyme for both wild type and mutant P145T enzyme with kaempferol. For subsequent kinetic analyses with quercetin, 0.25 µg of enzyme was used to incubate reaction mixture for 10 minutes and with kaempferol 0.5 µg of enzyme was incubated for 10 minutes.

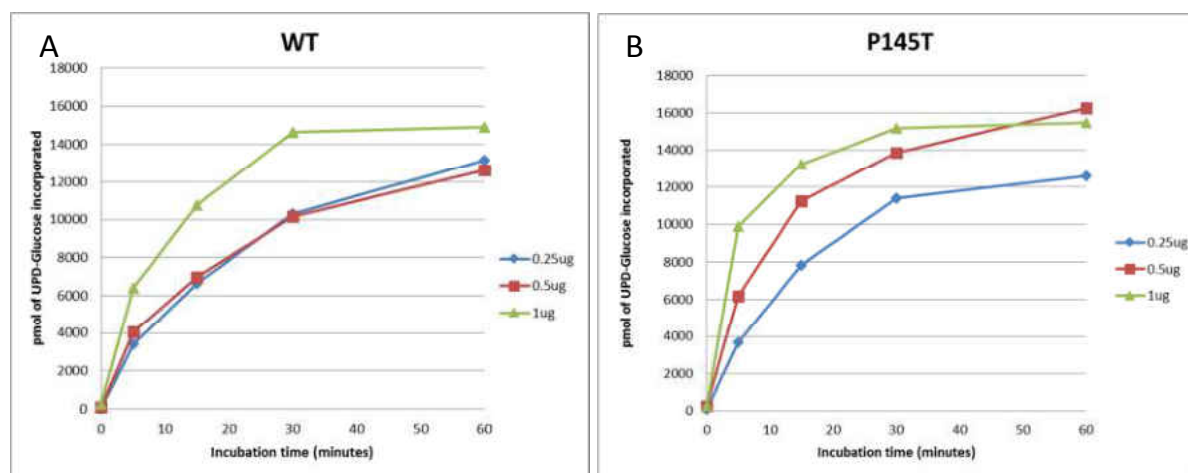


Fig. 2. 7 Time course study with different concentration of enzyme with quercetin. A) WT- wild type Cp3GT, B) mutant 145T. The reactions were linear for 15 minutes with 0.25 µg enzyme for both wild type Cp3GT and mutant P145T.

Activity screening results suggested that the wild type Cp3GT and mutant P145T had almost similar activity (Table 2.1) that needed further confirmation by a regular assay. The pmoles of UDP-glucose incorporated into the reaction product after incubating 0.25 µg of wild type Cp3GT for 30 minutes (Fig. 2.7, A) was almost equal to the pmoles of UDP-glucose incorporated into the reaction product after incubating 0.25 µg of mutant P145T for 30 minutes (Fig. 2.7, B). As a time course study is a regular biochemical assay, this confirmed the activity screening results with quercetin for both wild type Cp3GT and mutant P145T.

Time course study with kaempferol further confirmed activity screening assay for kaempferol. The mutant had almost 119% relative activity with kaempferol whereas wild type Cp3GT had almost 53% relative activity (Table 2.1). The time course study showed that the pmoles of UDP-glucose incorporated into the reaction product after incubating 0.5 µg of mutant P145T for 30 minutes is two times higher than the pmoles of UDP-glucose incorporated into the reaction product after incubating 0.5 µg of wild type Cp3GT enzyme. This suggests that an



increase in the activity of mutant P145T is due to the structural change brought about by a single point mutation.

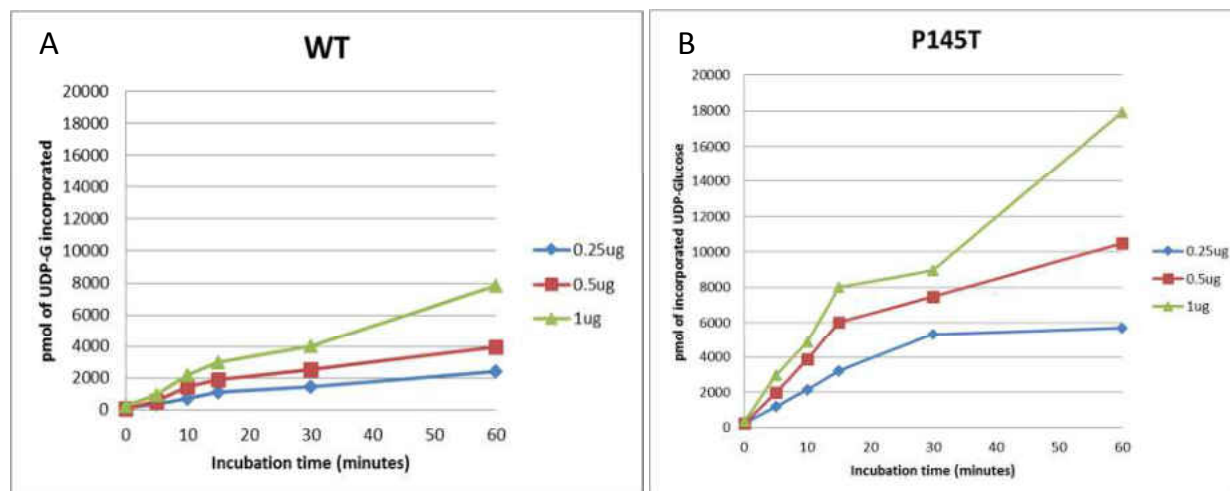


Fig. 2. 8 Time course study with different concentration of enzyme with kaempferol. A) WT-wild type Cp3GT, B) mutant 145T. The reactions were linear for 15 minutes with 0.5 µg enzyme for both wild type Cp3GT and mutant P145T.

### 3.3.1.2. Optimal temperature

The optimal reaction temperatures for wildtype and mutant P145T were determined by incubating the reaction mixture at different temperatures; 20, 30, 40, 50, 60, 70, and 80 °C

### A) Quercetin

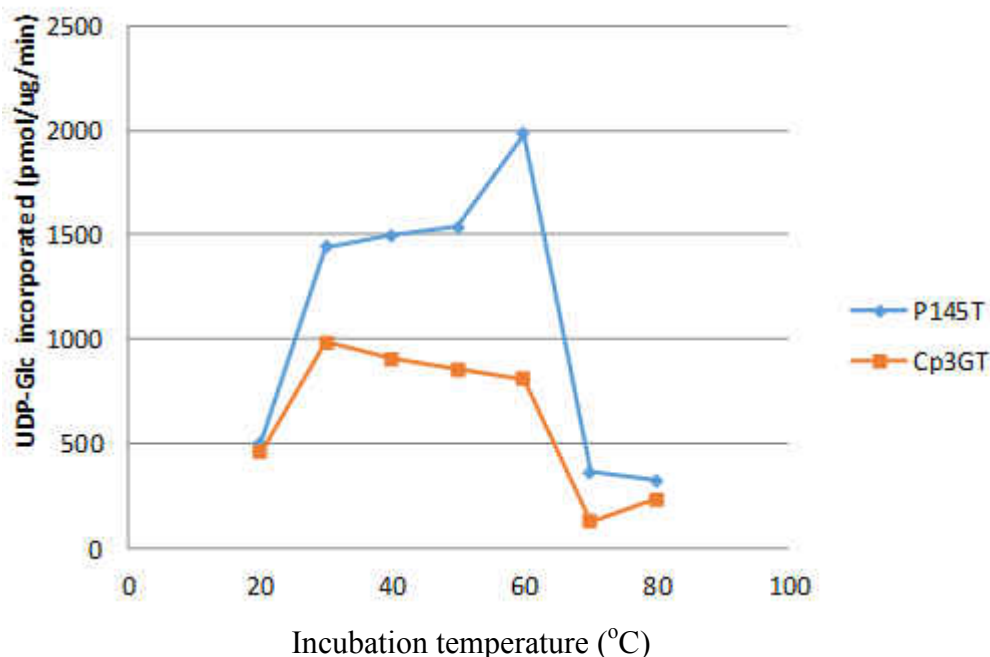


Fig. 2. 9 Effect of temperature on enzyme activity with quercetin. A reaction mixture of 75  $\mu$ L was incubated for 10 minutes at different temperatures with 0.25  $\mu$ g of enzyme. The optimum temperature for wild type Cp3GT was 30°C and for mutant 1-point peak was obtained for P145T at 60°C.

### B) Kaempferol

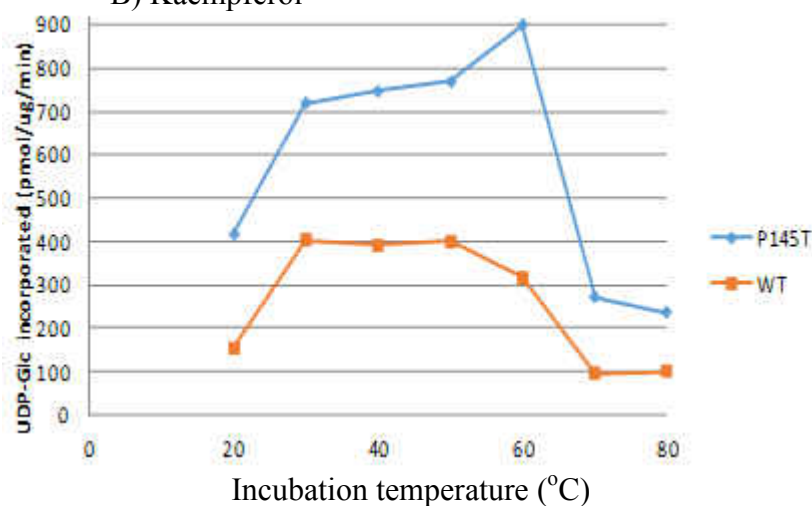


Fig. 2. 10 Effect of temperature on enzyme activity with kaempferol. A reaction mixture of 75  $\mu$ L was incubated for 10 minutes at different temperatures with 0.5  $\mu$ g of enzyme. The mutant P145T had optimal activity at 50-60°C whereas wild type Cp3GT had almost a broad maximal activity from 30-50°C.

A sharp peak was obtained for mutant P145T at 60°C with both quercetin and kaempferol as acceptor substrate. These results need to be confirmed as one sharp peak was observed at 60°C. The optimum temperature of wild type Cp3GT with quercetin was found to be 30°C as reported earlier [6], whereas the wild type enzyme showed almost constant activity from 30 °C to 50°C with kaempferol. The higher optimum temperature of mutant P145T could be due to single point mutation from proline to threonine that caused a change in the structure of the mutant P145T protein. Proline restricts the free rotation of the protein due to its cyclic side chain thereby affecting the interactions with other residues. However, as threonine has free side chains that might allow protein to rotate freely and involve in non-covalent interactions with other residues at the active side. This might cause a change in the structure of protein.

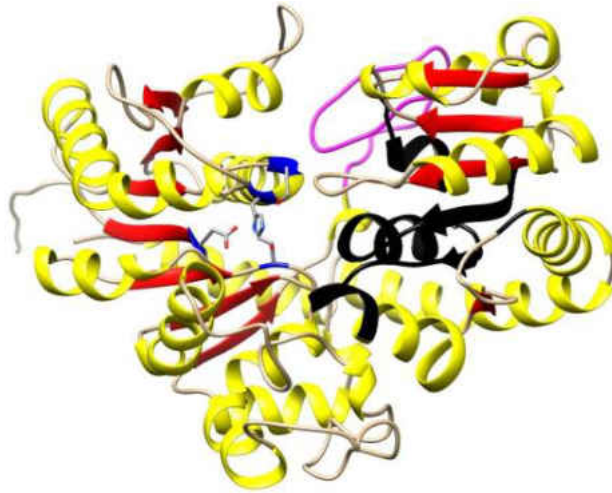
Activity screening suggested that mutant P145T had almost 2X higher relative activity with kaempferol (Table 2.1). The activity of mutant P145T was confirmed by the time course study (Fig. 2.10). This needs further confirmation with detailed kinetic study of the activity.

#### 3.4. Structural similarities and differences between wild type Cp3GT and P145T.

The 3-D structure of a protein gives an insight toward residues critical for the functioning of the protein as well as the residues that are present within the active site. The recombinant mutant P145T and wild type Cp3GT protein was modeled with *Vitis vinifera* 3-O-GT. VvGT glucosylates both flavonols and anthocyanidins [27]. It has 56% sequence identity and 83% homology with Cp3GT. Homology modeling was performed to predict the 3-D structure of recombinant wild type Cp3OGT and mutant P145T using different computational tools. Comparison of the 3-D structures with the crystal structure of VvGT reveals conserved 3-D

structures. Despite having 83% sequence similarity and 56% identity, Cp3GT and VvGT have similar 3-D structure. However, some differences in the structure were apparent (Fig. 2.12).

A.



B.

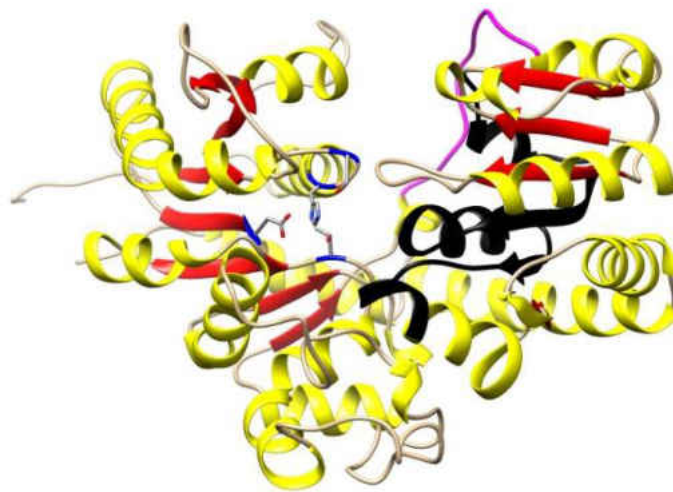


Fig. 2. 11 Homology modeling of mutant P145T with VvGT. A) Predicted tertiary structure of mutant P145T, and B) Tertiary structure of VvGT. Red represents the  $\beta$ -strands, yellow represents the  $\alpha$ -helices, black represents the PSPG box, and purple represents the flexible linker that connects N-terminal and C-terminal domain. Structures generated using Chimera1.3 (<https://www.cgl.ucsf.edu/chimera/>).

The tertiary structure of wild type Cp3GT and mutant P145T are made up of  $\alpha$ -helices,  $\beta$ -strands and coils and is globally similar with tertiary structure of VvGT (Fig. 2.2, 2.11). Like all other GTs of Family 1, the wild type Cp3GT and mutant P145T models showed a common GT-B fold topology that consisted of the two Rossmann-like domains that faced each other linked by a flexible linker region. The N-terminal region consisted of seven  $\beta$ -strands facing towards the catalytic cleft and the C-terminal domain had six  $\beta$ -strands. These secondary structures faced each other towards the center of the catalytic cleft. The predicted active site is formed in a cleft between the N-terminal domain and C-terminal domain linked by a flexible loop (Leu 254 to Gly 271). The PSPG box at the C-terminal domain had conserved amino acid residues that extend from tryptophan 345 to arginine 389.

Comparison of the 3-D model structure of wild type Cp3GT and mutant P145T has revealed several structural similarities and differences (Fig. 2.12). The overall 3-D structure of wild type Cp3GT and the mutant P145T is conserved. A difference in the structure due to a change in single amino acid is seen in the catalytic cleft shown by the red circle (Fig. 2.12 B and C). Acceptor and donor substrate binding pocket is formed in the cleft between N-terminal and C-terminal domain. The loop region (Leu254 to Gly271) as represented by purple color (Fig. 2.12) connects the N-terminal and C-terminal domains. A difference was observed in the position of the loop region in all of the predicted structures. Other differences in the orientation of helices and loops are indicated by arrows. These models suggested that a mutation of proline to threonine is causing structural change in the tertiary structure of protein. As proline's side chain restricts the rotation of the protein, substituting proline to threonine brought structural change in the protein model. This structural change might be the cause of altered substrate and regiospecificity as observed in experimental results.

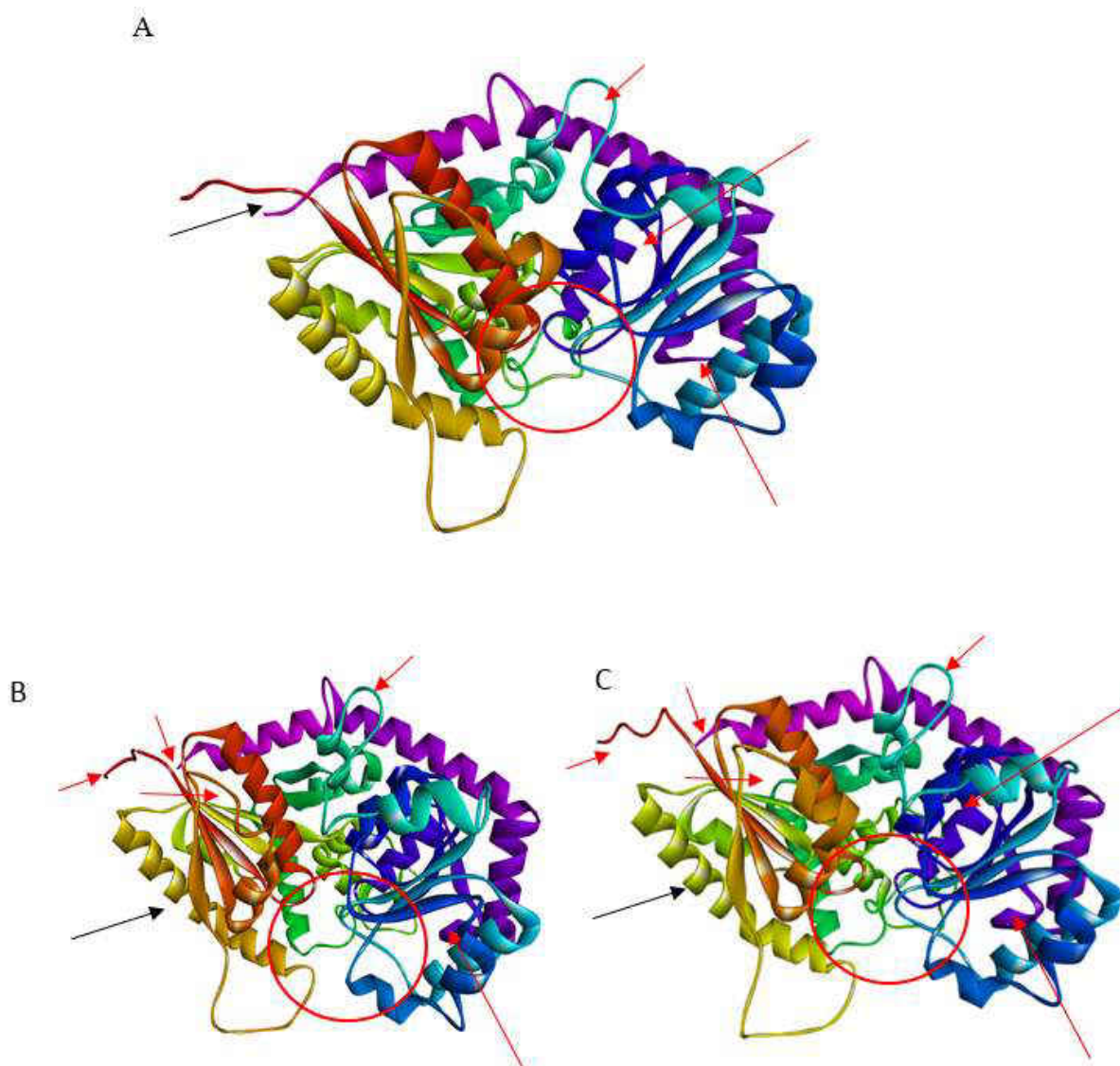


Fig. 2. 12 Structural similarities and differences. A) VvGT, B) wild type Cp3GT, and C) mutant P145T generated using EasyModeller by using VvGT as template. Differences were observed in the catalytic cleft that forms active site for the substrates. Red arrow and red circle indicate differences; black arrow indicates similarities. Differences were also observed in the loop region that connects the N-terminus and C-terminus.

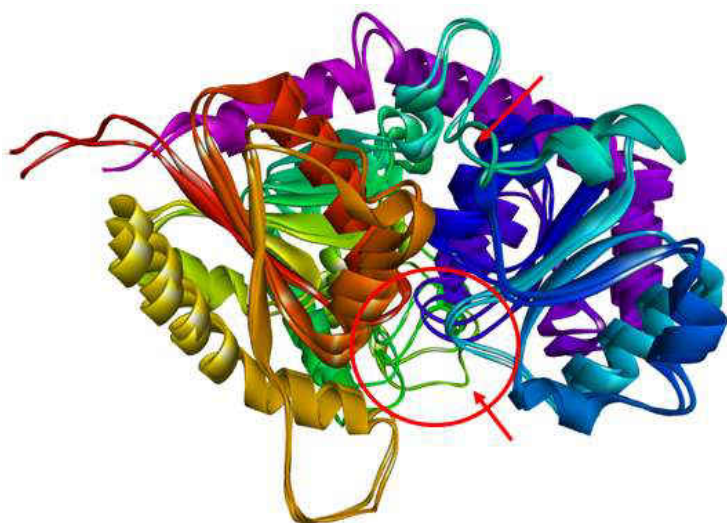


Fig. 2. 13 Superimposition of VvGT and Cp3GT. The 3-D structure of VvGT and Cp3GT is highly conserved. The red arrow/circle shows the significant differences seen in the modeled structures.

The superimposition of Cp3GT and VvGT also shows a somewhat conserved 3-D structure with some key differences (Fig. 2.13). The most significant difference observed was at the catalytic cleft that is formed between the N-terminal and C-terminal domain as shown by red circle. Another major difference observed was on the loop region. The flexible loop region connects the C-terminal and N-terminal domain and the catalytic site is formed in a cleft between the N-terminal and C-terminal domain. This difference observed between wild type Cp3GT and VvGT might contribute to different acceptor recognition pattern as VvGT glucosylates anthocyanidins and flavonols whereas Cp3GT glucosylates flavonols only.

The conservation of overall 3-D structures of these models, yet differences in substrate specificity, need a more rigorous analysis to resolve. Therefore, further *in silico* analyses were performed by models docking with potential substrates (Fig.2.14- 2.19).



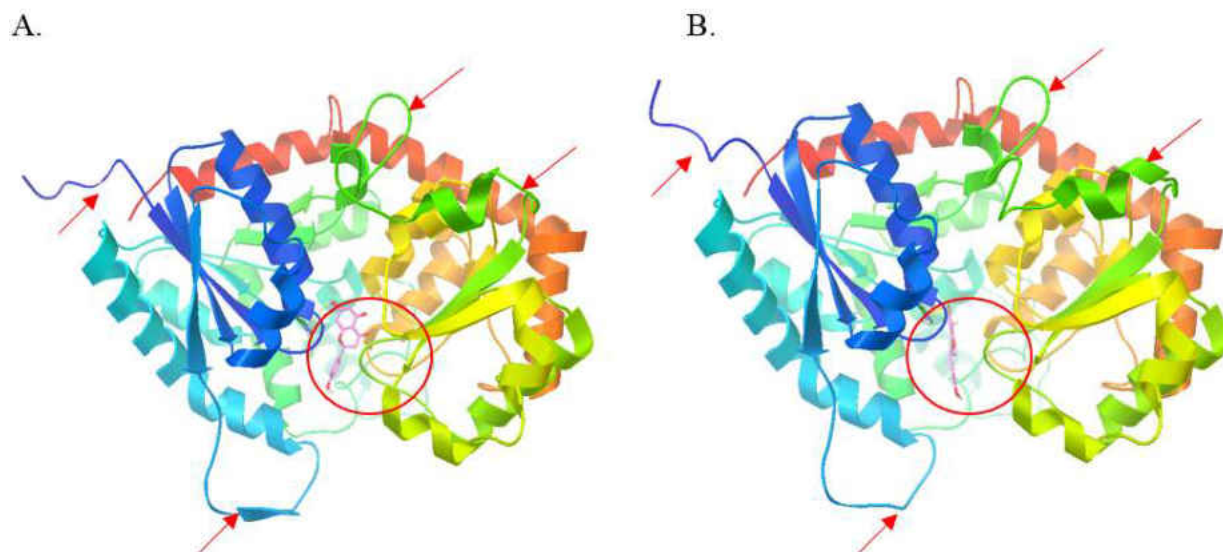


Fig. 2. 14 Docking of wild type Cp3GT and mutant P145T with naringenin. A) Wild type Cp3GT and B) Mutant P145T docked with Naringenin. The red arrows and the circle indicate a major difference observed in the structure of wild type Cp3GT and mutant P145T.

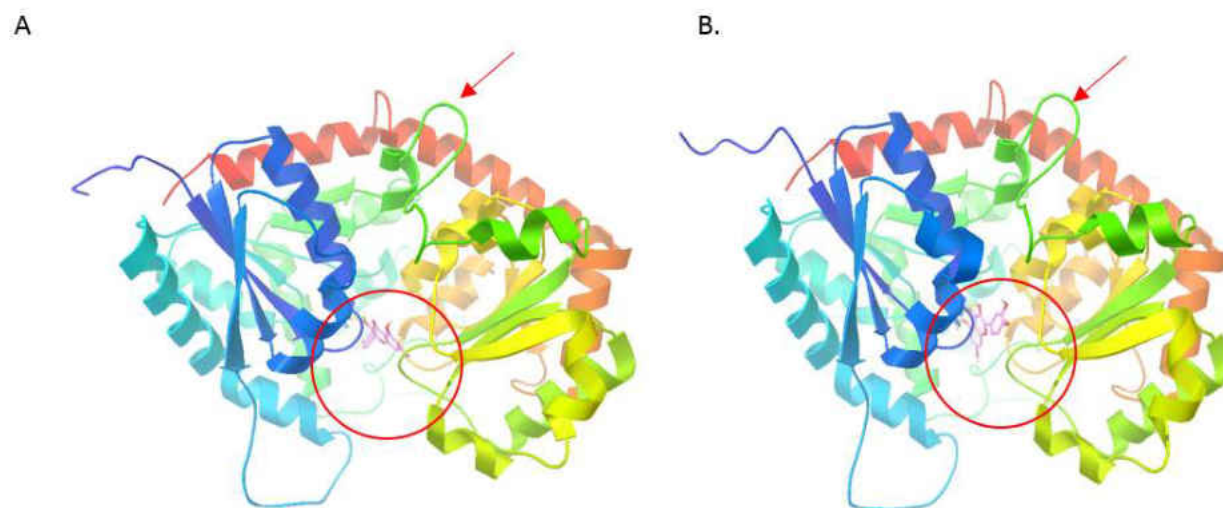


Fig. 2. 15 Docking of wild type Cp3GT and mutant P145T with quercetin. A) Wild type Cp3GT and B) Mutant P145T docked with quercetin. The red arrows and the circle indicate a major difference observed in the structure of wild type Cp3GT and mutant P145T.



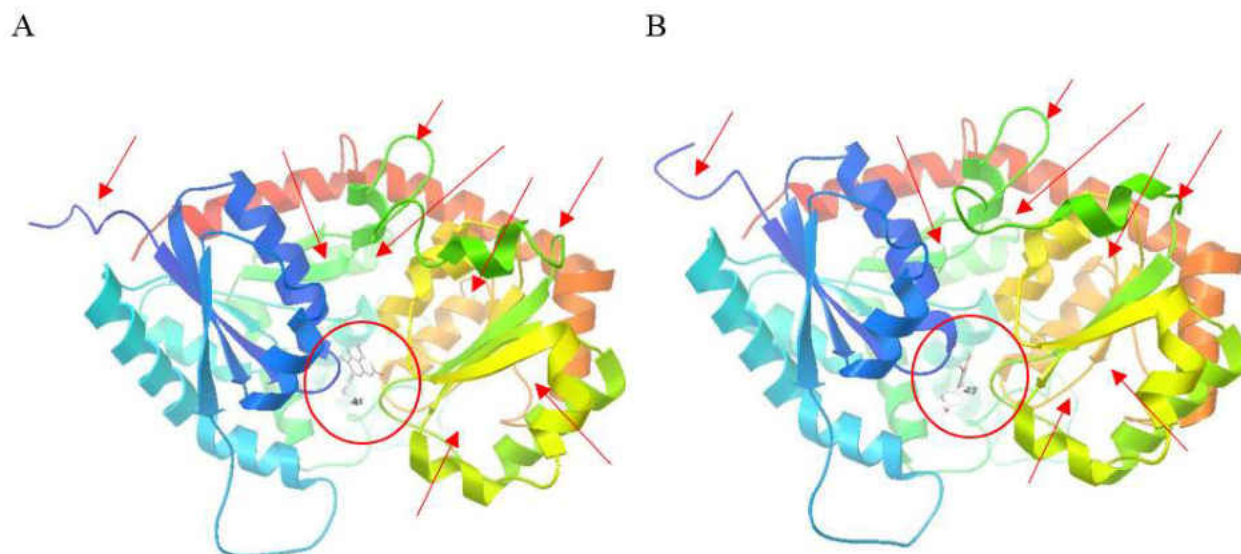


Fig. 2. 16 Docking of wild type Cp3GT and mutant P145T with kaempferol. A) Wild type Cp3GT and B) Mutant P145T docked with kaempferol. The arrows and circle represent the major differences observed between the wild type Cp3GT and the mutant P145T when docked with kaempferol.

The model structures of wild type Cp3GT and mutant P145T docked with different acceptor substrates revealed many similarities and differences. One significant difference was on the catalytic cleft formed between the N-terminal domain and the C-terminal domain, and the loop region that connects the N-terminal and C-terminal domains (Fig. 2.14-2.16). Another difference was the orientation of the acceptor substrate. Compared to wild type Cp3GT, the orientation of the acceptor substrates was different with the mutant P145T enzyme (indicated by the red circle in Fig.2.14- 2.16). The acceptor substrate is rotated by  $90^{\circ}$ . The most significant influence of the catalytic residue histidine on the active site is deprotonation of the  $-OH$  group of the acceptor substrate. The proximity of the catalytic residues His22, Ser20 and Asp122 might be the reason for differential orientation of the acceptor substrate at the active site. The orientation of the acceptor substrate at the active site might be the reason for altered substrate and

regiospecificity of wild type Cp3GT and mutant P145T as the orientation of the acceptor substrate affects the interaction of –OH groups with the catalytic residues and hence might influence the glucosylation reaction.

### 3.5. Molecular docking with acceptor substrates and protein-ligand interaction

Recombinant wild type Cp3GT and mutant P145T were docked *in silico* with acceptor substrates quercetin, kaempferol, and naringenin (Fig. 2.17-2.19). The catalytic dyad His-Asp is widely conserved among the flavonoid UGTs (Breton et al. 2012). The distance of these catalytic residues with the carbonyl and hydroxyl group of acceptor substrate is considered to be important for the acceptor substrate specificity as well as the enzyme catalyzed reaction (Offen et al. 2006). A distance within 5 Å between the catalytic residue His22 and the -OH of the acceptor molecule, side chain of Asp122 and Nε2 atom of His22 and side chain of Ser20 with the O-atom of 4-carbonyl group of the acceptor substrate indicates a likelihood of the sugar attaching reaction (Offen et al. 2006). Molecular docking of different acceptor substrates with the enzyme elucidates the residues within the binding site that are capable of forming hydrogen bond with –OH groups of the acceptor substrates. Hydrogen bonds between the –OH group of the sugar acceptor and the different residues within the active site of the enzyme are critical for the glucosylation reaction.

Crystal structures of these related GTs have shown a common acceptor recognition pattern. A catalytic mechanism of the formation of 3-*O*-glucosylated product has been recently postulated (Hiromoto et al. 2015). Histidine17 in the UDP-glucose: anthocyanidin 3-*O*-glucosyltransferase, UGT78K6 from *Clitoria ternatea*, acts as a Bronsted base that deprotonates the 3-OH group of the acceptor substrate (Hiromoto et al. 2015). Acidic residue Asp114 is

involved in charge stabilization. His-Asp catalytic activity is analogous to the Ser-His-Asp triad of serine hydrolases (reviewed in McIntosh and Owens 2016). After the deprotonation of the glucose accepting group of the acceptor molecule, the anomeric carbon, C1, of glucose undergoes nucleophilic attack by an oxyanion (Lairson et al. 2008). During the transition state, the negative charge of the leaving phosphate group of the donor may be stabilized by positively charged His at position 363 (Hiromoto et al. 2015). Similarly, it was suggested that Asp181 determines the orientation of the acceptor at the acceptor binding site and plays an important role in the differentiation of the regioselective glucosylations (Hiromoto et al. 2015).

The crystal structures of six GTs show that this catalytic dyad His17-Asp114 in UGT78K6 (*Clitoria ternatea*) is conserved as His22-Asp121 in UGT71G1 (Shao et al. 2005), His20-Asp119 in VvGT1 (Offen et al. 2006), His21-Asp125 in UGT85H2 (Li et al. 2007), and His19-Asp117 in AtUGT72B1 (Brazier-Hicks et al. 2007). In all of these GTs, histidine is located at a distance within 5 Å from C1 of UDP-glucose and the acidic residue aspartic acid. We generated structural models and docked different acceptors with the wild type Cp3GT and mutant P145T and looked at the distances within each models to determine if models would predict the likelihood of the glucosylation reactions. These predictions were compared with the experimental results and models were validated.

### 3.5.1. Docking with quercetin

In the docking models, the 3-OH group of the acceptor substrate quercetin was positioned at a distance of 2.88 Å from the Ne2 atom of the His22 catalytic residue in wild type Cp3GT whereas in mutant P145T it is located at a distance of 4.9 Å (Fig. 2.17). Serine 20 was located closer to the 4-carbonyl oxygen in Cp3GT (3.6 Å) compared to a distance of 5.6 Å in mutant

P145T. Asp 122 is located closer to the His22 (4.4 Å) in Cp3GT and 4.3 Å in mutant P145T.

Asp 122 is reported to be involved in charge stabilization after the deprotonation of the OH group of the acceptor.

Molecular docking with quercetin revealed many amino acid residues that could interact with the OH group of the acceptor. The closer proximity of catalytic residue His22 and Ser20 might be the reason of a slightly higher glucosylation activity of wild type Cp3GT with quercetin compared to mutant P145T. In wild type Cp3GT, Pro145 is closer to the 3-OH group of the acceptor and might be involved in non-covalent interactions such as Van der Waals forces.

A

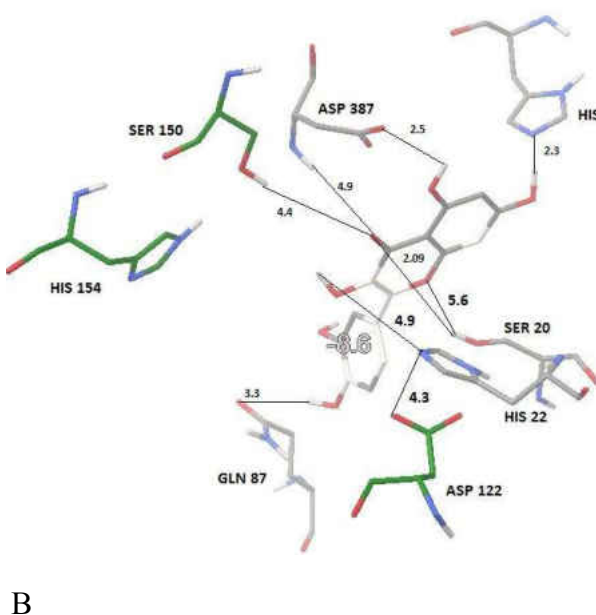
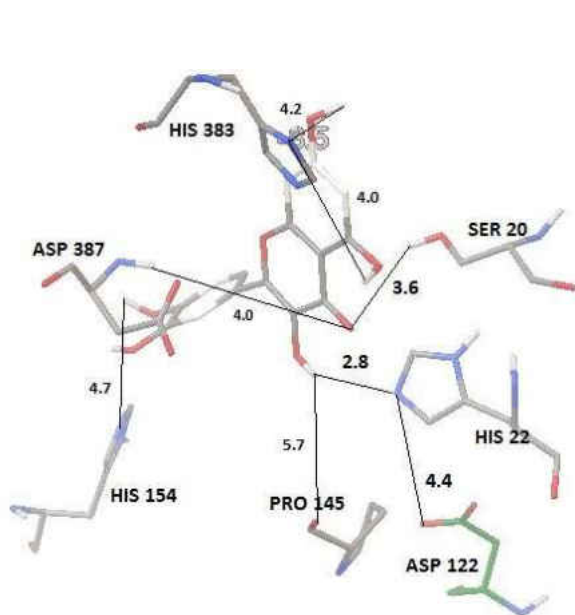


Fig. 2. 17 Molecular docking of wild type Cp3GT and mutant P145T with quercetin. A) Cp-3-O-GT and B) P145T with quercetin. The distance of residues predicted to be within the active site are measured in Å.

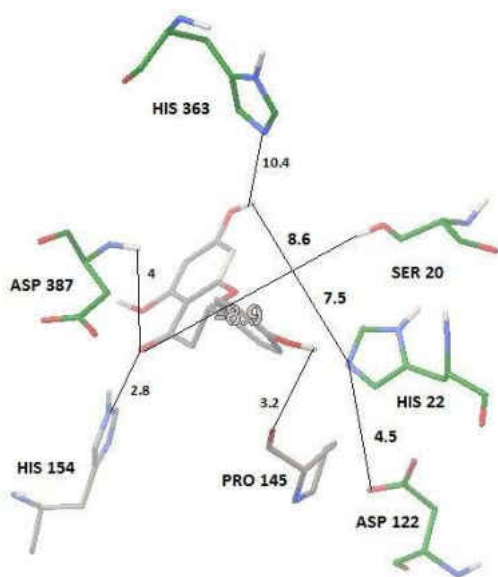
In mutant P145T, threonine145 does not seem to be involved in any type of interactions with quercetin in the models shown. This might be the reason for slight lower activity with mutant P145T. In wild type Cp3GT His154, Asp387, and His363 were found to be interacting with OH and the carbonyl group (Fig. 2.17, A). His154 is positioned at an H-bonding distance (4.7Å) from the 4'-OH group in the wild type Cp3GT model. The imidazole ring of His154 forms H-bond with the 4'-OH group in VvGT (Offen et al. 2006) which was observed in this model with Cp3GT. His363 was found to be present within the H-bonding distance from 7-OH group (4.2Å) as well as 5-OH group (4.03Å) in wild type Cp3GT model. His363 is involved in the stabilization of the charge of the leaving phosphate group of the sugar donor (Hiromoto et al. 2015). Its proximity within the active site suggests it might have a role in enzyme catalyzed reaction.

In the mutant P145T model docked with quercetin, His154 does not make any interaction with the acceptor unlike that seen in wild type Cp3GT model (Fig. 2.17). Other residues such as Asp387, Glu87, and Ser150 interact with the acceptor within H-bonding distance, but they do not make any contact within the 3-OH group of the acceptor. Ser20 forms an H-bond with carbonyl group of C-ring (2.09 Å), Glu87 H-bonds with 4'-OH (3.33 Å), Asp 387 is closer to 4-carbonyl of the C-ring (4.9 Å) and 5-OH (2.56), His363 closer to 7-OH group as in Cp3GT and Ser150 closer to 4-carbonyl oxygen of C-ring (4.4 Å). In this model, different residues are interacting with the acceptor substrate, however, as the catalytic residues are positioned at a greater distance from the 3-OH group and the 4-carbonyl group compared to wild type Cp3GT, the model might be predictive of the role of catalytic residues as well as role of His154 in orienting acceptor substrate within the active site.

### 3.5.2. Docking with naringenin

Docking with naringenin shows the residues located within the active site and the distance of the catalytic residues from the 7-OH group of the acceptor (Fig. 2.18). Mutant P145T is active with naringenin whereas there is no activity with wild type Cp3GT. In mutant P145T, the catalytic residue His22 is positioned at a distance of 4.67Å from the sugar accepting 7-OH group, whereas in Cp3GT the distance is 7.5Å. Ser20 is interacts with 4-carbonyl group of C-ring at a closer distance (5.1 Å) compared to that with wild type Cp3GT (8.6 Å). However, the distance of Asp 122 is similar in both wild type and mutant.

A.



B.

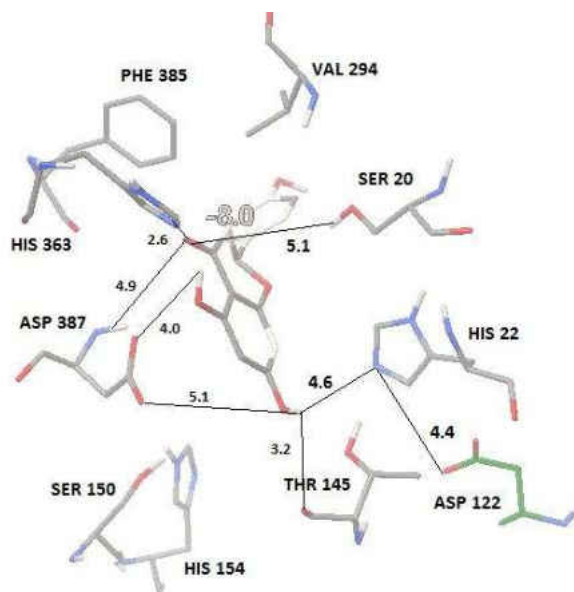


Fig. 2. 18 Molecular docking of wild type Cp3GT and mutant P145T with naringenin. A) Cp-3-*O*-GT and B) P145T with naringenin. The distance of residues predicted to be within the active site are measured in Å.

The closer proximity of the catalytic residue His22 to the 7-OH group of naringenin in the P145T model could likely be the reason for the activity of mutant P145T with naringenin. The involvement of serine20 in forming an H-bond with the 4-carbonyl group of naringenin could be a reason for the activity with mutant P145T as the involvement of side chain of Serine18 in VvGT (position specific residue in Cp3GT is Serine20) in forming an H-bond with the 4-carbonyl group of the acceptor substrate is considered to be important in orienting substrate within the active site for glucosylation reaction (Offen et al. 2006).

In mutant P145T, threonine145 may be involved in non-covalent interactions such as Van der Waals interaction with the 7-OH group as it is positioned closer (3.2 Å). Threonine in the mutant P145T model might be involved in a conformational change which might cause other residues to be involved in different interactions for enzyme activity such as in mutant P145T model, Asp387 is in closer proximity with 7-OH group (5.1 Å) and Asp387 is also positioned within H-bonding distance from the 5'-OH group (4.0 Å) and the 4-carbony group of C-ring (4.9 Å) whereas in wild type Cp3GT, Asp387 is likely to interact only with the 4-carbonyl group of C-ring. In wild type Cp3GT, only His154 seemed to be interacting with the 5-OH group (2.8 Å) while other residues do not appear to be involved in any kind of interactions.

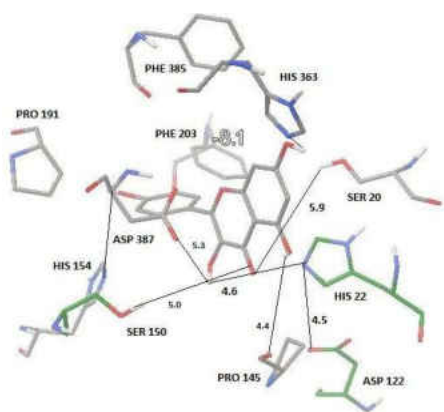
This suggests that along with the closer proximity of the catalytic residues His22, Ser20, and Asp122, other residues such as Asp387, Thr145, and His363 might have a role in the activity of the mutant P145T with naringenin. Interestingly, Thr145 was found to be in closer proximity to the sugar accepting 7-OH group. The activity of mutant P145T with naringenin could be due to structural change due to a single point mutation of proline to threonine. When model with lowest DOPE score did not agree with experimental results, additional models were examined and these

models were the most energy favorable models that predicted the likelihood of reaction with mutant P145T enzyme.

### 2.5.3. Docking with kaempferol

The activity of mutant P145T with kaempferol is almost 2 fold greater than the wild type Cp3GT (Table 2.1 and Fig. 2.8). The distance of catalytic residue His22 from the 3-OH group of kaempferol for mutant P145T is 2.9 Å whereas in wild type Cp3GT, it is 4.6 Å (Fig. 2.19). Similarly, Ser 20 is located more closely (5.0 Å) in mutant P145T from the 4-carbonyl group of C-ring compared to wild type Cp3GT (5.9 Å). Asp122, involved in charge stabilization, is located at a similar distance from the catalytic residue His22 in both wild type and mutant P145T.

A.



B.

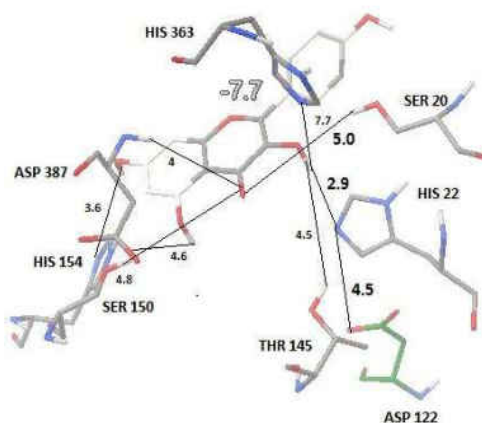


Fig. 2. 19 Molecular docking of wild type Cp3GT and mutant P145T with kaempferol. A) Cp3GT and B) P145T with kaempferol. The distance of residues predicted to be within the active site are measured in Å.



In the docking model, Thr145 in mutant P145T is closer to the 3-OH group (4.51 Å), whereas the corresponding residue in wild type Cp3GT, Pro145, is near to the 5-OH (4.4 Å) not the 3-OH. In mutant P145T, Ser150 is interacting with 4-carbonyl group of the C-ring (4.8 Å). In contrast, Ser150 is at a distance of 5.0 Å in wild type Cp3GT. His 154 is closer to the 7-OH group in mutant P145T (3.6 Å) whereas His 154 is interacting with the 4'-OH in wild type Cp3GT (3.9 Å). Asp387 is close to 5-OH group (4.6 Å) as well as 4-carbonyl group (4.0 Å) in mutant P145T, whereas in Cp3GT, Asp 387 is interacting with the 3-OH group (5.3 Å). These models predict structural change with this mutation that results in kaempferol docking differently.

The closer proximity of the catalytic residues as well the interactions of other residues such as Thr145, Asp387, His154, and Ser150 with the acceptor substrate suggests the likelihood of increased catalytic activity with mutant P145T. The experimental results with kaempferol are in agreement with these computational models that mutant P145T had increased activity with kaempferol. The energy minimization for mutant P145T for this model is -7.7kcal/mol whereas for the wild type it is -8.1kcal/mol. His145 and Serine150 in mutant P145T interact with the 4-carbonyl group and the 3-OH group of the acceptor molecule respectively to form an H-bond. The involvement of serine and histidine in forming an H-bond, although at a different location but in closer proximity to the acceptor substrate binding site, might be a reason for the increased activity as the roles of these residues in forming an H-bond for acceptor substrate specificity have been previously discussed [27]. Thr145 is interacting with the sugar accepting 3-OH group at an H-bonding distance whereas the corresponding residue in wild type Cp3GT is interacting with the OH group that is not involved in glucosylation reaction. This suggests Thr145 at the

closer proximity of the acceptor substrate might be responsible for structural change of the protein that can impact functional differences.

### 3.6. Insertion of thrombin site and heterologous expression

Primers for insertion of a thrombin cleavage site were designed (Table B.1). Site-directed mutagenesis PCR was done to clone a thrombin cleavage site into recombinant wild type Cp3GT and mutant P45T between the enzyme and C-myc/6x-His tags. This was done in preparation to crystallize proteins as well as to test the effect of tags on recombinant enzyme activity.

Site-directed mutagenesis PCR product of both wild type and mutant P145T was first transformed in *E.coli* (Fig. 2.20) and the positive transformants were re-streaked on LB agar plates and confirmed by colony PCR (Fig. 2.21). The insertion of thrombin site in-frame was confirmed by sequencing (Fig. 2.22- 2.23). Plasmid was extracted and linearized using SacI (Fig. 2.24) and work is in progress for its transformation into *Pichia pastoris* for recombinant expression.

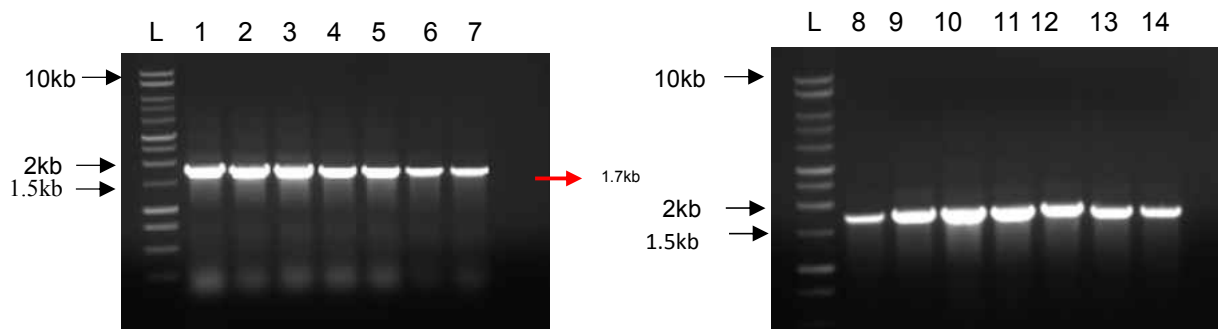


Fig. 2. 20 Colony PCR of positive transformants having thrombin site. Agarose gel electrophoresis of the product of Colony PCR of wild type Cp3GT and mutant P145T plasmid with thrombin site, transformed into competent *E.coli*. PCR was performed with AOX primers. L- 1kb DNA ladder, A) Lane 1 to 7- mutant P145T, and B) Lane 8-14 wild type Cp3GT having thrombin site.

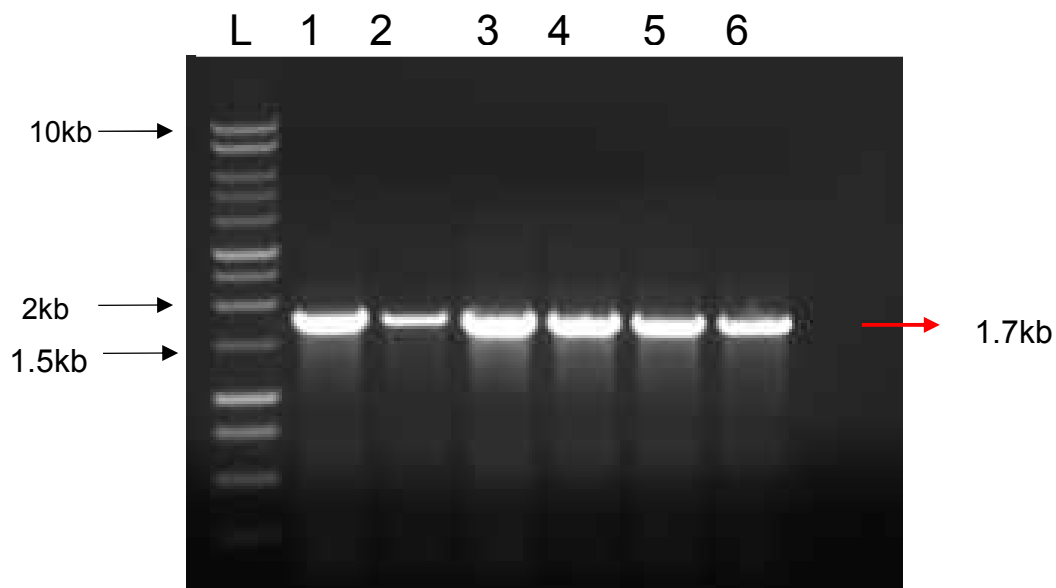


Fig. 2. 21 Colony PCR of positive transformants of *E.coli*. having thrombin cleavage site. Agarose gel electrophoresis of the product of the Colony PCR of wild type Cp3GT and mutant P145T plasmid positive transformants with thrombin site, transformed into competent *E.coli*. PCR was performed with AOX primers. L- 1kb DNA ladder, A) Lane 1 to 7- mutant P145T, and B) Lane 8-14 wild type Cp3GT having thrombin site.

P145T-pPICZA 1 ..... 10 20 30 40 50 60 70 80 90 100  
 ATGGCTCAACGCGAGTCTCAGCCCCGACCTCACATAGCTGTGCTGAATTTCCCTTCTCAACACACGCGCTCGTCCGTTCTTTCAATCATCAACGCGCTCG  
 M A Q T Q S Q P R P H I A V L N F P F S T H A S S V L S I I K R L  
 P145T- TCS 1 ..... 10 20 30 40 50 60 70 80 90 100  
 ATGGCTCAACGCGAGTCTCAGCCCCGACCTCACATAGCTGTGCTGAATTTCCCTTCTCAACACACGCGCTCGTCCGTTCTTTCAATCATCAACGCGCTCG  
 M A Q T Q S Q P R P H I A V L N F P F S T H A S S V L S I I K R L  
 P145T-pPICZA 101 ..... 110 120 130 140 150 160 170 180 190 200  
 CCGTCTCGGCAACCACTGCACCTGTTTACATTTCTTCAGCACTCCGCAATCCAACAGGCCCTTTTCTCCACTGGCCAAACAGCGTCATCTTCCAGCAATGT  
 A V S A P T A L F T F P S T P Q S N K A L F S T G Q Q R H L P S N V  
 P145T- TCS 101 ..... 110 120 130 140 150 160 170 180 190 200  
 CCGTCTCGGCAACCACTGCACCTGTTTACATTTCTTCAGCACTCCGCAATCCAACAGGCCCTTTTCTCCACTGGCCAAACAGCGTCATCTTCCAGCAATGT  
 A V S A P T A L F T F P S T P Q S N K A L F S T G Q Q R H L P S N V  
 P145T-pPICZA 201 ..... 210 220 230 240 250 260 270 280 290 300  
 AAAGCCTTACGACGATCCGATGGAGTCCCGGAAGGCCAGCTGTTCTCCGGGAAGCGTCAGGAAGATATCGAGCTGTTTCAATGCTGCTGATGCCAAC  
 K P Y D V S D G V P E G H V F S G K R Q E D I E L F M N A A D A N  
 P145T- TCS 201 ..... 210 220 230 240 250 260 270 280 290 300  
 AAAGCCTTACGACGATCCGATGGAGTCCCGGAAGGCCAGCTGTTCTCCGGGAAGCGTCAGGAAGATATCGAGCTGTTTCAATGCTGCTGATGCCAAC  
 K P Y D V S D G V P E G H V F S G K R Q E D I E L F M N A A D A N  
 P145T-pPICZA 301 ..... 310 320 330 340 350 360 370 380 390 400  
 TTCAGGAAGCAGTTGAGGCAGCAGTGGCCGAACTGGCAGGCCCTTGACTTTGTTGGTAAACAGATGCTTTCATTTGGTTTGTGCAGAGATGGCTCGAG  
 F R K A V E A A V A E T G R P L T C L V T D A F I W F A A E M A R  
 P145T- TCS 301 ..... 310 320 330 340 350 360 370 380 390 400  
 TTCAGGAAGCAGTTGAGGCAGCAGTGGCCGAACTGGCAGGCCCTTGACTTTGTTGGTAAACAGATGCTTTCATTTGGTTTGTGCAGAGATGGCTCGAG  
 F R K A V E A A V A E T G R P L T C L V T D A F I W F A A E M A R  
 P145T-pPICZA 401 ..... 410 420 430 440 450 460 470 480 490 500  
 AATGGAATAATGTCCCTTGGTTCCATGCTGGACCGTGGCCCAACTGCTCTCTGCTCATGTTTACACTGACATTATCAGGGACAAAATAGGCACCCA  
 E W N N V P W V P C W T A G P N S L S A H L Y T D I I R D K I G T Q  
 P145T- TCS 401 ..... 410 420 430 440 450 460 470 480 490 500  
 AATGGAATAATGTCCCTTGGTTCCATGCTGGACCGTGGCCCAACTGCTCTCTGCTCATGTTTACACTGACATTATCAGGGACAAAATAGGCACCCA  
 E W N N V P W V P C W T A G P N S L S A H L Y T D I I R D K I G T Q  
 P145T-pPICZA 501 ..... 510 520 530 540 550 560 570 580 590 600  
 AAGTCAAAATCAAGATCAACAAGTTTATTCACCTTATCCAGGAATGAATAAGGTACGCGTCGCCGACTTGCCTGAAGAGTTGTTTCCGGAGACTTTGGAT  
 S Q N Q D Q Q L I H F I P G M N K V R V A D L P E G V V S G D L D  
 P145T- TCS 501 ..... 510 520 530 540 550 560 570 580 590 600  
 AAGTCAAAATCAAGATCAACAAGTTTATTCACCTTATCCAGGAATGAATAAGGTACGCGTCGCCGACTTGCCTGAAGAGTTGTTTCCGGAGACTTTGGAT  
 S Q N Q D Q Q L I H F I P G M N K V R V A D L P E G V V S G D L D  
 P145T-pPICZA 601 ..... 610 620 630 640 650 660 670 680 690 700  
 TCAGTCTTTTCTGTTATGCTGCATCAATGGGAGCTCAGCTACCCAGGCGAGCTGCTTTTTCATCAACAGCTTTGAAGAGTTAGACCGCTGAGTTGACAA  
 S V F T S V M L H Q M G R Q L P K A A A V F I N S F E E L D P E L T  
 P145T- TCS 601 ..... 610 620 630 640 650 660 670 680 690 700  
 TCAGTCTTTTCTGTTATGCTGCATCAATGGGAGCTCAGCTACCCAGGCGAGCTGCTTTTTCATCAACAGCTTTGAAGAGTTAGACCGCTGAGTTGACAA  
 S V F T S V M L H Q M G R Q L P K A A A V F I N S F E E L D P E L T  
 P145T-pPICZA 701 ..... 710 720 730 740 750 760 770 780 790 800  
 ATCATCTCAAGACTAAATTCACAACAAGTTTCTCAGTGTGGCCCTTTCAAGCTACTACTAGCATCTGATCAGCAACCGTCTGTCGCAACTGATTTGGA  
 N H L K T R F N N K F L S V G P F K L L L A S D Q Q P S S A T D L D  
 P145T- TCS 701 ..... 710 720 730 740 750 760 770 780 790 800  
 ATCATCTCAAGACTAAATTCACAACAAGTTTCTCAGTGTGGCCCTTTCAAGCTACTACTAGCATCTGATCAGCAACCGTCTGTCGCAACTGATTTGGA  
 N H L K T R F N N K F L S V G P F K L L L A S D Q Q P S S A T D L D  
 P145T-pPICZA 801 ..... 810 820 830 840 850 860 870 880 890 900  
 TGATGAATATGGTTGCCGTGGCTGGACAAGCAGAAGAAGAAACCTGCTTCAGTGGCGTATGTCAAGCTTTGGTACAGTGGCAACACCATCTCCAAAC  
 D E Y G C L A W L D R Q K K K P A S V A Y V S F G T V A T P S P N  
 P145T- TCS 801 ..... 810 820 830 840 850 860 870 880 890 900  
 TGATGAATATGGTTGCCGTGGCTGGACAAGCAGAAGAAGAAACCTGCTTCAGTGGCGTATGTCAAGCTTTGGTACAGTGGCAACACCATCTCCAAAC  
 D E Y G C L A W L D R Q K K K P A S V A Y V S F G T V A T P S P N  
 P145T-pPICZA 901 ..... 910 920 930 940 950 960 970 980 990 1000  
 GAAATTGTGGCAATAGCAGAGGCCCTTGGGAAGCAATAAAGTGCCATTTATTTGGTCACTGAGACATAGGTCCGAGGCAAACTGCCAAATGGGTTCTTGG  
 E I V A I A E A L E A N K V P F I W S L R H R S Q A N L P N G F L  
 P145T- TCS 901 ..... 910 920 930 940 950 960 970 980 990 1000  
 GAAATTGTGGCAATAGCAGAGGCCCTTGGGAAGCAATAAAGTGCCATTTATTTGGTCACTGAGACATAGGTCCGAGGCAAACTGCCAAATGGGTTCTTGG  
 E I V A I A E A L E A N K V P F I W S L R H R S Q A N L P N G F L  
 P145T-pPICZA 1001 ..... 1010 1020 1030 1040 1050 1060 1070 1080 1090 1100  
 AAAGGACAAGATCAGATGGAATTGTGGTGGATTTGGGCCCCCGAGGTCAATGTTTTAGCACATGAAGCAGTTGGGGTCTTTGTAAACACATTTGTGGTGGGG  
 E R T R S D G I V V D W A P Q V N V L A H E A V G V F V T H C G W G  
 P145T- TCS 1001 ..... 1010 1020 1030 1040 1050 1060 1070 1080 1090 1100  
 AAAGGACAAGATCAGATGGAATTGTGGTGGATTTGGGCCCCCGAGGTCAATGTTTTAGCACATGAAGCAGTTGGGGTCTTTGTAAACACATTTGTGGTGGGG

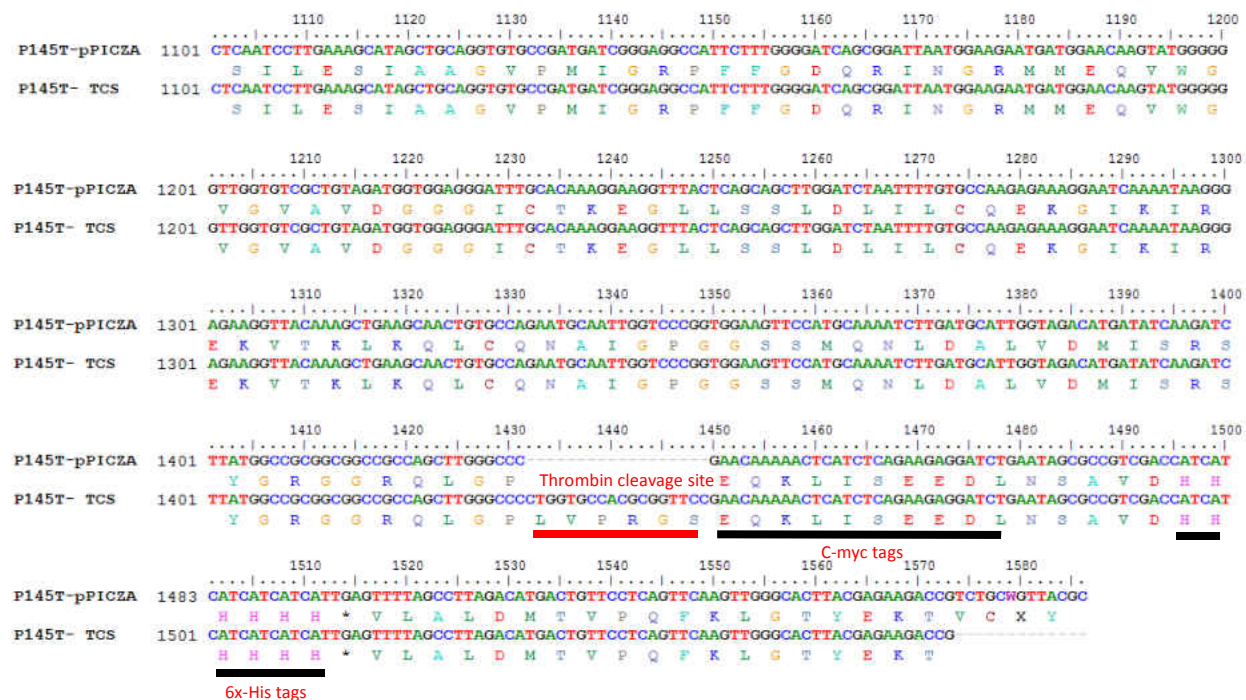


Fig. 2. 22 Sequencing of mutant P145T with thrombin site. Sequencing results of insertion of thrombin cleavage site (leu-val-pro-arg-gly-ser). P145T-pPICZA represents the mutant P145T sequence in pPICZA vector containing gene of interest as well as C-myc and His tags. P145T-TCS represents the consensus sequence obtained with 5' and 3' primers as well as internal primers of the mutant P145T with thrombin cleavage site inserted. Thrombin cleavage site and the tags are highlighted in red and black respectively.



WT-pPICZA 1 .....10 20 30 40 50 60 70 80 90 100  
 ATGGCTCAAAACGAGTCTCAGCCCCGACCTCACATAGCTGTGCTGAATTTCCCTTTCTCAACACAGCGCTCGTCCGTTCTTTCAATCATCAAAACGCTCG  
 M A Q T Q S Q P R P H I A V L N F P F S T H A S S V L S I I K R L

WT-TCS 1 .....10 20 30 40 50 60 70 80 90 100  
 ATGGCTCAAAACGAGTCTCAGCCCCGACCTCACATAGCTGTGCTGAATTTCCCTTTCTCAACACAGCGCTCGTCCGTTCTTTCAATCATCAAAACGCTCG  
 M A Q T Q S Q P R P H I A V L N F P F S T H A S S V L S I I K R L

WT-pPICZA 101 .....110 120 130 140 150 160 170 180 190 200  
 CCGTCTCGGCACCAACTGCAGTGTTCACATTCTTCAGCACTCCGCAATCCAACAGGCCCTTTTCTCCACTGGCCAAACAGCGTCATCTTCCGAGCAATGT  
 A V S A P T A L F T F F S T P Q S N K A L F S T G Q Q R H L P S N V

WT-TCS 101 .....110 120 130 140 150 160 170 180 190 200  
 CCGTCTCGGCACCAACTGCAGTGTTCACATTCTTCAGCACTCCGCAATCCAACAGGCCCTTTTCTCCACTGGCCAAACAGCGTCATCTTCCGAGCAATGT  
 A V S A P T A L F T F F S T P Q S N K A L F S T G Q Q R H L P S N V

WT-pPICZA 201 .....210 220 230 240 250 260 270 280 290 300  
 AAAGCCTTACGAGTATCCGATGGAATCCCGAAGGCCAGCTGTCTTCCCGGAAGCCGTGAGGAAGATATCGAGCTGTTTCATGAATGCTGCTGATGCCAAC  
 K P Y D V S D G V P E G H V F S G K R Q E D I E L F M N A A D A N

WT-TCS 201 .....210 220 230 240 250 260 270 280 290 300  
 AAAGCCTTACGAGTATCCGATGGAATCCCGAAGGCCAGCTGTCTTCCCGGAAGCCGTGAGGAAGATATCGAGCTGTTTCATGAATGCTGCTGATGCCAAC  
 K P Y D V S D G V P E G H V F S G K R Q E D I E L F M N A A D A N

WT-pPICZA 301 .....310 320 330 340 350 360 370 380 390 400  
 TTCAGGAAAGCAGTTGAGGCAGCAGTGGCCGAAACTGGCAGGCCCTTGACTTGTITGGTAACAGATGCTTTCAITTTGGTTTGTGTCAGAGATGGCTCGAG  
 F R K A V E A A V A E T G R P L T C L V T D A F I W F A A E M A R

WT-TCS 301 .....310 320 330 340 350 360 370 380 390 400  
 TTCAGGAAAGCAGTTGAGGCAGCAGTGGCCGAAACTGGCAGGCCCTTGACTTGTITGGTAACAGATGCTTTCAITTTGGTTTGTGTCAGAGATGGCTCGAG  
 F R K A V E A A V A E T G R P L T C L V T D A F I W F A A E M A R

WT-pPICZA 401 .....410 420 430 440 450 460 470 480 490 500  
 AATGGAATAATGTCCCTTGGGTTCATGCTGGCCCGTGGCCCAACTCTCTCTCTGCTCATCTTTACACTGACATTATCAGGGACAAAATAGGCACCCCA  
 E W N N V P W V P C W P A G P N S L S A H L Y T D I I R D K I G T Q

WT-TCS 401 .....410 420 430 440 450 460 470 480 490 500  
 AATGGAATAATGTCCCTTGGGTTCATGCTGGCCCGTGGCCCAACTCTCTCTCTGCTCATCTTTACACTGACATTATCAGGGACAAAATAGGCACCCCA  
 E W N N V P W V P C W P A G P N S L S A H L Y T D I I R D K I G T Q

WT-pPICZA 501 .....510 520 530 540 550 560 570 580 590 600  
 AAGTCAAAATCAAGATCAACAACTTATTCACCTTCAATCCAGGAATGAATAAGGTACCGCTGCCGACTTGCCTGAAGGAGTTGTTTCCGAGAGCTTGGAT  
 S Q N Q D Q Q L I H F I P G M N K V R V A D L P E G V V S G D L D

WT-TCS 501 .....510 520 530 540 550 560 570 580 590 600  
 AAGTCAAAATCAAGATCAACAACTTATTCACCTTCAATCCAGGAATGAATAAGGTACCGCTGCCGACTTGCCTGAAGGAGTTGTTTCCGAGAGCTTGGAT  
 S Q N Q D Q Q L I H F I P G M N K V R V A D L P E G V V S G D L D

WT-pPICZA 601 .....610 620 630 640 650 660 670 680 690 700  
 TCAGTCTTTTCTGTATGCTGCATCAAAATGGGACGTGAGCTACCCAGGCAGCTGCTGTTTTCATCAACAGCTTTGAAGAGTTAGACCCCTGAGTTGACAA  
 S V F S V M L H Q M G R Q L P K A A A V F I N S F E E L D P E L T

WT-TCS 601 .....610 620 630 640 650 660 670 680 690 700  
 TCAGTCTTTTCTGTATGCTGCATCAAAATGGGACGTGAGCTACCCAGGCAGCTGCTGTTTTTCATCAACAGCTTTGAAGAGTTAGACCCCTGAGTTGACAA  
 S V F S V M L H Q M G R Q L P K A A A V F I N S F E E L D P E L T

WT-pPICZA 701 .....710 720 730 740 750 760 770 780 790 800  
 ATCATCTCAAGACTAAATTCACCAACAAAGTTTCTCAGTGTGGCCCTTTCAAGCTACTACTAGCATGTGATCAGCAACCGTCGTCCGCAACTGATTTGGA  
 N H L K T K F N N K F L S V G P P K L L L A S D Q Q P S S A T D L D

WT-TCS 701 .....710 720 730 740 750 760 770 780 790 800  
 ATCATCTCAAGACTAAATTCACCAACAAAGTTTCTCAGTGTGGCCCTTTCAAGCTACTACTAGCATGTGATCAGCAACCGTCGTCCGCAACTGATTTGGA  
 N H L K T K F N N K F L S V G P P K L L L A S D Q Q P S S A T D L D

WT-pPICZA 801 .....810 820 830 840 850 860 870 880 890 900  
 TGATGAATATGGTTGCTTGGCTGGCTGGGACAGCAGAAGAAGAAACCTGCTTCAGTGGCGTATGTGAGCTTTGGTACAGTGGCAACACCATCTCCAAAC  
 D E Y G C L A W L D K Q K K K P A S V A Y V S F G T V A T P S P N

WT-TCS 801 .....810 820 830 840 850 860 870 880 890 900  
 TGATGAATATGGTTGCTTGGCTGGCTGGGACAGCAGAAGAAGAAACCTGCTTCAGTGGCGTATGTGAGCTTTGGTACAGTGGCAACACCATCTCCAAAC  
 D E Y G C L A W L D K Q K K K P A S V A Y V S F G T V A T P S P N

WT-pPICZA 901 .....910 920 930 940 950 960 970 980 990 1000  
 GAAATTTGTGGCAATAGCAGAGGCTTTGGAAGCAAAATAAGTGCCATTATTTGGTCACTGAGACATAGGTCGAGGCAAAATCTGCCAAATGGGTTCTGTG  
 E I V A I A E A L E A N K V P F I W S L R H R S Q A N L P N G F L

WT-TCS 901 .....910 920 930 940 950 960 970 980 990 1000  
 GAAATTTGTGGCAATAGCAGAGGCTTTGGAAGCAAAATAAGTGCCATTATTTGGTCACTGAGACATAGGTCGAGGCAAAATCTGCCAAATGGGTTCTGTG  
 E I V A I A E A L E A N K V P F I W S L R H R S Q A N L P N G F L

WT-pPICZA 1001 .....1010 1020 1030 1040 1050 1060 1070 1080 1090 1100  
 AAAGGACAGATCAGATGGAATTTGGTGGATTTGGGCCCCCAGGTCAATGTTTTAGCACATGAAGCAGTTGGGGTCTTTGTAAACACATTTGTGGTTGGG  
 E R T R S D G I V V D W A P Q V N V L A H E A V G V F V T H C G W G

WT-TCS 1001 .....1010 1020 1030 1040 1050 1060 1070 1080 1090 1100  
 AAAGGACAGATCAGATGGAATTTGGTGGATTTGGGCCCCCAGGTCAATGTTTTAGCACATGAAGCAGTTGGGGTCTTTGTAAACACATTTGTGGTTGGG  
 E R T R S D G I V V D W A P Q V N V L A H E A V G V F V T H C G W G

WT-pPICZA 1101 .....1110 1120 1130 1140 1150 1160 1170 1180 1190 1200  
 CTCATCTCTTGAAGCATAGCTGCAGGTGTGCCGATGATCGGGAGGCCATTCTTTGGGGATCAGCGGATTAATGGAAGATGATGGAACAAAGTATGGGGG  
 S I L E S I A A G V P M I G R P F F G D Q R I N G R M M E Q V W G

WT-TCS 1101 .....1110 1120 1130 1140 1150 1160 1170 1180 1190 1200  
 CTCATCTCTTGAAGCATAGCTGCAGGTGTGCCGATGATCGGGAGGCCATTCTTTGGGGATCAGCGGATTAATGGAAGATGATGGAACAAAGTATGGGGG  
 S I L E S I A A G V P M I G R P F F G D Q R I N G R M M E Q V W G

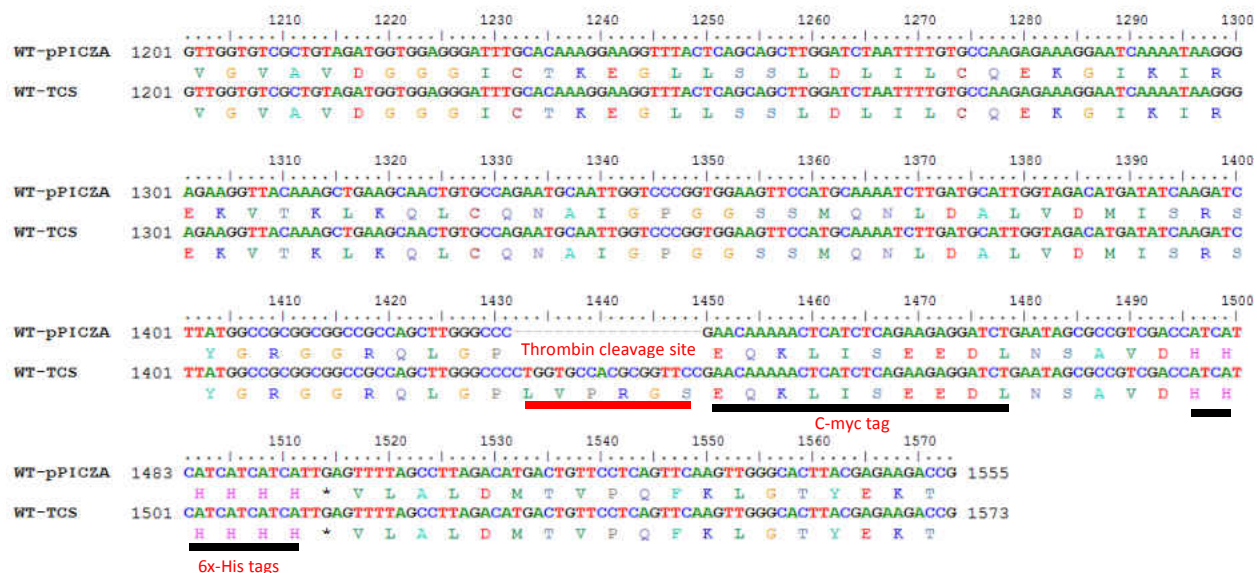


Fig. 2. 23 Sequencing of wild type Cp3GT with thrombin cleavage site. Sequencing results of insertion of thrombin cleavage site (leu-val-pro-arg-gly-ser). WT-pPICZA represents the wild type Cp3GT sequence in pPICZA vector containing gene of interest as well as C-myc and His tags. WT-TCS represents the consensus sequence obtained with 5' and 3' primers as well as internal primers of wild type Cp3GT with thrombin cleavage site inserted. Thrombin cleavage site and the tags are highlighted in red and black respectively.

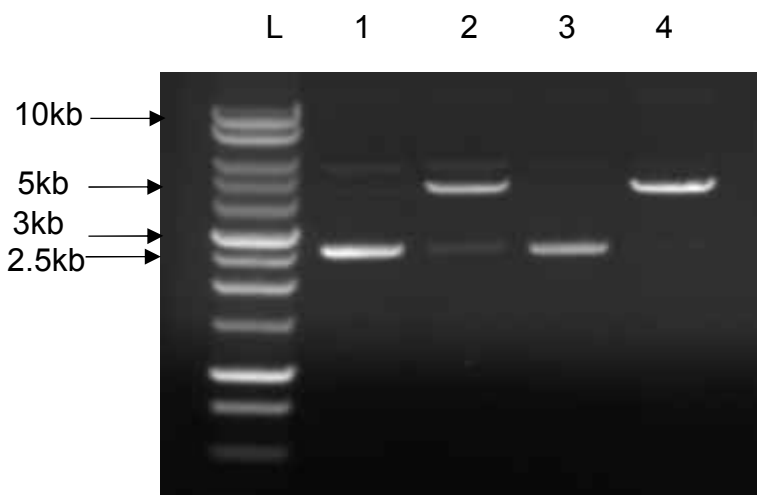


Fig. 2. 24 Digestion of Cp3GT plasmid in pPICZA and P145T plasmid in pPICZA having thrombin site with SacI. Agarose gel electrophoresis of the wild type Cp3GT and mutant P145T plasmid with thrombin site before and after digestion with SacI. L- 1kb DNA ladder, lane 1- non-linear mutant P145T plasmid, lane 2- linearized mutant P145T, lane 3- non-linear wild type Cp3GT, lane 4- linearized wild type Cp3GT.

#### 4. Conclusions

Substitution of proline to threonine brought structural change in the protein. Mutant P145T showed altered substrate as well as regiospecificity. Initial screening of mutant P145T with different aglycones showed increased activity with kaempferol and ability to glucosylate flavanones and flavones that do not contain 3-OH group for glucosylation. The activity with different acceptors quercetin, kaempferol, and naringenin were verified by identifying the products of the reaction. HPLC analysis for product identification showed that the mutant P145T glucosylates naringenin at the 7-OH position forming naringenin-7-*O*-glucoside (prunin); flavonols are glucosylated at 3-OH position. After confirming activity with hesperetin, and scutallerein detailed kinetic studies with quercetin, kaempferol, naringenin, hesperetin, and scutallerein should be done as kinetic studies elucidates the biochemical function of the enzyme with details about the catalytic efficiency of the enzyme, turnover number, and the affinity of the enzyme with the substrates (apparent  $K_m$ ).

Homology modeling and molecular docking was done to generate computational models and models were validated by experimental results. The catalytic residues His22, Ser20, and Asp122 were found to interact with acceptor substrate at the active site. Homology modeling with other acceptor substrates such as hesperetin and scutallerein as well as with the donor substrate UDP-glucose could further elucidate the substrate recognition pattern.

Thrombin cleavage site was inserted into the wild type Cp3GT and mutant P145T between the enzyme and the C-myc tags in order to be able to cleave off tags. In addition to testing biochemical activity with and without tags, the ability to cleave the tags makes it possible to try to crystallize protein. Crystal structures should be attempted for wild type Cp3GT and



mutant P145T in the presence of acceptor substrate with high recombinant enzyme activity. Out of six glucosyltransferases whose crystal structures have been published, VvGT and UGT78K6 glucosylates both anthocyanidins and flavonols whereas our enzyme Cp3GT is strictly a flavonol specific. Crystalizing wild type Cp3GT and mutant P145T that had altered substrate and regiospecificity could possibly explain the substrate recognition patterns of this flavonol specific GT.

### **Acknowledgements**

This research was supported by NSF Grant MCB-1120268 and an ETSU RDC grant awarded to C.A. M. as well as ETSU Graduate School Research Grant awarded to S.K.

## References

1. Owens DK, McIntosh CA. 2011. Biosynthesis and function of Citrus glycosylated flavonoids. In: The Biological Activity of Phytochemicals. Editor Gang DR. Recent Advances in Phytochemistry. Springer New York. p. 67-95.
2. Ferrer J-L, Austin M, Stewart C, Noel J. 2008. Structure and function of enzymes involved in the biosynthesis of phenylpropanoids. *Plant Physiology and Biochemistry* 46(3):356-370.
3. Buer CS, Imin N, Djordjevic MA. 2010. Flavonoids: new roles for old molecules. *Journal of Integrative Plant Biology* 52(1):98-111.
4. Weston LA, Mathesius U. 2013. Flavonoids: their structure, biosynthesis and role in the rhizosphere, including allelopathy. *Journal of Chemical Ecology* 39(2):283-297.
5. Harborne JB. 1993. The flavonoids- Advances in research since 1986. Chapman and Hall, London.
6. Owens DK, McIntosh CA. 2009. Identification, recombinant expression, and biochemical characterization of a flavonol 3-O-glucosyltransferase clone from *Citrus paradisi*. *Phytochemistry* 70(11):1382-1391.
7. Agati G, Brunetti C, Di Ferdinando M, Ferrini F, Pollastri S, Tattini M. 2013. Functional roles of flavonoids in photo protection: new evidence, lessons from the past. *Plant Physiology and Biochemistry* 72:35-45.
8. Brouillard R, Dangles O. 1994. Flavonoids and flower color. In: The Flavonoids. Springer. p. 565-588.
9. Weir TL, Park S-W, Vivanco JM. 2004. Biochemical and physiological mechanisms

mediated by allelochemicals. *Current Opinion in Plant Biology* 7(4):472-479.

10. McIntosh CA, Owens DK. 2016. Advances in flavonoid glycosyltransferase research: integrating recent findings with long-term Citrus studies. *Phytochemistry Reviews*. 1-17.
11. Kumar S, Pandey AK. 2013. Chemistry and biological activities of flavonoids: an overview. *The Scientific World Journal* 2013.
12. Wang X. 2009. Structure, mechanism and engineering of plant natural product glycosyltransferases. *FEBS letters* 583(20):3303-3309.
13. Jones P, Vogt T. 2001. Glycosyltransferases in secondary plant metabolism: tranquilizers and stimulant controllers. *Planta* 213(2):164-174.
14. Lim E-K, Bowles DJ. 2004. A class of plant glycosyltransferases involved in cellular homeostasis. *The EMBO Journal* 23(15):2915-2922.
15. Gachon CMM, Langlois-Meurinne M, Saindrenan P. 2005. Plant secondary metabolism glycosyltransferases: the emerging functional analysis. *Trends in Plant Science* 10(11):542-549.
16. Bowles D, Isayenkova J, Lim E-K, Poppenberger B. 2005. Glycosyltransferases: managers of small molecules. *Current Opinion in Plant Biology* 8(3):254-263.
17. Bhat WW, Dhar N, Razdan S, Rana S, Mehra R, Nargotra A, Dhar RS, Ashraf N, Vishwakarma R, Lattoo SK. 2013. Molecular characterization of UGT94F2 and UGT86C4, two glycosyltransferases from *Picrorhiza kurrooa*: comparative structural insight and evaluation of substrate recognition. *PloS one* 8(9):e73804.
18. Tiwari P, Sangwan RS, Sangwan NS. 2016. Plant secondary metabolism linked

glycosyltransferases: An update on expanding knowledge and scopes. Biotechnology Advances. [doi.org/10.1016/j.biotechadv.2016.03.006](https://doi.org/10.1016/j.biotechadv.2016.03.006)

19. Breton C, Imberty A. 1999. Structure/function studies of glycosyltransferases. Current Opinion in Structural Biology 9(5):563-571.
20. Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B. 2009. The Carbohydrate-Active EnZymes database (CAZy): an expert resource for glycogenomics. Nucleic Acids Research 37:233-238.
21. Campbell JA, Davies GJ, Bulone V, Henrissat B. 1997. A classification of nucleotide-diphospho-sugar glycosyltransferases based on amino acid sequence similarities. Biochemical Journal 326:929-939.
22. Coutinho PM, Deleury E, Davies GJ, Henrissat B. 2003. An evolving hierarchical family classification for glycosyltransferases. Journal of Molecular Biology 328(2):307-317.
23. Hughes J, Hughes MA. 1994. Multiple secondary plant product UDP-glucose glucosyltransferase genes expressed in cassava (*Manihot esculenta* Crantz) cotyledons. DNA Sequence 5(1):41-49.
24. Vogt T, Jones P. 2000. Glycosyltransferases in plant natural product synthesis: characterization of a supergene family. Trends in Plant Science 5(9):380-386.
25. Mackenzie PI, Owens IS, Burchell B, Bock KW, Bairoch A, Belanger A, Giguere SF, Green M, Hum DW, Iyanagi T. 1997. The UDP glycosyltransferase gene superfamily: recommended nomenclature update based on evolutionary divergence. Pharmacogenetics and Genomics 7(4):255-269.
26. Hu Y, Walker S. 2002. Remarkable structural similarities between diverse

glycosyltransferases. *Chemistry & Biology* 9(12):1287-1296.

27. Shao H, He X, Achnine L, Blount JW, Dixon RA, Wang X. 2005. Crystal structures of a multifunctional triterpene/flavonoid glycosyltransferase from *Medicago truncatula*. *The Plant Cell* 17(11):3141-3154.
28. Offen W, Martinez-Fleites C, Yang M, Kiat-Lim E, Davis BG, Tarling CA, Ford CM, Bowles DJ, Davies GJ. 2006. Structure of a flavonoid glucosyltransferase reveals the basis for plant natural product modification. *The EMBO journal* 25(6):1396-1405.
29. Osmani SA, Bak S, Møller BL. 2009. Substrate specificity of plant UDP-dependent glycosyltransferases predicted from crystal structures and homology modeling. *Phytochemistry* 70(3):325-347.
30. McIntosh CA, Latchinian L, Mansell RL. 1990. Flavanone specific 7-*O*-glucosyltransferase activity in *Citrus paradisi* seedlings: purification and characterization. *Archives of Biochemistry and Biophysics* 282(1): 50-57
31. Fukuchi-Mizutani M, Okuhara H, Fukui Y, Nakao M, Katsumoto Y, Yonekura-Sakakibara K, Kusumi T, Hase T, Tanaka Y. 2003. Biochemical and molecular characterization of a novel UDP-glucose: anthocyanin 3'-*O*-glucosyltransferase, a key enzyme for blue anthocyanin biosynthesis, from gentian. *Plant Physiology* 132(3):1652-1663.
32. Sawada Sy, Suzuki H, Ichimaida F, Yamaguchi M-a, Iwashita T, Fukui Y, Hemmi H, Nishino T, Nakayama T. 2005. UDP-glucuronic acid: Anthocyanin glucuronosyltransferase from red daisy (*Bellis perennis*) flowers enzymology and phylogenetic of a novel glucuronosyltransferase involved in flower pigment biosynthesis. *Journal of Biological Chemistry* 280(2):899-906.

33. Caputi L, Lim EK, Bowles DJ. 2008. Discovery of new biocatalysts for the glycosylation of terpenoid scaffolds. *Chemistry—A European Journal* 14(22):6656-6662.
34. Kramer CM, Prata RT, Willits MG, De Luca V, Steffens JC, Graser G. 2003. Cloning and regiospecificity studies of two flavonoid glucosyltransferases from *Allium cepa*. *Phytochemistry* 64(6):1069-1076.
35. Modolo LV, Blount JW, Achnine L, Naoumkina MA, Wang X, Dixon RA. 2007. A functional genomics approach to (iso) flavonoid glycosylation in the model legume *Medicago truncatula*. *Plant Molecular Biology* 64(5):499-518.
36. Brazier-Hicks MaER. 2005. Functional importance of the family 1 glucosyltransferase UGT72B1 in the metabolism of xenobiotics in *Arabidopsis thaliana*. *The Plant Journal* 42(4):556--566.
37. Li L, Modolo LV, Escamilla-Trevino LL, Achnine L, Dixon RA, Wang X. 2007. Crystal structure of *Medicago truncatula* UGT85H2 – insights into the structural basis of a multifunctional (iso) flavonoid glucosyltransferase. *Journal of Molecular Biology* 370(5):951-963.
38. Hiromoto T, Honjo E, Tamada T, Noda N, Kazuma K, Suzuki M, Kuroki R. 2013. Crystal structure of UDP-glucose: anthocyanidin 3-*O*-glucosyltransferase from *Clitoria ternatae*. *Journal of Synchrotron Radiation*. 20: 894-898.
39. Hiromoto T, Honjo E, Noda N, Tamada T, Kazuma K, Suzuki M, Blaber M, Kuroki R. 2015. Structural basis for acceptor-substrate recognition of UDP-glucose: anthocyanidin 3-*O*-glucosyltransferase from *Clitoria ternatea*. *Protein Science*. 24(3):395-407.
40. Jackson R, Knisley D, McIntosh C, Pfeiffer P. 2011. Predicting Flavonoid UGT Regioselectivity. *Advances in Bioinformatics*. Article ID 506583, doi

41. Knisley D, Seier E, Lamb D, Owens D, McIntosh C. 2009 A graph-theoretic model based on primary and predicted secondary structure reveals functional specificity in a set of plant secondary product UDP-glucosyltransferases. In Proceedings of the 2009 International Conference on Bioinformatics, Computational Biology, Genomics and Chemoinformatics, (BCBGC-09), (Loging, W. et al, eds.) pp. 65-72, ISRST.
42. Sarkar TR, Strong CL, Sibhatu MB, Pike LM, McIntosh CA. 2007. Cloning, expression and characterization of a putative flavonoid glucosyltransferase from grapefruit (*Citrus paradisi*) leaves. In: Concepts in Plant Metabolomics. Editors Nikolau B, Wurtele E Springer. p. 247-257.
43. Devaiah SP, Owens DK, Sibhatu MB, Sarkar TR, Strong CL, Mallampalli VK, Asiago J, Cooke J, Kiser S, Lin Z, Wamuch, A, Deborah H, Bruce EW, Berhow M, Pike LM, McIntosh CA. 2016. Identification, recombinant expression, and biochemical analysis of putative secondary product glucosyltransferases from *Citrus paradisi*. Journal of Agricultural and Food Chemistry 64(9):1957-1969.
44. Adepoju OA. 2014. Using site-directed mutagenesis to determine impact of amino acid substitution on substrate and regio-specificity of grapefruit flavonol-3-O glucosyltransferase. East Tennessee State University M.S. Thesis.
45. Sathanatham P. 2015. Impact of mutations of targeted serine, histidine, and glutamine residues in *Citrus paradisi* flavonol specific glucosyltransferase. East Tennessee State University M.S. Thesis.
46. Kuntal BK, Aparoy P, Reddanna P. 2010. EasyModeller: A graphical interface to MODELLER. BMC Research Notes 3(1):226.

## CHAPTER 3

### SUMMARY AND DIRECTIONS FOR FUTURE RESEARCH

In this research, site-directed mutagenesis, homology modeling and biochemical assay of recombinant wild type Cp3GT and mutant P145T enzyme with different flavonoid substrates was performed for the structure-function study of flavonol specific 3-*O*-glucosyltransferase from *Citrus paradisi* (Cp3GT). UDP-glucose; flavonoid 3-*O*-glycosyltransferase from *Vitis vinifera* (VvGT) glucosylates flavonols as well as anthocyanins at the 3-OH and the 7-OH positions respectively (Offen et al. 2006). VvGT and Cp3GT share 56% sequence identity and 87% homology. Cp3GT was modeled with VvGT to design a Cp3GT-P145T mutant by substituting proline145 of Cp3GT with threonine to test the hypothesis that mutation of proline in Cp3GT by threonine can alter substrate and regiospecificity of Cp3GT. Homology models showed that threonine 141 was closer to the substrate binding pocket of VvGT and the corresponding residue in Cp3GT was proline 145. Therefore, the models suggest a structural change. As the cyclic structure of proline's side chain provides proline an exceptional conformational rigidity that forces a bend in the structure of the protein, it was hypothesized that replacing the amino acid that imparts structural rigidity by a residue that allow free rotation resulting in different orientation or may expose protein to different interactions resulting in a change in the structure and hence the catalytic properties of enzyme.

The recombinant enzyme activity of the wild type Cp3GT and mutant P145T enzyme was assayed with 15 different flavonoid substrates. HPLC was performed to verify the activity assay with quercetin, kaempferol, and naringenin. Computational models were generated and molecular docking was performed to predict the location of amino acid residues that are critical



for enzyme activity and function. The computational models were validated by *in vitro* results. In other work, a thrombin cleavage site was inserted into the wild type Cp3GT and mutant P145T in between the enzyme and the C-myc tags in order to be able to cleave off tags.

The mutant P145T enzyme possessed altered substrate and regiospecificity, glucosylating flavones and flavanones along with flavonols. Acceptor substrate specificity was performed by the activity screening assay and activity was confirmed for quercetin, kaempferol and naringenin by product identification using HPLC, also by time course study for quercetin and kaempferol. Product identification with other potential acceptor substrates such as scutallerein and hesperetin should be performed as the mutant P145T showed activity with these acceptors in the screening assay. Because the screening assay is a highly sensitive assay that uses radioactive UDP-glucose and can give false positive results, the acceptor substrate specificity should be further confirmed by performing an assay with unlimiting non-radioactive UDP-glucose. In order to explain the mechanism of enzyme catalysis and function, the mutant P145T could further be biochemically characterized by detailed kinetic study of activity with quercetin, kaempferol, naringenin, hesperetin and scutallerein.

Models of wild type Cp3GT and mutant P145T were generated and molecular docking was used to generate computational models by docking wild type Cp3GT and mutant P145T with quercetin, kaempferol, and naringenin. This provided insight into the key residues that might have a role in the altered substrate and regiospecificity observed. It was found that the catalytic triad His-Asp-Ser was present at the active site of the enzyme within a distance of 5Å from different interacting groups of the acceptor substrate. The proximity of these catalytic residues with the acceptor substrate was used to compare to the *in vitro* results. The

computational models were further validated using experimental results supporting the role of different amino acids in the enzyme activity. Homology modeling and molecular docking further provided insight into the possible orientation of the substrate in active site of the enzyme and the change in the confirmation of enzyme due to the point mutation.

The most significant change in mutant P145T compared to wild type Cp3GT was observed at the region between the N and C domain that forms a cleft for binding donor and acceptor substrate. The models generated suggested the likelihood of the reaction due to the increased proximity of the catalytic residues with the –OH group of the acceptor. Threonine145 was found in close proximity to the sugar accepting group of kaempferol and naringenin in models with mutant P145T. This residue might be involved in non-covalent interactions such as Van der Waals forces causing a change in the structure of the protein within the active site. Similarly, molecular docking had revealed the roles of other amino acid residues that were present within the active site. Asp387, His154, Ser150, and Glu87 were found to interact with the acceptor substrate within the active site suggesting residues other than the catalytic triad might have a role in acceptor substrate recognition.

After confirming the enzyme activity with other substrates such as hesperetin and scutallerein, molecular docking analysis should be performed with these substrates as the mutant P145T had activity with these substrates in screening assay. In the future, wild type and mutant P145T could be docked with the sugar donor that could explain the roles of amino acid in the donor substrate recognition pattern. Docking with donor and acceptor substrate on the active site could also explain the roles of the amino acids in the enzyme activity and function as well the interaction and the orientation of the substrate at the active site. As active site is formed in a

narrow cleft between N-terminal and C-terminal domain, docking with both acceptor and donor might explain the mechanism of enzymatic reaction further as was done for VvGT (Offen et al. 2006).

The recombinant wild type Cp3GT expressed in *E.coli* was investigated for the effects of N-terminal vector-encoded thioredoxin and 6x His tags by comparing the thrombin-treated and untreated recombinant enzyme (Owens and McIntosh 2009). No effect on the enzyme activity was observed whether the tags were present or absent. However, changing the expression system from bacteria to yeast brought a change in the activity of recombinant enzyme with increased stability at higher pH values although the substrate preference for flavonols, effect of metals ions,  $K_i$  value for inhibitor UDP, temperature optimum and temperature stability were similar (Devaiah et al. 2016). This could be due to different chemistry and nature of C-myc/6x-His tags in the yeast expression system as well as position of C-myc tags in the yeast expression system whereas thioredoxin tags were at the N-terminus. In this research, a thrombin site has been inserted into both wild type Cp3GT and mutant P145T just before the C-myc tags in C-terminus. Biochemical characterization of wild type and mutant P145T enzyme in absence of tags should further provide insight into the effects of tags in the recombinant enzyme activity.

In addition to testing biochemical activity with and without tags, the ability to cleave the tags makes it possible to try to crystallize protein. Crystal structures should be attempted for native WT protein and protein crystallized in the presence of acceptor substrate with high recombinant enzyme activity. Out of six glucosyltransferases whose crystal structures have been published, VvGT and UGT78K6 glucosylates both anthocyanidins and flavonols whereas our enzyme Cp3GT is strictly a flavonol specific. Crystalizing wild type Cp3GT and mutant P145T

that had altered substrate and regiospecificity could possibly explain the substrate recognition patterns of this flavonol specific GT. Out of six different crystal structures of family 1 GTs, VvGT and UGT78K6 have specificity with anthocyanidins and flavonols whereas Cp3GT glucosylates flavonols only. VvGT and UGT78K6 have similar acceptor recognition pattern. Crystallizing Cp3GT could possibly explain the substrate specificity patterns of this flavonol specific GT. Mutant P145T had altered substrate and regiospecificity and models generated predicted the likelihood of reaction. Crystallizing P145T with kaempferol and naringenin could possibly help in the comparison of the pattern of acceptor substrate specificity and roles of different residues in altering the strict specificity of wild type Cp3GT. This could further explain the roles of threonine in the altered specificity of the enzyme. Along with this research, earlier work on some of the mutations in our lab should be considered for crystallization. For example a mutant S20L (serine mutated to leucine) and P297F (proline mutated to phenylalanine) completely abolished the enzyme activity (Adepoju 2014). Similarly, another double mutant S20G+T21S altered the substrate specificity with mutant glucosylating flavanones as well as flavonols (Sathanantham 2015). This could possibly explain the roles of different amino acids in the enzyme activity and better aid in the custom designing of enzyme for biotechnological and pharmaceutical applications to synthesize compounds of enhanced nutritional and health benefits.

## REFERENCES

- Adepoju OA. 2014. Using site-directed mutagenesis to determine impact of amino acid substitution on substrate and regio-specificity of grapefruit flavonol-3-O glucosyltransferase. East Tennessee State University M.S. Thesis.
- Agati G, Elisa A, Susanna P, Massimiliano T. 2012. Flavonoids as antioxidants in plants: location and functional significance. *Plant Science* 196: 67-76.
- Appel HM. 1993. Phenolics in ecological interactions: the importance of oxidation. *Journal of Chemical Ecology* 19(7):1521-1552.
- Asada K. 2006. Production and scavenging of reactive oxygen species in chloroplasts and their functions. *Plant Physiology* 141(2):391-396.
- Ashihara H, Clifford MN, Crozier A. 2006. *Plant Secondary Metabolites: Occurrence, Structure and Role in the Human Diet*. Oxford: Wiley-Blackwell pp-372
- Bais HP, Vepachedu R, Gilroy S, Callaway RM, Vivanco JM. 2003. Allelopathy and exotic plant invasion: from molecules and genes to species interactions. *Science* 301(5638):1377-1380.
- Barcelo J, Poschenrieder C. 2002. Fast root growth responses, root exudates, and internal detoxification as clues to the mechanisms of aluminium toxicity and resistance: a review. *Environmental and Experimental Botany* 48(1):75-92.
- Barreca D, Bellocco E, Caristi C, Leuzzi U, Gattuso G. 2012. Flavonoid distribution in neglected Citrus species grown in the Mediterranean basin. In *Handbook of Flavonoids: Dietary Sources, Properties, and Health Benefits*; Yamane, K., Kato, Y., Eds.; Nova Biomedical: New York pp 443–459.

- Baur R, Haribal M, Renwick J, Stadler E. 1998. Contact chemoreception related to host selection and oviposition behavior in the monarch butterfly, *Danaus plexippus*. *Physiological Entomology* 23(1):7-19.
- Benavente-García O, Castillo J, Marin FR, Ortuño A, Del Río JA. 1997. Uses and Properties of Citrus Flavonoids. *Journal of Agricultural and Food Chemistry* 45(12):4505-4515.
- Bennett RN, Wallsgrove RM. 1994. Secondary metabolites in plant defence mechanisms. *New Phytologist* 127(4):617-633.
- Berhow M, Tisserat B, Kanes K, Vandercook C. 1998. Survey of phenolic compounds produced in Citrus. USDA ARS Technical Bulletin, #1856, p.1-154.
- Bhat WW, Dhar N, Razdan S, Rana S, Mehra R, Nargotra A, Dhar RS, Ashraf N, Vishwakarma R, Lattoo SK. 2013. Molecular characterization of UGT94F2 and UGT86C4, two glycosyltransferases from *Picrorhiza kurrooa*: comparative structural insight and evaluation of substrate recognition. *PloS one* 8(9):e73804.
- Bieza K, Lois R. 2001. An Arabidopsis mutant tolerant to lethal ultraviolet-B levels shows constitutively elevated accumulation of flavonoids and other phenolics. *Plant Physiology* 126(3):1105-1115.
- Birt DF, Hendrich S, Wang W. 2001. Dietary agents in cancer prevention: flavonoids and isoflavonoids. *Pharmacology & Therapeutics* 90(2):157-177.
- Boly R, Gras T, Lamkami T, Guissou P, Serteyn D, Kiss R, Dubois J. 2011. Quercetin inhibits a large panel of kinases implicated in cancer cell biology. *International Journal of Oncology* 38(3):833-842.
- Bourne Y, Henrissat B. 2001. Glycoside hydrolases and glycosyltransferases: families and

- functional modules. *Current Opinion in Structural Biology* 11(5):593-600.
- Bowles D, Isayenkova J, Lim E-K, Poppenberger B. 2005. Glycosyltransferases: managers of small molecules. *Current Opinion in Plant Biology* 8(3):254-263.
- Bowles D, Lim E-K, Poppenberger B, Vaistij FE. 2006. Glycosyltransferases of lipophilic small molecules. *Annu. Rev. Plant Biol.* 57:567-597.
- Brazier-Hicks MaER. 2005. Functional importance of the family 1 glucosyltransferase UGT72B1 in the metabolism of xenobiotics in *Arabidopsis thaliana*. *The Plant Journal* 42(4):556--566.
- Breton C, Fournel-Gigleux S, Palcic MM. 2012. Recent structures, evolution and mechanisms of glycosyltransferases. *Current Opinion in Structural Biology* 22(5):540-549.
- Britsch L, Grisebach H. 1986. Purification and characterization of (2S) -flavanone 3-hydroxylase from *Petunia hybrida*. *European Journal of Biochemistry* 156(3):569-577.
- Britsch L, Heller W, Grisebach H. 1981. Conversion of flavanone to flavone, dihydroflavonol and flavonol with an enzyme system from cell cultures of parsley. *Zeitschrift für Naturforschung C* 36(9-10):742-750.
- Broughton WJ, Zhang F, Perret X, Staehelin C. 2003. Signals exchanged between legumes and Rhizobium: agricultural uses and perspectives. *Plant and Soil* 252(1):129-137.
- Brouillard R, Dangles O. 1994. Flavonoids and flower color. In: *The flavonoids*. Springer. p. 565-588.
- Buer CS, Imin N, Djordjevic MA. 2010. Flavonoids: New roles for old molecules. *Journal of Integrative Plant Biology* 52(1):98-111.

- Burghardt F, Knüttel H, Becker M, Fiedler K. 2000. Flavonoid wing pigments increase attractiveness of female common blue (*Polyommatus icarus*) butterflies to mate-searching males. *Naturwissenschaften* 87(7):304-307.
- Campbell JA, Davies GJ, Bulone V, Henrissat B. 1997. A classification of nucleotide-diphospho-sugar glycosyltransferases based on amino acid sequence similarities. *Biochemical Journal* 326:929-939.
- Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B. 2009. The Carbohydrate-Active EnZymes database (CAZy): an expert resource for glycogenomics. *Nucleic Acids Research* 37:233-238.
- Caputi L, Lim EK, Bowles DJ. 2008. Discovery of new biocatalysts for the glycosylation of terpenoid scaffolds. *Chemistry—A European Journal* 14(22):6656-6662.
- Castellarin SD, Pfeiffer A, Sivilotti P, Degan M, Peterlunger E, Di Gaspero G. 2007. Transcriptional regulation of anthocyanin biosynthesis in ripening fruits of grapevine under seasonal water deficit. *Plant, Cell & Environment* 30(11):1381-1399.
- Chalker-Scott L. 1999. Environmental significance of anthocyanins in plant stress responses. *Photochemistry and Photobiology* 70(1):1-9.
- Chen K, Ohmura W, Doi S, Aoyama M. 2004. Termite feeding deterrent from Japanese larch wood. *Bioresource Technology* 95(2):129-134.
- Chirumbolo S. 2010. The role of quercetin, flavonols and flavones in modulating inflammatory cell function. *Inflammation & Allergy-Drug Targets* 9(4):263-285.
- Chou C-H. 1999. Roles of allelopathy in plant biodiversity and sustainable agriculture. *Critical Reviews in Plant Sciences* 18(5):609-636.



- Christie PJ, Alfenito MR, Walbot V. 1994. Impact of low-temperature stress on general phenylpropanoid and anthocyanin pathways: enhancement of transcript abundance and anthocyanin pigmentation in maize seedlings. *Planta* 194(4):541-549.
- Clere N, Faure S, Carmen Martinez M, Andriantsitohaina R. 2011. Anticancer properties of flavonoids: roles in various stages of carcinogenesis. *Cardiovascular & Hematological Agents in Medicinal Chemistry* 9(2):62-77.
- Cody V. 1987. Crystal and molecular structures of flavonoids. *Progress in Clinical and Biological Research* 280:29-44.
- Cook NC, Samman S. 1996. Flavonoids—Chemistry, metabolism, cardioprotective effects, and dietary sources. *The Journal of Nutritional Biochemistry* 7(2):66-76.
- Coutinho PM, Deleury E, Davies GJ, Henrissat B. 2003. An evolving hierarchical family classification for glycosyltransferases. *Journal of Molecular Biology* 328(2):307-317.
- Croteau R, Kutchan TM, Lewis NG. 2000. Natural products (secondary metabolites). Chapter 24. *Biochemistry and Molecular Biology of Plants* p:1250-1319.
- Cushnie TPT, Lamb AJ. 2005. Antimicrobial activity of flavonoids. *International Journal of Antimicrobial Agents* 26(5):343-356.
- Cushnie TT, Cushnie B, Lamb AJ. 2014. Alkaloids: an overview of their antibacterial, antibiotic-enhancing and antivirulence activities. *International Journal of Antimicrobial Agents* 44(5):377-386.
- Devaiah SP, Owens DK, Sibhatu MB, Sarkar TR, Strong CL, Mallampalli VK, Asiago J, Cooke J, Kiser S, Lin Z, Wamuchu, A, Deborah H, Bruce EW, Berhow M, Pike LM, McIntosh

- CA. 2016. Identification, recombinant expression, and biochemical analysis of putative secondary product glucosyltransferases from *Citrus paradisi*. *Journal of Agricultural and Food Chemistry* 64(9):1957-1969.
- Diaz C, Saliba-Colombani V, Loudet O, Belluomo P, Moreau L, Daniel-Vedele F, Morot-Gaudry J-F, Masclaux-Daubresse C. 2006. Leaf yellowing and anthocyanin accumulation are two genetically independent strategies in response to nitrogen limitation in *Arabidopsis thaliana*. *Plant and Cell Physiology* 47(1):74-83.
- Dixon RA, Pasinetti GM. 2010. Flavonoids and isoflavonoids: from plant biology to agriculture and neuroscience. *Plant Physiology* 154(2):453-457.
- Duarte J, Vizcaíno FP, Utrilla P, Jiménez J, Tamargo J, Zarzuelo A. 1993. Vasodilatory effects of flavonoids in rat aortic smooth muscle. Structure-activity relationships. *General Pharmacology: The Vascular System* 24(4):857-862.
- Dubey VS, Ritu B, Rajesh L. 2003. An overview of the non-mevalonate pathway for terpenoid biosynthesis in plants. *Journal of Biosciences* 28(5): 637-646.
- Eisenreich W, Rohdich F, Bacher A. 2001. Deoxyxylulose phosphate pathway to terpenoids. *Trends in Plant Science* 6(2):78-84.
- Feild TS, Lee DW, Holbrook NM. 2001. Why leaves turn red in autumn. The role of anthocyanins in senescing leaves of red-osier dogwood. *Plant Physiology* 127(2):566-574.
- Ferrer JL, Austin MB, Stewart Jr C, Noel JP. 2008. Structure and function of enzymes involved in the biosynthesis of phenylpropanoids. *Plant Physiology and Biochemistry* 46(3):356-370.

- Firmin, JL, Wilson, KE, Rossen L, Johnston, A. 1986. Flavonoid activation of nodulation genes in *Rhizobium* reversed by other compounds present in plants. *Nature* 324, 90 - 92
- Foy C, Lee E, Rowland R, Devine T, Buzzell R. 1995. Ozone tolerance related to flavonol glycoside genes in soybean. *Journal of Plant Nutrition* 18(4):637-647.
- Fukuchi-Mizutani M, Okuhara H, Fukui Y, Nakao M, Katsumoto Y, Yonekura-Sakakibara K, Kusumi T, Hase T, Tanaka Y. 2003. Biochemical and molecular characterization of a novel UDP-glucose: anthocyanin 3'-O-glucosyltransferase, a key enzyme for blue anthocyanin biosynthesis, from gentian. *Plant Physiology* 132(3):1652-1663.
- Gachon CMM, Langlois-Meurinne M, Saindrenan P. 2005. Plant secondary metabolism glycosyltransferases: the emerging functional analysis. *Trends in Plant Science* 10(11):542-549.
- Gonzalez R, Ballester I, Lopez-Posadas R, Suárez M, Zarzuelo A, Martinez-Augustin O, Medina FSD. 2011. Effects of flavonoids and other polyphenols on inflammation. *Critical Reviews in Food Science and Nutrition* 51(4):331-362.
- Hagmann ML, Heller W, Grisebach H. 1983. Induction and Characterization of a Microsomal Flavonoid 3'-Hydroxylase from Parsley Cell Cultures. *European Journal of Biochemistry* 134(3):547-554.
- Hahlbrock K, Grisebach H. 1970. Formation of coenzyme A esters of cinnamic acids with an enzyme preparation from cell suspension cultures of parsley. *FEBS letters* 11(1):62-64.
- Hahlbrock K, Scheel D. 1989. Physiology and molecular biology of phenylpropanoid metabolism. *Annual Review of Plant Biology* 40(1):347-369.
- Hanasaki Y, Ogawa S, Fukui S. 1994. The correlation between active oxygens scavenging and

- antioxidative effects of flavonoids. *Free Radical Biology and Medicine* 16(6):845-850.
- Harborne JB. 1986. Nature, distribution and function of plant flavonoids. In: *Plant Flavonoids in Biology and Medicine: Biochemical, Pharmacological and Structure Activity Relationships*. Editors Cody V, Middleton E, Harborne JB. Alan R Liss Inc., New York. p15-24.
- Harborne J. 1988. Flavonoids in the environment: structure-activity relationships. *Progress in Clinical and Biological Research* 280:17-27.
- Harborne JB. 1991. The chemical basis of plant defense. In: *Plant Defenses Against Mammalian Herbivores*. Editors Palo RT, Robbins CT. CRC Press, Boca Raton, FL, USA. pp. 45-59.
- Harborne JB. 1993. *The Flavonoids – Advances in Research since 1986*. Chapman and Hall, London.
- Havsteen B. 1983. Flavonoids, a class of natural products of high pharmacological potency. *Biochemical Pharmacology* 32(7):1141-1148.
- Heller W, Forkmann G. 1988. Biosynthesis of flavonoids. In: *The Flavonoids: Advances in Research since 1980*. Editor Harborne JB. Boston, MA: Springer US. p. 399-425.
- Heller W, Forkmann G, Britsch L, Grisebach H. 1985. Enzymatic reduction of (+)-dihydroflavonols to flavan-3, 4-cis-diols with flower extracts from *Matthiola incana* and its role in anthocyanin biosynthesis. *Planta* 165(2):284-287.
- Heller W, Hahlbrock K. 1980. Highly purified “flavanone synthase” from parsley catalyzes the formation of naringenin chalcone. *Archives of Biochemistry and Biophysics* 200(2):617-619.

- Henrissat B. 1991. A classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochemical Journal* 280:309-316.
- Henrissat B, Bairoch A. 1993. New families in the classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochemical Journal* 293:781-788.
- Henrissat B, Bairoch A. 1996. Updating the sequence-based classification of glycosyl hydrolases. *Biochemical Journal* 316:695-696.
- Hernández I, Alegre L, Van Breusegem F, Munné-Bosch S. 2009. How relevant are flavonoids as antioxidants in plants? *Trends in Plant Science* 14(3):125-132.
- Hipskind J, Wood K, Nicholson RL. 1996. Localized stimulation of anthocyanin accumulation and delineation of pathogen ingress in maize genetically resistant to *Bipolaris maydis* race O. *Physiological and Molecular Plant Pathology* 49(4):247-56.
- Hiromoto T, Honjo E, Tamada T, Noda N, Kazuma K, Suzuki M, Kuroki R. 2013. Crystal structure of UDP-glucose: anthocyanidin 3-O-glucosyltransferase from *Clitoria ternatae*. *Journal of Synchrotron Radiation* 20: 894-898.
- Hiromoto T, Honjo E, Noda N, Tamada T, Kazuma K, Suzuki M, Blaber M, Kuroki R. 2015. Structural basis for acceptor-substrate recognition of UDP-glucose: anthocyanidin 3-O-glucosyltransferase from *Clitoria ternatea*. *Protein Science* 24(3):395-407.
- Hollman PC, Katan MB. 1998. Bioavailability and health effects of dietary flavonols in man. *Diversification in Toxicology—Man and Environment*. Springer. pp. 237-248.
- Hope WC, Welton AF, Fiedler-Nagy C, Batula-Bernardo C, Coffey JW. 1983. In vitro inhibition of the biosynthesis of slow reacting substance of anaphylaxis (SRS-A) and lipoxygenase activity by quercetin. *Biochemical Pharmacology* 32(2):367-371.

- Hu Y, Walker S. 2002. Remarkable structural similarities between diverse glycosyltransferases. *Chemistry & Biology* 9(12):1287-1296.
- Hughes J, Hughes MA. 1994. Multiple secondary plant product UDP-glucose glucosyltransferase genes expressed in cassava (*Manihot esculenta* Crantz) cotyledons. *DNA sequence* 5(1):41-49.
- Hungria M, Joseph CM, Phillips DA. 1991a. Anthocyanidins and flavonols, major nod gene inducers from seeds of a black-seeded common bean (*Phaseolus vulgaris* L.). *Plant Physiology* 97(2):751-758.
- Hungria M, Joseph CM, Phillips DA. 1991b. Rhizobium nod gene inducers exuded naturally from roots of common bean (*Phaseolus vulgaris* L.). *Plant Physiology* 97(2):759-764.
- Hungria M, Stacey G. 1997. Molecular signals exchanged between host plants and rhizobia: basic aspects and potential application in agriculture. *Soil Biology and Biochemistry* 29(5):819-830.
- Jackson R, Knisley D, McIntosh C, Pfeiffer P. 2011. Predicting Flavonoid UGT Regioselectivity. *Advances in Bioinformatics*. Article ID 506583, doi 10.1155/2011/506583.
- Jay M, Viricel MR, Gonnet JF. 2006. C-glycosylflavonoids. In: *Flavonoids: chemistry, biochemistry and applications*, Editors Andersen M, Markham KR. ISBN 0-8493-2021-6, p 857-915.
- Jones P, Vogt T. 2001. Glycosyltransferases in secondary plant metabolism: tranquilizers and stimulant controllers. *Planta* 213(2):164-174.
- Jourdan PS, McIntosh CA, Mansell RL. 1985. Naringin Levels in Citrus Tissues: II. Quantitative Distribution of Naringin in *Citrus paradisi* Macfad. *Plant Physiology* 77(4):903-908.

- Kabera JN, Semana E, Mussa AR, He X. 2014. Plant Secondary Metabolites: Biosynthesis, Classification, Function and Pharmacological Properties. *Journal of Pharmacy and Pharmacology* 2:377-392.
- Kaul TN, Middleton E, Ogra PL. 1985. Antiviral effect of flavonoids on human viruses. *Journal of Medical Virology* 15(1):71-79.
- Kesterson, JW, Hendrickson R. 1953. Naringin a bitter principle of grapefruit: occurrence, properties and possible utilization. University of Florida Agricultural Experiment Station Tech. Bull., 511. pp. 5–35.
- Kidd PS, Llugany M, Poschenrieder C, Gunsé B, Barceló J. 2001. The role of root exudates in aluminium resistance and silicon-induced amelioration of aluminium toxicity in three varieties of maize (*Zea mays* L.). *Journal of Experimental Botany* 52(359):1339-1352.
- Kieran P, MacLoughlin P, Malone D. 1997. Plant cell suspension cultures: some engineering considerations. *Journal of Biotechnology* 59(1):39-52.
- Klinder A, Shen Q, Heppel S, Lovegrove JA, Rowland I, Tuohy K. 2016. Impact of increasing fruit and vegetable and flavonoid intake on the human gut microbiota. *Food & Function* 7(4): 1715–2116
- Kobayashi H, Graven YN, Broughton WJ, Perret X. 2004. Flavonoids induce temporal shifts in gene-expression of nod-box controlled loci in *Rhizobium* sp. NGR234. *Molecular Microbiology* 51(2):335-347.
- Kolesnikov P, Zore S. 1957. Anthocyanin formation in wheat shoots, induced by visible and invisible ultraviolet light. *Doklady Akademii Nauk SSSR* 112(6):1079-1081.
- Kossalak RM, Bookland R, Barkei J, Paaren HE, Appelbaum ER. 1987. Induction of

- Bradyrhizobium japonicum* common nod genes by isoflavones isolated from *Glycine max*. Proceedings of the National Academy of Sciences 84(21):7428-7432.
- Koukol J, Conn EE. 1961. The metabolism of aromatic compounds in higher plants IV. Purification and properties of the phenylalanine deaminase of *Hordeum vulgare*. Journal of Biological Chemistry 236(10):2692-2698.
- Kovinich N, Saleem A, Arnason JT, Miki B. 2010. Functional characterization of a UDP-glucose:flavonoid 3-O-glucosyltransferase from the seed coat of black soybean (*Glycine max* (L.) Merr.). Phytochemistry 71(11–12):1253-1263.
- Kramer CM, Prata RT, Willits MG, De Luca V, Steffens JC, Graser G. 2003. Cloning and regiospecificity studies of two flavonoid glucosyltransferases from *Allium cepa*. Phytochemistry 64(6):1069-1076.
- Kühnau J. 1975. The flavonoids. A class of semi-essential food components: their role in human nutrition. World Review of Nutrition and Dietetics 24:117-191.
- Kumar S, Pandey AK. 2013. Chemistry and biological activities of flavonoids: an overview. The Scientific World Journal 2013 doi.org/10.1155/2013/162750
- Lairson L, Henrissat B, Davies G, Withers S. 2008. Glycosyltransferases: structures, functions, and mechanisms. Annual Review of Biochemistry 77(1):521-555
- Langenheim JH. 1994. Higher plant terpenoids: a phytocentric overview of their ecological roles. Journal of Chemical Ecology 20(6):1223-1280.
- Lee E-R, Kang G-H, Cho S-G. 2007. Effect of flavonoids on human health: old subjects but new challenges. Recent Patents on Biotechnology 1(2):139-150.
- Li J, Ou-Lee T-M, Raba R, Amundson RG, Last RL. 1993. Arabidopsis flavonoid mutants are



- hypersensitive to UV-B irradiation. *The Plant Cell* 5(2):171-179.
- Li L, Modolo LV, Escamilla-Trevino LL, Achnine L, Dixon RA, Wang X. 2007. Crystal structure of *Medicago truncatula* UGT85H2 – insights into the structural basis of a multifunctional (iso) flavonoid glycosyltransferase. *Journal of Molecular Biology* 370(5):951-963.
- Liang DM, Liu JH, Wu H, Wang BB, Zhu HJ, Qiao JJ. 2015. Glycosyltransferases: mechanisms and applications in natural product development. *Chemical Society Reviews* 44(22):8350–8374
- Lim E-K, Bowles DJ. 2004. A class of plant glycosyltransferases involved in cellular homeostasis. *The EMBO Journal* 23(15):2915-2922.
- Lincoln DE. 1985. Host-plant protein and phenolic resin effects on larval growth and survival of a butterfly. *Journal of Chemical Ecology* 11(11):1459-1467.
- Lincoln DE, Walla MD. 1986. Flavonoids from *Diplacus aurantiacus* leaf resin. *Biochemical Systematics and Ecology* 14(2):195-198.
- Liu J, Mushegian A. 2003. Three monophyletic superfamilies account for the majority of the known glycosyltransferases. *Protein Science* 12(7):1418-1431.
- Lo Piero, A. R., Puglisi, I., Rapisarda, P., & Petrone, G. 2005. Anthocyanins accumulation and related gene expression in red orange fruit induced by low temperature storage. *Journal of Agricultural and Food Chemistry* 53(23):9083-9088.
- Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B. 2014. The carbohydrateactive enzymes database (CAZy) in 2013. *Nucleic Acids Research* 42(1):490-495.

- Lorenc-Kukula K, Jafra S, Oszmianski J, Szopa J. 2005. Ectopic expression of anthocyanin 5-O-glucosyltransferase in potato tuber causes increased resistance to bacteria. *Journal of Agricultural and Food Chemistry* 53(2):272-281.
- Mackenzie PI, Owens IS, Burchell B, Bock KW, Bairoch A, Belanger A, Giguere SF, Green M, Hum DW, Iyanagi T, Lancet D. 1997. The UDP glycosyltransferase gene superfamily: recommended nomenclature update based on evolutionary divergence. *Pharmacogenetics and Genomics* 7(4):255-69.
- Mansell RL, McIntosh CA, Vest SE. 1983. An analysis of the limonin and naringin content of grapefruit juice samples collected from Florida State Test Houses. *Journal of Agricultural and Food Chemistry* 31(1):156-162.
- Martens S, Forkmann G, Matern U, Lukačín R. 2001. Cloning of parsley flavone synthase I. *Phytochemistry* 58(1):43-46.
- McIntosh CA, Latchinian L, Mansell RL. 1990. Flavanone-specific 7-O-glucosyltransferase activity in *Citrus paradisi* seedlings: purification and characterization. *Archives of Biochemistry and Biophysics* 282(1):50-57.
- McIntosh CA, Mansell RL. 1990. Biosynthesis of naringin in *Citrus paradisi*: UDP-glucosyltransferase activity in grapefruit seedlings. *Phytochemistry* 29(5):1533-1538.
- McIntosh CA, Mansell RL. 1997. Three-dimensional distribution of limonin, limonoate A-ring monolactone, and naringin in the fruit tissues of three varieties of *Citrus paradisi*. *Journal of Agricultural and Food Chemistry* 45(8):2876-2883.
- McIntosh CA, Owens DK. 2016. Advances in flavonoid glycosyltransferase research: integrating recent findings with long-term citrus studies. *Phytochemistry Reviews*. 1-17.

- Medina-Puche L, Cumplido-Laso G, Amil-Ruiz F, Hoffmann T, Ring L, Rodríguez-Franco A, Caballero JL, Schwab W, Muñoz-Blanco J, Blanco-Portales R. 2014. MYB10 plays a major role in the regulation of flavonoid/phenylpropanoid metabolism during ripening of *Fragaria* × *ananas* fruits. *Journal of Experimental Botany* 65(2):401-417.
- Middleton Jr E, Kandaswami C. 1994. The impact of plant flavonoids on mammalian biology: implications for immunity, inflammation and cancer. In: *The flavonoids*. JB Harborne Editor. London: Chapman and Hall. p 619-652
- Mo Y-Y, Geibel M, Bonsall RF, Gross DC. 1995. Analysis of sweet cherry (*Prunus avium* L.) leaves for plant signal molecules that activate the syrB gene required for synthesis of the phytotoxin, syringomycin, by *Pseudomonas syringae* pv *syringae*. *Plant Physiology* 107(2):603-612.
- Modolo LV, Li L, Pan H, Blount JW, Dixon RA, Wang X. 2009. Crystal structures of glycosyltransferase UGT78G1 reveal the molecular basis for glycosylation and deglycosylation of (iso) flavonoids. *Journal of Molecular Biology* 392(5):1292-1302.
- Mol J, Grotewold E, Koes R. 1998. How genes paint flowers and seeds. *Trends in Plant Science* 3(6):212-217.
- Moustafa E, Wong E. 1967. Purification and properties of chalcone-flavanone isomerase from soya bean seed. *Phytochemistry* 6(5):625-632.
- Mulvihill EE, Huff MW. 2016. Citrus flavonoids as regulators of lipoprotein metabolism and atherosclerosis. *Annual Review of Nutrition* 36(1).  
(<http://www.annualreviews.org/doi/abs/10.1146/annurev-nutr-071715-050718>)
- Newman DJ, Gordon MC. 2016. Natural products as sources of new drugs from 1981 to 2014. *Journal of Natural Products* 79(3): 629-661.

- Offen W, Martinez-Fleites C, Yang M, Kiat-Lim E, Davis BG, Tarling CA, Ford CM, Bowles DJ, Davies GJ. 2006. Structure of a flavonoid glucosyltransferase reveals the basis for plant natural product modification. *The EMBO journal* 25(6):1396-1405.
- Osbourn, AE, Lanzotti V. 2009. Plant-derived natural products. Synthesis, function and application. Springer. ISBN 978-0-387-85497-7.
- Osmani SA, Bak S, Møller BL. 2009. Substrate specificity of plant UDP-dependent glycosyltransferases predicted from crystal structures and homology modeling. *Phytochemistry* 70(3):325-347.
- Owens DK, McIntosh CA. 2009. Identification, recombinant expression, and biochemical characterization of a flavonol 3-O-glucosyltransferase clone from *Citrus paradisi* *Phytochemistry* 70(11):1382-1391.
- Owens DK, McIntosh CA. 2011. Biosynthesis and function of citrus glycosylated flavonoids. In: *The Biological Activity of Phytochemicals*. Editor Gang DR. Recent Advances in *Phytochemistry* Springer New York. 41: 67-95.
- Paquette S, Møller BL, Bak S. 2003. On the origin of family 1 plant glycosyltransferases. *Phytochemistry* 62(3):399-413.
- Peters N, Verma D. 1990. Phenolic compounds as regulators of gene expression in plant-microbe interactions. *Mol. Plant-Microbe Interact* 3:4-8.
- Peters NK, Frost JW, Long SR. 1986. A plant flavone, luteolin, induces expression of *Rhizobium meliloti* nodulation genes. *Science* 233(4767):977-980.
- Phillips DA, Joseph CM, Maxwell CA. 1992. Trigonelline and stachydrine released from alfalfa seeds activate NodD2 protein in *Rhizobium meliloti*. *Plant Physiology* 99(4):1526-1531.

- Potts JRM, Weklych R, Conn EE. 1974. The 4-hydroxylation of cinnamic acid by sorghum microsomes and the requirement for cytochrome P-450. *Journal of Biological Chemistry* 249(16):5019-5026.
- Quideau S, Deffieux D, Douat-Casassus C, Pouységu L. 2011. Plant polyphenols: chemical properties, biological activities, and synthesis. *Angewandte Chemie International Edition* 50(3):586-621.
- Ragg H, Kuhn DN, Hahlbrock K. 1981. Coordinated regulation of 4-coumarate:CoA ligase and phenylalanine ammonia-lyase mRNAs in cultured plant cells. *Journal of Biological Chemistry* 256(19):10061-10065.
- Randhir R, Lin Y-T, Shetty K. 2004. Phenolics, their antioxidant and antimicrobial activity in dark germinated fenugreek sprouts in response to peptide and phytochemical elicitors. *Asia Pacific Journal of Clinical Nutrition* 13(3):295-307.
- Ratty A, Das N. 1988. Effects of flavonoids on nonenzymatic lipid peroxidation: structure-activity relationship. *Biochemical Medicine and Metabolic Biology* 39(1):69-79.
- Rice EL. 1984. Allelopathy—an update. *The Botanical Review* 45(1):15-109.
- Robak J, Gryglewski RJ. 1988. Flavonoids are scavengers of superoxide anions. *Biochemical Pharmacology* 37(5):837-841.
- Robak J, Korbut R, Shridi F, Swies J, Rzadkowska-Bodalska H. 1987. On the mechanism of antiaggregatory effect of myricetin. *Polish journal of Pharmacology and Pharmacy* 40(3):337-340.
- Russell DW. 1971. The metabolism of aromatic compounds in higher plants X. Properties of the cinnamic acid 4-hydroxylase of pea seedlings and some aspects of its metabolic and

- developmental control. *Journal of Biological Chemistry* 246(12):3870-3878.
- Saija A, Scalese M, Lanza M, Marzullo D, Bonina F, Castelli F. 1995. Flavonoids as antioxidant agents: importance of their interaction with biomembranes. *Free Radical Biology and Medicine* 19(4):481-486.
- Saito K, Kobayashi M, Gong Z, Tanaka Y, Yamazaki M. 1999. Direct evidence for anthocyanidin synthase as a 2-oxoglutarate-dependent oxygenase: molecular cloning and functional expression of cDNA from a red form of *Perilla frutescens*. *The Plant Journal* 17(2):181-189.
- Sathanatham P. 2015. Impact of mutations of targeted serine, histidine, and glutamine residues in *Citrus paradisi* flavonol specific glucosyltransferase. East Tennessee State University M.S. Thesis.
- Sakagami H, Jiang Y, Kusama K, Atsumi T, Ueha T, Toguchi M, Iwakura I, Satoh K, Fukai T, Nomura T. 1999. Induction of apoptosis by flavones, flavonols (3-hydroxyflavones) and isoprenoid-substituted flavonoids in human oral tumor cell lines. *Anticancer Research* 20(1A):271-277.
- Salvayre R, Negre A, Affany A, Lenoble ML, Douste-Blazy L. 1987. Protective effect of plant flavonoids, analogs and vitamin E against lipid peroxidation of membranes. *Progress in Clinical and Biological Research* 280:313-316.
- Sarkar TR, Strong CL, Sibhatu MB, Pike LM, McIntosh CA. 2007. Cloning, expression and characterization of a putative flavonoid glucosyltransferase from grapefruit (*Citrus paradisi*) leaves. In: *Concepts in Plant Metabolomics*. Editors Nikolau B, Wurtele E Springer. p. 247-257.
- Sawada Sy, Suzuki H, Ichimaida F, Yamaguchi M-a, Iwashita T, Fukui Y, Hemmi H, Nishino T,

- Nakayama T. 2005. UDP-glucuronic acid: Anthocyanin glucuronosyltransferase from red daisy (*Bellis perennis*) flowers enzymology and phylogenetic of a novel glucuronosyltransferase involved in flower pigment biosynthesis. *Journal of Biological Chemistry* 280(2):899-906.
- Schröder J, Kreuzaler F, Schäfer E, Hahlbrock K. 1979. Concomitant induction of phenylalanine ammonia-lyase and flavanone synthase mRNAs in irradiated plant cells. *Journal of Biological Chemistry* 254(1):57-65.
- Seo DJ, Jeon SB, Oh H, Lee B-H, Lee S-Y, Oh SH, Jung JY, Choi C. 2016. Comparison of the antiviral activity of flavonoids against murine norovirus and feline calicivirus. *Food Control* 60:25-30.
- Shao H, He X, Achnine L, Blount JW, Dixon RA, Wang X. 2005. Crystal structures of a multifunctional triterpene/flavonoid glycosyltransferase from *Medicago truncatula*. *The Plant Cell* 17(11):3141-3154.
- Sharma R, Panigrahi P, Suresh C. 2014. In-Silico analysis of binding site features and substrate selectivity in plant flavonoid-3-O glycosyltransferases (F3GT) through molecular modeling, docking and dynamics simulation studies. *PloS one* 9(3):e92636.
- Simmonds MS. 2003. Flavonoid–insect interactions: recent advances in our knowledge. *Phytochemistry* 64(1):21-30.
- Sinnott ML. 1990. Catalytic mechanism of enzymic glycosyl transfer. *Chemical Reviews* 90(7):1171-1202.
- Stachel SE, Messens E, Van Montagu M, Zambryski P. 1985. Identification of the signal molecules produced by wounded plant cells that activate T-DNA transfer in *Agrobacterium tumefaciens*. *Nature* 318:624-629.

- Stafford HA, Lester HH. 1982. Enzymic and nonenzymic reduction of (+)-dihydroquercetin to its 3, 4,-diol. *Plant Physiology* 70(3):695-698.
- Stapleton AE, Walbot V. 1994. Flavonoids can protect maize DNA from the induction of ultraviolet radiation damage. *Plant Physiology* 105(3):881-889.
- Stotz G, Forkmann G. 1982. Hydroxylation of the B-ring of flavonoids in the 3'-and 5'-position with enzyme extracts from flowers of *Verbena hybrida*. *Zeitschrift für Naturforschung C* 37(1-2):19-23.
- Suematsu N, Hosoda M, Fujimori K. 2011. Protective effects of quercetin against hydrogen peroxide-induced apoptosis in human neuronal SH-SY5Y cells. *Neuroscience Letters* 504(3):223-227.
- Sutter A, Poulton J, Grisebach H. 1975. Oxidation of flavanone to flavone with cell-free extracts from young parsley leaves. *Archives of Biochemistry and Biophysics* 170:547-556.
- Tanaka Y, Sasaki N, Ohmiya A. 2008. Biosynthesis of plant pigments: anthocyanins, betalains and carotenoids. *The Plant Journal* 54(4):733-749.
- Taniguchi N, Honke K, Fukuda M. 2002. Glucosyltransferase. In: *Handbook of glycosyltransferases and related genes*. Editors Taniguchi N, Honke K, Fukuda M Springer. p. 3-10.
- Thill J, Miosic S, Gotame TP, Mikulic-Petkovsek M, Gosch C, Veberic R, Preuss A, Schwab W, Stampar F, Stich K. 2013. Differential expression of flavonoid 3'-hydroxylase during fruit development establishes the different B-ring hydroxylation patterns of flavonoids in *Fragaria* × *ananassa* and *Fragaria vesca*. *Plant Physiology and Biochemistry* 72:72-78.
- Thoison O, Sévenet T, Niemeyer HM, Russell GB. 2004. Insect antifeedant compounds from



- Nothofagus dombeyi* and *N. pumilio*. *Phytochemistry* 65(14):2173-2176.
- Tiwari P, Sangwan RS, Sangwan NS. 2016. Plant secondary metabolism linked glycosyltransferases: An update on expanding knowledge and scopes. *Biotechnology Advances*. <http://dx.doi.org/10.1016/j.biotechadv.2016.03.006>
- Treutter D. 2005. Significance of flavonoids in plant resistance and enhancement of their biosynthesis. *Plant Biology* 7(6):581-591.
- Torel J, Cillard J, Cillard P. 1986. Antioxidant activity of flavonoids and reactivity with peroxy radical. *Phytochemistry* 25(2):383-385.
- Tzeng S-H, Ko W-C, Ko F-N, Teng C-M. 1991. Inhibition of platelet aggregation by some flavonoids. *Thrombosis Research* 64(1):91-100.
- Van Breusegem F, Dat JF. 2006. Reactive oxygen species in plant cell death. *Plant Physiology* 141(2):384-390.
- Vogt T, Jones P. 2000. Glycosyltransferases in plant natural product synthesis: characterization of a supergene family. *Trends in Plant Science* 5(9):380-386.
- Wang X. 2009. Structure, mechanism and engineering of plant natural product glycosyltransferases. *FEBS letters* 583(20):3303-3309.
- Webb EC. 1992. Enzyme nomenclature. Recommendations of the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology on the Nomenclature and Classification of Enzymes. Academic Press.
- Webster G, Jain V, Davey M, Gough C, Vasse J, Denarie J, Cocking E. 1998. The flavonoid naringenin stimulates the intercellular colonization of wheat roots by *Azorhizobium*

- caulinodans*. Plant, Cell & Environment 21(4):373-383.
- Weir TL, Park S-W, Vivanco JM. 2004. Biochemical and physiological mechanisms mediated by allelochemicals. Current Opinion in Plant Biology 7(4):472-479.
- Weston LA, Mathesius U. 2013. Flavonoids: their structure, biosynthesis and role in the rhizosphere, including allelopathy. Journal of Chemical Ecology 39(2):283-297.
- Wink M. 2003. Evolution of secondary metabolites from an ecological and molecular phylogenetic perspective. Phytochemistry 64(1):3-19.
- Winkel-Shirley B. 2001. Flavonoid biosynthesis. A colorful model for genetics, biochemistry, cell biology, and biotechnology. Plant Physiology 126(2):485-493.
- Winkel-Shirley B. 2002. Biosynthesis of flavonoids and effects of stress. Current Opinion in Plant Biology 5(3):218-223.
- Yang Z-F, Bai L-P, Huang W-b, Li X-Z, Zhao S-S, Zhong N-S, Jiang Z-H. 2014. Comparison of *in vitro* antiviral activity of tea polyphenols against influenza A and B viruses and structure–activity relationship analysis. Fitoterapia 93:47-53.
- Yasuda S, Tahara S, Sakakibara A. 1975. The phenolic constituents of normal and reaction woods of Karamatsu, *Larix leptolepis* GORD. Research Bulletins of the College Experiment Forests Hokkaido University 32(1):55-62.
- Yonekura-Sakakibara K, Hanada K. 2011. An evolutionary view of functional diversity in family 1 glycosyltransferases. The Plant Journal 66(1):182-193.
- Yu O, Matsuno M, Subramanian S, Teixeira da Silva J. 2006. Flavonoid compounds in flowers: genetics and biochemistry. In: Floriculture, Ornamental and Plant Biotechnology Editors

Teixeira da Silva, J. A. Volume 1 p. 282-292.

Yuting C, Rongliang Z, Zhongjian J, Yong J. 1990. Flavonoids as superoxide scavengers and antioxidants. *Free Radical Biology and Medicine* 9(1):19-21.

Zhang Y, Huang M, Wang Q, Cheng J. 2016. Structure-guided unravelling: Phenolic hydroxyls contribute to reduction of acrylamide using multiplex quantitative structure–activity relationship modelling. *Food Chemistry* 199:492-501.

Ziegler J, Facchini PJ. 2008. Alkaloid biosynthesis: metabolism and trafficking. *Annual Review. Review of Plant Biology* 59: 735-769.

## APPENDICES

### APPENDIX A: RECIPES

#### **Acceptor substrates (aglycones) for enzyme assay 50nmol/5 $\mu$ L (1 mL)**

##### **Quercetin**

3.4 mg Quercetin

Dissolve in 1mL of ethylene glycol monomethylether (EGME)

Vortex to dissolve completely

Store at -20°C

##### **Kaempferol**

2.86 mg of Kaempferol

Dissolve in 1 mL of EGME

Vortex to dissolve completely

Store at -20°C

##### **Fisetin**

2.86 mg of Fisetin

Dissolve in 1mL of EGME

Vortex to dissolve completely

Store at -20°C

##### **Gossypetin**

3.18 mg of Gossypetin

Dissolve in 1mL of EGME

Vortex to dissolve completely

Store at -20°C

##### **4'methoxyflavonol**

2.68 mg of 4'methoxyflavonol

Dissolve in 1 mL of EGME

Vortex to dissolve completely

Store at -20°C

### **Dihydroquercetin**

3.04 mg of Dihydroquercetin  
Dissolve in 1mL of EGME  
Vortex to dissolve completely  
Store at -20°C

### **Naringenin**

2.72 mg of Naringenin  
Dissolve in 1 mL of EGME  
Vortex to dissolve completely  
Store at -20°C

### **Hesperetin**

3.02 mg of Hesperetin  
Dissolve in 1 mL of EGME  
Vortex to dissolve completely  
Store at -20°C

### **Eriodyctiol**

2.88 mg of Eriodyctiol  
Dissolve in 1 mL of EGME  
Vortex to dissolve completely  
Store at -20°C

### **Isosakuranetin**

2.86 mg of Isosakuranetin  
Dissolve in 1 mL of EGME  
Vortex to dissolve completely  
Store at -20°C

### **Luteolin**

2.86 mg of Luteolin  
Dissolve in 1 mL of EGME  
Vortex to dissolve completely  
Store at -20°C

### **Diosmetin**

3.0 mg of Diosmetin  
Dissolve in 1 mL of EGME  
Vortex to dissolve completely  
Store at -20°C

### **Cyanidin chloride**

2.87 mg of cyanidin chloride  
Dissolve in 1 mL of EGME  
Vortex to dissolve completely  
Store at -20°C

### **Apigenin**

2.7 mg of Apigenin  
Dissolve in 1mL of EGME  
Vortex to dissolve completely  
Store at -20°C

### **Scutallerein**

2.86 mg of Scutallerein  
Dissolve in 1 mL of EGME  
Vortex to dissolve completely  
Store at -20°C

### **4'-acetoxy-7-hydroxy-6-methoxy isoflavone**

3.26 mg of Diosmetin  
Dissolve in 1 mL of EGME  
Vortex to dissolve completely  
Store at -20°C

### **Sugar donor (UDP- glucose)**

*Non-radioactive for HPLC*

#### **100nmol UDP- $\alpha$ -D-glucose Disodium Salt (non-radioactive)/10 $\mu$ L, 1mL**

6.1 mg of UDP- $\alpha$ -D-glucose

Dissolve in 1000  $\mu$ L of sterile dH<sub>2</sub>O

Vortex before use and store at -20 °C

*Radioactive UDP-glucose, for activity screening assay*

#### **UDP-14<sup>C</sup>-glucose- 22,000 cpm/10 $\mu$ L, 400 $\mu$ L**

20 $\mu$ L of UDP-14<sup>C</sup> glucose (Specific activity=50  $\mu$ Ci/2.5 ml)

380 $\mu$ L of dH<sub>2</sub>O

Vortex well and verify cpm by adding 2 $\mu$ L of the mix to 2 ml scintillation cocktail

Store at -20 °C

*Kinetic assays and biochemical characterization*

#### **UDP-14<sup>C</sup>-glucose (radioactive) – 45,000cpm, 100nmol/10 $\mu$ L, 400 $\mu$ L**

6.78mg of UDP-glucose dissolve in 1mL of dH<sub>2</sub>O = 100nmol/9 $\mu$ L (Stock solution)

40 $\mu$ L of UDP-14<sup>C</sup> glucose (Specific activity=50  $\mu$ Ci/2.5 ml)

360 $\mu$ L of non-radioactive UDP-glucose (100nmols/9  $\mu$ L)

Vortex well and verify cpm by adding 2 $\mu$ L of the mix to 2 ml scintillation cocktail

Store at -20 °C

### **Media and Buffers**

*Protein expression*

#### **Low salt Luria broth (LB) liquid media (200 ml)**

1 g yeast extract (0.5%)

2 g tryptone (1%)

2 g sodium chloride (1%)

Dissolve ingredients in 180 ml dH<sub>2</sub>O

Adjust pH to 7.5 with 1M NaOH

Make up the volume to 200ml using miliQ dH<sub>2</sub>O

Sterilize by autoclaving for 20 minutes

After the solution gets cooled, add 50  $\mu$ L of filter sterilized zeocin from (100 mg/mL stock) to a final concentration of 25  $\mu$ g/ml media

Store at 4°C

### **Low salt LB-Agar plates (200 mL)**

1 g yeast extract (0.5%)

2 g tryptone (1%)

2 g sodium chloride (1%)

Dissolve ingredients in 150 mL dH<sub>2</sub>O

Adjust to pH=7.5 with 1M NaOH

Add 3 g agar (1.5%)

Make up volume to 200 mL with dH<sub>2</sub>O

Autoclave for 20 minutes

Allow to cool to ~55°C and add 50 µL of filter sterilized zeocin from (100 mg/mL stock)

Pour 25ml of media into plates and allow to solidify

Store at 4°C in the dark

### **YPD (200mL)**

2 g yeast extract (1%)

4 g peptone (2%)

Dissolve ingredients in 180 mL dH<sub>2</sub>O (Adjust final pH=6.5)

Autoclave for 20 minutes

Cool to room temperature

Add 20mL of autoclaved 20% dextrose (final concentration of 2%)

Add 200 µL of filter sterilized zeocin from (100 mg/mL) stock.

Store at 4°C

### **YPD-Agar Plate (200mL)**

2 g yeast extract (1%)

4 g peptone (2%)

Dissolve ingredients in 180 mL dH<sub>2</sub>O (Final pH=6.5)

Add 4 g agar (2%)

Autoclave for 20 minutes

Cool to about 55°C and add 20 mL of autoclaved 20% dextrose

Add 200 µL of filter sterilized zeocin from (100 mg/mL) stock

Mix well and pour 25ml in each plate and allow it to solidify

Store plates at 4°C in dark



### **YPDS**

2 g yeast extract (1%)  
4 g peptone (2%)  
36.44 g sorbitol  
Dissolve in 180 mL dH<sub>2</sub>O (Final pH=6.5)  
Autoclave for 20 minutes  
Allow to cool to room temperature  
Add 20 mL of autoclaved 20% dextrose  
Add 200 µL zeocin from (100 mg/mL) stock

### **YPDS-Agar plates (200 mL)**

2 g yeast extract (1%)  
4 g peptone (2%)  
36.44 g sorbitol  
Dissolve in 180 mL dH<sub>2</sub>O (Final pH=6.5)  
Add 4 g agar (2%)  
Autoclave for 20 minutes  
Allow to cool and add 20 mL of autoclaved 20% dextrose  
Add 200 µL zeocin from (100 mg/mL) stock  
Mix well and pour 25 mL into each plate and allow to solidify  
Pour media into plates and allow solidify  
Store plates at 4°C in dark

### **Zeocin (100mg/mL) stock solution (1mL)**

100mg zeocin  
Dissolve in 1mL dH<sub>2</sub>O  
Filter sterilize  
Store at -20°C in dark

### **10X Dextrose (20% Dextrose) (100 mL)**

20 g dextrose ( $\alpha$ -D-glucose)  
Dissolve in 100 mL dH<sub>2</sub>O  
Autoclave for 20min and store at room temperature

### **BMGY (1L)**

10 g yeast extract  
20 g peptone  
Dissolve in 700 mL sterile dH<sub>2</sub>O  
Autoclave for 20 min  
Allow to cool at room temperature

Add:

100 mL autoclaved 1M potassium phosphate buffer (pH=6.0)  
100 mL filter sterilized 1M 10X YNB  
100 mL filter sterilized 10X Glycerol  
2 mL filter sterilized 500X Biotin  
Store at 4°C

### **BMMY (1L)**

10 g yeast extract  
20 g peptone  
Dissolve in 700 mL dH<sub>2</sub>O  
Autoclave for 20 min  
Allow to cool at room temperature

Add:

100 mL autoclaved 1M potassium phosphate buffer (pH=6.0)  
100 mL filter sterilized 1M 10X YNB  
100 mL filter sterilized 10X Methanol (add methanol just before use)  
2mL 500X Biotin  
Store at 4°C

### **10X GY (10% glycerol) (1000ml)**

100 mL of 100Xglycerol  
Mix 900 mL dH<sub>2</sub>O  
Autoclave for 20 min  
Store at 4°C

### **10X YNB (500 mL)**

17 g yeast nitrogen base  
50 g ammonium sulfate  
Dissolve in 500 mL dH<sub>2</sub>O  
Heat slightly to dissolve the ingredients completely  
Filter sterilize the solution and store at 4°C

### **500X Biotin (0.02% Biotin) 100ml**

20 mg biotin  
Dissolve in 100 mL dH<sub>2</sub>O  
Filter sterilize the solution  
Store at 4°C

### **10X M (5% Methanol) (500 mL)**

25 mL methanol  
Add 475 mL of dH<sub>2</sub>O  
Filter sterilize  
Store at 4°C

### **1M potassium phosphate buffer, pH=6.0**

87.09g of dibasic potassium phosphate (K<sub>2</sub>HPO<sub>4</sub>)  
Dissolved in 500mL of dH<sub>2</sub>O

136.09g of monobasic potassium phosphate (KH<sub>2</sub>PO<sub>4</sub>)  
Dissolved in 1000mL of dH<sub>2</sub>O

Mix 132mL of 1M dibasic potassium phosphate (K<sub>2</sub>HPO<sub>4</sub>) + 868 mL of 1M monobasic potassium phosphate (KH<sub>2</sub>PO<sub>4</sub>)  
Make up pH of 6.0  
Autoclave for 20 minutes at liquid cycle and store at 4°C

### *Agarose gel electrophoresis*

### **0.8% Agarose Gel (50mL)**

0.4 g Agarose  
Dissolve ingredient in 50 mL 1X TAE buffer  
Microwave for 60 to 70 seconds until agarose dissolves  
Add dH<sub>2</sub>O to make up volume of 50mL  
Allow to cool and add 2.5µL of a 10mg/mL ethidium bromide  
Mix well and pour 40ml into gel casting tray

### **50X TAE buffer (1L)**

242 g Tris base  
57.1 mL glacial acetic acid  
Add 100 mL of 0.5M EDTA (pH=8.0)  
Dissolve the ingredients in ~700 mL dH<sub>2</sub>O  
Make up volume to 1L with dH<sub>2</sub>O and store at room temperature

### **1X TAE (1L)**

Take 20mL of 50X TAE stock and make up volume to 1L adding 980mL dH<sub>2</sub>O

### *SDS-PAGE electrophoresis*

#### **10X- Running buffer (1L)**

30.28g Tris base

144.14g glycine

10g SDS

Dissolve in 1000ml dH<sub>2</sub>O

Store at room temperature

#### **1X- Running buffer**

100mL of 10X stock

Add 900mL of dH<sub>2</sub>O

#### **4X-Separating buffer (1.5M Tris, pH 8.8) 100mL**

18.2g Tris base

Dissolve in 80ml H<sub>2</sub>O

Adjust pH to 8.8 using 6M HCl

make up volume to 100ml

Filter sterilize the solution

Store at 4°C

#### **4X-Stacking buffer (1M Tris, pH 6.8) 100mL**

12.1g Tris base

Dissolve in 80mL dH<sub>2</sub>O

Adjust pH to 6.8 with 6M HCl

Make up volume to 100 mL with dH<sub>2</sub>O

filter sterilize and store at 4°C

#### **4X SDS-PAGE Sample Buffer (10mL)**

2.4mL of 1M Tris, pH 6.8

0.8g SDS (sodium lauryl sulfate)

4mL 100% glycerol

4mg bromophenol blue

3.1mL dH<sub>2</sub>O

Store at room temperature

Add 50 µL of βME per mL of 4X sample buffer immediately before use.

#### *10% SDS-PAGE Gel*

#### **10% SDS (100 mL)**

10 g SDS

Dissolve in 100 mL dH<sub>2</sub>O

Store at room temperature

#### **10% APS (Ammonium persulfate)**

0.1g Ammonium persulfate

Dissolve in 1mL dH<sub>2</sub>O (use fresh each time)

#### **Separating gel**

	<b>1 gel</b>	<b>2 gel</b>
H <sub>2</sub> O.....	2.3ml	4.8ml
40% Acrylamide.....	1.3ml	2.5ml
1.5M Tris (pH 8.8).....	1.3ml	2.5ml
10% SDS.....	50ul	100ul
10% APS.....	50ul	100ul
TEMED.....	2ul	4ul

#### **Stacking gel**

	<b>1 gel</b>	<b>2 gel</b>
H <sub>2</sub> O.....	730ul	1.5ml
40% Acrylamide .....	130ul	260ul
1 M Tris (pH 6.8).....	130ul	250ul
10% SDS.....	10ul	20ul
10% APS.....	10ul	20ul
TEMED.....	1ul	2ul

*Western blot*

**10X Transfer buffer (1L)**

30.28g Tris base

144.14g glycine

Dissolve ingredients in 1000 mL dH<sub>2</sub>O

Store at 4°C

**1X Transfer buffer (1L)**

100 mL of 10X stock

Add 700 mL of dH<sub>2</sub>O

Add 200 mL of methanol

Mix and store at 4°C

**Blocking solution (50 ml)**

2.5 g non-fat milk powder

Dissolve in 50 ml of 1X TBS and add 50 µL of 20% sodium azide

**20% sodium azide (1 ml)**

20 mg sodium azide

Dissolve in 1 ml dH<sub>2</sub>O

Store at 4 °C

**5X TBS buffer (1L)**

40g NaCl

1g KCl

15g Tris base

Dissolve ingredients in 900mL dH<sub>2</sub>O

Adjust to pH=7.4 with 6N HCl

Make up to a final volume of 1L with dH<sub>2</sub>O

Store at room temperature

**Primary Antibody (1:2500 dilution)**

6 µL of anti-C-myc monoclonal antibody from mouse

Dissolve in 15 ml of 1X TBS

**Secondary Antibody (1:10,000 dilution)**

1.5 µL of anti-mouse IgG from goat

Dissolve in 15 ml of 1X TBS

**Alkaline phosphate (AP) buffer (pH 9.5) (500 ml)**

2.925 g NaCl

507.5 mg  $\text{MgCl}_2 \cdot 6\text{H}_2\text{O}$

50 ml of 1M CHES buffer (41.44g CHES dissolve in 200mL  $\text{dH}_2\text{O}$  to make 200mL of 1M CHES)

Dissolve ingredients in 475 ml  $\text{dH}_2\text{O}$

Adjust pH to 9.5

Make up the volume to 500ml using  $\text{dH}_2\text{O}$

Store at room temperature

**Nitro blue tetrazolium (NBT) solution (1mL)**

83mg NBT dissolved in 700uL of N, N-dimethyl formamide and 300uL of  $\text{dH}_2\text{O}$

Store at  $-20^\circ\text{C}$

**5-bromo-4-chloro-3-indolyl-phosphate (BCIP) solution (1mL)**

42mg BCIP dissolve in 1mL of N, N-dimethylformamide

Store at  $20^\circ\text{C}$

**0.5 M EDTA (pH 8.0) (500 ml)**

73.06 g of EDTA

Dissolve in 450 ml  $\text{dH}_2\text{O}$

Adjust the pH to 8.0

Make up the volume using  $\text{dH}_2\text{O}$

Store at room temperature

*Protein extraction and purification*

**Breaking buffer, pH 7.5 (500ml)**

50mM Sodium phosphate buffer (pH 7.5)

1mM PMSF (Phenyl methyl sulfonyl fluoride)

1mM EDTA

5% glycerol

5mM  $\beta\text{ME}$

(PMSF and  $\beta\text{ME}$  should be added before use)

0.65 g monobasic Sodium phosphate

5.5 g dibasic Sodium phosphate

145mg EDTA

Dissolve in 400ml  $\text{DH}_2\text{O}$

Adjust pH to 7.5

Add 50mL of 50% glycerol  
Make up volume to 500mL  
Store at 4°C

-Add 5ul of 200mM PMSF per ml of breaking buffer  
-Add 0.35μL of βME per mL of breaking buffer

**200mM PMSF (1mL)**

35mg of PMSF dissolved in 1mL methanol  
Store at -20°C

**50% glycerol (500ml)**

250mL of 100%  
Make up volume to 500mL adding dH<sub>2</sub>O  
Mix well and store at 4°C

**Equilibration/Wash buffer (pH=7.5) (500mL)**

50mM Sodium phosphate buffer (pH 7.5)  
300mM NaCl  
5mM βME

0.65 g monobasic sodium phosphate  
5.5 g dibasic sodium phosphate  
8.75 g sodium chloride  
Dissolve in 450mL dH<sub>2</sub>O  
Adjust pH to 7.5  
Make up volume to 500mL with dH<sub>2</sub>O  
Store at 4°C

Add 0.35μL of βME per mL of elution buffer prior to use

**Elution buffer, pH 7.5 (500 ml)**

50mM Sodium phosphate buffer (pH 7.5)  
300mM NaCl  
5mM βME  
150mM Imidazole

0.65 g monobasic sodium phosphate  
5.5 g dibasic sodium phosphate  
8.75 g sodium chloride  
5.1 g imidazole



Dissolve in 450ml dH<sub>2</sub>O and adjust pH to 7.5  
Make up the volume using dH<sub>2</sub>O and store at 4 °C

Add 0.35 µL of βME per mL of elution buffer prior to use

**Assay/ Final buffer, pH 7.5 (500 ml)**

50mM Sodium phosphate buffer (pH 7.5)  
14mM βME

0.65 g monobasic sodium phosphate  
5.5 g dibasic sodium phosphate  
Dissolve in 450ml dH<sub>2</sub>O and adjust pH to 7.5  
Make up the volume using dH<sub>2</sub>O and store at 4 °C  
Add 1µL of βME per mL of assay buffer prior to use

**IMAC column regeneration and storing**

20mM MES-KOH, pH 5.0 with 0.3M NaCl

0.39 g MES  
1.75 g NaCl

Dissolve in 100 ml dH<sub>2</sub>O  
Adjust pH to 5.0 using 1 M KOH  
Store at 4 °C

**IMAC Column storing solution (100 ml)**

20% ethanol containing 0.1% sodium azide

20ml of 100X ethanol  
Make volume to 100ml using dH<sub>2</sub>O  
Add 100mg of sodium azide  
Mix well and Store at 4 °C

## APPENDIX B: METHODS

### SDS-PAGE

Two 10% SDS-PAGE gels were prepared as described in Appendix A. Samples for SDS-PAGE were made by mixing 5 $\mu$ L of SDS-PAGE sample buffer (Appendix) with 10-15 $\mu$ L of protein samples. The mixture was boiled for 5 minutes to denature the proteins. The SDS-PAGE tank was filled with 1X Tris-glycine electrophoresis running buffer (Appendix). A 5 $\mu$ L volume of pre-stained low molecular weight marker (Bio-rad) was loaded in the first well followed by protein samples. The samples were electrophoresed for 90 minutes at 100V. One gel was stained with Coomassie blue overnight keeping on a shaker and destained using destaining solution.

### Western blotting

One of the two SDS-PAGE gels were used for western blot analysis to detect the proteins. A black and white transfer cassette was taken with black surface down. A white porous pad was placed over the cassette and then a filter paper above it followed by gel and then nitrocellulose membrane and then again white porous pad. In a Western blot tank, 1X ice cold transfer buffer was poured and the cassette was placed in the center of the tank. Electrophoresis was performed for 90 minutes at 100V. The nitrocellulose membrane was carefully removed and emerged in blocking solution for 2 hours at room temperature (or at 40C overnight). The membrane was then washed for five minutes thrice with 15 mL of 1x TBS buffer containing 15 $\mu$ L of 20% sodium azide. The primary antibody (Anti-C-myc monoclonal antibody from mouse, Sigma Aldrich) was diluted 1:2500 in 15mL of TBS buffer with 15mL of 20% Sodium azide. The membrane was immersed in primary antibody for 2 hours in a shaker and then washed again three times with TBS buffer. The membrane was then immersed in secondary antibody (Ant-mouse Immunoglobulin G-alkaline phosphatase conjugate from goat, Novagen) for 2 hours at

room temperature. Again the membrane was washed with TBS three time for five minutes each. The membrane was immersed in 15mL of alkaline phosphate buffer (AP buffer, pH 9.0) for five minutes for buffer exchange. In another dish, 15mL of AP buffer was taken and 60μL of BCIP and 60μL of NBT was added and mixed well and the membrane was immersed in this solution until the protein band appeared. The membrane was then transferred to a separate dish containing a mixture of 15 ml TBS and 60 μL of 5M EDTA in order to stop the reaction.

### Insertion of thrombin cleavage site

#### Primer design

Thrombin recognizes the consensus sequence leucine-valine-proline-arginine-glycine-serine, cleaving the peptide bond between arginine and glycine. Primers specific for thrombin site were designed (Table B.1) using Quickchange Primer Design web tool from Agilent Technologies.

([http://www.genomics.agilent.com/primerDesignProgram.jsp?&\\_requestid=994370](http://www.genomics.agilent.com/primerDesignProgram.jsp?&_requestid=994370)). Site-directed mutagenesis PCR was done to insert thrombin site in between the wild type cp3GT and mutant P145T enzyme and the C-myc tags.

Table B. 1 Primers for insertion of the thrombin site. Forward and reverse primers used for insertion of thrombin site between the enzyme and C-myc tags.

Primer name	Primer sequence	Length	T <sub>m</sub> (°C)
CSP212-1527R	5'-GAGATGAGTTTTTGTTCGGAACCGCGTGGCACCAGGGGCCCAAGCTGGCGG-3'	51	81.62
CSP211-1527F	5'-CCGCCAGCTTGGGCCCCTGGTGCCACGCGGTTCCGAACAAAACTCATCTC-3'	51	81.62

### Site-directed mutagenesis

Site-directed mutagenesis PCR was performed to insert thrombin site in between the wild type Cp3GT and mutant P145T enzyme and the C-myc tags using QuikChange<sup>R</sup> Lightning Site-Directed Mutagenesis Kit from Agilent Technologies. The reaction mixture contained:

#### P145T-thrombin cleavage site

5μL of 10X reaction buffer  
1.25μL of 125ng/μL Primer (forward)  
1.25μL of 125ng/μL Primer (reverse)  
1μL of dNTP mix  
1.5μL of Quick solution reagent:  
1 μL of Quick change DNA polymerase enzyme  
39μL of P145T DNA template (45ng)

#### Cp3GT-thrombin cleavage site

5μL of 10X reaction buffer  
1.25μL of 125ng/μL Primer (forward)  
1.25μL of 125ng/μL Primer (reverse)  
1μL of dNTP mix  
1.5μL of Quick solution reagent:  
1 μL of Quick change DNA polymerase enzyme  
39μL of Cp3GT DNA template (35ng)

PCR was run using the following conditions as mentioned on the QuikChange<sup>R</sup> Lightning Site-Directed Mutagenesis Protocol from Agilent Technologies.

Table B. 2 PCR conditions for insertion of thrombin site using site directed mutagenesis.

Step	Temperature (°C)	Time(s)	Cycles
Initiation	95	30	1
Denaturation	95	30	16
Annealing	55	60	
Elongation	68	600	
Extension	68	600	1

The PCR product was run on 0.8% agarose gel to confirm the band of appropriate size (Supplementary Fig. C.2).

#### DpnI digestion of parental DNA of the site-direct mutagenesis PCR product

The site directed mutagenesis PCR product of wild type Cp3GT and mutant P145T containing thrombin cleavage site in between the enzyme and the C-myc tags was treated with the restriction enzyme DpnI (Agilent Technologies) to digest the parental template DNA. 1μL of DpnI enzyme (10U/μL) was added to each of the 40μL of site-directed mutagenesis PCR products and mixed properly by pipetting up and down several times. Then the tubes were incubated at 37°C water bath for 1 hour. The digested product was run on 0.8% agarose gel (Supplementary Fig. C.3) to confirm the presence of intact plasmid of expected size 4.7 kb.

#### Transformation into *E.coli* competent cells

DpnI digested site-directed mutagenesis PCR product containing wild type Cp3GT and mutant P145T with thrombin cleavage site was transformed into *E.coli* competent cells (Invitrogen™) using heat shock method. Two tubes containing 50μL of competent cells were thawed on ice and 25μL of each was transferred to a different sterile Eppendorf tube. A 2μL aliquot of DpnI digested wildtype Cp3GT and mutant P145T plasmid with thrombin cleavage site was added in each tube and mixed gently (not by pipetting up and down). The tubes were incubated on ice for 30 minutes. Low Salt LB agar plates containing zeocin at a concentration of 25 μg/mL were made at the meantime and 42°C water bath was made ready for transformation. The tubes were subjected to heat shock by transferring to a pre-heated 42°C water bath for 30 seconds followed by incubation on ice for 2 minutes. A 250μL of S.O.C media (Invitrogen) was

added to each tubes and incubated at 37°C shaker for 1 hour at 250rpm. A 50µL and 100 µL of this transformed product was spreaded on different LB agar plates and incubated at 37°C overnight.

#### PCR screening of positive transformants

A total of seven colonies were selected for each wild type Cp3GT and mutant P145T with thrombin cleavage site and resuspended in PCR tubes containing 10 µL sterilized dH<sub>2</sub>O. A 5 µL aliquot of the mixture was used as the template for colony PCR to identify positive transformants and the remaining mixture was kept at 4°C. PCR reaction were run using the conditions shown on Table 4.

The PCR reaction was run using AOX primers (Table 3). The reactions contained:

1 µL of AOX1-5' forward primer (1µM)

1 µL of AOX1-3' reverse primer (1µM)

5µL of 2X GoTaq<sup>R</sup> green master mix (Promega)

Table B. 3 AOX primer sequence used for colony PCR

Primer	Sequence	T <sub>m</sub> (°C)
5' AOX1	5'-GACTGGTTCCAATTGACAAGC-3'	54.3
3' AOX1	5'-GCAAATGGCATTCTGACATCC-3'	54.8

Table B. 4 Colony PCR reaction conditions to identify the positive transformants.

Step	Temperature(°C)	Time (s)	Cycles
Initiation	94	300	1
Denaturation	94	30	30
Annealing	58	30	
Elongation	72	90	
Final elongation	72	300	1

The PCR products were electrophoresed on a 0.8% agarose gel (Fig. 2.18). Expected size of 1.7kb was obtained for all the colonies tested.

#### Re-streaking of positive transformants

Three different positive transformants each of wild type Cp3GT and mutant P145T with thrombin site was selected and re-streaked on six different LB agar plates containing 25µg/mL of zeocin and incubated at 37 °C overnight (3 plates each for wild type Cp3GT and mutant P145T). One colony from each plates was selected and PCR screening was repeated for the re-streaked positive colonies as described. Bands of size 1.7kb was obtained for each colonies tested (Fig. 2.11).

#### Miniprep plasmid extraction for sequencing

Three positive colonies from PCR screening each of wild type Cp3GT and mutant P145T with thrombin cleavage site were selected for further Miniprep plasmid extraction. A 5µL aliquot of the PCR product of the positive colonies were inoculated into 4mL of LB media containing zeocin at a concentration of 25µg/mL and incubated at 37oC overnight at 225rpm. 500µL of 40% glycerol and 500µL of each of the PCR product inoculated for plasmid extraction were mixed and the tubes were kept in -80°C.

GeneJET plasmid mini-prep kit was used for mini-prep plasmid extraction ([https://tools.thermofisher.com/content/sfs/manuals/MAN0012655\\_GeneJET\\_Plasmid\\_Miniprep\\_UG.pdf](https://tools.thermofisher.com/content/sfs/manuals/MAN0012655_GeneJET_Plasmid_Miniprep_UG.pdf), Thermo scientific). As directed on the user manual, RNAase was mixed to the Resuspension solution and kept at 4°C. A 4.3 mL volume of ethanol was added to the Wash solution. A 1.5mL volume of the bacterial cells grown overnight were harvested by centrifuging at 8000rpm for 2 min at room temperature. The supernatant was discarded and again 1.5mL of

bacterial cells were added and centrifuged and supernatant was discarded. A 500  $\mu$ L volume of the bacterial cells were used to make glycerol stocks and stored at  $-80^{\circ}\text{C}$ . Cell pellets were resuspended in 250 $\mu$ L of Resuspension solution. Vortexing was done to resuspend the cells completely. A 250 $\mu$ L volume of the Lysis solution was added and mixed thoroughly by inverting the tube up and down for 4-5 times. The solution then became viscous and slightly clear. At this stage, vortexing was completely avoided to prevent the shearing of the chromosomal DNA. A 300 $\mu$ L volume of Neutralization solution was added and mixed immediately and thoroughly by inverting the tubes 4-6 times. It is always important to mix thoroughly and gently after the addition of the Neutralization solution to avoid localized precipitation of bacterial cell debris. The neutralized bacterial lysate should become cloudy. Then it was centrifuged for 5 min at 13000 rpm to pellet the cell debris and chromosomal DNA. Then the supernatant was transferred to the GeneJet spin column by pipetting. Transferring of the white precipitate was completely avoided. Then the spin column was centrifuged for 1 minute at 13000 rpm. The flow through was discarded and the column was kept back to the same tube. A 500 $\mu$ L of wash solution was added to the column and centrifuged for 1 minute at 13000rpm. Again the flow through was discarded and again the column was centrifuged to avoid residual ethanol in plasmid preparation. The GeneJet spin column was transferred to a sterile 1.5mL centrifuge tube and 50 $\mu$ L of Elution buffer was added to the center of the spin column membrane to elute the plasmid DNA. It was incubated for 3 minutes at room temperature. The samples were then centrifuged for 2 min and again 50 $\mu$ L of Elution buffer was added and centrifuged for 2 minutes. The concentration of DNA was measured using Nanodrop at 260nm. The concentration of DNA obtained for four different samples are shown in Table below:



Table B. 5 Miniprep Plasmid Extraction. Concentration of DNA after Miniprep extraction of four different colonies of wild type Cp3GT and mutant P145T with thrombin cleavage site.

Recombinant mutant with thrombin cleavage site	Sample	Concentration (ng/ $\mu$ L)
P145T	P1	102.5
	P3	82.3
WT	W1	77.3
	W3	70.9

Four different samples P1, P3, W1 and W3 were sent to Molecular Biology core facility of University of Tennessee, Knoxville for sequencing (At least 3 $\mu$ g of DNA is required by UT, Knoxville for sequencing). AOX primers and gene specific internal primers were also send for sequencing. The samples were sent as following:

Table B. 6 Samples sent for sequencing. Concentration and volume of samples sent for sequencing.

Sample name	Conc. ( $\mu$ g/ $\mu$ L)	Volume( $\mu$ L)	Total DNA ( $\sim\mu$ g)
P1	0.1	50	5
P3	0.08	50	4.1
W1	0.077	50	3.8
W3	0.07	50	3.5

The sequencing results were analyzed with the Bioedit software and the presence of thrombin cleavage site within the frame confirmed. Sample P3 and W3 showed the residues within the frame (Fig. 5). Therefore P3 and W3 were used for further analysis.

#### Midiprep plasmid extraction for transformation into yeast

Wizard plus Midipreps DNA purification System (Promega) was used for midiprep DNA extraction and purification

(<https://www.promega.com/~media/files/resources/protocols/technical%20bulletins/0/wizard%20plus%20midipreps%20dna%20purification%20system%20protocol.pdf>). Glycerol stock of

*E. coli* cells containing wild type Cp3GT and mutant P145T with thrombin cleavage site from -80°C was inoculated in 4mL of LB media containing zeocin at a concentration of 25µg/mL and incubated at 37°C shaker overnight at 250rpm. A 100µL of overnight culture was transferred to 50mL of LB media containing zeocin at a concentration of 25µg/mL and incubated at 37°C overnight at 250rpm. A 40mL volume of overnight culture was centrifuged at 10,000 x g for 10 minutes at 4°C. The supernatant was poured out and the tube was blotted upside-down on a paper towel to remove excess media. The cell pellets were resuspended in 3mL of Resuspension solution completely by vortexing. A 3mL volume of Cell lysis solution was added to the tube and mixed well by inverting the tube upside down 4-5 times. A 3mL volume of Neutralization solution was added and mixed by inverting the tubes upside down for 4-5 times. The tubes were centrifuged at 14,000 x g for 15 minutes at 4°C. Then the supernatant was decanted carefully transferred to a different tube avoiding white precipitate.

A 10ml volume of Wizard® Midipreps DNA Purification Resin was added to the tubes and swirled to mix well. Each Midicolumn tip was inserted into the vacuum manifold port and the resin/DNA mixture into the Midicolumn. A vacuum of at least 15 inches of Hg was applied to pull the resin/DNA mix into the Midicolumn. When all of the sample passed through the column, 15 mL of Column wash solution was added to the Midicolumn and vacuum was applied to draw the solution through the Midicolumn. After washing off completely, another 15mL of Column wash solution was added to the Midicolumn and vacuum was applied again to draw the solution through the Midicolumn. In the meantime, nuclease free H<sub>2</sub>O was pre-heated at 65°C water bath.

The Midicolumn was removed from the vacuum port and the lower tip of the Midicolumn was cut as described in the protocol. After cutting the Midicolumn, it was transferred to a 1.5ml microcentrifuge tube. The Midicolumn was centrifuged at  $10,000 \times g$  in a microcentrifuge tube for 2 minutes to remove any residual Column Wash Solution. The Midicolumn was transferred again to a new microcentrifuge tube. A 300 $\mu$ l volume of preheated (65–70°C) Nuclease-Free Water was added to the Midicolumn and incubated for 1 minute. The DNA was eluted by centrifuging the Midicolumn at  $10,000 \times g$  for 20 seconds in a microcentrifuge. The Midicolumn was transferred to a different microcentrifuge tube and added 300 $\mu$ l of Nuclease-Free Water and centrifuged the sample at  $10,000 \times g$  in a microcentrifuge for 5 minutes to pellet the fines. The DNA-containing supernatant was carefully transferred to a clean microcentrifuge tube and the concentration of each sample was determined using a Nanodrop at 260nm. The concentration of the samples are shown in table below.

Table B. 7 Midiprep plasmid extraction and purification. Concentration of wild type Cp3GT and mutant P145T having thrombin cleavage site after midiprep extraction.

Sample	Concentration (ng/ $\mu$ L)
P3-a (P145T)	154.85
P3-b (P145T)	176.95
W3-a (Cp3GT)	174.3
W3-b (Cp3GT)	139.3

#### Linearization using SacI for transformation in to yeast

The plasmid containing wild type Cp3GT and mutant P145T having thrombin cleavage site was linearized by restriction enzyme SacI (Promega) for efficient transformation into yeast. The plasmid DNA extracted by midiprep at a concentration of 5-10 $\mu$ g was used for linearization. The reaction mixture consisted of the followings:

#### P145T plasmid

Midiprep plasmid DNA-	100 $\mu$ L
10X Buffer J-	34 $\mu$ L
100X BSA-	3 $\mu$ L
SacI-	3 $\mu$ L
Sterile H <sub>2</sub> O-	200 $\mu$ L
Total volume-	340 $\mu$ L

#### Cp3GT plasmid

Midiprep plasmid DNA-	70 $\mu$ L
10X Buffer J-	34 $\mu$ L
100X BSA-	3 $\mu$ L
SacI-	3 $\mu$ L
Sterile H <sub>2</sub> O-	230 $\mu$ L
Total volume-	340 $\mu$ L

A 1.5 $\mu$ L aliquot of SacI enzyme was first added to the reaction mixture containing plasmid DNA, 10X Buffer J, 100X BSA and sterile water. The mixture was incubated at 37°C water bath for 4 hours and additional 1.5 $\mu$ L of SacI was added further and incubated overnight. A 1 $\mu$ L aliquot of the digested sample as well as 1 $\mu$ L aliquot of undigested plasmid DNA was mixed with 4 $\mu$ L of sterile water and 1 $\mu$ L of 6X dye and run through 0.8% agarose gel to ensure the complete digestion (Fig. 2.10). The reaction was then stopped by incubating the reaction mixture in a 65°C water bath to heat inactivate the SacI enzyme.

#### Phenol-chloroform extraction of linearized plasmid DNA

A 250 $\mu$ L volume of the linearized DNA was taken for phenol-chloroform extraction to purify and concentrate the DNA for transformation into yeast. A 250 $\mu$ L volume of cold phenol-chloroform-isoamyl alcohol solution was added (1:1 volume ratio) to the linearized DNA sample and the mixture was vortexed for 30 seconds and centrifuged at 13000rpm for 1 minute. The volume of upper aqueous phase was measured and transferred to a different microcentrifuge

tube.

Volume of P145T plasmid with thrombin site- 200 $\mu$ L

Volume of Cp3GT plasmid with thrombin site- 200L

A 1/10 volume of sterile 3M sodium acetate was added to each sample. A 20 $\mu$ L volume of 3M sodium acetate was added to each sample. The final volume of the sample was 220 $\mu$ L. To this, 2.5 times the volume of ice cold ethanol was added to the final volume of the samples (220 $\mu$ L x 2.5 = 550 $\mu$ L of 100% ethanol). The samples were then incubated at -80°C for 1 hour followed by centrifugation at 14,000rpm for 15 minutes at 4°C. The supernatant was carefully decanted and the DNA pellet was washed with 250 $\mu$ L of 80% ice cold ethanol and centrifuged at 14,000rpm for 10 minutes at 4°C. The supernatant was carefully poured off and the tubes were dried by inverting the tubes over a Kimwipe. The pellets were dissolved in 20 $\mu$ L of sterile water and tubes were vortexed and centrifuged for 20 seconds to spin down the DNA. 1 $\mu$ L of the purified DNA was resuspended in 19 $\mu$ L of sterile water and concentration was measured using Nanodrop at 260nm. The concentration of different samples after phenol-chloroform extraction is shown in the table B.8.

Table B. 8 Phenol chloroform extraction. Concentration of phenol-chloroform extracted wild type Cp3GT and mutant P145T plasmid having thrombin cleavage site.

Sample	Concentration ng/ $\mu$ L
P3-a (P145T)	2033.0
W3-b (Cp3GT)	2722.7

#### Competent cell preparation for transformation of DNA into yeast

A starter X-33 Mut<sup>+</sup> yeast cells stored at -80°C was inoculated into a 3mL YPD media and incubated at 30°C for 22 hours at 250rpm. A 200 $\mu$ L of overnight culture was transferred to

200mL of YPD media in a 1L flask and incubated overnight at 30°C at 250rpm. The cells were grown to an OD<sub>600</sub> of 1.3-1.4. The overnight culture was transferred to 4 different 50mL centrifuge tubes and centrifuged for at 2,000 x g for 5 minutes at 4°C. The supernatant was poured off and the cell pellets were resuspended in 200mL of ice cold sterile water and centrifuged to wash off the pellet. The pellets were again resuspended in 8mL of ice cold 1M sorbitol and centrifuged at 2000 x g for 5 minutes at 4°C. The supernatant was poured off and the cell pellets were again resuspended in 400µL of ice cold 1M sorbitol and transferred to sterile 2mL microcentrifuge tubes placed on ice. Electroporation cuvettes (Bio-rad) were placed on ice. An 80µL volume of the yeast cells was mixed with 1µg of the purified DNA. In a separate microcentrifuge tube incubated on ice. The mixture was transferred to Electroporation cuvettes placed on ice and incubated on ice for 5 minutes. The mixture was pulsed at 1.5kV in the Electroporation cuvettes using the Bio-Rad micro pulser<sup>TM</sup> using setting for *Pichia pastoris*. A 1 mL aliquot of ice cold sorbitol was added immediately to the cuvettes and the entire contents were transferred to a 15mL sterile glass tube and incubated at 30°C incubator for 90 minutes without shaking. Aliquots of 25µL, 50µL, 75µL, and 100µL of the transformed cells were spreaded on YPDS plates containing 100µg/mL of zeocin. The plates were incubated at 28°C for 48-72 hours. Colonies were seen after 48 hours of incubation. Three colonies from each plate were re-streaked in YPDS plates and incubated at 28°C for 72 hours.

## APPENDIX C: SUPPLEMENTARY FIGURES

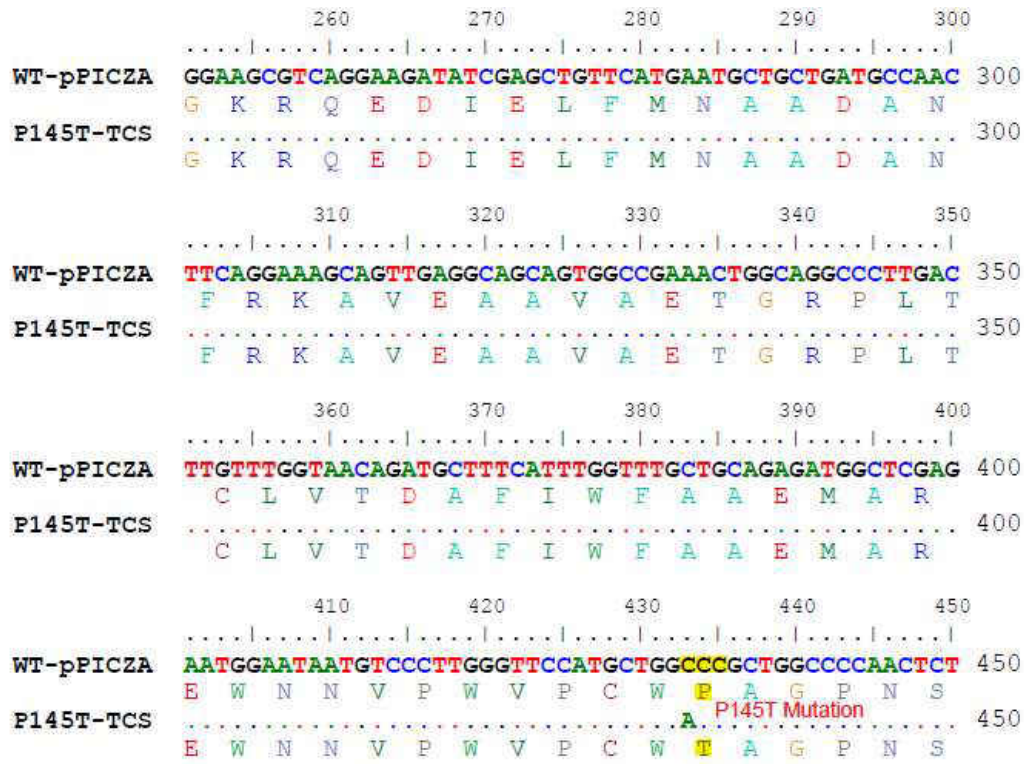


Fig. C. 1 Multiple sequence alignment of wild type Cp3GT and VvGT using Biomedit.

The base pairs highlighted with yellow shows the only amino acid difference between the wild type Cp3GT and mutant P145. Cytosine is substituted to Adenosine.

### Insertion of thrombin site

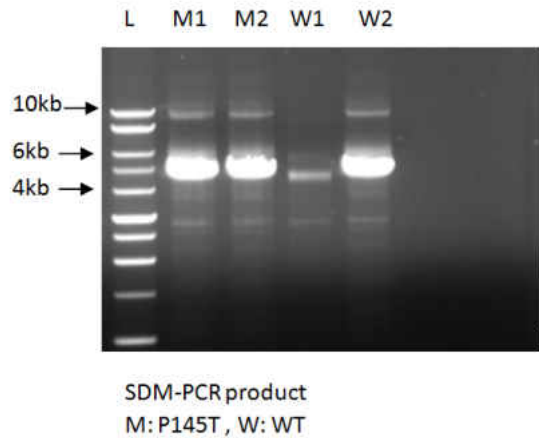


Fig. C. 2 Electrophoresis of site directed mutagenesis PCR product.

Agarose gel electrophoresis of site-directed mutagenesis PCR product to insert thrombin cleavage site in to the wild type Cp3GT and mutant P4145T in between the enzyme and C-myc tags. L- 1kb DNA ladder, M1 and M2- mutant P145T, W1 and W2- Wild type Cp3GT.

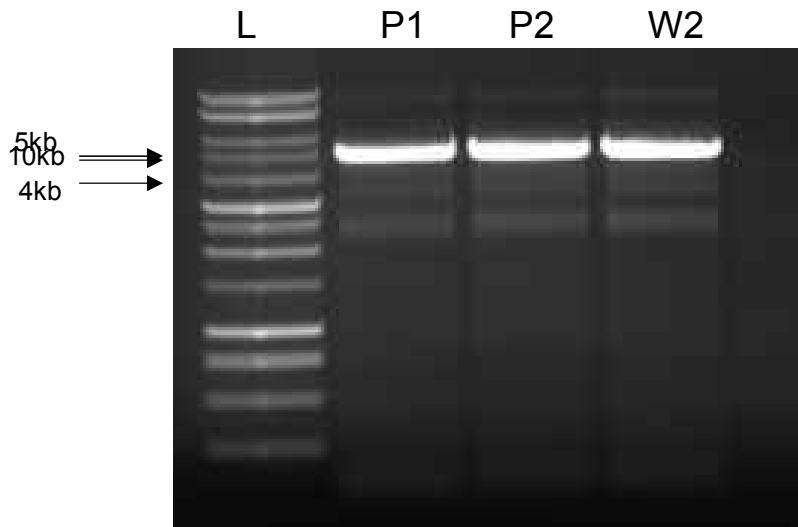


Fig. C. 3 DpnI digestion of site-directed PCR product to insert thrombin site.

Agarose gel electrophoresis of DpnI digested product of site-directed mutagenesis PCR having thrombin cleavage site in between the enzyme and C-myc tags. L- 1kb DNA ladder, P1 and P2- mutant P145T, W2 and W2- Wild type Cp3GT.



## VITA

### SANGAM KANDEL

- Personal Data:                      Date of Birth: July 14, 1989  
    Place of Birth: Bharatpur, Nepal
- Education:                            B. Tech, Biotechnology  
    Kathmandu University, Dhulikhel,  
    Nepal, 2008-2012
- M.S. Biology  
    East Tennessee State University, Johnson City,  
    TN, USA. 2014-2016.
- Professional Experience:          Graduate Research Assistant  
    Biological Sciences, East Tennessee State University  
    Dr. McIntosh Lab  
    2015-2016
- Teaching Assistant  
    Department of Biological Sciences,  
    East Tennessee State University, Johnson City, TN, USA.  
    2014
- Presentations:                      Sangam Kandel, Shivakumar P. Devaiah, Cecilia A. McIntosh.  
    “Structural and Functional Analysis of Grapefruit Flavonol-  
    Specific-3-*O*-GT Mutant P145T. Presented at 55<sup>th</sup> Annual  
    Meeting of Phytochemical Society of North America, Davis, CA,  
    USA. August, 2016.
- Sangam Kandel, Shivakumar P. Devaiah, Cecilia A. McIntosh.  
    “Biochemical characterization of a Cp3GT mutant P145T and  
    study of the tags effect on GT activity”. Presented at Appalachian  
    Student Research Forum, Tennessee, USA. April, 2016.
- Sangam Kandel, Shivakumar P. Devaiah, Cecilia A. McIntosh.  
    “Structural and functional analysis of Grapefruit Cp3GT mutant  
    P145T”. Presented at 54<sup>th</sup> Annual Meeting of Phytochemical  
    Society of North America, Urbana-Champaign, IL, USA. August,  
    2015.
- Sangam Kandel, Shivakumar P. Devaiah, Cecilia A. McIntosh.

“Effect of mutant P145T on the substrate and regiospecificity of a flavonol specific-3-*O*-glucosyltransferase from *Citrus paradisi*”. Presented at the Appalachian Student Research Forum, Tennessee, USA. April, 2015.

Sangam Kandel. “Effect of point mutation P145T on the substrate and regiospecificity of a flavonol specific 3-*O*-glucosyltransferase from *Citrus paradisi*”. Seminar presented to the Department of Biological Sciences, East Tennessee State University, Johnson city, TN, USA. April, 2015.

Honors and Award:

**Best Poster Presentation-Natural Sciences**  
Appalachian Student Research Forum, ETSU  
April 2016

**Graduate Student Research Grant**  
School of Graduate Studies and the ETSU Graduate Council  
March 2016

**Frank and Mary Loewus Student Travel Award**  
Phytochemical Society of North America  
August 2016 and August 2015

**Graduate and Professional Students Association Travel Award**  
Graduate and Professional Students Association,  
East Tennessee State University  
August 2016 and August 2015

Professional Mentorships:

Member Phytochemical Society of North America (PSNA)  
Member ETSU Graduate and Professional Student Association (GPSA)