

Université Bordeaux 1

Année 2007

Thèse n° 3347

THESE

pour le

DOCTORAT DE L'UNIVERSITE BORDEAUX 1

Mention : Sciences du Vivant, Géosciences, Sciences de l'Environnement

Option : Sciences des Aliments

Présentée et soutenue publiquement

le 15 février 2007

Par **Alain-Pierre PETIT**

Né le 18 mai 1978 à Bergerac

**Etude structurale de la dihydroflavonol 4-reductase,
enzyme clé de la biosynthèse des flavonoïdes
chez *Vitis vinifera***

Membres du jury :

A. AUBRY	Directeur de recherche, Université H.Poincaré, Nancy	Rapporteur
J-M. MERILLON	Professeur, Université Bordeaux 2	Rapporteur
S. DELROT	Professeur, Université Bordeaux 2	Président
S. QUIDEAU	Professeur, Université Bordeaux 1	Examineur
JM. SCHMITTER	Professeur, Université Bordeaux 1	Examineur
B. GALLOIS	Directeur de recherche, Université Bordeaux 1	Directeur de thèse
M. BARTHE	Directrice du service technique du CIVB	Membre invité

REMERCIEMENTS

Ces travaux de recherche ont été réalisés dans le Laboratoire de Biophysique Structurale, UMR 5471 – CNRS, sous la direction de Monsieur A. Brisson. Je lui exprime toute ma gratitude pour m'avoir accueilli au sein de son laboratoire.

Les recherches ont été menées sous la direction de Monsieur B. Gallois. Je tiens à lui exprimer mon profond respect et à lui témoigner ma reconnaissance pour ses précieux conseils, son esprit critique et son soutien.

Je remercie Messieurs S. Delrot, S. Quideau et JM. Schmitter pour avoir accepté de composer le jury de ma thèse, et Messieurs A. Aubry et JM. Mérillon pour l'honneur qu'ils me font en acceptant d'être les rapporteurs de mon travail de recherche.

Je tiens à exprimer ma gratitude à Mademoiselle B. Langlois d'Estaintot pour l'ampleur et la qualité du travail de cristallisation qu'elle a effectué tout au long de ses trois années.

Je remercie Monsieur T. Granier pour son aide précieuse et pour m'avoir initié à la cristallographie des protéines.

J'adresse mes remerciements à Monsieur J. Chaudière pour ses conseils avisés, tant en enzymologie qu'en chimie organique. Je lui suis très reconnaissant et lui témoigne un profond respect.

Un grand merci à Monsieur C. Manigand pour son soutien et ses encouragements ainsi que pour les conversations scientifiques prolifiques que nous avons eues.

Je remercie Messieurs P. Picard et B. Busetta pour leurs participations et leurs conseils bénéfiques.

Je remercie également tous les membres de l'équipe Biologie de la Vigne de l'UMR Physiologie et Biotechnologie Végétale de l'INRA et tout particulièrement Monsieur S. Hamdi, co-directeur de cette thèse pour sa confiance et son implication enthousiaste et dynamique dans ces travaux.

Un grand merci à Madame C. Le Henaff pour son accueil au sein de son laboratoire, sa gentillesse et la qualité de ses conseils scientifiques.

J'adresse mes remerciements au Conseil Interprofessionnel du Vin de Bordeaux pour son aide financière durant ces trois années de thèse.

Enfin, je dédie ce mémoire à mes parents, ma sœur et mes proches qui m'ont toujours encouragé et soutenu et tout particulièrement Christel pour son soutien de tous les instants.

ABBREVIATIONS

ADN	Acide désoxyribonucléique
ADNc	Acide désoxyribonucléique complémentaire
AnoCmpl	Complétude sur les données anormales
AnoMlt	Multiplicité anormale
ANR	Anthocyanidine réductase
ANS	Anthocyanidine synthase
ARN	Acide ribonucléique
ARNm	Acide ribonucléique messenger
ATP, ADP, AMP	Adénosine 5' tri-, di-, mono- phosphate
ATPase	ATP synthase
Bis Tris	Bis(2-hydroxyéthyl)amino-tris(hydroxyméthyl)méthane
CC	Coefficient de corrélation
CHI	Chalcone isomérase
CHS	Chalcone synthase
CoA	Coenzyme A
DFR	Dihydroflavonol réductase
DFR-1C	Simple mutant de DFR (N133D)
DFR-4A	Quadruple mutant de DFR (N133D, Q135T, Q138T et Y142I)
DHK	Dihydrokaempférol
DHM	Dihydromyricétine
DHQ	Dihydroquercétine
DNase	Désoxyribonucléase
dNTP	Désoxynucléotide 5' triphosphate (dATP, dCTP, dGTP, dTTP)
DO	Densité optique
DTPA	Diéthylène triamine pentaacétate
DTT	Dithiotréitol
EDTA	Ethylène diamine tétraacétate
ESRF	European Synchrotron Radiation Facility ou Installation Européenne de Rayonnement Synchrotron
FLS	Flavonol synthase
FOM	Figure de mérite
F3H	Flavanone-3-hydroxylase
F3'H	Flavonoïde 3'-hydroxylase
F3'5'H	Flavonoïde 3',5'-hydroxylase
Glc	Glucose
GPC	Chromatographie par perméation de gel
HEPES	Acide éthane-sulfonique hydroxyéthyl-piperazine
HLH	Protéine à motif hélice-boucle-hélice

HPLC	Chromatographie liquide haute performance
IPTG	Isopropyl- β -thiogalactopyranoside
LAR	Leucoanthocyanidine réductase
LC-MS	HPLC couplée à la spectrométrie de masse
LC-MS/MS	Spectrométrie de masse couplée à une LC-MS
LDL	Lipoprotéines à faible densité
LDOX	Leucoanthocyanidine dioxygénase
MALDI-TOF	Désorption/ionisation par laser assistée par matrice avec analyseur à temps de vol
MES	Acide morpholino-2-éthanesulfonique
MM	Masse molaire
MS	Spectrométrie de masse
NADH	Nicotinalide adénine dinucléotide (forme réduite)
NADPH	Nicotinalide adénine dinucléotide phosphate (forme réduite)
NADP ⁺	Nicotinalide adénine dinucléotide phosphate (forme oxydée)
NO	Monoxyde d'azote
PCR	Réaction de polymérisation en chaîne
PEG	Polyéthylène glycol
PPiase	Pyrophosphatase inorganique
PRP	Protéine riche en proline
Ranom	Facteur de confiance anomal
Rsym	Facteur de confiance global
SDR	Réductase à chaîne courte
SDS	Sodium dodécyl sulfate
SSCR	Carbonyl reductase de <i>Sporobolomyces salmonicolor</i>
PAGE	Electrophorèse sur gel de polyacrylamide
TEMED	N,N,N',N'-tétraméthyl-éthylènediamine
TFA	Acide trifluoro-acétique
TNF α	Facteur de nécrose tumorale
Tris	Tris(hydroxyméthyl)aminométhane acétate
UDP	Uridyl diphosphate
UFGT	Uridine diphosphate glucose-flavonoïde 3-O glucosyl transférase
UV	Ultraviolet
VLDL	Lipoprotéines de très basse densité
2D	Bidimensionnel
3D	Tridimensionnel

TABLE DES MATIERES

INTRODUCTION.....	1
I. Vins et économie mondiale.....	1
I.1. Production mondiale de vins.....	1
I.2. Consommation mondiale.....	1
I.3. Equilibre Production – Consommation mondiale.....	2
II. Vins et santé.....	4
II.1. L'influence de l'alcool dans le « paradoxe français ».....	4
II.2. Participation des composés polyphénoliques à la prévention des maladies cardio-vasculaires.....	5
II.3. Le vin rouge : un remède préventif de bon goût ?.....	8
III. Les composés phénoliques des raisins et des vins.....	8
III.1. Les stilbènes, composés non-flavonoïdes.....	10
III.2. Les flavonoïdes.....	10
IV. Evolution des composés phénoliques au cours du vieillissement des vins rouges.....	15
IV.1. Synthèse des nouveaux pigments.....	16
IV.2. Autres pigments.....	19
IV.3. Incidences de structures complexes sur le vieillissement du vin.....	19
IV.4. Vieillissement des vins et propriétés organoleptiques.....	21
V. Voies de biosynthèse des flavonoïdes.....	23
V.1. Les enzymes ubiquitaires.....	28
V.2. Voie latérale des flavonols.....	32
V.3. Voie latérale des anthocyanes.....	33
V.4. Voie latérale des tanins condensés.....	36
V.5. Fonction et mécanismes du transport vacuolaire des flavonoïdes.....	38
VI. Régulation moléculaire de la voie de biosynthèse des flavonoïdes.....	39
VI.1. Généralités.....	39
VI.2. Approche structurale des protéines Myb.....	40
VII. Conclusion et Projet de thèse.....	42

RESULTATS	45
I. Première approche structurale par analyse bioinformatique.....	54
I.1. Etude du site de fixation du co-facteur	54
I.2. Etude du site catalytique	54
I.3. Reconnaissance et fixation du substrat	54
II. La dihydroflavonol réductase (DFR).....	55
II.1. Expression hétérologue de la DFR	55
II.2. Production de la DFR sauvage et conditions de purification	56
II.3. Caractérisation enzymatique.....	63
II.4. Cristallisation et qualité de diffraction	75
II.5. Etudes structurales	79
II.6. Validation de la structure obtenue	90
II.7. Structure de la DFR sauvage	92
II.8. Justification du choix des mutations.....	104
III. Détermination de structure d’analogues de substrat.....	105
III.1. Traitement des données	106
III.2. Descriptions structurales	110
 CONCLUSIONS ET PERSPECTIVES	 121
 MATERIEL ET METHODES	 141
I. Matériel biologique	141
I.1. Banque d’ADNc.....	141
I.2. Souches bactériennes <i>Escherichia coli</i> :	141
I.3. Souche de levure <i>Pichia pastoris</i>	141
II. Biologie Moléculaire	141
II.1. Amplification génique ou PCR (“polymerase chain reaction”)	141
II.2. Clonage moléculaire	141
II.3. Analyse de l’ADN par électrophorèse sur gel d’agarose.....	141
II.4. Extraction d’ADN plasmidique	141
II.5. Transformation de bactéries chimio-compétentes	141
II.6. Transformation de <i>Pichia pastoris</i>	141
II.7. Mutagenèse dirigée	141
III. Méthodes d’expression et d’analyse des protéines.....	141
III.1. Expression des protéines chez <i>E. coli</i>	141
III.2. Marquage de protéine au Sélénium	141
III.3. Préparation d’extraits cellulaires d’ <i>E. coli</i>	141
III.4. Electrophorèse monodimensionnelle en conditions dénaturantes (SDS-PAGE)..	141
III.5. Electrophorèse sur gel d’acrylamide en conditions non dénaturantes (PAGE) ...	141
III.6. Purification des protéines recombinantes.....	141

IV. Méthodes d'étude enzymatique	141
IV.1. Contrôles	141
IV.2. Activité et colorimétrie	141
IV.3. Enzymologie et fluorimétrie	141
V. Cristallisation.....	141
V.1. Cristallisation en goutte suspendue par diffusion de vapeur	141
V.2. Cryoconservation des cristaux.....	141
VI. Cristallographie	141
VI.1. Diffraction des Rayons X et traitement des données	141
VI.2. Méthodes de détermination des phases	142
VI.3. Amélioration des cartes de densité électronique.....	145
VI.4. Validité du modèle	147
REFERENCES BIBLIOGRAPHIQUES	149
ANNEXES	163

INTRODUCTION

INTRODUCTION

I. Vins et économie mondiale

I.1. Production mondiale de vins

La production mondiale de vin a culminé en début de décennie 1980 où elle a atteint 333,6 millions d'hectolitres par année. Après une régression importante dont le paroxysme fut atteint en 1995, la production mondiale n'a depuis jamais été inférieure au niveau enregistré cette année là, et ce, sous l'influence notamment de l'accroissement des surfaces de culture dans les pays du Nouveau Monde. La production mondiale de vin reste dominée par l'Europe qui représentait près de 71 % de celle-ci en 2004. A cette même période, l'Amérique (16 %), l'Asie (4,5 %) et l'Océanie (5,1 %), constituaient les autres principaux bassins de la production mondiale de vins.

I.2. Consommation mondiale

Après avoir atteint une valeur maximale en fin des années 1970 à près de 286 millions d'hectolitres, la consommation mondiale de vin a fortement régressé jusqu'en 1996 (**Tableau 1**). Depuis, un renversement lent de tendance a pu être mesuré grâce à l'élévation de la consommation des pays néo-producteurs tels que l'Asie et l'Océanie, et des pays non producteurs tels que le Royaume-Uni et les pays scandinaves. Les pays producteurs et consommateurs préférentiels de vin (France) connaissent quant à eux une chute de la consommation individuelle même si celle-ci reste la plus élevée au monde (*3^{ème} assemblée générale de l'Organisation Internationale de la Vigne et du Vin, 2005*).

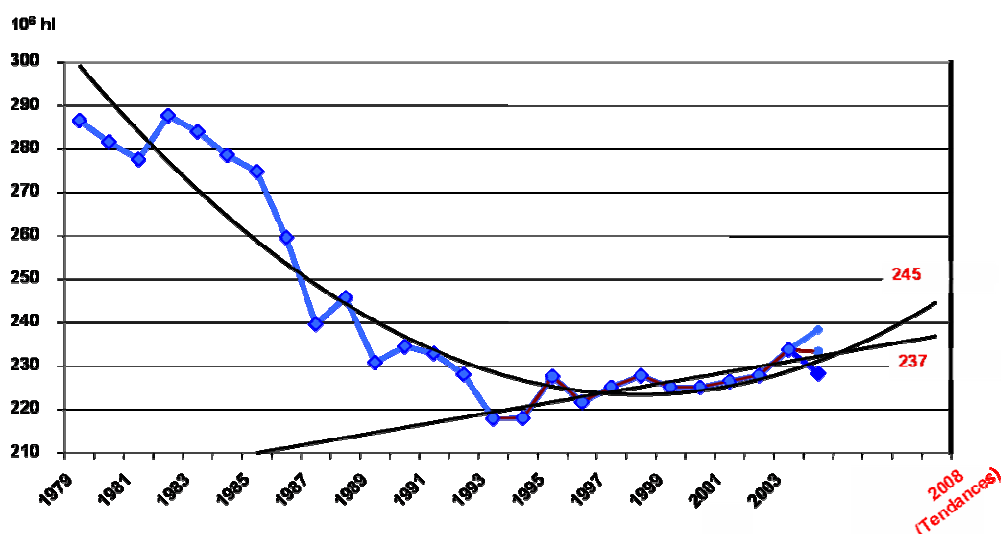


Tableau 1 : Evolution de la consommation mondiale de vin

I.3. Equilibre Production – Consommation mondiale

Différentes études témoignent du dynamisme du marché mondial du raisin. Sa croissance semble vouloir se poursuivre puisque le marché mondial en 2002 se situe à +10,4 % par rapport à la moyenne 1996-2000 (**Tableau 2**). Cependant, l'examen des contributions continentales sur l'évolution du marché mondial du raisin montre que la fraction relative de l'Europe à la croissance des échanges mondiaux a été divisée par deux depuis la fin de la décennie 1980, passant de 2/3 à 1/3 de la croissance du marché mondial.

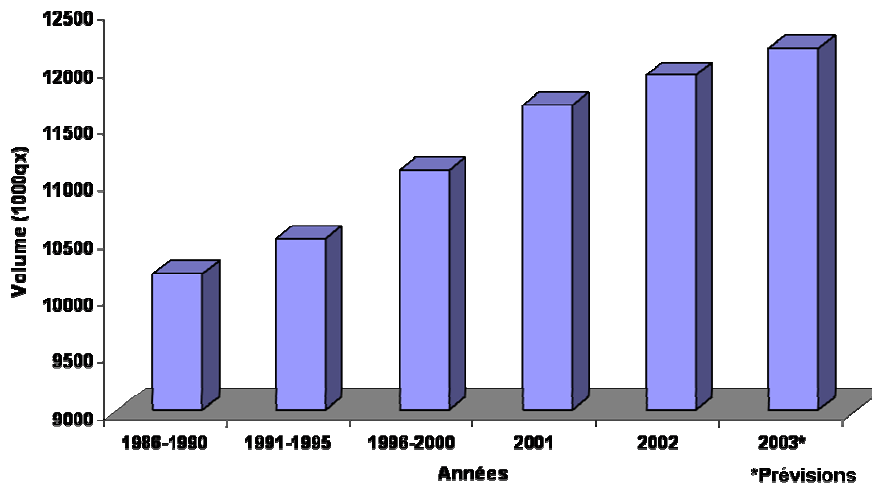


Tableau 2 : Evolution de la consommation mondiale de raisin. (3^{ème} assemblée générale de l'Organisation Internationale de la Vigne et du Vin, 2005)

Cette chute s'explique en partie par l'apparition des pays néo-producteurs, mais faiblement consommateurs du fait de la faiblesse de leurs marchés intérieurs. Il existe donc un excédent de production mettant en évidence la nécessité des échanges commerciaux entre les principaux pays producteurs, et une nécessité pour les pays exportateurs, adossés à un marché intérieur dominant, d'optimiser la qualité afin d'inciter à l'achat des produits nationaux (**Tableau 3**).

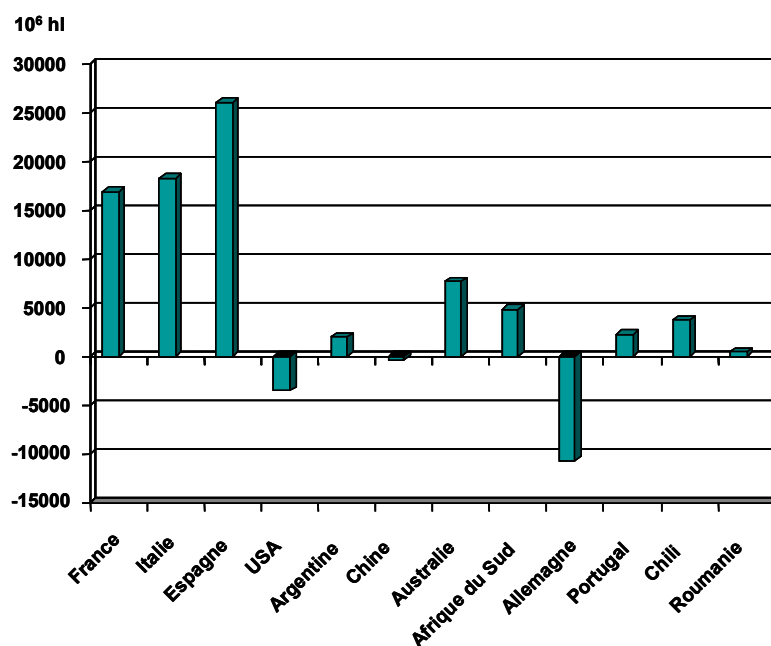


Tableau 3 : Différence entre la production et la consommation de vin pour les principaux pays producteurs. (3^{ème} assemblée générale de l'Organisation Internationale de la Vigne et du Vin, 2005)

En effet, les vins issus des pays néo-producteurs (tel que le Chili) séduisent par leurs qualités intrinsèques (fruité, légers...) mais aussi par leurs reproductibilités au cours des années. Cette qualité est imputable aux conditions naturelles du pays, telles que le climat et le sol. Le climat de type méditerranéen tempéré, avec des étés longs et secs, une amplitude de variation des températures quotidiennes qui dépasse les 15°C et une humidité relative de 55 à 60 %, permet une formation saine des racines et une accumulation de sucres et d'arômes pendant une saison de longue maturation. De plus, la majorité des sols dans la région vinicole du Chili par exemple est d'origine alluviale ou volcanique. Leurs variations en nutriments contribuent elles aussi à la qualité régionale des vins.

Enfin, les vins chiliens possèdent les taux les plus hauts de flavanols, molécules antioxydantes, qui leur confèrent un atout supplémentaire, d'ordre sanitaire.

L'implication du vin dans la diminution des risques des maladies cardio-vasculaires paraît donc être un moyen moderne de susciter l'intérêt du consommateur actuel ou à venir, et ainsi, de stimuler la consommation de vin.

II. Vins et santé

II.1. L'influence de l'alcool dans le « paradoxe français »

En 1992, sont exposés les résultats d'une étude comparative portant sur la probabilité d'apparition de maladies cardio-vasculaires chez deux populations sujettes aux mêmes facteurs de risques (sédentarité, tabagisme, manque d'exercice...). Cette étude démontre que la consommation modérée d'alcool sous forme de vin réduit de 40 % l'apparition de crises cardiaques par rapport à une population identique, qui consomme tout autre type d'alcool (*Renaud et al, 1992*). Ce phénomène est appelé le « paradoxe français ». Différentes études épidémiologiques ont depuis confirmé ces résultats (*Gronbaek et al, 2000*). Elles traduisent un risque plus faible d'apparition des maladies cardio-vasculaires pour une population consommatrice de vin. Globalement, une consommation modérée de vin permet une réduction de l'hypertension, prévient l'apparition de l'infarctus du myocarde, de l'ischémie coronarienne et des accidents vasculaires cérébraux ischémiques (*Rasmussen et al, 2005*), par opposition aux buveurs excessifs et aux abstinents. Plus précisément, la consommation modérée d'alcool (uniquement sous forme de vin) entraîne l'augmentation du taux plasmatique des lipoprotéines de haute densité de type 2 et 3, et la synthèse de leurs apolipoprotéines A1 et A2 (*Imhof et al, 2003*), réduisant ainsi la formation d'athérosclérose.

L'influence de l'alcool sur la dégradation du cholestérol dans le foie n'explique que pour moitié son effet protecteur, l'autre moitié étant imputable à son action sur les paramètres de l'hémostase à savoir la coagulation et la fibrinolyse.

L'ingestion d'alcool est associée (*Martin et al, 2001*) à la diminution du fibrinogène plasmatique, favorisée par les synthèses accrues du facteur VII, de l'activateur du plasminogène, et par la stimulation de la fibrinolyse. L'effet protecteur de l'alcool peut également être lié en partie à la réduction de l'agrégation des plaquettes par inhibition de la phospholipase A2. L'inhibition de cette dernière réduirait la libération d'acide arachidonique des membranes des plaquettes et la production de thromboxane A2, qui est un puissant agrégant plaquettaire (*De Lange et al, 2004 ; Salem et al, 2005*). Enfin, *Dimmit et al (1998)* stipulent que l'effet bénéfique de l'alcool sur la coagulation sanguine atteint une limite variable selon la quantité et le type de boisson ingérée.

Par ailleurs, de multiples études expérimentales (*Hertog et al, 1997*) indiquent que le jus de raisin ou des extraits de vin provoquent les mêmes résultats que le vin lui-même, ce qui tend à montrer que les composés spécifiques du vin, les polyphénols, sont impliqués dans cet effet protecteur.

II.2. Participation des composés polyphénoliques à la prévention des maladies cardio-vasculaires

II.2.1. Effets antioxydants et inhibition de l'oxydation des lipoprotéines de basse densité

La chaîne respiratoire produit de manière continue des espèces oxygénées hautement réactives comme l'anion superoxyde (O_2^-) ou le radical hydroxyle ($OH\cdot$), espèces endommageant des lipides, des protéines et des acides nucléiques et impliquées dans diverses pathologies et le vieillissement cellulaire. Les polyphénols ingérés piègent les radicaux libres par transfert d'électrons. Cette faculté n'est possible que grâce à la stabilité relative du groupe phénoxy ainsi créé, et à la formation de composés semi-quinoniques (*Jovanovic, 1998*). Ainsi, les polyphénols protègent les lipoprotéines à faible densité (LDL) de l'oxydation par les radicaux libres. Ces dernières, lorsqu'elles sont oxydées, jouent un rôle important dans le développement de l'athérosclérose via l'accumulation de « noyaux » lipidiques sur la face interne des cellules endothéliales des vaisseaux sanguins. (*Aviram et al, 2002 ; Howard et al, 2002*). Enfin, de récents travaux décrivent les effets de l'association des polyphénols aux LDL conduisant, via une cascade réactionnelle, à la diminution du taux de LDL circulant (*Zern et al, 2005*) (**Figure 1**). Parmi les processus mis en jeu, citons :

- ✓ la diminution du cholestérol absorbé,
- ✓ la synthèse accrue des ARNm des récepteurs aux LDL,
- ✓ une modification des lipoprotéines de très basse densité (VLDL),
- ✓ la réduction de triglycérides plasmatiques.

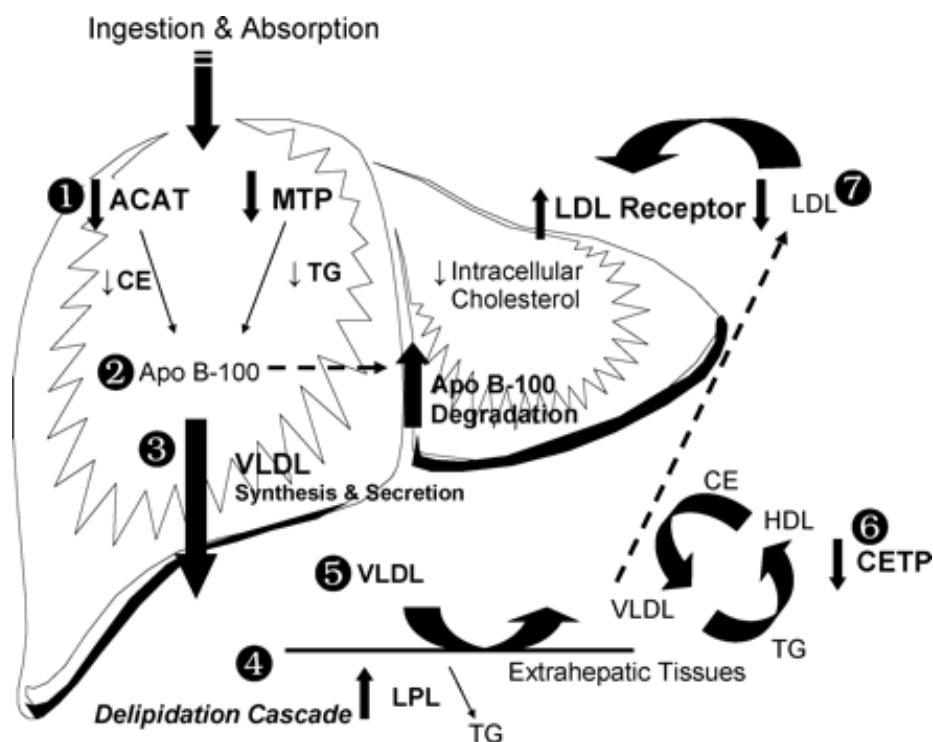


Figure 1 : Description des effets induits par les polyphénols sur le métabolisme hépatique du cholestérol et des lipoprotéines (Zern et al, 2005).

Les polyphénols diminuent l'activité de l'acyl-CoA cholestérol acyltransférase (ACAT) et de la protéine de transfert microsomal (MTP) impliquées dans l'assemblage des lipides dans les lipoprotéines de très basse densité (VLDL) (1). Par manque de constituants lipidiques, la dégradation de l'apo B-100 est accrue (2). La diminution de substrats disponibles stimule dans le même temps la diminution de la synthèse et de la sécrétion des VLDL (3). La voie de délipidation est aussi stimulée par la diminution des VLDL circulants (4). Le remaniement des VLDL (par diminution des apolipoprotéines E), induit l'activité des lipoprotéines lipases (LPL) et diminue ainsi la concentration en triglycérides (TG) et VLDL plasmatiques (5). L'activité de la protéine de transfert des esters de cholestérol (CETP) (6) est aussi affectée par la diminution de substrats disponibles (VLDL). La diminution des VLDL circulants entraîne la réduction du taux de LDL. Enfin, l'altération des concentrations du cholestérol hépatique engendre une stimulation des récepteurs aux LDL pour maintenir l'homéostasie du cholestérol hépatique.

II.2.2. Inhibition de la réponse inflammatoire dans l'athérosclérose

Les processus inflammatoires jouent un rôle essentiel lors de toutes les phases du développement de l'athérosclérose. Les flavonoïdes sont impliqués dans la régulation des monocytes via l'inhibition de l'adhésion induite par le TNF α (facteur de nécrose tumorale α), et ce, en diminuant l'expression des molécules d'adhésions de type 1 aux cellules vasculaires (*Sen et al 2001*). Ces molécules sont des glycoprotéines de surface, inductibles lors des processus d'adhésion cellulaire. Néanmoins, le mécanisme moléculaire permettant cet effet spécifique sur les glycoprotéines d'adhésion des cellules vasculaires et non intercellulaires n'est pas encore élucidé.

II.2.3. Action vaso-relaxante de l'endothélium

L'endothélium joue un rôle essentiel dans la régulation du système vasculaire via la synthèse de monoxyde d'azote (NO). Ce dernier est responsable de la vaso-relaxation de l'endothélium, et contribue à la réduction de l'agrégation plaquettaire et du flux de protéines plasmatiques athérogéniques dans les artères (*Duffy et al, 2003*). Les composés polyphénoliques stimulent la synthèse de NO par mobilisation du Ca²⁺ et activation de la voie des tyrosine kinases (*Martin et al, 2002*). Ils sont donc responsables de la vaso-relaxation de l'endothélium. Ces études ne permettent toutefois pas de définir quelles classes de polyphénols (flavan-3-ols, acides phénoliques ou proanthocyanes) sont plus particulièrement impliquées (*Fitzpatrick et al, 2000*).

II.2.4. Réduction de l'agrégation plaquettaire

La nature du vin consommé est une composante principale dans la réduction de l'agrégation plaquettaire (*Demrow et al 1995*). Seule l'absorption de vin rouge, et non celle de vin blanc, possède un effet significatif sur l'activité plaquettaire et sur la formation du thrombus. Le vin rouge possède donc des composés efficaces contre l'agrégation plaquettaire. Diverses hypothèses sont émises pour expliquer ce phénomène, telles que le potentiel inhibiteur des proanthocyanidines sur la 5-lipoxygénase ou l'effet conjugué de la quercétine et de la catéchine sur l'agrégation plaquettaire lorsque celle-ci est induite par le collagène (*Murphy et al, 2003 ; Pignatelli et al, 2000*).

II.3. Le vin rouge : un remède préventif de bon goût ?

Les résultats conjugués des études épidémiologiques et des travaux scientifiques précédents démontrent largement l'impact de la consommation de vin rouge sur la santé. Néanmoins, les effets trouvés sont inexistantes lors de l'absorption de tout autre alcool (vin blanc,...) et ne peuvent être détectés lors de comportements abstinent et excessif. Rappelons qu'une surconsommation d'alcool entraîne de nombreuses maladies neurologiques mais qu'une absorption modérée de vin réduit les risques de démence de 80 % et de maladie d'Alzheimer de 75 % (*Orgogozo et al 1997*). Enfin, nous devons signaler une réduction de la prolifération des cellules du cancer du sein et de la prostate associée à la prise, toujours modérée de substances phénoliques (*Kampa et al, 2000*).

L'impact sur la santé de la consommation modérée de vin rouge, et donc de l'absorption de composés polyphénoliques, ne peut plus être mis en doute. Néanmoins, les travaux précédents ne permettent pas de définir quelle(s) classe(s) de composés polyphénoliques est (sont) plus particulièrement impliquée(s). De telles études sont actuellement en cours. Leurs difficultés de mise en œuvre résident dans la diversité des composés à étudier.

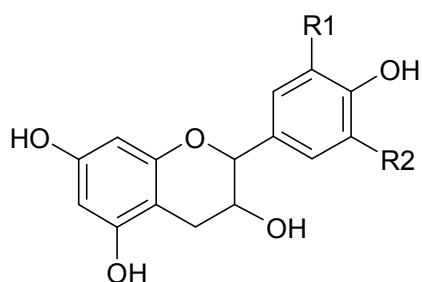
III. Les composés phénoliques des raisins et des vins

Les composés phénoliques sont présents dans tous les tissus végétaux, où ils exercent des fonctions très diverses. Leur synthèse a lieu dans tous les organes, mêmes jeunes, et dérive d'une voie du métabolisme du sucre. Les premiers travaux portant sur les polyphénols du raisin datent des années 1965-1969 (*Ribereau-Gayon, 1965, 1969*). Ces composés interviennent dans la couleur des vins rouges et participent à leur saveur et à leurs arômes. Ils se divisent en plusieurs familles et comprennent majoritairement les stilbènes, les anthocyanes, les tanins et les flavonols (**Figure 2**).

Ces différentes familles sont décrites dans les paragraphes III-1 et III-2. Sont rapportés pour chacune d'elles, leurs structures, leurs rôles ainsi que différentes précisions les concernant telles que leurs localisations et teneurs dans la plante.

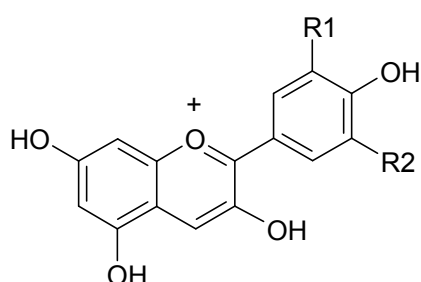
Lors de la vinification en rouge, les composés phénoliques se solubilisent dans le moût. Deux familles, les anthocyanes et les tanins, jouent un rôle particulier quant à la qualité d'un vin. En effet, elles conduisent lors de leur vieillissement à de nouveaux complexes dont les structures évoluent au cours de la macération (*Somers, 1966*). Ces complexes sont d'importance. Ils sont responsables de la couleur et des qualités organoleptiques des vins.

Dans un nouveau paragraphe, nous nous intéresserons à la composition des vins durant leur vieillissement et présenterons les différentes réactions qui y sont associées.

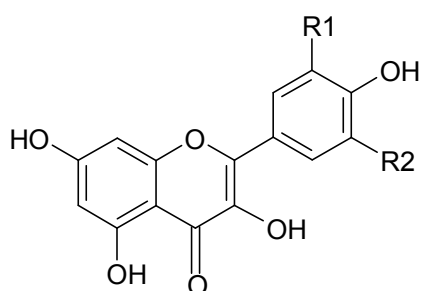


Flavan-3-ols

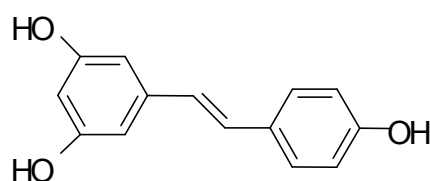
Molécule élémentaire des tanins condensés



3-OH Anthocyanidines



Flavonols



Trans-resvératrol

Les stilbènes se caractérisent par un squelette $C_6-C_2-C_6$. Cette classe de composés est ici représentée par l'un de ses membres : le Trans-resvératrol

Figure 2 : Les grandes classes de polyphénols (Winkel-Shirley, 2001)

Issus de la voie de biosynthèse des phénylpropanoïdes qui interviennent dans la formation de plusieurs familles de métabolites secondaires, les composés phénoliques sont synthétisés suite à une succession de réactions de condensation décarboxylative d'une molécule de coumaryl-CoA et de trois molécules de malonyl-CoA. Ils englobent à la fois les stilbènes caractérisés par un squelette carboné de type $C_6-C_2-C_6$ et les flavonoïdes (anthocyanes, flavonols, tanins condensés) de structure $C_6-C_3-C_6$ (**Figure 2**).

III.1. Les stilbènes, composés non-flavonoïdes

Différentes formes de molécules existent. Outre les aglycones (entités non couplées à des sucres), on rencontre également leurs équivalents glycosylés comme le trans-picéide et l'astringine. Le resvératrol (3,4',5-trihydroxystilbène) est la molécule aglycone la plus largement étudiée de par l'influence de ses propriétés sur la santé humaine (*Bavaresco et al, 1999*). Le resvératrol et ses analogues préviennent des maladies cardio-vasculaires. Les mécanismes moléculaires mis en jeu sont similaires à ceux décrits au paragraphe II.2. Outre ces propriétés cardio-vasculaires, les stilbènes présentent une action chimio-protectrice vis-à-vis de certains cancers en interférant sur la transduction du signal. Ils modulent ainsi la synthèse des protéines du cycle cellulaire et induisent l'apoptose des lignées cancéreuses (*Ulrich, 2005*).

La teneur en resvératrol est faible (1-3 mg/L) dans le vin rouge (*Vitrac et al, 2005*), comparativement à celle de sa forme glycosylée (le trans-picéide), qui avoisine 30 mg/l. Si comme le sont les flavonoïdes, le trans-picéide est déglycosylé par les β -glucosidases des cellules épithéliales de l'intestin grêle, la concentration en resvératrol devient, dans l'organisme, très supérieure à celle initialement contenue dans le vin (*Nemeth et al, 2003*). Cette possible augmentation de la concentration en resvératrol peut expliquer ainsi l'influence bénéfique de l'absorption de vin sur la santé.

Chez la vigne, les stilbènes sont des constituants des organes ligneux tels que les racines et les tiges. Ils sont inductibles lors d'un stress dans les fruits et les feuilles. Ils agissent dès lors comme des phytoalexines lors des mécanismes de résistance à certains pathogènes ou aux UV. La synthèse de resvératrol dans la baie est catalysée par la stilbène synthase, enzyme qui utilise les mêmes substrats (coumaryl-CoA et malonyl-CoA) que la chalcone synthase pour la production de chalcone, composé précurseur des flavonoïdes. Bien qu'un stress stimule la production de resvératrol, celle-ci diminue donc lors de la maturation de la baie, compte tenu de la compétition entre les deux enzymes. Il s'ensuit, à partir de la véraison, une diminution de la réponse à un stress. Ceci explique une plus grande vulnérabilité de la vigne à partir de ce stade de développement (*Versari et al, 2001*).

III.2. Les flavonoïdes

III.2.1. Les anthocyanes

Les anthocyanes regroupent l'ensemble des molécules aglycones (les anthocyanidines) et leurs hétérosides, les anthocyanines. Ces composés possèdent en commun un squelette flavylum (ou 2-phényl-1-benzopyrilium) et se différencient par la nature des fonctions R1 et R2 du cycle B (**Figure 3**), le nombre de glucoses estérifiés, le taux d'acylation de ces glucoses et la nature des acides estérifiés.

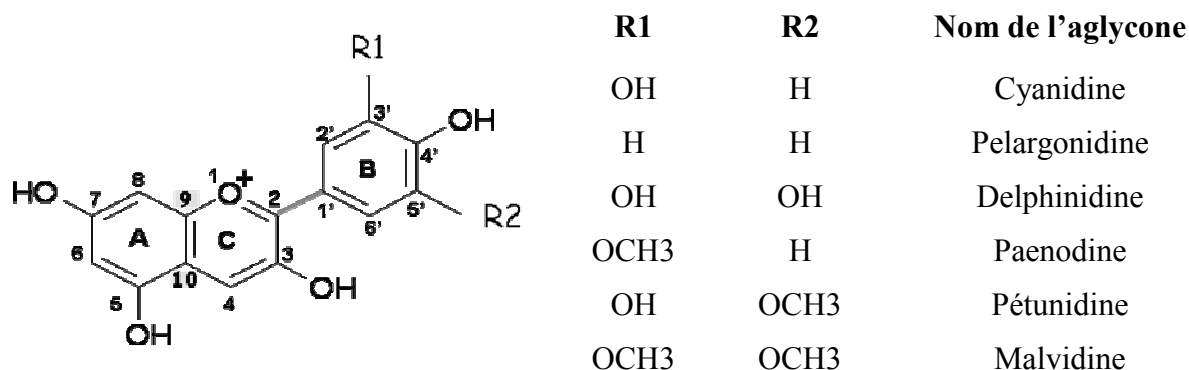


Figure 3 : Structure des anthocyanidines du raisin et du vin

Chez la vigne, 16 anthocyanines ont pu être identifiées. Les substituants et le nom de leur forme aglycone sont présentés **Figure 3**. Leur synthèse a lieu dans les feuilles, les tendrons et dans la pellicule de la baie. Elle est activée à la véraison, se stabilise par la suite avant de régresser en fin de maturation (*Roggero et al, 1986*). La quantité et la nature des différentes anthocyanines varient en fonction du cultivar, du climat et de la région de production. Leur synthèse peut être stimulée suite à un stress abiotique. Ces composés sont responsables de multiples fonctions chez la plante, telles la capture des espèces réactives oxygénées, la prévention contre les herbivores et la protection contre l'excès de lumière ou la chute des températures en réduisant ainsi le risque de photo-inhibition (*Schaefer et al, 2006*).

Les anthocyanes, en solution aqueuse, existent sous la forme d'un mélange de quatre structures résonantes qui se différencient en fonction du pH. Chacune d'elles présente une gamme d'absorption de la lumière différente. Pour un pH compris entre 1 et 3, le cation flavylum est rouge, alors que la pseudo base carbinole générée à pH 4-5 est incolore, et que la base quinoïdale formée à pH 7-8 est de couleur bleue/violette (**Figure 4**). La forme carbinole est en équilibre avec la chalcone, jaune à un pH équivalent (*Lapidot et al, 1999*).

Dans les cépages rouges, à partir de la véraison, le « cation flavylum » est la forme prépondérante compte tenu du pH de la baie. Il confère donc à cette dernière sa coloration. Dans les baies de cépages blancs, aucun anthocyane n'est détecté (*Boss et al, 1996*) bien que leurs précurseurs soient présents. Ceci résulte de la non expression du gène codant pour l'UDP glucose-flavonoïde 3-O-glucosyltransferase (UFGT), enzyme, nous le verrons, responsable de la synthèse des anthocyanes.

Lorsqu'ils sont ingérés par l'homme, ces composés se retrouvent dans l'intestin, milieu de pH élevé où ils sont absorbés par le système sanguin. Les anthocyanines se présentent donc sous forme de bases quinoïdales qui inhibent significativement la peroxydation des lipides. Ceci montre la réactivité de telles formes malgré la perte de l'ion flavylum.

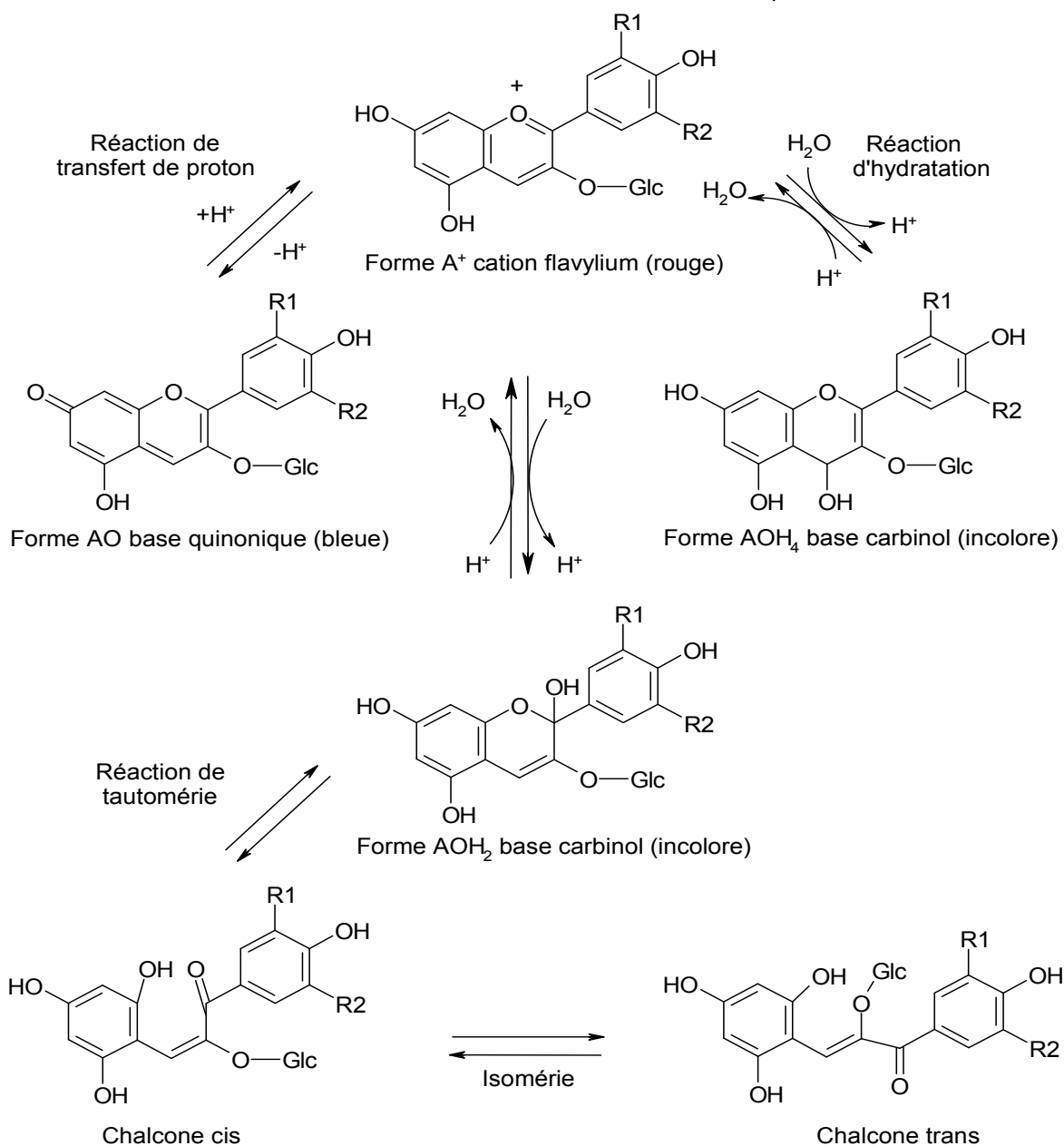


Figure 4 : Structures résonantes des anthocyanes en fonction du pH (Brouillard et al, 1978)

III.2.2. Les tanins condensés ou proanthocyanidines

De par leur composition chimique, les tanins peuvent être répartis en deux familles.

Les tanins hydrolysables, gallotanins et ellagitanins (Mingshu et al, 2006), sont composés d'un cœur polyol, souvent le glucose, où les fonctions hydroxyles sont estérifiées par l'acide gallique ou un de ses dérivés proches. Ces composés ne représentent qu'un faible pourcentage des tanins totaux. Ils n'appartiennent pas à la famille des flavonoïdes et ne seront donc pas plus longuement explicités dans ce mémoire.

Les tanins condensés sont des composés phénoliques de masse moléculaire comprise entre 500 et 3000 Da. Ils sont stockés dans les pellicules, les pépins et les rafles. Dans la pellicule, ils sont localisés dans les couches externes de l'hypoderme sous forme d'amas (*Park et al, 1995*). La période d'accumulation des tanins condensés a lieu dès la formation du fruit, la teneur maximale se situe lors de la véraison (*Downey et al, 2003*). Leur synthèse peut être stimulée lors d'un stress environnemental causé par une carence en nutriments, la sécheresse ou une variation de température ou d'intensité lumineuse (*Lawler et al, 1997*). Ils sont associés à un mécanisme de défense contre les herbivores (*Forkner et al, 2004*) en inhibant par précipitation certaines enzymes. Les tanins condensés peuvent donc entraîner des carences nutritionnelles chez certains animaux herbivores qui les ingèrent (*Beart et al, 1985*).

Ils sont issus de la polymérisation d'entités moléculaires, les flavan-3-ols qui possèdent la structure générale des flavonoïdes. Ces composés se distinguent par leur nombre d'unités monomériques et le type de liaison les reliant entre-elles, soit C4-C6 ou C4-C8 lors d'une polymérisation ordonnée, c'est-à-dire sans l'intervention de radicaux libres (**Figure 5**). Le nombre d'unités peut varier de 1 à 30. Leur caractérisation est possible grâce à leur capacité à libérer des anthocyanes en milieu acide, à chaud, par rupture de la liaison intermonomérique. Dans le raisin, on distingue deux types de proanthocyanidines suivant la nature de l'anthocyane libéré. D'une part, les procyanidines, polymères de catéchine et d'épicatéchine (**Figure 6**), qui libèrent de la cyanidine ; d'autre part, les prodelfphinidines, polymères de gallocatéchine et d'épigallocatéchine, qui libèrent de la delphinidine et qui ne diffèrent des composés précédents que par une fonction hydroxyle supplémentaire sur le C3 du cycle B (*Jourdes, 2003*).

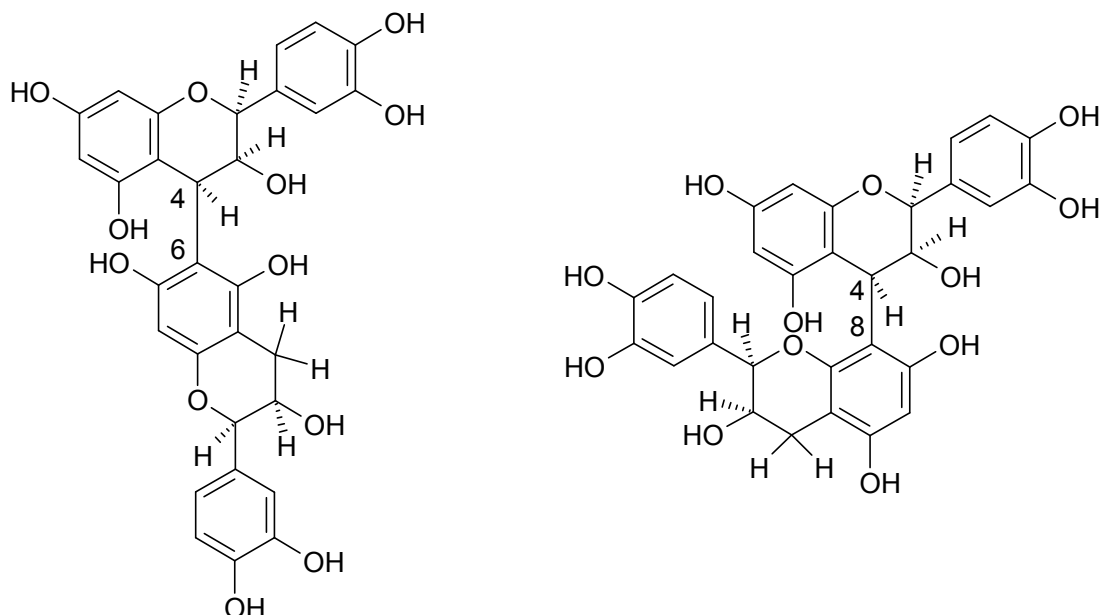
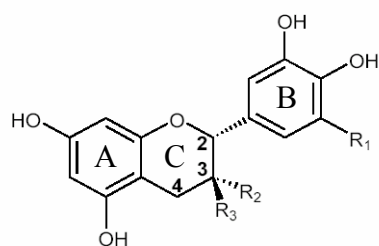


Figure 5 : Exemple de polymérisation ordonnée C4-C6 ou C4-C8 des proanthocyanidines conduisant à la formation de tanins condensés dimériques (*Galvin, 1993*).



R1	R2	R3	Procyanidine
H	H	OH	(+) Catéchine
H	OH	H	(-) Epicatéchine
OH	H	OH	(+) Gallocatéchine
OH	OH	H	(-) Epigallocatéchine

Figure 6 : Structure des flavan-3-ols, principaux monomères des tanins condensés du raisin (Jourdes, 2003)

Les tanins peuvent interagir avec les protéines, telles que les protéines de la salive riches en prolines (PRPs), pour former des complexes insolubles. Les interactions se font par la formation de liaisons non covalentes, le plus souvent, de type hydrophobe. La complexation peut être renforcée par des liaisons hydrogène, par exemple entre le carbonyle de la fonction amine secondaire de la proline et les OH phénoliques (Haslam *et al*, 1994). Les interactions tanins-protéines dépendent des caractéristiques des tanins : taille, structure et charge. Les PRPs contribuent notamment à une défense de l'organisme contre les tanins en formant des complexes insolubles et en empêchant ainsi leurs interactions avec d'autres entités biologiques (Lu *et al*, 1998). Les tanins possèdent aussi la propriété de réagir avec les PRPs, en modifiant leur état et leurs propriétés lubrifiantes. Selon la nature et la concentration des tanins, la conséquence peut être une sensation harmonieuse ou au contraire une agressivité révélée qui correspond à l'amertume ou astringence (Mirabel, 2000).

III.2.3. Les flavonols

Les mousses et les fougères contiennent des taux importants de flavonols qui représentent de par leur ancienneté et leur universalité un groupe fondamental de molécules au sein de la famille des flavonoïdes (Winkel-Shirley *et al*, 2002).

La structure des flavonols se distingue de celle des anthocyanes par la présence d'une fonction cétone supplémentaire sur le carbone 4 du cycle C. Quatre flavonols sont majoritairement présents dans le raisin : le kaempférol, la quercétine, la myricétine et l'isorhamnétine (Figure 7). Les dérivés de la quercétine sont toujours prédominants. Ceux de la myricétine semblent être quant à eux, spécifiques aux variétés de raisins rouges.

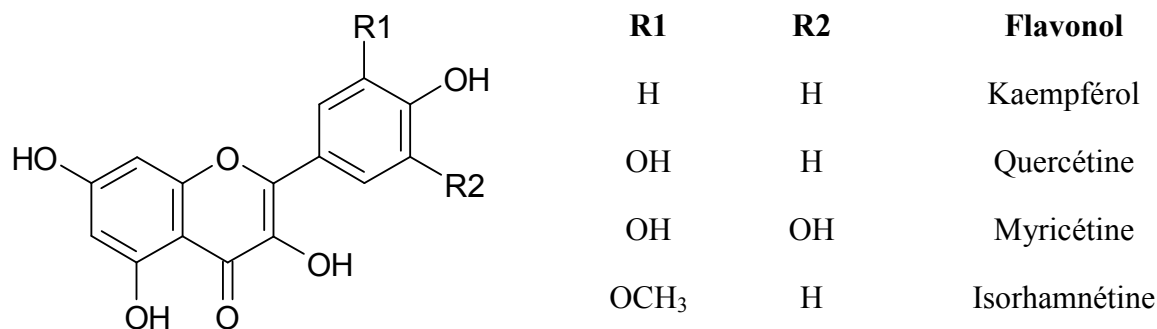


Figure 7 : Flavonols majoritaires chez le raisin

Chez la vigne, les flavonols sont synthétisés durant la phase de développement comprise entre l'apparition et le mûrissement de la baie. Pendant cette période, les flavonols sont stockés dans l'épiderme de cette dernière. Leur accumulation, dans les pépins, reste faible (*Bogs et al, 2006*).

Ces composés agissent en tant que molécules protectrices des ultraviolets permettant ainsi de limiter les effets tissulaires et cellulaires engendrés par l'irradiation. Néanmoins, différentes études, chez pétunia et *Arabidopsis thaliana*, montrent que les ultraviolets induisent la synthèse de flavonols enrichis en groupements hydroxyle. L'hydroxylation de ces composés n'affectant ni leur capacité antioxydante, ni l'absorption des ultraviolets (*Ryan et al, 2001*), la réponse induite par un stress lumineux reste un phénomène complexe à caractériser.

IV. Evolution des composés phénoliques au cours du vieillissement des vins rouges

Lors de son vieillissement, tant la couleur que les propriétés organoleptiques d'un vin évoluent. Ceci est lié notamment à la complexité des réactions chimiques qui s'établissent entre composés phénoliques, anthocyanes et tanins condensés plus particulièrement. Les anthocyanes libres sont en effet des molécules instables, qui, soit se dégradent, soit interagissent avec les tanins condensés. Les travaux de *Somers en 1966* mettent en évidence l'existence de complexations diverses entre anthocyanes et tanins condensés durant le vieillissement des vins. Différents modes réactionnels ont été caractérisés (*Wang et al, 2003*). Ils conduisent, via la création de nouveaux pigments, à des modifications de couleur en fonction du temps :

- les condensations directes entre tanins condensés et anthocyanes entraînent, selon leur type, l'apparition d'une composante rouge ou orange,
- la condensation par l'intermédiaire de l'acétaldéhyde conduit à l'apparition d'une composante mauve.

IV.1. Synthèse des nouveaux pigments

IV.1.1. Condensation de type T-A

Les tanins condensés peuvent, après protonation, former un carbocation ($C4^+$) et réagir de manière réversible avec des sites nucléophiles tels que ceux correspondant aux positions C6 et C8 d'une molécule d'anthocyane pour conduire à l'espèce T-A⁺. Lors de cette réaction, l'anthocyanes est présente soit sous sa forme de base carbinole, soit éventuellement chalcone après réaction de tautomérisation (**Figure 8**). La conservation du vin, en cuve étanche ou en bouteille, est favorable à ce type de condensation (*Galvin, 1993*).

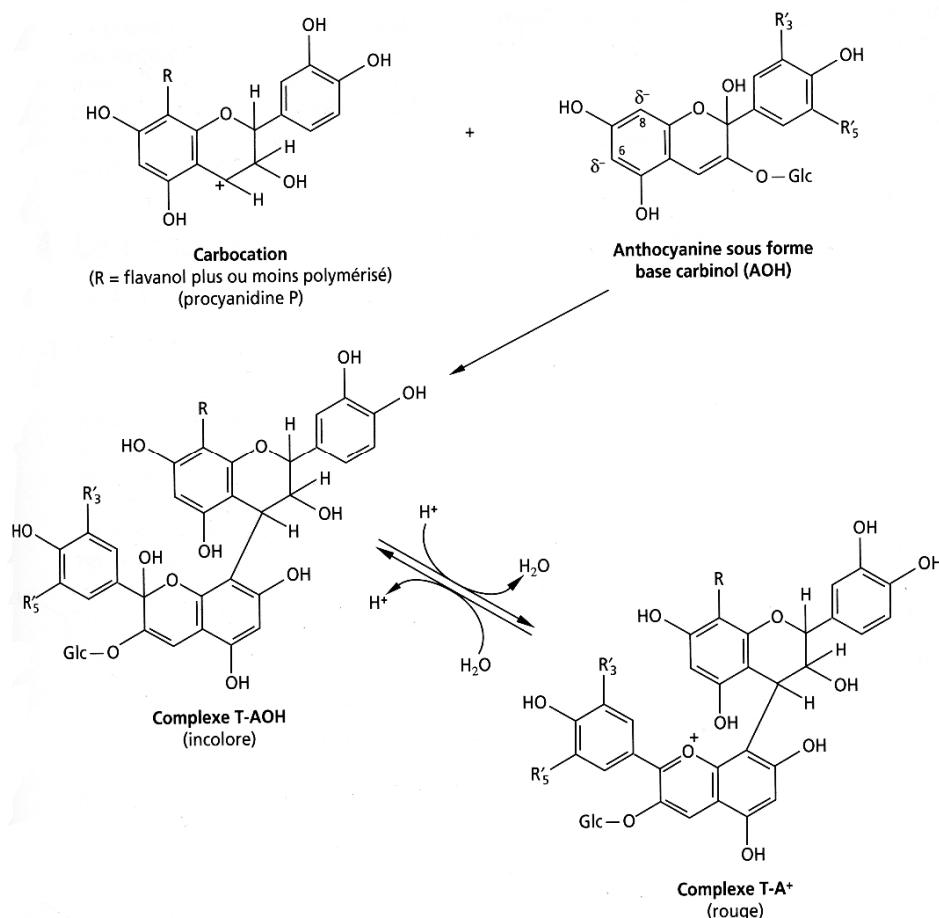


Figure 8 : Condensation directe de type T-A

IV.1.2. Condensation de type A-T

Les anthocyanes sont, lors de cette réaction, sous forme cationique et réagissent sur les sites électro-négatifs des protanthocyanidines (C6 ou C8) en formant une flavène incolore. La présence d'oxygène ou d'un milieu oxydant réintroduit une charge positive sur le cycle C qui restaure la coloration de ce composé (*Galvin, 1993*). A mesure que la réaction progresse, l'équilibre entre la forme flavylum et la base hémiacétal se déplace vers le cation flavylum (**Figure 9**). Dans le vin, ce mécanisme apparaît lors de l'écoulage à la fin de la cuvaison. Le vin « prend » de la couleur après aération.

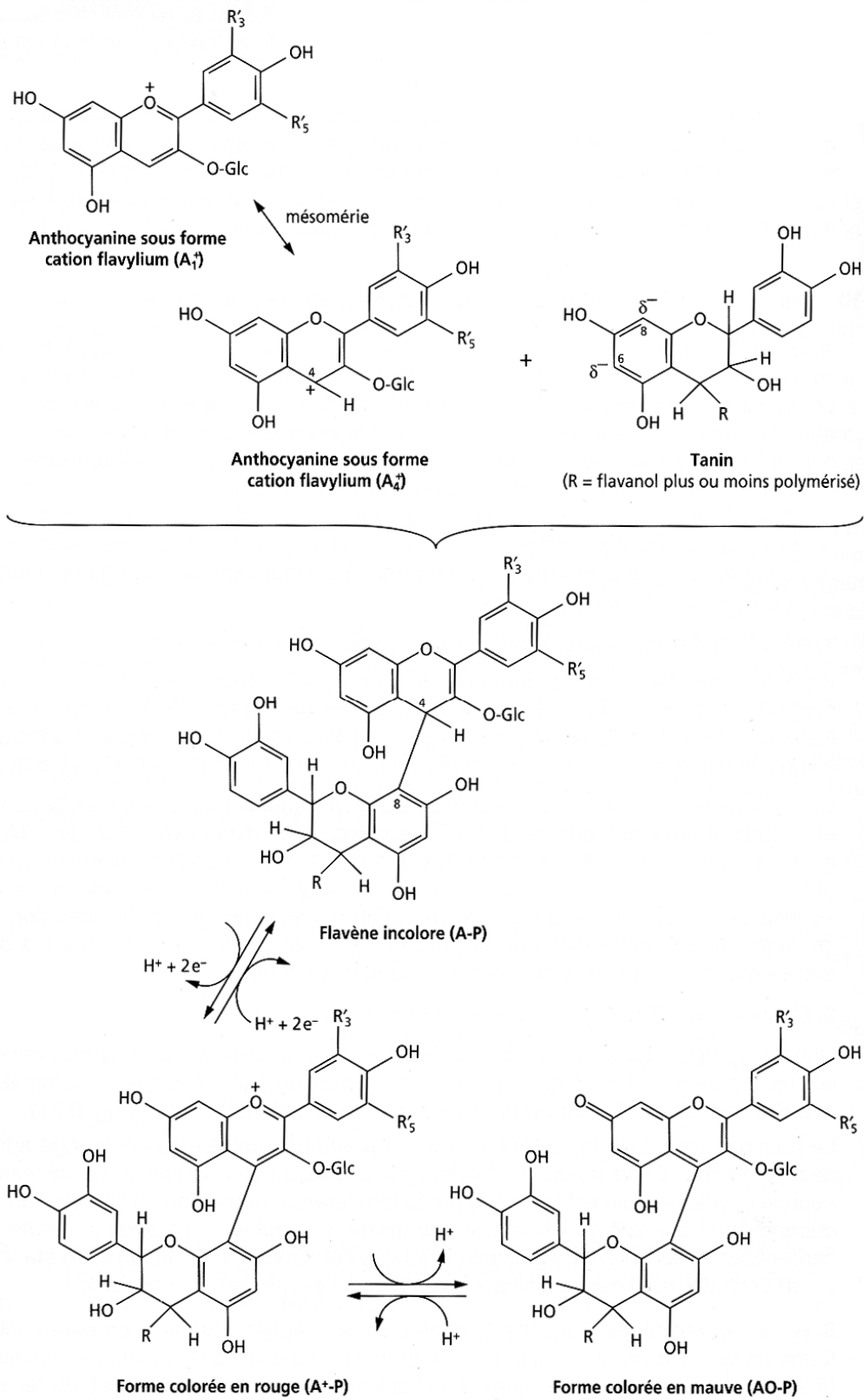


Figure 9 : Condensation directe de type A-T

IV.1.3. Condensation avec un pont éthyle

L'éthanal (ou acétaldéhyde) participe aux réactions de condensation pour aboutir à des structures polymériques où les noyaux A des flavanols et des anthocyanes sont reliés par l'intermédiaire d'un pont éthyle résultant de la formation d'un complexe intermédiaire flavanol-éthanal.

En milieu acide, la fonction carbonyle de l'éthanal est polarisée. Le carbocation ainsi formé réagit avec les sites nucléophiles (C4 et C8) des flavanols et des anthocyanes présents sous forme neutre (base carbinol) (**Figure 10**) (*Timberlake, 1976*). La liaison privilégiée est celle entre deux carbones C8. Néanmoins, la réaction dépend du pH du milieu et des taux de flavanols et d'anthocyanes susceptibles de réagir. Dans la phase de vieillissement du vin, cette réaction a lieu lors de l'oxydation ménagée qui accompagne l'élevage en barrique. La couleur du vin s'intensifie, change de nuance et devient plus sombre en quelques mois de conservation.

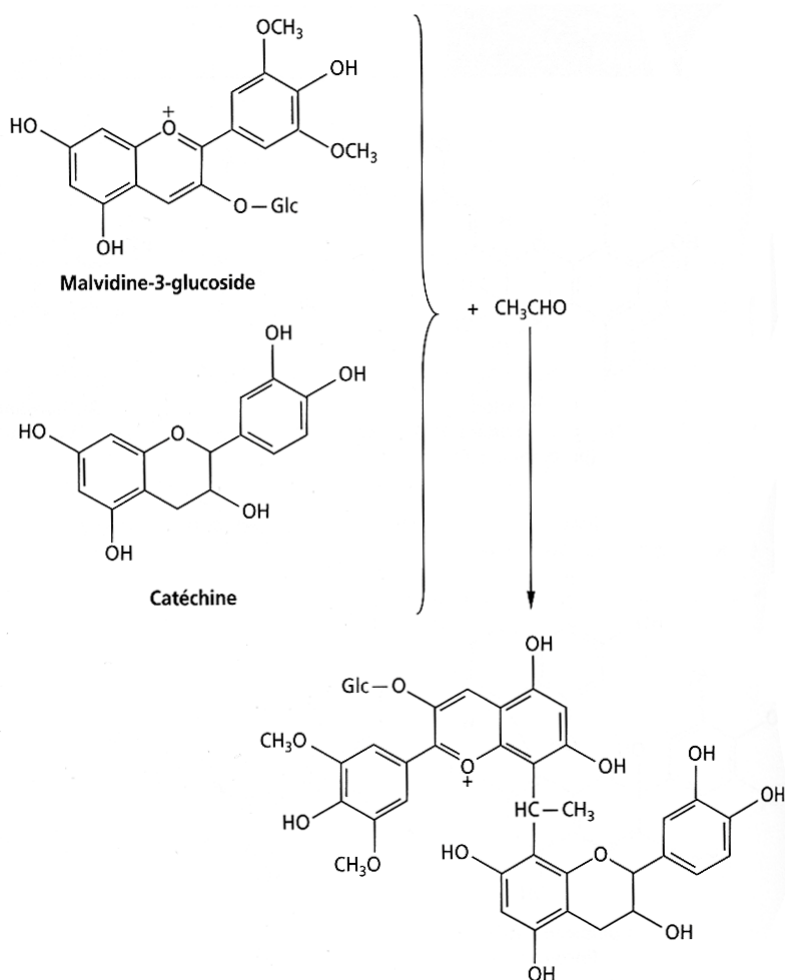


Figure 10 : Réaction entre catéchine et malvidine-3-glucoside en milieu acide et en présence d'éthanal

IV.2. Autres pigments

Lors du vieillissement du vin, les polyphénols incolores (tanins condensés, flavones, flavonols...) interagissent, via des interactions hydrophobes, avec les noyaux flavyliums riches en électrons π des anthocyanes. Il en résulte un nouveau type d'association. Cette réaction est décrite comme l'étape préliminaire à la formation lente, durant le vieillissement, de composés colorés plus complexes (*Escribano-Bailon, 1996*). Diverses expériences prouvent que ces nouveaux pigments sont moins sensibles aux variations de pH et à l'oxydation par le dioxyde de soufre.

IV.3. Incidences de structures complexes sur le vieillissement du vin

Les réactions précédemment décrites induisent une diminution du taux d'anthocyanes libres lors du vieillissement au profit de nouveaux pigments. La diminution du taux d'anthocyanes libres devrait conduire à une perte de coloration liée à la formation de composés hémiacétals. Cette perte n'est pas observée compte tenu de la formation de structures complexes colorées qui modifient et renforcent la couleur initiale. Le vin prend une coloration pourpre.

Les réactions précédentes sont aussi responsables de l'évolution des propriétés organoleptiques telles que l'astringence et l'amertume (*Perez-Magarino et al, 2004*) : l'œnologue parle de ces changements de propriétés en termes d'« arrondissement » et d'« assouplissement » du vin.

L'ensemble des modifications observées lors du vieillissement du vin, tant au niveau de la couleur que des propriétés gustatives, résulte d'une chimie complexe dont les mécanismes réactionnels commencent seulement à être décrits avec précision. Cette chimie est à notre connaissance spécifique au vin. Elle résulte notamment de la structure particulière des anthocyanes qui autorise leur complexation avec les tanins condensés.

Les anthocyanes du vin, de par leur simplicité, sont très sensibles à l'oxydation. Lorsqu'ils sont extraits de leur milieu naturel et dissous dans une solution aqueuse d'acidité équivalente à celle du vin, ils perdent leur coloration (*Escribano-Bailon et al, 1996*). En effet, l'ion flavylium subit l'attaque nucléophile de l'eau et est converti, par une réaction d'hydratation, en composé hémiacétal (**Figure 11**).

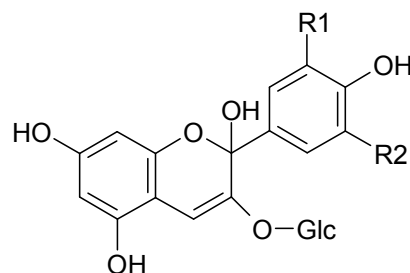


Figure 11 : Hémiacétal résultant de l'hydratation d'une molécule d'anthocyane

Toutefois, en présence d'acétaldéhyde, composé issu du métabolisme secondaire des levures lors de la fermentation (*Medina et al, 2005*), cette sensibilité à l'oxydation est fortement diminuée. Ceci est dû à l'accessibilité de la fonction hydroxyle en C5 du cycle A qui confère aux anthocyanes du vin une réactivité chimique spécifique (*Brouillard et al, 2003*). Le groupement phloroglucinol (1,3,5-trihydroxybenzène) permet la formation d'un hétérocycle supplémentaire (D) entre la position C4 et la fonction hydroxyle en C5 (*Fulcrand et al, 1998*). La molécule ainsi formée, *vitisine*, (**Figure 12**) est, de part son aromaticité, peu sensible à l'oxydation et reste donc principalement sous forme « flavylium » (*Bakker et al, 1997*).

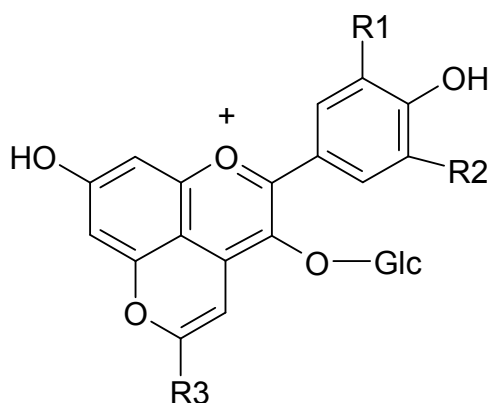


Figure 12 : Formule de la Vitisine

De part leurs particularités physico-chimiques, des anthocyanes plus complexes, tels que ceux présents chez *Matthiola incana* par exemple, sont moins sensibles à l'oxydation.

Deux raisons expliquent cette stabilité :

- ✓ l'empilement des cycles aromatiques (esters de cinnamoyl) sur chaque face du cycle C limite l'attaque nucléophile de l'eau sur l'ion flavylium,
- ✓ la formation de liaisons hydrogènes entre le radical malonyl du O-malonyl-glucose en C5 et la fonction hydroxyle en C7 (**Figure 13**) entraîne une constante d'acidité supérieure à la constante d'hydratation, protégeant ainsi le pigment (*Brouillard et al, 2003*).

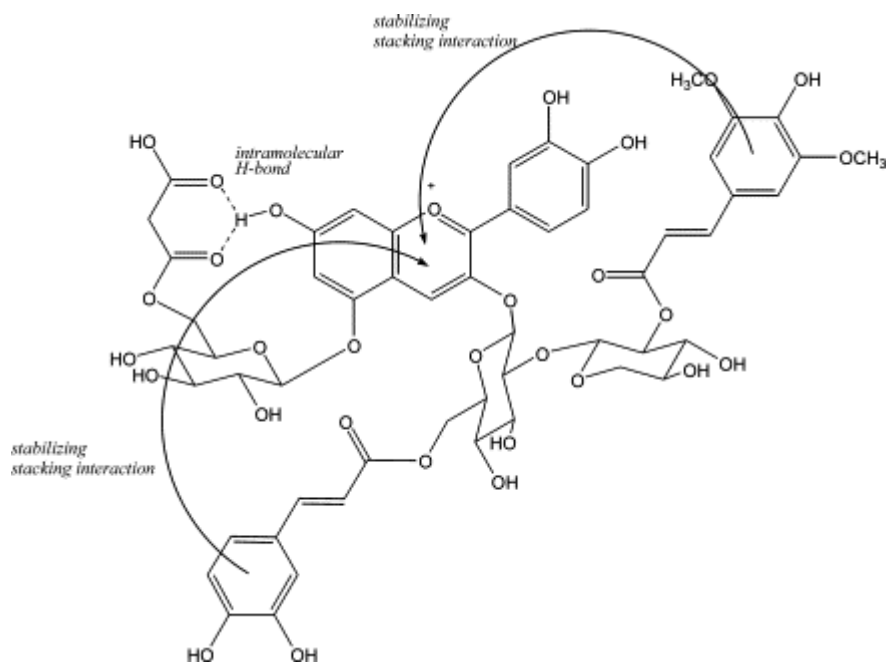


Figure 13 : Structure d'un anthocyane de *Matthiola incana*

IV.4. Vieillessement des vins et propriétés organoleptiques

La relation existant entre la composition chimique d'un vin et ses propriétés gustatives demeure un sujet critique en œnologie. De nombreux travaux sont entrepris afin de définir le rôle particulier de chaque constituant dans l'élaboration des propriétés sensorielles.

L'amertume et la sensation tactile d'astringence sont induites par la teneur en monomères de flavan-3-ol (catéchine, épicatechine, épigallocatechine, épicatechine gallate et épigallocatechine gallate) et par celle de leurs oligomères respectifs que constituent les tanins condensés (*Brossaud et al, 2001 ; Gawel, 1998*). Nous l'avons vu, la formation de ces derniers est un mécanisme complexe, soumis à la fois, à la synthèse des différents constituants monomériques, mais aussi à l'existence de multiples réactions chimiques intervenant dans le vieillissement des vins (degré de polymérisation, degré d'oxydation, substitutions...). La conséquence principale de cette complexité est le recensement à ce jour de près de 4000 flavonoïdes chez les plantes (*Cheyrier, 2005*). Au-delà de ce nombre, les valeurs de pH, les pourcentages d'éthanol et de sucre et enfin la viscosité des vins sont autant de paramètres susceptibles de moduler les sensations d'astringence et d'amertume (*Lesschaeve et al, 2005*). Les monomères sont plus amers qu'astringents et inversement, les oligomères sont plus astringents qu'amers. Les **Figures 14 et 15** résument l'influence du degré d'oligomérisation, du type de monomères et de la nature des liaisons entre monomères sur la perception de l'astringence et de l'amertume.

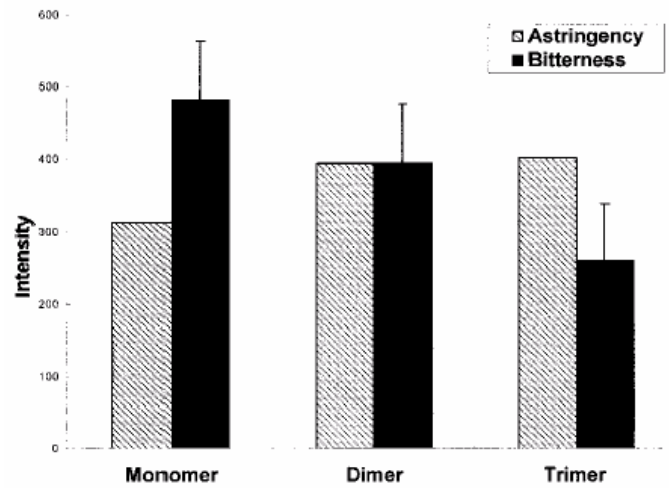


Figure 14 : Influence du degré de polymérisation (Peleg et al, 1999)

Stimulus	Chemical structure
(+)-catechin	(+)-cyanidol-3
(-)-epicatechin	cis-2-[3,4-dihydroxyphenyl]-3,4-dihydro-2H-1-benzopyran-3,5,7-triol
dimer B3	catechin (4→8) catechin
dimer B6	catechin (4→6) catechin
dimer B4	catechin (4→8) epicatechin
trimer C2	catechin (4→8) catechin (4→8) catechin
trimer C	catechin (4→8) catechin (4→8) epicatechin

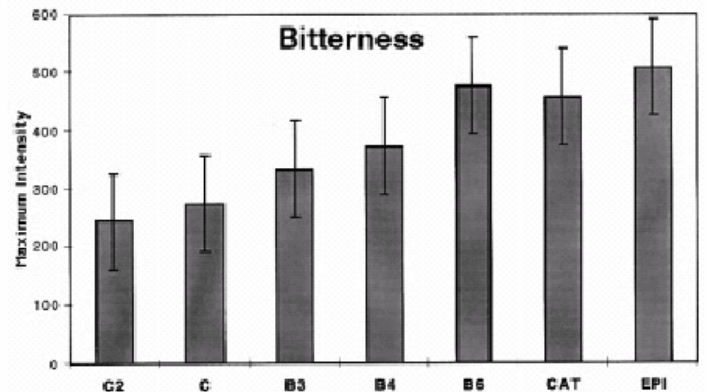
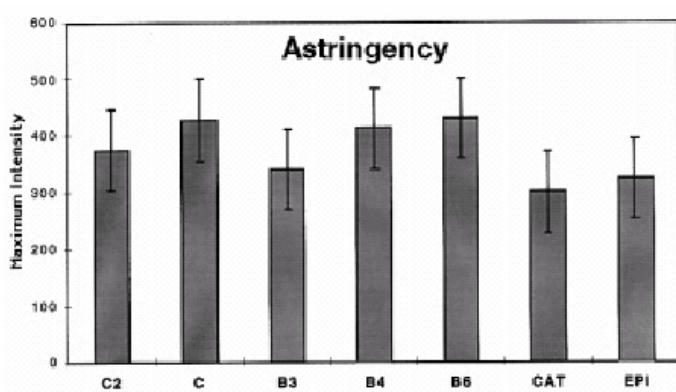


Figure 15 : Influence des différents oligomères sur la perception de l'astringence et de l'amertume (Peleg et al, 1999)

Enfin, l'analyse par anova (analyse statistique à multi variables) des constituants du vin ne conclut pas à un quelconque rapport entre teneur et nature d'un composé spécifique et une propriété sensorielle donnée (*Boselli et al, 2004*). Les connaissances actuelles ne permettent en aucun cas d'associer à des caractéristiques gustatives précises une composition donnée en flavonoïdes, et cela pour deux raisons.

La première réside dans la difficulté de mesurer objectivement les sensations élémentaires que sont l'astringence et l'amertume. Ces dernières varient beaucoup selon les dégustateurs et leur séparation reste très diffuse (*Mirabel, 2000*).

La seconde réside dans la difficulté d'analyse des composés du vin. Les techniques de chimie analytique permettant difficilement la caractérisation d'un polymère complexe, la majorité des travaux portent sur des composés monomériques ou des oligomères de faible poids moléculaire qui peuvent être identifiés individuellement (*Cheyrier et al, 2005*). Ce problème est renforcé par le fait que seul 50 % des polyphénols présents dans un vin de deux ans d'âge sont connus (anthocyanes, flavonols, flavanols incluant catéchines et épicatechines,...). La seconde moitié des molécules présentes, issues de l'élevage et du vieillissement, restent à ce jour inconnues (*Caderni et al, 2000*).

Dans ce paragraphe, nous avons souligné la particularité de la vigne. Cette dernière découle de la variété, de la teneur, et de la faculté d'évolution de ses différents constituants en fonction du temps. La synthèse des diverses classes de molécules y est possible grâce à l'existence d'enzymes. Dans le paragraphe suivant, nous présentons le rôle spécifique de chaque enzyme et son expression spatio-temporelle. Une courte description de leur structure tridimensionnelle y est associée dès lors que cette dernière est connue. Une de ces enzymes, la dihydroflavonol-4-reductase est étudiée de manière plus approfondie, et les motivations qui guident son choix comme protéine d'étude dans ce travail sont explicitées.

V. Voies de biosynthèse des flavonoïdes

Les précurseurs de la voie de biosynthèse des flavonoïdes sont le malonyl-CoA et le coumaryl-CoA, issus respectivement du métabolisme des sucres et de la voie des phénylpropanoïdes. La première étape enzymatique est catalysée par la chalcone synthase (CHS), qui aboutit à un produit jaune, la chalcone. Pour la majorité des plantes, la chalcone ne s'accumule pas. Elle est métabolisée, via de multiples étapes enzymatiques, en différentes classes de flavonoïdes : les flavanones, les dihydroflavonols et finalement les anthocyanes. D'autres classes de flavonoïdes telles que les isoflavones, les flavones, les auronnes, les pro-anthocyanidines et les flavonols sont issues de voies de biosynthèse latérales dont les précurseurs sont les intermédiaires de formation des anthocyanes (**Figure 16**) (*Schijlen et al, 2004*).

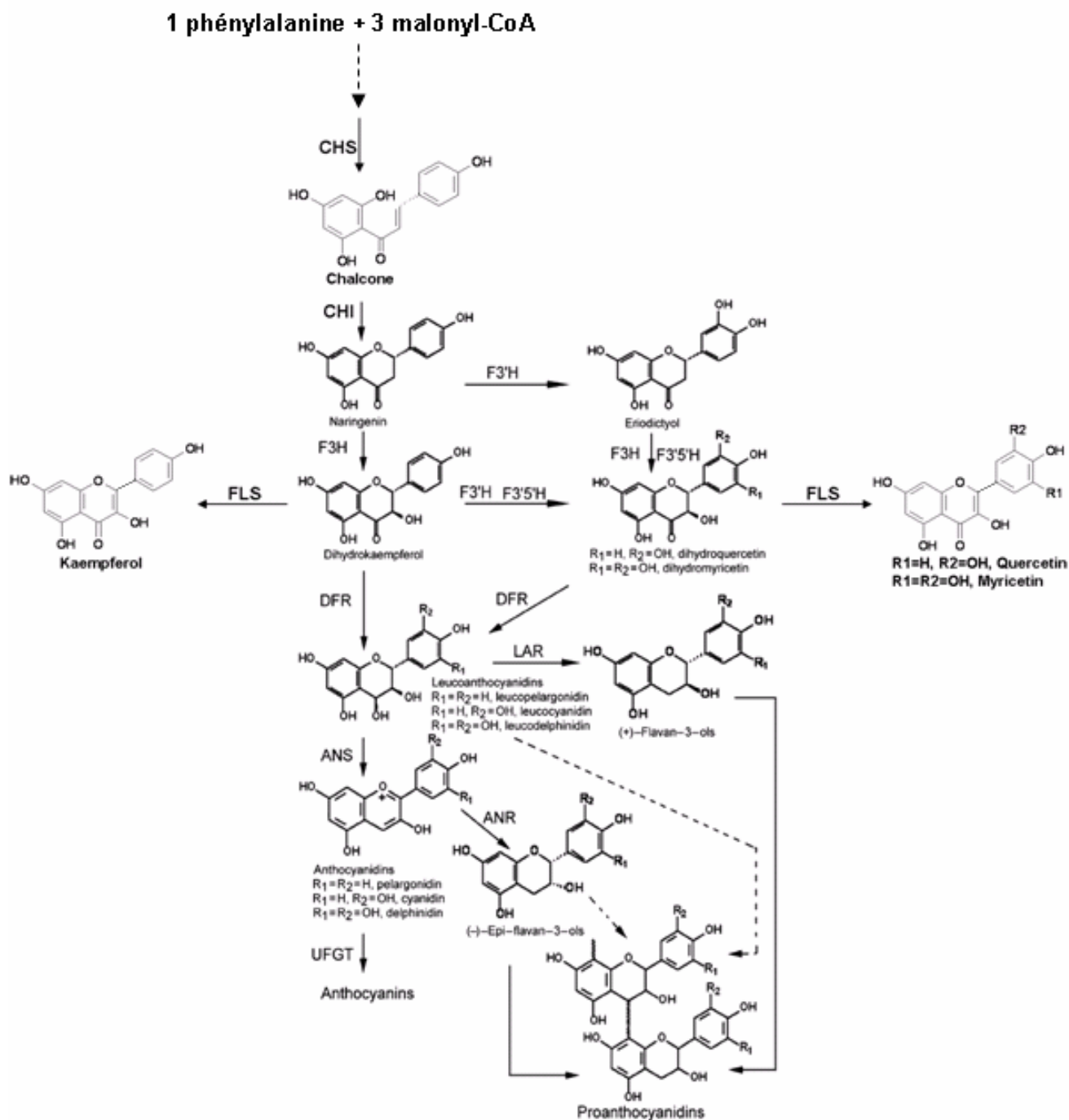


Figure 17 : Schéma général de la voie de biosynthèse des flavonoïdes

CHS, chalcone synthase; CHI, chalcone isomérase; F3H, flavanone-3-hydroxylase; F3'H, flavonoïde 3'-hydroxylase; F3'5'H, flavonoïde 3',5'-hydroxylase; FLS, flavonol synthase; DFR, dihydroflavonol réductase; LAR, leucoanthocyanidine réductase; ANS, anthocyanidine synthase; ANR, anthocyanidine réductase; UFGT ou 3-GT, uridine diphosphate glucose-flavonoïde 3-O glucosyl transférase.

Les gènes codant pour les enzymes impliquées dans la voie de biosynthèse des anthocyanes possèdent, chez la vigne, un profil similaire de transcription. Les feuilles de taille moyenne (5 cm) et les pépins contiennent les plus hauts taux de transcrits (**Figure 18**). Une expression significative de ces mêmes gènes (phénylalanine ammonia lyase (PAL), chalcone synthase (CHS), chalcone isomérase (CHI), flavanone-3-hydroxylase (F3H), dihydroflavonol réductase (DFR) et anthocyanidine synthase (ANS)) est toutefois mesurée dans de nombreuses parties de la plante : jeunes feuilles, vrilles, tiges vertes, racines et fleurs (**Boss et al, 1996**). Les taux les plus bas sont, quant à eux, détectés dans les feuilles matures et la pellicule de la baie.

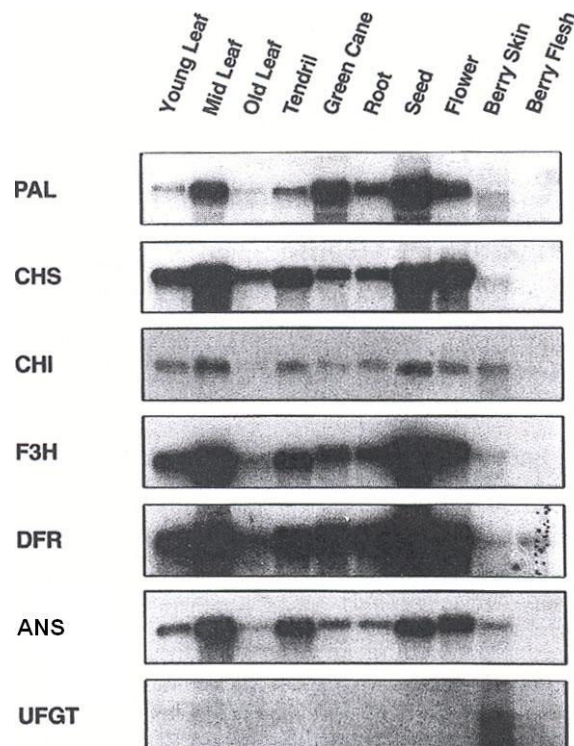


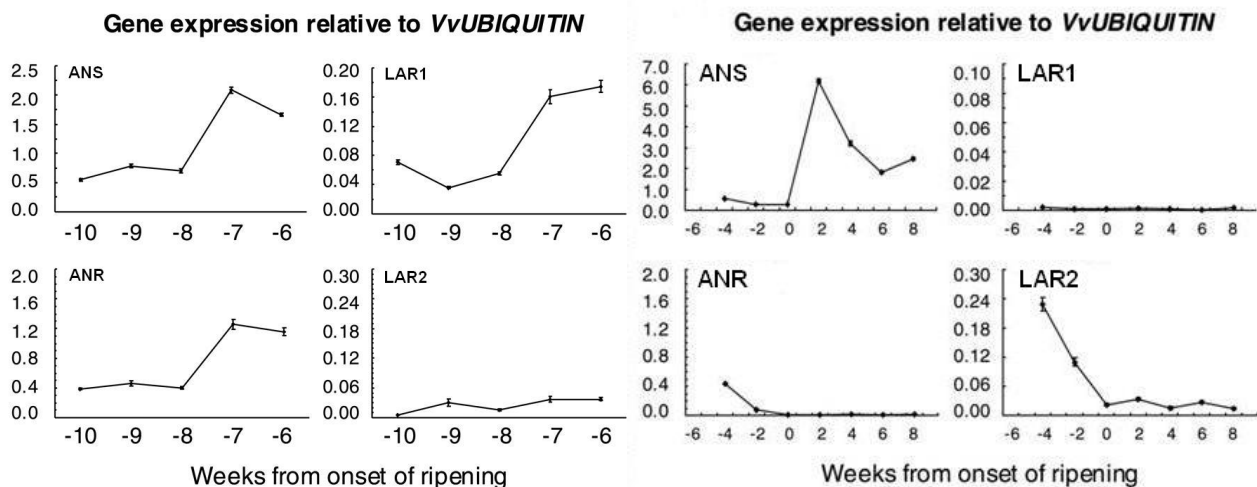
Figure 18 : Analyse par northern-blot des ARNs totaux extraits de différents tissus de baie rouge de cultivar Shiraz (Boss et al, 1996)

PAL, phénylalanine ammonia lyase ; CHS, chalcone synthase; CHI, chalcone isomérase; F3H, flavanone-3-hydroxylase; DFR, dihydroflavonol réductase; ANS, anthocyanidine synthase; UFGT, uridine diphosphate glucose-flavonoïde 3-O glucosyl transférase.

Si l'analyse par northern-blot est effectuée plus en détail, il apparaît que seule l'expression de la DFR a pu être détectée dans la partie charnue de la baie. L'uridine diphosphate glucose-flavonoïde 3-O glucosyl transférase (UFGT ou 3-GT) est, quant à elle, localisée uniquement dans la pellicule, zone d'accumulation des anthocyanes. La production de ces derniers est donc, chez la vigne, régulée par l'UFGT. Ceci est une caractéristique de la vigne. L'UFGT ne contrôle pas la production d'anthocyanes chez les autres plantes. Dans le cas du pétunia par exemple, ce sont la F3H et la DFR qui contrôlent cette biosynthèse.

Durant la première partie du développement de la plante, les principaux flavonoïdes sont les proanthocyanidines et les flavonols, composés dont les taux vont diminuer fortement durant le mûrissement de la baie au profit des anthocyanes (*Jaakola et al, 2002*). Ceci s'accorde avec une activité de la flavonol synthase (FLS), forte durant les premières phases de développement de la plante puis déclinant rapidement à partir de la véraison. Une régulation temporelle est donc mise en jeu dans la plante pour limiter une utilisation compétitive des dihydroflavonols dans la synthèse des anthocyanes et des flavonols (*Forkmann et al, 1999*).

Lors de la synthèse des proanthocyanidines, les gènes codant la leucoanthocyanidine réductase (LAR) et l'anthocyanidine réductase (ANR) ont un profil d'expression qui concorde avec l'accumulation de ces composés. Ces gènes sont exprimés à la floraison, soit 10 semaines avant le mûrissement. Leur transcription n'augmente significativement que deux semaines après cette date (**Figure 19**). Dans la pellicule, l'accumulation des proanthocyanidines est achevée à la véraison. L'expression forte de l'ANS à cette période implique une synthèse différenciée dans le temps des proanthocyanidines et des anthocyanes (**Figure 20**). Ces derniers sont biosynthétisés dès que la synthèse des proanthocyanidines est achevée (*Bogs et al, 2005*).



Figures 19 et 20 : Analyse de l'expression des gènes codant pour ANS, ANR, LAR1, et LAR2 durant les premiers stades de développement (19) et le développement global (20) de la baie de raisin par PCR en temps réel (*Bogs et al, 2005*).

La LAR étant codée par deux gènes LAR1 et LAR2, l'analyse des profils d'expression a été effectuée sur les deux gènes .

V.1. Les enzymes ubiquitaires

Tous les gènes impliqués dans ce processus, à l'exception de ceux codant pour l'ANR (*Xie et al, 2003*) et de la LAR (*Tanner et al, 2003*), qui ont été découverts plus récemment, ont été identifiés et clonés chez *Vitis vinifera* en 1994 (*Sparvoli en 1994*).

V.1.1. Chalcone Synthase (CHS)

Cette enzyme (EC : 2.3.1.74) produit la 4,2',4',6'-tétrahydrochalcone. Cette synthèse implique la condensation séquentielle d'une molécule de coumaryl et de trois molécules malonyl-CoA (**Figure 21**). Après fixation du coumaryl, chaque étape de condensation commence par la décarboxylation d'un malonyl-CoA. Le carbanion acétyl-CoA résultant agit comme nucléophile pour l'élongation de la chaîne. Au terme de cette triple réaction, un intermédiaire tétracétide est formé. Il se circularise en noyau aromatique hydroxylé par une condensation de Claisen. Durant la cyclisation, le groupe méthylène proche de la moitié coumaryl attaque le carbone du thioester. La fermeture du cycle est possible par le transfert interne d'un proton depuis le carbone nucléophile jusqu'à l'oxygène du carbonyle. Le cycle ainsi formé est le cycle A du polyphénol. La totalité de la réaction a été décrite en 1999 (*Ferrer et al, 1999*).

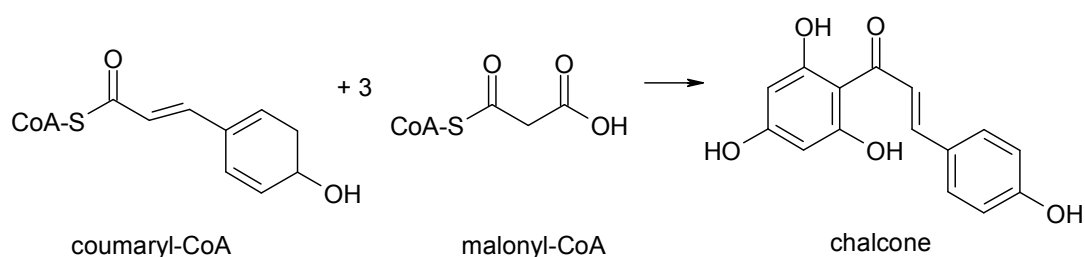


Figure 21 : Schéma réactionnel simplifié de la chalcone synthase

La structure tridimensionnelle de la CHS a été déterminée par diffraction des rayons X à la résolution de 1,7 Å (*Ferrer et al, 1999*). Le monomère est composé de deux domaines. Le domaine principal est situé en N-terminal. Il se caractérise par un motif pseudo-symétrique $\alpha\beta\alpha\beta\alpha$ rencontré, par exemple, chez la 3-cétoacyl-CoA thiolase de *Saccharomyces cerevisiae*. Le domaine catalytique (C-terminal) fait apparaître trois cavités interconnectées constituant respectivement le tunnel de fixation du CoA, la poche de fixation des molécules de malonyl-CoA et une poche de cyclisation (**Figure 22**).

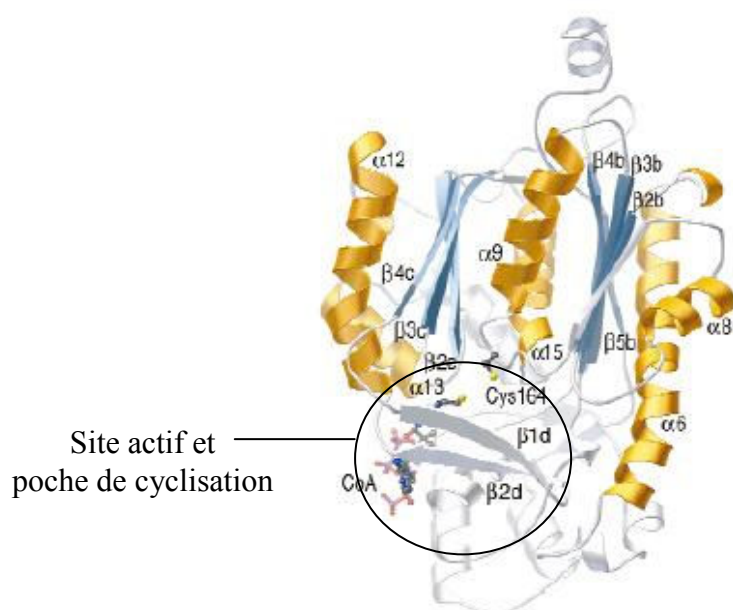


Figure 22 : Représentation en rubans d'un monomère de chalcone synthase

La cystéine 164 impliquée dans la fixation et le transfert de la structure polycétonique, le site de fixation du CoA et la poche de cyclisation délimitée par les brins $\beta 1d$ et $\beta 2d$ sont localisés par un cercle (**Ferrer et al, 1999**).

V.1.2. Chalcone isomérase (CHI)

La cyclisation intramoléculaire de la chalcone en 2S-naringénine est une réaction spontanée (**Figure 23**). Toutefois, en présence de la CHI (EC : 5.5.1.6), cette réaction est stéréospécifique, (isomère S) et d'un rendement très supérieur (facteur 10^7) (**Jez et al, 2000**). La structure centrale de la CHI comporte un large feuillet β ($\beta 3a$ - $\beta 3f$) surmonté d'une série d'hélices α ($\alpha 1$ - $\alpha 7$) (**Figure 24**). L'analyse de la structure 3D révèle l'importance de liaisons hydrogène entre une molécule d'eau et la fonction cétone de la naringénine. Ainsi, après l'attaque nucléophile de l'oxygène 2' sur la double liaison α/β insaturée, une molécule d'eau stabilisée par la tyrosine 106 agit comme un acide et stabilise l'énolate. Un intermédiaire flav-3-en-4-ol est créé qui, après réaction de tautomérisation, devient de la naringénine.

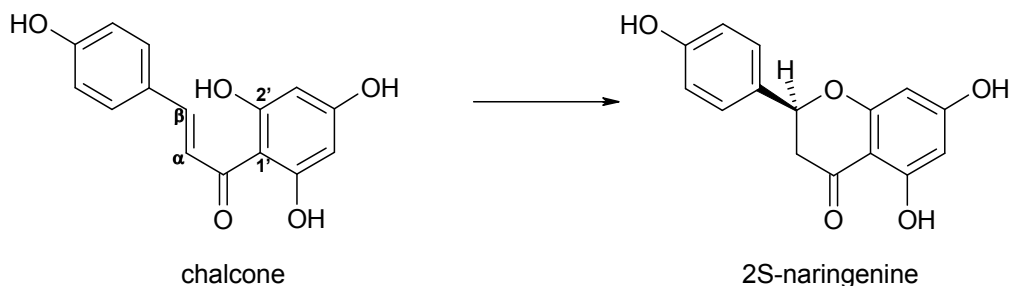


Figure 23 : Formation de la 2S-naringénine catalysée par la chalcone isomérase (Jez et al, 2000)

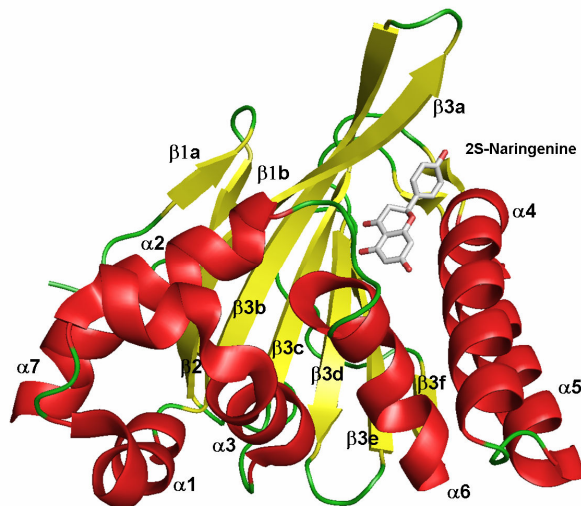


Figure 24 : Représentation en rubans de la chalcone isomérase (Jez *et al*, 2000)

V.1.3. Flavonoïdes hydroxylases (*F3H*, *F3'H*, *F3'5'H*)

Les flavanones sont converties par la flavanone 3 β -hydroxylase (*F3H*) (EC : 1.14.11.9) en dihydroflavonols par hydroxylation de leur position 3 (cf **Figure 17**). Cette enzyme est classée parmi les dioxygénases, 2-oxoglutarate dépendantes. Différents co-facteurs sont nécessaires à son activité : le 2-oxoglutarate, le dioxygène, l'ion Fe²⁺, et l'ascorbate (*Springob et al*, 2003). La *F3H* catalyse par exemple l'hydroxylation de la 2S-naringénine en 2R-3R dihydrokaempférol (DHK). Ce dernier peut être alors hydroxylé soit en position 3' (par la *F3'H*) (EC : 1.14.13.21) soit en positions 3' et 5' (par la *F3'5'H*) (EC : 1.14.13.88) pour former respectivement la 2R-3R dihydroquercétine (DHQ) et la 2R-3R dihydromyricétine. Les dihydroflavonols ainsi formés sont à la base de la différenciation des anthocyanes existant chez les plantes (pélagonidine, cyanidine, delphinidine). La *F3'H* et la *F3'5'H* sont quant à elles des cytochromes P450 (*Ueyama et al*, 2002) utilisant le NADPH et O₂ comme substrats. Aucune structure 3D de ces enzymes n'est à ce jour connue.

V.1.4. Dihydroflavonol réductase (*DFR*)

La *DFR* (EC : 1.1.1.219) catalyse la réduction stéréospécifique des 2R,3R-dihydroflavonols en 2R,3S,4S-leucoanthocyanidines ou flavan-3,4-diols par oxydation du NADPH (**Figure 17**). Les leucoanthocyanidines sont les précurseurs des anthocyanes mais aussi des proanthocyanidines par la voie métabolique des catéchines (LAR). La *DFR* occupe donc une place centrale, elle conduit à la double voie de synthèse des anthocyanes et des tanins condensés. Plusieurs séquences codantes de *DFRs* issues d'organismes différents ont été clonées. Leur profil d'expression chez la plante a été largement étudié. Il apparaît similaire à celui des autres enzymes communes à la voie de biosynthèse des flavonoïdes (*Shimada et al*, 2004 ; *Gollop et al*, 2002) (**Figure 18**).

Au sein d'une espèce donnée, la séquence nucléotidique de la DFR peut être rencontrée sous la forme d'un gène unique ou de plusieurs gènes. Ce dernier cas est rencontré par exemple chez *Lotus japonicus* (6 gènes, **Shimada et al, 2005**) et chez *Medicago truncatula* (2 gènes, **Xie et al, 2004 (b)**).

De nombreux travaux ont montré que cette enzyme ne réduit pas de façon équivalente les différents dihydroflavonols. C'est ainsi par exemple, que les DFRs de *Pétunia* et de *Cymbidium* ne réduisent pas efficacement le DHK. Ces espèces ne peuvent donc pas produire de fleurs colorées à base de pigment orange (pélagonidine) même si les enzymes F3'H et F3'5'H sont absentes (**Johnson et al, 1999**). Cette inactivité demeure à ce jour inconnue. Les trois principaux substrats, structuralement identiques à l'exception du nombre et de la position des fonctions hydroxyles sur le cycle B, jouent très certainement un rôle dans la spécificité de substrat observée. Des expériences de transfection chez *Pétunia* ont suggéré qu'une région de 27 acides aminés (segment 130-156) est impliquée dans la fixation du substrat (**Johnson et al, 2001**).

L'analyse des relations entre taux d'activité pour un substrat donné et séquence en acides aminés du segment 130-156 tend à montrer l'importance de la nature du résidu 133. La présence d'une asparagine à cette position est associée à une réduction préférentielle du DHK. Sa mutation en aspartate est quant à elle corrélée à une forte réduction de DHQ (**Shimada et al, 2005 ; Xie et al, 2004 (b)**). Diverses difficultés expérimentales rendent difficile la compréhension du phénomène. Parmi celles-ci, citons l'instabilité des leucoanthocyanidines en solution (**Nakajima et al, 2001**), l'expression de DFR sous forme généralement inactive (**Martens et al, 2002**), et l'obtention d'une forme instable de la DFR lors de toute tentative de purification (**Forkmann et al, 1987**).

Aucune structure de DFR n'est à ce jour publiée. Toutefois, l'analyse de la séquence primaire permet, grâce aux outils de la bio-informatique de classer la DFR parmi la superfamille des épimérasés-réductases à chaînes courtes (SDR). Un ancêtre possible de la DFR pourrait être l'UDP galactose-4-épimérase (**Baker et al, 1992**), dont la structure a été résolue par cristallographie (**Thoden et al, 1996**). Chaque monomère est constitué de deux domaines. Le premier, en N-terminal, est responsable de la fixation du co-facteur, alors que le second (C-terminal) est impliqué dans la fixation du substrat. La conformation du domaine N-terminal est de type Rossmann fold (7 brins β en configuration parallèle avec, de part et d'autre du feuillet ainsi formé, trois hélices α , **Figure 25**). Le domaine C-terminal est constitué de cinq brins β et quatre hélices α . Sa conformation est spécifique, aucune SDR ne possédant le même repliement en partie C-terminale.

Cette dernière remarque rend délicate toute prédiction de structure 3D de la partie C-terminale de la DFR. Seule la partie N-terminale, de par la conservation du Rossmann fold, peut être aisément prédite en dépit d'une faible homologie de séquence.

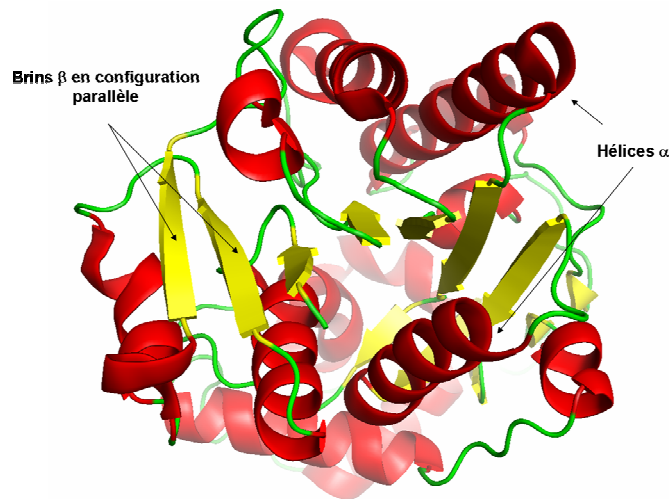


Figure 25 : Représentation en rubans d'un modèle de repliement de Rossmann composé d'un feuillet β (jaune) et de trois hélices α (rouge) de part et d'autres de celui-ci. Ce modèle est issu de la structure de l'alcool deshydrogénase de *Entamoeba histolytica* (Shimon et al, 2006)

V.2. Voie latérale des flavonols

V.2.1. Flavonol synthase, (FLS)

Cette enzyme est directement responsable de la biosynthèse des flavonols (kaempferol, quercétine, myricétine) (**Figure 26**). Elle conduit à la formation d'une double liaison entre les atomes C2 et C3 du cycle C. La FLS (EC : 1.14.11.23) appartient, comme la F3H, à la famille des dioxygénases à 2-oxoglutarate. Seule la FLS de *Citrus unshiu* a été la cible d'une étude fonctionnelle. Les résidus His221, His227 et Asp223 interviennent dans les liaisons de coordination avec le fer. Arg287 est responsable de la fixation du 2-oxoglutarate. Les résidus Gly68, His75, Gly261 et Pro207 n'ont pas de rôle clairement défini, cependant, leurs mutations affectent fortement l'activité de la protéine (Wellmann et al, 2002). Aucune structure tridimensionnelle de FLS n'a été à ce jour déterminée.

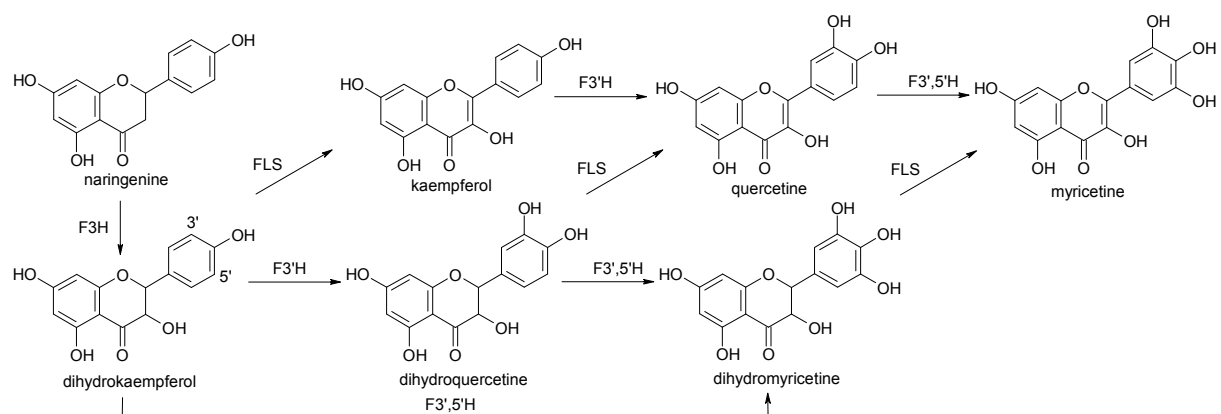


Figure 26 : Voies métaboliques des dihydroflavonols et de leurs flavonols correspondants (Ueyama et al, 2002).

V.3. Voie latérale des anthocyanes

V.3.1. Anthocyanidine synthase (ANS)

L'anthocyanidine synthase (EC : 1.14.11.19) (appelée également leucoanthocyanidine dioxygénase, LDOX) oxyde un substrat incolore, les leucoanthocyanidines, en anthocyanidines colorées (**Figure 27**), composés précurseurs des anthocyanes et des proanthocyanidines par la voie des épicatechines (**Figure 17**). Parallèlement, des études *in vitro* ont montré que cette enzyme est susceptible, dans certaines conditions, de convertir les leucoanthocyanidines en dihydroflavonols, voire en flavonols. Pour cela, une déshydratation ou une énoisation d'un produit intermédiaire, le flav-3-ène-3,4-diol, est nécessaire (**Turnbull et al, 2000 ; Springob et al, 2003**).

L'ANS a été isolée et clonée chez de multiples organismes. Cette enzyme appartient à la famille des dioxygénases à 2-oxoglutarate. La structure 3D du complexe enzyme-substrat-cofacteur a été déterminée chez *Arabidopsis thaliana* (**Wilmouth et al, 2002**). La leucoanthocyanidine étant de nature instable, le substrat utilisé lors de la co-cristallisation est la DHQ. Une sous-unité protéique est composée de 13 brins β parmi lesquels huit ($\beta 5$ - $\beta 12$) forment un tonneau délimitant une cavité hydrophobe à l'extrémité de laquelle se trouve le site actif (**Figure 28**). Une hélice ($\alpha 12$) et deux brins ($\beta 3$ et $\beta 5$) complètent l'extrémité N-terminale du tonneau. La région C-terminale est composée d'une longue boucle (309-333) et d'une hélice α ($\alpha 17$) qui ferment le site catalytique. L'atome de fer est en coordinance octaédrique. Il est lié aux chaînes latérales des acides aminés His232, His288, et Asp234 et possède une interaction bidentate avec le 2-oxoglutarate. La sixième liaison de coordination est avec une molécule d'eau (W400) (**Figure 29**). L'analyse de la structure met en évidence la présence de deux molécules de DHQ. L'une est située au site actif. Elle interagit, via l'oxygène de sa fonction cétone en C4, avec l'hydrogène de la fonction carboxylique en C1 du 2-oxoglutarate. Son cycle A participe à l'empilement de cycles aromatiques avec la chaîne latérale de la Phe304. Enfin, cette molécule est stabilisée par des liaisons hydrogène via les fonctions hydroxyles du cycle B et la chaîne latérale de la Tyr142. La seconde molécule de substrat est, quant à elle, proche du site actif. Sa présence résulte, selon les auteurs, d'un artéfact cristallographique. Elle se trouve sur la voie d'entrée du substrat et traduit vraisemblablement une position intermédiaire entre l'extérieur de la protéine et le site actif.

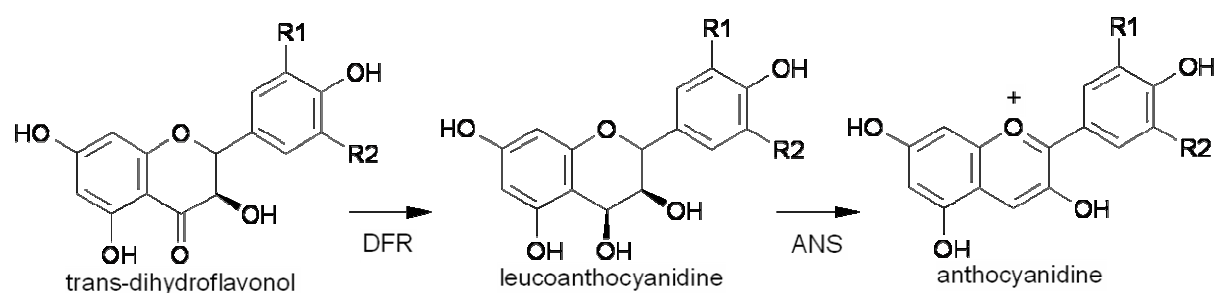


Figure 27 : Synthèse des anthocyanidines

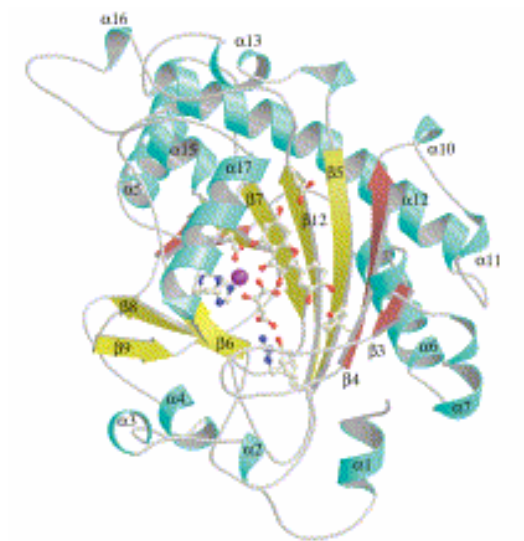


Figure 28 : Représentation en rubans du complexe de l'ANS d'*Arabidopsis thaliana* (Wilmouth et al, 2002). L'atome de fer est matérialisé en violet. Les hélices α sont en bleu. Les brins β constituant le tonneau sont en jaune et les autres brins en rouge. Substrat et co-facteur sont représentés en « ball and sticks ».

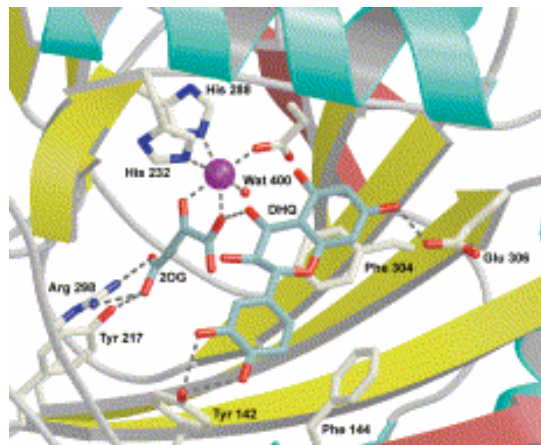


Figure 29 : Géométrie du site actif de l'ANS d'*Arabidopsis thaliana* (Wilmouth et al, 2002).

L'ANS catalyse l'oxydation du cycle C du substrat. L'étape préliminaire est la réaction du 2-oxoglutarate et du dioxygène avec l'ion Fe^{2+} . Cette étape aboutit à la libération de succinate, de CO_2 , et d'un intermédiaire ferryl réactif ($Fe(IV)=O \leftrightarrow Fe(III)-O^{\bullet}$). Celui-ci est responsable de l'oxydation du flavan-3,4-diol (*Turnbull et al, 2004*) par insaturation des carbones 2 et 3. Le composé ainsi formé est le 2-flaven-3,4-diol. Néanmoins, seul l'isomère 3-flaven-2,3-diol est réellement présent car, thermodynamiquement plus stable au pH cytosolique (*Nakajima et al, 2001*). L'isomérisation spontanée de la double liaison 2,3 en 3,4 s'effectue par élimination du groupement hydroxyl en C4 et insertion d'un même groupement en C2. Le produit coloré ne peut apparaître à ce stade que si ce composé est soumis à un pH acide. Cette condition favorise la perte de l'anion hydroxyle en C2 et la formation de l'ion flavylum.

V.3.2. Flavonoïde 3-O-glucosyltransférase (3-GT ou UFGT)

Les flavonoïdes et les anthocyanidines possèdent un groupement hydroxyle libre en C3 du cycle C, responsable de leur instabilité à un pH physiologique (*Forkmann et al, 1999*).

La 3-GT (EC : 2.4.1.115) catalyse le transfert de glucose depuis l'UDP-glucose sur la fonction 3-hydroxyle, augmentant ainsi la solubilité de ces composés dans l'eau. Les pigments ainsi formés peuvent donc être accumulés dans les vacuoles au sein d'un milieu aqueux. La structure 3D de la 3-GT de *Vitis vinifera* a été résolue récemment (**Figure 30**) (*Offen et al, 2006*). Une sous unité est constituée de deux domaines de type Rossmann ($\beta/\alpha/\beta$), comprenant respectivement les résidus 7-250 et les résidus 260-437. Elle se termine par une hélice α (résidus 437-456) qui ramène la partie C-terminale au proche voisinage de la partie N-terminale.

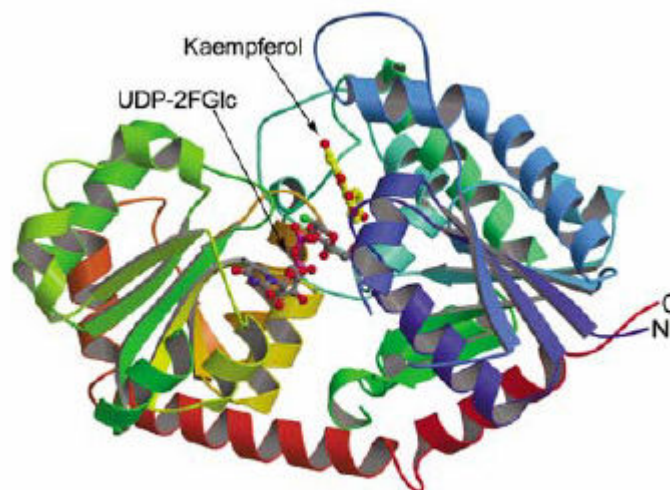


Figure 30 : Représentation en rubans de la 3-GT co-cristallisée en présence d'UDP-2-deoxy-2-fluoro glucose et de kaempferol (*Offen et al, 2006*).

La **Figure 31** décrit la géométrie du site actif. Elle montre en particulier l'importance de l'His20 qui joue le rôle de base de Brønsted dans l'activation de l'hydroxyle en C3 en vue d'une attaque nucléophile en C1 du glucose.

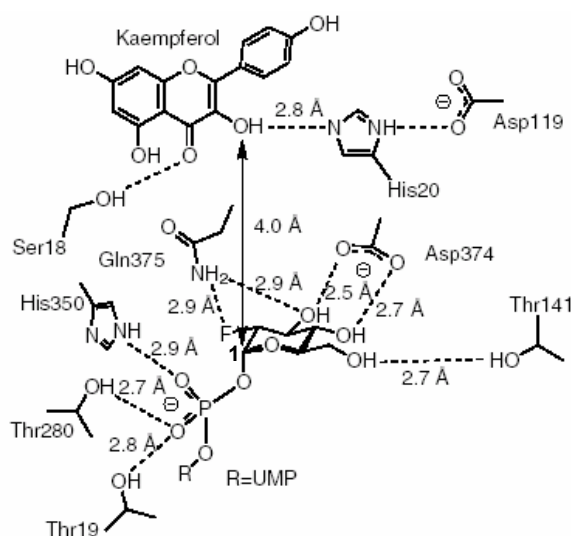


Figure 31 : Interactions au sein du site catalytique de la 3-GT (Offen et al, 2006).

V.4. Voie latérale des tanins condensés

La voie de biosynthèse des proanthocyanidines donnait lieu il y a encore peu de temps à forte controverse dans la bibliographie. Ce débat est aujourd'hui clos. Deux enzymes sont reconnues comme directement responsables de la formation des proanthocyanidines : la LAR, (leucoanthocyanidine réductase) et l'ANR (anthocyanidine réductase).

V.4.1. Leucoanthocyanidine réductase (LAR)

La LAR (EC : 1.17.1.3) catalyse la synthèse des 2R,3R catéchines (**Figure 17**) (2,3-trans-flavan-3-ols) par réduction des 2R,3R,4S-leucoanthocyanidines. Cette activité a été définie après la purification de l'enzyme à partir de feuilles de *Desmodium uncinatum* (Tanner et al 2003). Plus récemment, les gènes codant pour la LAR de *Vitis vinifera* ont été séquencés (Bogs et al, 2005). Néanmoins, ni les mécanismes réactionnels d'élimination de la fonction hydroxyle en C4, ni la structure 3D de cette enzyme ne sont connus. Cependant, par homologie de séquence, la LAR peut être considérée comme un membre de la famille des déshydrogénases à courtes chaînes, ce qui implique un repliement de type Rossmann comme celui que nous avons suggéré précédemment pour la DFR.

V.4.2. Anthocyanidine réductase (ANR)

L'anthocyanidine réductase (EC : 1.3.1.77) est une enzyme à NADPH codée par le gène BANYULS. Seuls les gènes codant pour les ANRs d'*Arabidopsis thaliana*, de *Medicago truncatula* et de *Vitis vinifera* sont connus (Devic et al, 1999 ; Xie et al, 2003 ; Bogs et al, 2005). Cette enzyme est responsable de la synthèse des épicatechines à partir des anthocyanidines (Xie et al, 2003). L'hétérocycle C, chiral pour les anthocyanidines, possède deux doubles liaisons O1-C2 et C3-C4 (Figure 32). L'ANR réduit les positions C2 et C3 pour former des flavan-3-ols entraînant une inversion de stéréochimie du carbone C3. A ce jour, ni les mécanismes réactionnels, ni la structure 3D de l'enzyme ne sont connus. La caractérisation de son activité demeure incomplète.

Toutefois un mécanisme réactionnel complexe a été proposé. Il est schématisé Figure 32. Nous en rappelons brièvement les principales étapes.

Selon une première hypothèse (Ia), l'ion hydrure, fourni par le co-facteur, permet la formation de 2R flav-3-en-3-ol par attaque nucléophile en C2. Un second ion hydrure est nécessaire pour réduire le carbone C4. Enfin, une protonation de la position C3 entraîne la réduction de la double liaison conduisant ainsi à la formation de 2R, 3R-2,3-cis-flavan-3-ols.

Selon une seconde hypothèse (Ib), l'ANR transforme l'intermédiaire flav-3-en-3-ol en sa forme tautomère flavan-3-oxonium qui est alors réduite grâce au NADPH pour former un flavan-3-ol. La configuration 3R résulte de l'attaque stéréospécifique de l'hydrure, ce qui explique la formation unique d'épicatechines (Xie et al, 2004).

Une troisième voie (IIa) est possible. L'attaque par un hydrure de la position C4 conduit à la formation du flav-2-en-3-ol, qui par réduction de la double C2-C3 par le NADPH aboutit à la formation de flavan-3-ol.

Enfin, une dernière voie de synthèse est proposée (IIb). Après attaque de la position C4 par un hydrure, une réduction de la position C2 conduit à la forme tautomère flavan-3-oxonium. Le produit final est alors obtenu par une réduction de la fonction 3-cétone.

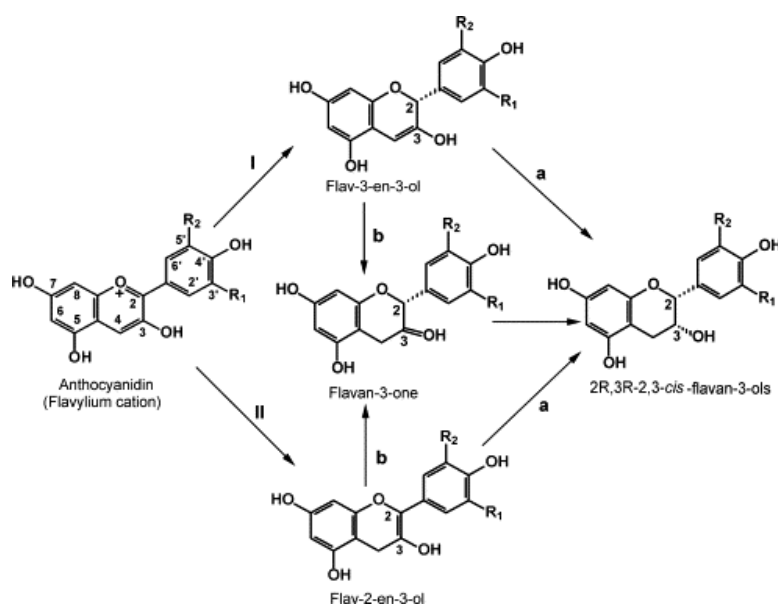


Figure 32 : Voies possibles de conversion des anthocyanidines en 2,3-cis-flavan-3-ols (Xie et al, 2004).

V.5. Fonction et mécanismes du transport vacuolaire des flavonoïdes

Les flavonoïdes, et particulièrement leurs formes hydrophobes aglycones, sont, de par leur forte réactivité chimique, toxiques pour la machinerie cellulaire. Afin de protéger la cellule, ils sont donc extraits du cytoplasme immédiatement après leur synthèse (*Debeaujon et al, 2001*). Les flavonoïdes peuvent être soit transférés dans les vacuoles (cas des anthocyanes, des proanthocyanidines, des phytoalexines, et des flavonols glycosylés) soit, excrétés vers la barrière cellulaire (cas des flavonols glycosylés polyméthylés) (*Grotewold et al, 1998*). Les vacuoles, occupant 90 % du volume cellulaire, représentent un espace de stockage suffisamment grand pour que la teneur en flavonoïdes soit adéquate avec la fonction de protection contre les prédateurs, les pathogènes ou encore les ultraviolets.

Trois processus de « détoxification » autorisent le transfert des flavonoïdes mais aussi des xénobiotiques tels que les toxines pathogènes, les pesticides, etc (**Figure 33**).

Le premier (A) nécessite l'acylation d'anthocyanes afin de permettre leur transport vers la vacuole. Ce mécanisme a été démontré chez le persil et la carotte. Néanmoins, le degré d'implication du groupement acyle dans l'interaction avec le transporteur reste à définir (*Hopp et al, 1987 ; Matern et al, 1986*).

Le deuxième mécanisme (B) permet un transport actif des flavonoïdes grâce à un antiport X/H^+ transmembranaire dépendant du gradient de protons. Ce dernier est généré par deux pompes intégrées au tonoplaste : H^+ -ATPase et H^+ -PPiase (*Klein et al, 1996*). Ce mécanisme est responsable, chez l'orge, du transport de l'isovitexine (flavone glycosylée) dans la vacuole.

Le troisième (C) fait appel aux enzymes glutathion-S-transferases ayant comme co-facteur l'ATP. Ces dernières lient de façon covalente les groupements fonctionnels des flavonoïdes à transporter sur des molécules hydrophiles telles que le glucose, le malonate, le glutathion ou le glucuronate (*Gachon et al, 2005*). Les complexes ainsi formés sont alors exportés depuis le cytosol vers les vacuoles par des transporteurs membranaires (*Klein et al, 2000*). Chez le maïs par exemple, la cyanidine-3-glucoside est conjuguée à une molécule de glutathion via la glutathion-S-transférase codée par le gène *BZ2*. Le transfert dans les vacuoles s'effectue grâce à une pompe transmembranaire (GS-X) (*Alfenito et al, 1998*). Si le gène *BZ2* est muté, les anthocyanes s'accumulent dans le cytosol et confèrent au maïs une couleur bronze issue de l'oxydation des pigments. Dans les tissus de plante, seul un faible pourcentage d'anthocyanes conjugués au glutathion est mesuré. Ce dernier est donc probablement clivé et recyclé dès l'entrée de l'anthocyane dans la vacuole (*Mol et al, 1998*).

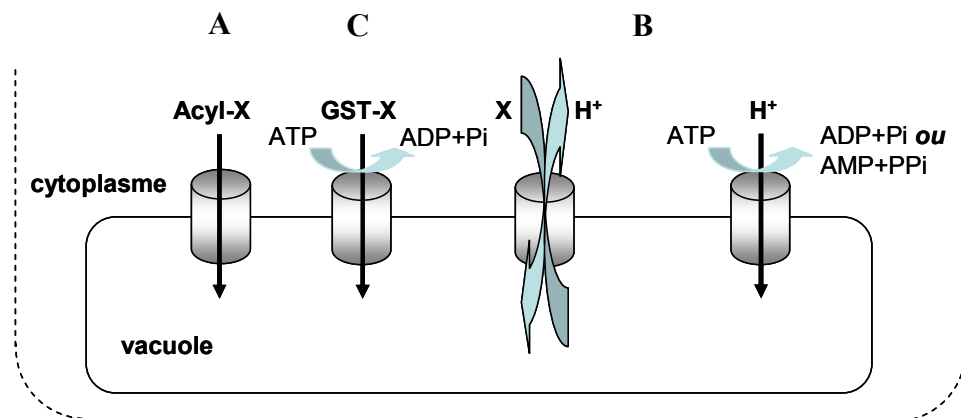


Figure 33 : Schéma récapitulatif des différents modes de transport des anthocyanes (X) dans la vacuole.

VI. Régulation moléculaire de la voie de biosynthèse des flavonoïdes

Nous avons montré dans le paragraphe précédent la complexité de la voie métabolique des flavonoïdes. Elle conduit à la synthèse d'un très grand nombre de composés qui influent sur le développement de la plante (pigmentation et protection). De nombreuses enzymes sont mises en jeu. La transcription de ces dernières est régie par différents facteurs que nous rappelons ci-dessous.

VI.1. Généralités

Les facteurs de transcription impliqués dans la régulation de la biosynthèse des anthocyanes et des proanthocyanidines peuvent être divisés en plusieurs catégories : les protéines Myb, les protéines à motif HLH (hélice-boucle-hélice) et les protéines WD40 (*Koes et al, 2005*). Bien que les facteurs Myb puissent se fixer aux séquences régulatrices du promoteur, ils ne peuvent activer seuls la transcription du gène. Pour cela, un co-activateur HLH est indispensable. Ce dernier recruterait un facteur inconnu qui se fixerait à un deuxième élément en cis du promoteur (*Hernandez et al, 2004*). Le complexe, Myb-HLH a largement été étudié chez le pétunia, le maïs et le muflier. Dans ces organismes, il joue le rôle de régulateur de la biosynthèse des seuls anthocyanes. Chez *Arabidopsis thaliana*, il régule également la synthèse des tanins condensés dans les graines (*Vom Endt et al, 2002*). Les facteurs WD40 sont, quant à eux, fortement conservés y compris dans des organismes qui ne peuvent synthétiser les anthocyanes tels les algues, les champignons ou les animaux. Leur fonction reste à préciser. Des expériences sur les mutants *tgl1* (*Arabidopsis thaliana*) et *an11* (pétunia) suggèrent que ces facteurs activeraient de façon post-traductionnelle le complexe Myb-HLH (*De Vetten et al, 1997*).

La **Figure 34** schématise la régulation des gènes impliqués dans la voie de biosynthèse des anthocyanes. Cette dernière est divisée en trois catégories qui se différencient par leur mode d'activation. Bg1, par exemple, représente un ensemble de gènes codant pour les seules enzymes matérialisées en violet. Les facteurs de transcription sont mentionnés sous leur nom commun : bHLH, Myb et WD40.

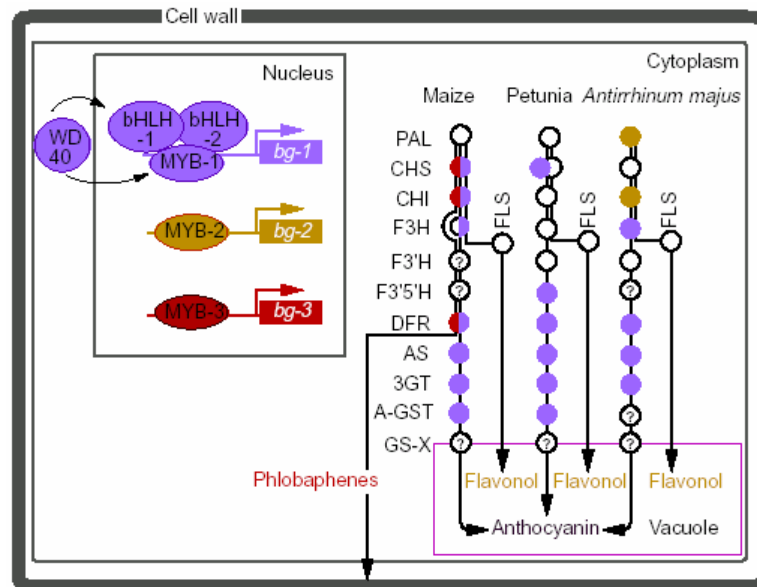


Figure 34 : Description de la coopération des facteurs de transcription impliqués dans la voie de biosynthèse des anthocyanes (Mol et al, 1998)

Chez *Vitis vinifera*, le gène *VvMYB5a* codant pour une protéine appartenant à la sous-famille des Myb-R2R3 vient d'être identifié (Deluc et al, 2006). Ce gène est principalement activé durant les premières étapes du développement de la baie dans la pellicule, la chair et les pépins. Sa surexpression chez le tabac stimule la biosynthèse de flavonols, d'anthocyanes et de tanins condensés mais affecte le métabolisme des lignines. Ce gène est donc impliqué dans le contrôle de plusieurs branches de la voie métabolique des phénylpropanoïdes.

VI.2. Approche structurale des protéines Myb

La régulation des gènes impliqués dans la biosynthèse des flavonoïdes est donc régie par un mécanisme complexe dans lequel différents facteurs de transcription sont impliqués. Une recherche bibliographique témoigne de l'abondance de ces derniers. Le choix du végétal étudié, le tissu considéré et son stade de développement sont autant de paramètres impliqués dans l'association des facteurs de transcription régulant une enzyme donnée (Marles et al, 2003).

A ce jour, aucune structure 3D de facteur de transcription de vigne n'a été décrite. La structure de facteurs Myb d'autres organismes a néanmoins été déterminée.

Ces protéines contiennent deux régions distinctes : un domaine C-terminal variable, responsable de la régulation de la transcription, et un domaine hautement conservé localisé en partie N-terminale. Ce dernier domaine est composé de trois régions consécutives R1, R2, et R3 (**Figure 35**). R1 stabilise l'interaction entre séquence d'ADN et protéine (*Howe et al, 1991*). Il n'est toutefois pas présent chez les végétaux. R2 et R3 sont impliqués directement dans la fixation de l'ADN (*Ogata et al, 1994*). Chaque domaine est constitué de trois hélices α séparées par des boucles qui facilitent l'insertion de la protéine dans l'ADN double brin. La troisième hélice est impliquée dans la reconnaissance de la courte séquence de l'ADN cible lors de la fixation du facteur sur la séquence promotrice. Le repliement des facteurs Myb est induit par le caractère amphiphile des hélices, et par la présence de trois acides aminés tryptophane hautement conservés (*Morii et al, 1999*). Leur mutation en alanine ou en acides aminés hydrophiles conduit à une forte diminution de la capacité de fixation de l'ADN (*Ogata et al, 1995*).

Parmi les trois répétitions, seule R2 possède une cavité à l'intérieur du cœur hydrophobe. Celle-ci est impliquée dans la fixation du facteur Myb à l'ADN et dans son activité de trans-activation (*Ogata et al, 1996*).

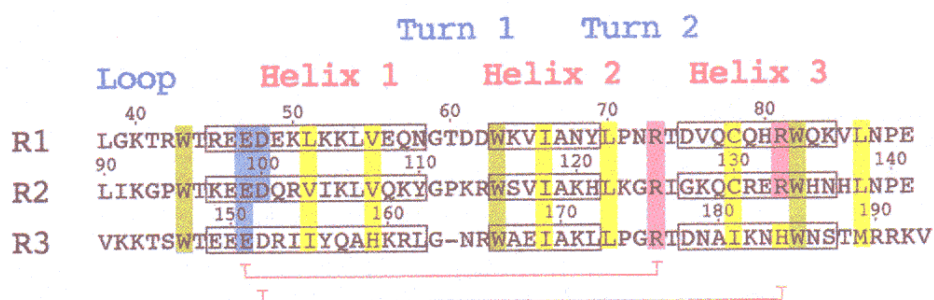


Figure 35 : Alignement de R1, R2, R3 du facteur c-Myb chez la souris

Les acides aminés conservés (W), et ceux impliqués dans le cœur hydrophobe sont respectivement colorés en vert et jaune (*Ogata et al, 1995*).

Les unités R2 et R3 indépendantes dans la forme libre, interagissent par leurs boucles N-terminales lorsque l'ADN est présent (**Figure 36**). La région R2R3 est fortement liée à la séquence spécifique de l'ADN par de multiples interactions entre la première et la seconde boucle, et le squelette phosphorylé de l'ADN. La troisième hélice se place quant à elle dans le grand sillon de l'ADN.

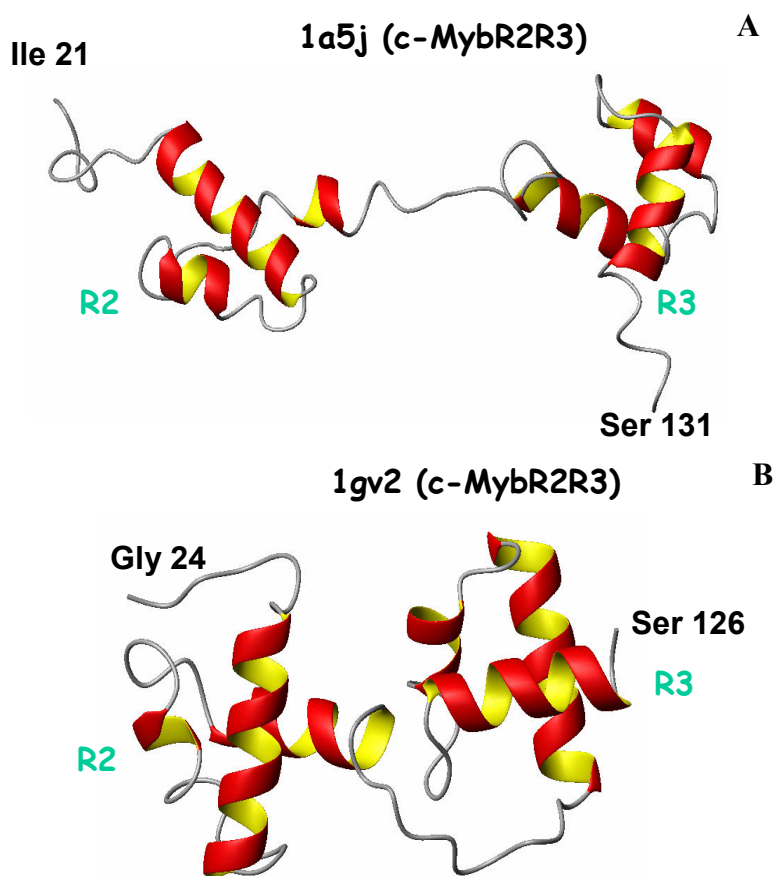


Figure 36 : Structure de R2/R3 sous forme libre (A) ou complexé avec de l'ADN lors de la co-cristallisation (B)

La figure B ne permet pas de visualiser l'ADN car le fichier PDB 1gv2 ne contient pas les coordonnées atomiques de celui-ci.

VII. Conclusion et Projet de thèse

Les flavonoïdes constituent une classe de composés très largement étudiés compte tenu de leurs propriétés originales. Chez la plante, ils sont impliqués dans des fonctions essentielles telles que la pigmentation, la reproduction, la croissance et la protection contre les pathogènes. Chez l'homme, lorsqu'ils sont ingérés, ils contribuent à la lutte contre de nombreuses maladies et ont un rôle de prévention, notamment de par leurs propriétés antioxydantes.

Parce qu'ils participent fortement à l'élaboration des propriétés sensorielles d'un vin, la filière vini-viticole souhaite disposer d'outils de suivi, voir de maîtrise de la typicité phénolique d'un vin. Y parvenir nécessite une parfaite connaissance des mécanismes tant enzymatiques que chimiques qui régissent la synthèse de ces composés dans la baie de raisin, et leurs transformations lors du processus de vieillissement du vin.

Le scientifique dispose de très nombreux outils (biochimiques, biophysiques,...) permettant d'appréhender les propriétés fonctionnelles d'une enzyme donnée et les différents paramètres susceptibles de les modifier. Les techniques de détermination de structures tridimensionnelles de ces enzymes sont l'un de ces outils. Nombre de travaux ont en effet montré l'étroite relation existant entre structure et fonctions de ces molécules. Nous nous proposons donc d'utiliser la technique de diffraction des rayons X pour l'étude structure-fonctions d'une enzyme impliquée dans la voie de biosynthèse des flavonoïdes : la dihydroflavonol 4-reductase de *Vitis vinifera*.

Les travaux de thèse s'inscrivent dans cette logique. Cette enzyme a été choisie de par sa position stratégique au sein de la voie métabolique des flavonoïdes. Ces travaux sont conduits à partir de banques d'ADNc établies à divers stades du développement de baies de raisin par le groupe Biologie de la Vigne (*UMR 619*). Ils conduiront à terme, à la synthèse d'analogues de substrat permettant en outre la détection rapide, spécifique et quantitative, des enzymes concernées dans le développement de la baie de raisin.

Dans ce manuscrit sont présentés les résultats obtenus pour cette enzyme. Ces derniers portent sur sa surexpression chez *E. coli* et sur ses conditions de purification qui sont les travaux préalables à toute étude structurale. La protéine ainsi purifiée a permis l'obtention d'un cristal de qualité suffisante pour la résolution de sa structure 3D. Cette structure a, par la suite, servi de modèle pour l'étude de l'influence des groupements hydroxyles du cycle B de différents analogues de substrat sur la fixation de ces derniers à l'enzyme. Parallèlement, la caractérisation des acides aminés impliqués dans la spécificité de l'enzyme pour son substrat a été entreprise via la surexpression de DFRs mutées. Les propriétés enzymatiques de ces protéines ont été comparées à celles de la DFR sauvage.

Enfin, l'ensemble des résultats est discuté et comparé à ceux publiés dans la littérature. Nous y discutons notamment la méthode de détection des produits enzymatiques, l'influence du résidu 133, et enfin la théorie expliquant la spécificité de substrat.

RESULTATS

RESULTATS

I. Première approche structurale par analyse bioinformatique

Deux séquences de DFR de *Vitis vinifera* sont disponibles dans la banque de données EMBL (<http://www.ebi.ac.uk/>). Elles sont accessibles sous les références P51110 (cépage *Lambrusco*) et P93799 (cépage *Cabernet sauvignon*). Leurs séquences primaires sont composées de 337 résidus et ne diffèrent que de trois acides aminés. Leur alignement (Figure 37) montre la position et la nature des résidus divergents.

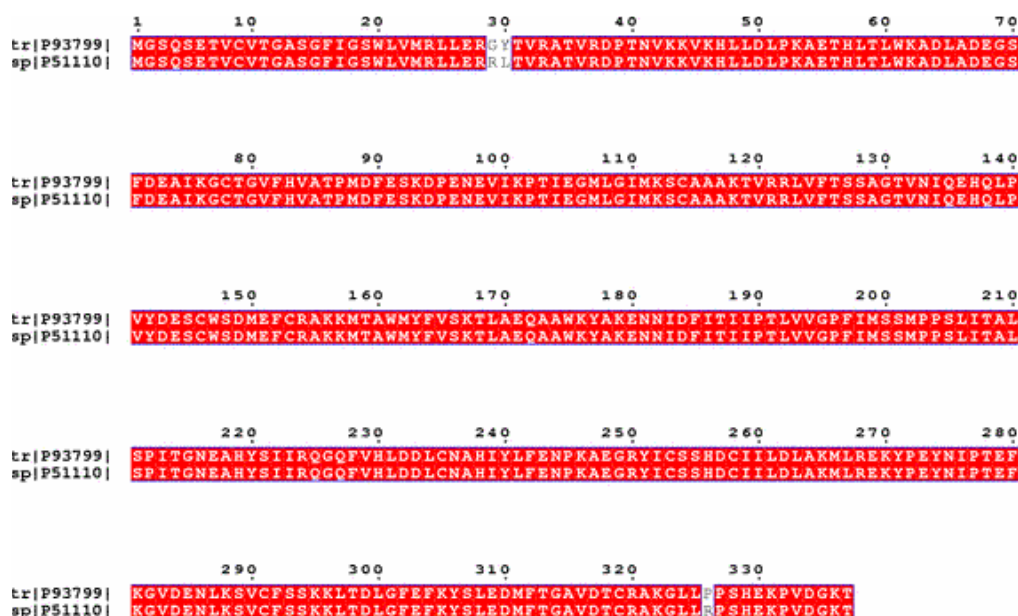


Figure 37 : Alignement des séquences P51110 et P93799

L'ensemble du travail effectué dans ce mémoire concerne la séquence protéique P93799, la banque d'ADNc utilisée lors du clonage de la DFR étant celle du cépage *Cabernet sauvignon* (cf : chapitre Matériel et méthodes). Avant toute investigation des propriétés enzymatiques et structurales de cette enzyme, nous avons souhaité utiliser les outils offerts par la bio-informatique pour une première caractérisation de celle-ci.

Lorsqu'aucune structure tridimensionnelle de protéine n'est déterminée, cette dernière peut néanmoins être estimée à l'aide d'outils informatiques tels que : Swiss-Model disponible à l'adresse suivante (<http://swissmodel.expasy.org//SWISS-MODEL.html>). Une condition doit néanmoins être respectée : la séquence primaire de la protéine dont on recherche la structure doit posséder une identité de séquence voisine ou supérieure à 30% avec celle d'une protéine de structure connue. Aucune prédiction de structure de ce type n'a pu être réalisée pour la DFR en 2003 (année où débuta ce projet de recherche). En effet, à cette date, la meilleure identité de séquence mesurée était de 19%. Elle concernait l'UDP-galactose 4 épimérase d'*E. coli*. Afin d'obtenir de premiers éléments structuraux, il nous a donc fallu nous tourner vers l'utilisation d'autres types d'outils.

Le serveur « NCBI conserved domain search » permet, à partir de la séquence primaire d'une protéine, l'identification de domaines souvent responsables d'une fonction spécifique donnée (<http://130.14.29.110/Structure/cdd/cdd.shtml>). Pour cela, il fait appel à une banque de données dans laquelle chaque domaine prédéfini a été identifié par un alignement multiple. Nous avons mis à profit l'existence d'un tel serveur pour rechercher la présence d'un de ces domaines dans la séquence primaire de la DFR. Cette démarche a permis d'identifier un domaine commun aux déshydrogénases à courtes chaînes et aux épimérases. Parmi l'ensemble des protéines appartenant à ces familles, les plus fortes identités de séquence trouvées avec la DFR concernent: l'UDP-galactose 4 épimérase (code pdb : 1LRJ) (*Thoden et al, 1996*), la tropinone réductase 2 (code pdb : 2AE2) (*Yamashita et al, 1999*) et l'alcool déshydrogénase (code pdb : 1A4U) (*Benach et al, 1998*). Les identités de séquence sont respectivement de 19%, 10% et 6% (BioEdit : <http://www.mbio.ncsu.edu/BioEdit/bioedit.html>). Alors que nous nous attendions à une identité de séquence forte entre ces trois protéines, celle-ci ne dépasse pas 4%. Ce dernier point illustre le fait que des protéines puissent posséder une fonction identique, sans toutefois présenter une identité de séquence importante.

Les structures 3D de ces protéines présentent des caractéristiques communes. Leur repliement met en évidence deux domaines distincts:

- ✓ Un premier domaine, N-terminal, s'étend sur quelques 200 acides aminés. Son repliement est de type Rossmann. Il est caractérisé par la succession de 6 motifs de type β/α (**Figure 38**).
- ✓ Un second domaine, C-terminal, est de longueur variable. Sa conformation semble être propre à chaque enzyme.

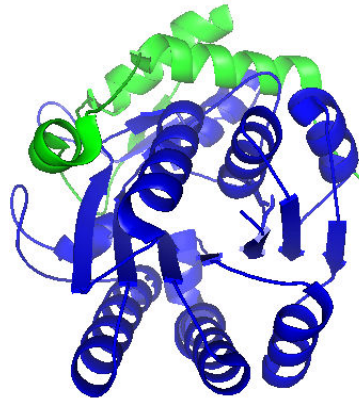


Figure 38 : Structure 3D de l'UDP galactose 4 épimérase (Thoden et al, 1996)

Le repliement de Rossmann, composé de six brins β en configuration parallèle et de six hélices α , trois de chaque côté du feuillet β , est ici représenté en bleu.

La **figure 39** atteste de la bonne superposition des domaines N-terminaux, comme l'indiquent les valeurs des rms moyens calculés à partir des seules positions des atomes de carbone C α équivalents :

- ✓ UDP gal 4 épimérase vs alcool deshydrogénase : **1,62 Å**
- ✓ UDP gal 4 épimérase vs tropinone réductase 2 : **1,97 Å**
- ✓ Alcool deshydrogénase vs tropinone réductase 2 : **1,52 Å**

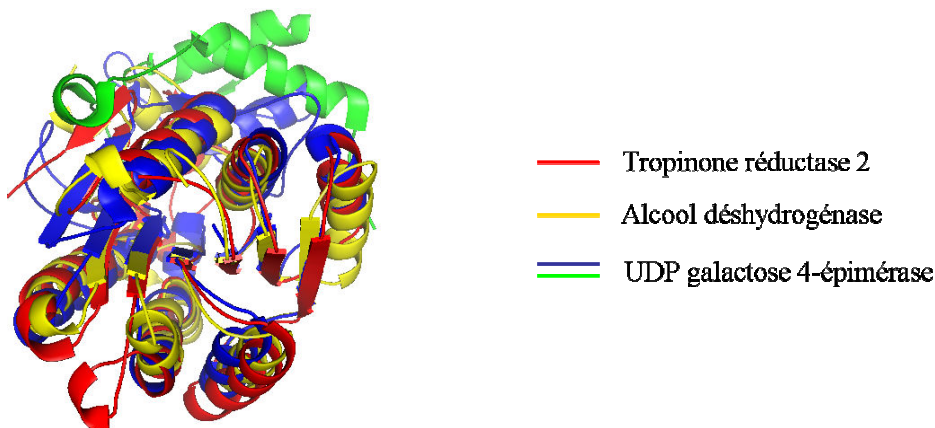


Figure 39 : Représentation en rubans des structures de référence superposées. La superposition est obtenue à l'aide du logiciel graphique Pymol (DeLano, 2002).

Une analyse plus approfondie des structures primaires des domaines N-terminaux met en évidence la présence de motifs caractéristiques associés à la fonction commune des enzymes : le premier, riche en glycine, permet la fixation du co-facteur (NADPH), le second, est lié à l'activité enzymatique et est composé de trois acides aminés : une sérine en amont et distante des deux autres, une tyrosine et une lysine séparées l'une de l'autre par trois résidus quelconques.

I.1. Recherche chez la DFR du site de fixation du co-facteur

Dans l'ensemble des structures de protéines de la famille des SDRs, le site de reconnaissance du NADPH est un site riche en glycine, de type GxGxxG ou Gx(x)GxxxG (Kleiger *al*, 2002). Ce motif se retrouve dans l'UDP-galactose 4-épimérase, les trois résidus glycine sont en position 8, 10 et 13. Toutefois, ce motif peut être dégénéré dans d'autres protéines de la famille. Dans l'alcool déshydrogénase et la tropinone réductase 2, seuls deux résidus glycine sont trouvés en position 15 et 18 pour la première et en positions 17 et 22 pour la seconde. Seule une superposition des structures 3D, **Figure 40**, met en évidence la stricte conservation de la géométrie spatiale de ce motif. Ceci nous permet d'identifier le troisième partenaire. L'alanine 13 remplace la première glycine dans l'alcool déshydrogénase et l'arginine 19 substitue la seconde glycine dans la tropinone réductase 2. La faible identité de séquence entre ces protéines (<4%) n'autorisait pas l'identification de ces résidus à partir d'un simple alignement de séquences primaires.

Le motif de reconnaissance du NADPH, constitué de trois glycines, peut donc ne pas être strictement conservé. Par superposition de la séquence primaire de la DFR de *Vitis vinifera* à celle des trois protéines précédentes, ce motif serait dégénéré et constitué des résidus Ala13, Gly15 et Gly18. Néanmoins, la présence d'une glycine en position 12 de la séquence DFR laisse à penser que le site de fixation du NADPH serait de type GxxGxxG. Pour que ce motif joue son rôle, cela implique une stricte superposition des trois résidus glycine à celle trouvée pour les trois enzymes de référence et donc, chez la vigne, la présence d'une boucle plus étendue (un acide aminé supplémentaire) entre les deux premières glycines.

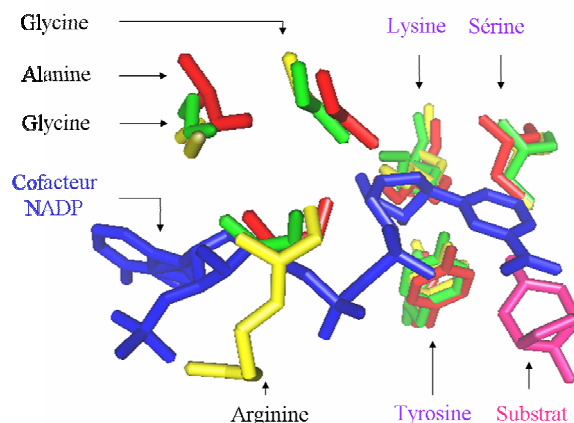


Figure 40 : Représentation du site de fixation du NADPH et de la triade catalytique des trois enzymes de référence

Le co-facteur est ici coloré en bleu et le substrat de la tropinone réductase 2 (seule enzyme ayant été co-cristallisée avec son substrat) en magenta. Le code couleur utilisé pour l'identification des enzymes est identique à celui de la **Figure 39**.

I.2. Recherche du site catalytique chez la DFR

Le motif catalytique représenté par la triade d'acides aminés (S...YxxxK) est parfaitement conservé dans les trois protéines modèles. Malgré le non-alignement des séquences primaires, nous pouvons remarquer par superposition des structures 3D (**Figure 40**) que les résidus impliqués dans le processus enzymatique se superposent parfaitement, rendant compte là encore d'une géométrie spatiale parfaitement conservée.

Une analyse attentive de la séquence primaire de la DFR fait apparaître un motif YxxxK entre les positions 163 et 167 qui laisse à penser que la tyrosine 163 et la lysine 167 sont deux des trois résidus impliqués dans la triade catalytique. La longueur du segment entre la sérine et la tyrosine étant variable, il est délicat, à ce stade, d'identifier quelle sérine dans la séquence de la DFR constitue le troisième résidu.

I.3. Eléments de structure secondaire de la partie N-terminale de la DFR

A ce stade du travail, il est fortement probable que la DFR de *Vitis vinifera* appartienne à la famille des SDR comme l'indiquait la recherche de domaines conservés effectuée sur le serveur NCBI. Afin de vérifier que le domaine N-terminal de la DFR adopte un repliement de type Rossmann, nous avons entrepris une prédiction des éléments de structure secondaire pouvant être présents. Pour cela, nous avons utilisé plusieurs programmes (Predator (*Frishman & Argos, 1995*), Porter (*Pollastri & McLysaght, 2005*), Psipred (*McGuffin et al, 2000*) et Gor IV (*Garnier et al, 1996*)) afin de minimiser la variabilité de prédiction.

Les différentes prédictions obtenues sont présentées dans la **Figure 41** :

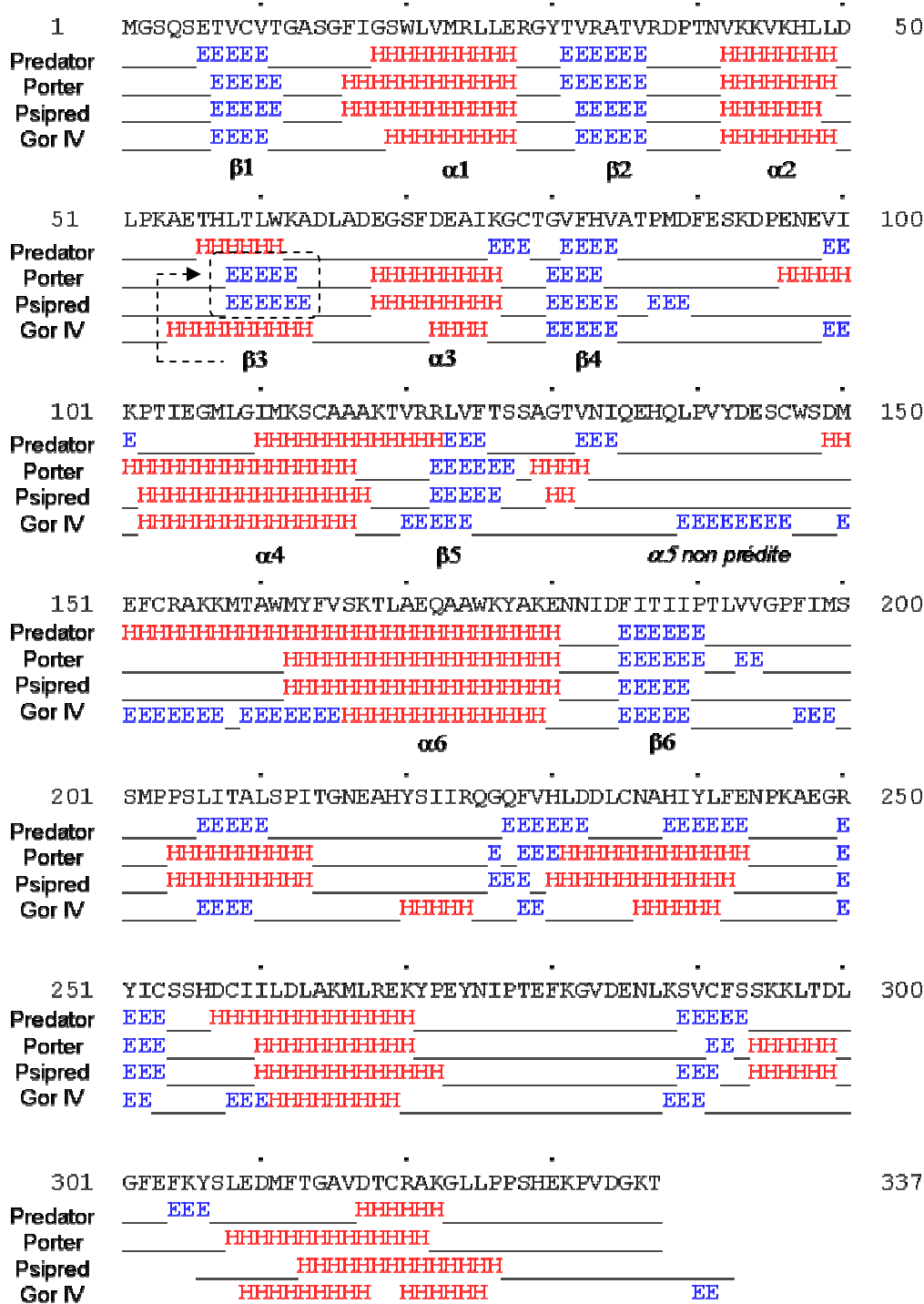


Figure 41 : Prédiction des éléments de structure secondaire de la DFR
 Les hélices α et les brins β sont respectivement représentés en bleu et rouge.

L'analyse des résultats tend à montrer que la partie N-terminale (200 premiers acides aminés) est constituée d'une succession de motifs β/α caractéristiques du repliement de type Rossmann attendu, ce qui renforce une fois de plus le fait que la DFR de *Vitis vinifera* peut appartenir à la famille des SDR.

I.4. Alignement de séquences de DFR

Parmi l'ensemble des séquences primaires de DFR disponibles dans les diverses banques de données, nous en avons considéré dans ce travail uniquement quelques unes représentatives de la diversité existant au sein du règne végétal. Le programme utilisé pour réaliser l'alignement multiple est : ClustalW : <http://www.ch.embnet.org/software/ClustalW.html>. L'alignement obtenu est présenté sur les deux pages suivantes. La numérotation adoptée est celle de la séquence P93799 de *Vitis vinifera*.

Un examen rapide témoigne que :

- ✓ le site de reconnaissance du NADPH (résidus 12 à 18) au sein de la famille des DFRs est un motif ubiquitaire de type GxxGxxG,
- ✓ la triade catalytique (S128, Y163, K167) est strictement conservée. Le choix de la sérine 128 sera justifié lors de la description structurale de la protéine étudiée.

Sur cet alignement, sont également surlignés, le segment 131-156, les résidus 133 et 227. Les raisons de leur mise en évidence sont explicitées dans le chapitre II.7.4 (Architecture du site catalytique).

DIHYDROFLAVONOL 4-REDUCTASES

	10	20	30	40	50	60	70	80	90																																																																																	
<i>Vitis vinifera</i>	-----MG	SQSE	TV	CV	T	CA	S	F	I	G	S	W	L	V	M	R	L	L	E	R	G	T	V	R	A	T	V	R	-	D	P	T	N	V	K	V	K	H	L	L	D	L	P	K	A	E	T	H	L	T	L	M	K	A	D	L	A	D	E	G	S	F	D	E	A	I	K	G	T	G	V	F	H	V	A	T	P	M	D	F										
<i>Malus domestica</i>	-----MG	SE	S	V	C	V	T	C	A	S	F	I	G	S	W	L	V	M	R	L	L	E	H	G	T	V	R	A	T	V	R	-	D	P	T	N	V	K	V	K	H	L	L	D	L	P	K	A	E	T	H	L	T	L	M	K	A	D	L	A	D	E	G	S	F	D	E	A	I	K	G	S	G	V	F	H	V	A	T	P	M	D								
<i>Rosa hybrida</i>	-----MA	S	E	S	V	C	V	T	C	A	S	F	I	G	S	W	L	V	M	R	L	L	D	R	G	T	V	R	A	T	V	R	-	D	P	A	N	K	V	K	V	N	H	L	L	D	L	P	K	A	A	T	H	L	T	L	M	K	A	D	L	A	E	E	S	F	D	E	A	I	K	G	T	G	V	F	H	V	A	T	P	M								
<i>Camellia sinensis</i>	-----MK	D	S	V	A	S	A	T	A	S	A	P	T	C	V	T	C	A	A	G	F	I	G	S	W	L	V	M	R	L	L	E	R	G	T	V	R	A	T	V	R	-	D	P	A	N	L	K	V	K	V	H	L	L	D	L	P	K	A	D	T	N	L	T	L	M	K	A	D	L	N	E	E	S	F	D	E	A	I	K	G	S	G	V	F	H	V	A	T	P
<i>Glycine max</i>	-----MG	S	A	S	A	T	C	V	T	C	A	A	G	F	I	G	S	W	L	V	M	R	L	L	E	R	G	T	V	R	A	T	V	R	-	D	P	A	N	M	K	V	K	H	L	L	E	L	P	G	A	K	S	K	L	S	I	M	K	A	D	L	A	E	E	S	F	D	E	A	I	K	G	S	G	V	F	H	V	A	T									
<i>Arabicopsis thal.</i>	-----MV	S	Q	K	E	T	C	V	T	C	A	S	F	I	G	S	W	L	V	M	R	L	L	E	R	G	T	V	R	A	T	V	R	-	D	P	G	N	L	K	V	Q	H	L	L	D	L	P	N	A	K	T	Q	L	T	L	M	K	A	D	L	S	E	E	S	F	D	D	A	I	K	G	S	G	V	F	H	V	A											
<i>Gerbera hybrida</i>	-----ME	E	D	S	P	A	T	C	V	T	C	A	A	G	F	I	G	S	W	L	V	M	R	L	L	E	R	G	T	V	R	A	T	V	R	-	D	P	G	L	K	V	K	H	L	L	E	L	P	K	A	Q	T	N	L	K	I	M	K	A	D	L	T	Q	E	G	S	F	D	E	A	I	K	G	S	G	V	F	H											
<i>Medicago trunc. 1</i>	-----MV	K	L	H	T	T	C	V	T	C	A	S	F	I	G	S	W	L	V	M	R	L	L	E	R	G	T	V	R	A	T	V	R	-	D	P	E	N	L	K	V	K	H	L	L	E	L	P	G	A	K	G	K	L	S	I	M	K	A	D	L	G	E	E	S	F	D	E	A	I	K	G	S	G	V	F														
<i>Daucus Carota</i>	-----MG	S	A	A	K	T	C	V	T	C	A	A	G	F	I	G	S	W	L	V	M	R	L	L	E	R	G	T	V	R	A	T	V	R	-	D	P	G	N	P	K	V	K	H	L	L	E	L	P	G	A	K	T	N	L	T	I	M	N	A	D	L	T	E	E	S	F	D	E	A	I	K	G	S	G	V														
<i>Lotus japo. 1</i>	-----MG	S	V	P	E	T	C	V	T	C	A	A	G	F	I	G	S	W	L	V	M	R	L	L	E	R	G	T	V	R	A	T	V	R	-	D	P	A	N	M	K	V	K	H	L	L	E	L	P	E	A	K	T	L	T	L	M	K	A	D	L	A	E	E	S	F	D	E	A	I	K	G	S	G																
<i>Lotus japo. 2</i>	-----MG	S	A	A	K	T	C	V	T	C	A	A	G	F	I	G	S	W	L	V	M	R	L	L	E	R	G	T	V	R	A	T	V	R	-	D	P	A	N	M	K	V	K	H	L	L	E	L	P	G	A	K	T	N	L	T	I	M	N	A	D	L	T	E	E	S	F	D	E																					
<i>Lotus japo. 3</i>	-----MS	S	E	S	E	T	C	V	T	C	A	A	G	F	I	G	S	W	L	V	M	R	L	L	E	R	G	T	V	R	A	T	V	R	-	D	P	A	N	M	K	V	K	H	L	L	E	L	P	D	A	K	T	L	S	I	M	K	A	D	L	A	E	E	S	F	D																							
<i>Lotus japo. 5</i>	-----MA	S	E	V	H	S	V	D	A	H	S	P	K	T	P	T	C	A	A	G	F	I	G	S	W	L	V	M	R	L	L	E	R	G	T	V	R	A	T	V	R	-	D	P	E	N	Q	K	V	K	H	L	L	E	L	P	D	A	K	T	L	S	I	M	K	A	D	L	A	E	E																			
<i>Solanum tuberosum</i>	-----MG	S	V	S	E	T	C	V	T	C	A	A	G	F	I	G	S	W	L	V	M	R	L	L	E	R	G	T	V	R	A	T	V	R	-	D	P	D	N	M	K	V	K	H	L	L	E	L	P	G	A	N	S	K	L	S	I	M	K	A	D	L	G	E	E																									
<i>Medicago trunc. 2</i>	-----MG	S	I	A	E	T	C	V	T	C	A	A	G	F	I	G	S	W	L	V	M	R	L	L	E	R	G	T	V	R	A	T	V	R	-	D	P	D	N	M	K	V	K	H	L	L	E	L	P	K	A	S	T	H	L	T	L	M	K	A	D	L	A	E																										
<i>Citrus sinensis</i>	-----MG	T	E	A	E	T	C	V	T	C	A	A	G	F	I	G	S	W	L	V	M	R	L	L	E	R	G	T	V	R	A	T	V	R	-	D	P	D	N	M	K	V	K	H	L	L	E	L	P	K	A	S	T	H	L	T	L	M	K	A	D	L	S	E																										
<i>Populus tremul.</i>	-----MA	S	E	A	V	H	A	P	P	V	P	T	C	V	T	C	A	A	G	F	I	G	S	W	L	V	M	R	L	L	E	R	G	T	V	R	A	T	V	R	-	D	P	E	N	M	K	V	K	H	L	L	E	L	P	K	A	D	T	N	L	T																												
<i>Petunia hybrida</i>	-----ES	K	D	P	E	N	E	V	I	K	P	T	I	E	G	M	L	G	I	M	K	S	C	A	A	A	K	T	V	R	R	L	V	F	T	S	A	G	T	V	N	I	Q	E	H	Q	L	P	V	D	E	S	C	M	S	D	M	E	F	C	R	A	K	K	M	T	A	M	M	F																				

133

	100	110	120	130	140	150	160	170	180	190																																																																																				
<i>Vitis vinifera</i>	ESK	D	P	E	N	E	V	I	K	P	T	I	E	G	M	L	G	I	M	K	S	C	A	A	A	K	T	V	R	R	L	V	F	T	S	A	G	T	V	N	I	Q	E	H	Q	L	P	V	D	E	S	C	M	S	D	M	E	F	C	R	A	K	K	M	T	A	M	M	F	V	S	K	T	L	A	E	Q	A	A	M	K	Y	A	K	E	N	N	I	D	F	I	T	I	I
<i>Malus domestica</i>	ESK	D	P	E	N	E	V	I	K	P	T	I	E	G	M	L	G	I	M	K	S	C	A	A	A	K	T	V	R	R	L	V	F	T	S	A	G	T	V	N	I	Q	E	H	Q	L	P	V	D	E	S	C	M	S	D	M	E	F	C	R	S	V	K	M	T	G	M	M	F	V	S	K	T	L	A	E	Q	A	A	M	K	Y	A	K	E	N	N							
<i>Rosa hybrida</i>	ESK	D	P	E	N	E	V	I	K	P	T	I	E	G	M	L	G	I	M	K	S	C	A	A	A	K	T	V	R	R	L	V	F	T	S	A	G	T	V	N	V	E	E	H	Q	L	P	V	D	E	S	N	M	S	D	E	F	C	R	R	V	K	M	T	G	M	M	F	V																									
<i>Camellia sinensis</i>	ESK	D	P	E	N	E	V	I	K	P	T	I	E	G	M	L	G	I	M	K	S	C	A	A	A	K	T	V	R	R	L	V	F	T	S	A	G	T	V	N	V	E	E	H	Q	L	P	V	D	E	N	M	S	D	E	F	C	R	R	V	K	M	T	G	M	M	F																											
<i>Glycine max</i>	ESK	D	P	E	N	E	V	I	K	P	T	I	E	G	M	L	G	I	M	K	S	C	A	A	A	K	T	V	R	R	L	V	F	T	S	A	G	T	V	N	V	E	E	H	Q	L	P	V	D	E	N	M	S	D	E	F	C	R	R	V	K	M	T	G	M	M																												
<i>Arabicopsis thal.</i>	ESK	D	P	E	N	E	V	I	K	P	T	I	E	G	M	L	G	I	M	K	S	C	A	A	A	K	T	V	R	R	L	V	F	T	S	A	G	T	V	N	V	E	E	H	Q	L	P	V	D	E	N	M	S	D	E	F	C	R	R	V	K	M	T	G	M																													
<i>Gerbera hybrida</i>	ESK	D	P	E	N	E	V	I	K	P	T	I	E	G	M	L	G	I	M	K	S	C	A	A	A	K	T	V	R	R	L	V	F	T	S	A	G	T	V	N	V	E	E	H	Q	L	P	V	D	E	N	M	S	D	E	F	C	R	R	V	K	M	T	G																														
<i>Medicago trunc. 1</i>	ESK	D	P	E	N	E	V	I	K	P	T	I	E	G	M	L	G	I	M	K	S	C	A	A	A	K	T	V	R	R	L	V	F	T	S	A	G	T	V	N	V	E	E	H	Q																																																	

DIHYDROFLAVONOL 4 - REDUCTASES

	200	210	220	230	240	250	260	270	280	290							
<i>Vitis vinifera</i>	VGPFIMSMP	PPSLITL	SPITGNE	AHYSI	IRQG	FVHLDD	LCMAHI	YLFENP	KAEGRY	ICSSHD	CIILDL	AKMLRE	KYPEYNI	PTEFKG	VDENLK	SVCFSS	SSKK
<i>Malus domestica</i>	IGPFLMPS	MPPSLIT	LSPILRN	ESHYGI	IKQG	FVHLDD	LCLSHI	YLIEHP	KAEGRY	ICSSHD	ATIHEL	VKMLRE	KYPEYNI	PTKFKG	IDDNLE	IPVHFS	SSKK
<i>Rosa hybrida</i>	IGPFLMPS	MPPSLIT	LSPITGNE	SHYSI	IKQG	FVHLDD	LCQSHI	YLIEHP	KAEGRY	ICSSHD	ATIHEL	IAKLLG	KYPEYNI	PTKFKG	EENL	PKVHFS	SSKK
<i>Camellia sinensis</i>	VGPFIMPT	FPSSLIT	ALSPITR	NEGHYSI	IKQG	FVHLDD	LCESHIF	LYERPE	QAEGRY	ICSSHD	ATIHDIA	KMLREK	WPEYNI	PTKFKG	DKDL	PVVSFS	SSKK
<i>Glycine max</i>	VGPFILMT	SMPPSLIT	ALSPITR	NEGHYSI	IKQG	FVHLDD	LCMAHI	FLYEQAA	KAAGRY	ICSSHD	ATIITIS	KFLRP	KYPEYNI	PTKFKG	VIQLE	LVRVFS	SSKK
<i>Arabidopsis thal.</i>	VGPFITST	FFPSSLIT	ALSLITG	NEAHYSI	IKQG	FVHLDD	LCMAHI	FLYEQAA	KAAGRY	ICSSHD	ATIHQ	LAKI	IKKWPE	EYIPTK	FGIDEL	ELPIVS	FS
<i>Gerbeza hybrida</i>	VGPFILPT	MPPSSLIT	ALSPITG	NEAHYSI	IKQG	FVHLDD	LCMAHI	FLYERPE	KAEGRY	ICSSHD	ATIHDIA	KLINTK	KYPEYNI	PTKFKG	NIPE	DELVRF	SSKK
<i>Medicago trunc.</i> 1	VGPFILPT	MPPSSLIT	ALSPITG	NEAHYSI	IKQG	FVHLDD	LCMAHI	FLYERPE	KAEGRY	ICSSHD	ATIHDIA	KLINTK	KYPEYNI	PTKFKG	NIPE	DELVRF	SSKK
<i>Daucus Carota</i>	VGPFILPT	MPPSSLIT	ALSPITG	NEAHYSI	IKQG	FVHLDD	LCMAHI	FLYERPE	KAEGRY	ICSSHD	ATIHDIA	KLINTK	KYPEYNI	PTKFKG	NIPE	DELVRF	SSKK
<i>Lotus japo.</i> 1	VGPFILPT	MPPSSLIT	ALSPITG	NEAHYSI	IKQG	FVHLDD	LCMAHI	FLYERPE	KAEGRY	ICSSHD	ATIHDIA	KLINTK	KYPEYNI	PTKFKG	NIPE	DELVRF	SSKK
<i>Lotus japo.</i> 2	VGPFILPT	MPPSSLIT	ALSPITG	NEAHYSI	IKQG	FVHLDD	LCMAHI	FLYERPE	KAEGRY	ICSSHD	ATIHDIA	KLINTK	KYPEYNI	PTKFKG	NIPE	DELVRF	SSKK
<i>Lotus japo.</i> 3	VGPFILPT	MPPSSLIT	ALSPITG	NEAHYSI	IKQG	FVHLDD	LCMAHI	FLYERPE	KAEGRY	ICSSHD	ATIHDIA	KLINTK	KYPEYNI	PTKFKG	NIPE	DELVRF	SSKK
<i>Lotus japo.</i> 5	VGPFILPT	MPPSSLIT	ALSPITG	NEAHYSI	IKQG	FVHLDD	LCMAHI	FLYERPE	KAEGRY	ICSSHD	ATIHDIA	KLINTK	KYPEYNI	PTKFKG	NIPE	DELVRF	SSKK
<i>Solanum tuberosum</i>	VGPFILPT	MPPSSLIT	ALSPITG	NEAHYSI	IKQG	FVHLDD	LCMAHI	FLYERPE	KAEGRY	ICSSHD	ATIHDIA	KLINTK	KYPEYNI	PTKFKG	NIPE	DELVRF	SSKK
<i>Medicago trunc.</i> 2	VGPFILPT	MPPSSLIT	ALSPITG	NEAHYSI	IKQG	FVHLDD	LCMAHI	FLYERPE	KAEGRY	ICSSHD	ATIHDIA	KLINTK	KYPEYNI	PTKFKG	NIPE	DELVRF	SSKK
<i>Citrus sinensis</i>	VGPFILPT	MPPSSLIT	ALSPITR	NEAHYSI	IKQG	FVHLDD	LCMAHI	FLYERPE	KAEGRY	ICSSHD	ATIHDIA	KLINTK	KYPEYNI	PTKFKG	NIPE	DELVRF	SSKK
<i>Populus tremul.</i>	VGPFILPT	MPPSSLIT	ALSPITG	NEAHYSI	IKQG	FVHLDD	LCMAHI	FLYERPE	KAEGRY	ICSSHD	ATIHDIA	KLINTK	KYPEYNI	PTKFKG	NIPE	DELVRF	SSKK
<i>Petunia hybrida</i>	VGPFILPT	MPPSSLIT	ALSPITG	NEAHYSI	IKQG	FVHLDD	LCMAHI	FLYERPE	KAEGRY	ICSSHD	ATIHDIA	KLINTK	KYPEYNI	PTKFKG	NIPE	DELVRF	SSKK
<i>Vitis vinifera</i>	LTDLGF	FEFKYS	LEDMFT	GAVD	TCRA	KGLLP	---	PSHKP	VDGKT	---	---	---	---	---	---	---	---
<i>Malus domestica</i>	LREIG	FEFKYS	LEDMFT	VGAVD	ACKR	AGLLP	---	IPAEK	TEAAEES	NLVD	KVG	---	---	---	---	---	---
<i>Rosa hybrida</i>	LLETF	FEFKYS	LEDMFT	VGAVD	ACKR	AGLLP	---	PPRVE	KQEVDES	SVVR	KVGTG	---	---	---	---	---	---
<i>Camellia sinensis</i>	LIGM	FEFKYS	LEDMFT	RGAID	TCRE	KGLLP	---	HSF	AMP	VNGN	KV	---	---	---	---	---	---
<i>Glycine max</i>	ITDLG	FEFKYS	LEDMFT	GAID	TCR	KGLLP	---	KPA	EG	LFTK	PGET	PV	NAMHK	---	---	---	---
<i>Arabidopsis thal.</i>	LITDMG	FNFKYS	LEDMFT	GAID	TCR	KGLLP	---	VS	LSY	QSI	SEI	KT	KNEN	ID	VKT	GDG	LTDG
<i>Gerbeza hybrida</i>	LITDMG	FNFKYS	LEDMFT	GAID	TCR	KGLLP	---	YST	IR	NH	ING	HV	NG	VH	YI	KN	DD
<i>Medicago trunc.</i> 1	IKDLG	FEFKYS	LEDMFT	GAID	TCR	KGLLP	---	---	---	---	---	---	---	---	---	---	---
<i>Daucus Carota</i>	LITDMG	FTFKY	NLEDM	FKGAI	ETCR	EKGLLP	---	---	---	---	---	---	---	---	---	---	---
<i>Lotus japo.</i> 1	IKDMG	FQFKY	LTLEDM	FKGAI	DACR	EKGLLP	---	---	---	---	---	---	---	---	---	---	---
<i>Lotus japo.</i> 2	IKDMG	FEFKYS	LEDMFT	GAID	TCR	EKGLLP	---	---	---	---	---	---	---	---	---	---	---
<i>Lotus japo.</i> 3	IKDMG	FEFKYS	LEDMFT	GAID	TCR	EKGLLP	---	---	---	---	---	---	---	---	---	---	---
<i>Lotus japo.</i> 5	ITDLG	FQFKY	LTLEDM	FKGAI	ETCR	EKGLLP	---	---	---	---	---	---	---	---	---	---	---
<i>Solanum tuberosum</i>	LMDMG	FQFKY	LTLEDM	FKGAI	ETCR	QKLLP	---	---	---	---	---	---	---	---	---	---	---
<i>Medicago trunc.</i> 2	ITDLG	FIFKYS	LEDMFT	GAID	TCR	EKGLLP	---	---	---	---	---	---	---	---	---	---	---
<i>Citrus sinensis</i>	LITDLG	FEFKYS	LEDMFT	GAID	TCR	AKGLLP	---	---	---	---	---	---	---	---	---	---	---
<i>Populus tremul.</i>	LITDLG	FEFKYS	LEDMFT	FAGAV	ETCR	EKGLLP	---	---	---	---	---	---	---	---	---	---	---
<i>Petunia hybrida</i>	LITDMG	FQFKY	LTLEDM	FKGAI	ETCR	QKLLP	---	---	---	---	---	---	---	---	---	---	---

Code Swiss-Prot/TrEMBL: *Vitis vinifera*: P93799; *Malus domestica*: Q9XES5; *Rosa hybrida*: Q41156; *Camellia sinensis*: Q9S787; *Glycine max*: Q9SRJ5; *Arabidopsis thaliana*: Q56XQ1; *Gerbera hybrida*: F51105; *Medicago truncatula*: Q6TQ11; *Daucus carota*: Q9SP52; *Lotus japonicus* *DFR1*: Q403H0; *Lotus japonicus* *DFR2*: Q403G9; *Lotus japonicus* *DFR3*: Q403G8; *Lotus japonicus* *DFR5*: Q403G5; *Solanum tuberosum*: Q6WG03; *Medicago truncatula* 2 : Q6TQ10; *Citrus sinensis*: Q6R130; *Populus tremuloides*: Q8GV52; *Pefunia hybrida*: P14720.

I.5. Reconnaissance et fixation du substrat

Les études bio-informatiques précédentes montrent que la géométrie spatiale du site de reconnaissance du co-facteur et du site catalytique présents dans les DFRs, et de façon plus générale dans les SDRs, est conservée. Elles ne permettent par contre pas de comprendre, pour une séquence donnée, les différences d'activité rencontrées en fonction de la nature des substrats dans la littérature. Rappelons par exemple que la DFR de *Pétunia* ne réduit pas efficacement le DHK (*Johnson et al, 2001*) alors que l'enzyme DFR2 de *Lotus japonicus* le réduit préférentiellement par rapport au DHQ et au DHM (*Shimada et al, 2005*).

Une possible explication, issue de l'étude effectuée sur l'isocitrate déshydrogénase (*Mesecar et al, 1997*), autre enzyme de la famille des SDRs, est avancée dans ce dernier paragraphe. Pour cette enzyme, le changement de co-facteur entraîne une diminution drastique du pouvoir catalytique. Cette diminution résulte de l'apparition de nouvelles interactions protéine co-facteur (issues de la substitution de l'adénine par l'hypoxanthine) qui conduisent à la perte de la géométrie optimale entre les deux protagonistes réactionnels (**Figure 42**).

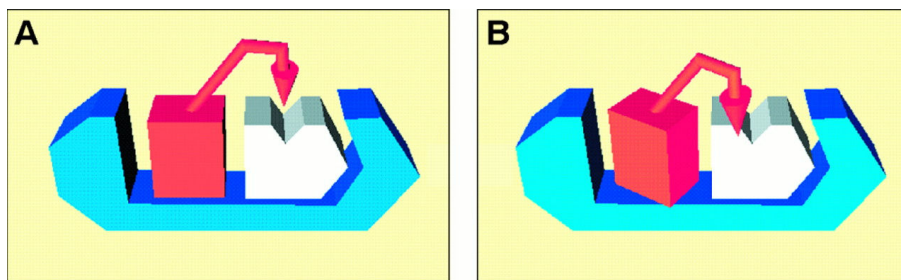


Figure 42 : Schéma représentant l'influence d'une modification du substrat dans l'alignement des deux protagonistes (*Mesecar et al, 1997*).

Le schéma A représente l'alignement optimal du NADP^+ (co-produit naturel de l'isocitrate déshydrogénase) avec son substrat. Le schéma B traduit l'influence de la substitution du co-produit par le NHDP^+ (analogue du NADP^+ où l'adénine est remplacée par l'hypoxanthine) sur sa fixation au site actif.

Il est possible d'imaginer qu'un tel résultat puisse être obtenu en gardant le co-facteur inchangé et en modifiant la nature du substrat. Après avoir caractérisé l'activité catalytique et la structure tridimensionnelle du complexe ternaire DFR- NADP^+ -DHQ, nous tenterons de vérifier ce dernier point à partir de l'étude de nouveaux complexes dans lesquels nous modifierons la nature du substrat.

II. La dihydroflavonol réductase (DFR)

II.1. Expression hétérologue de la DFR

Une construction moléculaire contenant la séquence codante de la DFR (EC : 1.1.1.219) de *Vitis vinifera*, assortie d'une étiquette histidine et d'un motif de reconnaissance par une endoprotéase (protéase Xa) permettant l'élimination de l'étiquette après purification, a été réalisée (**Figure 43**). Cette construction a été utilisée pour transformer deux souches bactériennes d'*E. coli* (X11-blue et M15[pREP4]). Deux autres constructions moléculaires contenant des mutations ponctuelles dans la séquence codante de la DFR ont également été créées. Ainsi, outre la protéine sauvage, une forme simple mutant N133D (**DFR-1C**) et une forme quadruple mutant (N133D, Q135T, Q138T et Y142I) (**DFR-4A**) ont été obtenues. Les raisons ayant motivé l'expression de ces deux mutants sont détaillées dans le chapitre II-8.



Figure 43 : Schéma de la DFR surexprimée dans la bactérie

Parallèlement, l'expression hétérologue de la DFR, sans étiquette ni motif de reconnaissance par une endoprotéase, a été entreprise chez la levure *Pichia pastoris*. Cette construction a été réalisée car elle permet d'une part la sécrétion de la protéine directement dans le milieu de culture (ce qui simplifie le protocole de purification) et d'autre part, la possible observation de modifications post-traductionnelles qui pourraient se révéler d'importance pour l'activité enzymatique de la DFR.

Néanmoins, seuls les résultats concernant l'expression hétérologue chez *E. coli* seront présentés dans ce manuscrit. L'étude de l'expression chez la levure étant encore en cours de réalisation.

II.2. Production de la DFR sauvage et conditions de purification

II.2.1. Conditions d'expression hétérologue chez *E. coli*

Pour déterminer les conditions de culture qui procurent le meilleur rendement, différents essais ont été effectués en fonction non seulement du temps et de la température d'induction mais aussi de la souche bactérienne. Les bactéries qui ont été testées présentent en effet un matériel génétique susceptible d'activer le promoteur T5 associé à la DFR. Le **Tableau 4** indique la quantité de DFR purifiée par chromatographie d'affinité en fonction des paramètres précédents. La quantification a été effectuée par la méthode de Bradford (*Bradford, 1976*).

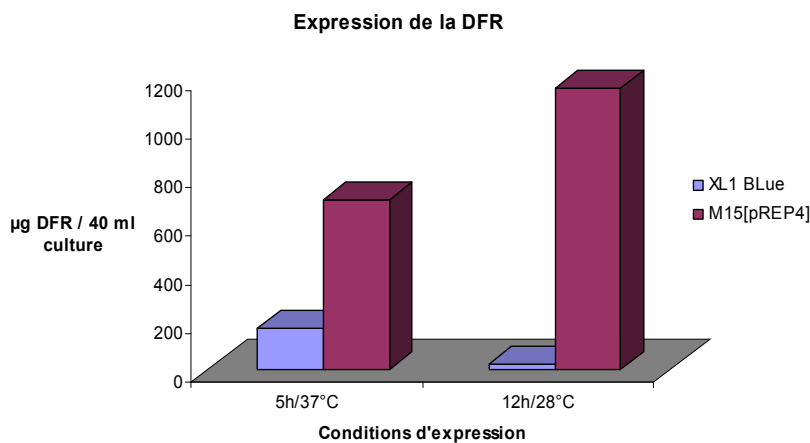


Tableau 4 : Optimisation de la surexpression de la DFR

Quelles que soient les conditions de temps et de température utilisées, les bactéries M15[pRep4] permettent une surexpression plus élevée que les bactéries XL1 Blue.

Le rendement de production dans les bactéries M15[pREP4] est inférieur de près de 40% lorsque l'expression est réalisée à la température optimale de croissance théorique de 37°C pendant 5 heures (temps au bout duquel la phase stationnaire est atteinte) par comparaison avec le rendement obtenu à plus basse température, 28°C et pour une plus longue durée, 12 heures. Par conséquent, le protocole retenu tout au long de notre travail est celui pour lequel la plus grande quantité de protéine accessible a été obtenue, c'est-à-dire dans les bactéries M15[pREP4] après 12 heures d'expression à 28°C. Il est à noter que, de par sa basse température d'induction, ce mode opératoire permet d'éviter une étape supplémentaire de renaturation de la protéine, rendue nécessaire par l'utilisation de conditions dénaturantes.

Il est probable que la différence de production observée soit due au fait qu'une température d'induction trop élevée puisse entraîner l'accumulation d'une fraction de la protéine d'intérêt dans les corps d'inclusion. La quantité de protéine étant estimée après extraction en conditions non dénaturantes, et purification par chromatographie d'affinité, la fraction de protéine présente dans les corps d'inclusions, n'est pas extraite, et n'est donc pas quantifiée.

II.2.2. Purification par chromatographie d'affinité

Les conditions de purification utilisées pour extraire du lysat bactérien la protéine d'intérêt sont celles préconisées par le fournisseur de la colonne de purification : 20 mM de tampon phosphate pH 7,4, 500 mM de NaCl et 500 mM d'imidazole. La **Figure 44** permet d'estimer après séparation par électrophorèse SDS-Page 10 % la pureté de l'échantillon élué.

Dans le lysat protéique, deux protéines majoritaires sont observées. La première est voisine de 17 kDa et n'est pas retenue sur la colonne. Il s'agit du lysozyme ajouté lors de l'étape d'extraction des protéines totales. La seconde, dont la masse moléculaire est similaire à celle attendue pour la DFR recombinante, est retenue sur la colonne par interaction avec les ions Ni^{2+} adsorbés sur la matrice, et est éluée par une forte concentration en imidazole (300 mM). Dans la fraction éluée, seule cette protéine est présente.

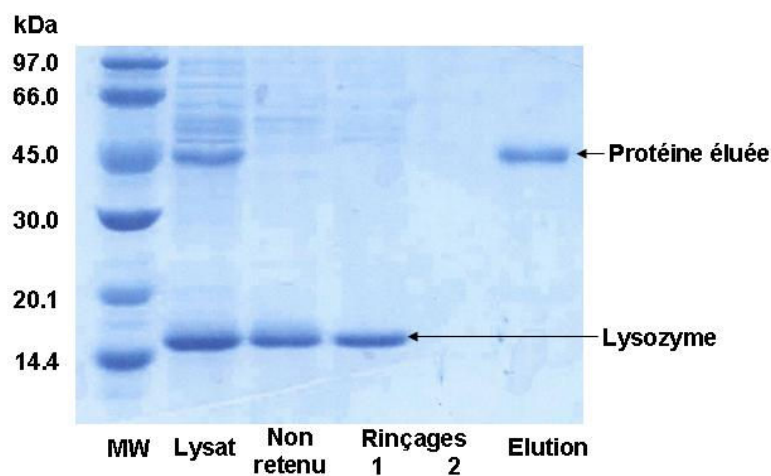


Figure 44 : Séparation par électrophorèse en conditions dénaturantes des différents échantillons obtenus lors de la chromatographie d'affinité

Les conditions utilisées pour la chromatographie d'affinité entraînent instantanément après purification, la précipitation de l'échantillon et ce malgré toute modification, de pH, l'ajout d'agents réducteurs, de chélateurs ou de détergents doux. Si après conservation à 4°C pendant 12 heures, cette fraction est analysée par chromatographie par perméation de gel (exclusion stérique), la totalité de l'échantillon est présente au volume mort de la colonne. L'échantillon est présent sous une forme oligomérisée dont la masse moléculaire est en dehors de la capacité séparatrice de la colonne utilisée (~ 600 kDa).

La **Figure 45** représente l'élugramme typique obtenu lors d'une étape de chromatographie de perméation de gel dans un tampon Tris 100 mM pH 7,4, NaCl 500 mM et imidazole 300 mM.

Les différents aliquots (A8-B7), définis sur le chromatogramme de la **Figure 45** ont été analysés par électrophorèse en conditions dénaturantes (**Figure 46**). Un même profil de séparation étant obtenu pour tous les échantillons indique que la protéine est présente sous de nombreuses formes oligomériques dans ces conditions de purification.

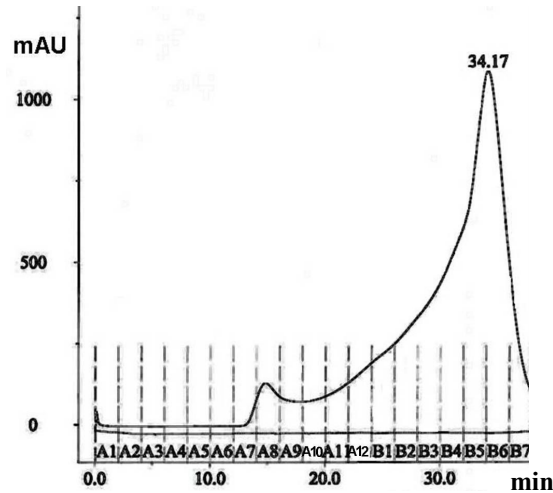


Figure 45 : Profil de séparation, par chromatographie gel perméation à 280 nm, de la fraction purifiée par chromatographie d'affinité.

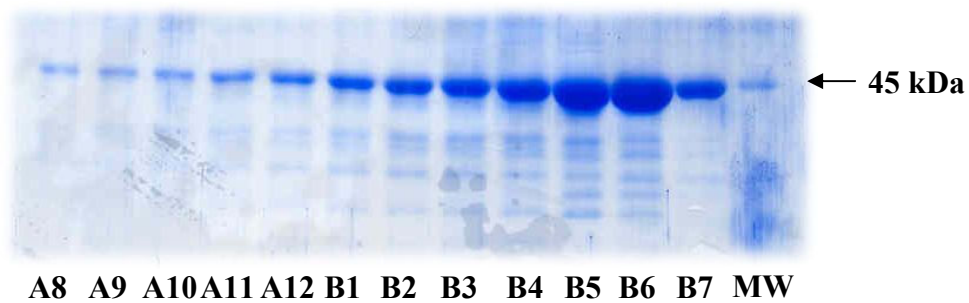


Figure 46 : Electrophorèse dénaturante des différents aliquots

II.2.3. Premières conditions de stabilité et identification de la DFR

Après de nombreux essais, une première condition de stabilité a été obtenue (**Figure 47**). La protéine est présente très majoritairement sous sa forme monomérique (temps de rétention : 32 min) si toutes les étapes de purification sont effectuées dans un tampon Tricine 10 mM pH 7,3 associé à une forte concentration en NaCl (500 mM). Une faible proportion de dimères persiste toutefois. Cette dernière est détectée pour un temps de rétention voisin de 28 min.

A ce stade du travail, et dans les conditions déterminées, la protéine reste stable pendant trois semaines. Il est à noter que seule la conjugaison de l'ensemble des bons paramètres (nature du tampon, pH, salinité) conduit à la stabilité de la protéine. Le non respect de l'un d'entre eux, annule l'effet recherché.

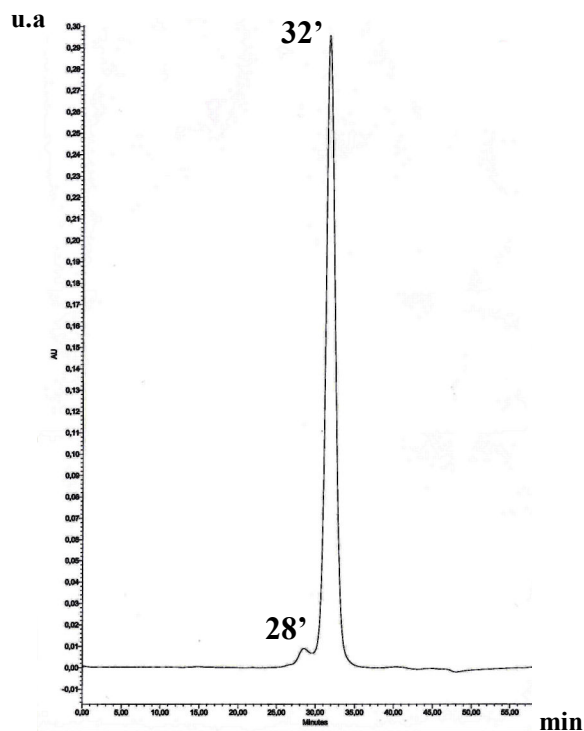


Figure 47 : Analyse par gel perméation à 280 nm de la « DFR » après deux étapes de purification

La stabilité obtenue a permis de contrôler la nature et la pureté du produit par spectrométrie de masse MALDI-TOF (**Figure 48**) et électrophorèse bidimensionnelle (**Figure 49**).

L'interprétation du spectre de masse indique que la protéine surexprimée a une masse de 39879 Da. Cette valeur, si l'on tient compte de l'incertitude expérimentale voisine de 0,1 %, confirme la présence de la DFR (masse théorique : 39847.7 Da). Le pic de masse 19965 Da correspond à l'espèce doublement chargée $[M+2H]^+/2$. Dans ce spectre, la présence d'un pic à 17828 Da met en évidence la présence d'un contaminant dans l'échantillon.

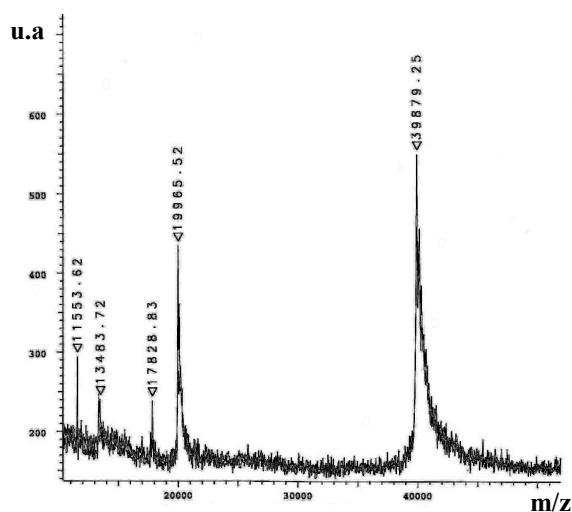


Figure 48 : Analyse qualitative par MALDI-TOF de l'échantillon purifié.

Le gel 2D (**Figure 49**), révélé au nitrate d'argent, démontre que l'impureté se trouve uniquement à l'état de traces dans l'échantillon. Le spot anormalement étendu correspond à la DFR déposée en quantité importante. Sa forme résulte du mode opératoire utilisé. En effet, le dessalage de l'échantillon, l'utilisation de tampon Tris et la variation de pH imposée lors de l'électrofocalisation sont autant de paramètres qui, nous l'avons montré, participent à la précipitation de l'enzyme.

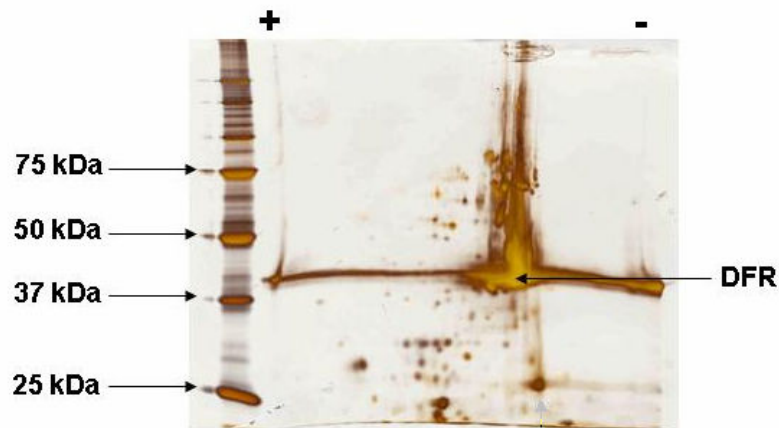


Figure 49 : Estimation de la pureté de l'échantillon par électrophorèse bidimensionnelle révélée au nitrate d'argent.

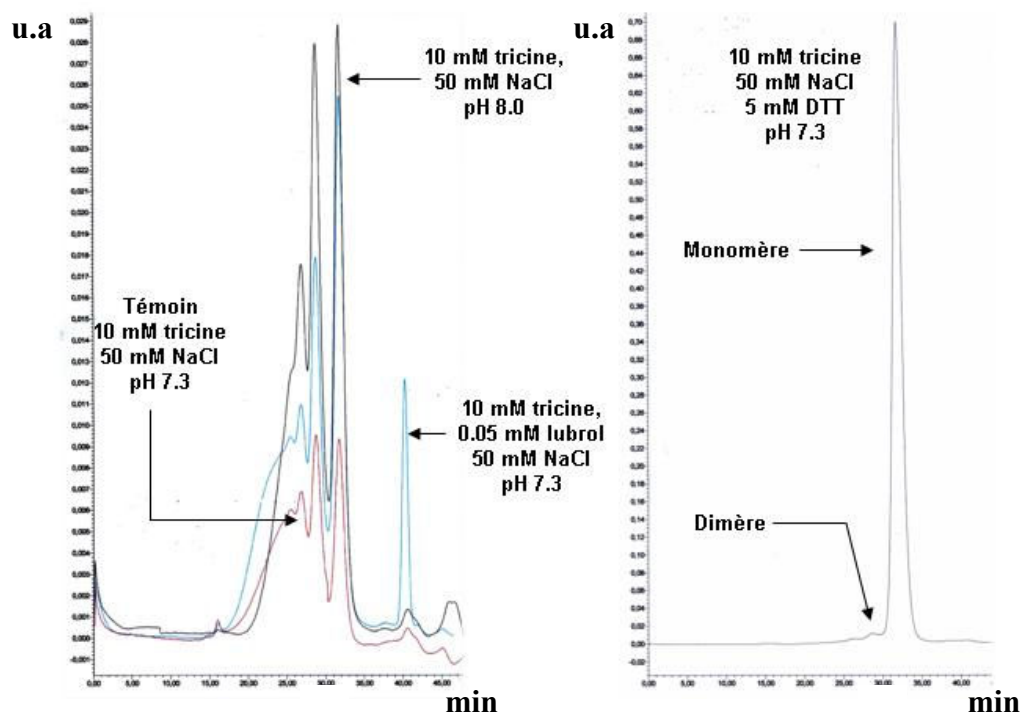
II.2.4. Optimisation des conditions de stabilité

La cristallisation, étape préalable à toute détermination de structure par diffraction des rayons X, est synonyme de l'obtention à l'état solide d'un arrangement périodique de molécules dans les trois directions de l'espace. La qualité du cristal obtenu dépend de la cinétique de croissance cristalline. Celle-ci doit, le plus souvent, être lente, une condition qui pour être remplie, nécessite une parfaite stabilité de la protéine. Cette condition est satisfaite dans notre cas. Toutefois, la concentration importante en sel (500 mM) peut être un facteur limitant la cristallisation pour deux raisons essentielles :

- ✓ la première est liée au diagramme de phase. Une forte concentration en sel peut conduire directement à l'observation de précipités avant même le temps d'équilibration de la goutte par diffusion de vapeur.
- ✓ la seconde est qu'une forte concentration en un sel donné dans la solution de protéine peut masquer l'effet produit par d'autres sels utilisés en moindre quantité dans la solution de cristallisation.

Nous avons donc recherché de nouvelles conditions de stabilité, en imposant une concentration plus faible en sel. Les tests ont été effectués avec 250 mM, 100 mM et 50 mM de sel. Les résultats étant équivalents pour ces trois concentrations, nous ne présentons ici que ceux obtenus pour la teneur en NaCl la plus faible.

Pour 50 mM de sel, aucun résultat satisfaisant n'est obtenu quels que soient les variations de pH, l'ajout d'additifs (glycérol, EDTA, ascorbate, lubrol), de substrats et/ou de coproduit (NADP⁺). Les chromatogrammes représentés **Figure 50a** en témoignent dans le cas d'un changement de pH ou de l'ajout de détergent doux. Seul l'ajout de DTT, à une concentration finale de 5 mM, permet la conservation de la forme monomérique de la DFR (**Figure 50b**).



Figures 50a et 50b : Influence de différents paramètres sur l'oligomérisation de la DFR contrôlée par chromatographie d'exclusion stérique

L'emploi de DTPA, comme autre chélateur à la place du DTT, annule l'effet observé alors que l'utilisation de β -mercaptoéthanol, agent réducteur, permet sa conservation. Ceci nous amène à penser que seule la fonction réductrice (et donc non chélatrice) du DTT est responsable de la stabilité de la DFR sous sa forme monomérique, en limitant l'oxydation des cystéines et donc la formation de ponts disulfures.

Les conditions d'obtention d'une forme stable sont identiques quelle que soit la DFR étudiée : sauvage ou mutée. Elles s'appliquent également dans le cas de la purification d'une protéine marquée aux sélénométhionines dont nous rappelons la nécessité d'obtention dans le paragraphe suivant.

II.2.5. Contrôle de l'expression de la DFR sélénée

Ne disposant pas d'une structure modèle (l'homologie de séquence entre la DFR et toute autre protéine de la famille des SDRs étant inférieure à 20 %), la méthode dite du remplacement moléculaire ne peut être utilisée pour résoudre la structure, c'est-à-dire pour obtenir une première estimation des phases. La méthode retenue est basée sur l'incorporation d'un atome à fort signal anomal (méthode de diffusion anormale à longueurs d'onde multiples). Elle consiste à substituer lors de la culture bactérienne les méthionines de chaque sous unité de l'enzyme par des séléno-méthionines. Pour un phasage efficace, le taux de substitution doit être le plus élevé possible. Pour ce faire, nous avons utilisé la méthode de marquage décrite dans le chapitre « Matériel et Méthodes ».

Afin d'estimer le taux d'insertion des atomes de sélénium, les deux formes de DFR, marquée et non marquée, ont été soumises à un clivage protéolytique (trypsine). Les produits de digestion ont été séparés par microchromatographie sur colonne C18 (Zip-Tip). Chaque peptide a été identifié par sa masse (MALDI-TOF), et le taux de marquage a pu être déterminé pour toutes les méthionines détectées. Le taux d'incorporation de séléno-méthionines est de 100 % pour tous les peptides de la forme marquée.

A titre d'exemple, les spectres trouvés pour le peptide [327-341] sont reportés sur la **Figure 51**. La présence de sélénium est vérifiée par la signature isotopique caractéristique de cet atome, distincte de celle du soufre.

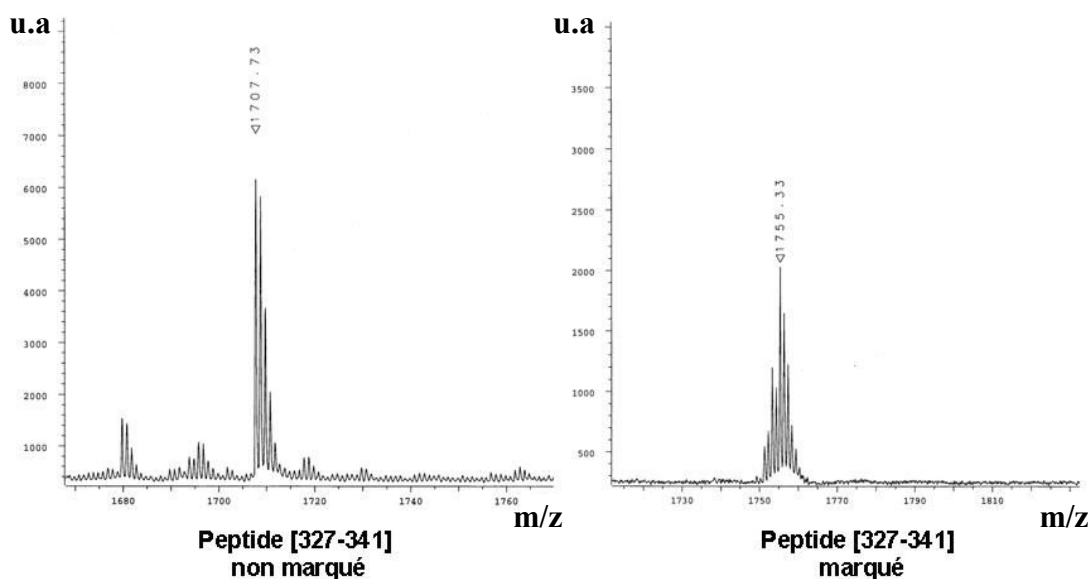


Figure 51 : Contrôle de l'incorporation de la séléno-méthionine au sein du peptide marqué [327-341] par spectrométrie de masse

Nous verrons lors de la résolution de la structure de la DFR, que seules les positions atomiques de 23 atomes de sélénium ont pu être déterminées sur les 26 composant l'unité asymétrique.

II.3. Caractérisation enzymatique

II.3.1. Rappel de la méthode de dosage

En présence du cofacteur nicotinique NADPH, l'enzyme catalyse la réduction des dihydroflavonols en leucoanthocyanidines. Ces produits de réaction sont incolores et ne peuvent être séparés des substrats par méthode HPLC en phase inverse sur colonne C18. Les leucoanthocyanidines peuvent néanmoins être oxydées en anthocyanidines par simple chauffage en milieu organique. A pH acide, les anthocyanidines (cyanidine, pélargonidine ou delphinidine en fonction du nombre et de la position des fonctions hydroxyles sur le cycle B) sont colorées et peuvent être quantifiées (voir chapitre « Matériel et Méthodes »). La **Figure 52** résume les différents composés métabolisés au cours de ces réactions.

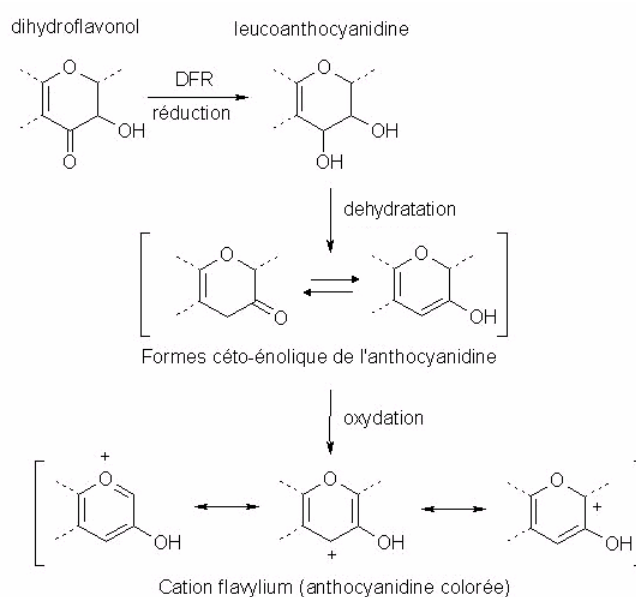


Figure 52 : Oxydation en milieu acide des leucoanthocyanidines formées

Pour déterminer la nature et estimer la quantité des pigments obtenus, un couplage HPLC-spectrométrie de masse (LC-MS) a été effectué à 214 nm.

II.3.2. Etude LC-MS des témoins réactionnels

L'analyse des spectres obtenus en présence de DHQ (*Fluka, Réf: 78666*) et de tous les réactifs, excepté l'enzyme, montre deux pics principaux. Le premier, majoritaire, observé pour un temps de rétention de 11,89 min, est attribué à la DHQ (**Figure 53-A**). Le second, de moindre surface, traduit vraisemblablement la présence d'un second produit que nous n'avons pas réussi à identifier. Le pic correspondant (temps de rétention de 7,40 min) n'est pas observé lors de l'analyse d'un échantillon de DHQ commerciale sans autre réactif et n'ayant subi aucune étape de chauffage ni d'acidification, mais réapparaît après traitement de l'échantillon. Ceci tend à montrer qu'il s'agit d'un produit de dégradation.

Les spectres obtenus en présence de DHK (*Arbonova, Ref: 480-20-6*) possèdent un profil de LC-MS identique. Le pic majoritaire (13,18 min) (**Figure 53-C**) st celui du DHK. Son « produit de dégradation » est élué à 7,34 min.

II.3.3. Etude LC-MS des essais enzymatiques

La DFR associée au mélange réactionnel provoque l'apparition d'une couleur fuchsia ou rouge lorsque les substrats utilisés sont respectivement la DHQ (**2**) et le DHK (**2'**). Les chromatogrammes (**Figures 53-B/D**) indiquent que la réaction n'est que partielle dans les conditions expérimentales utilisées puisque la présence du pic associé au substrat persiste. Obtenir une réaction plus complète nécessiterait l'acidification du milieu réactionnel qui favoriserait la réduction du substrat, mais une telle condition demeure incompatible avec la stabilité de la DFR.

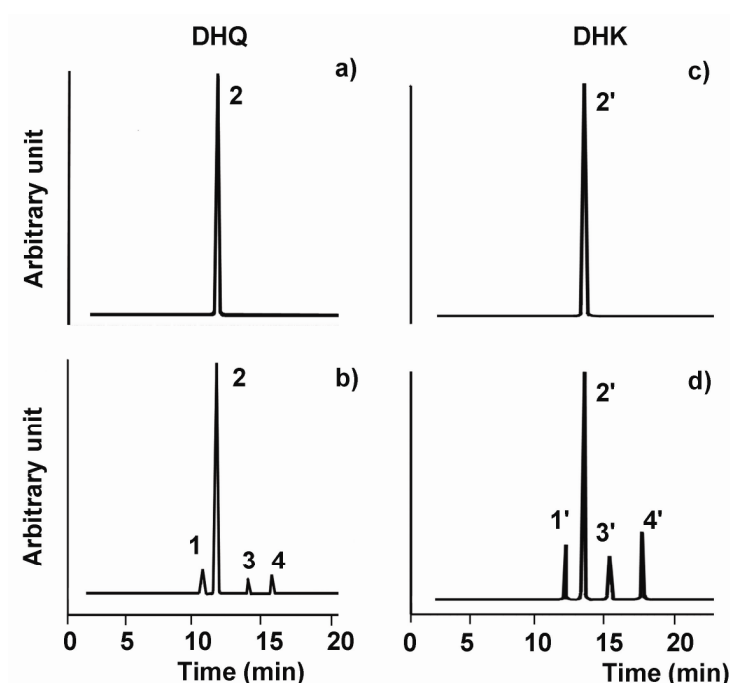


Figure 53 : Représentation schématique des chromatogrammes obtenus à 214 nm
Les pics correspondant aux produits de dégradation des substrats ne sont pas mentionnés.

Trois produits apparaissent majoritairement (**Figure 53-B**):

- ✓ la cyanidine (**1**) (MM : 287 Da)
- ✓ la 3-O-butyl-cyanidine (**3**) (MM : 343 Da) : dont la nature a été déterminée par fragmentation (LC-MS/MS)
- ✓ un dérivé du 3-O-butyl-cyanidine (**4**) comportant deux unités de masse supplémentaire (MM : 345 Da)

La différence de masse existante entre les produits (1) et (2) traduit une réaction d'éthérisation de la cyanidine par le butanol et libération d'une molécule d'eau. Nous avons confirmé cette hypothèse en utilisant deux alcools différents lors de l'étape d'oxydation. Lors de l'utilisation de 1-propanol, la différence de masse observée est celle d'un radical propyl associé à la cyanidine (Δ masse = 43 Da). L'emploi de tertio-butanol entraîne une forte réduction de la surface des pics correspondant aux dérivés éthérisés. Ceci montre que l'éthérisation s'effectue sur la fonction hydroxyle en position 3 du cycle C, compte tenu de l'encombrement stérique du tertio-butanol par rapport aux deux autres alcools primaires employés.

Un comportement similaire (**Figure 53-D**) est observé si le substrat utilisé est le DHK. Les trois produits sont respectivement :

- ✓ la pélargonidine (**1'**) (MM : 271 Da)
- ✓ la 3-O-butyl-pélargonidine (**3'**) (MM : 327 Da)
- ✓ un dérivé du 3-O-butyl-pélargonidine (**4'**) (MM : 329 Da) comportant deux unités de masse supplémentaires.

Le tableau ci-dessous (**Tableau 5**) résume, en fonction du substrat et de l'alcool utilisés, les temps de rétention et les masses molaires des différents composés mis en évidence.

Composé	Masse molaire (Da)	Temps de rétention (min)
DHQ	304	11,88
cyanidine	287	10,92
3-O-butyl-cyanidine	343	14,14
3-O-butyl-cyanidine +2	345	15,89
3-O-propyl-cyanidine	329	13,24
3-O-propyl-cyanidine +2	331	14,69
3-O-tbutyl-cyanidine	343	14,00
3-O-tbutyl-cyanidine +2	345	non détecté

Composé	Masse molaire (Da)	Temps de rétention (min)
DHK	289	13,18
pélargonidine	271	11,74
3-O-butyl-pélargonidine	327	15,00
3-O-butyl-pélargonidine +2	329	17,54
3-O-propyl-pélargonidine	313	14,11
3-O-propyl-pélargonidine +2	315	non détecté
3-O-tbutyl-pélargonidine	327	15,00
3-O-tbutyl-pélargonidine +2	329	17,50

Tableau 5 : Caractéristiques des composés mis en évidence

II.3.4. Méthode de quantification

Si l'on se rapporte aux protocoles expérimentaux publiés dans la littérature (et rappelés dans le chapitre « Matériel et Méthodes »), l'activité de la DFR est quantifiée par spectrométrie UV-visible à 550 nm, longueur d'onde correspondant à l'absorption maximale de la cyanidine. Un tel protocole ne peut être employé s'il on veut comparer l'activité de notre enzyme vis-à-vis des deux substrats étudiés. En effet, la pélargonidine absorbe à 534 nm, longueur d'onde significativement différente de celle de la cyanidine (**Figure 54**). Il résulterait du calcul une sous estimation de la concentration en pélargonidine. En outre, une estimation de l'activité de la DFR basée sur la seule prise en compte de son produit direct (cyanidine ou pélargonidine) ne peut être effectuée compte tenu de la présence de produits secondaires (produits d'éthérisation) dont certains, par ailleurs, ne sont pas colorés. Seule l'utilisation du tertio-butanol réduisant de façon importante, nous l'avons vu, l'apparition des produits éthérisés pourrait être envisagée si n'apparaissait pas systématiquement un produit non identifié (18,95 min) et ce, quel que soit le substrat utilisé.

Afin de nous affranchir des difficultés mentionnées précédemment, nous avons donc choisi d'estimer l'activité de la DFR par quantification des substrats (et non des produits). La longueur d'onde choisie est de 214 nm. La surface du pic supplémentaire mentionné lors de l'étude des échantillons témoins n'étant pas affectée par l'ajout de DFR, celui-ci ne peut être associé à un dérivé de substrat métabolisable. Ce pic sera donc négligé lors du calcul des consommations de substrats que nous avons choisi comme méthode de quantification.

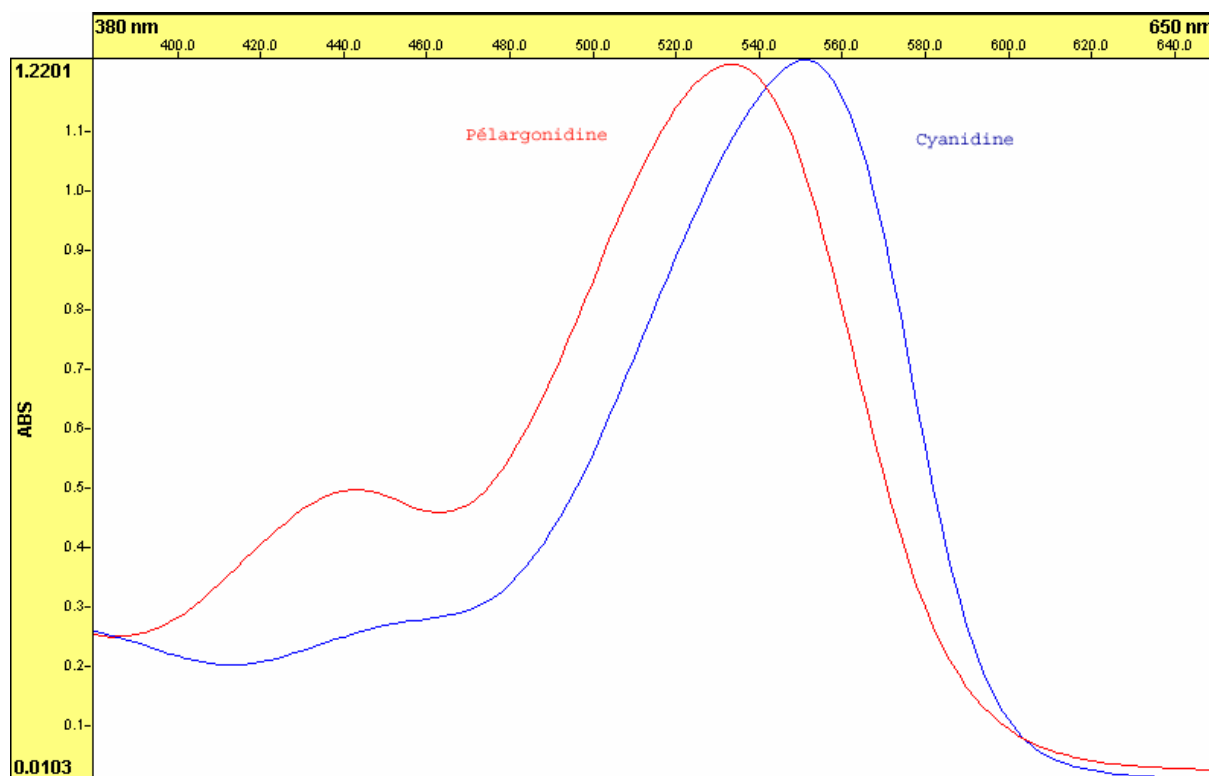


Figure 54 : Spectres de la « cyanidine » et de la « pélargonidine »

II.3.5. Activité catalytique et affinité de substrats

La **Figure 55** illustre l'influence des mutations et de la nature du substrat sur la coloration des échantillons. Le **Tableau 6** quantifie les différences observées si l'on utilisait les protocoles publiés en considérant comme seul produit présent la cyanidine ou la pélargonidine. Le **Tableau 7**, quant à lui, chiffre la proportion de substrat non réduit par l'enzyme.

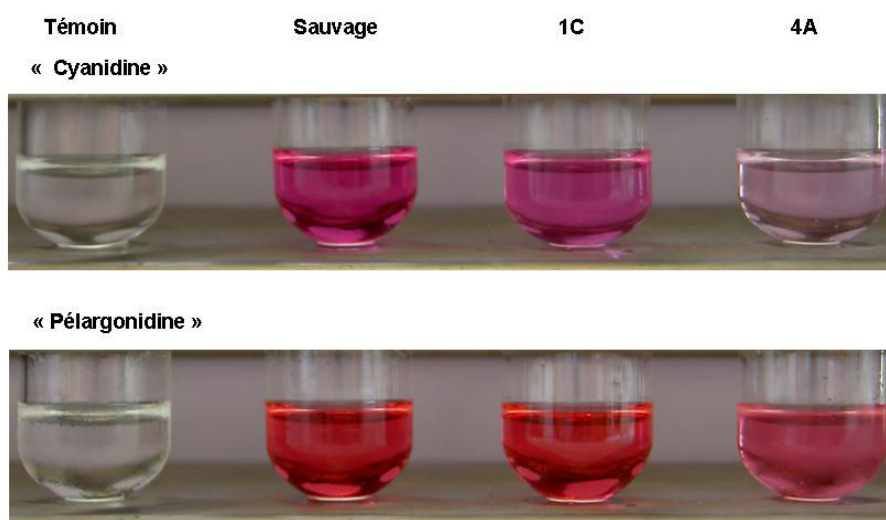


Figure 55 : Relations entre couleur de l'échantillon, substrat et nature de la DFR

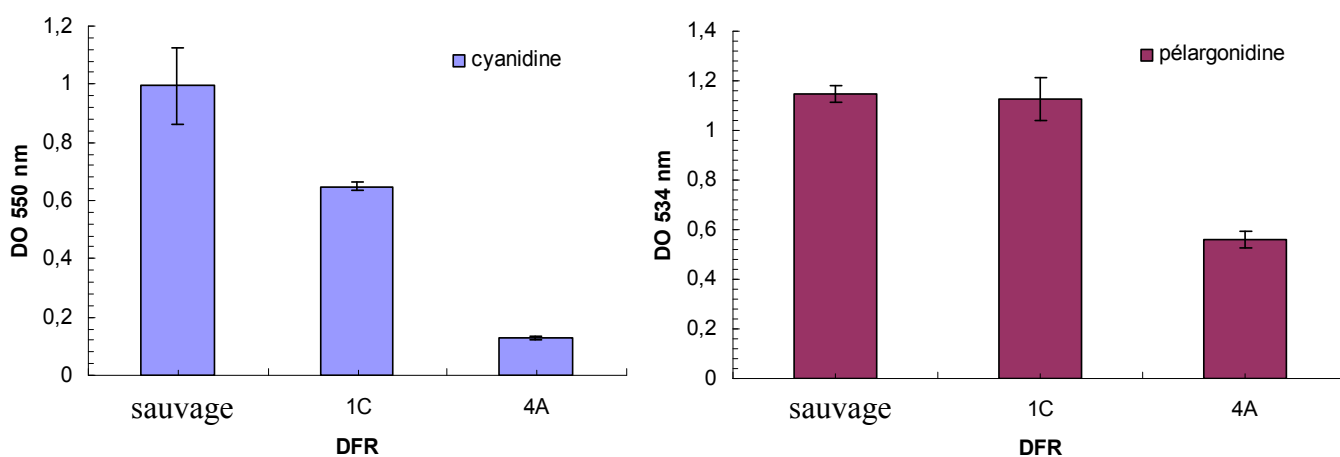


Tableau 6 : Quantification des pigments à 550 nm et 534 nm

Les écarts types sur les valeurs numériques rapportées sont calculés à partir d'expériences menées en quadruplet.

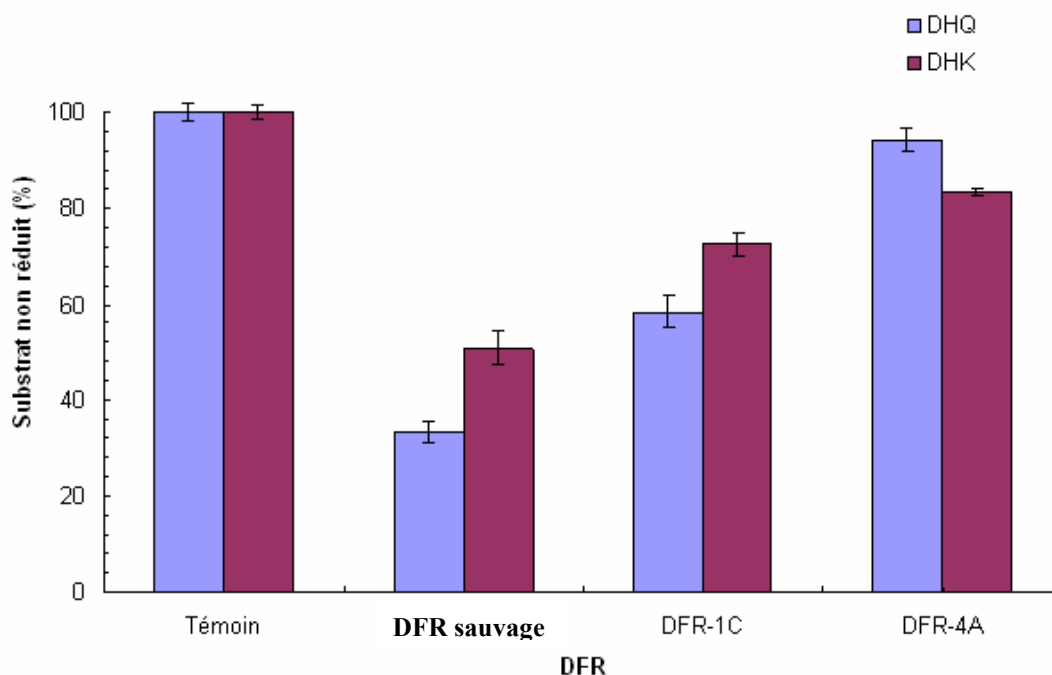


Tableau 7 : Pourcentage de substrat résiduel

A la vue des données du **Tableau 6**, la mutation 1C ne semble affecter que la consommation de DHQ, un résultat qui est contredit par les données du **Tableau 7**. Ce dernier montre que l'activité est en réalité diminuée vis-à-vis des deux substrats. Elle reste toutefois moindre dans le cas du DHK. En ce qui concerne le quadruple mutant, les intensités de coloration ne sont pas contradictoires avec la consommation de substrat. L'activité est néanmoins encore diminuée par rapport à celle mesurée pour le simple mutant.

Le **Tableau 8** reporte les consommations en substrat. Ces dernières traduisent simplement les pourcentages de substrat réduit.

	DHQ réduit (%)	DHK réduit (%)	Affinité relative DHQr / DHKr
DFR sauvage	66.8	49.3	1.35
DFR-1C	41.6	27.3	1.52
DFR-4A	5.68	16.4	0.35
Activité relative de la DFR-1C (%WT)	62.3	55.5	
Activité relative de la DFR-4A (%WT)	8.50	33.3	

Tableau 8 : Activités enzymatiques et affinités relatives

La mutation de l'asparagine 133 en aspartate (**DFR-1C**) est associée à une diminution relative de l'activité catalytique de l'ordre de 40 %. Toutefois, cette seule mutation n'affecte pas profondément l'affinité relative DHQ/DHK de l'enzyme.

L'activité du quadruple mutant (N133D, Q135T, Q138T et Y142I) est par contre particulièrement affectée, et ceci notamment pour la DHQ.

Ces derniers résultats (mutant **4A**), comparés à ceux du simple mutant, tendent à montrer l'importance des mutations 135, 138 et 142. Ces trois mutations affectent notablement l'affinité de substrat. La DFR réduit alors préférentiellement le DHK.

Il serait intéressant de préciser les différentes constantes michaeliennes (K_m , V_m) de la réaction enzymatique ainsi que l'influence de l'ordre de fixation des substrats. Pour cela, la méthode la plus simple aurait été de suivre l'oxydation du NADPH à 340 nm, longueur d'onde maximale d'absorption de ce co-facteur. Ce protocole expérimental est néanmoins difficilement applicable, les différents substrats et produits absorbant dans le même domaine de longueurs d'onde. Pour y parvenir, nous avons donc choisi, comme méthode d'analyse, le suivi de la fluorescence du NADPH.

II.3.6. Détermination d'un protocole expérimental par fluorescence

II.3.6.1. Synthèse de leucoanthocyanidines

La détermination des différents paramètres de fluorescence (longueurs d'onde d'excitation et d'émission, gamme de concentrations, largeur d'ouverture des fentes) nécessite la parfaite connaissance des différents constituants du mélange réactionnel. Or, nous ne disposons pas du produit de réaction, la leucoanthocyanidine, composé instable. Il nous a donc fallu le synthétiser chimiquement. La réaction choisie est la réduction totale dans l'éthanol absolu de la DHQ par le borohydrure de sodium (NaBH_4). Cette réaction a été suivie par spectrophotométrie UV-visible en incrémentant le ratio $[\text{NaBH}_4] / [\text{DHQ}]$. Les résultats sont présentés **Figure 56**.

La présence de points isobestiques à 255 et 305 nm traduit la réduction directe de la DHQ en leucoanthocyanidine (absence d'intermédiaires réactionnels). Les spectres obtenus montrent que la leucoanthocyanidine absorbe de façon spécifique à 330 nm. L'absorbance est maximale pour 1,5 équivalents de NaBH_4 , et décroît à plus forte concentration. Cette décroissance est probablement induite soit par dégradation soit par modification chimique de la leucoanthocyanidine.

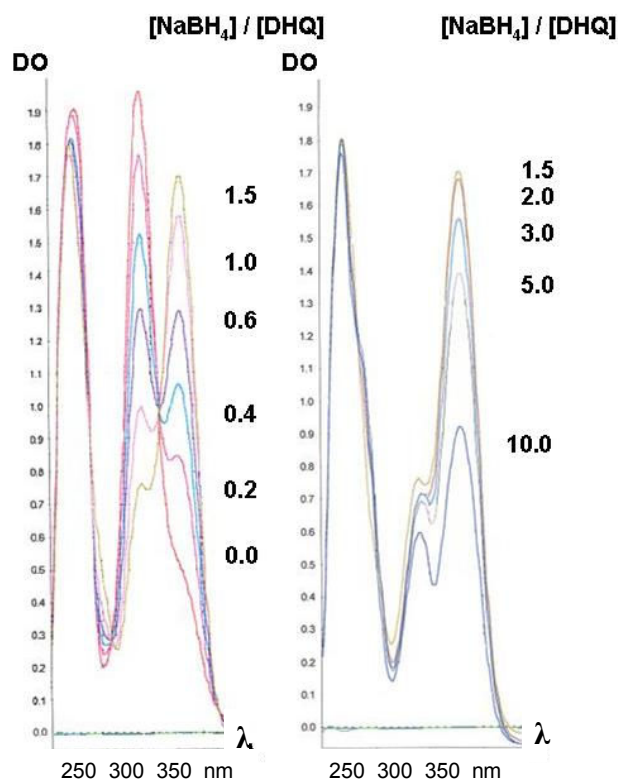


Figure 56 : Suivi de la réduction de la DHQ par le NABH₄ par spectrométrie UV

Pour confirmer l'obtention de leucoanthocyanidine, le produit de réaction a été oxydé par chauffage en milieu acide. Il en résulte l'apparition d'un produit de coloration rouge, la cyanidine ou l'un de ses dérivés. L'intensité de coloration des différents échantillons a été mesurée par spectrophotométrie à 550 nm (**Figure 57**). Chaque point du graphique a été mesuré quatre fois.

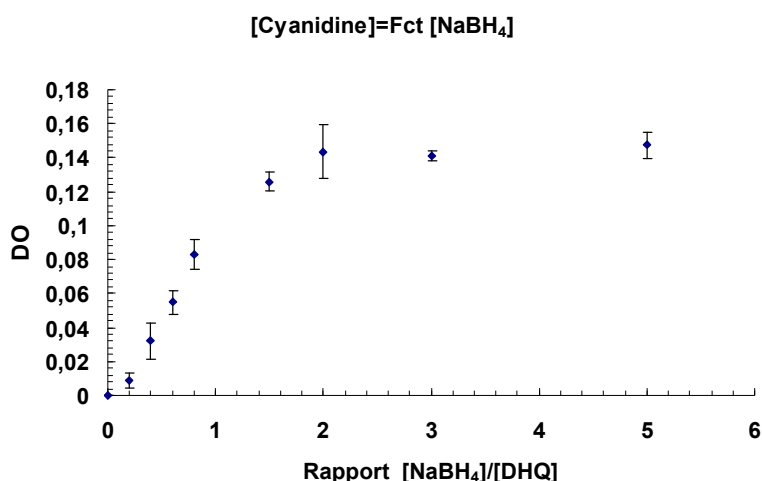


Figure 57 : Influence du nombre d'équivalents de NABH₄ sur la formation de pigments

La courbe obtenue décrit une augmentation quasi linéaire pour un rapport de concentration compris entre 0 et 1,5. Pour un rapport supérieur, la coloration atteint une valeur plateau, la DHQ est alors totalement réduite. Une preuve supplémentaire aurait été de suivre la réduction de la fonction carbonyle par spectrométrie infrarouge. Ceci n'a pas été possible compte tenu de la faible solubilité des dihydroflavonols et du volume limité de l'échantillon à étudier.

II.3.6.2. Longueurs d'onde d'excitation et d'émission

Tous les essais, en présence ou non d'enzyme, ont été effectués à une température thermostatée de 30°C. Les λ_{\max} d'excitation et d'émission du NADPH sont respectivement de 340 et 450 nm. Dans ces conditions, la fluorescence est linéaire jusqu'à une concentration maximale en fluorophore de 15 μM . Cette concentration est supposée trop faible pour calculer l'influence de la concentration en NADPH sur la vitesse enzymatique de la DFR. Aussi avons-nous recherché une autre longueur d'onde d'excitation.

Pour toute longueur d'onde d'excitation plus courte que 340 nm, la fluorescence est soumise à plusieurs effets tels que l'extinction dynamique (quenching proportionnel aux teneurs en DHQ et leucocyanidine) ou l'influence du tampon d'activité. Ce dernier doit satisfaire deux impératifs : être compatible avec la viabilité de l'enzyme, et jouer son rôle à pH acide afin que la réaction ait lieu. A pH alcalin, la fonction hydroxyle en position 7 du cycle A est déprotonée pour conduire à un phénolate, qui est stabilisé par délocalisation de la charge négative jusqu'au carbonyle du cycle B (**Figure 58**). Le caractère électrophile du carbone C4 est alors fortement diminué et rend l'attaque nucléophile du NADPH impossible. Parmi les tampons susceptibles d'être utilisés dans une gamme de pH comprise entre 5,8 et 7,2, seul le Bis-Tris permet d'observer une réduction par l'enzyme. La composition du « polytampon » retenue est donc : 50 mM Tricine, 50 mM Bis-Tris, 50 mM NaCl et 1 mM DTT. Néanmoins, l'utilisation de longueurs d'onde d'excitation inférieures à λ_{\max} s'est avérée rapidement délicate de par la présence de quenching dynamique, fonction de plusieurs paramètres.

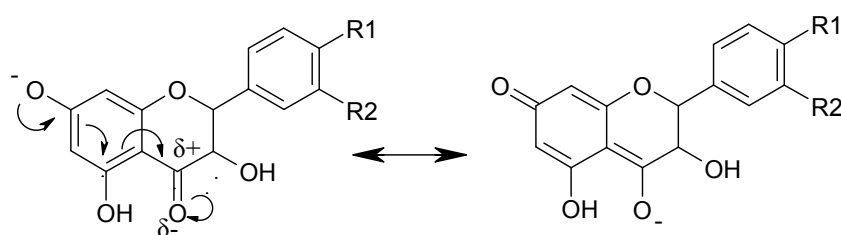


Figure 58 : Formation de dérivés (1,6) semiquinoniques à pH alcalin

Des tests de linéarité de fluorescence ont donc été effectués avec une longueur d'onde d'excitation comprise entre 373 et 390 nm. L'ensemble des résultats trouvés est décrit dans la **Figure 59**. La concentration maximale en NADPH applicable aux différents essais enzymatiques est de 50 μM , valeur très largement supérieure à la limite préalablement trouvée (15 μM à λ_{ex} de 340 nm).

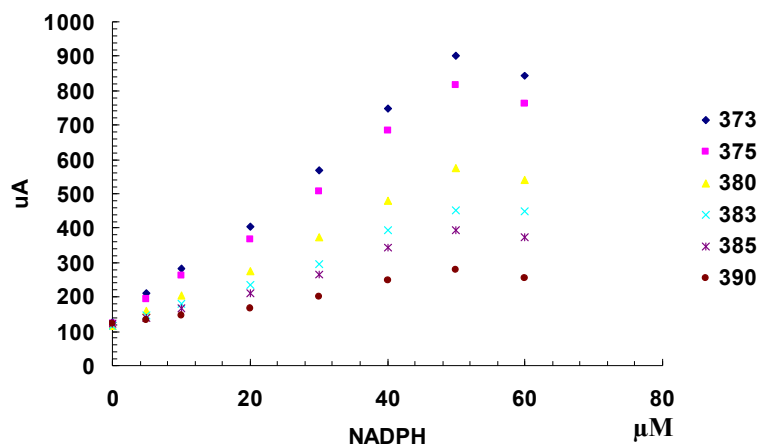


Figure 59 : Fluorescence du NADPH en fonction de sa concentration et pour différentes longueur d'onde d'excitation
Les fentes des monochromateurs sont de 2,5 et 10,5 nm.

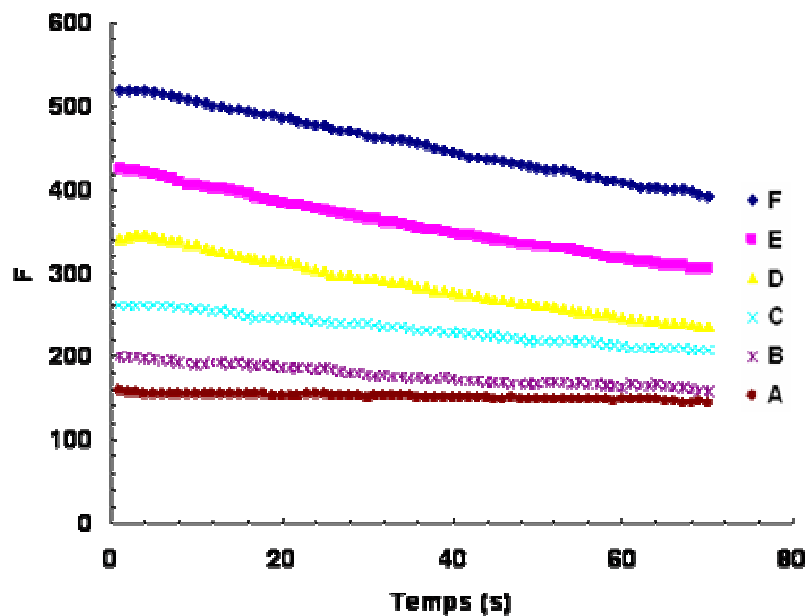
II.3.6.3. Influences de la DHQ et de la leucocyanidine dans les conditions expérimentales déterminées

L'influence de la concentration tant en substrat qu'en produit a été étudiée pour diverses concentrations en NADPH comprises entre 0 et 50 μM et pour trois longueurs d'onde excitatrices différentes (373, 375 et 380 nm). Ces mesures nous ont permis de montrer l'absence de toute extinction dynamique associée à une concentration de ces composés comprise entre 0 et 90 μM .

L'obtention d'une concentration supérieure en NADPH permettant d'observer une linéarité de fluorescence, et l'absence d'effet quancheur nous ont conduit à retenir comme longueur d'onde excitatrice une valeur de λ supérieure à 340 nm. Nous avons retenu comme valeur expérimentale : 380 nm.

II.3.7. Résultats enzymatiques obtenus

Un exemple de courbe de consommation du NADPH est présenté sur la **Figure 60** pour différentes concentrations en NADPH variant de 5 à 50 μM et une concentration constante en DHQ (90 μM). Dans tous les cas, la courbe de fluorescence en fonction du temps témoigne de la consommation du NADPH.



CINETIQUE	DHQ (μM)	NADPH (μM)	DFR (nM)
A	90	5	5
B	90	10	25
C	90	20	25
D	90	30	50
E	90	40	50
F	90	50	50

Figure 60 : Cinétiques mesurées pour diverses concentrations de NADPH

Le **Tableau 9** ci-dessous représente les valeurs calculées de la vitesse initiale de la réaction. Pour obtenir ces valeurs, les pentes mesurées à l'origine ont été corrigées par le facteur de dilution de l'enzyme.

Cinétique	A	B	C	D	E	F
Pente réelle	-1,559	-1,130	-1,690	-1,6742	-1,7959	-1,9291

Tableau 9 : Vitesses initiales mesurées

Une généralisation de ce travail consiste à faire varier séparément les concentrations en substrat ou en co-facteur. Les valeurs des vitesses initiales calculées sont représentées **Figure 61**.

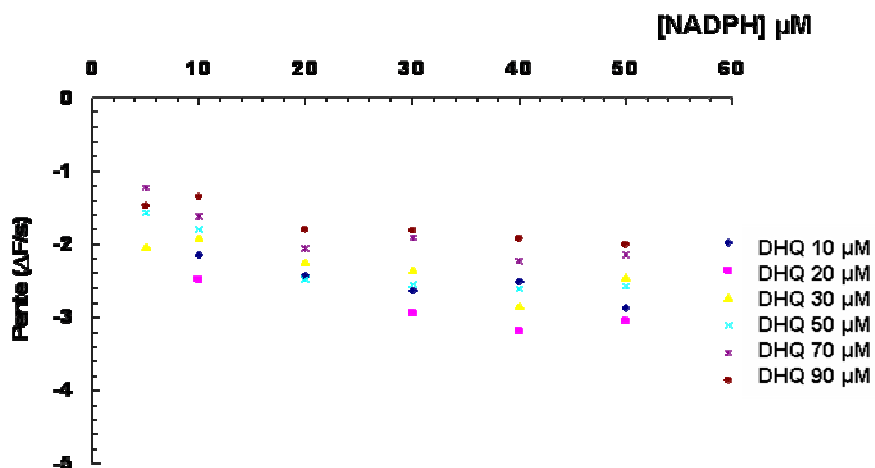


Figure 61a : Vitesse initiale fonction de la concentration en NADPH (μM) pour plusieurs teneurs en DHQ (10 à 90 μM).

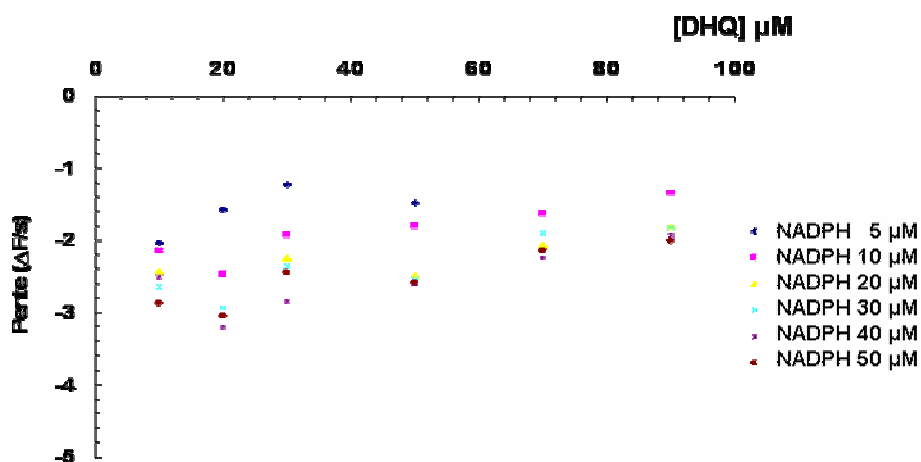


Figure 61b : Vitesse initiale fonction de la concentration en DHQ (μM) pour plusieurs teneurs en NADPH (5 à 50 μM).

Bien que n'ayant pas estimé la reproductibilité des mesures, les vitesses initiales ($\sim 14 \mu\text{M}\cdot\text{min}^{-1}$) apparaissent indépendantes de la concentration en substrat et en co-facteur. Ces données suggèrent une réaction d'ordre 0 vis-à-vis des concentrations en NADPH et DHQ, résultat improbable. Une possible explication peut être tirée des travaux de **Nelsestuen et al (1997)** qui montrent qu'une analyse michaelienne est à exclure si l'enzyme n'est pas diluée de façon homogène dans la solution. La concentration en substrat non lié à l'enzyme n'est alors pas uniforme dans la solution. Dans ce cas précis, la protéine agrégée fait chuter fortement la concentration du substrat dans son voisinage immédiat d'où une cinétique limitée uniquement par le flux de substrat.

Un contrôle effectué par la suite par GPC tend à montrer que le polytampon provoquerait l'agrégation de l'enzyme, lorsque celle-ci est fortement diluée (100 nM). L'influence du BisTris avait pourtant été estimée préalablement aux essais enzymatiques, mais uniquement dans le cas d'une concentration notable en enzyme (1 mM). Des contrôles supplémentaires, détaillés dans la discussion, semblent, par ce résultat, visiblement nécessaires afin de mieux appréhender ce phénomène.

Nous nous sommes interrogés par la suite sur l'influence de l'étiquette « Histidine » dans le phénomène d'agrégation de l'enzyme à pH acide. En effet, la DFR n'est pas stable à ce pH, condition pourtant essentielle à la réduction enzymatique des dihydroflavonols. Les cycles imidazoles de l'étiquette sont des ions imidazolium qui apportent à l'enzyme six charges supplémentaires. Il est possible que ces charges localisées ponctuellement en partie N-terminale puissent affecter la stabilité de l'enzyme. Afin de tester cette hypothèse, nous avons donc décidé de cliver l'étiquette histidine (21 acides aminés - **Figure 43**). Toutes les tentatives se sont soldées par un échec. En effet, les conditions expérimentales d'activité de la protéase Xa sont en tout point incompatibles avec celles nécessaires à la stabilité de la DFR (pH, natures du tampon et des sels).

Il semble donc que, dans les conditions expérimentales retenues, (étiquette « Histidine » - pH acide), la détermination des constantes enzymatiques soit difficilement réalisable. Pour s'affranchir de l'étiquette, tout en ayant un protocole de purification simple, nous tentons d'exprimer la DFR par un organisme eucaryote : la levure *Pichia pastoris*.

II.4. Cristallisation et qualité de diffraction

II.4.1. Cristallisation de la DFR sauvage

La recherche des conditions de cristallisation s'est effectuée en utilisant différents kits de cristallisation commerciaux (Hampton Research). Chaque kit contient 48 ou 96 conditions différentes. Nous avons testé les kits suivants : Crystal screen 1 et 2, Crystal Lite, PEG ion, Wizard 1 et 2, Index et MembFac. Dans chaque cas, nous avons utilisé :

- ✓ soit la DFR seule,
- ✓ soit le complexe avec cinq équivalents de NADP^+ **ou** cinq équivalents de DHQ (complexe binaire),
- ✓ soit le complexe avec cinq équivalents de NADP^+ **et** cinq équivalents de DHQ (complexe ternaire).

La technique utilisée est celle de la goutte suspendue. Les conditions expérimentales sont rappelées dans le chapitre « Matériel et Méthodes ». Seule la présence du substrat et du co-produit a permis d'observer les premiers cristaux. Les conditions d'obtention sont reportées dans le **Tableau 10** ci-dessous.

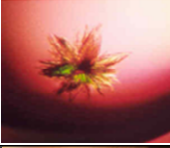
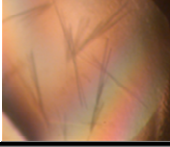
Screening	DFR	Condition de cristallisation	Forme cristalline	Photo
Crystal screen 2	NaCl 50 mM DTT 5 mM Tricine 10 mM pH = 7,3 DFR = 4,5 mg/ml	KNa-tartrate 200 mM (NH ₄) ₂ SO ₄ 2M Na-citrate 100 mM pH = 5,6	Polycristaux Assemblage de losanges	
Index	NaCl 50 mM DTT 5 mM Tricine 10 mM pH = 7,3 DFR = 4,5 mg/ml	NaCl 200 mM PEG 3350 25% Hepes 100 mM pH = 7,5	Aiguilles	

Tableau 10 : Premières conditions de cristallisation

L'optimisation des conditions de cristallisation a été conduite pour les deux formes cristallines obtenues (polycristaux et aiguilles). Dans le premier cas, nous n'avons pu nous affranchir de l'agrégation des cristaux. Pour cette forme, la recherche de conditions optimales a été abandonnée, la seconde ayant rapidement donné des résultats encourageants. Typiquement, les cristaux apparaissent au bout de 48 heures et atteignent leur taille optimum dans les deux semaines suivantes. Au-delà de cette période, on observe une dégradation des cristaux, ce qui impose leur congélation afin de les conserver.

Le **Tableau 11** récapitule l'ensemble des conditions testées lors de l'optimisation des aiguilles.

DFR	Paramètres constants du réservoir	Paramètres variables du réservoir
DFR 8 mg/ml	Hepes 100 mM pH 7.5	PEG (4000, 3000,3350) 19 à 29% NaCl 50 à 200mM
DFR 8 mg/ml	Hepes 100 mM pH 7.5 NaCl 200mM	PEG (2000, 3000, 3350) 25 à 35%
DFR 8 mg/ml	Hepes 100 mM pH 7.5 NaCl 80 mM	PEG (3350-3000) 28-1% à 17-12%
DFR 15 à 50 mg/ml	Hepes 100 mM pH 7.5	PEG 3350 19 à 30% NaCl 40 à 140 mM
DFR 15 mg/ml	Hepes 100 mM pH 7.5 NaCl 100mM	PEG 3350 27 à 30%
DFR 13.4 mg/ml	Hepes 100 mM pH 7.5 « Sel » 110mM	NaCl, NaBr, KBr, NaI PEG 3350 32.5 à 36%

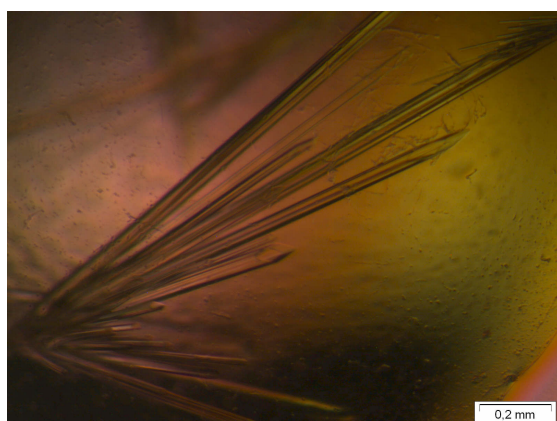
Tableau 11 : Paramètres testés lors de l'optimisation

La DFR est toujours associée à cinq équivalents de substrat et de co-facteur à l'exception de la 5^{ème} phase d'optimisation où cette valeur varie entre 2 et 15.

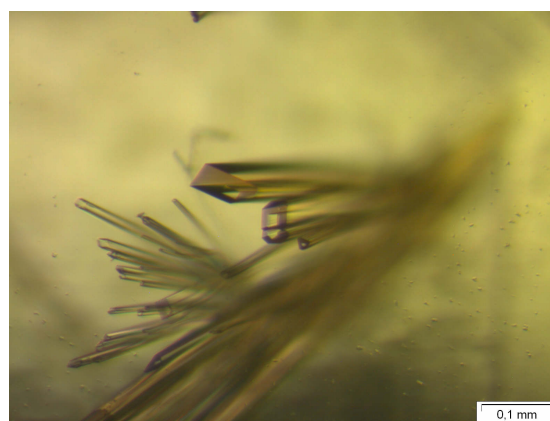
Les conditions de cristallisation ayant permis l'obtention des meilleurs cristaux (**Figure 62**), tant de la DFR sauvage que sélénée, sont reportées ci-dessous :

Substrat	Cofacteur	Concentration en DFR	Réservoir	Résolution (Å)
DHQ / EtOH	NADP ⁺	15 mg/ml	NaCl 90 mM PEG 3350 24 % Hepes 100 mM pH 7,5	1,8
<i>DHQ / EtOH</i>	<i>NADP⁺</i>	<i>15.5 mg/ml DFR sélénée</i>	<i>NaCl 100 mM PEG 3350 24.5 % Hepes 100 mM pH 6,9</i>	<i>2,4</i>

Ces cristaux ont été testés tant au laboratoire que lors de nos différents passages sur rayonnement synchrotron (ESRF). Ils diffractent respectivement à 1,8Å (DFR sauvage) et 2,4Å (DFR sélénée).



DFR sauvage (L = 1.27 mm ; l = 0.15 mm)



DFR-Sélénée (L = 0,5 mm ; l = 0.08 mm)

Figure 62 : Photographies des cristaux de DFR sauvage et DFR-Sélénée

II.4.2. Co-cristallisation de la DFR avec différents complexes (substrat-cofacteur)

Afin d'étudier l'influence, sur l'affinité, des fonctions hydroxyles présentes en différentes positions des cycles B et C de chaque substrat, deux démarches complémentaires ont été exploitées. D'une part nous avons co-cristallisé la DFR avec différents complexes. D'autre part, nous avons entrepris la co-cristallisation de DFR mutées en présence des complexes DHQ/NADP⁺ ou DHK/NADP⁺. Dans ce dernier cas, nous sommes parvenus à ce jour à cristalliser uniquement le quadruple mutant, et ce, sous deux faciés distincts : losanges et aiguilles. Dans un souci de regrouper l'ensemble des données de cristallisation dans un même paragraphe, nous mentionons ci-dessous pour chaque cas les conditions expérimentales retenues ainsi que la meilleure résolution cristalline obtenue. L'absence de photographies de cristaux pour une condition donnée traduit un échec de cristallisation.

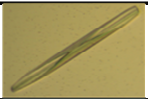


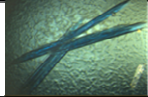
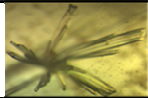
Substrat	Co-facteur	Composition du réservoir	Dimensions (mm)	Résolution maximale (Å)	Cristal
Dihydroquercetine	NADPH	NaCl 70 mM PEG 3350 27% Hepes 100 mM pH = 7,2	0,90 x 0,06	2,0	
Quercetine	NADP	NaCl 50 mM PEG 3350 29% Hepes 100 mM pH = 6,9	0,15 x 0,06	2,8	
Dihydrokaempferol	NADP	-	-	-	-
Kaempferol	NADP	-	-	-	-
Myricetine	NADP	NaCl 50 mM PEG 3350 29% Hepes 100 mM pH = 6,9	0,40 x 0,25	2,0	
Luteolin	NADP	-	-	-	-
Eriodictyol	NADPH	NaCl 100 mM PEG 3350 31% Hepes 100 mM pH = 7,2	0,42 x 0,07	2,5	
Eriodictyol	NADP	NaCl 60 mM PEG 3350 30% Hepes 100 mM pH = 7,2	0,33 x 0,08	1,9	
Apigenine	NADP	-	-	-	-
Chrysine	NADP	-	-	-	-

Tableau 12 : Essais de co-cristallisation effectués

Tous les tests ont été effectués avec une solution de DFR sauvage à 15 mg/ml dissoute dans 10 mM Tricine pH 7,3, 50 mM NaCl, 5 mM DTT et 3 mM NaN₃.

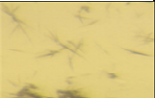
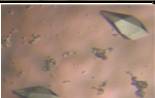

Substrat	Co-facteur	Composition du réservoir	Dimensions (mm)	Résolution maximale (Å)	Cristal
Dihydroquercetine	NADP	NaCl 100 mM PEG 3350 21% Tris 100mM pH 8,0	-	-	
Dihydroquercetine	NADP	NaCl 200 mM PEG 3350 25% Bis-Tris 100 mM pH = 6,5	0,17 x 0,06	3,6	
Dihydrokaempferol	NADP	LiSO ₄ 200 mM PEG 4000 30% Tris 100mM pH 8,5	-	-	

Tableau 13 : Cristaux de mutants 4A obtenus

Les résultats obtenus pour deux des conditions précédentes conduisent à des micro-cristaux dont la qualité cristalline n'a pas encore été déterminée.

II.5. Etudes structurales

Une première tentative de détermination de structure a été entreprise dès lors que nous avons obtenu les premiers cristaux de DFR sauvage diffractant à une résolution acceptable. Dans notre cas, celle-ci était de 2,5Å. La technique dite du remplacement moléculaire a été employée. Le modèle moléculaire utilisé correspond aux deux cents premiers acides aminés de la carbonyl réductase, de *Sporobolomyces salmonicolor*, enzyme dont la structure a été résolue en 2005 et présentant une homologie de séquence de 26% avec la DFR (**Kamitori et al, 2005**), Aucune solution n'a néanmoins été mise en évidence avec ce modèle. Nous n'avons pas persévéré dans cette voie, l'échec étant pratiquement attendu. Rappelons en effet les premiers résultats issus de l'analyse bioinformatique qui montraient la faible homologie de séquence (19 % maximum) entre les membres de la famille des déshydrogénases à courtes chaînes de structure connues (en 2003) et la DFR. Nous avons donc très vite opté pour l'utilisation de la méthode de diffusion anormale multi-longueurs d'onde et l'expression d'une protéine « sélénée ».

II.5.1. Collecte et traitement de données de la protéine sélénée

Les collectes de données à différentes longueurs d'onde ont été obtenues sur la ligne de lumière BM30A (ESRF-Grenoble). Le choix des longueurs d'onde a été déduit de l'analyse d'un spectre de fluorescence X du cristal étudié, enregistré au voisinage du seuil d'absorption K du sélénium. L'utilisation du programme CHOOCH (**Evan et al, 2001** ; <http://www.gwyndafevans.co.uk/>) permet de déduire les valeurs f'' (composante anormale du facteur de diffusion) et f' (composante de dispersion du facteur de diffusion) en fonction de la longueur d'onde (**Figure 63**). La première longueur d'onde expérimentale choisie ($\lambda_1 = 0,98074\text{\AA}$) est déduite de la valeur de l'énergie pour laquelle la composante f'' est maximale ($E_1 = 12642\text{eV}$). La seconde ($\lambda_2 = 0,98089\text{\AA}$) est déduite de la valeur de l'énergie ($E_2 = 12640\text{eV}$) pour laquelle la valeur de f' est minimale. Enfin, une troisième longueur d'onde de collecte est retenue. Cette dernière ($\lambda_3 = 0,97319\text{\AA}$) correspond à une énergie éloignée du seuil d'absorption du sélénium ($E_3 = 12740\text{eV}$), pour laquelle il n'y a plus de variation notable de f' et f'' .

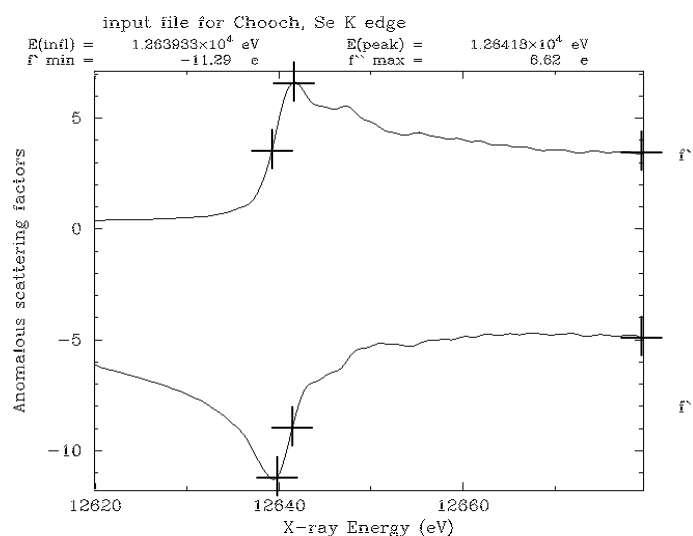


Figure 63 : Coefficients f' et f'' déduits du spectre de fluorescence du sélénium en fonction de la longueur d'onde

Les collectes à ces trois longueurs d'onde ont été effectuées sur le même cristal qui diffracte à une résolution maximale de 2,369Å. La distance cristal-détecteur est de 185 mm. Les conditions expérimentales de collecte pour 90 images différentes sont les suivantes :

- ✓ oscillation ($\Delta\phi$) : 1° par image,
- ✓ temps d'exposition : 45 s par image.

La méthode de collecte choisie est telle que les données de diffraction aux trois longueurs d'onde ont été mesurées successivement pour un même domaine angulaire de rotation, ceci afin de minimiser l'influence d'une dégradation du cristal due à l'irradiation.

Le traitement des données (Mosflm : *Leslie et al, 2001*) ; <http://www.mrc-lmb.cam.ac.uk/harry/mosflm/>) a permis l'inventaire des éléments de symétrie du cristal et par conséquent, la mise en évidence de son appartenance au système cristallin orthorhombique. Les paramètres de maille déterminés sont les suivants :

$$a = 88,06 \text{ \AA} \quad b = 89,71 \text{ \AA} \quad c = 93,29 \text{ \AA} \quad \alpha = 90^\circ \quad \beta = 90^\circ \quad \gamma = 90^\circ$$

L'utilisation du programme Scala (CCP4) (*Evans et al, 1993*) permet, pour une collecte donnée, une mise à l'échelle de l'ensemble des intensités diffractées afin de minimiser la décroissance de l'intensité diffractée en fonction du temps d'irradiation. Elle permet également dans le cas d'une collecte multiple, une mise à l'échelle des intensités entre jeux de données. Les intensités de réflexions équivalentes ainsi calculées sont ensuite moyennées. Dans notre cas, afin de ne pas perdre le signal anomal du facteur de diffusion, nous considérons comme différentes les réflexions d'une même paire de Friedel ($I_{hkl} \neq I_{-h-k-l}$). Les intensités trouvées sont regroupées par domaine de résolution. Les **Tableaux 14** ($\lambda_1, \lambda_2, \lambda_3$) résument pour chaque domaine, mais aussi pour la totalité du jeu de donnée, les statistiques d'intégration les plus pertinentes. Parmi celles-ci, retenons par exemple :

- ✓ différents facteurs de confiance (Rsym, Rfull et Ranomal),
- ✓ complétude et multiplicité,
- ✓ rapport $I / \sigma(I)$.

Dès lors que plusieurs réflexions équivalentes ont été mesurées, on peut traduire la dispersion des mesures obtenues pour toutes les réflexions équivalentes par un facteur de confiance global appelé Rsym.

$$R_{sym}(I) = \frac{\sum_{hkl} \sum_i |I_i(hkl) - \overline{I(hkl)}|}{\sum_{hkl} \sum_i I_i(hkl)}$$

Toutes les réflexions, que celles-ci aient traversé totalement ou non la sphère d'Ewald, interviennent dans le calcul de Rsym. Seules les réflexions du premier type sont prises en considération dans le calcul de Rfull. Pour chaque jeu de données, Rsym est égal à Rfull avec un écart inférieur à 10 %, ce qui rend compte d'une bonne intégration des données et donc une estimation correcte de la mosaïcité.

Le facteur de confiance anomal (Ranom) traduit la différence d'intensité entre les paires de Friedel. Une valeur du facteur de confiance Ranom supérieure à celle de Rsym signifie donc une contribution significative du diffuseur anomal. L'examen du **Tableau 14** montre par exemple que, pour le premier jeu de données, le signal anomal n'est conséquent que pour une résolution inférieure à 2,78 Å. A plus haute résolution, l'incertitude de la mesure influe de manière prépondérante dans la différence d'intensité entre paires de Friedel.

$$R_{\text{anomalous}} = \frac{\sum (\langle I^+ \rangle - \langle I^- \rangle)}{\sum (\langle I^+ \rangle + \langle I^- \rangle)}$$

La complétude traduit le rapport entre le nombre de réflexions mesurées et le nombre de réflexions théoriquement mesurables. Cette valeur doit rester proche de 100 %. Ceci est vrai quelque soit le jeu de données. Néanmoins, la complétude sur les seules données anormales (AnoCmpl) décroît très fortement dans la dernière couronne de résolution (2,33-2,45Å). Elle avoisine 30 à 40 % selon le jeu de données, valeur très insuffisante. Il en va de même de la multiplicité anormale (AnoMlt) dont la valeur est proche de l'unité, valeur signifiant l'absence d'observations multiples des réflexions d'une même paire de Friedel. Cette dernière condition traduit une incertitude de mesure plus élevée pour ces réflexions qui, pour cette raison, pourraient être exclues du jeu de données.

Le rapport signal sur bruit (Mn(I)/sd) n'est pas un paramètre limitant. Il est toujours supérieur à 2 (valeur minimale requise) dans tout le domaine de résolution.

En conclusion, la seule valeur limitante reste pour ces trois jeux de données, la comparaison entre facteurs de confiance qui conduit à suspecter la validité du signal anomal pour les hautes résolutions.

λ_1	N	1/d ²	Dmin(A)	Rsym	Rfull	Ranom	Nanom	Mn(I)/sd	Nmeas	C _{poss}	M _{plct}	AnoCmpl	AnoMlt
1	0.0184	7.36	0.035	0.033	0.087	668	34.8	3119	88.0	3.4	83.7	2.0	
2	0.0369	5.21	0.040	0.037	0.082	1584	34.1	7301	95.7	3.9	99.6	2.1	
3	0.0553	4.25	0.033	0.027	0.056	2107	34.5	9494	97.6	4.0	99.8	2.1	
4	0.0737	3.68	0.040	0.034	0.056	2535	32.1	11212	98.4	4.1	99.9	2.1	
5	0.0922	3.29	0.050	0.036	0.059	2900	26.9	12681	98.9	4.1	100.0	2.1	
6	0.1106	3.01	0.055	0.046	0.067	3208	21.4	13874	99.1	4.1	100.0	2.1	
7	0.1291	2.78	0.069	0.059	0.081	3537	16.2	15186	99.3	4.1	100.0	2.1	
8	0.1475	2.60	0.108	0.092	0.101	3780	11.5	16048	99.4	4.1	100.0	2.1	
9	0.1659	2.45	0.134	0.114	0.120	3786	7.9	14001	99.5	3.6	93.2	1.9	
10	0.1843	2.33	0.139	0.112	0.149	1737	4.5	1610	94.8	1.8	40.6	1.1	
overall			0.048	0.039	0.068	25842	19.7	104526	94.8	3.8	89.7	2.0	

λ_2

	N	1/d ²	Dmin(A)	Rsym	Rfull	Ranom	Nanom	Mn(I)/sd	Nneas	C _{poss}	M _{plct}	AnoC _{mpl}	AnoM _t
1	0.0185	7.36	0.034	0.031	0.062	671	35.0	3108	87.6	3.4	84.1	2.0	
2	0.0369	5.20	0.039	0.035	0.060	1584	34.3	7292	95.6	3.9	99.2	2.1	
3	0.0554	4.25	0.034	0.028	0.043	2109	34.8	9530	97.5	4.0	99.5	2.1	
4	0.0739	3.68	0.039	0.032	0.043	2534	32.3	11219	98.4	4.1	99.9	2.1	
5	0.0923	3.29	0.050	0.035	0.046	2912	26.7	12715	98.8	4.1	99.9	2.1	
6	0.1108	3.00	0.059	0.052	0.054	3219	20.9	13925	99.1	4.1	100.0	2.1	
7	0.1293	2.78	0.074	0.064	0.066	3544	15.5	15203	99.3	4.1	100.0	2.1	
8	0.1477	2.60	0.117	0.094	0.088	3785	10.8	16044	99.4	4.1	100.0	2.1	
9	0.1662	2.45	0.142	0.120	0.112	3794	7.3	13970	99.4	3.6	93.3	1.9	
10	0.1847	2.33	0.155	0.128	0.148	1736	4.2	1624	94.7	1.8	40.4	1.1	
overall			0.049	0.040	0.054	25888	19.3	104630	94.7	3.8	89.7	2.0	

 λ_3

	N	1/d ²	Dmin(A)	Rsym	Rfull	Ranom	Nanom	Mn(I)/sd	Nneas	C _{poss}	M _{plct}	AnoC _{mpl}	AnoM _t
1	0.0188	7.30	0.033	0.032	0.050	734	35.9	3307	91.4	3.4	89.4	2.0	
2	0.0375	5.16	0.044	0.037	0.052	1635	33.9	7495	96.9	3.9	99.9	2.1	
3	0.0563	4.21	0.037	0.033	0.037	2177	33.5	9714	98.3	4.0	100.0	2.1	
4	0.0751	3.65	0.046	0.029	0.037	2604	30.3	11426	99.9	4.0	99.9	2.1	
5	0.0939	3.26	0.056	0.047	0.042	2970	24.4	12826	99.2	4.1	100.0	2.1	
6	0.1126	2.98	0.065	0.066	0.052	3312	18.2	14198	99.4	4.1	100.0	2.1	
7	0.1314	2.76	0.083	0.083	0.064	3619	13.1	15374	99.5	4.1	100.0	2.1	
8	0.1502	2.58	0.137	0.132	0.095	3903	9.0	16378	99.6	4.1	100.0	2.1	
9	0.1690	2.43	0.172	0.173	0.124	3882	6.0	14028	99.6	3.6	93.1	1.9	
10	0.1877	2.33	0.309	0.523	0.170	1684	3.3	1503	94.0	1.0	30.4	1.1	
overall			0.054	0.043	0.049	26520	17.7	106249	94.8	3.7	89.6	2.0	

Tableaux 14 : Statistiques obtenues pour les trois collectes

II.5.2. Détermination du groupe d'espace

L'examen des extinctions selon les rangées principales de la maille permet la détermination du groupe d'espace dans lequel cristallise le composé. La **Figure 64** montre que les intensités diffractées d'indice de Miller impaire sont nulles, ce qui traduit la présence d'un axe hélicoïdal (2_1) dans les directions a et b de la maille cristalline.

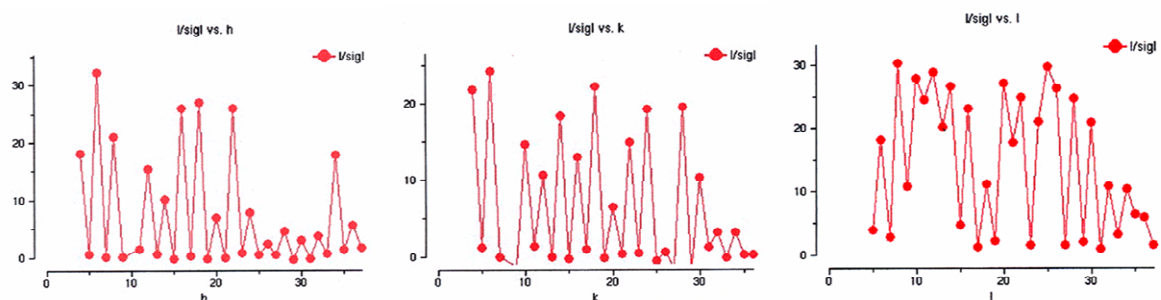


Figure 64 : Analyse des intensités diffractées dans les trois directions principales de la maille

Les réflexions d'indices $\ell = 11, 13, 23, 25$ présentant des valeurs $I / \sigma(I)$ positives, le groupe d'espace doit être $P2_12_12$. La détermination des positions des atomes de sélénium n'a néanmoins pu être menée à bien que dans le groupe d'espace $P2_12_12_1$. Il en a été de même de la résolution de la structure. Le groupe d'espace $P2_12_12_1$ a donc été retenu sans que nous puissions réellement donner d'explication à la contradiction trouvée. Pour un tel groupe, un calcul du coefficient de Matthews (2,6 pour 49,5 % de solvant) (*Matthews et al, 1968*) indique que deux molécules de DFR constituent l'unité asymétrique. L'orientation de l'une par rapport à l'autre est par la suite déterminée au moyen de la fonction de « self rotation ». Cette dernière met en évidence la présence d'un axe non cristallographique d'ordre 2 parallèle à la direction $\vec{a} + \vec{b}$.

II.5.3. Positions atomiques des atomes de sélénium et premières cartes de densité électronique

Shelx (*Sheldrick et al, 2002* ; <http://shelx.uni-ac.gwdg.de/SHELX/>) est une chaîne de programmes dévolu à la résolution et l'affinement de structures cristallographiques. Trois de ces programmes, (Shelx C, D et E) sont utilisés dans le cas des macromolécules biologiques.

ShelxC analyse les jeux de données enregistrés à différentes longueurs d'onde et génère un fichier global utilisé par la suite par les autres programmes. Il permet notamment le calcul d'un coefficient de corrélation (CC) qui traduit la qualité des jeux de données en estimant, en fonction de la résolution, les différences anormales trouvées. Dans notre cas, ce coefficient chute au dessous d'une valeur communément admise (CC = 30 %) pour une résolution supérieure à 2,8 Å. La diffusion anormale n'est plus significative au-delà de cette résolution. Les réflexions impliquées sont donc exclues des jeux de données. Ce résultat confirme l'analyse faite précédemment lors de la discussion sur la valeur relative des facteurs de confiance R_{sym} et R_{anom} .

L'utilisation des méthodes directes et l'interprétation des cartes de Patterson (ShelxD) permettent la localisation des atomes lourds. A partir des positions atomiques trouvées, ShelxE fait une estimation des phases, calcule les premières cartes de densité électronique et améliore les phases trouvées grâce à de premiers calculs de modification de densité de type « aplatissement de solvant ».

Parmi les solutions déterminées par ShelxD, une solution se dégage plus particulièrement. Elle se caractérise par une valeur élevée d'une figure de mérite associée à la bonne superposition des fonctions de Patterson calculées. Les positions et les facteurs d'occupation des atomes de sélénium associés à cette solution sont rapportés dans le tableau ci-dessous (**Tableau 15**). Parmi ces positions, seules les 23 premières sont conservées pour un premier calcul des densités électroniques compte tenu d'une valeur significative de leur facteur d'occupation.

Cette valeur est à rapprocher du nombre d'atomes de sélénium effectivement contenus dans l'unité asymétrique. Treize séléno-méthionines étant présentes dans une molécule et la valeur du coefficient de Matthews indiquant que deux molécules composent l'unité asymétrique, seuls trois atomes de sélénium n'ont donc pas pu être positionnés.

					X	Y	Z	F _{oc}	B
HETATM	1	S	HAT	1	87.444	61.114	6.928	1.000	20.00
HETATM	2	S	HAT	2	93.133	67.046	-28.934	0.980	20.00
HETATM	3	S	HAT	3	65.324	74.084	-33.214	0.972	20.00
HETATM	4	S	HAT	4	74.331	70.399	13.733	0.954	20.00
HETATM	5	S	HAT	5	67.144	68.760	1.115	0.912	20.00
HETATM	6	S	HAT	6	83.221	79.628	-36.332	0.895	20.00
HETATM	7	S	HAT	7	84.722	86.907	-23.641	0.872	20.00
HETATM	8	S	HAT	8	79.061	88.745	10.999	0.846	20.00
HETATM	9	S	HAT	9	88.062	61.611	-44.654	0.764	20.00
HETATM	10	S	HAT	10	92.421	66.300	23.149	0.745	20.00
HETATM	11	S	HAT	11	68.499	78.183	1.763	0.738	20.00
HETATM	12	S	HAT	12	75.243	84.886	-24.241	0.710	20.00
HETATM	13	S	HAT	13	96.202	76.845	-39.499	0.602	20.00
HETATM	14	S	HAT	14	64.041	71.374	11.900	0.577	20.00
HETATM	15	S	HAT	15	77.133	57.978	17.652	0.553	20.00
HETATM	16	S	HAT	16	102.489	68.550	-27.774	0.548	20.00
HETATM	17	S	HAT	17	70.878	75.232	-33.710	0.510	20.00
HETATM	18	S	HAT	18	86.244	51.953	6.300	0.504	20.00
HETATM	19	S	HAT	19	89.319	74.302	-42.868	0.474	20.00
HETATM	20	S	HAT	20	79.955	64.954	20.940	0.450	20.00
HETATM	21	S	HAT	21	100.913	67.202	-25.715	0.418	20.00
HETATM	22	S	HAT	22	78.093	83.990	11.728	0.417	20.00
HETATM	23	S	HAT	23	82.494	89.449	-33.406	0.390	20.00
HETATM	24	S	HAT	24	69.695	58.850	-10.021	0.190	20.00
HETATM	25	S	HAT	25	70.616	62.666	-20.952	0.183	20.00
HETATM	26	S	HAT	26	84.717	66.189	-10.184	0.163	20.00
HETATM	27	S	HAT	27	113.661	76.857	-32.105	0.159	20.00
HETATM	28	S	HAT	28	101.342	40.737	7.615	0.101	20.00
HETATM	29	S	HAT	29	90.203	69.665	-12.848	0.095	20.00
HETATM	30	S	HAT	30	72.284	67.313	-28.873	0.089	20.00
HETATM	31	S	HAT	31	84.894	92.954	19.438	0.021	20.00

Tableau 15 : Positions des atomes de sélénium

II.5.4. Calcul de la première carte de densité et choix de l'énantiomère

La qualité des cartes de densité calculées par ShelxE est estimée par une valeur du contraste qui doit croître en fonction des cycles d'affinement (**Figure 65**). Un contraste élevé est la signature d'une parfaite localisation des molécules au sein de la maille cristalline. Sa valeur est calculée pour la solution déterminée et son inverse. La courbe de contraste la plus élevée permet la sélection de la structure recherchée.

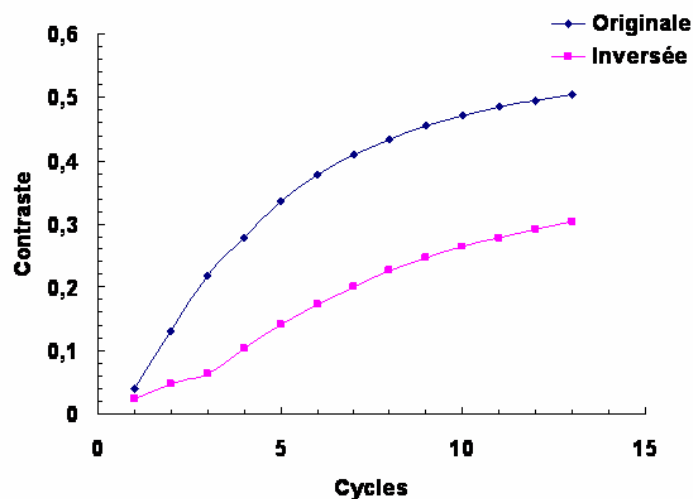


Figure 65 : Evolution de la valeur du contraste pour la solution directe et son inverse

II.5.5. Redéfinition de la qualité des cartes de densité par les programmes Solve / Resolve

La Figure 66 montre la faible qualité des cartes issues de ShelxE. Seules certaines positions atomiques semblent définies. La carte obtenue ne permet pas la visualisation de la chaîne principale. Toutefois, prise dans sa globalité, elle met en évidence l'organisation des molécules au sein de l'unité asymétrique et la présence d'un axe de symétrie non cristallographique d'ordre 2.

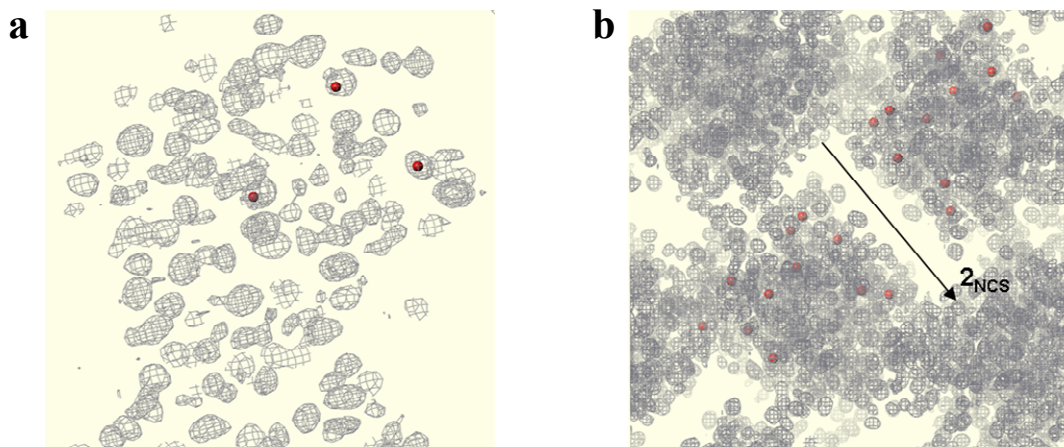


Figure 66 : Cartes électroniques de type $(F_o.e^{i\phi^{calc}})$ obtenues par le programme ShelxE

Les sphères rouges matérialisent la position des atomes de sélénium. a : détails de la carte au voisinage de la position de certains atomes de sélénium. b : carte plus globale montrant une première organisation des monomères entre eux et la présence d'une pseudo-symétrie d'ordre 2. L'axe non cristallographique d'ordre 2 schématisé, est défini par un point de coordonnées orthogonales $-6,893 ; -16,943 ; -11,086$, et par les valeurs des cosinus directeurs suivants : $0,722 ; -0,692 ; 0,004$.

L'amélioration des cartes de densité électronique a été effectuée grâce à l'utilisation du logiciel Resolve (*Terwilliger, 2004* ; <http://www.solve.lanl.gov/>) qui utilise différentes techniques (aplatissement de solvant, moyennisation des densités). A partir de ces cartes de densité, Resolve a construit un modèle protéique constitué de 72 % des acides aminés présents dans une molécule. Pour 41 % d'entre eux, les chaînes latérales ont pu être positionnées. La **Figure 67** illustre la qualité du modèle obtenu.

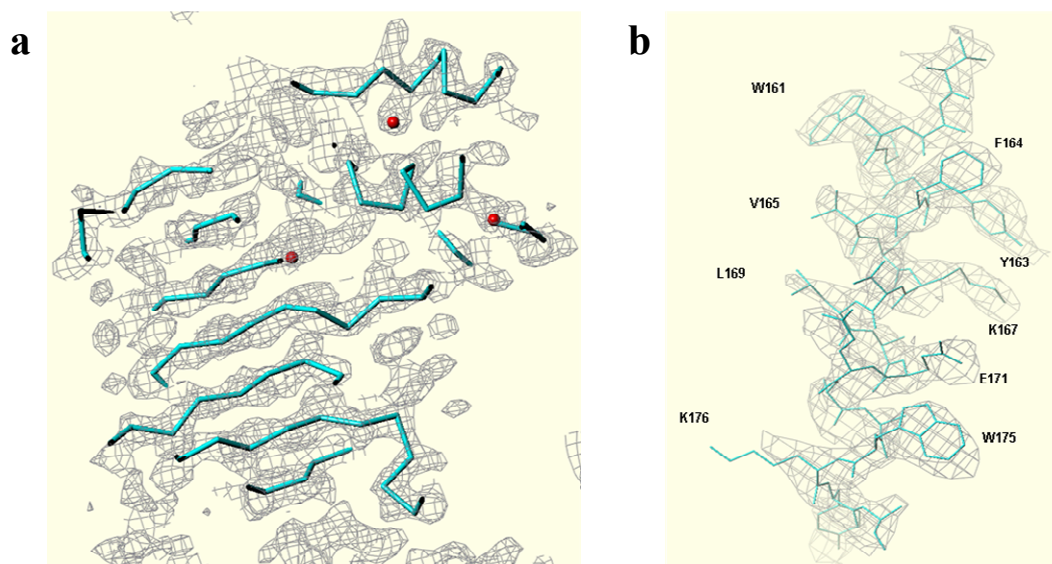


Figure 67 : Exemples de densités électroniques de type $(F_o \cdot e^{i\phi_{calc}})$ obtenues par le programme Resolve

a : Observation de la chaîne principale dans une région riche en brins β . b : La qualité de la densité électronique obtenue au voisinage d'une hélice α met en évidence l'orientation de nombreuses chaînes latérales.

II.5.6. Premiers affinements du modèle

Le modèle obtenu par Resolve a été vérifié et régularisé à l'aide du logiciel graphique Turbo-Frodo (*Roussel et al, 1990*). Les positions atomiques trouvées ont été affinées en utilisant le programme Refmac (*Murshudov et al, 1997*). Elles permettent de calculer de nouvelles cartes de densité électronique, plus précises. Leur visualisation permet d'améliorer le modèle existant (modifications et/ou introductions de nouvelles positions atomiques) qui, ainsi complété, est de nouveau affiné. Ce processus est conduit de façon itérative jusqu'à ce que soit minimisé l'écart existant entre facteurs de structure calculés et expérimentaux. Pour cela, Refmac ajuste les positions atomiques tout en maintenant des contraintes géométriques telles que longueurs de liaison, valeur des angles, contacts de van der Waals, etc.

Le **Tableau 16** décrit l'évolution de différents paramètres qui traduisent la bonne marche de l'affinement et l'amélioration du modèle :

- ✓ la figure de mérite (FOM), dont la valeur doit tendre vers 1 en fin d'affinement,
- ✓ les facteurs de confiance Rfactor overall, et Rfree (formules ci-dessous).

Ces deux facteurs traduisent l'accord entre facteurs de structure observés et calculés en prenant en compte soit l'intégralité des réflexions mesurées, soit 5 % des réflexions choisies aléatoirement et non prises en compte volontairement dans l'affinement. Les valeurs de ces deux paramètres décroissent dans le cas d'un affinement correct. En fin d'affinement, la valeur de Rfactor overall est de l'ordre de 20 %, celle de Rfree reste légèrement supérieure.

$$R_{Factor} = \frac{\sum_{hkl} \left| |F_{obs}| - k |F_{calc}| \right|}{\sum_{hkl} |F_{obs}|} \quad R_T^{free} = \frac{\sum_{hkl \subset T} \left| |F_{obs}| - k |F_{calc}| \right|}{\sum_{hkl \subset T} |F_{obs}|}$$

- ✓ les écarts quadratiques moyens (rms) sur les longueurs et les angles de liaisons traduisent la distribution des valeurs trouvées par rapport aux valeurs théoriques.

Cycle	FOM	Roverall	Rfree	rms-Liaison (Å)	rms-Angle (°)
0	0,534	0,409	0,413	0,060	0,809
1	0,692	0,296	0,364	0,010	1,379
2	0,733	0,276	0,339	0,009	1,267
3	0,735	0,274	0,332	0,009	1,283
4	0,737	0,272	0,330	0,009	1,232
5	0,745	0,265	0,325	0,010	1,360

Tableau 16 : Evolution des statistiques données par Refmac en cours d'affinement

Dans le tableau précédent :

- ✓ les valeurs données pour le cycle 0 sont celles obtenues avant tout affinement (calcul effectué sur les données issues de Resolve).
- ✓ pour le cycle 1, les valeurs sont celles calculées après affinement des positions globales des deux monomères dans la maille (affinement par groupes rigides). Les facteurs d'agitation thermique sont à cette étape, constants et identiques.
- ✓ pour les cycles de 2 à 5, les positions atomiques et les facteurs d'agitation thermique sont affinés individuellement. Les résultats présentés sont ceux issus d'un nouvel affinement après correction du modèle par observation graphique des cartes de densité électronique obtenues au cycle précédent.

A ce stade du travail, il ne nous est pas paru nécessaire de poursuivre l'affinement relatif à l'enzyme sélénée mais plutôt d'utiliser le modèle obtenu pour affiner la structure de l'enzyme native. Diverses raisons peuvent être évoquées :

- ✓ le repliement des deux molécules doit être identique, que la protéine soit sélénée ou non,
- ✓ la meilleure résolution du jeu de données relatif à l'enzyme native (1,81 Å au lieu de 2,40 Å).

II.5.7. Détermination de la structure native

La qualité de divers cristaux de DFR non sélénés a été testée sur la même ligne de lumière (BM30A) que précédemment. Deux collectes de données ont été effectuées (longueur d'onde utilisée : $\lambda = 0,9824 \text{ \AA}$) sur le meilleur cristal qui diffractait à la résolution de 1,75 Å. Les deux jeux de données ont été enregistrés dans les mêmes conditions (angle d'oscillation par image : $\Delta\phi = 1,2^\circ$; domaine angulaire total : 100° ; distance cristal – détecteur : 130 mm). Seul le temps d'exposition change d'une collecte à l'autre (10 s et 45 s), ceci afin de tenir compte des différences d'intensité observées à basse et haute résolutions et d'obtenir une plus grande précision de mesure.

II.5.8. Intégration des données

Le protocole de traitement des données est identique à celui décrit précédemment. Chaque jeu de données est indexé à l'aide du programme Mosflm dans le groupe d'espace $P2_12_12_1$ (paramètres de maille : $a = 87,80 \text{ \AA}$, $b = 89,83 \text{ \AA}$, $c = 93,15 \text{ \AA}$, $\alpha = 90^\circ$, $\beta = 90^\circ$, $\gamma = 90^\circ$). La réduction des données s'est effectuée à l'aide du logiciel Scala. Les intensités ainsi obtenues pour chaque jeu ont été mises à l'échelle puis moyennées entre elles. L'analyse critique des résultats nous conduit à exclure les données pour lesquelles le rapport $Mn(I)/sd$ est inférieur à 2, valeur limite communément acceptée. La résolution maximale retenue est alors de 1,81 Å (**Tableau 17**).

N	$1/d^2$	Dmin(A)	Rsym	Mn(I)/sd	%poss	C%poss	Mlplct
1	0.0305	5.72	0.044	44.5	91.4	91.4	6.4
2	0.0610	4.05	0.046	46.1	97.9	95.6	7.3
3	0.0916	3.30	0.054	41.3	98.7	97.0	7.8
4	0.1221	2.86	0.067	31.2	99.1	97.7	8.1
5	0.1526	2.56	0.093	22.7	99.4	98.2	8.2
6	0.1831	2.34	0.145	13.4	99.1	98.4	4.7
7	0.2137	2.16	0.134	10.1	99.7	98.7	4.0
8	0.2442	2.02	0.183	6.8	99.8	98.9	3.9
9	0.2747	1.91	0.313	3.9	99.7	99.0	3.8
10	0.3052	1.81	0.590	2.0	98.1	98.9	3.3
overall		0.069	16.7		98.9	98.9	5.3

Tableau 17 : Intégration des données après coupure à 1,81 Å

II.5.9. Reprise de l'affinement

Un nouvel affinement est conduit. Il prend en compte, non seulement ce nouveau jeu d'intensités mais aussi l'ensemble des positions atomiques trouvées lors de l'exploitation des données relatives à la protéine séléniée. Bien évidemment, dans ce dernier fichier, sont considérées des méthionines et non des séléno-méthionines. Les valeurs des différents paramètres obtenus après visualisation des cartes de densité et affinement sont reportées dans le **Tableau 18** et **Figure 68**.

Cycle	FOM	Roverall	Rfree	rms-Liaison (Å)	rms-Angle (°)
0	0,539	0,371	0,366	0,011	0,781
1	0,648	0,302	0,329	0,012	1,405
2	0,656	0,289	0,325	0,010	1,218
3	0,674	0,275	0,313	0,010	1,358
4	0,703	0,224	0,278	0,010	1,373
5	0,712	0,218	0,271	0,010	1,354
6	0,736	0,206	0,251	0,009	1,259
7	0,745	0,196	0,246	0,013	1,457
8	0,752	0,193	0,244	0,012	1,429

Tableau 18 : Evolution des statistiques données par Refmac en cours d'affinement

Un facteur d'agitation thermique isotrope a été calculé pour chaque atome, et ce, pour les huit cycles d'affinement

Les données relatives au cycle 0 ont été calculées à partir des coordonnées atomiques du modèle de départ (structure séléniée). Aucun cycle d'affinement n'a été effectué.

Les fortes variations observées à la fin du cycle 1 traduisent les modifications des positions atomiques découlant de la prise en compte d'un nouveau jeu de données et de la meilleure résolution de celui-ci. Les cycles 2, 3, 4 sont associés à la prise en compte des molécules de substrat et co-facteur, des molécules de solvant. Les deux cycles suivants sont relatifs à de simples régularisations de structure après examen des cartes de densité électronique. Le cycle 7 rend compte de la prise en considération des atomes d'hydrogène. Le cycle 8 traduit un dernier affinement et mentionne les résultats définitifs associés à la structure déposée à la PDB (code **2c29**).

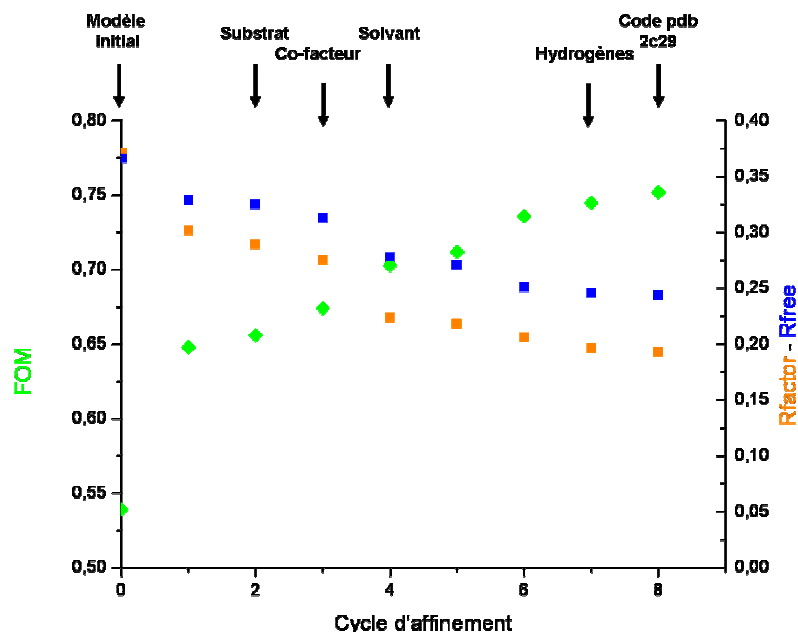


Figure 68 : Evolution des principaux paramètres relatifs à l'affinement

II.6. Validation de la structure obtenue

La structure affinée a été validée par un ensemble de programmes regroupés sous l'appellation Biotech Validation Suite for Protein Structures (*Vriend & Sander, 1993*) disponible à l'adresse internet suivante : <http://biotech.ebi.ac.uk/>.

Si l'on suppose un ensemble de N résultats, obéissant à une distribution gaussienne centrée sur une valeur théorique donnée, il est possible de discuter de la validité de chacun d'eux à partir d'un paramètre appelé « Z_{score} » défini comme suit :

$$\checkmark Z_{score} = \frac{x_i - \bar{x}}{\sigma} \text{ où } x_i \text{ est le paramètre discuté, } \bar{x} \text{ et } \sigma \text{ sont respectivement la valeur théorique et l'écart type de la distribution.}$$

L'écart quadratique moyen des valeurs Z_{score} s'écrit : $rms(Z_{score}) = \sqrt{\frac{\sum Z^2}{N}}$

Ce dernier paramètre traduit la crédibilité que l'on peut accorder à la mesure de x_i . Plus sa valeur est proche de 0, plus ce paramètre est déterminé avec précision. Le Tableau 19 liste les $rms(Z_{score})$ obtenus pour différents paramètres inhérents à l'affinement. L'ensemble des valeurs proches de 0 traduit la validité du modèle obtenu.

Paramètre contrôlé	Angles de liaison	Déviations des angles de liaison	Longueur des liaisons	Déviations des longueurs des liaisons
rms(Z_{score})	0,691	1,386	0,550	0,013

Tableau 19 : rms sur les Z_{score} obtenus sur les angles et longueurs de liaison

Un deuxième critère de validité est obtenu en traçant le diagramme de Ramachandran et en analysant le nombre de couples angulaires Phi, Psi s'écartant notablement des domaines associés à chaque élément de structure secondaire. On rappelle pour cela que la liaison peptidique est plane et rigide, et que la chaîne principale polypeptidique n'a que deux degrés de liberté de rotation. Ces derniers sont estimés par les angles dièdres ϕ ($C'_{(n-1)}-N-C_{\alpha}-C$) et ψ ($N-C_{\alpha}-C-N'_{(n+1)}$) dont les valeurs sont étroitement corrélées à la nature des structures secondaires auxquelles appartient le résidu considéré.

La **Figure 69** représente le diagramme de Ramachandran de la DFR dans lequel les domaines angulaires autorisés pour chaque élément de structure secondaire sont matérialisés en rouge. Un examen rapide de ce diagramme montre qu'un nombre très limité de ces résidus (résidus glycine exceptés) se trouve à l'extérieur des zones ainsi définies.

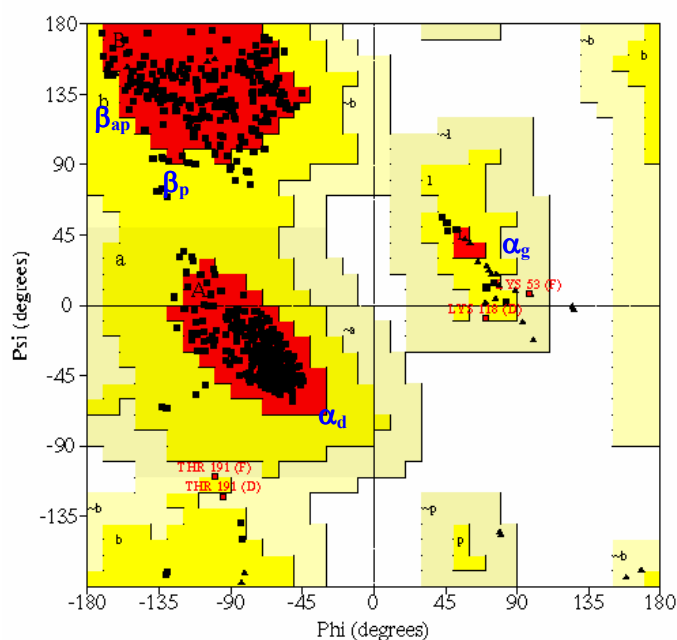


Figure 69 : Diagramme de Ramachandran
(αd ou g = hélices α droites ou gauches, βp ou ap = brin β parallèle ou antiparallèle).

Divers autres paramètres directement associés à chaque résidu permettent de critiquer plus précisément la validité de la structure obtenue. Ils sont issus de l'utilisation du programme Procheck. Dans un souci de clarté, ces résultats sont reportés en annexe I. Leur analyse ne montre que peu d'écarts notables à une géométrie acceptable.

II.7. Structure de la DFR sauvage

II.7.1. Repliement

Le repliement d'un monomère de DFR est détaillé dans le diagramme topologique représenté ci-dessous (**Figure 70**). La structure peut être décrite par deux domaines principaux.

Le premier en position N-terminale est composé d'un feuillet β coiffé de part et d'autre par une série d'hélices α . C'est un motif de type α/β communément nommé « repliement de Rossmann ». C'est la détection de celui-ci par les méthodes d'analyse bio-informatique qui, rappelons-le, nous avait permis une première approche structurale de l'enzyme. Le feuillet β se compose de sept brins β parallèles ($\beta 1$ à $\beta 7$) qui constituent le cœur hydrophobe de la protéine. Sept hélices coiffent de part et d'autre le feuillet. Le repliement de Rossmann de la DFR se caractérise néanmoins de par la présence d'un 7^{ème} brin β et de deux hélices α ($\alpha 9$, $\alpha 11$) qui modifient sensiblement la symétrie habituelle de ce domaine.

Le deuxième domaine, en position C-terminale, est composé de quatre hélices α ($\alpha 7$, $\alpha 8$, $\alpha 10$, $\alpha 12$), de deux hélices 3_{10} (3a et 3b) et de cinq brins β . Ces derniers forment deux petits feuillets β distincts ($\beta 2'\beta 3'\beta 4'$ et $\beta 1'\beta 5'$).

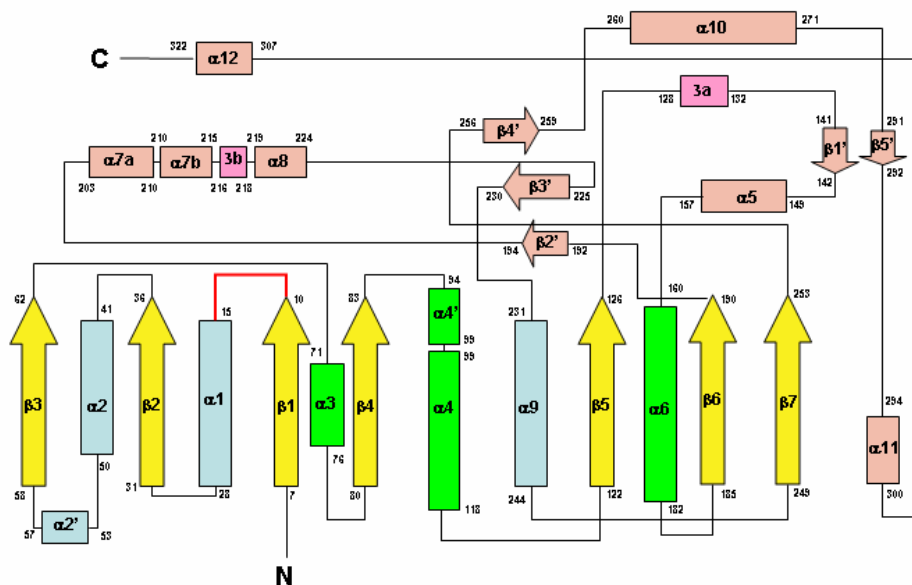


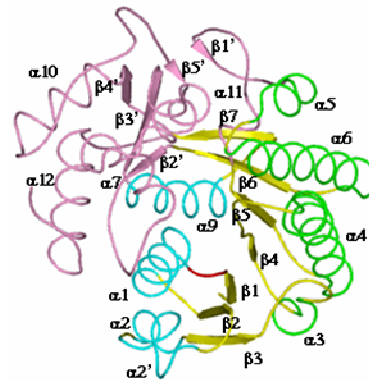
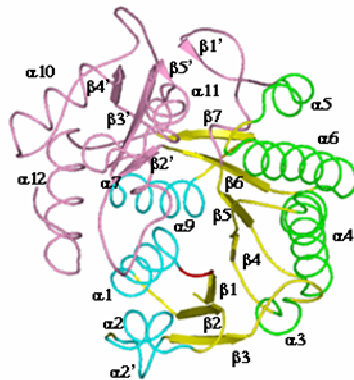
Figure 70 : Topologie de la DFR

Le feuillet β du domaine N-terminal est coloré en jaune. Les hélices α situées au dessus et au dessous du plan du feuillet sont respectivement de couleur verte ou bleue. Le domaine C-terminal est représenté en magenta et les hélices 3_{10} en rose. Les numéros des acides aminés commençant et terminant chaque élément de structure secondaire sont mentionnés.

Une comparaison de la structure 3D de la DFR à celles d'autres structures connues de la famille des SDRs (déshydrogénases à courtes chaînes) a été effectuée grâce au serveur DALI (<http://www.ebi.ac.uk/dali/>). Le meilleur alignement est celui obtenu avec la structure de la carbonyl reductase, de *Sporobolomyces salmonicolor*, SSCR (code d'entrée à la pdb : 1Y1P) (*Kamitori et al, 2005*). Cette dernière possède 26 % d'identité de séquence avec la DFR sauvage si l'on considère uniquement les 301 résidus se superposant. La déviation moyenne entre positions des carbones α correspondants est de 2,4 Å.

La **Figure 71** illustre la qualité de la superposition des structures secondaires de la DFR et de la SSCR et le **Tableau 20** traduit, pour les domaines de bonne superposition définis par Dali, les écarts calculés entre les positions des carbones α .

DFR



DFR-SSCR

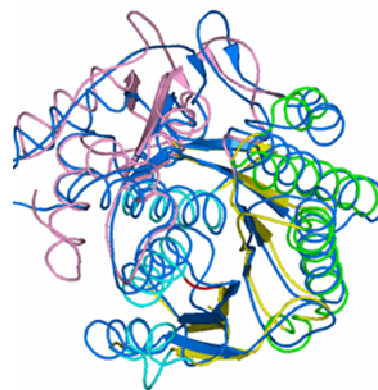


Figure 71 : Vue stéréoscopique en rubans de la DFR seule ou superposée à la SSCR (bleu foncé)

SSCR	DFR	Éléments de structure secondaire	RMS C α
S13-K48	E6-N41	β 1- α 1- β 2	1,74
L49-W56	K44-L51	α 2	2,45
K59-G62	P52-E55	α 2'	3,56
R63-V69	H57-A63	β 3'	2,09
D71-N99	D64-S92	α 3- β 4	2,02
K100-E103	D94-N97	α 4'	3,29
V104-I139	V99-I134	α 4- β 5-3a	1,05
V144-L156	H137-D149	β 1	3,55
I159-L165	M150-K156	α 5	3,90
Q171-P197	K157-I183	α 6	2,10
T200-P216	D184-S200	β 6- β 2	1,29
Q219-P231	S201-I213	α 7	3,39
E234-M242	T214-I222	3 β - α 8	2,39
P243-R268	R224-G249	β 3'- α 9	1,44
R270-G304	R250-D284	β 7- β 4'- α 10	3,42
Q305-L316	L287-T298	β 5'- α 11	3,27
E317-K320	L300-E303		3,41
G326-G339	F304-D317	α 12	2,12

Tableau 20 : Éléments de structure secondaire alignés par DALI et rms sur les C α correspondants.

Dali ne donnant pas les rms associés à chaque domaine, ces derniers ont été calculés à l'aide du logiciel Swiss-PdbViewer (*Guex et al, 1997*).

II.7.2. Site de fixation du Co-facteur

Le NADP⁺ est à l'interface des deux domaines. L'adénosine est située à proximité de la face supérieure du feuillet β tandis que le cycle nicotinamide interagit avec le substrat à proximité du brin β 2' (**Figure 72**).

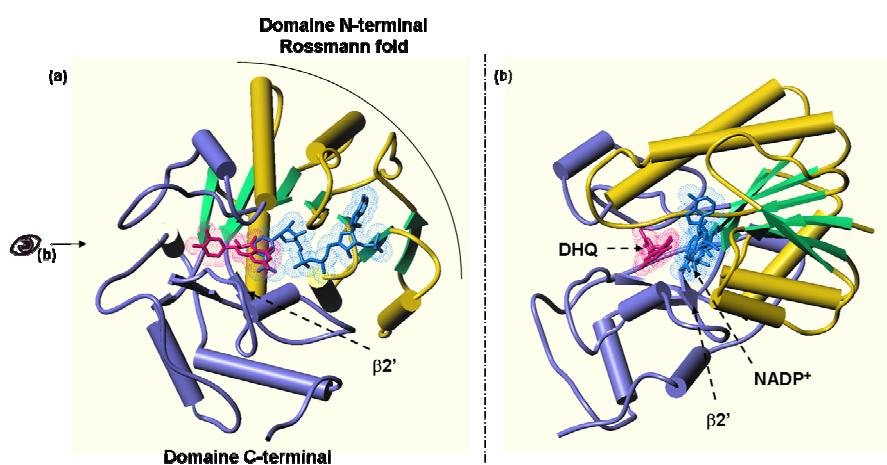


Figure 72 : Fixation du NADP⁺ au sein de la DFR

Le co-facteur est lié à la protéine par des interactions hydrophobes et un réseau de liaisons hydrogène et électrostatiques (**Figure 73**). Ces liaisons impliquent les résidus du domaine riche en glycine (Gly12, Ser14, Gly15, Phe16, Ile 17 et Gly18) de la boucle β 1- α 1, les résidus des boucles β 2- α 2 (Arg37, Asp 38, Lys 44), β 3- α 3 (Asp 64, Leu 65), β 4- α 4 (Ala 85, Thr 86, Pro 87), les résidus des hélices α 6 (Tyr 163, Lys 167) et α 7a (Ser 205), et enfin des brins β 6 et β 2' (Pro 190 et Val 193).

Le groupement pyrophosphate du co-facteur interagit avec les atomes de la chaîne principale du motif de reconnaissance du NADP⁺ par des liaisons hydrogène directes ou par l'intermédiaire d'une molécule d'eau hautement conservée dans ce type de repliement (Rossmann) (*Bottoms et al, 2002*).

Le phosphate (proche de l'adénosine) interagit quant à lui avec la DFR par des liaisons électrostatiques via les résidus Arg 37 et Lys 44. Le ribose associé adopte une conformation C2'-endo, et « anti » par rapport à l'adénosine.

Des interactions hydrophobes impliquant les résidus Leu 65, Ala 85 et Pro 87 sont créées avec la face A de l'adénosine. Des interactions de même nature existent entre la chaîne latérale de l'Arg 37 et la face B. La fonction amine (N6) établit une liaison hydrogène avec la fonction carboxyle de l'Asp 64. Les autres atomes d'azote de l'adénosine forment des liaisons hydrogène avec des molécules d'eau ou avec les atomes de la chaîne principale (Leu 65).

Le mononucléotide, constitué du nicotinamide, du ribose et du phosphate, présente quant à lui les caractéristiques suivantes :

- ✓ une conformation pro-S pour le ribose,
- ✓ l'existence d'interactions hydrophobes entre la face A du nicotinamide et les chaînes latérales des résidus Ile 17, Pro 190 et Val 193. (La face B interagit avec le substrat). La DFR peut ainsi être considérée comme une enzyme stéréospécifique de type B (pro-S)

La différence essentielle entre la DFR et la SSCR vis-à-vis de la fixation du co-facteur concerne les liaisons entre le groupement monophosphate et les résidus Ser 14, Arg 37 et Lys 44. Celles-ci sont présentes pour la SSCR (Asn 21, Arg 44 et Lys 48), néanmoins, la présence de l'Asn 21 et la délétion de trois résidus entre Arg 44 et Lys 48 par rapport à la DFR ont pour conséquence, une exposition au solvant du groupe monophosphate plus importante pour la SSCR que pour la DFR.

Nous étudierons par la suite le rôle des résidus impliqués dans la stabilisation du cycle B et l'influence de ses derniers vis-à-vis de l'activité et de la spécificité de la DFR. Les résultats répondant à cette problématique sont présentés dans la discussion.

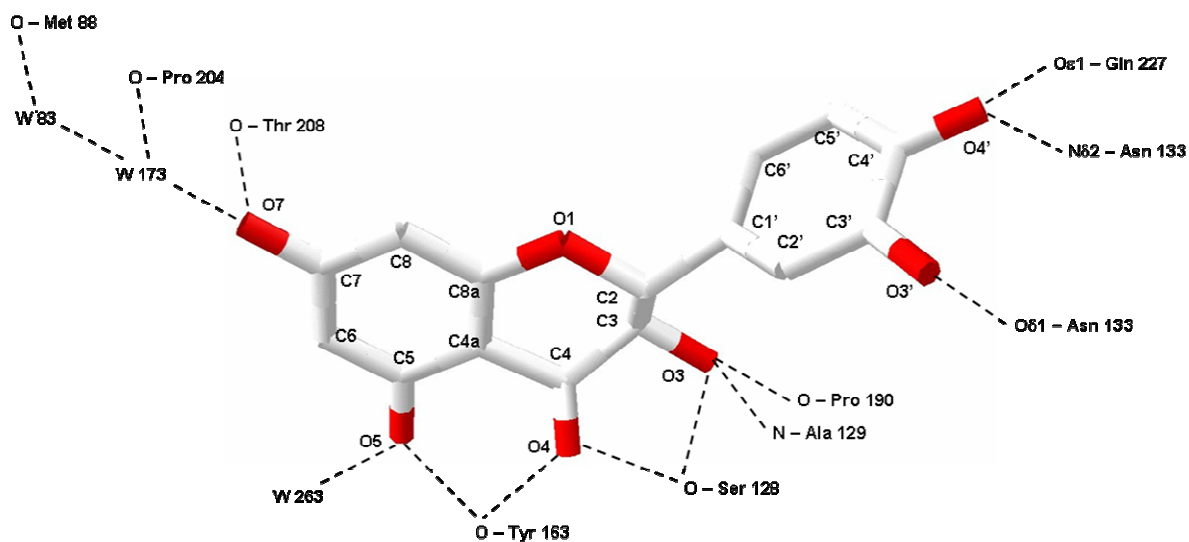


Figure 74 : Interactions polaires de la DHQ avec la DFR

La présence de deux atomes chiraux au sein du substrat implique l'existence de quatre stéréoisomères possibles : 2R3R, 2S3S, 2R3S et 2S3R. Nous avons essayé de tirer profit du fait que la DFR a été co-cristallisée avec un mélange racémique de DHQ, pour identifier si une conformation interagissait de façon préférentielle avec la DFR. La résolution maximale des cristaux obtenue (1,8 Å) n'est malheureusement pas suffisante pour discriminer avec certitude la conformation de la molécule fixée. Un examen attentif des cartes de densité électronique (**Figure 75**) suggère néanmoins une fixation préférentielle des conformations 2R3R et 2S3S (faible densité résiduelle en Fo-Fc pour ces deux énantiomères).

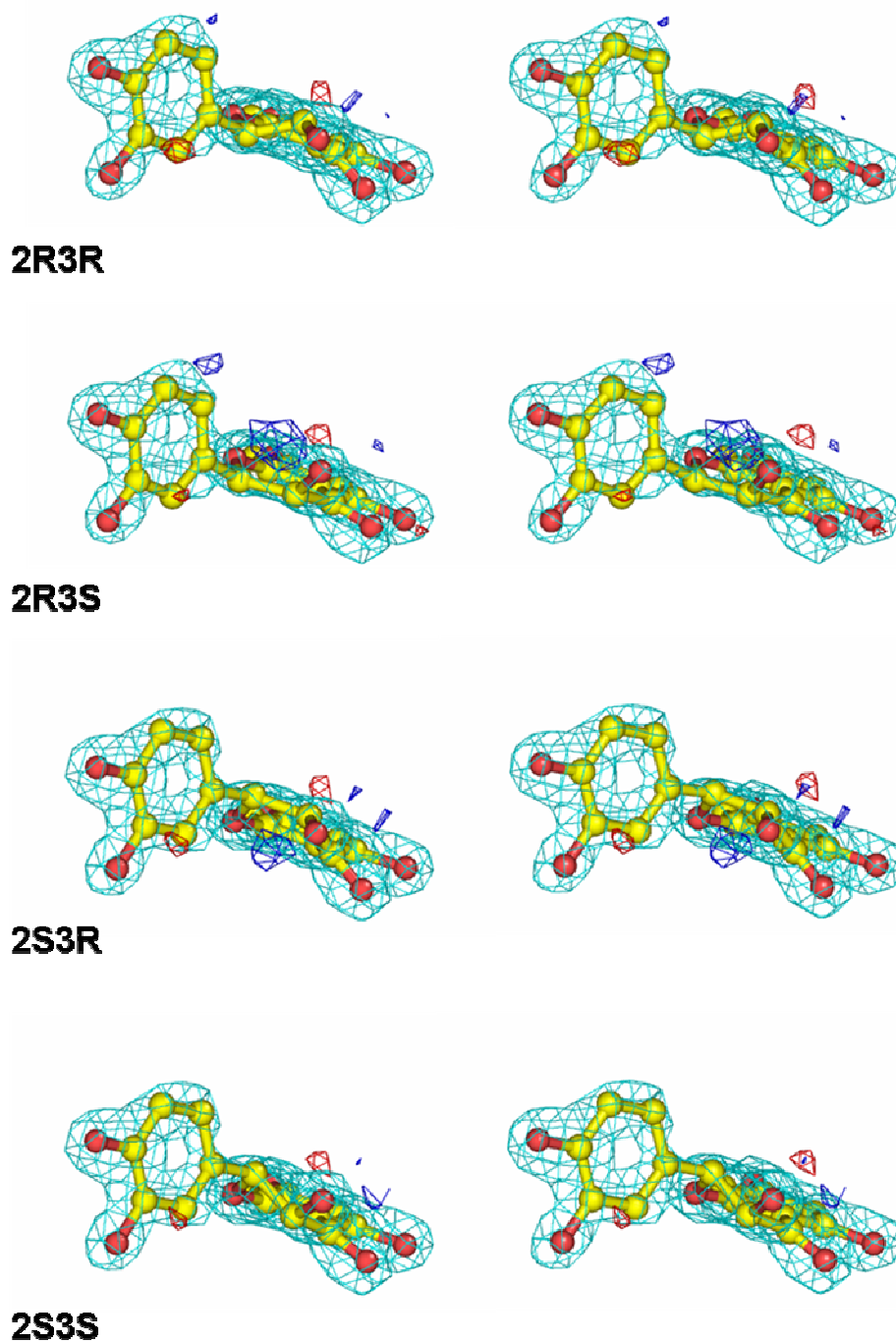


Figure 75 : Vue stéréoscopique des conformations possibles de la DHQ et cartes de densité électronique associées

Les cartes de densité électronique de type $(2F_o - F_c)$ contourées à 3σ sont représentées en bleu clair. Les cartes de type $F_o - F_c$ sont colorées en bleu foncé et rouge respectivement selon le niveau considéré -3σ et 3σ .

II.7.4. Architecture du site catalytique

Prises isolément, les molécules de NADP^+ et de DHQ ont des surfaces de contact égales à 906 \AA^2 et 479 \AA^2 . Dès lors qu'elles sont liées à la DFR, leurs surfaces accessibles (41 \AA^2 et $6,8 \text{ \AA}^2$) au solvant sont très réduites. Les deux protagonistes sont pris en tenaille à l'interface des deux domaines (**Figure 72**). Le recouvrement du site actif (**Figure 76**) est établi par les interactions faibles suivantes :

- ✓ une liaison hydrogène directe entre la Thr 159- $\text{O}\gamma_1$ et l'Ala 218-O,
- ✓ deux liaisons hydrogène s'établissant via une molécule d'eau : Thr 159-O-W178-Ser 221- $\text{O}\gamma$ et Thr 159-N-W154-Ser 221- $\text{O}\gamma$,
- ✓ deux liaisons hydrophobes impliquant les chaînes latérales des résidus Phe 164 et Ile 222, et Met 88 et Pro 204.

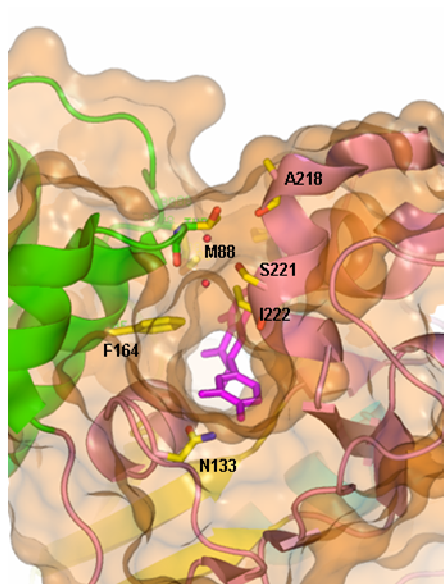


Figure 76 : Recouvrement du site actif

A l'une des extrémités de l'enzyme, une boucle ordonnée coiffant le cycle B de la DHQ achève l'enfouissement des substrats. Elle interagit avec le reste de l'enzyme via son brin $\beta 1'$ qui forme, avec le brin $\beta 5'$, un feuillet β parallèle. Grâce à divers travaux de biologie moléculaire, cette boucle (résidus 131-156), est suspectée être à l'origine de la spécificité de substrat (*Johnson et al, 2001*). La seule résolution de structure ne peut corroborer une telle conclusion. Néanmoins, celle-ci prouve que, non seulement ce segment prend part à la géométrie du site actif, mais aussi qu'il peut interagir directement avec le cycle B du substrat via l'Asn 133 (**Figure 77**).

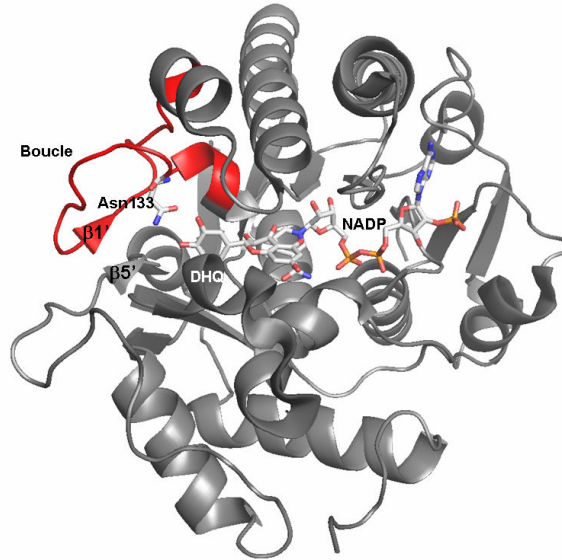


Figure 77 : Boucle délimitant une extrémité du site actif pour la DFR de *Vitis vinifera*

N'ayant pu jusqu'alors obtenir des cristaux de DFR en absence d'une ou des deux molécules de substrat, nous ne pouvons pas préciser si la conformation que nous venons de décrire pour l'holoenzyme est celle adoptée par l'apoenzyme. Un examen de l'ensemble des structures de SDRs publiées ne permet pas non plus de conclure sur ce point. En effet, différentes possibilités ont pu être observées dès lors que les structures d'apoenzyme, de complexes binaires et/ou ternaires sont connues :

- ✓ Pour la SSCR (1Y1P), le domaine C-terminal de l'apoenzyme est désordonné, suggérant qu'une fraction importante de cette dernière est mobile et que les sites de fixation aux substrats sont aisément accessibles au solvant. Par opposition, le complexe binaire (SSCR-NADP⁺) est parfaitement ordonné (*Kamitori et al, 2005*).
- ✓ La 3-D-hydroxybutyrate deshydrogénase de *Pseudomonas fragi* (1WMB, 1X1T) (*Ito et al, 2006*) présente quant à elle, une extrémité C-terminale désordonnée après fixations du NAD⁺ et du substrat. Par opposition, l'apoenzyme est ordonnée et décrit une conformation ouverte propice à la fixation des différents substrats.
- ✓ Enfin, certaines enzymes telles que la R-alcool deshydrogénase de *Lactobacillus brevis* (1ZK2-1ZK4) (*Schlieben et al, 2005*) sont parfaitement ordonnées. Aucune différence structurale importante n'est observée entre apo et holoenzyme.

Dans le complexe ternaire, l'arrangement tridimensionnel au niveau du site catalytique est similaire à celui observé dans les autres structures de protéines de la famille des SDRs. Pour rappel, l'analyse bio-informatique nous avait permis de conclure à la superposition des résidus de la triade de plusieurs SDRs, ceci en dépit de la faible homologie de ces dernières. Une vue du centre réactionnel est montrée **Figure 78**. Les groupements hydroxyle de Ser 128 et Tyr 163 interagissent par des liaisons hydrogène avec l'atome O4 de la DHQ alors que le groupement amine de la lysine 167 interagit avec les groupements hydroxyle 2' et 3' du ribose du groupement nicotinamide. Les positions relatives du substrat et du co-facteur conduisent à la géométrie d'interaction entre le carbone C4 du nicotinamide et la DHQ dont les caractéristiques sont les suivantes :

- ✓ une distance de 2,99 Å entre le C4 du nicotinamide et le C4 du substrat,
- ✓ un angle dièdre de 81° entre le plan du cycle nicotinique (6 atomes) et le plan délimité par les atomes N1 (NADP⁺), C4 (NADP⁺) et C4 (DHQ),
- ✓ un angle (C4(DHQ)-C4(NADP⁺)-N1(NADP⁺)) de 105°, valeur compatible avec celle (102°-119°) calculée avec différentes méthodes (*Yamashita et al, 1999 et références incluses*).

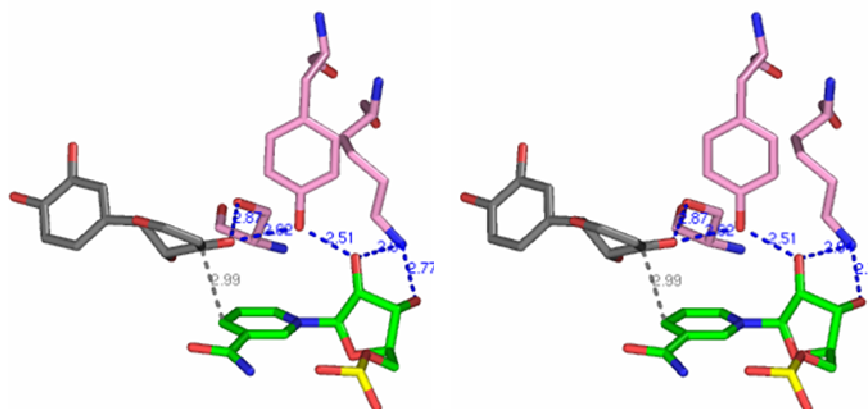


Figure 78 : Vue stéréographique montrant la géométrie du site actif

Cette géométrie est favorable à un transfert d'hydrure entre le cycle nicotinamide et le carbone C4 du substrat. Quant au transfert de proton, il est favorisé par une diminution du pKa de la tyrosine 163, induite par les différentes interactions entre la lysine 167 et le ribose proche du groupement nicotinamide. La tyrosine 163 possède donc le rôle de base catalytique (**Figure 79**).

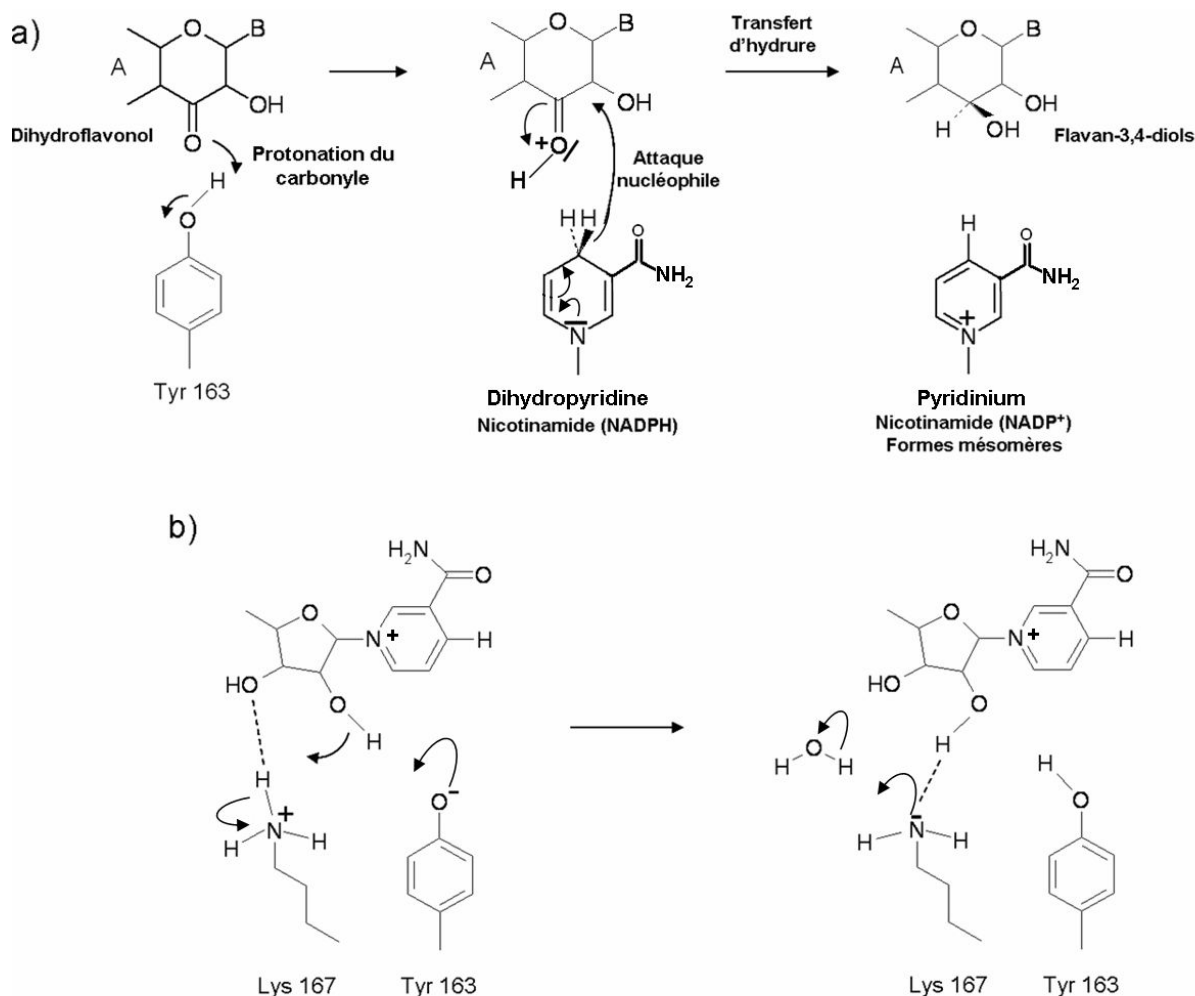


Figure 79 : Mécanisme enzymatique : hypothèse de catalyse acido-basique concertée, par le duo Tyrosine / Lysine

a) Réduction des dihydroflavonols ; b) Re-protonation de la tyrosine 163

Ce mécanisme a été proposé pour les SDR, enzymes où la triade catalytique est composée des résidus S-Y-K. Il n'est cependant pas exclu que l'activation électrophile du carbone carbonyle C₄ implique une liaison hydrogène entre l'oxygène carbonyle et le groupement phénol de la tyrosine.

Nous pouvons remarquer en outre que la lysine 167 interagit avec une molécule d'eau (W82), dont la localisation peut s'expliquer par la présence de la proline 102, qui limite la prolongation de l'hélice α_4 et permet ainsi l'initiation du transfert de proton entre le solvant et la lysine 167. Ceci aboutit à terme, au transfert d'un proton de la lysine 167 sur la tyrosine 163 via le ribose du nicotinamide (**Figure 79 et 80**).

Contrairement à ce qui est observé pour les autres SDRs, ni la molécule d'eau W82, ni ses quatre voisines ne sont en contact direct avec le solvant, ce qui est incompatible, dans le cas d'une structure figée, avec l'hypothèse d'un apport de proton issu du solvant. (*Filling et al, 2002 ; Schlieben et al, 2005*).

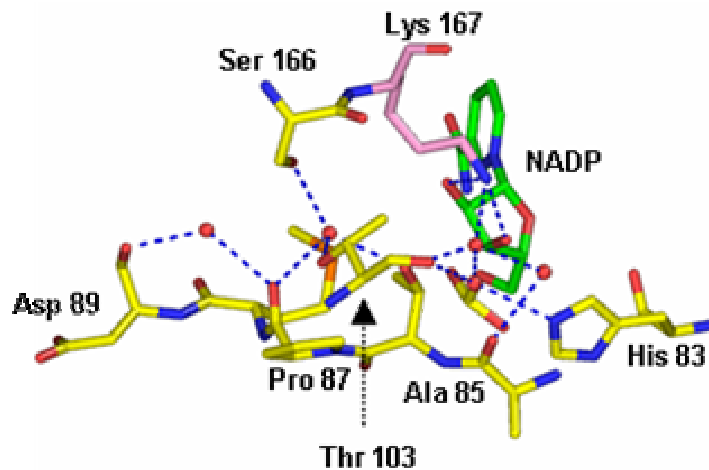


Figure 80 : Molécules d'eau impliquées dans le transfert de protons

La résolution obtenue (1,81 Å) lors de l'étude des cristaux du complexe ternaire DFR sauvage-DHQ-NADP⁺ nous a permis de décrire avec précision :

- ✓ le repliement général de l'enzyme
- ✓ la localisation et la géométrie du site actif
- ✓ la conformation de chaque molécule de substrat et les interactions substrat-enzyme
- ✓ la géométrie d'interaction DHQ-NADP⁺

Cette étude structurale est d'importance dans la mesure où aucune structure tridimensionnelle de dihydroflavonol réductase n'avait à ce jour été décrite. Les résultats obtenus confirment que cette enzyme appartient à la famille des déshydrogénases - réductases à courtes chaînes. L'obtention de ces résultats ouvre la voie à l'étude des relations structure-fonction.

Il ne nous a pas échappé, lors de l'analyse structurale que :

- ✓ la molécule de substrat se liait notamment à l'enzyme via les groupements hydroxyles de son cycle B qui, par leur nombre et leur position, caractérisent la nature même du substrat,
- ✓ ces groupements hydroxyles interagissent avec deux résidus de l'enzyme. Le premier (Gln 227), étant strictement conservé dans toutes les séquences de DFR, ne peut donc jouer un rôle important dans la fixation privilégiée d'un substrat donné. Le second (Asn 133), souvent muté en aspartate, appartient à la boucle 131-156 et est souvent décrit dans la littérature (*Johnson et al, 2001 ; Shimada et al, 2005*) comme responsable de la sélectivité de substrat.

Un article décrivant la majeure partie de ce travail a été soumis pour publication. Il est reproduit *in extenso* en Annexe II.

Compte tenu de la forte spécificité de substrat mentionnée dans l'introduction et le chapitre « caractérisation enzymatique », il devient important d'estimer l'influence de ces interactions quant à la fixation d'un substrat donné et à l'activité de l'enzyme correspondante. Deux voies d'études distinctes ont été envisagées :

- ✓ la première consiste à évaluer l'importance du nombre des fonctions hydroxyle du cycle B et donc d'envisager l'étude structurale de divers substrats complexés à la DFR sauvage,
- ✓ la seconde à pour but d'étudier l'importance de la nature des résidus constituant la boucle (**Figure 77**) sur la fixation d'un substrat donné. Pour cela, a été envisagé l'étude structurale de DFR mutées complexées à différents substrats donnés.

II.8. Justification du choix des mutations

II.8.1. Mutant 1C

Les études bibliographiques ont souligné la relation existante entre la nature du résidu 133 (Asn ou Asp) et le substrat préférentiellement réduit. Ces études concluent à une réduction préférentielle du DHK lorsqu'un résidu asparagine est présent. Nos résultats concernant l'activité enzymatique de la DFR sauvage contredisent ce dernier point.

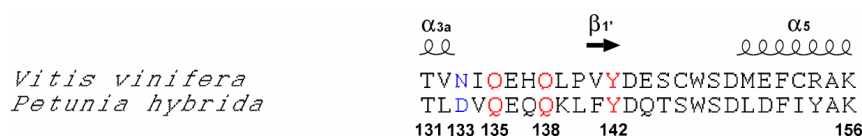
Les études structurales ayant montré l'implication de ce résidu dans la fixation du substrat, il est donc nécessaire d'étudier son influence réelle, tant d'un point de vue enzymatique que structural.

Les résultats obtenus montrent une réduction globale de l'activité dès lors qu'un aspartate est présent en position 133, et s'opposent à ceux découlant des observations mentionnées dans la littérature.

II.8.2. Mutant 4A

Pour expliquer que la DFR de vigne diffère de celles des autres plantes dans le sens où elle réduit préférentiellement la DHQ, il faut concevoir qu'elle a en commun avec la DFR de Pétunia des résidus (autres que celui présent en position 133) qui limitent très fortement la réduction du DHK. Rappelons en effet, que la DFR de Pétunia ne réduit pas le DHK.

L'alignement de séquences entre Vigne et Pétunia dans la région 131-156 est présenté ci-dessous.



Dans ce segment, les résidus communs sont en position 135, 138 et 142. Selon le raisonnement précédent, ces résidus susceptibles d'annuler la réduction du DHK, doivent être mutés pour que la DFR de Vigne adopte un comportement identique à celui des autres DFRs.

Le choix des substitutions est basé sur les résultats issus de l'étude de plusieurs DFRs de *Lotus japonicus* (*Shimada et al, 2005*). La référence choisie est la séquence de la DFR5 qui, parmi toutes les séquences étudiées du cluster, est celle qui montre une plus forte activité pour la DHQ.

Cette démarche conduit à adopter les substitutions suivantes : Q135T, Q138P et Y142I. Nous y avons ajouté la substitution N133D qui, à en croire la littérature, renforce également une réduction privilégiée de la DHQ.

Ce mutant, comparativement au précédent, doit privilégier une réduction de DHK. Les résultats expérimentaux confortent ce dernier point dès lors que l'on s'intéresse au rapport relatif des consommations en DHQ et DHK. Néanmoins, force est de constater que la réduction globale de substrats est affectée par ces trois dernières mutations.

III. Détermination de structure d'analogues de substrat

Les travaux menés dans ce chapitre constituent une première approche de l'étude de l'importance du nombre des fonctions hydroxyle du cycle B sur la fixation du substrat au site actif de la DFR sauvage. Cette étude aurait dû être effectuée, en toute rigueur, avec comme substrats, le dihydrokaempférol (un groupement hydroxyle en 4') et la dihydromyricétine DHM (trois groupements hydroxyle en 3', 4', 5'). Ceci n'a pu être réalisé pour les raisons suivantes :

- ✓ aucun cristal n'a pu être obtenu en présence de dihydrokaempférol (cf : cristallisation et qualité de diffraction),
- ✓ l'utilisation de DHM n'a pas été retenue dans un premier temps, compte tenu de la distribution limitée de ce produit (un seul distributeur (APIN-Chemicals, UK) et de son prix (500\$ / mg).

Trois substrats, dont les structures sont rappelées **Figure 81**, ont été utilisés :

- ✓ l'ériodictyol, de structure en tout point équivalente à celle de la DHQ (carbones 2 et 3 saturés, et deux fonctions hydroxyle en position 3' et 4') mais dépourvu de fonction hydroxyle en position 3.
- ✓ la myricétine, comportant quatre fonctions hydroxyle en position 3, 3', 4', 5', et une double liaison entre les carbones 2 et 3 (cycles A-C plans).

- ✓ la quercétine, de structure identique à la myricétine, mais comportant uniquement deux fonctions hydroxyles (positions 3' et 4').

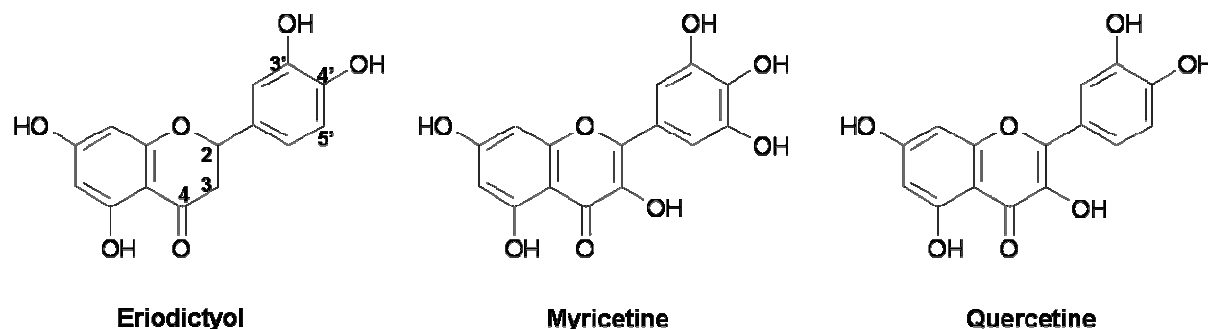


Figure 81 : Formules chimiques des substrats mis en jeu

Les conditions de cristallisation et la qualité des cristaux obtenus ont été précisées précédemment. Les collectes d'intensités diffractées relatives aux cristaux ayant pour substrats l'ériodictyol et la myricétine ont été enregistrées à l'ESRF (Grenoble) sur la ligne de lumière ID29 ($\lambda = 1,0400 \text{ \AA}$). Les données relatives à la quercétine ont été mesurées sur la ligne de lumière ID14-2D ($\lambda = 0,933 \text{ \AA}$). Dans ce chapitre, sont présentés tout d'abord les statistiques d'intégration des trois jeux de données, les paramètres de maille et groupe d'espace de chaque complexe, une brève description de la méthode utilisée pour résoudre la structure, et enfin la description structurale de chaque complexe.

III.1. Traitement des données

Une définition des différents paramètres permettant de caractériser la qualité d'un jeu de données a été donnée lorsque nous avons exposé les résultats obtenus sur la DFR sauvage complexée à la DHQ. Nous ne reviendrons donc pas sur celle-ci et rapportons uniquement les valeurs numériques obtenues selon un format déjà rencontré.

III.1.1. Eriodictyol-NADP⁺-DFR

Les données ont été collectées pour une rotation totale du cristal de 180°. Le temps d'exposition et l'angle d'oscillation ($\Delta\phi$) par image sont respectivement de 0,3 s et de 1°.

N	1/d ²	Dmin(A)	Rsym	Rfull	Mn(I)/sd	%poss	C%poss	Mlplct
1	0.0227	6.64	0.038	0.031	37.1	75.5	75.5	5.9
2	0.0454	4.70	0.059	0.049	32.5	80.2	78.5	6.0
3	0.0680	3.83	0.061	0.046	27.5	81.7	80.0	5.2
4	0.0907	3.32	0.080	0.071	20.1	83.4	81.2	5.5
5	0.1134	2.97	0.127	0.126	13.3	85.0	82.3	5.5
6	0.1361	2.71	0.222	0.244	8.7	85.9	83.2	5.5
7	0.1587	2.51	0.391	0.490	5.7	87.3	84.0	5.5
8	0.1814	2.35	0.599	0.795	4.0	87.6	84.7	5.4
9	0.2041	2.21	0.961	1.354	2.8	88.7	85.3	5.4
10	0.2268	2.10	1.378	1.654	2.0	89.1	85.9	5.3
Overall			0.136	0.128	10.9	85.9	85.9	5.4

Tableau 21 : Intégration des données après coupure à 2,0 Å

Lors de la collecte, les intensités diffractées ont été mesurées jusqu'à une résolution de 1,9 Å. Toutefois, un examen rapide de l'évolution du rapport Mn(I)/sd montre qu'il a été nécessaire d'exclure les réflexions mesurées à une résolution supérieure à 2,1 Å.

Le groupe d'espace trouvé (P2₁2₁2₁, système orthorhombique) et les paramètres de maille ($a = 87,82 \text{ \AA}$, $b = 90,14 \text{ \AA}$, $c = 93,29 \text{ \AA}$, $\alpha = 90^\circ$, $\beta = 90^\circ$, $\gamma = 90^\circ$) sont semblables à ceux du cristal modèle. Deux molécules indépendantes sont présentes dans la maille.

Cet isomorphisme nous a conduit à considérer que l'organisation moléculaire était identique dans les deux mailles cristallines. L'affinement de la structure a été conduit en utilisant le programme Refmac et en considérant comme modèle de départ les positions atomiques du dimère trouvé dans la structure de référence. L'examen des cartes de densité électronique différence ont conduit à positionner sans ambiguïté co-facteur, molécules de substrat et de solvant. Le tableau ci-dessous (**Tableau 22**) mentionne les principales caractéristiques obtenues en fin d'affinement.

Cycle	FOM	Roverall	Rfree	rms-Liaison (Å)	rms-Angle (°)
10	0,785	0,190	0,245	0,009	1,253

Tableau 22 : Statistiques en fin d'affinement

III.1.2. Myricétine-NADP⁺-DFR

Les caractéristiques de la collecte de données sont équivalentes à celles précédemment utilisées (temps d'exposition par image, 0,3 s, $\Delta\phi = 1^\circ$). Le domaine angulaire total enregistré est de 120°. Ce jeu de données a été indexé dans le groupe d'espace P2₁ (système monoclinique) avec pour paramètres de maille : $a = 47,23 \text{ \AA}$, $b = 177,95 \text{ \AA}$, $c = 92,59 \text{ \AA}$, $\alpha = 90^\circ$, $\beta = 104,77^\circ$, $\gamma = 90^\circ$. La résolution maximale est de 2,06 Å de résolution.

N	1/d ²	Dmin(A)	Rsym	Rfull	Mn(I)/sd	%poss	C%poss	Mlplct
1	0.0236	6.51	0.042	0.037	23.9	97.2	97.2	3.3
2	0.0471	4.61	0.073	0.062	22.4	99.2	98.5	3.2
3	0.0707	3.76	0.066	0.051	22.8	99.8	99.1	3.3
4	0.0943	3.26	0.059	0.047	19.7	99.9	99.4	3.3
5	0.1178	2.91	0.076	0.059	15.4	100.0	99.5	3.3
6	0.1414	2.66	0.112	0.091	10.9	100.0	99.6	3.3
7	0.1650	2.46	0.154	0.127	7.8	100.0	99.7	3.3
8	0.1885	2.30	0.231	0.197	5.3	100.0	99.8	3.3
9	0.2121	2.17	0.477	0.413	3.4	99.9	99.8	3.2
10	0.2356	2.06	0.505	0.421	2.1	94.1	99.0	2.5
Overall			0.093	0.073	10.7	99.0	99.0	3.2

Tableau 23 : Intégration des données après coupure à 2,06 Å

Dans ce cas précis, un calcul du coefficient de Matthews montre la présence de quatre molécules indépendantes dans la maille ($V_m = 2,4 \text{ \AA}^3/\text{Da}$ soit 47,3 % du volume de la maille occupée par le solvant).

La structure a été résolue par remplacement moléculaire en utilisant le programme Molrep (*Vagin, 1997*). Cette méthode consiste à choisir un modèle moléculaire de structure 3D connue et semblable à celle étudiée, puis à l'orienter et le translater dans la maille cristalline afin de trouver la position qu'il occupe. La recherche de l'orientation et de la translation est effectuée par superposition des cartes de Patterson calculées respectivement à partir des intensités diffractées (collecte de données) et des positions du modèle. Les coordonnées atomiques d'une molécule de DFR sont utilisées comme modèle. La résolution de structure a été immédiate en utilisant les paramètres par défaut donnés par le programme. Parmi les solutions de rotation proposées, quatre se dégagent plus particulièrement. Elles ont été retenues dans la recherche des opérations de translation. Après calcul des translations, une solution est évidente. Elle se caractérise par un facteur de confiance de 0,41 et un facteur de corrélation de 0,59. Les positions atomiques ainsi obtenues ont permis un premier calcul des densités électroniques suffisamment correctes pour procéder à l'affinement de la structure. Les statistiques de fin d'affinement sont présentées dans le **Tableau 24** ci-dessous.

Cycle	FOM	Roverall	Rfree	rms-Liaison (Å)	rms-Angle (°)
10	0,789	0,190	0,263	0,014	1,934

Tableau 24 : Statistiques en fin d'affinement

III.1.3. Quercétine-NADP-DFR

Les intensités diffractées ont été mesurées pour un domaine angulaire total de 100°. Le temps d'exposition et l'angle d'oscillation par image sont respectivement de 40 s et de 1°. Le cristal ayant une moindre qualité que les deux précédents, la collecte de données a été effectuée jusqu'à 2,6 Å de résolution. L'exploitation des intensités diffractées n'a pu, quant à elle, être effectuée que jusqu'à la résolution de 2,9 Å.

N	1/d ²	Dmin(Å)	Rsym	Rfull	Mn(I/sd)	%poss	C%poss	Mlplct
1	0.0119	9.17	0.060	0.058	23.5	99.8	99.8	7.0
2	0.0238	6.48	0.072	0.067	20.6	100.0	99.9	7.6
3	0.0357	5.29	0.096	0.087	16.6	100.0	100.0	7.6
4	0.0476	4.59	0.090	0.081	18.6	100.0	100.0	7.7
5	0.0595	4.10	0.096	0.084	18.3	100.0	100.0	7.7
6	0.0713	3.74	0.117	0.104	15.4	100.0	100.0	7.7
7	0.0832	3.47	0.168	0.149	11.2	100.0	100.0	7.7
8	0.0951	3.24	0.233	0.204	8.5	100.0	100.0	7.6
9	0.1070	3.06	0.329	0.294	6.2	100.0	100.0	7.5
10	0.1189	2.90	0.414	0.366	4.9	100.0	100.0	7.4
Overall			0.149	0.133	12.4	100.0	100.0	7.6

Tableau 25 : Intégration des données après coupure à 2,9 Å

Les données de diffraction ont été indexées dans une maille hexagonale (groupe d'espace : P6₁22) dont les paramètres sont : $a = b = 174,94 \text{ \AA}$, $c = 290,17 \text{ \AA}$, $\alpha = \beta = 90^\circ$, $\gamma = 120^\circ$. Si l'on admet que 55 % du volume de la maille est occupé par le solvant (coefficient de Matthews : 2,7 A³/Da), l'unité asymétrique est constituée de six molécules indépendantes. En dépit de très nombreux efforts, nous n'avons pu trouver par les méthodes de remplacement moléculaire aucune solution satisfaisante. L'analyse de la fonction de Patterson, lorsque les données sont exploitées jusqu'à 8 Å de résolution, met en évidence la présence d'un pic d'intensité élevée, dont la hauteur est voisine de 45 % de celle du pic à l'origine. Ce pic (de coordonnées fractionnées, U = 0,333, V = 0,666, W = 0,000) indique très vraisemblablement la présence d'un vecteur de pseudo-translation. Si ce pic est utilisé comme traduisant une translation de réseau d'une maille plus petite (Navaza *et al*, 1998), les paramètres de cette dernière sont : $a' = b' = 101,0 \text{ \AA}$, $c' = 290,2 \text{ \AA}$, $\alpha = \beta = 90^\circ$, $\gamma = 120^\circ$. Les nouveaux indices h'k'l' des réflexions se déduisent des anciens, hkl, par les transformations suivantes : $h' = h/3 - k/3$; $k' = h/3 + 2k/3$ et $l' = l$. Cette réduction de maille aboutit à la prise en compte d'une réflexion sur 3 et, dans le même temps, à un nombre de molécules indépendantes moindres (2 au lieu de 6). Ce nouveau jeu de données a été utilisé pour résoudre la structure. Parmi les solutions de rotation, deux se détachent nettement. Ces deux solutions ont été retenues pour un calcul de translations. La solution trouvée conduit à un facteur de corrélation égal à 0,30 et un facteur de confiance de 0,69. La paire de molécules ainsi trouvée a par la suite été utilisée comme modèle pour la recherche des deux autres paires dans la maille initiale. La position de ces derniers a pu être déterminée sans grande ambiguïté (facteur de corrélation : 0,58 ; facteur de confiance : 0,53).

L'affinement de la structure a été conduit selon une méthode en tous points identique à celle utilisée lors des déterminations structurales précédentes. L'utilisation du programme Refmac conduit en fin d'affinement aux différentes caractéristiques suivantes :

Cycle	FOM	Roverall	Rfree	rms-Liaison (Å)	rms-Angle (°)
10	0,639	0,272	0,363	0,014	2,543

Tableau 26 : Statistiques en fin d'affinement

III.2. Descriptions structurales

Dans ce chapitre, nous présenterons uniquement les résultats structuraux utiles à une discussion de l'influence de la nature d'un substrat quant à la fixation de celui-ci. Nous nous attacherons à préciser plus particulièrement la géométrie des interactions substrat-protéine ainsi que les modifications de structure induites au voisinage du site actif par la fixation de molécules de substrat de nature différente. Nous sommes néanmoins conscient que des différences locales de conformation de la chaîne protéique peuvent être relevées au sein d'une même structure ou de structures homologues en d'autres lieux qu'au site actif. Ces différences, essentiellement dues à des modifications de l'arrangement moléculaire ou du degré de solvation, restent néanmoins limitées et ne présentent qu'un faible intérêt par rapport à la problématique retenue.

Seuls les fichiers relatifs aux structures concernant l'ériodictyol (pdb code : 2NNL) et la myricétine (pdb code : 2IOD) ont été déposés à la Protein Data Bank. La moindre qualité cristalline du jeu de données relatif à la quercétine, associée à une valeur encore élevée du facteur de confiance, nous a conduit à repousser le dépôt des coordonnées atomiques de cette dernière structure.

III.2.1. Modifications conformationnelles de la chaîne polypeptidique

Une première estimation des modifications induites par la fixation de molécules de substrat différentes peut être déduite des différences observées après superposition de molécules caractérisant chaque complexe. Cette superposition a été effectuée à l'aide du programme Coot (*Emsley et al, 2004*). La **Figure 82** traduit les écarts de position entre carbones α . La structure prise comme référence est celle du complexe ternaire DFR-NADP⁺-DHQ.

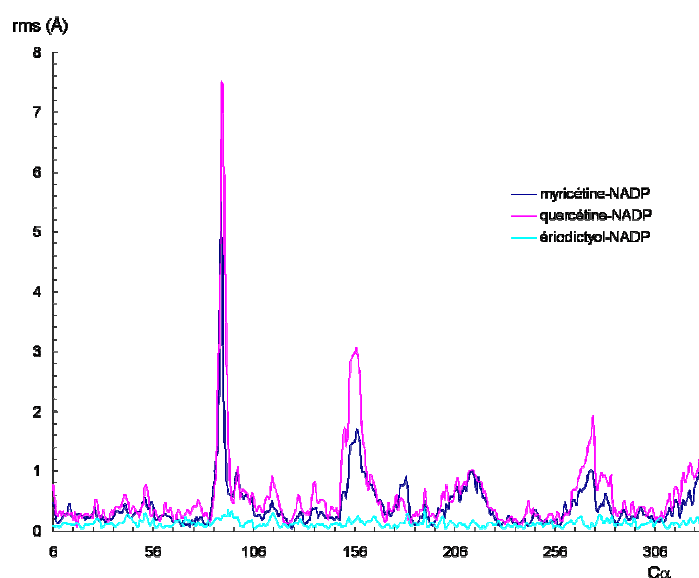


Figure 82 : Ecarts de position entre carbones α

Aucune modification importante n'est observée sur la chaîne principale de la protéine dans le cas où l'ériodictyol est utilisé comme substrat. L'écart moyen entre positions atomiques est de 0,13 Å avec une valeur maximum de 0,35 Å.

Si l'on considère comme significatif tout écart de position supérieur à 0,5 Å, un changement de conformation de la chaîne protéique est observé pour cinq segments distincts que l'on retrouve tant dans la structure relative à la myricétine que dans celle relative à la quercétine. Ces segments, ainsi que les écarts moyens sur les positions des carbones alpha correspondants, sont reportés dans le **Tableau 27** ci-dessous :

Domaines	1	2	3	4	5
Résidus (DFR-Myricétine)	87-106	150-169	200-222	265-282	322-329
Ecart moyen (Å)	1,28	0,97	0,71	0,64	0,79
Résidus (DFR-Quercétine)	87-106	150-169	198-218	264-284	313-329
Ecart moyen (Å)	1,73	1,61	0,74	1,01	0,85

Tableau 27 : Segments significativement altérés par la fixation de la myricétine ou de la quercétine par rapport à la référence

Il est à noter que la conformation du segment 131-156, délimitant le site actif (cf : structure de référence) et décrit dans littérature comme responsable de la spécificité de substrat (*Johnson et al, 2001*) n'est globalement pas affectée (à l'exception des résidus 151-156) par la nature du substrat co-cristallisé. Les écarts moyens calculés sur les positions des carbones alpha de ce segment sont respectivement 0,5 et 0,8 Å pour les structures co-cristallisées avec la myricétine et la quercétine.

La localisation des cinq segments précédents est mise en évidence **Figure 83** dans le cas de la myricétine.

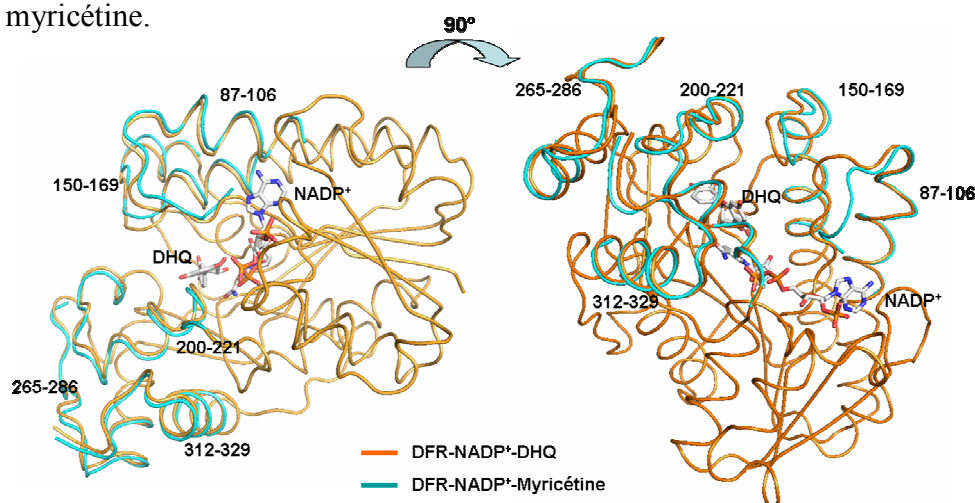


Figure 83 : Superposition des structures du modèle et de l'étude (DFR-NADP⁺-Myricétine)

Les segments 1 et 2 appartiennent au domaine N-terminal. Ils constituent à eux deux une des extrémités de la pince qui stabilise les molécules de substrat. Les segments 3, 4 et 5 appartiennent au domaine C-terminal. A lui seul, le segment 3 constitue la seconde extrémité de la pince. Les segments 4 et 5 sont plus éloignés du site actif et plus particulièrement exposés au solvant. Les modifications structurales observées pour les segments 1 et 2 mettent principalement en évidence une réorientation locale de l'enzyme. Pour les autres segments, la valeur élevée de l'écart moyen calculé traduit non pas un changement de conformation mais une dilatation de l'enzyme, associée, nous le verrons, à une augmentation de volume de la poche catalytique.

III.2.2.DFR-NADP⁺-Eriodictyol

Comme le laissait supposer l'analyse précédente des écarts de positions atomiques entre carbones α , la structure de ce complexe est en tout point identique à celle de la structure de référence. La géométrie de la molécule de substrat est conservée. Sa position et son orientation au sein du site actif le sont également. Ni les interactions substrat-protéine, ni les interactions substrat – co-facteur ne sont modifiées. Les valeurs de quelques paramètres (**Tableau 28**) représentatifs de la géométrie d'interaction sont fortement conservées.

Substrat co-cristallisé à la DFR		Eriodictyol	DHQ
Interactions cycle B - DFR	O(C3')H... O8 ₁ (Asn 133)	2,51 ± 0,02 Å	2,54 ± 0,01 Å
	O(C4')... HN8 ₂ (Asn 133)	3,03 ± 0,11 Å	3,01 ± 0,06 Å
	O(C4')H... O8 ₁ (Gln 227)	2,46 ± 0,06 Å	2,68 ± 0,18 Å
Géométrie du substrat	O1 -C2-C2'-C3'	-67,98 ± 1,81°	-66,22 ± 0,2°
	O1-C2-C3-C4	-57,82 ± 2,58°	-59,90 ± 1,06°
Transfert d'hydrure	C4 (substrat)-C4(NADP ⁺)-N1 (NADP ⁺)	105,83 ± 0,04°	108,72 ± 2,94°
	C4 (substrat)-C4(NADP ⁺)	3,17 ± 0,05 Å	2,94 ± 0,08 Å

Tableau 28 : Interactions Eriodictyol – DFR

Les valeurs indiquées sont les valeurs moyennes des paramètres considérés pour chaque molécule.

En conclusion, l'absence de la fonction hydroxyle en C3 n'induit aucune différence dans la géométrie de fixation quelle que soit la molécule de substrat considérée (ériodictyol ou dihydroquercétine). Elle ne modifie en rien l'orientation de la chaîne latérale du résidu Ser128 qui, par contre, n'établit plus qu'une liaison hydrogène avec la fonction carbonyle en C4. La distance entre donneur et accepteur est conservée : $2,94 \pm 0,21$ Å (valeur de référence : $3,24 \pm 0,04$ Å).

L'absence ou non de fonction hydroxyle en C3 n'entraînant aucune modification dans la géométrie de fixation, nous pouvons considérer que la réorientation de cette fonction associée à la planéité du cycle C ne sera pas déterminante quant à l'étude de la fixation de la myricétine en lieu et place de la dihydromyricétine.

III.2.3.DFR-NADP⁺-Myricétine

L'analyse des cartes de densité électronique révèle la présence de deux molécules de myricétine en interaction avec la DFR.

III.2.3.1. Géométrie des substrats

La géométrie de la première molécule est équivalente à celle de la molécule de référence (dihydroquercétine). L'angle dièdre (O1-C2-C2'-C3') correspondant à l'angle formé par le cycle B par rapport au plan des cycles A-C est de -70° . La position et l'orientation de cette molécule sont en tout point similaires à celles trouvées pour la dihydroquercétine ou l'ériodictyol. Les valeurs caractérisant les interactions substrat-protéine et substrat - co-facteur sont reportées dans le **Tableau 29**.

Substrat co-cristallisé à la DFR		Myricétine	DHQ
Interactions cycle B - DFR	O(C3')H... O δ_1 (Asn 133)	2,62 ± 0,12 Å	2,54 ± 0,01 Å
	O(C4')... HN δ_2 (Asn 133)	3,18 ± 0,08 Å	3,01 ± 0,06 Å
	O(C4')H... O ϵ_1 (Gln 227)	2,61 ± 0,05 Å	2,68 ± 0,18 Å
	O(C5')H... O ϵ_1 (Gln 227)	2,66 ± 0,08 Å	-
Géométrie du substrat	O1 -C2-C2'-C3'	-70,12 ± 1,21°	-66,22 ± 0,2°
	O1-C2-C3-C4	0,00°	-59,90 ± 1,06°
Transfert d'hydure	C4 (substrat)-C4(NADP ⁺)-N1(NADP ⁺)	88,80 ± 1,47°	108,72 ± 2,84°
	C4 (substrat)-C4(NADP ⁺)	3,28 ± 0,05 Å	2,94 ± 0,08 Å

Tableau 29 : Interactions Myricétine1-DFR

Les valeurs indiquées sont les valeurs moyennes des paramètres considérés pour chaque molécule.

La seconde molécule est quasiment plane, l'angle entre le cycle B et les cycles A et C est de $7,3^\circ$.

Les deux molécules de substrat sont pratiquement parallèles entre elles, l'angle formé par les cycles A-C est égal à $13,3^\circ$ et la distance entre plans est comprise entre 3,3 et 3,8 Å (**Figure 84**). Toutefois, le recouvrement n'est pas total, l'angle formé par les directions C6-C2 de chaque molécule est de $102,4^\circ$.

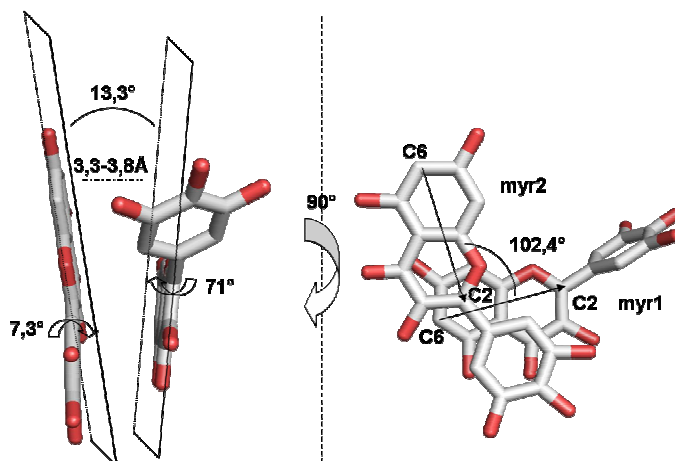
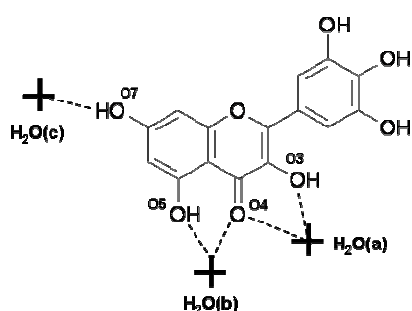


Figure 84 : Positions et interactions des deux molécules de myricétine

La molécule 2 est positionnée à l'interface des deux domaines N et C terminaux. Elle est toutefois moins enfouie que la première molécule. Son environnement proche est constitué, d'un côté du plan A-C de la molécule 1 et de l'autre côté, des acides aminés des segments 1, 2 et 3 responsables de la prise en tenaille des substrats dans la structure de référence. Elle interagit notamment via des interactions hydrophobes avec les résidus Thr 86, Met 88, Thr 159, Tyr 163, Phe 164, Lys 167 et Ile 222. Quelques liaisons hydrogène (**Tableau 30**) sont néanmoins observées. Elles s'établissent principalement avec la serine 128, le ribose associé au groupement nicotinamide et, du fait d'un enfouissement moindre, avec des molécules d'eau du solvant.



Substrat co-cristallisé à la DFR		Myricétine
Interactions H ₂ O-DFR	O(C3)H...O(H ₂ O)a	2,80 ± 0,15 Å
	O(C4)H...O(H ₂ O)a	3,28 ± 0,16 Å
	O(C4)H...O(H ₂ O)b	2,87 ± 0,03 Å
	O(C5)H...O(H ₂ O)b	3,54 ± 0,08 Å
	O(C7)H...O(H ₂ O)c	2,54 ± 0,17 Å
Interactions cycle B - DFR	O(C3')H...Oγ(Ser 128)	2,28 ± 0,04 Å
	O(C4')H...Oγ(Ser 128)	2,42 ± 0,06 Å
Interactions cycle B - NADP ⁺	O(C4')H...NO2 ⁺	3,32 ± 0,13 Å
	O(C5')H...NO2 ⁺	2,85 ± 0,17 Å
Géométrie du substrat	O1 -C2-C2-C3'	173,94 ± 3,59°
	O1-C2-C3-C4	0,00°

Tableau 30 : Interactions de type liaisons hydrogène Myricétine2-DFR

Les valeurs indiquées sont les valeurs moyennes des paramètres considérés pour chaque monomère.

La présence de cette deuxième molécule de substrat au voisinage du site actif contraint la géométrie de l'enzyme. Outre le fait qu'elle induit un changement de conformation des segments 1 et 2, elle limite, de par son encombrement stérique, « l'effet tenaille » observé préalablement entre les domaines N et C terminaux. Les interactions inter domaines décrites dans la structure modèle ne sont plus observées (pertes de liaisons hydrogène et des interactions hydrophobes) comme le montre l'évolution des distances listées ci-dessous :

Interactions inter domaines	DFR-NADP ⁺ -Myricétine	DFR-NADP ⁺ -DHQ
(Thr 159)O _γ H-O(Ala 218)	4,68 ± 0,18 Å	2,53 ± 0,06 Å
(Thr 159)O-W-O _γ (Ser 221)	2,57 ± 0,13 Å 3,82 ± 0,40 Å	2,80 ± 0,06 Å 2,67 ± 0,06 Å
(Thr 159)N-W-O _γ (Ser 221)	4,23 ± 0,48 Å 2,91 ± 0,88 Å	2,84 ± 0,13 Å 2,72 ± 0,33 Å

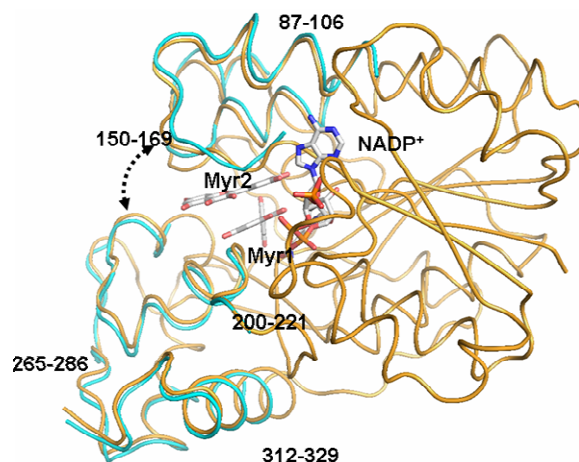


Figure 85 : Interactions inter domaines

Nous allons donc maintenant nous intéresser à l'influence de la fixation de cette deuxième molécule de substrat sur la structure de la chaîne polypeptidique des segments 1 et 2 de la protéine.

III.2.3.2. Modifications induites par la deuxième molécule de substrat

L'analyse des modifications des angles de torsion ϕ , ψ (Kleywegt plot) (*Kleywegt, 1997*) pour le segment 1 (résidus 87 à 106) (**Figure 86**) met en évidence de larges variations angulaires pour les résidus 87 à 92 impliqués dans la boucle comprise entre le brin β_4 et l'hélice α_4' . Ces résidus demeurent néanmoins dans des conformations énergétiquement favorables. Les modifications observées n'entraînent pas une restructuration du segment considéré mais seulement une réorientation de la boucle à ce niveau (**Figure 87**) sans modification notable des angles de torsion χ_1 et χ_2 des chaînes latérales.

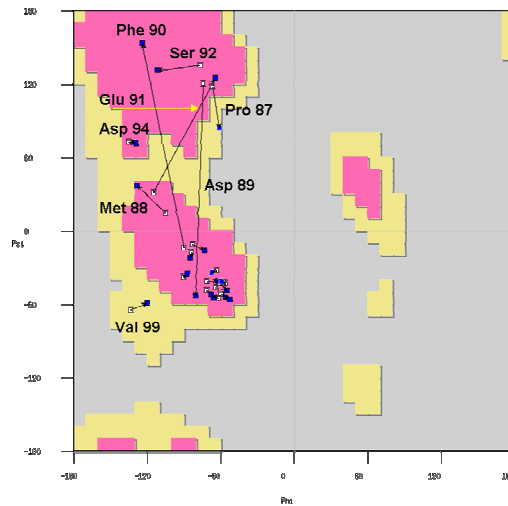


Figure 86 : Diagramme de Ramachandran pour le segment 87-106

Les symboles blancs et bleus représentent respectivement les couples ϕ , ψ d'un résidu donné dans la structure de référence et dans celle obtenue en présence de myricétine.

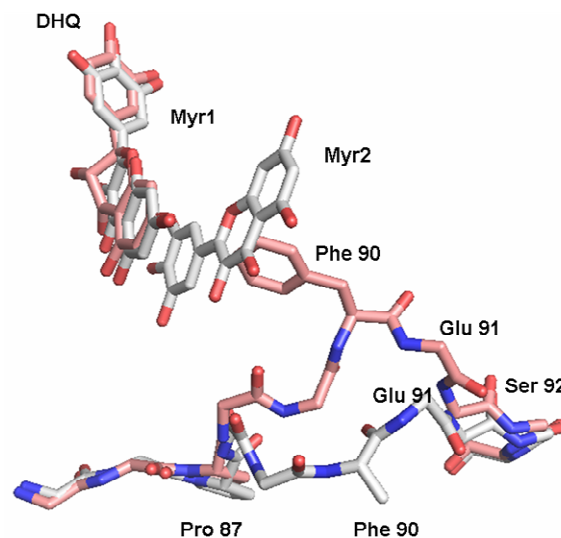


Figure 87 : Comparaison des domaines 85-94 (la structure de référence est colorée en magenta).

Une analyse identique a été effectuée pour le deuxième domaine (résidus 150 à 169). Dans ce cas, aucune différence significative n'est observée en ce qui concerne les angles ϕ , ψ de chaque résidu. Il n'y a donc pas de réorientation de la chaîne principale. Les modifications essentielles concernent l'orientation des chaînes latérales, notamment celle des résidus 156, 157, 163, 164 (**Figure 88**)

La présence d'une deuxième molécule de substrat ne modifie pas l'orientation de la lysine 167 impliquée dans la triade catalytique. Elle provoque par contre une réorientation de la chaîne latérale de la tyrosine 163, autre résidu impliqué dans la réaction enzymatique. L'interaction entre le groupement phénol de cette tyrosine et le carbonyle en position 4 de la molécule 1 n'étant plus observée, le transfert de proton ne peut être effectué. Cette modification structurale annihile donc toute activité de l'enzyme.

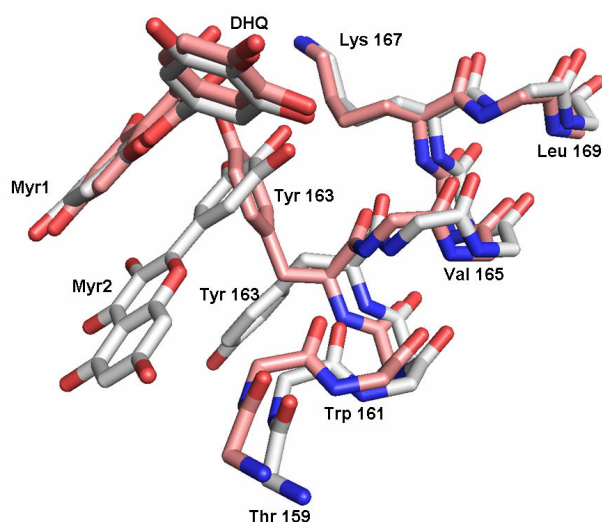


Figure 88 : Comparaison des domaines 150-169 (structure de référence en magenta)

La co-cristallisation de la DFR avec la myricétine dans des conditions équivalentes à celles utilisées pour le complexe de référence provoque la fixation d'une seconde molécule qui entraîne une altération de la géométrie du site actif. Ce résultat était inattendu et ne peut guère être expliqué à ce stade du travail.

L'analyse de la structure montre néanmoins que la modification du degré d'hydroxylation du cycle B ne modifie en rien la fixation de la première molécule de myricétine, cette dernière se superposant parfaitement à la DHQ. Pour le site actif, la géométrie du complexe ternaire n'est pas modifiée (à l'exception des perturbations structurales induites par la présence d'une seconde molécule de substrat). La présence d'un troisième groupement hydroxyle ne change pas la conformation du segment 131-156 qui, nous l'avons vu, délimite le site actif et participe fortement, via de nombreuses liaisons hydrogène, à la fixation et à l'orientation de la molécule de substrat. En absence de la deuxième molécule de substrat, qui perturbe l'organisation de la triade catalytique (Tyr 163 réorientée), rien ne semble s'opposer à l'activité catalytique optimale de l'enzyme.

III.2.4.DFR-NADP⁺-Quercétine

Tout comme dans le cas de la myricétine, deux molécules de substrats sont observées dès lors que la quercétine est co-cristallisée à l'enzyme.

III.2.4.1. Géométrie des substrats

La géométrie de la première molécule de quercétine se superpose elle aussi à la molécule de dihydroquercétine de référence. L'angle dièdre (O1-C2-C2'-C3') correspondant à l'angle formé entre le cycle B et le plan des cycles A-C est de -86°, valeur peu éloignée de celle observée (-66°) pour la dihydroquercétine. La position et l'orientation de cette première molécule de substrat sont en tout point similaires à celles trouvées précédemment. Les valeurs caractérisant les interactions substrat-protéine et substrat-cofacteur sont reportées dans le **Tableau 31**.

Substrat co-cristallisé à la DFR		Quercétine	DHQ
Interactions cycle B - DFR	O(C3')H... Oδ ₁ (Asn 133)	2,61 ± 0,17	2,54 ± 0,01
	O(C4')... HNδ ₂ (Asn 133)	3,07 ± 0,43	3,01 ± 0,06
	O(C4')H...Oδ ₁ (Asn 133)	2,65 ± 0,35	-
	O(C3')H...Oε ₁ (Gln 227)	2,76 ± 0,05	-
Géométrie du substrat	O1 -C2-C2'-C3'	-85,80 ± 7,24	-66,22 ± 0,2
	O1-C2-C3-C4	0,00	-59,90 ± 1,06
Transfert d'hydrure	C4 (substrat)-C4(NADP ⁺)-N1 (NADP ⁺)	88,81 ± 3,85	108,72 ± 2,84
	C4 (substrat)-C4(NADP ⁺)	3,29 ± 0,04	2,94 ± 0,08

Tableau 31 : Interactions Quercétine1-DFR

Les valeurs indiquées sont les valeurs moyennes des paramètres considérés pour chaque molécule.

La rotation du cycle B de la seconde molécule par rapport au plan moyen des cycles A-C, est de 25,5°. Cette désorientation est moindre que celle observée pour la première molécule de substrat, et plus importante que celle de la seconde molécule de myricétine.

Comme dans le cas de la myricétine, les deux molécules de substrat sont pratiquement parallèles entre elles, l'angle formé par les cycles A-C est égal à 19,1°. La distance inter planaire est comprise entre 3,2 et 3,5 Å (**Figure 89**). Les directions C6-C2 des deux molécules sont pratiquement colinéaires, l'angle formé par ces directions n'excède pas 5°. La deuxième molécule se déduit de la première par une rotation de 180° autour d'un axe parallèle à la direction C6-C2. Cette molécule est sensiblement à la même place que la seconde molécule de myricétine, son orientation diffère par contre de 90°. Enfin, la deuxième molécule de substrat est isolée du solvant par les parties hydrophobes des résidus suivants : Thr 86, Met 88, Tyr 163, Phe 164, Lys 167, Pro 204, His 219 et Ile 222. Elle interagit malgré tout avec l'enzyme via des liaisons hydrogène établies à partir des groupements hydroxyle du cycle B (**Tableau 32**).

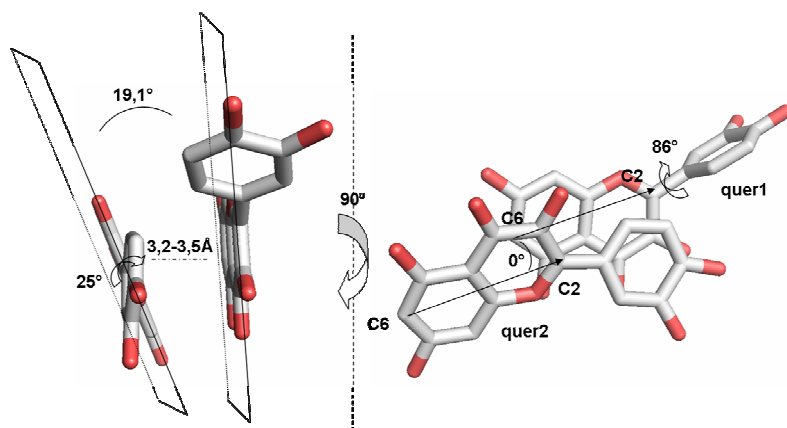


Figure 89 : Positions et interactions des deux molécules de quercétine

Substrat co-cristallisé à la DFR		Quercétine
Interactions cycle B - DFR	O(C3')H...O γ (Ser 128)	2,30 ± 0,04 Å
	O(C4')H...O γ (Ser 128)	2,41 ± 0,03 Å
Géométrie du substrat	O1 -C2-C2'-C3'	-24,26 ± 1,30°
	O1-C2-C3-C4	0,00°

Tableau 32 : Interactions de type liaisons hydrogène Quercétine2-DFR

Les valeurs indiquées sont les valeurs moyennes des paramètres considérés pour chaque monomère.

III.2.4.2. Modifications induites par la deuxième molécule de substrat

Les modifications induites par la seconde molécule de substrat sont équivalentes à celles observées dans la structure DFR-NADP⁺-Myricétine :

- ✓ la conformation du segment 87-106 est modifiée par des valeurs différentes des couples φ , ψ de certains résidus, notamment 88 à 92.
- ✓ les modifications observées pour le segment 150-169 ne concernent que des réorientations de chaînes latérales (résidus 156, 157 et 163).

Comme dans le cas de la myricétine, le seul résidu de la triade catalytique dont l'orientation est modifiée est la tyrosine 163. Sa réorientation est imposée par la présence de la deuxième molécule de substrat, qui génère un encombrement stérique et s'oppose à l'interaction entre le groupement hydroxyle de la tyrosine et la fonction carbonyle présente en C4 de la molécule 1.

En outre, la faible résolution du cristal ne permettant pas un positionnement précis des molécules de solvant, nous ne pouvons estimer « l'ouverture » de la DFR que par la distance séparant la chaîne latérale de la thréonine 159 de l'atome d'oxygène de l'alanine 218. Cette dernière, proche de 7 Å, confirme l'influence de la seconde molécule de substrat sur la réduction de l'effet tenaille observé.

Ces travaux ont, là encore, été effectués dans des conditions de co-cristallisation identiques à celles utilisées lors de l'étude du complexe ternaire de référence (5 molécules de substrat pour une molécule d'enzyme). Dans la configuration obtenue, la DFR ne peut être active de par l'introduction au voisinage du site actif d'une seconde molécule de substrat. Si l'on rejette l'idée que l'apparition de cette dernière molécule est associée à un artefact de cristallisation, on ne peut que s'interroger sur l'influence de l'insaturation présente entre les carbones 2 et 3 du cycle C vis-à-vis de la présence de cette deuxième molécule. L'aromaticité des composés et la distance existant entre les plans des cycles A-C laissent supposer l'existence d'interactions π - π entre les molécules de substrat induites par le recouvrement intermoléculaire des orbitales p des cycles conjugués. Ces interactions pourraient être à l'origine de la fixation de la seconde molécule de substrat.

CONCLUSIONS & PERSPECTIVES

CONCLUSIONS ET PERSPECTIVES

Les travaux présentés dans ce mémoire avaient pour but l'étude structurale de la dihydroflavonol 4-réductase de vigne et la caractérisation des mécanismes structure-fonction de cette enzyme qui, chez les plantes, est déterminante pour la biosynthèse de flavonoïdes d'intérêt que sont les anthocyanes et les tanins condensés. La principale difficulté de ce travail était l'obtention après purification d'une enzyme stable en solution. Rappelons qu'un tel résultat n'avait pu être obtenu en dépit de très nombreux travaux menés à partir tant d'extraits naturels de plante que de sur-expressions hétérologues. Le clonage du gène d'intérêt a été élaboré à partir d'une banque d'ADNc de baie de raisin, de cépage *Cabernet Sauvignon*, au stade véraison. L'optimisation des conditions d'expression chez *E. coli* nous a permis d'obtenir un rendement élevé favorable à la recherche des paramètres impliquant la stabilité de la DFR après purification.

Dans un premier chapitre, nous avons démontré l'activité de la protéine. Pour ce travail, nous avons optimisé un protocole déjà publié, reliant la réduction des dihydroflavonols à la production de pigments par oxydation chimique des leucoanthocyanidines. L'étude détaillée des composés métabolisés lors de cette dernière étape a conclu à l'existence d'un mélange de plusieurs complexes que nous avons identifiés par spectrométrie de masse. La mesure de l'activité de la DFR, basée sur l'existence d'un seul produit (cyanidine ou pélargonidine) était donc erronée. Pour nous affranchir de ce problème, nous avons montré qu'il fallait quantifier l'activité de la DFR par étude de la consommation de ses substrats.

Néanmoins, une caractérisation enzymatique précise ne peut être réalisée en utilisant le mode opératoire précédent. Nous nous sommes attachés à la mise au point d'un protocole expérimental capable de suivre l'activité enzymatique en fonction du temps. La méthode retenue ne pouvait pas être la spectrométrie UV, l'ensemble des produit (leucoanthocyanidine) et réactifs (NADPH, dihydroflavonols) absorbant dans cette gamme de longueurs d'onde. Nous avons opté pour la fluorimétrie. Les résultats obtenus laissent supposer une cinétique d'ordre 0 pour les deux substrats c'est-à-dire une vitesse initiale indépendante de la concentration en substrats.

De tels résultats traduiraient :

- soit une oxydation résiduelle du NADPH en dehors de toute activité enzymatique.
- soit une agrégation de l'enzyme dans les conditions expérimentales retenues. Dans ce cas précis, rappelons que la cinétique serait limitée par l'apport de substrat.

La première hypothèse ne peut être retenue, la vitesse initiale (14 μM de substrat par minute) traduisant l'existence d'une réelle activité enzymatique.

Différentes études peuvent être proposées pour contrôler la stabilité de l'enzyme dans un polytampon contenant du Bis-Tris et de la Tricine :

- ✓ Un contrôle de l'état d'oligomérisation de la DFR par diffusion de lumière.
- ✓ Une mesure de l'activité de l'enzyme après greffage de cette dernière sur support solide afin de s'affranchir de tout phénomène d'agrégation dans les conditions expérimentales retenues.

D'autres travaux doivent permettre de confirmer la viabilité de l'enzyme dans ce polytampon :

- ✓ Un contrôle de son activité par réduction totale des dihydroflavonols grâce à un système régénérateur de NADPH tel le système G_6PDH-G_6P (*Agarwal et al, 2005*).
- ✓ Un contrôle de l'influence de l'étiquette histidine en milieu acide sur la solubilité de l'enzyme par sur-expression de la DFR de Vigne sans étiquette.
- ✓ Une prise en compte des modifications post-traductionnelles (présentes dans tout organisme eucaryote) et hypothétiquement impliquées dans l'activité enzymatique. C'est la raison pour laquelle nous avons initié l'expression de l'enzyme chez la levure *Pichia pastoris*.

L'activité enzymatique trouvée, bien que partiellement caractérisée, tend à montrer que la protéine exprimée adopte un repliement compatible avec sa fonction biologique.

La faible homologie de séquence avec les autres SDRs de structure connue ne nous a pas permis de résoudre le problème du phasage par remplacement moléculaire. Nous avons choisi, comme méthode de résolution structurale, la diffusion anormale à longueurs d'onde multiples. Cette démarche a permis de décrire la première structure de DFR à 1,8 Å de résolution.

La géométrie de la triade catalytique étant identique à celle communément observée au sein de la famille des SDRs, le mécanisme réactionnel est parfaitement connu. Nous en avons rappelé les différentes étapes. Une interrogation subsiste malgré tout. Elle concerne l'origine du proton, qui, via la lysine 167, va reprotoner la chaîne latérale de la tyrosine 163. En effet, au sein du complexe ternaire, la chaîne latérale de la lysine est en contact uniquement avec quelques molécules d'eau piégées dans la poche catalytique et nullement accessibles au solvant. Le « proton relay » est impossible dans cette configuration. Deux hypothèses mécanistiques peuvent être soulevées :

- ✓ soit la re-protonation de la tyrosine s'effectue par un mécanisme spécifique à la DFR,
- ✓ soit la lysine est accessible au solvant dans une structure ouverte de l'apoenzyme, existante dès lors qu'il y a libération des produits (leucoanthocyanidines et NADP⁺).

Cette seconde hypothèse semble renforcée par l'existence d'une structure partiellement désordonnée de l'apoenzyme SSCR. Cette structure contraste fortement avec celles des complexes binaire et ternaire dans lesquelles la lysine est non exposée au solvant tout comme dans notre enzyme.

Si l'obtention d'un cristal d'apoDFR demeure à ce jour un échec, deux méthodes peuvent être envisagées pour contrôler expérimentalement soit l'accessibilité de la chaîne latérale de la lysine par les molécules de solvant, soit les changements conformationnels associés à l'activité catalytique. Ces méthodes, décrites ci-dessous, sont basées sur les propriétés de la diffraction des neutrons :

- ✓ Tout d'abord, l'inaccessibilité de la lysine au solvant dans le complexe ternaire peut être confirmée par diffraction des neutrons après dialyse contre une solution deutérée d'un volume de DFR incubé au préalable avec son substrat (DHQ) et son co-facteur (NADP⁺). La substitution de l'hydrogène par le deutérium, parce qu'elle réduit la diffusion incohérente des atomes d'hydrogène, permet la réduction du bruit de fond existant dans les expériences de diffraction neutronique. Une telle technique devrait permettre de vérifier l'accessibilité de la chaîne latérale de la lysine 167, qui pourrait être deutérée (*Meilleur et al, 2006*).
- ✓ Pour confirmer un possible réarrangement conformationnel, il est nécessaire d'analyser en fonction du temps la position atomique des hydrogènes présents sur les chaînes latérales des acides aminés. Pour des échelles de temps variant de la picoseconde à la nanoseconde, le mouvement des hydrogènes, répartis sur l'ensemble de la protéine, peut être assimilé aux modifications conformationnelles de celle-ci. La diffusion élastique incohérente des neutrons autorise une telle étude qui a été employée dans le cas de la dihydrofolate réductase de *E. coli* (*Tehei et al, 2006*).

La détermination du complexe ternaire nous a néanmoins permis la caractérisation des résidus et des interactions impliqués dans la fixation du substrat au sein du site actif. Afin d'étudier l'influence de la nature des résidus sur la fixation des substrats, nous avons surexprimé deux mutants, 1C et 4A. Les résultats enzymatiques effectués sur ces derniers témoignent d'une divergence importante entre les activités attendues et observées.

Le mutant 1C présente une activité globale moindre que celle de la DFR sauvage, et ce quelque soit le substrat métabolisé. Les rapports des activités entre simple mutant et DFR sauvage sont respectivement de 0,62 et 0,55 pour la DHQ et le DHK mais l'activité relative, $[DHQ]_{\text{réd.}} / [DHK]_{\text{réd.}}$, reste quasiment inchangée. La seule substitution de l'asparagine 133 n'a donc pas entraîné une modification significative de l'affinité, ce qui est contradictoire avec les résultats publiés dans la littérature. Rappelons que l'activité enzymatique mesurée sur cinq DFRs d'un même cluster (*Lotus japonicus*) avait permis de conclure (**Shimada et al, 2005**) en l'existence de deux familles d'enzymes se différenciant par la nature du résidu 133. Les conclusions émises étaient basées sur la seule nature du résidu 133 et ne prenaient pas en compte les différences de séquence observées au proche voisinage de ce résidu. Nos résultats semblent montrer que l'acide aminé 133, pris isolément, ne joue pas un rôle prépondérant dans la sélectivité de substrat bien qu'il soit fortement impliqué dans l'activité de l'enzyme. L'hypothèse que des acides aminés, autres que ce seul résidu, interviennent dans la régulation de l'activité catalytique est renforcée par les résultats obtenus sur les DFRs de *Cymbidium* et *Pétunia*. Ces deux plantes ne réduisent ni l'une ni l'autre le DHK (**Johnson et al, 1999 et 2001**) bien que la nature du résidu 133 (asparagine ou aspartate) ne soit pas identique.

L'activité du quadruple mutant est encore moindre. Les rapports des activités entre quadruple mutant et DFR sauvage sont respectivement de 0,08 et 0,33 pour la DHQ et le DHK. Les mutations 135, 138 et 142 ont donc réduit l'efficacité de l'enzyme dans son ensemble. Ces trois mutations ont néanmoins affecté significativement la spécificité de l'enzyme. La réduction du DHK est alors favorisée comme le montre les valeurs des rapports $[DHQ]_{\text{réd.}} / [DHK]_{\text{réd.}}$ égales à 0,35 et 1,52 pour respectivement le simple et le quadruple mutant. Ce dernier travail de mutagenèse dirigée tend à montrer que la vigne possède un caractère de type « Pétunia ». On peut donc supposer que l'activité mesurée vis-à-vis du DHK est fortement réduite en dépit de la présence d'une asparagine en position 133.

L'obtention des cristaux des mutants 1C et 4A est primordiale pour comprendre l'influence des mutations 135, 138 et 142 sur la fixation des substrats au sein du site actif. Néanmoins, à ce jour :

- ✓ nous n'avons pas pu obtenir de cristaux du mutant 1C en dépit de nombreux tests de différentes conditions de cristallisation
- ✓ la qualité cristalline du mutant 4A reste à optimiser, la résolution actuelle n'excédant pas 3,6 Å.

Les difficultés rencontrées n'étaient pas attendues. Les mutations affectant des résidus enfouis au sein de l'enzyme, nous pensons que les conditions optimales de cristallisation seraient voisines de celles de l'enzyme sauvage. Les résidus mutés conduisant à une réduction de l'activité enzymatique, nous ne pouvons exclure qu'ils soient responsables, par un mécanisme restant à expliquer, d'une moindre affinité du substrat au site actif. D'un point de vue structural, cela se traduirait par une diminution du facteur d'occupation du site actif impliquant au sein du cristal, la présence de molécules d'enzymes non complexées, entités que nous savons difficiles à cristalliser.

Diverses tentatives de cristallisation de ces mutants sont en cours d'expérimentation :

- ✓ une incorporation ordonnée des substrats à l'enzyme avant toute tentative de cristallisation,
- ✓ une dilution plus importante des composés polyphénoliques lors de leur incubation à la DFR. Cette dilution limiterait les associations aspécifiques entre polyphénols et molécules de protéine dictées par de nombreuses interactions de type π et/ou hydrophobe (*Pascal et al, 2006*). Elle permettrait par conséquent de réduire la turbidité observée dès lors que nous mettons en solution la DFR avec ses substrats. L'apparition de cette turbidité est en effet contraignante pour la recherche des conditions de cristallisation. Elle gêne une distribution homogène de la solution protéique par le robot de cristallisation, affectant ainsi la qualité du mélange réalisé avec la solution du réservoir.

A la suite de cette première étude, différents mutants sont en cours de clonage au laboratoire pour préciser le mécanisme de reconnaissance des substrats. Parmi ceux-ci, citons les mutations N133L et N133A permettant d'estimer les influences de polarité et d'encombrement stérique sur la fixation des substrats. D'autres mutations devront être étudiées. Elles affecteront les résidus du segment 131-156 qui, nous l'avons montré, bordent le site actif et influent sur l'activité enzymatique.

Le rôle des interactions entre DFR et groupements hydroxyle du cycle B du substrat, a, quant à lui, été estimé lors de l'analyse des résultats structuraux décrits dans la troisième partie du manuscrit. Plusieurs substrats ou analogues de substrats, différents entre eux par le nombre et la position de groupements hydroxyle du cycle B, ont été incubés avec la DFR.

Plusieurs cristaux de qualité satisfaisante ont été obtenus. Parmi ceux-ci, un cristal s'est avéré essentiel à l'avancée de notre problématique, celui de la protéine incubée avec le NADP⁺ et la myricétine. Cet analogue de substrat porte trois groupements hydroxyle en position 3', 4' et 5' du cycle B. La structure affinée montre que deux molécules de myricétine se fixent à chaque molécule de DFR. La première est à la même place que celle trouvée pour la DHQ dans la structure de référence. Elle adopte en outre la même orientation au sein du site actif. Ceci nous a permis de conclure que le nombre de groupements hydroxyle présents sur le cycle B ne modifiait pas la géométrie de fixation du substrat. Il est donc improbable que la spécificité de substrat soit uniquement régie par les seules interactions existantes entre le cycle B et le résidu 133. Cette conclusion pourra être confirmée dès lors que nous décrirons une structure avec comme substrat, le DHK ou la DHM. Les constatations précédentes montrent en outre que le mécanisme réactionnel diffère de celui proposé pour l'isocitrate déshydrogénase dans le chapitre I où une modification de la nature du substrat entraînait une réorientation de celui-ci, et par là même une activité différente.

La présence d'une seconde molécule de substrat est, quant à elle, potentiellement induite par la planéité du cycle C. La perte de la conformation demi-chaise semble responsable de la formation d'interactions de type π entre les deux molécules. Elle est également observée avec un second analogue de substrat plan : la quercétine. De par son encombrement stérique, cette seconde molécule empêche toute activité enzymatique. Elle implique une rotation de la chaîne latérale de la tyrosine 163, résidu de la triade catalytique.

L'analyse des structures de la troisième partie du manuscrit montre en outre que les flavonols peuvent constituer des inhibiteurs de la DFR, grâce à la structure plane des cycles A-C et leur faculté d'auto association. Ce résultat renforce les conclusions obtenues lors de l'étude de l'aldose reductase humaine (*Harrison et al, 1997*) montrant que l'association de composés aromatiques par « empilement » π peut inhiber l'activité enzymatique.

A ce stade du travail, les relations structure-fonctions de la DFR ne sont pas totalement élucidées. Les travaux entrepris doivent être poursuivis afin de déterminer les résidus impliqués dans la synthèse privilégiée d'un flavan-3,4-diol donné, composé qui est, rappelons-le, à la base de la diversité des anthocyanes et des tanins condensés. A terme, la simple lecture d'une séquence primaire de DFR permettrait d'appréhender l'activité de l'enzyme. De la lecture des séquences de DFR de vigne associées à différents cépages et déposées dans la banque de données TIGR, il serait possible de définir certains croisements privilégiés.

La détermination des structures 3D des enzymes impliquées dans la voie de biosynthèse des flavonoïdes ouvre en outre une voie de recherche plus appliquée, consistant à mettre au point des outils de détection de ces enzymes à différents stades du développement de la baie. De tels outils permettraient au viticulteur un suivi simple et rapide de l'expression de ces enzymes (voir à plus long terme, leur maîtrise, dès lors que les mécanismes d'action sont connus). La connaissance de la structure tridimensionnelle du site actif autorise par exemple la mise au point d'analogues de substrat spécifiques marqués.

MATERIEL & METHODES

MATERIEL ET METHODES

I. Matériel biologique

I.1. Banque d'ADNc

Les ARN totaux ont été extraits de grappes de raisins de cépage *Cabernet sauvignon* et une banque d'ADNc a été produite par le laboratoire de Biologie de la Vigne (UMR 619, INRA - Bordeaux).

I.2. Souches bactériennes *Escherichia coli*

- **XL1 blue**

Son phénotype est : recA1 nda1 gyrA96 thi-1 hsdR17 supE44 relA1 lac[F proAB lacIqZΔM15 Tn10(Tetr)]

La bactérie permet, lorsqu'elle contient un vecteur répliatif, une synthèse importante du plasmide et de l'ADN d'intérêt.

- **M15 [pRep4]**

Son phénotype est : Nal^S Str^S Rif^S Thi⁻ Lac⁻ Ara⁺ Gal⁺ Mtl⁻ F⁻ RecA⁺ Uvr⁺ Lon⁺

Elle permet une expression régulée grâce au vecteur de « répression » pRep4. Cette bactérie contient le gène de résistance à la kanamycine

Toutes les souches bactériennes sont cultivées à 37°C dans du milieu LB (Luria-Bertani) contenant 1 % (m/v) de bactotryptone, 0,5 % (m/v) d'extrait de levure, 1 % (m/v) de NaCl en présence ou non d'antibiotique approprié.

I.3. Souche de levure *Pichia pastoris*

- **GS115** : his4, phénotype mut⁺.

Pichia pastoris est une levure méthylotrophe. La souche GS115 possède une mutation sur le gène de l'histidinol déshydrogénase qui bloque toute synthèse d'histidine et par conséquent, permet le criblage des transformants sur milieu sélectif.

Les levures sont cultivées à 30°C soit dans un milieu complet YPD (Yeast extract – Peptone – Dextrose), soit dans des milieux minimum. L'YPD est composé d'extraits de levure à 1 % (m/v) et de bactopeptone à 2 % (m/v) auxquels est ajoutée une source de carbone, du glucose à 2 % (m/v).

Composition des milieux minimum :

- ✓ Pour l'inoculation : « MM : *Minimal Methanol* » composé d'YNB à 1,34 %, de biotine à 4.10^{-5} % et de méthanol à 0,5 %.
- ✓ Pour la pré-culture : « BMGY : *Buffered Glycerol-complex Medium* », composé d'extrait de levure à 1 %, de peptone à 2 %, d'YNB à 1,34 %, de biotine à 4.10^{-5} %, de glycérol à 1 % et tamponné par une solution de potassium phosphate, pH 6,0 à 100 mM.
- ✓ Pour l'induction de l'expression de la protéine d'intérêt : « BMMY » identique au BMGY mais où le glycérol est substitué par du méthanol à 0,5 %.

Le milieu « YNB : *Yeast nitrogen base with ammonium sulfate without amino acids* » (Invitrogen), utilisé pour la préculture et pour l'induction, est un mélange défini de sels, vitamines et de source azotée favorable à la croissance des levures.

II. Biologie Moléculaire

II.1. Amplification génique ou PCR (“polymerase chain reaction”)

II.1.1. Choix des amorces

Les amorces choisies s'hybrident avec les extrémités 5' et 3' de la séquence à amplifier. Elles sont déterminées à partir des séquences présentes dans les banques de données génomiques (Tigr, Swiss-prot).

- ✓ http://www.tigr.org/tigr-scripts/tgi/T_index.cgi?species=grape
- ✓ <http://www.expasy.ch/sprot/>

Des sites de restriction ont été ajoutés à chaque extrémité de ces amorces afin de pouvoir lier ultérieurement la séquence codante dans le vecteur d'expression. Ces sites sont rappelés dans le **Tableau 33**. Les amorces ont été synthétisées par Proligo.

Amorce	Séquence 5' - 3'	Site de restriction inséré	T _m (°C)	% (G+C)
DFR sens (<i>E. coli</i>)	TAAT aggcct ATGGGTTACAAAGTGAAACCG	Stu I	62	44
DFR anti sens (<i>E. coli</i>)	TAAT aagctt CTAGGTCTTGCCATCTAC	Hind III	57	39
DFR sens (<i>P. pastoris</i>)	TAAT ctcgag ATGGGTTACAAAGTGAAACC	Xho I	60	42
DFR anti sens (<i>P. pastoris</i>)	TAAT ggatcc CTAGGTCTTGCCATCTACAGG	BamH I	63	48

Tableau 33 : Récapitulatif des amorces utilisées

Les bases relatives à chaque site de restriction sont mentionnées en minuscule (caractères gras). Les quatre bases en 5' ont été ajoutées afin de ne pas initier la réplication au niveau du site de restriction.

II.1.2. ADN polymérase

La Pfu ADN polymérase issue de *Pyrococcus furiosus* (Promega) a été utilisée pour amplifier les fragments d'ADN car elle réalise une synthèse de haute fidélité. Pour les PCR de contrôle, l'enzyme utilisée est une Taq polymérase, de *Thermus aquaticus*, produite à l'UMR 619, Physiologie et Biotechnologie Végétales (Domaine de la Grande Ferrade, Villenave d'Ornon). Elle réalise une amplification moins fidèle.

La réaction de PCR (*Saiki et al, 1985*) est réalisée dans un tampon d'incubation contenant 20 mM Tris-HCl pH 8,8, 10 mM KCl, 10 mM (NH₄)₂SO₄, 2 mM MgSO₄, 0,1 mg/ml BSA, 0,1 % Triton X-100 (v/v) (Promega) pour la Pfu polymérase et 50 mM Tris-HCl pH 8,3, BSA 0,5 mg/ml, ficoll 0,5 %, saccharose 1 %, KCl 30 mM, MgCl₂ 3 mM et tartrazine 1mM pour la Taq polymérase.

II.1.3. Conditions de PCR

Les amplifications sont réalisées dans un thermocycleur iCycler (Biorad). Après une étape de 3 min à 95°C, le mélange réactionnel est soumis à 35 cycles d'amplification comprenant une étape de dénaturation des matrices de 30 s à 95°C, une étape d'hybridation des amorces de 30 s à une température variable selon le couple d'amorces utilisé et une étape de synthèse par amplification des amorces de 1 min à 72°C. Le cycle se termine par une étape de 3 min à 72°C.

Le mélange réactionnel comprend (dans le tampon d'incubation) 250 µM de chaque dNTP, 250 µM de chaque amorce, 10 à 50 ng d'ADN matriciel et 1 unité d'ADN polymérase dans un volume final de 50 µl.

II.1.4. Purification des produits de PCR

La purification des produits de PCR permet d'éliminer les sels et les autres réactifs présents dans le mélange d'amplification. Cette étape est réalisée à l'aide du kit « PCR Product Purification » (Qiagen) basé sur l'adsorption de l'ADN sur une colonne de silice.

II.2. Clonage moléculaire

La première étape du clonage moléculaire consiste à amplifier le fragment d'intérêt par PCR avec des amorces permettant l'insertion des sites de restriction appropriés pour l'intégration dans le vecteur d'expression choisi. Après purification, le fragment amplifié est sous-cloné, c'est-à-dire inséré dans un premier vecteur, le pGEM-T Easy, et sa séquence est contrôlée par séquençage.

La deuxième étape consiste en la digestion du pGEM-T Easy par des endonucléases spécifiques reconnaissant les sites de restriction ajoutés lors de l'étape d'amplification. Cette étape permet de libérer le fragment d'intérêt qui est ensuite ligaturé dans le vecteur final préalablement linéarisé.

II.2.1. Sélection des plasmides

- ✓ **pGEM-T Easy** (Promega) : le plasmide pGEM-T est utilisé comme vecteur de sous-clonage de séquences amplifiées par la technique de PCR. Ce vecteur est initialement sous forme ouverte (coupé par EcoRV) et additionné de thymidine aux extrémités 3' terminales. Le multisite de clonage est situé entre les promoteurs T7 et SP6 des bactériophages correspondants, dans le fragment d'ADN codant pour le peptide α de la β -galactosidase. L'inactivation du peptide α par insertion de l'insert permet la discrimination rapide des clones positifs.
- ✓ **pQE30-Xa** (Qiagen) : ce vecteur appartient à la famille des plasmides de type pDS et a été optimisé afin de permettre une forte expression de la protéine d'intérêt. Pour cela, une séquence promotrice du phage T5 et un site de fixation ribosomique de type RBSII ont été ajoutés. Pour faciliter la purification, une séquence codante pour une étiquette composée de six histidines est présente à l'extrémité N-terminale du multisite de clonage. Un site de reconnaissance de la protéase Xa (IEGR), intercalé entre l'étiquette histidine et la protéine traduite, permet la coupure de l'étiquette libérant ainsi une protéine « sauvage ». Le gène *bla* codant pour la β -lactamase lui confère la résistance à l'ampicilline.

- ✓ **pHIL-S1** (Invitrogen) : ce vecteur permet la sécrétion de la protéine dans le milieu de culture grâce à la séquence PHO1. L'expression est contrôlée grâce au gène AOX1 codant pour une alcool oxydase qui, en présence de méthanol, permet une synthèse importante de la protéine d'intérêt.

Les cartes des vecteurs utilisés au cours de l'étude sont représentées sur la **Figure 90**.

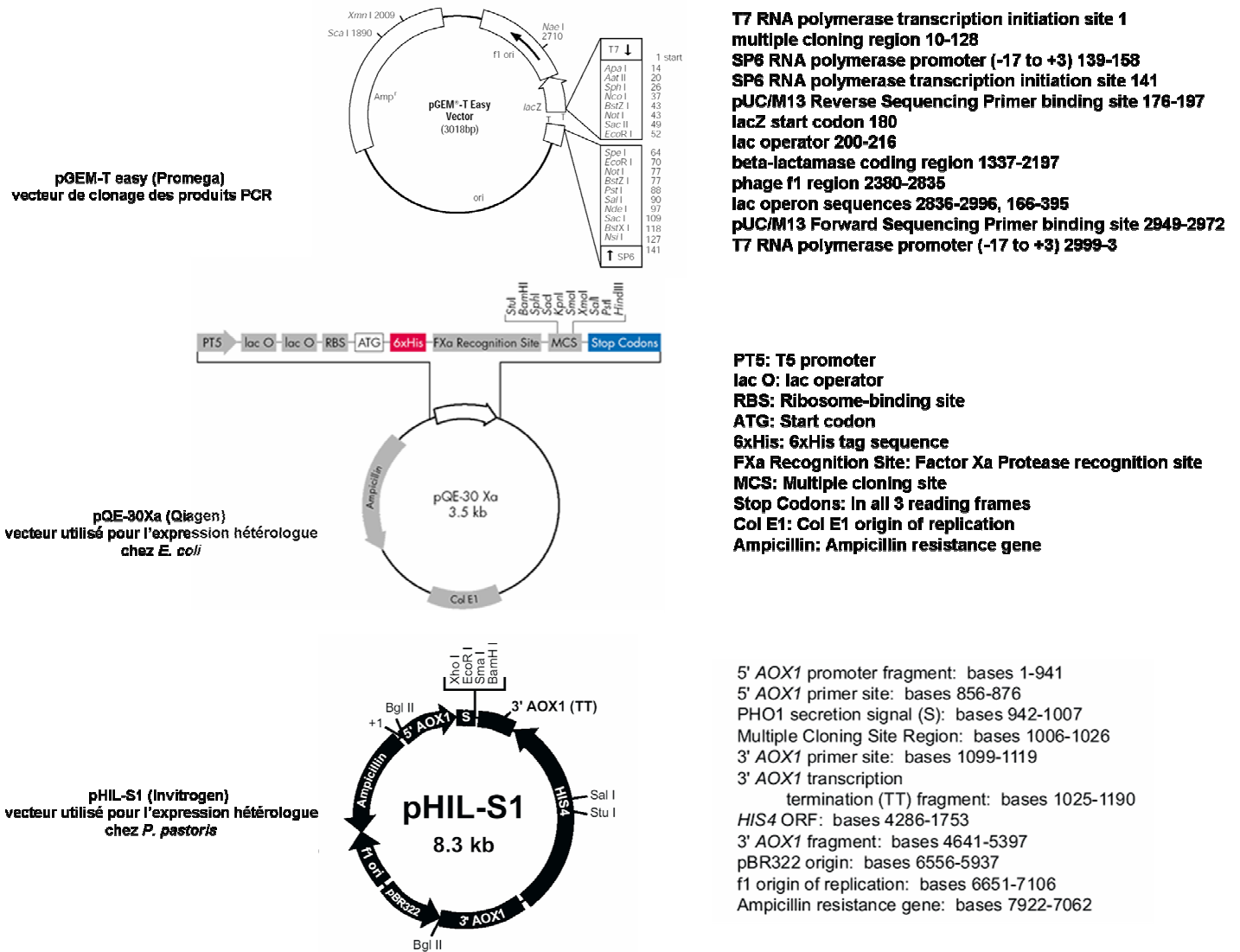


Figure 90 : Cartes des plasmides

II.2.2. Ajout d'une déoxyadénosine en 3' du fragment de PCR

La Pfu polymérase n'ajoute pas de déoxyadénosine en 3' lors de l'amplification du fragment. Or, il est nécessaire d'ajouter ce nucléotide afin de permettre la ligature de l'insert dans pGEM-T Easy. Pour cela, un mélange composé du produit de PCR purifié (7 µl), du tampon (10X) réactionnel de la Taq polymérase, de dATP à 0,2 mM et de 5 unités de Taq polymérase est incubé à 70°C pendant 30 minutes.

II.2.3. Quantification des fragments d'ADN

Tous les produits obtenus sont quantifiés et leur qualité est évaluée à l'aide d'un spectrophotomètre GeneQuant 8000 (Amersham) à 260 nm. Des mesures d'absorbance, à 230 et 280 nm sont effectuées pour évaluer respectivement les présences de sucres et de protéines dans les extraits. Un aliquote d'ADN est considéré comme pur pour des rapports 260/230 supérieur à 2 et 260/280 supérieur ou égal à 1,8.

II.2.4. Ligature de fragment de PCR dans un vecteur

Le rapport molaire vecteur/insert utilisé est de 1/3. Pour un volume réactionnel de 20 µl, 30 fmoles de vecteur et 90 fmoles d'insert sont mélangées avec 5 unités d'ADN T4 ligase (Promega) et incubés pendant 16 heures à 4°C dans le tampon préconisé (30 mM Tris-HCl pH 7,8, 10 mM MgCl₂, 10 mM DTT, 1 mM ATP).

II.2.5. Digestion par des endonucléases de restriction

Les digestions sont réalisées pour libérer l'insert à l'intérieur du vecteur de sous-clonage. Elles sont effectuées dans un volume total de 10 µl en présence d'1/10^{ème} du tampon 10 X préconisé par le fournisseur et de 2 unités d'endonucléase permettant de digérer un site de restriction donné (New England Biolabs).

II.2.6. Séquençage des fragments d'ADN

Les ADN sont séquencés en double brin par la méthode de Sanger (*Sanger & Coulson, 1975*) à l'aide des amorces utilisées pour la réaction de PCR lorsqu'il s'agit d'ADN amplifié ou à l'aide d'amorces universelles pour les ADN insérés dans des plasmides (prestation de service par la société Génome express).

II.3. Analyse de l'ADN par électrophorèse sur gel d'agarose

Les produits d'amplification (ainsi que les produits de digestion) sont analysés par électrophorèse sur gel d'agarose (*Sambrook et al, 1989*). Les gels sont composés d'un mélange de tampon de migration TAE (40 mM Tris, 20 mM acétate de sodium pH 8,3, 1 mM EDTA), de 1,0 % d'agarose et de bromure d'éthidium afin de visualiser l'ADN sous UV.

Un microlitre de tampon de charge (0,05 % bleu de bromophénol, 10 % glycérol) est ajouté à 10 µl d'échantillon et le mélange est déposé dans le gel d'agarose. Après migration dans le tampon TAE avec un champ électrique de 100 V, les fragments d'ADN sont visualisés sous lumière ultra-violette à 320 nm. Les gels sont numérisés à l'aide d'un appareil d'acquisition d'image (Gel Doc, Biorad) et analysés par le logiciel Gel Doc Quantity One (Biorad). Un marqueur de masse moléculaire (100 pb ou 1kb Ladder – Promega) permet de déterminer la taille des fragments d'ADN.

II.4. Extraction d'ADN plasmidique

L'ADN plasmidique est extrait des cellules bactériennes en utilisant le kit « QIAprep PCR » purification kit (Qiagen) selon les recommandations du fournisseur. La culture bactérienne (2 à 5 ml) est sédimentée et le culot de cellules est soumis à l'extraction. L'ADN est élué dans 50 µl d'eau.

II.5. Transformation de bactéries chimio-compétentes

Le vecteur contenant l'insert d'intérêt est ajouté dans un volume ne dépassant pas 5 % de celui des cellules compétentes et sa quantité est limitée à 50 ng pour 100 µl de bactéries. Après un temps de latence dans la glace de 30 minutes, le mélange subit un choc thermique à 42°C pendant 90 secondes. Puis les bactéries sont transférées dans la glace pendant 2 minutes avant ajout de 800 µl de milieu LB. Le mélange est incubé dans une étuve à 37°C sous agitation lente pendant 1 h. Une fraction variable (suivant l'efficacité de transformation) sera étalée sur milieu sélectif et incubé à 37°C pendant 12 h.

II.6. Transformation de *Pichia pastoris*

Le protocole de traitement des cellules et d'électroporation utilisé est identique à celui décrit par *Wu et al, (2004)*. La pré-culture est reprise à température ambiante dans une solution composée de 100 mM de LiAc, 10 mM de DTT, 0,6 M de sorbitol et 10 mM de Tris-HCl pH 7,5 puis lavée à trois reprises dans une solution glacée de sorbitol à 1 M dont le volume est calculé afin d'obtenir une concentration finale de cellules de 10^{10} / ml. Les cellules sont mélangées à 3 ng d'ADN puis subissent un choc électrique de 1,5 kV, 25 µF, 186 Ω à l'aide d'un électroporateur (Biorad Genepulser). Les cellules sont ensuite diluées dans 1 ml de sorbitol à 1 M et mises en culture sur milieu RDB (1 M sorbitol, 1 % dextrose, 4×10^{-5} % biotine, 1,34 % YNB et 2 % agar).

II.7. Mutagenèse dirigée

Le protocole utilisé pour l'obtention de mutants est celui décrit dans le kit « Quikchange Site-Directed Mutagenesis » (Stratagene). Ce protocole permet l'incorporation de mutations par PCR.

Le **Tableau 34** détaille les amorces utilisées pour la synthèse des mutants.

Amorce	Séquence 5'- 3'	Tm (°C)	%(G+C)
DFR-1C sens	TCCGCAGGAACTGTGgatATTCAAGAACACCAA	63	45
DFR-1C anti sens	TTGGTGTTCTTGAATatcCACAGTTCCTGCGGA	63	45
DFR-4A sens	TCCGCAGGAACTGTGgatATTactGAACACcctCTG CCAGTGataGATGAAAGCTGCTGG	74	50
DFR-4A anti sens	CCAGCAGCTTTCATCtatCACTGGCAGaggGTGTT CagtAATatcCACAGTTCCTGCGGA	74	50

Tableau 34 : amorces utilisées pour la synthèse des mutants

Les conditions expérimentales de PCR pour amplifier l'insert associé au plasmide sont similaires à celles décrites précédemment à l'exception de la polymérase qui est une Turbo Pfu (Stratagene). Le cycle commence par une étape de dénaturation de 1 min à 95°C suivie de 12 cycles comprenant une étape de dénaturation de 30 s à 95°C, une étape d'hybridation de 1 min à 55°C et une étape d'élongation de 5 min à 68°C.

L'ADN matriciel méthylé est ensuite digéré pendant 1 heure à 37°C par l'ajout d'1 unité d'endonucléase Dpn1. L'ADN mutant est ensuite inséré par transformation dans la souche M15 [pREP4] et contrôlé par séquençage. La **Figure 91** résume les différentes étapes permettant l'obtention de mutants.

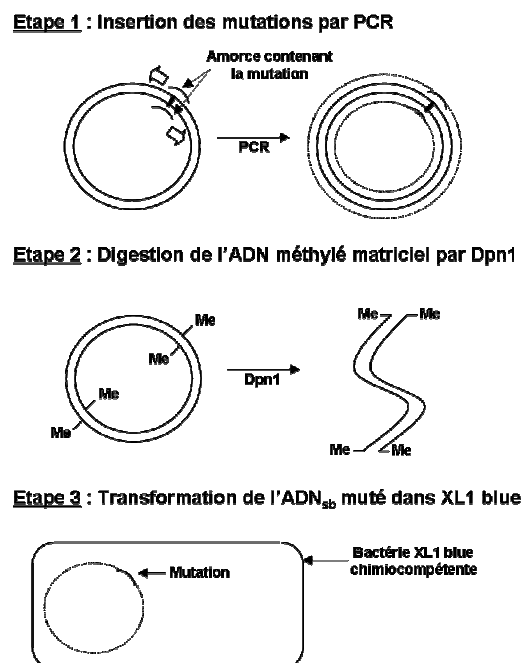


Figure 91 : Mutagenèse dirigée

III. Méthodes d'expression et d'analyse des protéines

III.1. Expression des protéines chez *E. coli*

Une colonie de bactérie transformée est mise en culture à 37°C sous agitation (200 rpm) dans le milieu LB contenant les antibiotiques adéquats jusqu'à l'obtention d'une DO_{600nm} de 0,5 à 0,7. L'expression de la protéine est induite par l'ajout d'IPTG à une concentration finale de 1 mM et maintenue pendant 12 heures à 30°C sous agitation.

III.2. Marquage de protéine au Sélénium

La méthode expérimentale utilisée permet l'incorporation de méthionine sélénée sans le recours à l'utilisation d'une souche d'*E. coli* auxotrophe. La préculture est effectuée 12h à 37°C sous agitation (200 rpm) dans le milieu minimum M9 (50 mM NaH₂PO₄, 20 mM KH₂PO₄, 8,5 mM NaCl, 20 mM NH₄Cl, 2 mM MgSO₄, 0,4 % glucose (m/v), 0,1 mM CaCl₂) auquel sont ajoutés les antibiotiques nécessaires pour maintenir la pression de sélection. La culture (milieu M9 + 15 mM vitamine B1) est inoculée au 1/50^{ème} et maintenue sous agitation jusqu'à l'obtention d'une DO_{600nm} de 0,5. Sont ajoutés stérilement au milieu 0,75 mM de lysine, thréonine et phénylalanine, 0,4 mM de leucine, isoleucine, valine et 0,25 mM de sélénio-méthionine. Au bout de 15 minutes, 1mM d'IPTG est ajouté pour initier l'expression et celle-ci est maintenue 12 heures à 30°C.

III.3. Préparation d'extraits cellulaires d'*E. coli*

Les cellules sont récoltées par centrifugation à 5000 g, pendant 20 min à 4°C puis resuspendues dans du tampon de lyse (50 mM NaH₂PO₄, 300 mM NaCl, 10 mM imidazole pH 8,0). L'ajout de lysozyme à 1 mg/ml et des périodes de sonication (Branson digital sonifier 250) de 10 s, puissance 30 %, pendant une durée totale de 3 minutes permettent la lyse bactérienne. Le surnageant, contenant la fraction cytoplasmique, est obtenu après centrifugation à 16000 g pendant 15 min.

III.4. Electrophorèse monodimensionnelle en conditions dénaturantes (SDS-PAGE)

L'échantillon à analyser est mélangé à un tampon de charge composé de 62,5 mM Tris-HCl pH 6,8, 10 % glycérol (v/v), 2 % (m/v) SDS, 1,25x10⁻² % (m/v) bleu de bromophénol et de 5 % (v/v) de β-mercaptoéthanol. Les protéines sont dénaturées et réduites par chauffage à 100°C pendant 5 min.

Les échantillons de protéines et le marqueur de masse moléculaire (LMW marker, Amersham) sont déposés sur un gel de polyacrylamide-SDS 10 % (m/v) vertical. Ce gel est composé d'un gel de séparation (10 % (m/v) acrylamide / bis-acrylamide (37,5:1), 0,375 M Tris-HCl, 0,1 % SDS (m/v), 0,1 % persulfate d'ammonium (m/v), 0,05 % TEMED (m/v), pH 8,8) surmonté d'un gel de concentration (4 % (m/v) acrylamide / bis-acrylamide (37,5:1), 0,125 M Tris-HCl, 0,1 % SDS (m/v), 0,1 % persulfate d'ammonium (m/v), 0,05 % TEMED (m/v), pH 6,8).

Le tampon d'électrophorèse est composé de 25 mM Tris-HCl, de 192 mM glycine et de 0,1 % SDS (m/v) pH 8,3. Un courant électrique de 35 mA est appliqué au gel durant 1 heure afin d'obtenir une séparation optimale des protéines contenues dans l'échantillon.

Enfin, la coloration des gels est effectuée pendant 1 heure dans une solution de coloration (45 % méthanol (v/v), 9 % acide acétique (v/v), 0,04 % bleu de coomassie R250 (m/v)) et leur révélation est effectuée durant la nuit dans la solution suivante : 30 % méthanol (v/v), 5 % acide acétique (v/v).

III.5. Electrophorèse sur gel d'acrylamide en conditions non dénaturantes (PAGE)

Le tampon de charge est composé de 84 mM Tris-HCl pH 6,8, 13 % (v/v) glycérol, $1,6 \cdot 10^{-3}$ % m/v de bleu de bromophénol et est mélangé à l'échantillon avec un volume équivalent.

Le mélange est déposé sur gel de polyacrylamide 10 % (m/v). Le gel de séparation est composé d'acrylamide / bis-acrylamide 10 % (m/v), 0,187 M Tris-HCl, 0,05 % persulfate d'ammonium (m/v), 0,05 % TEMED (m/v), pH 8,8 et le gel de concentration d'acrylamide / bis-acrylamide 8 % (m/v), 0,125 M Tris-HCl, 0,1 % persulfate d'ammonium (m/v), 0,05 % TEMED (m/v), pH 6,8. Le tampon de migration est composé de 45 mM Tris-Base pH 8,4, 40 mM acide borique, et de 1,25 mM EDTA. Les conditions de migration, coloration, décoloration et identification sont similaires à celles utilisées pour les électrophorèses en conditions dénaturantes.

III.6. Purification des protéines recombinantes

III.6.1. Chromatographie d'affinité sur matrice agarose – acide iminodiacétique

La purification par chromatographie d'affinité utilisée est basée sur la sélectivité et l'affinité de la matrice agarose – acide iminodiacétique pour les biomolécules ayant un motif 6 histidines lorsque celle-ci est saturée en nickel. Quatre volumes de colonne (HiTrap Chelating HP Amersham) sont chargés avec du tampon 10 mM tricine pH 7,3, 500 mM NaCl, 5 mM imidazole avant l'injection du lysat bactérien. Une étape de rinçage avec 6 volumes de ce même tampon est réalisée après injection du lysat, permettant ainsi l'élimination des protéines aspécifiques. Enfin, 4 volumes d'un tampon identique à l'exception de la concentration en imidazole (300 mM) sont utilisés pour l'élution des protéines recombinantes.

III.6.2. Chromatographie d'exclusion stérique ou gel perméation

La colonne utilisée (Superdex 200 Amersham) permet la séparation des molécules de taille comprise entre 10^4 et 5.10^5 Da. L'appareil de chromatographie liquide (WATERS 2690 Separation Molecule) est calibré de manière à obtenir un débit de 0,5 ml/min. La quantification est effectuée à l'aide du détecteur UV, WATERS 286 Tunable Absorbance Detector. Le tampon utilisé est composé de 10 mM tricine pH 7,3, NaCl 500 mM. Pour connaître le temps de rétention de la protéine d'intérêt, 20 µg d'échantillon sont injectés et le profil d'élution est ensuite comparé à ceux de protéines standards injectées dans les mêmes conditions. Une fois le temps de rétention de la protéine recombinante déterminé, sa purification est réalisée (sous condition d'une stabilité convenable).

IV. Méthodes d'étude enzymatique

IV.1. Contrôles

IV.1.1. Séparation des polyphénols par chromatographie liquide à haute pression

La chromatographie effectuée en phase inverse permet de séparer divers composés polyphénoliques en fonction de leur polarité. Cette étape est effectuée sur une colonne de type Atlantis dC18 (Waters) avec un débit de 1ml /min. L'appareillage est constitué d'un module principal Waters 600E et d'un détecteur UV Photodiode Array Detector 996 programmé à 214 nm. Le gradient utilisé est décrit dans le **Tableau 35** ci-dessous :

	H ₂ O-0,1 % TFA	Acétonitrile-0,08 % TFA
t = 0	90 %	10 %
t = 20 min	10 %	90 %

Tableau 35 : Gradient utilisé pour la séparation des polyphénols

IV.1.2. Synthèse et caractérisation des leucoanthocyanidines

IV.1.2.1. Caractérisation par spectrophotométrie UV

Lors de l'étude de l'activité enzymatique par fluorimétrie, il a été nécessaire d'estimer l'influence d'une concentration donnée de leucoanthocyanidines sur la fluorescence du NADPH. Les leucoanthocyanidines étant instables dans le temps, il était nécessaire d'obtenir un produit synthétisable extemporanément. La voie de synthèse choisie est celle de la réduction totale de dihydroflavonols par un réducteur fort : le borohydrure de sodium (NaBH₄).

Une gamme d'équivalents (0 à 5) de NaBH₄ a été testée afin de réduire dans l'éthanol absolu la dihydroquercetine sans la participation d'un catalyseur enzymatique. Les résultats ont été exploités en UV avec un spectrophotomètre UV-visible SAFAS 190 DES entre 200 et 400 nm

IV.1.2.2. Caractérisation par obtention des proanthocyanidines

Le protocole utilisé est détaillé dans la littérature par *Dedaldechamp et al, 1999*. Ce dernier prouve la synthèse des leucoanthocyanidines par dosage des pigments obtenus par oxydation chimique de celles-ci. Après extraction des polyphénols dans 2 volumes d'acétate d'éthyle, le solvant organique est évaporé et les polyphénols redissous dans une solution de n-butanol/HCl 95/5 (v/v). Ce mélange est incubé 30 min à 90°C pour accélérer la formation des pigments (anthocyanes). Les pigments sont dosés par spectrométrie UV-visible à la longueur d'onde d'absorption maximale de ces derniers, 550 nm pour la cyanidine et 534 nm pour la pélargonidine.

IV.2. Activité et colorimétrie

L'activité de la DFR a tout d'abord été estimée par dosage des pigments. Le mélange réactionnel est composé de 1mM de substrat (2R-3R DHQ ou 2R-3R DHK) et de 1mM de NADPH dans une solution tampon composée de 50 mM Bis-Tris, 50 mM Tricine pH 6.5, 50 mM NaCl et 5 mM DTT. La réaction enzymatique est initiée par l'ajout de 1µM de DFR. Le mélange est incubé à 30°C pendant 30 minutes. L'activité est révélée par transformation chimique des leucoanthocyanidines en proanthocyanidines par la méthode décrite dans le paragraphe précédent IV.1.2.2.

IV.3. Enzymologie et fluorimétrie

L'activité réductrice de la DFR est quantifiée en mesurant la variation de fluorescence du NADPH. Les longueurs d'ondes d'excitation et d'émission utilisées sont respectivement 380 et 450 nm. Les fentes du spectrofluorimètre (LS 55, Perkin Elmer) sont ajustées en excitation et en émission à 2,5 et 10,5 nm respectivement. L'enzyme est conservée dans 10 mM tricine pH 7,3, 50 mM NaCl, 1 mM DTT, 50 % glycérol, et un aliquot est dialysé extemporanément contre 10 mM tricine pH 7,3, 50 mM NaCl et 1 mM DTT. L'ensemble des tests enzymatiques a été effectué avec 50 nM de DFR à température constante : 30°C. Le tampon d'activité est un polytampon composé de 50 mM Bis-Tris, 50 mM Tricine pH 6,5 et 50 mM NaCl.

V. Cristallisation

V.1. Cristallisation en goutte suspendue par diffusion de vapeur

La recherche de premières conditions de cristallisation a été effectuée en utilisant les « screenings » commerciaux avec un robot de cristallisation de type : Honeybee 961 (Cartesian). Les plaques utilisées sont de type Greiner 96 réservoirs et 3 puits satellites par réservoir. Le volume dispensé par le robot est de 300 nl par goutte (150 nl de protéine, 150 nl de solution de réservoir). Chaque réservoir est rempli par un volume de 100 μ l de solution de cristallisation.

L'optimisation des conditions de cristallisation est effectuée à 20°C sur boîte de culture de cellule de type Linbro (24 puits). Les puits sont remplis d'une solution appelée réservoir (500 μ l ou 1 ml). Une goutte composée pour moitié de la solution protéique (1 μ l) et pour moitié de la solution du réservoir (1 μ l) est suspendue sous lamelle de microscope scellée par de la graisse de silicone (**Figure 92**).

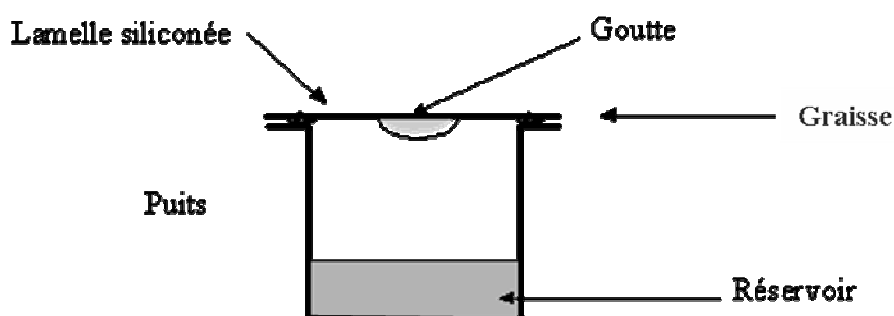


Figure 92 : Système de cristallisation par diffusion de vapeur en goutte suspendue (McPherson et al, 1990)

Par diffusion simple des espèces volatiles, l'équilibre s'établit entre la goutte et le réservoir jusqu'à ce que les concentrations en sels soient identiques dans la goutte et le réservoir. La protéine ne diffuse pas et se trouve concentrée. L'état de sursaturation s'établit donc progressivement par augmentation simultanée de la concentration de la protéine et de l'agent précipitant, conduisant à l'obtention de cristaux (**Figure 93**).

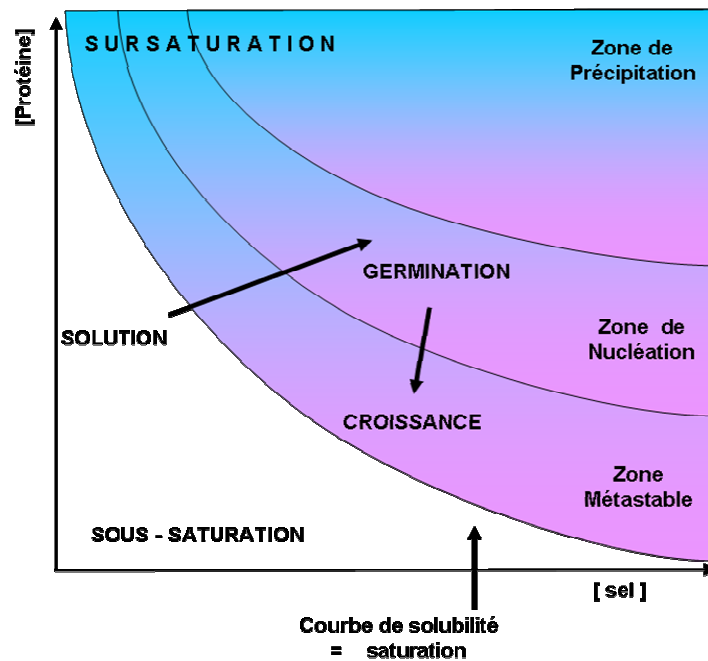


Figure 93 : Diagramme de phase et changement et évolution du degré de sursaturation

V.2. Cryo-conservation des cristaux

Les rayons X entraînent la formation de radicaux libres qui détruisent l'ordre cristallin. Dans ces conditions, les dommages causés peuvent conduire à l'arrêt prématuré de la diffraction par le cristal, ou du moins, à une réduction progressive de la limite de diffraction. Pour limiter ce phénomène, les données sont collectées sous un flux d'azote à 100 K. Avant toute collecte, les cristaux sont montés dans une boucle en nylon et trempés dans 4 gouttes de cryo-solution et ce, pendant 1 min pour chaque goutte. Les cryo-solutions sont composées de solution de réservoir à laquelle est ajoutée une concentration croissante en glycérol (5 %, 10 %, 15 % et 20 %) : le pourcentage le plus élevé en cryoprotectant empêchant la formation des cristaux de glace.

VI. Cristallographie

VI.1. Diffraction des Rayons X et traitement des données

Dés lors que les conditions de cristallisation permettent l'obtention de monocristaux exploitables en diffraction des rayons X, les procédures de détermination de structure peuvent être utilisées. Bien que les dimensions des cristaux de protéine soient le plus souvent inférieures à 0.5 mm, les informations structurales que sont susceptibles de livrer ces derniers sont essentielles. Cela nécessite d'enregistrer un jeu de données complet, c'est-à-dire de collecter l'ensemble des intensités diffractées dans un domaine angulaire donné (*Arndt & Wonacott, 1977*), dicté par la symétrie du cristal.

Les intensités diffractées sont proportionnelles aux facteurs de structure $\overline{F(\vec{s}^*)}$, traduisant l'amplitude de l'onde diffractée par l'ensemble des atomes de la maille. Elles sont par conséquent corrélées à la densité électronique des macromolécules dans le cristal par les relations suivantes :

$$\rho(r) = \int_S \overline{F(\vec{s})} \exp(-2i\pi\vec{s}\cdot\vec{r}) d\vec{s}$$
$$\overline{F(\vec{s})} = \left| \overline{F(\vec{s})} \right| \exp(i2\pi\alpha(\vec{s}))$$
$$\left| \overline{F(\vec{s})} \right| = c\sqrt{I(\vec{s})}$$

où $\overline{F(\vec{s})}$ et $\left| \overline{F(\vec{s})} \right|$ sont respectivement le facteur de structure et son module ; α : phase du facteur de structure, \vec{s} : vecteur de l'espace réciproque liant le nœud (h,k,l) à l'origine ; \vec{r} : vecteur du réseau cristallin définissant le point de coordonnées (x,y,z) dans la maille cristalline et S : le volume de l'espace réciproque collecté.

Bien que la qualité cristalline, c'est à dire le degré d'ordre de l'arrangement cristallin, soit le facteur essentiel qui détermine la qualité et la résolution d'un jeu de données, l'appareillage joue également un rôle essentiel. L'accès au rayonnement synchrotron, source intense de rayons X, permet l'étude de systèmes biologiques complexes (tel par exemple la sous-unité 50S du ribosome ; *Ban et al, 2000*). L'utilisation d'une telle source de rayonnement améliore de façon significative la résolution des données. La longueur d'onde expérimentale étant modifiable, cette source permet également la détermination des phases grâce à l'utilisation de méthodes dites de dispersion anormale à longueurs d'onde multiples. L'accès à un tel rayonnement est donc fortement recommandé.

L'utilisation de divers programmes, MARXDS (*Kabsch, 1988*), DENZO/SCALEPACK (*Otwinowski, 1997*), MOSFLM (*Leslie, 2001*) propose une liste de types de réseau, paramètres de maille et groupes d'espace potentiels et permet une évaluation semi-automatique des intensités diffractées. Ces programmes calculent l'orientation du cristal et intègrent les intensités diffractées en corrigeant ces dernières de différents effets tels que la polarisation du faisceau ou les effets d'irradiation. Après mise à l'échelle des différentes images, une liste finale d'intensités diffractées ($|F_o|^2 = I_o$) est obtenue.

VI.2. Méthodes de détermination des phases

Les intensités diffractées $I(hkl)$ ayant été mesurées, l'amplitude des facteurs de structure est déterminée par la relation $I(hkl) = |F(hkl)|^2$. Malheureusement, aucune information n'est connue en ce qui concerne les phases. Différentes techniques ont été développées pour résoudre ce problème. Certaines sont basées sur le remplacement isomorphe (SIR, SIRAS, MIR, MIRAS, MAD ; *Green et al, 1954 ; Rossmann, 1960*), les autres utilisent les structures 3D déjà déterminées de protéines semblables (Remplacement moléculaire ; *Hope, 1957 ; Rossmann et Blow, 1962 ; Rossmann, 1990*).

VI.2.1. Remplacement isomorphe (SIR/MIR)

La méthode du remplacement isomorphe multiple consiste à insérer un ou plusieurs atomes lourds (atomes dont le nombre d'électrons est élevé) dans le cristal de la protéine native puis à résoudre le problème des phases à partir de la contribution de ces seuls atomes. Ceci est obtenu de différentes façons : par trempage du cristal de protéine native dans des solutions de sels d'atomes lourds, par co-cristallisation, par exposition du cristal à une atmosphère de gaz noble ou par substitution des méthionines par des sélénométhionines. La difficulté de la méthode réside dans l'obtention d'un cristal du "dérivé lourd" le plus isomorphe possible du cristal natif. Dans ce cas, une carte dite de Patterson différence peut être obtenue.

L'expression suivante (dans laquelle $\overline{F_p}, \overline{F_{pH}}, \overline{F_H}$ sont respectivement les facteurs de structure de la protéine native, du dérivé lourd et de l'atome lourd) représente les coefficients utilisés dans la fonction de Patterson différence, révélant la structure de l'atome lourd :

$$F_H^2 \approx F_p^2 + F_{pH}^2 - 2F_{pH}F_p = (F_{pH} - F_p)^2$$

Le tracé et l'interprétation des différentes sections de cette carte (**Harker, 1956**) permettent d'obtenir les positions de l'atome lourd. Une fois connues et affinées ces positions, les phases de la protéine peuvent être calculées dans le cas d'un seul dérivé lourd (SIR) :

$$\rho_p = \rho_H + \cos^{-1} \left[\left(F_{pH}^2 - F_p^2 - F_H^2 \right) / 2F_pF_H \right]$$

La présence du terme cosinus dans l'expression ci-dessus implique que le problème des phases ne soit pas totalement résolu. Deux solutions, ρ_1 et ρ_2 , sont possibles. L'indétermination peut être levée de différentes manières :

- ✓ en enregistrant des données sur au moins deux dérivés lourds différents (MIR)
- ✓ en ajoutant une information structurale partielle issue du remplacement moléculaire (SIR-MR)
- ✓ par utilisation de données anormales si les collectes de données sont suffisamment précises (SIRAS, MIRAS).

VI.2.2. Diffusion anormale à longueurs d'onde multiples (MAD)

Si la protéine contient des diffuseurs anormaux, les différences d'intensité d'une paire de Bijvoet : $|F_h(+)|^2$ et $|F_h(-)|^2$ peuvent être exploitées pour déterminer les phases. La méthode utilisée est basée sur le fait que la diffusion anormale est fonction de la longueur d'onde. Différents jeux de données doivent être collectés à au moins trois longueurs d'onde distinctes sur un même cristal. Ces longueurs d'onde optimisent la différence d'intensités entre paires de Bijvoet. L'accès au rayonnement synchrotron, source de rayons X à longueurs d'onde variables, est alors indispensable. Wayne Hendrickson et ses collègues furent les premiers à tirer avantage de cette méthode pour résoudre la structure d'une protéine (**Hendrickson et al, 1988**). Cette dernière doit contenir un diffuseur anormal présentant un fort signal, résultat obtenu à partir de certains atomes lourds (Hg, Au, Pt et autres) ou en exprimant à partir de micro-organismes une protéine contenant des sélénio-méthionines à la place des méthionines. La présence d'un seul atome de Sélénium (numéro atomique : 34) est suffisante pour résoudre la structure d'une protéine de quelques 150 acides aminés par la méthode MAD (**Hendrickson et al, 1990 ; Leahy et al, 1992**).

VI.2.3. Remplacement moléculaire

La méthode du remplacement moléculaire (MR) tire bénéfice de l'existence de structures 3D de protéine déjà résolues. La connaissance de la structure primaire de la protéine à étudier autorise un alignement de séquences avec des séquences de protéines de structure connues. L'obtention d'une identité de séquence supérieure à 35 % permet d'identifier des protéines modèles, susceptibles de conduire à la solution de phasage recherchée, bien que différentes difficultés puissent être rencontrées. La fonction de Patterson est à la base d'une détermination de structure par remplacement moléculaire. Elle donne des informations sur les distances intermoléculaires présentes dans le cristal à partir des seules intensités diffractées. La méthode dite du remplacement moléculaire a pour but de déterminer la rotation (C) et la translation (\vec{t}) que l'on doit appliquer à un modèle X' de structure connue pour que ses régions les plus semblables à celles de la molécule recherchée X puissent se superposer dans l'espace réel de par la relation : $X = [C]X' + \vec{t}$

Deux fonctions de Patterson doivent pour cela être calculées : celle déduites des intensités mesurées du cristal de structure inconnue, et celle déduite des positions atomiques du modèle de structure connue placé dans une maille fictive. Les angles de rotation (exprimés dans un système de coordonnées polaires ou eulériennes) qui optimisent la corrélation (fonction de "Cross rotation") entre les deux fonctions de Patterson doivent être appliqués au modèle 3D pour orienter celui-ci correctement dans la maille cristalline inconnue. Une autre fonction de corrélation, dite de "self rotation" prend en compte la seule fonction de Patterson déduite des intensités diffractées. Cette fonction présente des maxima pour des valeurs angulaires correspondant à des symétries locales au sein du cristal.

$$P_{\text{modèle}}(\vec{u}) = \int_S \rho(\vec{r}) \cdot \rho(\vec{r} + \vec{u}) d\vec{r} \quad \text{où : } \vec{u} = (u, v, w), \vec{r} = (x, y, z)$$

$$P_{\text{cristal}}(\vec{u}) = \frac{1}{V} \sum_S |F(\vec{s})|^2 \cos[2\pi(\vec{u} \cdot \vec{s})] \quad \text{où : } \vec{u} = (u, v, w), \vec{s} = (h, k, l)$$

$$R_{\text{cross}}(C) = \int P_{\text{cristal}}(\vec{u}) \cdot P_{\text{modèle}}(C\vec{u}) d\vec{u}$$

$$R_{\text{self}}(C) = \int P_{\text{cristal}}(\vec{u}) \cdot P_{\text{cristal}}(C\vec{u}) d\vec{u}$$

(avec \vec{u} , \vec{s} et \vec{r} : les vecteurs de position respectivement dans l'espace de Patterson, l'espace réciproque et l'espace réel et C la matrice de rotation)

L'étape suivante consiste à déterminer le vecteur de translation (t_x, t_y, t_z) qui positionne le modèle correctement orienté dans la structure inconnue. Elle est réalisée en utilisant une fonction de Patterson calculée à partir des seuls vecteurs intermoléculaires existant en deux motifs moléculaires associés par un élément de symétrie.

$$P_{1,2}(\vec{u}) = \int \rho_1(\vec{r}) \cdot \rho_2(\vec{r} + \vec{u}) d\vec{r} \text{ où : } \vec{u} = (u, v, w), \vec{r} = (x, y, z)$$

$$T(\vec{t}) = \int P_{1,2}(\vec{u}, \vec{t}) \cdot P_{\text{cristal}}(\vec{u}) d\vec{u} \text{ où } \vec{t} = (t_x, t_y, t_z)$$

Cette fonction de translation $T(\vec{t})$ est maximale dès lors que les solutions de translation considérées sont correctes. Dès lors que la (ou les) différente(s) solution(s) de rotation et de translation est(sont) trouvée(s), l'empilement moléculaire en résultant doit être réaliste (non recouvrement des molécules entre elles, contacts intermoléculaires cohérents).

VI.3. Amélioration des cartes de densité électronique

De façon générale, toute information "de bon sens" touchant la structure cristalline recherchée et susceptible d'être utilisée (traduite en langage informatique) peut améliorer les cartes de densité électronique.

Applatissement de solvant : Les molécules de solvant, comblant l'espace existant au sein du cristal entre molécules protéiques, sont le plus souvent désordonnées d'une maille à l'autre. Une amélioration des phases estimées peut être obtenue en considérant constante la densité résiduelle (associée aux molécules de solvant) des cartes de densité électronique préalablement calculées à partir de la position des atomes lourds. Cela nécessite de différencier les régions attribuées aux molécules de protéine des autres régions de la maille cristalline en calculant le pourcentage de solvant présent dans le cristal (*Matthews, 1968*) et la forme d'une enveloppe délimitant la région protéique.

Moyennisation : Dès lors que l'unité asymétrique contient deux ou plusieurs entités moléculaires d'une même protéine, ces dernières sont en principe identiques. Les densités électroniques associées à chaque entité doivent donc l'être également. Cette information, lorsqu'elle est utilisée, permet d'améliorer considérablement le modèle obtenu par moyennisation de cartes de densité électronique trouvées différentes entre elles de par la prise en compte d'un modèle très incomplet.

VI.3.1. Affinement et estimation des erreurs

Une fois le problème des phases résolu et les cartes de densité électronique améliorées en utilisant les diverses techniques connues, les nouvelles cartes calculées présentent au niveau des entités protéiques un niveau de densité plus élevé. L'interprétation de ces cartes est effectué à l'aide de divers programmes performants tant au niveau graphique que de capacité de calcul (O, *Jones et al, 1991* ; Turbo Frodo, *Roussel et al, 1990* ; ...). La première étape consiste à construire un modèle protéique en ne considérant que les atomes de carbone Ca. S'en suit une reconstruction de la chaîne principale, puis des chaînes latérales en s'aidant de la séquence primaire. Un tel processus conduit à l'élaboration d'un modèle le plus complet possible respectant au mieux les cartes de densité. Un tel modèle contient toujours un grand nombre d'erreurs et doit être affiné au regard des intensités observées et de contraintes imposées à partir d'éléments structuraux ou chimiques déjà connus.

VI.3.2. Affinement restreint

Lors de la détermination des structures de "petites molécules", la résolution obtenue est généralement importante ($d < 0.7 \text{ \AA}$). En découle un affinement de structure dans lequel les positions atomiques sont parfaitement définies (le rapport du nombre d'intensités collectées sur le nombre de paramètres est bien supérieur à 10). Dans le cas de molécules plus importantes (protéines par exemple), le nombre de paramètres à déterminer et le nombre d'intensités diffractées collectées sont de même ordre de grandeur. Le type d'affinement est donc différent. Une protéine pouvant être considérée comme un polypeptide de 20 types différents d'acides aminés, les éléments de structure 3D déjà obtenues sur ces petites molécules à partir de données expérimentales précises (*Engh & Huber, 1991*) peuvent être paramétrisées et utilisées durant l'affinement d'une structure de protéine. Une fonction énergie E , sommant différents termes pondérés incluant une contribution issue de la diffraction des Rayons X, et diverses contributions d'origine géométrique, peut être minimisée :

$$\begin{aligned}
 E = & \sum_{hkl} w(hkl) \left\{ |F_{obs}(hkl)| - |F_{calc}(hkl)| \right\}^2 + \sum_{\substack{dist \\ j}} w_D (d_j^{idéal} - d_j^{modèle})^2 + \sum_{\substack{angles \\ i}} w_A (\alpha_i^{idéal} - \alpha_i^{modèle})^2 \\
 & + \sum_{\substack{chiral \\ m}} w_C (V_m^{idéal} - V_m^{modèle})^2 + \sum_{\substack{plans \\ n}} \sum_{\substack{atomes \\ coplanaires \\ p}} w_P (m_n r_{p,n} - d_n)^2 + \sum_{\substack{angles \\ torsion \\ \chi}} w_\chi (X_\chi^{idéal} - X_\chi^{modèle})^2 \\
 & + \sum_{\substack{contacts \\ intermol. \\ q}} w_{ci} (d_q^{min} - d_q^{modèle})^4
 \end{aligned}$$

VI.4. Validité du modèle

Chaque cycle d'affinement de la structure doit être validé afin d'éviter certaines erreurs introduites soit par le cristallographe soit par l'affinement lui-même. Un facteur de confiance, appelé Rfree (*Brünger, 1992*) calculé sur un jeu de réflexions exclues de l'affinement, est un parfait indicateur de la validité du modèle puisque son calcul n'est pas basé directement sur les paramètres affinés. Ce facteur doit diminuer en fonction de l'amélioration du modèle et donc de la convergence de l'affinement.

D'autres vérifications peuvent être prises en considération. Citons à titre d'exemple la recherche de longueurs et angles de liaison interdits, d'interactions intermoléculaires non cohérentes, de conformations de chaînes latérales non usuelles, ... Toutes ces erreurs doivent être systématiquement recherchées à l'aide de différents programmes (WHATIF, *Vriend 1990* ; Procheck, *Morris et al, 1992*, SFCHECK, *Vaguine et al, 1999*) afin de ne pas les conserver dans les cycles ultérieurs d'affinement où elles deviennent très difficilement corrigibles. Enfin, la déposition des coordonnées atomiques à la "Protein Data Bank" (<http://www.rcsb.org>) peut être effectuée, permettant ainsi une mise à disposition de la structure déterminée auprès de la communauté scientifique.

**REFERENCES
BIBLIOGRAPHIQUES**

REFERENCES BIBLIOGRAPHIQUES

Agarwal A. K., J. A. Richard. (2005) Minireview: Cellular Redox State Regulates Hydroxysteroid Dehydrogenase Activity and Intracellular Hormone Potency *Endocrinology* **146** (6): 2531–2538

Alfenito MR, Souer E, Goodman CD, Buell R, Mol J, Koes R, Walbot V (1998) Functional complementation of anthocyanin sequestration in the vacuole by widely divergent glutathione S-transferases. *Plant Cell* **10**: 1135-1149

Arndt UW, Wonacott AJ (1977) The rotation method in crystallography (Arndt, U. W., & Wonacott, A. J., Ed.), North Holland, Amsterdam

Aviram M, Fuhrman B (2002) Wine flavonoids protect against LDL oxidation and atherosclerosis. *Ann N Y Acad Sci* **957**: 146-161

Ban N, Nissen P, Hansen J, Moore PB, Steitz TA (2000) The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. *Science* **289**: 905–920

Baker ME, Blasco R (1992) Expansion of the mammalian 3 beta-hydroxysteroid dehydrogenase/plant dihydroflavonol reductase superfamily to include a bacterial cholesterol dehydrogenase, a bacterial UDP-galactose-4-epimerase, and open reading frames in vaccinia virus and fish lymphocystis disease virus. *FEBS Lett* **301**: 89-93

Bakker J, Timberlake, C.F. (1997) Isolation, identification, and characterization of new color-stable anthocyanins occurring in some red wines. *J. Agric. Food Chem.* **45**: 35-43

Bavaresco L, Fregoni C, Cantu E, Trevisan M (1999) Stilbene compounds: from the grapevine to wine. *Drugs Exp Clin Res* **25**: 57-63

Beart JE, Lilley T H and Haslam E. (1985) Plant polyphenols - secondary metabolism and chemical defense: Some observations. *Phytochemistry* **24**: 33-38

Benach J, Atrian S, Gonzalez-Duarte R, Ladenstein R. (1998). The refined crystal structure of *Drosophila lebanonensis* alcohol dehydrogenase at 1.9 Å resolution. *J. Mol. Biol.* **282**: 383-399

Bogs J, Downey MO, Harvey JS, Ashton AR, Tanner GJ, Robinson SP (2005) Proanthocyanidin synthesis and expression of genes encoding leucoanthocyanidin reductase and anthocyanidin reductase in developing grape berries and grapevine leaves. *Plant Physiol* **139**: 652-663

Bogs J, Ebadi A, McDavid D, Robinson SP (2006) Identification of the flavonoid hydroxylases from grapevine and their regulation during fruit development. *Plant Physiol* **140**: 279-291

- Boselli E, Boulton RB, Thorngate JH, Frega NG** (2004) Chemical and sensory characterization of DOC red wines from Marche (Italy) related to vintage and grape cultivars. *J Agric Food Chem* **52**: 3843-3854
- Boss PK, Davies C, Robinson SP** (1996) Expression of anthocyanin biosynthesis pathway genes in red and white grapes. *Plant Mol Biol* **32**: 565-569
- Bottoms CA, Smith PE, Tanner JJ** (2002) A structurally conserved water molecule in Rossmann dinucleotide-binding domains. *Protein Sci* **11**: 2125-2137
- Bradford MM** (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**: 248-254
- Brossaud F, Cheynier, V., and Noble A.C.** (2001) Bitterness and astringency of grape and wine polyphenols. *Australian Journal of Grape and Wine Research* **7**: 33-39
- Brouillard R, Chassaing S, Fougèrouse A** (2003) Why are grape/fresh wine anthocyanins so simple and why is it that red wine color lasts so long? *Phytochemistry* **64**: 1179-1186
- Brouillard R, Cheminat A** (1988) Flavonoids and plant color. *Prog Clin Biol Res* **280**: 93-106
- Brouillard R, Dubois J.E.** (1978) Chemistry of Anthocyanin Pigments. 3. Relaxation amplitudes in pH-Jump experiments. *J.Am.Chem.Soc* **100**: 6202-6205
- Brünger AT** (1992). Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature* **355**, 472-475
- Caderni G, De Filippo C, Luceri C, Salvadori M, Giannini A, Biggeri A, Remy S, Cheynier V, Dolara P** (2000) Effects of black tea, green tea and wine extracts on intestinal carcinogenesis induced by azoxymethane in F344 rats. *Carcinogenesis* **21**: 1965-1969
- Cheynier V** (2005) Polyphenols in foods are more complex than often thought. *Am J Clin Nutr* **81**: 223S-229S
- Cheynier V, Prieur C, Guyot S, Moutounet M.** (1997) The structures of tannins in grapes and wines and their interactions with proteins. *American Chemical Society Symposium* **661**: 81-93
- De Lange DW, Hijmering ML, Lorscheid A, Scholman WL, Kraaijenhagen RJ, Akkerman JW, van de Wiel A** (2004) Rapid intake of alcohol (binge drinking) inhibits platelet adhesion to fibrinogen under flow. *Alcohol Clin Exp Res* **28**: 1562-1568
- De Lange DW, Scholman WL, Kraaijenhagen RJ, Akkerman JW, van de Wiel A** (2004) Alcohol and polyphenolic grape extract inhibit platelet adhesion in flowing blood. *Eur J Clin Invest* **34**: 818-824
- De Lange DW, van de Wiel A** (2004) Drink to prevent: review on the cardioprotective mechanisms of alcohol and red wine polyphenols. *Semin Vasc Med* **4**: 173-186

De Vetten N, Quattrocchio F, Mol J, Koes R (1997) The an11 locus controlling flower pigmentation in petunia encodes a novel WD-repeat protein conserved in yeast, plants, and animals. *Genes Dev* **11**: 1422-1434

Debeaujon I, Peeters AJM, Léon-Kloosterziel KM, and Koornneef M. (2001). The *TRANSPARENT TESTA12* gene of *Arabidopsis* encodes a multidrug secondary transporter-like protein required for flavonoid sequestration in vacuoles of the seed coat endothelium. *Plant Cell* **13**, 853–871.

Dedaldechamp F, Uhel C (1999) Induction of anthocyanin synthesis in nonpigmented grape cell suspensions by acting on DFR substrate availability or precursors level. *Enzyme and Microbial Technology* **25**: 316-321

DeLano WL. (2002) The PyMOL Molecular Graphics System. DeLano Scientific, San Carlos, CA, USA, <http://pymol.sourceforge.net/>

Dell'Agli M, Busciala A, Bosisio E (2004) Vascular effects of wine polyphenols. *Cardiovasc Res* **63**: 593-602

Deluc L, Barrieu F, Marchive C, Lauvergeat V, Decendit A, Richard T, Carde JP, Merillon JM, Hamdi S (2006) Characterization of a grapevine R2R3-MYB transcription factor that regulates the phenylpropanoid pathway. *Plant Physiol* **140**: 499-511

Demrow HS, Slane PR, Folts JD (1995) Administration of wine and grape juice inhibits in vivo platelet activity and thrombosis in stenosed canine coronary arteries. *Circulation* **91**: 1182-1188

Devic M, Guilleminot J, Debeaujon I, Bechtold N, Bensuade E, Koornneef M, Pelletier G, Delseny M. (1999) The BANYULS gene encodes a DFR-like protein and is a marker of early seed coat development. *Plant J* **19**: 387–398

Dimmitt SB, Rakic V, Puddey IB, Baker R, Oosttryck R, Adams MJ, Chesterman CN, Burke V, Beilin LJ (1998) The effects of alcohol on coagulation and fibrinolytic factors: a controlled trial. *Blood Coagul Fibrinolysis* **9**: 39-45

Dixon RA, Paiva NL (1995) Stress-Induced phenylpropanoid metabolism. *Plant Cell* **7**: 1085-1097

Downey MO, Harvey, J.S., Robinson, S.P. (2003) Analysis of tannins in seeds and skins of shira grapes throughout berry development. *Aus J Grape and Wine Res* **9**: 15-27

Duffy SJ, Vita JA (2003) Effects of phenolics on vascular endothelial function. *Curr Opin Lipidol* **14**: 21-27

Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**: 2126-2132

Engh R, Huber R (1991) Accurate bond and angle parameters for X-ray structure refinement. *Acta Cryst.* **A47**, 392-400

Escribano-Bailon T, Dangles, O., Brouillard, R. (1996) Coupling reactions between flavylium ions and catechin. *Phytochemistry* **41**: 1583-1592

Evans G (2001) CHOOCH: a program for deriving anomalous-scattering factors from X-ray fluorescence spectra. *J. Appl. Cryst.* **34**: 82-86

Evans PR (1993) CCP4 study weekend data collection processing. 114-122

Ferrer JL, Jez JM, Bowman ME, Dixon RA, Noel JP (1999) Structure of chalcone synthase and the molecular basis of plant polyketide biosynthesis. *Nat Struct Biol* **6**: 775-784

Filling C, Berndt KD, Benach J, Knapp S, Prozorovski T, Nordling E, Ladenstein R, Jornvall H, Oppermann U (2002) Critical residues for structure and catalysis in short-chain dehydrogenases/reductases. *J Biol Chem* **277**: 25677-25684

Fitzpatrick DF, Fleming RC, Bing B, Maggi DA, O'Malley RM (2000) Isolation and characterization of endothelium-dependent vasorelaxing compounds from grape seeds. *J Agric Food Chem* **48**: 6384-6390

Forkmann G, Heller, W. (1999) Biosynthesis of flavonoids. *Comprehensive Natural Products Chemistry*: 713-748

Forkmann G, Ruhnau, B. (1987) Distinct substrate specificity of dihydroflavonol 4 reductase from flowers of *Petunia hybrida*. *Z.Naturforsch* **42**: 1146-1148

Forkner WJ, McArthur R.J., Lill J.T. (2004) Condensed tannins as anti-herbivore defences in leaf-chewing herbivore communities of *Quercus*. *Ecol Entomol* **29**: 174-187

Frishman D, Argos P. (1995) Knowledge-based secondary structure assignment. *Proteins : structure, function and genetics*, **23**, 566-579.

Fulcrand H, Benabdeljalil C, Rigaud J, Cheynier V, Moutounet M (1998) A new class of wine pigments generated by reaction between pyruvic acid and grape anthocyanins. *Phytochemistry* **47**: 1401-1407

Gachon CM, Langlois-Meurinne M, Henry Y, Saindrenan P (2005) Transcriptional co-regulation of secondary metabolism enzymes in *Arabidopsis*: functional and evolutionary implications. *Plant Mol Biol* **58**: 229-245

Galvin C. (1993) Etude de certaines réactions de dégradation des anthocyanes et de leur conservation avec les flavanols; conséquences sur la couleur des vins, Doctorat (UB2), Œnologie-Ampélogie. Université de Bordeaux 2

Garnier J, Gibrat JF, Robson B. (1996) GOR secondary structure prediction method version IV, *Methods in Enzymology*, R.F. Doolittle Ed., **266**, 540-553.

Gawel R (1998) Red wine astringency: a review. *Aus J Grape and Wine Res* **4**: 74-95

Gollop R, Even S, Colova-Tsolova V, Perl A (2002) Expression of the grape dihydroflavonol reductase gene and analysis of its promoter region. *J Exp Bot* **53**: 1397-1409

Green DW., Ingram VM., Perutz M. (1954) Proc. Roy. Soc. **A225**, 287-307

Gronbaek M, Becker U, Johansen D, Gottschau A, Schnohr P, Hein HO, Jensen G, Sorensen TI (2000) Type of alcohol consumed and mortality from all causes, coronary heart disease, and cancer. *Ann Intern Med* **133**: 411-419

Grotewold E, Chamberlin M, Snook M, Siame B, Butler L, Swenson J, Maddock S, Clair GS, Bowen B (1998) Engineering secondary metabolism in maize cells by ectopic expression of transcription factors. *Plant Cell* **10**: 721-740

Guex N, Peitsch M.C. (1997) SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* **18**: 2714-2723

Harker D. (1956) The determination of the phases of the structure factors of non-centrosymmetric crystals by the method of double isomorphous replacement. *Acta Cryst.* **A38**, 1-9

Harrison D.H., Bohren K.M., Ringe D., Gabbay K.H. (1997) The alrestatin double-decker: binding of two inhibitor molecules to human aldose reductase reveals a new specificity determinant. *Biochemistry* **36** (51):16134-40

Haslam E. (1994) Polyphenols-protein interactions. *Biochem.* **1974**, 135 : 285-288

Hendrickson WA, Smith JL, Phizackerley RP, Merritt EA (1988) Crystallographic structure analysis of lamprey hemoglobin from anomalous dispersion of synchrotron radiation. *Proteins* **4**(2), 77-88

Hendrickson WA, Horton JR, LeMaster DM (1990) Selenomethionyl proteins produced for analysis by multiwavelength anomalous diffraction (MAD): a vehicle for direct determination of three-dimensional structure. *EMBO J.* **9**(5), 1665-1672

Hendrickson WA (1991) Determination of macromolecular structures from anomalous diffraction of synchrotron radiation. *Science* **254**: 51-58

Hernandez JM, Heine GF, Irani NG, Feller A, Kim MG, Matulnik T, Chandler VL, Grotewold E (2004) Different mechanisms participate in the R-dependent activity of the R2R3 MYB transcription factor C1. *J Biol Chem* **279**: 48205-48213

Hertog MG, Feskens EJ, Kromhout D (1997) Antioxidant flavonols and coronary heart disease risk. *Lancet* **349**: 699

Holton TA, Cornish EC (1995) Genetics and biochemistry of anthocyanin biosynthesis. *Plant Cell* **7**: 1071-1083

Hopp W, Seitz, H.U. (1987) The uptake of acylated anthocyanin into isolated vacuoles from a cell suspension culture of *Daucus carota*. *Planta* **170**: 74-85

Hoppe W. (1957). Die 'Faltmolekülmethode' - eine neue Methode zur Bestimmung der Kristallstruktur bei ganz oder teilweise bekannter Molekülstruktur. *Acta Cryst.* **10**, 750-751

- Howard A, Chopra M, Thurnham D, Strain J, Fuhrman B, Aviram M** (2002) Red wine consumption and inhibition of LDL oxidation: what are the important components? *Med Hypotheses* **59**: 101-104
- Howe KM, Watson RJ** (1991) Nucleotide preferences in sequence-specific recognition of DNA by c-myb protein. *Nucleic Acids Res* **19**: 3913-3919
- Imhof A, Koenig W** (2003) Alcohol inflammation and coronary heart disease. *Addict Biol* **8**: 271-277
- Ito K, Nakajima Y, Ichihara E, Ogawa K, Katayama N, Nakashima K, Yoshimoto T** (2006) D-3-hydroxybutyrate dehydrogenase from *Pseudomonas fragi*: molecular cloning of the enzyme gene and crystal structure of the enzyme. *J Mol Biol* **355**: 722-733
- Jaakola L, Maatta K, Pirttila AM, Torronen R, Karenlampi S, Hohtola A** (2002) Expression of genes involved in anthocyanin biosynthesis in relation to anthocyanin, proanthocyanidin, and flavonol levels during bilberry fruit development. *Plant Physiol* **130**: 729-739
- Jez JM, Bowman ME, Dixon RA, Noel JP** (2000) Structure and mechanism of the evolutionarily unique plant enzyme chalcone isomerase. *Nat Struct Biol* **7**: 786-791
- Johnson ET, Ryu S, Yi H, Shin B, Cheong H, Choi G** (2001) Alteration of a single amino acid changes the substrate specificity of dihydroflavonol 4-reductase. *Plant J* **25**: 325-333
- Johnson ET, Yi H, Shin B, Oh BJ, Cheong H, Choi G** (1999) *Cymbidium hybrida* dihydroflavonol 4-reductase does not efficiently reduce dihydrokaempferol to produce orange pelargonidin-type anthocyanins. *Plant J* **19**: 81-85
- Jones TA, Zou JY, Cowan SW, Kjeldgaard M** (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Cryst.* **A47**(2), 110-119
- Jourdes M.** (2003) Réactivité, Synthèse, Couleur et Activité Biologique d'Ellagitannins C-Glycosidiques et Flavano-Ellagitannins. Thèse de Doctorat, Sciences Chimiques, Université Bordeaux 1
- Jovanovic** (1998) Antioxydant properties of flavonoids : reduction potentials and electron transfert reactions of flavonoid radicals. *In* M Dekker, ed, *Flavonoids in Health and Disease*, New York, pp 137-161
- Kabsch W** (1988) Evaluation of single-crystal X-ray diffraction data from a position-sensitive detector. *J. Appl. Cryst.* **21**, 916-924
- Kamitori S, Iguchi A, Ohtaki A, Yamada M, Kita K** (2005) X-ray structures of NADPH-dependent carbonyl reductase from *Sporobolomyces salmonicolor* provide insights into stereoselective reductions of carbonyl compounds. *J Mol Biol* **352**: 551-558

- Kampa M, Hatzoglou A, Notas G, Damianaki A, Bakogeorgou E, Gemetzi C, Kouroumalis E, Martin PM, Castanas E** (2000) Wine antioxidant polyphenols inhibit the proliferation of human prostate cancer cell lines. *Nutr Cancer* **37**: 223-233
- Kleiger G, Eisenberg D.** (2002) GXXXG and GXXXA Motifs Stabilize FAD and NAD(P)-binding Rossmann Folds Through C^α-H...O Hydrogen Bonds and van der Waals Interactions *Journal of Molecular Biology.* **323**, (1), 69-76
- Klein M, Martinoia E, Hoffmann-Thoma G, Weissenbock G** (2000) A membrane-potential dependent ABC-like transporter mediates the vacuolar uptake of rye flavone glucuronides: regulation of glucuronide uptake by glutathione and its conjugates. *Plant J* **21**: 289-304
- Klein M, Weissenbock G, Dufaud A, Gaillard C, Kreuz K, Martinoia E** (1996) Different energization mechanisms drive the vacuolar uptake of a flavonoid glucoside and a herbicide glucoside. *J Biol Chem* **271**: 29666-29671
- Kleywegt GJ** (1997) Validation of protein models from C α coordinates alone. *J Mol Biol* **273**: 371-376
- Koes R, Verweij W, Quattrocchio F** (2005) Flavonoids: a colorful model for the regulation and evolution of biochemical pathways. *Trends Plant Sci* **10**: 236-242
- Lapidot T, Harel S, Akiri B, Granit R, Kanner J** (1999) PH-dependent forms of red wine anthocyanins as antioxidants. *J Agric Food Chem* **47**: 67-70
- Lattman E** (1985) Diffraction methods for biological macromolecules. Use of the rotation and translation functions. *Methods Enzymol* **115**: 55-77
- Lawler IR, Foley, W.J., Woodrow, I.E.** (1997) The effects of elevated CO₂ atmospheres on the nutritional quality of Eucalyptus foliage and its interaction with soil nutrient and light availability. *Oecologia* **109**: 59-68
- Leahy DJ, Hendrickson WA, Aukhil I, Erickson HP** (1992) Structure of a fibronectin Type III domain from tenascin phased by MAD analysis of the selenomethionyl protein. *Science* **258**(5084), 987-991
- Leslie AGW** (2001) MOSFLM Program for Autoindexing and Integrating X-ray Diffraction Data, Version 6.11. MRC Laboratory of Molecular Biology, Hills Road, Cambridge, England
- Lesschaeve I, Noble AC** (2005) Polyphenols: factors influencing their sensory properties and their effects on food and beverage preferences. *Am J Clin Nutr* **81**: 330S-335S
- Li M, Kai Y, Qiang H, Dongying J** (2006) Biodegradation of gallotannins and ellagitannins. *J Basic Microbiol* **46**: 68-84
- Lu Y., Bennick A.** (1998). Interaction of tannin with human salivary proline-rich proteins. *Arch Oral Biol* **43**: 717-728

- Marles MA, Ray H, Gruber MY** (2003) New perspectives on proanthocyanidin biochemistry and molecular regulation. *Phytochemistry* **64**: 367-383
- Martens S, Teeri T, Forkmann G** (2002) Heterologous expression of dihydroflavonol 4-reductases from various plants. *FEBS Lett* **531**: 453-458
- Martin S, Andriambelason E, Takeda K, Andriantsitohaina R** (2002) Red wine polyphenols increase calcium in bovine aortic endothelial cells: a basis to elucidate signalling pathways leading to nitric oxide production. *Br J Pharmacol* **135**: 1579-1587
- Martin S, Diebolt M, Andriantsitohaina R** (2001) Moderate alcohol consumption and cardiovascular diseases. *Pathol Biol (Paris)* **49**: 769-774
- Matern U, Heller, W.** (1986) Efficient uptake of flavonoids into parsley vacuoles requires acylated glycosides. *Planta* **167**
- Matthews BW** (1968) Solvent content of protein crystals. *J.Mol.Biol* **33**: 491-497
- McGuffin LJ, Bryson K, Jones DT.** (2000) The PSIPRED protein structure prediction server. *Bioinformatics*. **16**, 404-405.
- McPherson A** (1990) Current approaches to macromolecular crystallization. *Eur J Biochem*. **189**: 1-23
- Medina K, Boido E., Carrau F.** (2005) Yeast interactions with anthocyanins during red wine fermentation. *Am J Enol Vitic* **56**: 104-109
- Meilleur F, Snell EH, Myles DA.** (2006) A quasi-Laue neutron crystallographic study of D-xylose isomerase. *Eur Biophys J* **35, 7** : 601-609
- Mesecar AD., Stoddard BL., Koshland DE.** (1997) Orbital Steering in the Catalytic Power of Enzymes: Small Structural Changes with Large Catalytic Consequences. *Science* **277**: 202-206
- Mingshu L, Kai Y, Qiang H, Dongying J.** (2006) Biodegradation of gallotannins and ellagitannins. *Journal of Basic Microbiology*. **46, 1** : 68
- Mirabel M.** (2000) Caractéristiques chimiques et organoleptiques des tanins des raisins de *Vitis vinifera* Var. Merlot et Cabernet Sauvignon issus de différents terroirs bordelais. Thèse de Doctorat, Œnologie-Ampélogie. Université de Bordeaux 2.
- Mol J, Grotewold, E., Koes, R.** (1998) How genes paint flowers and seeds. *Plant Science* **3**: 213-217
- Morii H, Uedaira H, Ogata K, Ishii S, Sarai A** (1999) Shape and energetics of a cavity in c-Myb probed by natural and non-natural amino acid mutations. *J Mol Biol* **292**: 909-920
- Morris AL, MacArthur MW, Hutchinson EG, Thornton JM** (1992) Stereochemical quality of protein structure coordinates. *Proteins* **12**: 345-64

Murphy KJ, Chronopoulos AK, Singh I, Francis MA, Moriarty H, Pike MJ, Turner AH, Mann NJ, Sinclair AJ (2003) Dietary flavanols and procyanidin oligomers from cocoa (*Theobroma cacao*) inhibit platelet function. *Am J Clin Nutr* **77**: 1466-1473

Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* **53**: 240-255

Nakajima J, Tanaka Y, Yamazaki M, Saito K (2001) Reaction mechanism from leucoanthocyanidin to anthocyanidin 3-glucoside, a key reaction for coloring in anthocyanin biosynthesis. *J Biol Chem* **276**: 25797-25803

Navaza J, Panepucci EH, Martin C (1998) On the use of strong Patterson function signals in many-body molecular replacement. *Acta Crystallogr D Biol Crystallogr* **54**: 817-821

Nelsestuen GL, Martinez MB (1997) Steady state enzyme velocities that are independent of [enzyme]: an important behavior in many membrane and particle-bound states. *Biochemistry* **36**: 9081-9086

Nemeth K, Plumb GW, Berrin JG, Juge N, Jacob R, Naim HY, Williamson G, Swallow DM, Kroon PA (2003) Deglycosylation by small intestinal epithelial cell beta-glucosidases is a critical step in the absorption and metabolism of dietary flavonoid glycosides in humans. *Eur J Nutr* **42**: 29-42

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. *Embo J* **25**: 1396-1405

Ogata K, Kanei-Ishii C, Sasaki M, Hatanaka H, Nagadoi A, Enari M, Nakamura H, Nishimura Y, Ishii S, Sarai A (1996) The cavity in the hydrophobic core of Myb DNA-binding domain is reserved for DNA recognition and trans-activation. *Nat Struct Biol* **3**: 178-187

Ogata K, Morikawa S, Nakamura H, Hojo H, Yoshimura S, Zhang R, Aimoto S, Ametani Y, Hirata Z, Sarai A, et al. (1995) Comparison of the free and DNA-complexed forms of the DNA-binding domain from c-Myb. *Nat Struct Biol* **2**: 309-320

Ogata K, Morikawa S, Nakamura H, Sekikawa A, Inoue T, Kanai H, Sarai A, Ishii S, Nishimura Y (1994) Solution structure of a specific DNA complex of the Myb DNA-binding domain with cooperative recognition helices. *Cell* **79**: 639-648

Orgogozo JM, Dartigues JF, Lafont S, Letenneur L, Commenges D, Salamon R, Renaud S, Breteler MB (1997) Wine consumption and dementia in the elderly: a prospective community study in the Bordeaux area. *Rev Neurol (Paris)* **153**: 185-192

Otwinowski Z, Minor W (1997) Processing of X-ray data diffraction data collected in oscillation mode. *Meth. Enzymol.* **276**, 307-326

Park HS, Fougère-Rifot, M., and Bouard, J. (1995) Les tanins vacuolaires de la baie de *Vitis vinifera* L. var. Merlot à maturité. *Oenologie* **95**: 115-118

Pascal C., Poncet-Legrand C., Vernhet A. (2006) Mechanisms of interaction between human salivary praline rich proteins and flavan-3-ols. *Polyphenols communications*

Peleg H, Gacon K., Noble A.C (1999) Bitterness and astringency of flavan-3-ol monomers, dimers and trimers. *Journal of the Science of food and agriculture* **79**: 1123-1128

Perez-Magarino S, Gonzalez-San Jose ML (2004) Evolution of flavanols, anthocyanins, and their derivatives during the aging of red wines elaborated from grapes harvested at different stages of ripening. *J Agric Food Chem* **52**: 1181-1189

Peters DJ, Constabel CP (2002) Molecular analysis of herbivore-induced condensed tannin synthesis: cloning and expression of dihydroflavonol reductase from trembling aspen (*Populus tremuloides*). *Plant J* **32**: 701-712

Pignatelli P, Pulcinelli FM, Celestini A, Lenti L, Ghiselli A, Gazzaniga PP, Violi F (2000) The flavonoids quercetin and catechin synergistically inhibit platelet function by antagonizing the intracellular production of hydrogen peroxide. *Am J Clin Nutr* **72**: 1150-1155

Pollastri G, McLysaght A. (2005) Porter: a new, accurate server for protein secondary structure prediction. *Bioinformatics*, **21** (8), 1719-20.

Rasmussen SE, Frederiksen H, Struntze Krogholm K, Poulsen L (2005) Dietary proanthocyanidins: occurrence, dietary intake, bioavailability, and protection against cardiovascular disease. *Mol Nutr Food Res* **49**: 159-174

Renaud S, de Lorgeril M (1992) Wine, alcohol, platelets, and the French paradox for coronary heart disease. *Lancet* **339**: 1523-1526

Ribereau-Gayon P (1965) The color of wines. *Aliment Vie* **53**: 232-248

Ribereau-Gayon P, Milhe JC (1969) Flavanes of grapes. *C R Acad Sci Hebd Seances Acad Sci D* **268**: 1813-1815

Ribereau-Gayon P, Stonestreet E (1965) Determination of anthocyanins in red wine. *Bull Soc Chim Fr* **9**: 2649-2652

Roggero JP, Coen, S., and Ragonnet, B. (1986) High performance liquid chromatography survey on changes in pigment content in ripening grapes of Syrah. An approach to antocyanin metabolism. *AM J Enol Vit* **37**: 77-83

Rossmann MG. (1960) The accurate determination of the position and shape of heavy-atom replacement groups in proteins. *Acta Cryst.* **13**, 221-226

Rossmann MG, Blow DM. (1962) *Acta Cryst.* **15**, 24-31

Rossmann MG. (1990) The molecular replacement method. *Acta Cryst.* **A46**(2), 73-82

Roussel A, Fontecilla-Camps, J.C. (1990) TURBO-FRODO: Bordeaux, France: XV IUCr Congress. *Acta Cryst.* **A46**: C66-C67

- Ryan KG, Swinny EE, Winefield C, Markham KR** (2001) Flavonoids and UV photoprotection in *Arabidopsis* mutants. *Z Naturforsch [C]* **56**: 745-754
- Saiki RK, Scharf S, Faloona F, Mullis KB, Horn KB, Erlich HA, Arnheim N.** (1985) Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science*, **230**, 1350-1354
- Salem RO, Laposata M** (2005) Effects of alcohol on hemostasis. *Am J Clin Pathol* **123 Suppl**: S96-105
- Sambrook J, Fritsch EF, Maniatis T.** (1989). *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
- Sanger F and Coulson AR.** (1975). A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *J. Mol. Biol.* **94**: 441-44
- Schaefer HM, Rolshausen G** (2006) Plants on red alert: do insects pay attention? *Bioessays* **28**: 65-71
- Schijlen EG, Ric de Vos CH, van Tunen AJ, Bovy AG** (2004) Modification of flavonoid biosynthesis in crop plants. *Phytochemistry* **65**: 2631-2648
- Schlieben NH, Niefind K, Muller J, Riebel B, Hummel W, Schomburg D** (2005) Atomic resolution structures of R-specific alcohol dehydrogenase from *Lactobacillus brevis* provide the structural bases of its substrate and cosubstrate specificity. *J Mol Biol* **349**: 801-813
- Sen CK, Bagchi D** (2001) Regulation of inducible adhesion molecule expression in human endothelial cells by grape seed proanthocyanidin extract. *Mol Cell Biochem* **216**: 1-7
- Sheldrick GM** (2002) Macromolecular phasing with SHELXE. *Z Kristallogr.* **217**: 644-650
- Shimada N, Sasaki R, Sato S, Kaneko T, Tabata S, Aoki T, Ayabe S** (2005) A comprehensive analysis of six dihydroflavonol 4-reductases encoded by a gene cluster of the *Lotus japonicus* genome. *J Exp Bot* **56**: 2573-2585
- Shimada S, Takahashi K, Sato Y, Sakuta M** (2004) Dihydroflavonol 4-reductase cDNA from non-anthocyanin-producing species in the Caryophyllales. *Plant Cell Physiol* **45**: 1290-1298
- Shimon LJ, Goihberg E, Peretz M, Burstein Y, Frolow F** (2006) Structure of alcohol dehydrogenase from *Entamoeba histolytica*. *Acta Crystallogr D Biol Crystallogr* **62**: 541-547
- Somers TC** (1966) Wine tannins--isolation of condensed flavonoid pigments by gel-filtration. *Nature* **209**: 368-370
- Sparvoli F, Martin C, Scienza A, Gavazzi G, Tonelli C** (1994) Cloning and molecular analysis of structural genes involved in flavonoid and stilbene biosynthesis in grape (*Vitis vinifera* L.). *Plant Mol Biol* **24**: 743-755

Springob K, Nakajima J, Yamazaki M, Saito K (2003) Recent advances in the biosynthesis and accumulation of anthocyanins. *Nat Prod Rep* **20**: 288-303

Tanner GJ, Francki KT, Abrahams S, Watson JM, Larkin PJ, Ashton AR (2003) Proanthocyanidin biosynthesis in plants. Purification of legume leucoanthocyanidin reductase and molecular cloning of its cDNA. *J Biol Chem* **278**: 31647-31656

Tehei M., Daniel R., Zaccai G. (2006) Fundamental and biotechnological applications of neutron scattering measurements for macromolecular dynamics *Eur Biophys J* **35, 7**: 551-558

Terwilliger T (2004) SOLVE and RESOLVE: automated structure solution, density modification and model building. *J Synchrotron Radiat* **11**: 49-52

Thoden JB, Frey PA, Holden HM (1996) High-resolution X-ray structure of UDP-galactose 4-epimerase complexed with UDP-phenol. *Protein Sci* **5**: 2149-2161

Thomas JL, Duax WL, Addlagatta A, Kacsoh B, Brandt SE, Norris WB (2004) Structure/function aspects of human 3beta-hydroxysteroid dehydrogenase. *Mol Cell Endocrinol* **215**: 73-82

Timberlake CB (1976) Interaction between anthocyanins, phenolic compounds and acetaldehyde and their significance in red wines. *Am. J. Enol. Vitic* **27**: 97-105

Turnbull JJ, Nakajima J, Welford RW, Yamazaki M, Saito K, Schofield CJ (2004) Mechanistic studies on three 2-oxoglutarate-dependent oxygenases of flavonoid biosynthesis: anthocyanidin synthase, flavonol synthase, and flavanone 3beta-hydroxylase. *J Biol Chem* **279**: 1206-1216

Turnbull JJ, Sobey, W.J., Aplin, R.T., Hassan, A., Firmin, J.L., Schofield, C.J. and Prescott, A.G. (2000) Are anthocyanidins the immediate products of anthocyanidin synthase. *Chem. Commun.:* 2473-2474

Ueyama Y, Suzuki, K., Tanaka, Y. (2002) Molecular and biochemical characterization of torenia flavonoid 3' hydroxylase and flavone synthase II and modification of flower color by modulating the expression of the genes. *Plant Science* **163**: 253-263

Ulrich S (2005) Molecular mechanisms of the chemopreventative effects of resveratrol and its analogs in carcinogenesis. *Molecular Nutrition & Food Research.* **49**: 452-461.

Vagin AA (1997) MOLREP: an automated program for molecular replacement. *J. Appl. Cryst.* **30**: 1022-1025

Vaguine AA, Richelle J, Wodak SJ (1999) SFCHECK: a unified set of procedures for evaluating the quality of macromolecular structure-factor data and their agreement with the atomic model. *Acta Crystallogr D Biol Crystallogr*, **55 (Pt 1)**, 191-205

Versari A, Parpinello GP, Tornielli GB, Ferrarini R, Giulivo C (2001) Stilbene compounds and stilbene synthase expression during ripening, wilting, and UV treatment in grape cv. Corvina. *J Agric Food Chem* **49**: 5531-5536

- Vitrac X, Bornet A, Vanderlinde R, Valls J, Richard T, Delaunay JC, Merillon JM, Teissedre PL** (2005) Determination of stilbenes (delta-viniferin, trans-astringin, trans-piceid, cis- and trans-resveratrol, epsilon-viniferin) in Brazilian wines. *J Agric Food Chem* **53**: 5664-5669
- Vom Endt D, Kijne JW, Memelink J** (2002) Transcription factors controlling plant secondary metabolism: what regulates the regulators? *Phytochemistry* **61**: 107-114
- Vriend G** (1990) WHAT IF: A molecular modeling and drug design program. *J. Mol. Graph.* **8**: 52-56
- Vriend G, Sander C** (1993) Quality control of protein models: directional atomic contact analysis. *J. Appl. Cryst.* **26**, 47-60.
- Wang H, Race EJ, Shrikhande AJ** (2003) Anthocyanin transformation in Cabernet Sauvignon wine during aging. *J Agric Food Chem* **51**: 7989-7994
- Wellmann F, Lukacin R, Moriguchi T, Britsch L, Schiltz E, Matern U** (2002) Functional expression and mutational analysis of flavonol synthase from *Citrus unshiu*. *Eur J Biochem* **269**: 4134-4142
- Wilmouth RC, Turnbull JJ, Welford RW, Clifton IJ, Prescott AG, Schofield CJ** (2002) Structure and mechanism of anthocyanidin synthase from *Arabidopsis thaliana*. *Structure* **10**: 93-103
- Winkel-Shirley B** (2001) Flavonoid biosynthesis. A colorful model for genetics, biochemistry, cell biology, and biotechnology. *Plant Physiol* **126**: 485-493
- Winkel-Shirley B** (2002) Biosynthesis of flavonoids and effects of stress. *Curr Opin Plant Biol* **5**: 218-223
- Wu S, Letchworth GJ** (2004) High efficiency transformation by electroporation of *Pichia pastoris* pretreated with lithium acetate and dithiothreitol. *Biotechniques* **36**: 152-154
- Xie DY, Sharma SB, Paiva NL, Ferreira D, Dixon RA** (2003) Role of anthocyanidin reductase, encoded by BANYULS in plant flavonoid biosynthesis. *Science* **299**: 396-399
- Xie DY, Sharma SB, Dixon RA** (2004) (a) Anthocyanidin reductases from *Medicago truncatula* and *Arabidopsis thaliana*. *Arch Biochem Biophys* **422**: 91-102
- Xie DY, Jackson LA, Cooper JD, Ferreira D, Paiva NL** (2004) (b) Molecular and biochemical analysis of two cDNA clones encoding dihydroflavonol-4-reductase from *Medicago truncatula*. *Plant Physiol* **134**: 979-994
- Yamashita A, Kato H, Wakatsuki S, Tomizaki T, Nakatsu T, Nakajima K, Hashimoto T, Yamada Y, Oda J** (1999) Structure of tropinone reductase-II complexed with NADP⁺ and pseudotropine at 1.9 Å resolution: implication for stereospecific substrate binding and catalysis. *Biochemistry* **38**: 7630-7637

Zern TL, Fernandez ML (2005) Cardioprotective effects of dietary polyphenols. J Nutr 135: 2291-2294

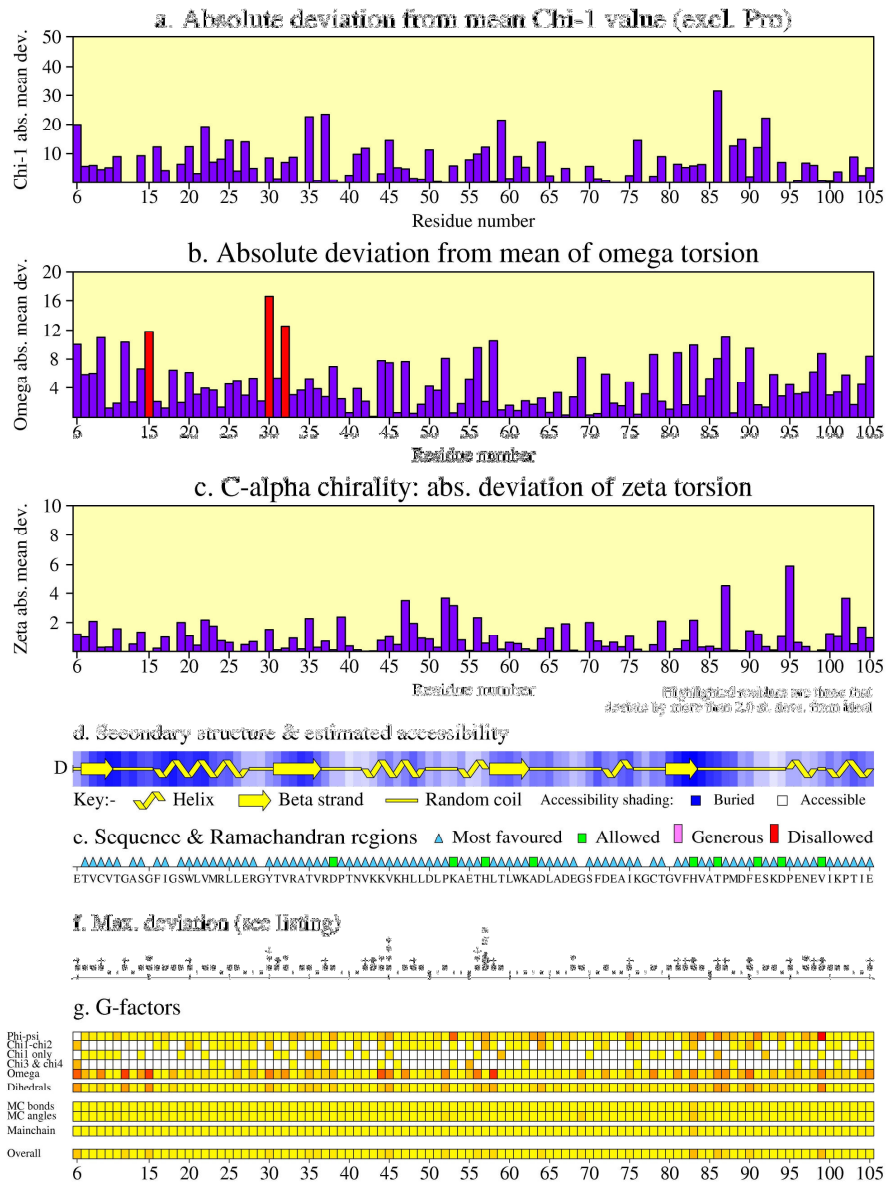
ANNEXES

ANNEXES

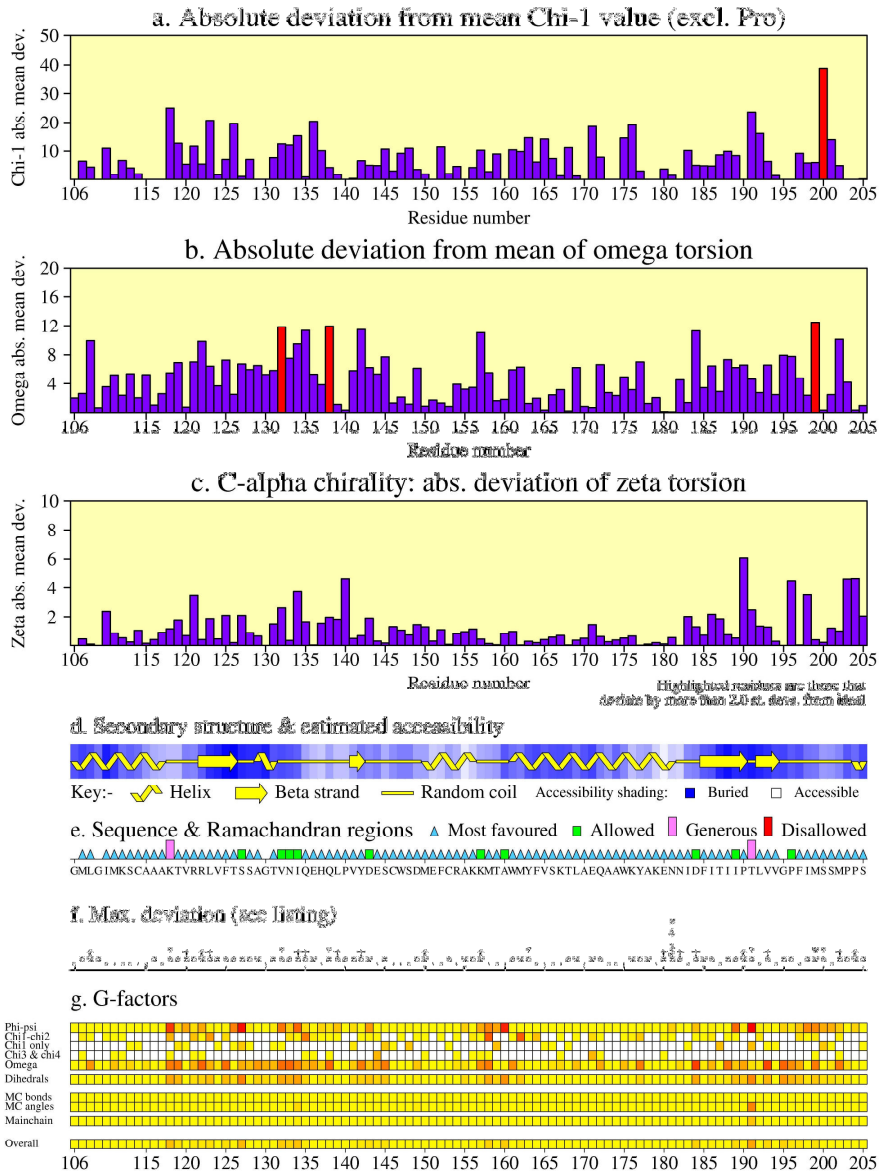
Annexe 1 : Analyse de la qualité stéréochimique de chaque résidu de la DFR sauvage obtenue à l'aide programme Procheck. Les paramètres étudiés sont les déviations par rapport aux angles χ_1 , ω , ξ , la structure secondaire, l'accessibilité estimée, la qualité de la conformation, la déviation maximale et le facteur de qualité G.

Les trois premiers tableaux (**a,b,c**) caractérisent l'écart à la moyenne théorique des angles de torsion χ_1 (chaîne latérale), ω (liaison peptidique), et ξ (C_α) ce dernier correspondant à l'angle formé par la structure tétraédrique C_α -N-C- C_β . Le tableau **d** traduit par une représentation schématique la structure secondaire de la protéine et l'exposition théorique des résidus au solvant. Le graphique **e** identifie la région du diagramme de Ramachandran où chaque résidu est placé. L'histogramme **f** montre la déviation maximale des paramètres stéréochimiques de chaque acide aminé de la structure. (Un astérisque traduit une déviation standard de 1σ . Le signe (+), symbolise une déviation de $1/2\sigma$). Enfin, le tableau **g** représente la valeur des G-factors pour différentes propriétés stéréochimiques de chaque résidu.

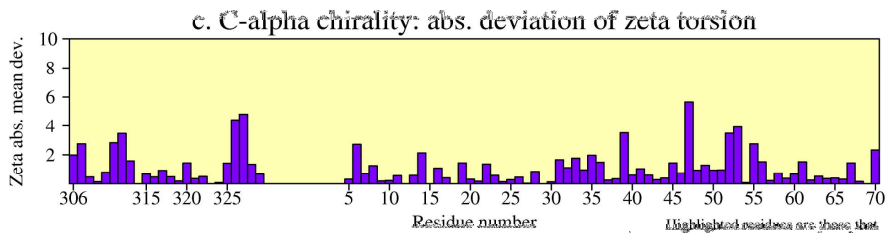
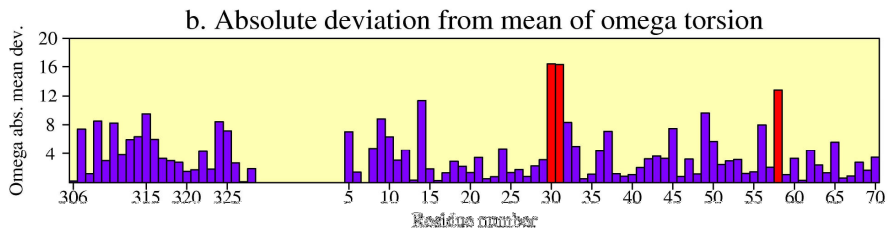
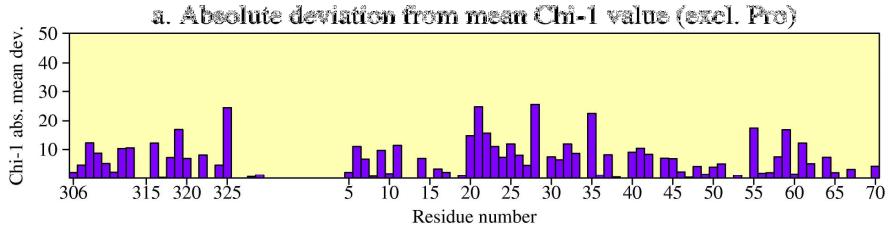
Residue properties



Residue properties



Residue properties

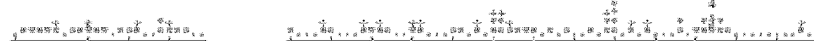


Highlighted residues are those that deviate by more than 2.0 st. dev. from the mean

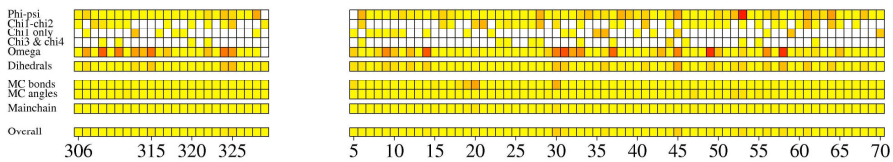
d. Secondary structure & estimated accessibility



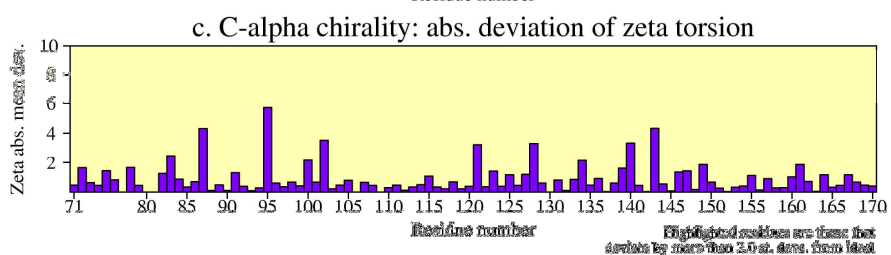
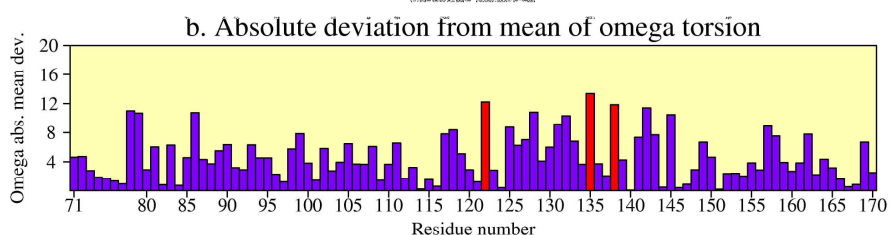
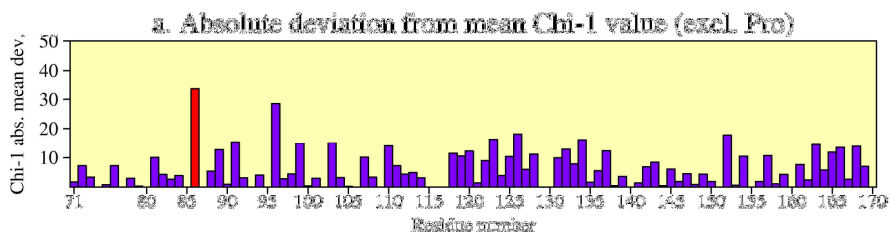
f. Max. deviation (see listing)



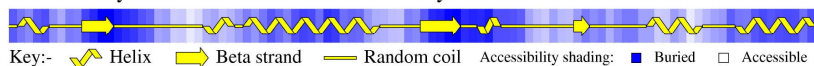
g. G-factors



Residue properties



d. Secondary structure & estimated accessibility

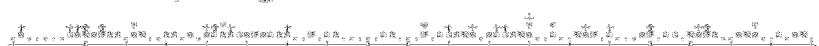


Key:- Helix Beta strand Random coil Accessibility shading: Buried Accessible

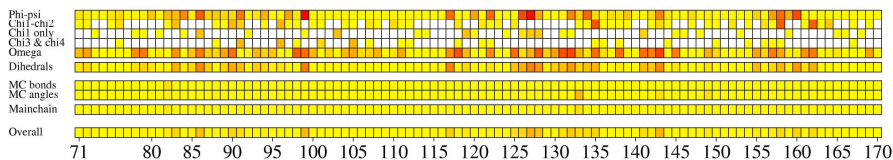
e. Sequence & Ramachandran regions Most favoured Allowed Generous Disallowed

FDEAIKGGCTGVFHVATPMDEESKDPENEVIKPTIEGMLGIMKSCAAAKTVRRILVFTSSAGTVNIOFHQLPVYDESCWSDMEFCRAKMTAWMYFVSKTLA

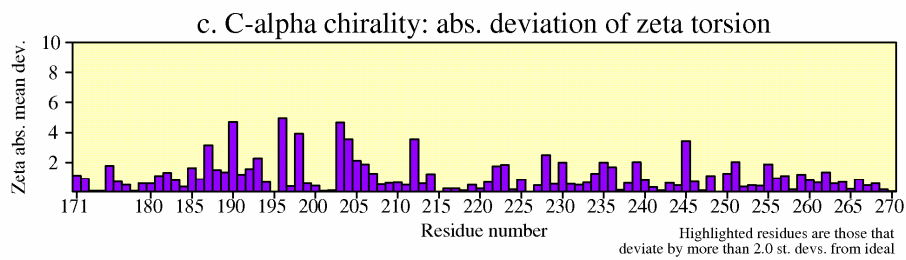
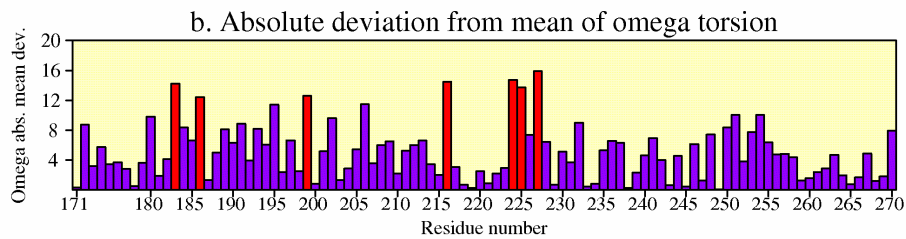
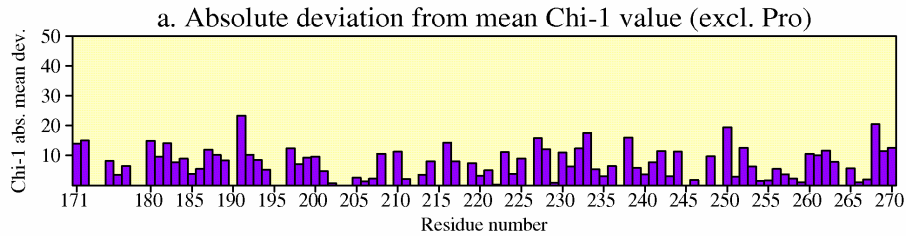
f. Max. deviation (see listing)



g. G-factors



Residue properties



d. Secondary structure & estimated accessibility



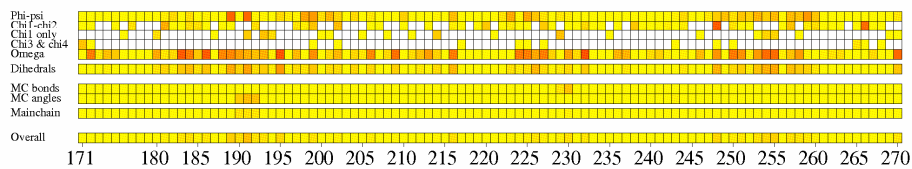
e. Sequence & Ramachandran regions



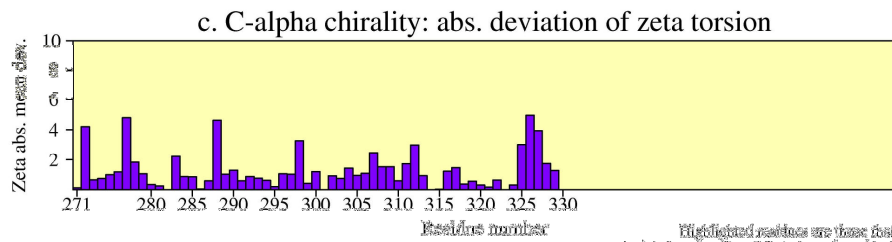
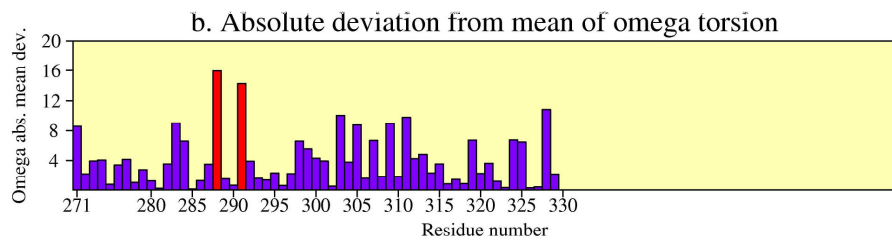
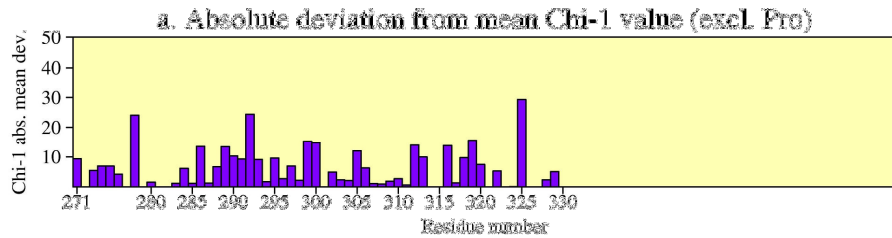
f. Max. deviation (see listing)



g. G-factors

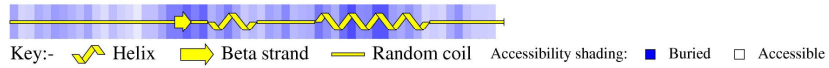


Residue properties



Highlighted residues are those that deviate by more than 2.0 st. dev. from the

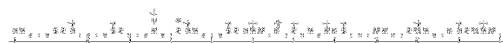
d. Secondary structure & estimated accessibility



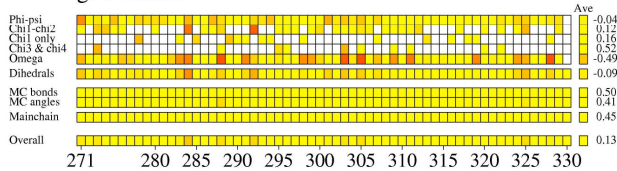
e. Sequence & Ramachandran regions Most favoured Allowed Generous Disallowed

YPEYNI PTEFKGVDFENLKSVCFS SSKLTDLGFEPKYSLEDMFTGAVDTCRAGLLPPSHE

f. B-factor deviation (see listing)



g. G-factors



CRYSTAL STRUCTURE OF A GRAPE DIHYDROFLAVONOL-4-REDUCTASE, A KEY ENZYME IN THE FLAVONOID BIOSYNTHESIS.

Pierre Petit,^a Thierry Granier,^a Béatrice Langlois d'Estaintot,^a Claude Manigand,^a Katell Bathany,^b Jean-Marie Schmitter,^b Virginie Lauvergeat,^c Saïd Hamdi^c and Bernard Gallois^{a,1}

^a Unité de Biophysique Structurale, UMR CNRS 5471, Bât. B8, Avenue des Facultés, Université Bordeaux 1, 33405 Talence Cedex, France

^b Institut Européen de Chimie et de Biologie, UMR CNRS 5144, 2 Rue Robert Escarpit, 33607 Pessac, France

^c UMR Ecophysiologie et Génomique Fonctionnelle de la Vigne, Université Bordeaux 1, Université Bordeaux 2, Institut National de la Recherche Agronomique, Institut des Sciences de la Vigne et du Vin, 71 Avenue Edouard Bourlaux, 33883 Villenave d'Ornon, France

¹ To whom correspondence should be addressed. E-mail b.gallois@ubs.u-bordeaux1.fr; fax: 33 (0)5 40 00 22 00

ABSTRACT

The NADPH-dependent enzyme dihydroflavonol-4-reductase (DFR) catalyses a late step in the biosynthesis of anthocyanins and condensed tannins, two flavonoid classes of importance to plant survival and human nutrition. This enzyme has been widely investigated in many plant species, but little is known about its structural and biochemical properties. To provide a basis for detailed structure-function studies, the crystal structure of *Vitis vinifera* DFR, heterologously expressed in *Escherichia coli*, has been determined at 1.8 Å resolution. The 3D structure of the ternary complex obtained with NADP⁺ and dihydroquercetin, one of the DFR substrates, presents common features with the short chain dehydrogenase family, i.e. a N-term domain adopting a Rossmann fold, and a variable C-term domain which participates to the substrate binding.

The structure confirms the importance of the 125-158 region which lines the substrate binding site and enlightens the role of a specific residue at position 133 (Asn or Asp), assumed to control

substrate recognition. The activity of the wild type enzyme and its variant N133D has been quantified in vitro, using dihydroquercetin or dihydrokaempferol. Our results demonstrate that the position 133 cannot be solely responsible for the recognition of the B cycle hydroxylation pattern of dihydroflavonols.

INTRODUCTION

Flavonoids constitute a large class of phenolic compounds widely studied, due to their important properties. They are present in most higher plants, and involved in relevant functions such as pigmentation, reproduction, growth and protection against pathogens. When ingested they may act in disease treatment and prevention, mainly due to their antioxidant capabilities. Many clinical effects have been ascribed to these compounds, including anticancer, anti-inflammatory and antimicrobial properties (Ross and Kasum, 2002; Halliwell et al. 2005) and a putative role has been found against coronary heart disease (CHD) (De Lorgeril et al 1996, Hertog et al. 1995).

Flavonoids have received much attention in many plants, including grape. Owing to its high flavonoid level, the moderate consumption of red wine potentially protects against CHD (Dell'agli et al. 2004) and flavonoids strongly influence the sensorial properties of this beverage. Two major classes of flavonoids are found in grape: anthocyanins (ANs) and proanthocyanidins (PAs). In *Vitis vinifera*, ANs are 3-monoglucoside derivatives of the aglycones delphinidin, cyanidin, peonidin and malvidin. They provide the color of the grapes and therefore of young red wine. PAs, also called condensed tannins, result from the polymerization of flavan-3-ol units (mainly 2R,3S-catechin and 2R,3R-epicatechin). Extracted during red wine making, they bestow sensorial properties such as bitterness and astringency and play an important role in long-term color stability when they react with ANs.

Flavonoids derive from phenylalanine and malonyl-coenzyme A. Their biosynthetic pathway has been largely investigated (Winkel-Shirley 2001) and most of the enzymes that control each single step have been identified from a molecular genetic point of view. The syntheses of both ANs and PAs follow the same pathway up to a last common step, driven by dihydroflavonol 4-reductase (DFR, EC 1.1.1.219). This enzyme catalyses the NADPH-dependent reduction of dihydroflavonols to flavan 3,4-diols. These in turn are substrates (Figure 1) to the next steps of the AN and PA biosyntheses, which involve either the anthocyanidin synthase (ANS) and the 3-O-glycosyltransferase (UGT) or the anthocyanidin reductase (ANR) and the leucocyanidin reductase (LAR). Further insights have been gained into this pathway by the recent 3D structure determination of ANS from *Arabidopsis thaliana* (Wilmouth et al., 2002) and of UGT from *Vitis vinifera* (Offen et al. 2006).

To understand the biochemical and structural properties of DFR, single or multiple gene(s) have been cloned from various plant species, i.e. maize (O'Reilly et al, 1985), petunia (Beld et al, 1989), gerbera (Helariutta et al, 1993), tomato (Bongue-Bartelsman et al, 1994), grape (Sparvoli et al 1994), medicago (Charrier et al, 1995; Xie et al, 2004), perilla (Chen et al, 1998; Gong et al, 1997), rosa

(Tanaka et al, 1995), glories (Inagaki et al, 1999), aspen (Peters et al, 2002), malus (Fischer et al, 2003), wheat (Himi and Noda, 2004) and lotus (Shimada et al, 2005).

To our knowledge, DFR activity was reported for the first time in the case of a synthesized enzyme from *Zea mays* (O'Reilly et al, 1985). Later on, the activity of various DFRs was deduced from plant protein extracts or phyto-chemical analyses of mutants and/or transgenic plants, until the successful expression of a functional DFR sequence from *Gerbera* (Martens et al, 2002). However, despite much effort, data on the biochemical properties of the enzyme are scarce. The purified DFR protein was described as very unstable (Forkmann and Ruhnau, 1987), which may explain why the crystal structure of the enzyme remains to be determined.

Up to now, the only structural information concerning DFRs has been deduced from sequence analysis. The DFR sequence contains most of the motives present in the short-chain dehydrogenase/reductase (SDR) superfamily, including the invariant YXXXXK motif (where X means any amino acid residue), a catalytic serine and a highly conserved N-terminal glycine rich nucleotide binding motif (Jörnvall et al., 1995). Consequently, DFR folding should present a large α/β single conserved domain named "Rossmann fold", like any other member of the superfamily.

However, sequence analysis cannot fully elucidate why most DFRs accept dihydroflavonols of different hydroxylation patterns (dihydrokaempferol, DHK; dihydroquercetin, DHQ and dihydromyricetin, DHM) as substrates, whereas DFRs from *Petunia* and *Cymbidium* do not efficiently act on DHK. A possible explanation for this particular case (Johnson et al, 2001) pinpointed an amino acid region susceptible to determine substrate specificity in the DFR sequences (Beld et al, 1989). Recent investigations support the latter finding: DFR sequences presenting an asparagine residue at position 134 (*Gerbera* numbering) reduce DHK more readily than DHQ, whereas those presenting an aspartic acid show a marked preference for DHQ (Johnson et al 2001, Xie et al, 2004, Shimada et al, 2005). The above assumption about the amino acid region determining the substrate binding site was only deduced from an analysis of sequence alignment and no 3D detailed description of the interactions between the substrate and the enzyme is available to date.

In this paper, we report the successful expression of a grape DFR cDNA leading to the synthesis of an active DFR enzyme from *Vitis vinifera* (Swiss-Prot accession number P93799_VITVI). Optimization of the elution conditions during purification and of the storage solution allowed us to keep the protein stable long enough in order not only to test its activity but also to co-crystallize it in the presence of DHQ and NADP⁺. Diffracting crystals to 1.8 Å resolution were obtained. The structure was solved by the multi wavelengths anomalous dispersion method using a Se-Met derivative, leading to the first 3D description of a DFR protein. Furthermore, in the light of the previous investigations cited above and the present structure analysis, we constructed the N133D (grape numbering) mutant and assayed its activity with respect to DHQ and DHK in order to compare it with that of the wild enzyme.

RESULTS

Enzyme activity

Analysis by mass spectrometry of the various peaks separated by HPLC showed four major products (Figure 2). Compounds eluting as peaks 1 and 3 displayed a red coloration whereas peaks 2 and 4 corresponded to colourless products. For DHQ, peak 1 is assigned to cyanidin (m/z 287.3), peak 2 to DHQ (m/z 305.0), peak 3 to O-butyl-cyanidin (m/z 343.2), and peak 4 to O-butyl-cyanidin plus $2H^+$ (m/z 345.2).

The reaction was found identical when DHK was used as the substrate. Peak 1' corresponds to pelargonidin (m/z 271.3), peak 2' to DHK (m/z 289.0), peak 3' to the O-butyl-pelargonidin (m/z 327.2) and peak 4' to the addition of $2H^+$ to the O-butyl derivative (m/z 329.1).

Substitution of n-butanol by other alcohols confirmed the assignation of peaks 3, 3', 4 and 4' to etherification products of the anthocyanidins. When n-propanol was used instead of n-butanol in the incubation process, m/z values of these peaks were shifted by 14 Daltons, i.e. the mass difference between the alcohols. When n-butanol was replaced by t-butanol, m/z values of products remained identical, but peaks area were strongly reduced, a result consistent with the change in steric hindrance of the alcohol.

The above data bring evidence that the purified enzyme is active, whatever the substrate considered (DHQ or DHK). Nevertheless, the efficiency was higher for DHQ than for DHK (Figure 2e). Because of the presence of various product derivatives, enzyme activity was estimated from the peak area of the remaining substrate.

Overall structure of DFR

A topological diagram with the secondary structure of DFR and a 3D representation of the ternary complex are shown in Figure 3. DFR contains 15 α helices with 152 amino acids (45% of the amino acid residues), 2 short 3_{10} helices with 8 residues (2.4% of the sequence) and 12 β strands with 52 residues (15.4 % of the amino acid residues).

A 3D alignment search using the server DALI (Holm and Park, 2000) shows that the best superposition was obtained with the structure of carbonyl reductase from *Sporobolomyces salmonicolor* (SSCR) co-crystallized with NADPH (Kamitori et al. 2005, pdb code 1y1p): an overall identity scored 26 % for 301 aligned residues and an r.m.s. deviation of 2.4 Å was obtained on aligned $C\alpha$ positions. Superimposition of both structures reveals the common features DFR shares with SSCR (Figure 4-A). Like the latter, it can be described as a two-domain protein:

- a large N-terminal domain which adopts a Rossmann fold: seven β strands form a large parallel β sheet and compose the hydrophobic core of the domain, flanked by seven α helices. Like

SSCR, DFR presents one β strand ($\beta 7$) and two α helices ($\alpha 9$ and $\alpha 11$) which disrupt the symmetry of the two halves β - α - β - α - β of a Rossmann fold.

- a small C-terminal domain composed of four α helices ($\alpha 7$, $\alpha 8$, $\alpha 10$, $\alpha 12$), two short 3_{10} helices (3a and 3b) and five β strands which form two small sheets ($\beta 2'\beta 3'\beta 4'$ and $\beta 1'\beta 5'$).

NADP⁺ binding site

The coenzyme binding site is formed by a cleft spanning the C-terminal face of the Rossmann fold domain (Figure 4-B). The co-factor NADP⁺ is bound in an extended conformation as seen in most SDR type proteins, through a tight network of hydrogen bonds and electrostatic interactions, as well as hydrophobic contacts. In the present structure, these interactions involve residues of the $\alpha 1$ - $\beta 1$ loop with the glycine-rich motif (G12, S14, G15, F16, I17, G18), residues of loop $\alpha 2$ - $\beta 2$ (R37, D38, K44), loop $\alpha 3$ - $\beta 3$ (D64, L65), loop $\alpha 4$ - $\beta 4$ (A85, T86, P87), residues of helix $\alpha 6$ (Y163, K167), residues of strands $\beta 6$ and $\beta 2'$ (P190, V193), and helix $\alpha 7a$ (S205). The adenine and nicotinamide rings are about 15 Å apart and their planes make an angle of 79°.

The central di-phosphate group of the NADP⁺ molecule binds the main chain atoms of the glycine-rich motif through direct hydrogen bonding or via a water molecule highly conserved in Rossmann dinucleotide-binding domains (Bottoms et al 2002).

As far as the adenine mononucleotide phosphate moiety is concerned, the mono-phosphoester group is tightly bound by electrostatic interactions to charged residues R37 and K44. The attached ribose displays the C2'-endo puckering conformation and adopts an anti-conformation with respect to the adenine moiety. The latter stacks its A face against L65, A85 and P87 hydrophobic residues and its B face against the aliphatic and guanidinium groups of the R37 side chain. The N6 amino group is hydrogen-bonded to the carboxylic end of D64. As for the other nitrogen atoms of the aromatic ring, all of them are involved in hydrogen bonding either with water molecules or main chain (L65) atoms.

The elements of the nicotinamide mononucleotide phosphate moiety present the following features:

- the nicotinamide planar group has a syn (pro-S) conformation with respect to its attached ribose ring, and the latter adopts a C2'-endo puckering conformation;
- the A face of the nicotinamide ring stacks against side chains of hydrophobic residues I17, P190 and V193, whereas its B face interacts with the substrate. DFR can thus be considered as a B (pro-S) type enzyme.

The most noticeable differences between the co-factor binding sites of DFR and SSCR concern the interactions of the mono-phosphoester group with residues S14, R37 and K44: these interactions are conserved in SCCR (N21, R44 and K48), but the presence of N21 in SSCR, as well as deletion of 3 residues between R44 and K48, leads to a slightly different position of the phosphate group, which is more exposed to solvent in the SSCR structure than in the DFR one.

Finally, the co-factor appears to be largely embedded within the protein: when considered alone, its solvent accessible surface is 906 Å², whereas when considered bound to the protein, the surface drops to 56 Å², and 41 Å² when taking into account the presence of the substrate.

Substrate binding site

Refinement of the structure was carried out with the different R/S conformations of the tetrahedral carbon atoms in positions 2 and 3 of cycle C. The resolution limit of 1.8 Å was not sufficient to discriminate with certainty which combination of R/S conformations should be attributed to the bound DHQ molecule in the present structure. However, the best electron density maps were obtained with the 2R3R or 2S3S conformations.

The configuration of the substrate displays the following characteristics:

- the fused rings A and C are almost co-planar, the overall rms deviation from the plane defined by rings A and C is 0.23 Å, with the greatest value observed for the asymmetrical carbon atoms C3 (0.48 Å) and C2 (0.33 Å), and the oxygen O4 (0.56 Å).
- ring B is rotated around the C2-C1' bond by an angle of 80° with respect to the mean plane of rings A and C.

Figure 4-C shows the DHQ binding site with the nicotinamide part of the NADP⁺ co-factor. The DHQ molecule is tightly bound to the protein and interacts with the nicotinamide ring. All its hydroxyl groups establish hydrogen bonds with the protein or water molecules. The only oxygen atom which does not participate in hydrogen bonding is O1 from cycle C. Ring C carbonyl interacts with the hydroxyls of S128 and Y163 of the catalytic triad (see below). Across the catalytic site, cycle C interacts with hydrophobic residues (A129, F164, L192, I222). Cycle A is relatively exposed to the solvent and is hydrogen-bonded to side chains of residues T208 or Y163 and water molecules. Both hydroxyl groups of cycle B are hydrogen-bonded to side chains of residues N133 and Q227, whereas the carbon atoms interact with A129 and I222 side chains via hydrophobic contacts. These interactions account for the rotation (80°) of plane B relative to the mean plane of fused rings A and C.

Like co-factor NADP⁺, the substrate is almost completely embedded inside the active site: its solvent accessible area drops from 479 to 6.8 Å² when free and bound to the protein, respectively.

Catalytic site architecture

In the ternary complex, the 3D arrangement at the catalytic site is similar to what is observed in other structures of SDR type proteins. A close-up view of the reaction centre is shown in Figure 5-A. The hydroxyl groups of Y163 and S128 interact via hydrogen bonds with DHQ O4, whereas the K167 amino group interacts with the hydroxyl groups 2' and 3' of the NADP⁺ ribose moiety. The relative positions of the substrate and co-factor lead to the expected interactions between the C4 of the nicotinamide ring and DHQ:

- a C4-C9 distance of 2.99 Å,

- a dihedral angle between the planes formed by the nicotinamide ring and that formed by N1(NADP⁺), C4(NADP⁺) and C4(DHQ) close to 90° (81°),

- an angle (C4(DHQ)-C4(NADP⁺)-N1(NADP⁺)) of 105°; a value within the range (102°-119°) computed by various methods (see Yamashita et al. 1999 and references therein) and compatible with a hydride transfer from the nicotinamide ring and the carbon atom of the substrate carbonyl group.

Finally, the nicotinamide ring and the DHQ molecule stack on top of each other in a favorable manner, with the planes formed by the atoms C10, C4, O4 and C3 of the DHQ molecule and those of the nicotinamide ring being only 28° apart.

Besides the interactions with the hydroxyl groups of the nicotinamide ribose, the amino group of K167 interacts with a water molecule which belongs to a cluster of 5 water molecules. These molecules fill a hydrophilic cavity and do not have access to the bulk solvent in the present structure.

DISCUSSION

Although DFR is one of the most investigated enzymes of the flavonoid biosynthetic pathway, its 3D structure remained to be determined. Not only the 3D description of the DFR-NADP⁺-DHQ ternary complex confirms that the enzyme belongs to the SDR family but it also reveals the exact conformation of the C-terminal domain enclosing the substrate binding site. Therefore, the protein - substrate interactions reported here may offer new insights into the key parameters driving the specificity of the substrate.

Conformational changes of members of the SDR family upon co-factor and/or substrate binding have been observed in several cases:

- in SSCR, a large part of the apo-enzyme C-terminal end is disordered, suggesting that this part of the protein is very mobile, and that the binding sites are easily accessible to the solvent, whereas the structure of the binary complex NADP⁺-enzyme was fully ordered;

- the structures of D-3-hydroxybutyrate dehydrogenase from *Pseudomonas fragi* (Ito et al., 2006) showed that the C-terminal end is disordered when both NAD⁺ and substrate bind the enzyme, whereas the apo form is fully ordered and exhibits an open conformation.

Nevertheless, such conformational changes are not systematic. Schlieben et al. (2005) observed that the apo and ternary complexes of the R-specific alcohol dehydrogenase from *Lactobacillus brevis* have very close and fully ordered structures.

With DFR, we have unfortunately failed to date to crystallize the apo-enzyme and the binary complex DFR/NADPH. In the ternary complex, the very low solvent accessibility surfaces of co-factor and substrate account for the fact that both are almost completely wrapped up by the protein within the cleft formed between the N and C-terminal parts. However, the junction of these two parts of the protein above the binding sites is achieved by relatively weak interactions (see Figure 5-B), i.e., two

hydrophobic contacts involving side chains of F164 and I222, M88 and P204, only one direct hydrogen bond (T159OG1 - A218O) and two hydrogen bond interactions mediated via a single water molecule each (T159O – W178 – S221OG; T159N – W154 – S221OG).

In the absence of any other DFR structure (apo-enzyme or binary complex), our structural data cannot shed light on the conformational changes of the protein upon substrate and co-factor binding, nor on the order of binding of the latter prior to the catalytic reaction. Even so, the interactions responsible for the embedding of the substrate seem to be remarkably weak and should not need much energy to be disrupted upon binding/release of the substrate.

From sequence alignment (Beld et al, 1989) and molecular constructions using chimeric DFR genes (Johnson et al, 2001), several authors have identified the 131 – 156 amino acid sequence as being responsible for the specificity of the substrate. This segment is fully ordered in the ternary complex. It interacts with the rest of the enzyme through its central $\beta 1'$ strand, which establishes a short (3 hydrogen bonds) parallel β sheet with strand $\beta 5'$. At one end of the molecule, it closes the cleft in which both the co-factor and the substrate are embedded. More precisely, in a plane roughly perpendicular to the C2 – C1' bond of the DHQ molecule, it forms, a large loop which covers DHQ ring B. Our structural data confirm the importance of this 27 amino acid fragment which not only takes part in the substrate binding site geometry, but also directly interacts with the DHQ molecule. The OD1 and ND2 side chain atoms of N133 are hydrogen-bonded to the 3' and 4' hydroxyl groups of the substrate respectively. These results represent the first structural evidence of the particular role of this amino acid which has been previously described to affect the substrate preference. Interestingly, the N133L mutant was shown in DFR gerbera to abolish cyanidin or delphinidin synthesis (Johnson et al, 2001), a result which was thought to be due to the bulkier and non-polar side chain of the leucine residue preventing any DHK or DHM binding.

Several DFRs from *L. japonicus* (Shimada et al, 2005) or *M. truncatula* (Xie et al, 2004) which present an asparagine at position 133 reduce DHK more efficiently than DHQ, unlike others (including those of petunia and *Cymbidium*) which have an aspartate at the same position. Such an observation reinforces the putative importance of the N133 residue. Nevertheless, as shown in Figure 2b, our data do not corroborate this result. In *Vitis vinifera*, the reduction is more efficient (66 %) for DHQ than for DHK (50 %). Similar behaviour is observed in the N133D mutant. The reduction is lowered for both substrates (41 % and 27 % for DHQ and DHK respectively) and is still in favour of DHQ. This seems to indicate that the N133D mutation is not the only relevant mutation to specifically block or significantly lower DHK reduction. These results are consistent with the fact that H bond interactions between ring B of the substrate and the surrounding amino acids could be important parameters influencing the affinity and/or the catalytic activity in stabilizing the molecule more or less within a given active site:

- the two hydroxyl groups of ring B plead more for an effective stabilization of the substrate than the sole DHK hydroxyl;

- whatever the substrate, the N133D mutation should not significantly modify the H-bond network, this point still pleads in favour of DHQ binding in the present environment, but remains to be verified. N133D DFR mutant crystallization is in progress.

This assumption about the H-bond importance is not refuted by earlier results (Johnson et al, 2001). The substitution of asparagine to leucine could either disrupt H-bonding or generate steric hindrance for the substrate binding. The N133L and N133A mutants are presently under study to shed light on this point.

In the present structure, the substrate 4'-hydroxyl group is also H-bonded to the O ϵ 1 atom of Q227. This does not affect the above mentioned line of argument as this hydroxyl group is conserved whatever the substrate. At the most, it underlines the importance of Q227 which is strictly conserved in all DFR sequences.

As far as the catalytic mechanism is concerned, the DFR structure shows that the catalytic site exhibits all the geometrical characteristics observed in other SDR proteins.

The presence of a cluster of water molecules in the vicinity of residue K167 is also a common feature to several SDRs. For instance, it was observed in the structure of 3 β /17 β hydroxysteroid dehydrogenase (3 β /17 β -HSD, Filling et al. 2002, Benach et al., 2002) or R-specific alcohol dehydrogenase (RADH, Schlieben et al. 2005), and its role has been underlined. The presence of an asparagine at position 103 (DFR numbering) in 3 β /17 β -HSD and RADH was found to be responsible for a kink in helix α 4. This kink favors the binding of a water molecule to the main chain carbonyl of the asparagine and the amino group of the catalytic lysine. This water molecule is the first of a cluster which transfers a proton between the bulk solvent and the catalytic lysine.

In several respects, the situation in the present structure is different from that described previously:

- In the *Vitis vinifera* DFR structure, there is a threonine at position 103. Helix α 4 also exhibits a sharp kink, but which is one turn ahead of that observed in 3 β /17 β -HSD and RADH. This kink is due to a proline in position 102: the helix break also allows a water molecule to bind the main chain carbonyl group of T103 at a position equivalent to that observed in 3 β /17 β -HSD and RADH. From this point of view, the SSCR structure is much closer to that of DFR since the proline is conserved, although an alanine is observed at position 103.

- None of the water molecules of the cluster has direct access to the bulk solvent. This situation does not completely fulfill the conditions required for the proton relay mechanism described by Filling et al (2002).

METHODS

Cloning and expression

The full-length coding sequence for the DFR protein was PCR amplified with a post-veraison cDNA library from *Vitis vinifera* cabernet sauvignon as template. PCR was performed using the oligonucleotide primers TaaTaggcctATGGGTTACAAAGTGAAACCG and TaaTaagcttCTAGGTCTTGCCATCTAC so that the PCR product contained the *Stu*I and *Hind*III restriction sites at 5' and 3' termini respectively. The PCR program was 35 cycles of 95°C (30 sec)/52°C (30 sec)/72°C (60 sec) followed by a final extension for 2 min at 72°C, using Pfu (Promega). The amplified gene product was ligated into pGEM-T Easy (Promega), sequenced and subcloned into the pQE-30Xa bacterial expression vector (Qiagen). The resulting plasmid was moved into *E. coli* strain M15(pREP4). Cultures were grown at 37°C for 16 hours on 1 % (w/v) Luria-Bertani agar supplemented with 100 µg/mL ampicillin and 25 µg/mL kanamycin.

A single colony was used to inoculate 1 l LB medium supplemented with 100 µg/mL ampicillin and 25 µg/mL kanamycin at 37°C until DO_{600} reached 0.6. Then 1 mM IPTG was introduced to initiate over-expression. The culture was incubated overnight at 30°C by shaking at 160 rev/min. The bacterial culture was harvested by centrifugation and resuspended in ice-cold lysis buffer (50 mM NaH_2PO_4 , 300 mM NaCl, 10 mM Imidazole, 1 mg/mL HEW lysozyme, pH 8). Cells were sonicated for 5 min (Branson Digital Sonifier) and debris were removed by centrifugation at 9000 g for 30 min at 4°C.

To produce selenomethionine (SeMet) substituted DFR, cells were cultured in minimal medium containing SeMet.

The DFR N133D mutant was constructed using the Stratagene QuickChange Site-Directed mutagenesis kit. The following primer and the respective reverse complements were used: TCC GCAGGAAGTGTGGATATTCAAGAACACCAA.

Purification

The purification procedures for SeMet-substituted and both native and mutated DFRs were identical. The supernatant was loaded onto a HiTrap Chelated HP 5-mL column (Amersham Biosciences) equilibrated with 10 mM Tricine, 10 mM imidazole, 500 mM NaCl, pH 7.3 and the protein was eluted with 10mM Tricine, 300 mM imidazole, 500 mM NaCl, at pH 7.3. To separate oligomeric states from monomers, different buffers were tested. Among them, only tricine avoids a precipitated state. The elution fraction was thus loaded onto a Superdex-200 gel filtration column (Amersham Biosciences) equilibrated with 10 mM Tricine, 500 mM NaCl, pH 7.3. For crystallization assays, the final fraction of interest was dialyzed against 10 mM Tricine, 50 mM NaCl, 5 mM DTT, pH 7.3 (buffer A), then concentrated by ultrafiltration. The protein concentration was determined using a theoretical molar extinction coefficient of $41850 M^{-1} cm^{-1}$ at 280 nm, and the purity of the sample was assessed by SDS-PAGE and MALDI-mass spectroscopy. The observed mass of the protein, native DFR plus tag, was

39879 Da, a value which, owing to accuracy of MALDI-MS in this mass range fairly agrees with the integrity of the gene already confirmed by DNA sequencing (theoretical mass 39847.7 Da). The yield for the purified protein was approximately 35 mg per liter of bacterial cell culture. The protein was concentrated to 30 mg/mL in buffer A and stored at 4°C.

Crystallization and data collection

DFR crystals grew at room temperature by hanging drop vapor diffusion in presence of racemic DHQ and NADP⁺ (Fluka, Sigma-Aldrich). Best native DFR crystals are obtained with a 15 mg/ml protein solution in buffer A, supplemented with a solution composed of 5 equivalents of DHQ in ethanol and 5 equivalents of NADP⁺ in buffer A. The protein and co-factors solutions are mixed with an equal volume of reservoir containing 90 mM NaCl, 24 % PEG 3350, 100 mM Hepes pH 7.5 and 3 mM NaN₃. Crystals appear within 48 hours and grow to their final size within 2 weeks. Crystals were cryoprotected by soaking in the crystallizing solution supplemented with increasing amounts of glycerol up to a final concentration of 25%.

Data sets of native and selenomethionyl crystals were collected at 100 K with a MarResearch CCD detector system from the synchrotron radiation source at the ESRF beam line BM30A. Data were processed and scaled using MOSFLM (Leslie et al, 1986) and SCALA from the CCP4 suite (Collaborative Computational Project Number 4, 1994). Data collection statistics are shown in Table 1.

Enzyme assay and LC-MS analysis of DFR reaction products

When Tricine (pH 7.3) was considered as the enzymatic buffer, no significant activity was observed. At this pH, a possible hydroxyl group deprotonation occurring at position 7 of the substrate together with a subsequent charge delocalization up to the C4 position could lead to the formation of a phenolate group. To conjugate the protein stability and a higher enzymatic efficiency, a lower pH was required. Good results were obtained with Bis-Tris and Tricine used together as the buffer solution. Thus the assay mixture contained 50 mM Bis-Tris, 50 mM Tricine (pH 6.5), 50 mM NaCl, 5 mM DTT to which 1 mM NADPH and either 1 mM of (2R,3R)-DHQ or (2R,3R)-DHK (ArboNova, Turku, Finland) were added. The reaction was initiated by addition of 1 μM enzyme and was terminated after 30 min at 30°C by extraction with ethyl acetate (twice). Since leucoanthocyanidin, the natural DFR product obtained by reduction of DHQ, is unstable in solution, the residue was dissolved after evaporation in 1 mL of a mixture of n-BuOH-HCl (95:5, v/v) and incubated at 95°C for 30 min to form cyanidin (Stafford et al, 1982). HPLC was performed using an Atlantis dc18 5 μm 4.6 x 250 mm column (Waters) at a flow rate of 1.0 mL min⁻¹ and a 20 min linear elution gradient from 10% acetonitrile in 0.08% TFA to 90 % acetonitrile in 0.08 % TFA. The wavelength used for detection was 214 nm (996 Photodiode Array Detector, Waters).

LC/MS analysis was performed on a ESI-ion trap mass spectrometer (LCQ Advantage, Thermo Electron) operated in positive ion mode, coupled to a Surveyor HPLC system (Thermo Electron), using the same column as above with a 20 min linear elution gradient from 10% acetonitrile in 0.05% TFA to 90 % acetonitrile in 0.05 % TFA. The 1.0 mL min⁻¹ flow rate was split post-column, a flow of 0.2 mL min⁻¹ only being sent to the ESI source.

Crystal structure determination

According to the unit cell size, the asymmetric unit is expected to be composed of 2 DFR molecules, which yields a Matthews coefficient of 2.4 Å³/Da and a solvent content of 49%. A self rotation Patterson map calculation performed with the program Polar Refine (CCP4 program suite) confirms the presence of a 2-fold non-crystallographic symmetry axis parallel to the diagonal of the unit cell (**a,b**) plane (polar angles $\Omega = 90.02$, $\Phi = 133.87$, $X = 179.17$).

MAD phasing was performed at 2.8 Å resolution using the programs SHELXD (Schneider and Sheldrick, 2002) and SHELXE (Sheldrick, 2002). Twenty-two selenium atoms (i.e. 11 per molecule) were located in the peak dataset. Starting from these positions, the phases were generated using the program SOLVE (Terwilliger and Berendzen, 1999) and further improved by solvent flattening and histogram matching with the program RESOLVE (Terwilliger, 2000, 2003). From the resulting electron density map, RESOLVE could build a model consisting of 2 poly-alanine/glycine chains of 298 residue each, which corresponds to 88% of the protein structure. This model was further modified and completed using the graphics Turbo-Frodo (Roussel et al. 1990). Prior to any refinement, each chain contained 320 residues out of 337. The main chain atoms of residues 1-5, 91-93, 227, 268 and 331-337 were missing, as well as 33 side chains.

The dataset obtained at 1.8 Å on the native crystal was then used for further structure refinement. The latter was performed using the REFMAC program (Murshudov et al. 1997) within the CCP4 program suite and was controlled throughout with the programs PROCHECK (Laskowski et al. 1993), SFCHECK (Vaguine et al 1999) and WHATIF (Vriend and Sander 1993). Protein stereochemical parameters were taken from Engh and Huber (1991). Hetero compound (NADP⁺ and DHQ) coordinates were taken from the Hic-up server (Kleywegt and Jones 1998). The four stereo-isomeric forms of DHQ, with the carbon atoms in positions 2 and 3 adopting R or S configurations, were built via geometric restraints generated with REFMAC. After several cycles of refinement and model building, the residual R factor dropped from 0.451 (freeR = 0.458 and fom = 0.375) to 0.294 (freeR = 0.328, fom = 0.689). When NADP⁺ and DHQ molecules were added, the Rfactor dropped to 0.269 (freeR = 0.300, fom = 0.713). Then, 576 water molecules were positioned in well-defined positive (mFo-DFc) residual densities with a lower cut-off of 3 σ , if they participated in hydrogen bonds to the protein, the co-substrates or other water molecules. Final statistics are shown in Table 2. Atomic coordinates have been deposited to the Protein Data bank (www.rcsb.org/pdb/) as entry 2c29. Figures were drawn with PyMol (DeLano, W. L., 2002).

Quality of the model

For each protein molecule of the asymmetric unit, a few residues could not be positioned in the (2mFo-DFc) electron density map: i.e. residues 1-5 and 330-337 for chain A, and residues 1-4 and 331-337 for chain B. For both chains, however, 18 side chains could not be fully built. Both chains of the asymmetric unit were NCS restrained and appear to be almost identical, with an overall r.m.s. deviation between the main chain atom positions of 0.34 Å. A slight difference in main chain conformations for region K112-V120 is visible due to crystal packing intermolecular contacts. A Ramachandran plot shows 91% of the residues belonging to the most favorable regions, 8% to additionally allowed regions and only 0.7% to generously allowed regions.

ACKNOWLEDGEMENTS

We thank the European Synchrotron Radiation Facility for provision of synchrotron radiation facilities and the staff of the FIP-BM30A beamline for their kind assistance during the data collection. We also thank Prof. Serge Delrot (Institut de la Vigne et du Vin, Villenave d'Ornon, France) for critical reading of the manuscript and Prof. Jean Chaudière (Université Bordeaux I, Talence, France) for fruitful discussions. This work is supported by the CIVB (Conseil Interprofessionnel du Vin de Bordeaux) grant N° 8915.

FOOTNOTES

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantcell.org) is: Bernard Gallois (b.gallois@ubs.u-bordeaux1.fr).

REFERENCES

- Beld, M., Martin, C., Huits, H., Stuitje, A.R. and Gerats, A.G.** (1989). Flavonoid synthesis in *Petunia hybrida*: partial characterization of dihydroflavonol-4-reductase genes. *Plant Mol. Biol.* **13**, 491-502.
- Benach, J., Filling, C. Oppermann, U. C., Roversi, P., Bricogne, G., Berndt, K. D., Jörnvall, H. and Ladenstein, R.** (2002). Structure of bacterial 3 β /17 β -hydroxysteroid dehydrogenase at 1.2 Å resolution: A model for multiple steroid recognition. *Biochemistry* **41**, 14659-14668.
- Bongue-Bartelsman, M., O'Neill, S.D., Tong, Y. and Yoder, J.I.** (1994). Characterization of the gene encoding dihydroflavonol 4-reductase in tomato. *Gene* **138**, 153-157.
- Bottoms, C.A., Smith, P.E. and Tanner, J.J.** (2002). A structurally conserved water molecule in Rossmann dinucleotide-binding domains. *Protein Sci.* **11**, 2125-2137.
- Charrier, B., Coronado, C., Kondorosi, A. and Ratet, P.** (1995). Molecular characterization and expression of alfalfa (*Medicago sativa L.*) flavanone-3-hydroxylase and dihydroflavonol-4-reductase encoding genes. *Plant Mol. Biol.* **29**, 773-786.
- Chen, M., SanMiguel, P. and Bennetzen, J.L.** (1998). Sequence organization and conservation in sh2/a1-homologous regions of sorghum and rice. *Genetics* **148**, 435-443.
- Collaborative Computational Project, Number 4 (1994). The CCP4 suite: program for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **50**, 760-763.
- DeLano, W.L.** (2002). The PyMOL Molecular Graphics System. *DeLano Scientific*, San Carlos, CA, USA.
- Dell'agli, M., Busciala, A., Bosisio, E.** (2004). Vascular effects of wine polyphenols. *Cardiovasc. Res.* **63**, 593-602.
- De Lorgeril, M., Salen, P., Martin, J.L., Mamelle, N., Monjaud, I., Touboul, P. and Delaye, J.** (1996). Effect of a mediterranean type of diet on the rate of cardiovascular complications in patients with coronary artery disease. Insights into the cardioprotective effects of certain nutrients, *J. Am. Coll. Cardiol.*, **28**, 1103-1108.
- Engh, R.A. and Huber, R.** (1991). Accurate bond and angle parameters for X-ray protein structure refinement. *Acta Crystallogr. A Foundations of Crystallogr.* **47**, 392-400.
- Filling, C., Berndt, K.D., Benach, J., Knapps, S., Prozorovski, T., Nordling, E., Ladenstein, R., Jörnvall, H. and Oppermann, U.** (2002). Critical residues for structure and catalysis in short-chain dehydrogenases/reductases. *J. Biol. Chem.*, **277**, 25677-25684.
- Fischer, T.C., Halbwirth, H., Meisel, B., Stich, K. and Forkmann, G.** (2003). Molecular cloning, substrate specificity of the functionally expressed dihydroflavonol 4-reductases from *Malus domestica* and *Pyrus communis* cultivars and the consequences for flavonoid metabolism. *Arch. of Biochem. and Biophys.*, **412**, 223-230.
- Forkmann, G. and Ruhnau, B.** (1987). Distinct substrate specificity of dihydroflavonol 4-reductase from flowers of *Petunia hybrida*. *Z. Naturforsch.*, **42C**, 1146-1148.
- Gawel, R.** (1998). Red wine astringency: A review. *Austral. J. Grape Wine Res.*, **4**, 74-95.
- Gong, Z., Yamazaki, M., Sugiyama, M., Tanaka, Y. and Saito, K.** (1997). Cloning and molecular analysis of structural genes involved in anthocyanin biosynthesis and expressed in a form-specific manner in *Perilla frutescens*. *Plant Mol. Biol.* **35**, 915-927.
- Halliwell, B., Rafter, J., Jenner, A.** (2005). Health promotion by flavonoids, tocopherols, tocotrienols, and other phenols: direct or indirect effects? Antioxidant or not? *Am. J. Clin. Nutr.* **81**, 268s-276s.
- Helariutta, Y., Elomaa, P., Kotilainen, M., Seppanen, P. and Teeri, T.H.** (1993). Cloning of cDNA coding for dihydroflavonol-4-reductase (DFR) and characterization of dfr expression in the corollas of *Gerbera hybrida* var. Regina (Compositae). *Plant Mol. Biol.* **22**, 183-193.
- Hertog, M.G., Kromhout, D., Aravanis, C., Blackburn, H., Buzina, R., Fidanza, F., Giampaoli, S., Jansen, A., Menotti, A., Nedejkovic, S., et al.** (1995). Flavonoid intake and long-term risk of coronary heart disease and cancer in the seven countries study. *Arch. Intern. Med.*, **155**, 381-386.
- Himi, E. and Noda, K.** (2004). Isolation and location of three homologous dihydroflavonol-4-reductase (DFR) genes of wheat and their tissue-dependent expression. *Journal of Exp. Bot.*, **55**, 365-375.

- Holm, L and Park, J.** (2000). DaliLite workbench for protein structure comparison. *Bioinformatics* **16**, 566-567.
- Inagaki, Y., Johzuka-Hisatomi, Y., Mori, T., Takahashi, S., Hayakawa, Y., Peyachoknagul, S., Ozeki, Y. and Iida, S.** (1999). Genomic organization of the genes encoding dihydroflavonol 4-reductase for flower pigmentation in the Japanese and common morning glories. *Gene.*, **226**, 181-188.
- Ito, K., Nakajima, Y., Ichihara, E., Ogawa, K., Katayama, N., Nakashima, K. and Yoshimoto, T.** (2006). D-3-Hydroxybutyrate dehydrogenase from *Pseudomonas fragi*: molecular cloning of the enzyme gene and crystal structure of the enzyme. *J. Mol. Biol.* **355**, 722-733.
- Johnson, E.T., Ryu, S., Yi, H., Shin, B., Cheong, H. and Choi, G.** (2001). Alteration of a single amino acid changes the substrate specificity of dihydroflavonol 4-reductase. *Plant J.* **25**, 325-333.
- Jörnvall, H., Persson, B., Krook, M., Atrian, S., Gonzalez-Duarte, R., Jeffery, J. and Ghosh, D.** (1995). Short-chain dehydrogenase/reductase (SDR). *Biochemistry* **34**, 6003-6013.
- Kamitori, S., Iguchi, A., Ohtaki, A., Yamada, M. and Kita, K.** (2005). X-ray structure of NADPH-dependent carbonyl reductase from *Sporobolomyces salmonicolor* provide insights into stereoselective reductions of carbonyl compounds. *J. Mol. Biol.* **352**, 551-558.
- Kleywegt, G. J. and Jones, T. A.** (1998). Databases in protein crystallography. (CCP4 Proceedings). *Acta Crystallogr. D Biol. Crystallogr.* **54**, 1119-1131.
- Laskowski, R. A., MacArthur, M. W., Moss, D. S. and Thornton, J. M.** (1993). PROCHECK : a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283-291.
- Leslie, A. G., Brick, P. and Wonacott, A. J.** (1986) Daresbury Laboratory Information quarterly for protein crystallography *CCP4 News* **18**, 33-39.
- Martens, S., Teeri, T. and Forkmann, G.** (2002). Heterologous expression of dihydroflavonol 4-reductases from various plants. *FEBS Letters* **531**, 453-458.
- Murshudov, G. N., Vagin, A.A. and Dodson, E.J.** (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.* **53**, 240-255.
- Offen, W., Martinez-Fleites, C., Yang, M., Kiat-Lim, E., Davis, B.G., Tarling, C.A., Ford, C.M., Bowles, D.J. and Davies, G.J.** (2006). Structure of a flavonoid glucosyltransferase reveals the basis for plant natural product modification. *The EMBO J.* **25**, 1396-1405).
- O'Reilly, C., Shepherd, N.S., Pereira, A., Schwarz-Sommer, Z., Bertram, I., Robertson, D.S., Peterson, P.A. and Saedler, H.** (1985). Molecular cloning of the *al* locus of *Zea mays* using the transposable elements *En* and *Mu1*. *EMBO J.*, **4**, 877-882.
- Peters, D.J. and Constabel, C.P.** (2002). Molecular analysis of herbivore-induced condensed tannin synthesis: cloning and expression of dihydroflavonol reductase from trembling aspen (*Populus tremuloides*). *Plant J.*, **32**, 701-712.
- Ross, J.A. and Kasum, C.M.** (2002). Dietary flavonoids: bioavailability, metabolic effects, and safety. *Annu. Rev. Nutr.*, **22**, 19-34.
- Roussel, A., Fontecilla-Camps, J. C. and Cambillau, C.** (1990). TURBO-FRODO: A new program for protein crystallography and modeling. XV IUCr Congress, Bordeaux, France. C66-C67.
- Sheldrick, G. M.** (2002). Macromolecular phasing with SHELXE. *Z. Kristallogr.* **217**, 644-650.
- Shimada, N., Sasaki, R., Sato, S., Kaneko, T., Tabata, S., Aoki, T. and Ayabe, S.** (2005). A comprehensive analysis of six dihydroflavonol 4-reductases encoded by a gene cluster of the *Lotus japonicus* genome. *J. Exp. Bot.*, **56**, 2573-2585.
- Schlieben, N. H., Niefind, K., Müller, J., Riebel, B., Hummel, W. and Schomburg, D.** (2005). Atomic resolution structures of R-specific alcohol dehydrogenase from *Lactobacillus brevis* provide the structural bases of its substrate and cosubstrate specificity. *J. Mol. Biol.* **349**, 801-813.
- Schneider, T. R. and Sheldrick, G. M.** (2002). Substructure solution with SHELXD. *Acta Crystallogr. D Biol. Crystallogr.* **58**, 1772-1779.
- Sparvoli, F., Martin, C., Scienza, A., Gavazzi, G. and Tonelli, C.** (1994). Cloning and molecular analysis of structural genes involved in flavonoid and stilbene biosynthesis in grape (*Vitis vinifera L.*). *Plant Mol. Biol.* **24**, 743-755.
- Stafford, H.A. and Lester, H. H.** (1982). Enzymic and nonenzymic reduction of (+)-Dihydroquercetin to its 3,4-diol. *Plant Physiol.*, **70**, 695-698.
- Tanaka, Y., Fukui, Y., Fukuchi-Mizutani, M., Holton, T.A., Higgins, E., and Kusumi, T.** (1995). Molecular cloning and characterization of *Rosa hybrida* dihydroflavonol 4-reductase gene. *Plant Cell Physiol.*, **36**, 1023-1031.

- Terwilliger, T.C. and Berendzen, J.** (1999). Automated MAD and MIR structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **55**, 849-861.
- Terwilliger, T.C.** (2000). Maximum-likelihood density modification. *Acta Crystallogr. D Biol. Crystallogr.* **56**, 965-972.
- Terwilliger, T.C.** (2003). Automated main-chain model-building by template matching and iterative fragment extension. *Acta Crystallogr. D Biol. Crystallogr.* **59**, 38-44.
- Vaguine, A. A., Richelle, J. and Wodak, S. J.** (1999). SFCHECK: a unified set of procedures for evaluating the quality of macromolecular structure-factor data and their agreement with the atomic model. *Acta Crystallogr. D Biol. Crystallogr.* **55**, 191-205.
- Vriend, G. and Sander, C.** (1993). Quality control of protein models: directional atomic contact analysis. *J. Appl. Cryst.* **26**, 47-60.
- Wilmouth, R.C., Turnbull, J.J., Welford, R.W., Clifton, I.J., Prescott, A.G. and Schofield, C.J.** (2002). Structure and mechanism of anthocyanidin synthase from *Arabidopsis thaliana*. *Structure*, **10**, 93-103.
- Winkel-Shirley, B.** (2001). Flavonoid biosynthesis. A colorful model for genetics, biochemistry, cell biology, and biotechnology. *Plant Physiol.*, **126**, 485-493.
- Xie, D-Y., Jackson, L.A., Cooper, J.D., Ferreira D. and Paiva, N.L.** (2004). Molecular and biochemical analysis of two cDNA clones encoding dihydroflavonol 4-reductase from *Medicago truncatula*. *Plant Physiol.* **134**, 979-994.
- Yamashita, A. Kato, H. Wakatsuki, S. Tomizaki, T. Nakatsu, T. Nakajima, K. Hashimoto, T. Yamada, Y. and Oda, J.** (1999). Structure of tropinone reductase-II complexed with NADP⁺ and pseudotropine at 1.9 Å resolution: implication for stereospecific substrate binding and catalysis. *Biochemistry* **38**, 7630-7.

TABLE 1. Data collection statistics

	Native	Se-Met		
		Peak	Edge	remote
Cell parameters				
a (Å)	87.818	88.068	88.068	88.068
b (Å)	89.927	89.714	89.714	89.714
c (Å)	93.191	93.288	93.288	93.288
Wavelength	0.97319	0.98074	0.98090	0.97320
Resolution range	16.43 – 1.81 (1.91 – 1.81)	22.57 – 2.33 (2.45 – 2.33)	22.57 – 2.33 (2.45 – 2.33)	22.57 – 2.31 (2.43 – 2.31)
Total Nbr of reflections	353174 (31321)	104516 (1602)	104626 (1604)	106245 (1525)
Nbr of unique reflections	66981 (4378)	27567 (649)	27589 (650)	28226 (622)
Completeness (%)	98.9 (98.1)	94.7 (94.7)	94.6 (94.6)	94.8 (94.8)
Data redundancy	5.3 (3.3)	3.8 (1.8)	3.8 (1.8)	3.7 (1.8)
$\langle I/\sigma(I) \rangle$	16.7 (2.0)	20.5 (4.7)	20.1 (4.3)	17.9 (3.3)
R _{merge}	0.07 (0.59)	0.045 (0.137)	0.045 (0.153)	0.053 (0.306)
R _{anom}		0.067 (0.149)	0.052 (0.146)	0.048 (0.169)

Table 2. Refinement statistics

Resolution range	16.43 – 1.81 (1.86 – 1.81)
Nbr of reflections	63444 (4378)
Completeness	98.6 (94.3)
R _{work}	0.192 (0.341)
R _{free}	0.244 (0.376)
r.m.s.d. bond lengths (Å)	0.012
r.m.s.d. bond angles (°)	1.43
No of amino acids	650
No of solvent molecules	576

FIGURE LEGENDS

Fig. 1. Final steps of the flavonoid biosynthetic pathway leading to the synthesis of anthocyanins and condensed tannins (proanthocyanidins). DFR, dihydroflavonol 4-reductase; ANR, anthocyanidin reductase; ANS, anthocyanidin synthase; LAR, leucoanthocyanidin reductase; UFGT, 3-O-glycosyltransferase.

Fig. 2. Schematic HPLC chromatograms of the reaction products after acid treatment at 95°C of both the substrate (DHQ or DHK) and NADPH in absence (a and c) or in presence (b and d) of the enzyme. Relative activity (e) of the wild type and the mutated DFRs towards DHQ and DHK.

Fig. 3. (A) DFR topology. The NADPH binding domain (N-term domain) is composed of a seven strand β sheet (yellow) surrounded by cyan (above) and green (beneath) α helices. The C-terminal region is colored pink. Next to the N-terminal, the glycine-rich motif constituting the NADPH binding site is colored in red.

(B) Stereo view representing the DFR 3D structure with its cofactor NADP⁺ and substrate DHQ bound to the active site. Coloring of secondary structure elements is the same as in (A). The bound NADP⁺ and DHQ are shown by a ball-and-stick model with light and dark grey carbon atoms respectively.

Fig. 4.

(A) Superimposition of the structure of DFR and SSCR as performed by The DALI server. SSCR is colored in blue. The most important discrepancy is observed for the C-terminal domain.

(B) NADP⁺ binding site of DFR. Thin black lines depict the superimposed corresponding region of the SSCR structure co-crystallized with NADPH. Red spheres and black crosses represent water molecules belonging to the DFR and SSCR structures respectively. Carbon atoms of the catalytic triad (S128, Y163 and K167) are colored pink whereas those of the nicotinamide part of NADP⁺ are colored green.

(C) Stereo view of the substrate interactions with the protein and the NADP⁺. Same color code as for (B). DHQ carbon atoms are colored dark grey.

Fig. 5.

(A) Geometry of the catalytic triad site. Same color code as in Figures 4.

(B) Stereo view of the substrate binding site. DHQ is wrapped up by the N and C-terminal parts of the protein: contacts are established via hydrophilic interactions or hydrophobic residues of both N and C-terminal domains.

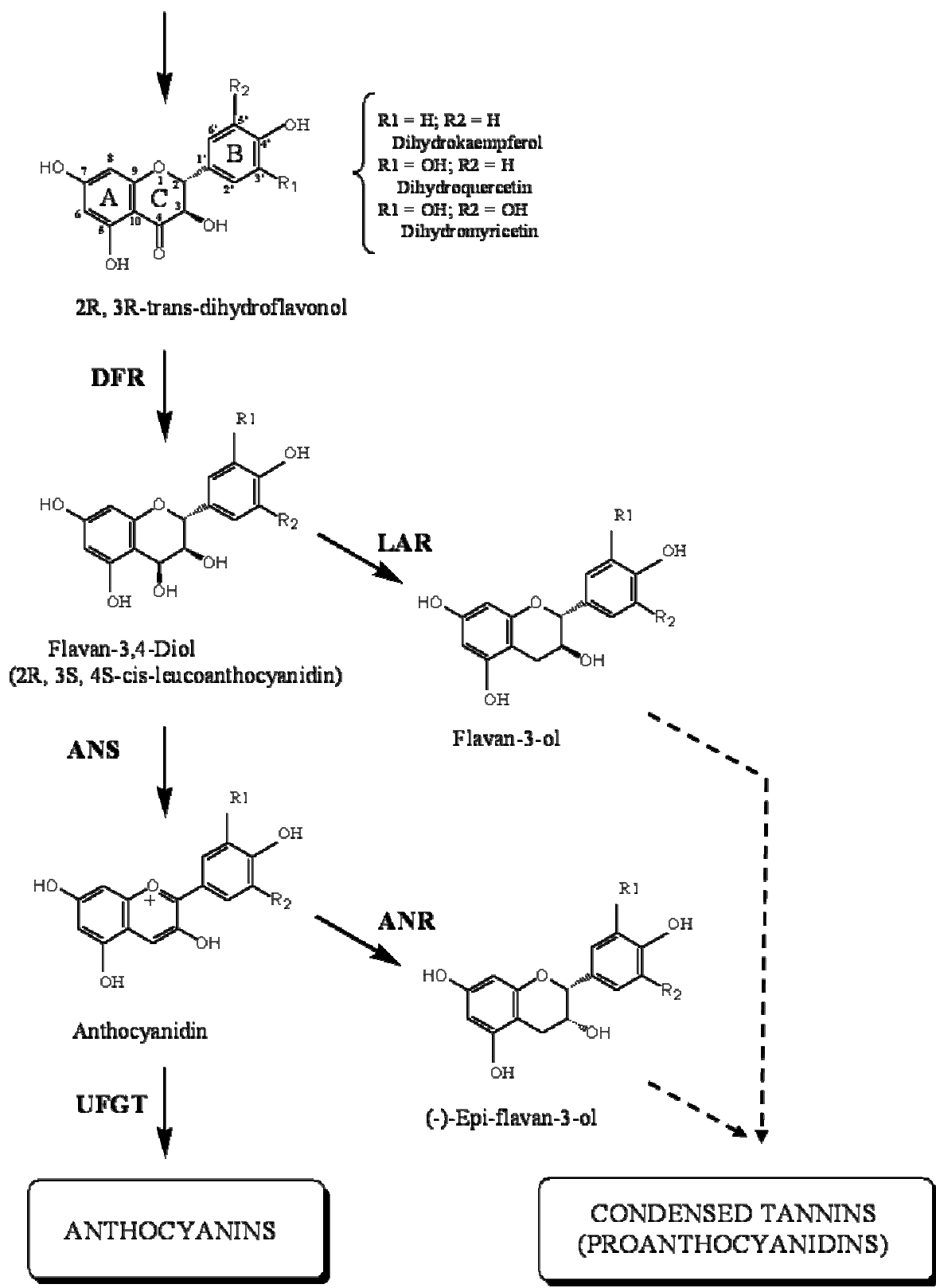


Figure1

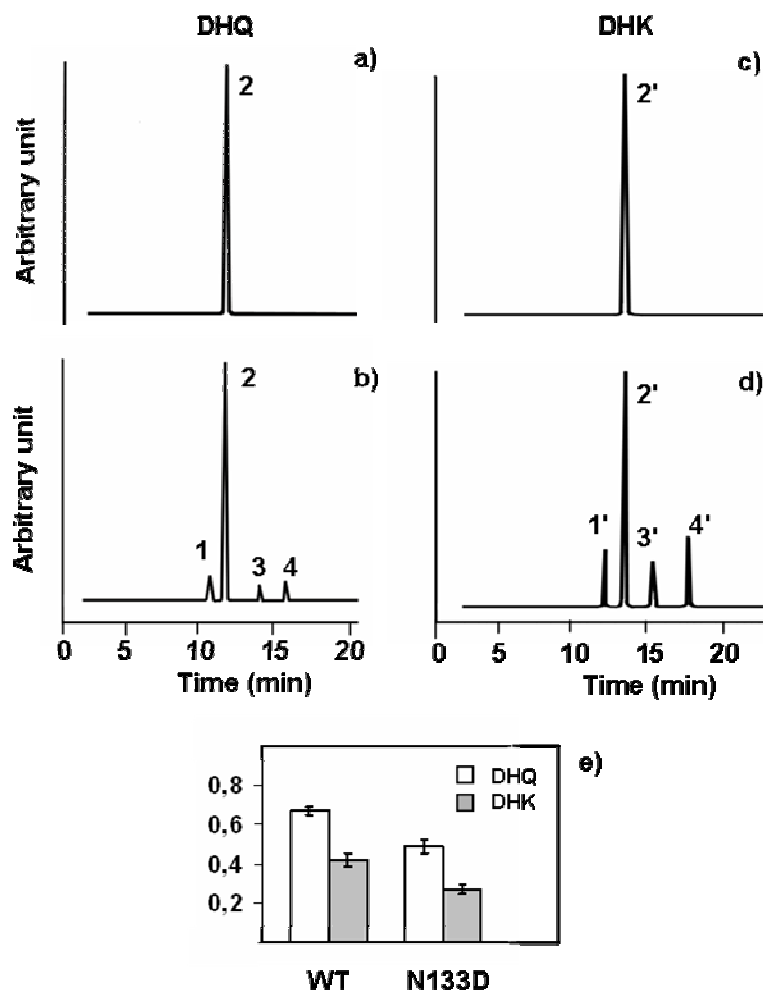
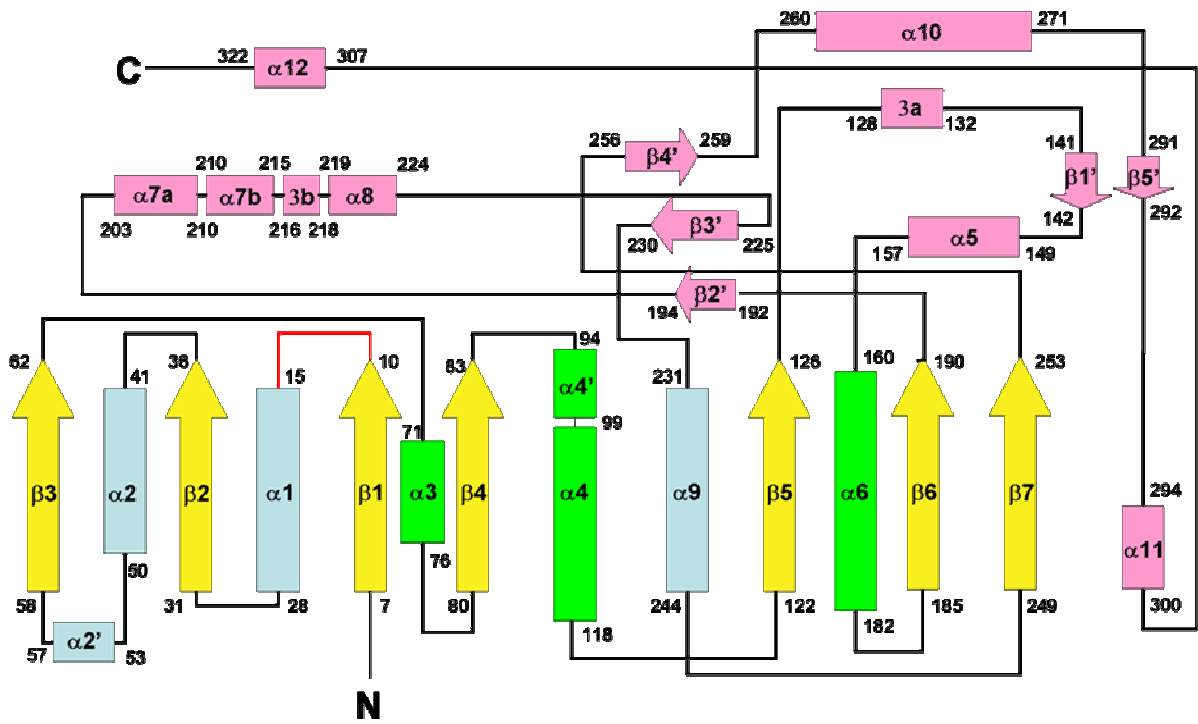
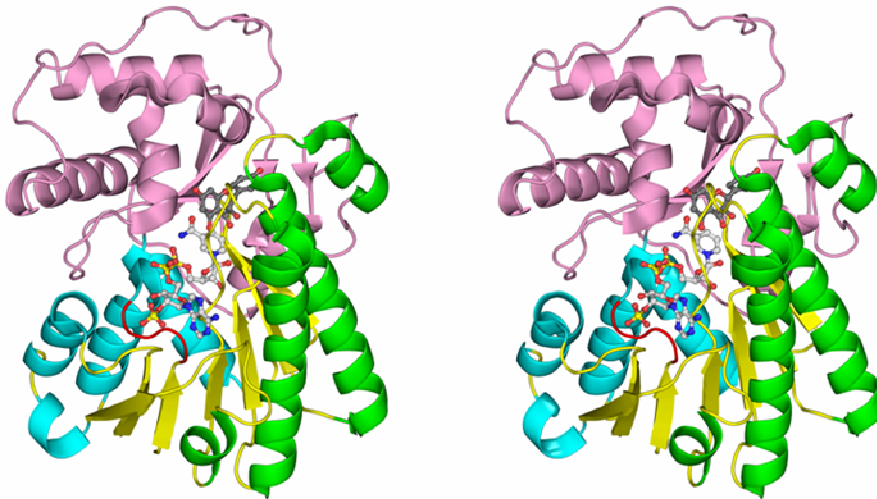


Figure 2

A**B****Figure 3**

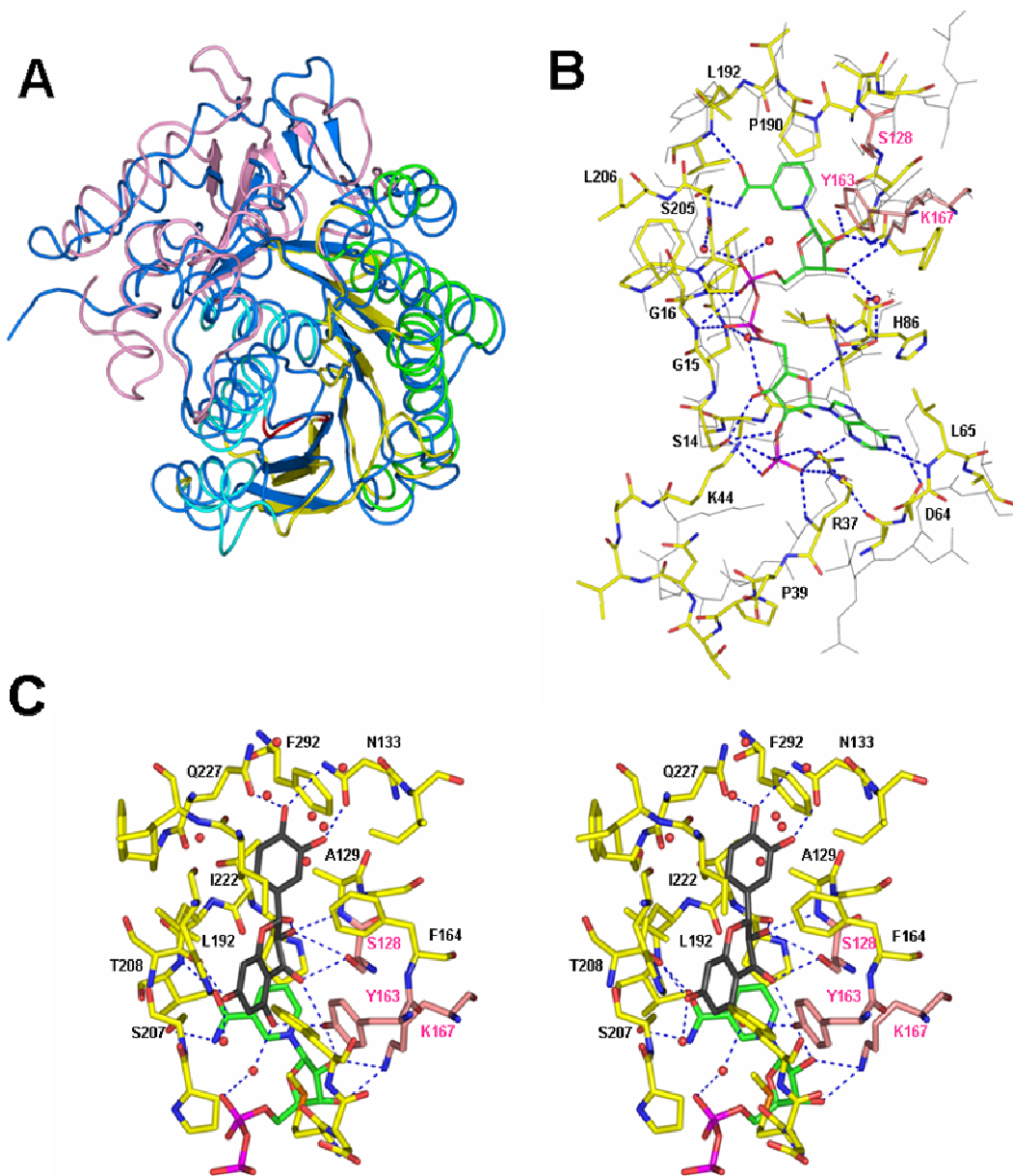
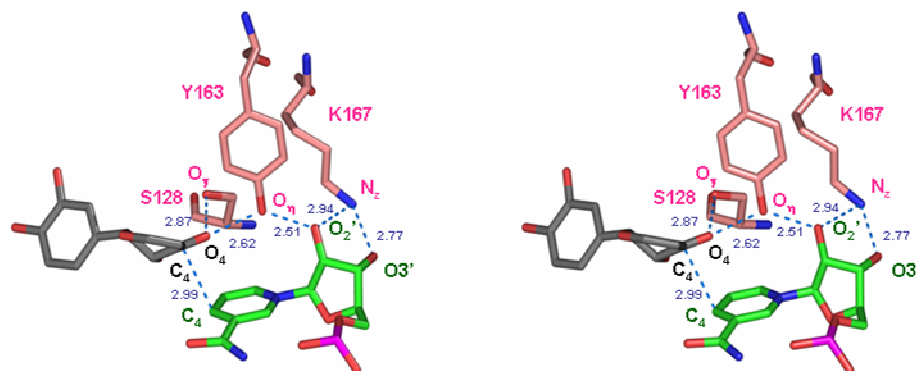
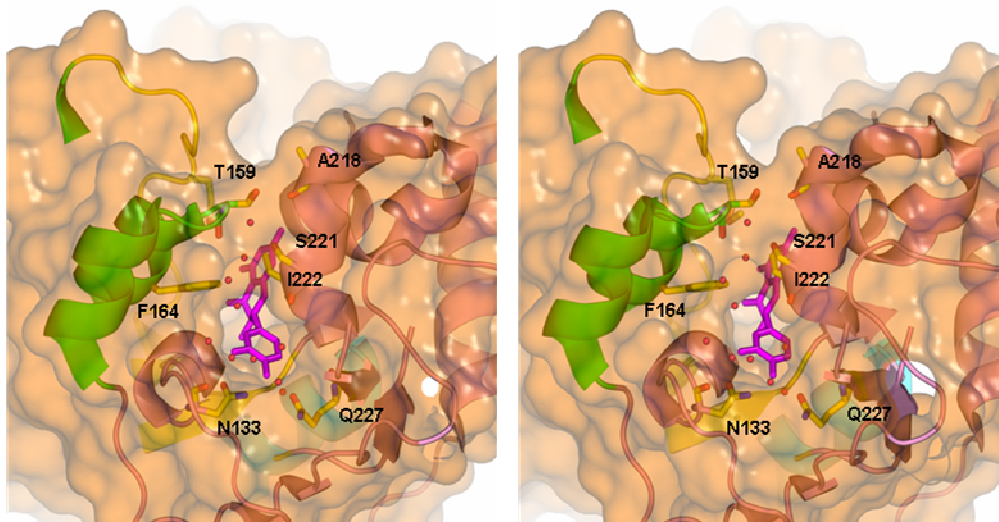


Figure 4

A**B****Figure 5**