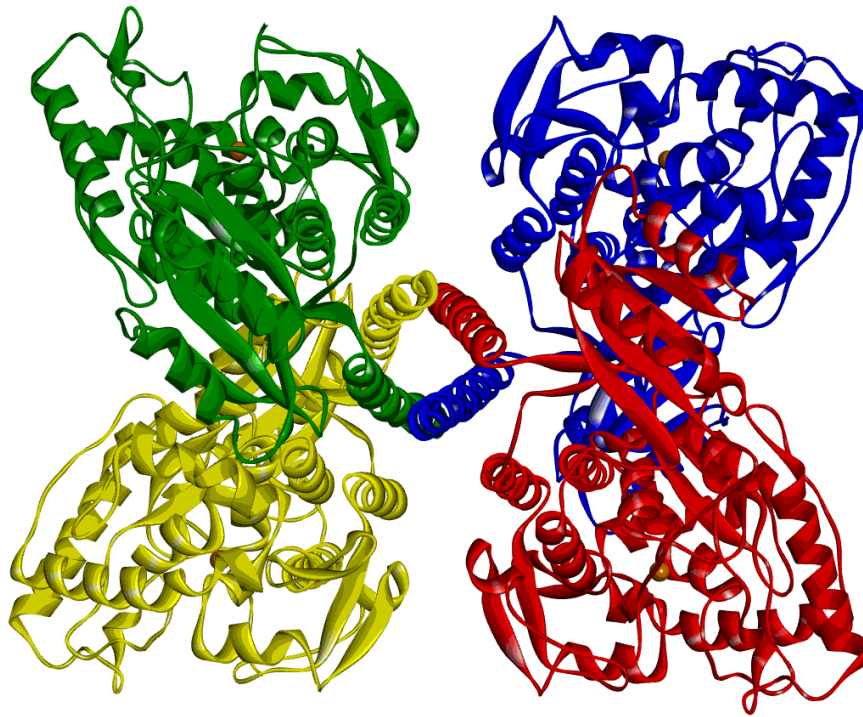


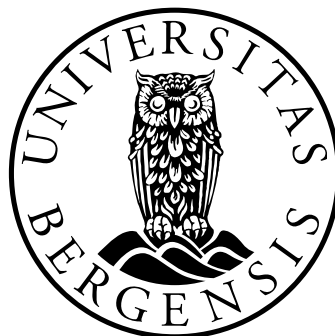
# *Identification of pharmacological chaperones for Phenylalanine Hydroxylase*

*A virtual screening approach to discover novel drug candidates for treatment of phenylketonuria*



Master thesis in Pharmacy

*Gerd Marie Eskerud Harris*



Centre for Pharmacy and  
Department of Biomedicine

University of Bergen

*May 2014*



## **ACKNOWLEDGEMENTS**

The work presented in this Master thesis was performed in the Biorecognition group at the Department of Biomedicine at the University of Bergen from August 2013 to May 2014.

First and foremost, I would like to thank my supervisor, Knut Teigen, for his guidance, support and engagement throughout this project, and for always being so positive and encouraging. I would also like to thank my co-supervisor, Aurora Martinez, for appreciated feedback in the final stages of the writing process.

My thanks also go to the Biorecognition group, especially to Magnus Hole, for appreciated guidance in the lab, and to Rikke Landsvik Berg, for the collaboration throughout our Master projects. I also wish to thank Erlend Hodneland for giving me statistical advice.

The last year has been exiting, challenging and rewarding, and the submission of this thesis marks the end of my pharmacy education. Thank you, kull09, for five memorable years in Bergen and Norwich. A special thanks to Synnøve and Linda ó I will miss sharing everyday life with you.

Finally, I would like to thank my parents for supporting me, like you always do.

*Gerd Marie Eskerud Harris*

May 2014

# TABLE OF CONTENTS

<b>ABSTRACT.....</b>	<b>2</b>
<b>ABBREVIATIONS.....</b>	<b>3</b>
<b>1 INTRODUCTION .....</b>	<b>5</b>
1.1 Phenylalanine hydroxylase .....	5
1.2 The aromatic amino acid hydroxylases .....	11
1.3 Phenylketonuria .....	13
1.4 Protein folding and misfolding diseases.....	14
1.5 Pharmacological chaperones .....	15
1.6 Computational drug design.....	17
<b>2 AIMS OF THE PROJECT.....</b>	<b>18</b>
<b>3 MATERIALS AND METHODS.....</b>	<b>19</b>
3.1 Materials .....	19
3.2 Target-based virtual screening.....	22
3.3 Experimental methods.....	28
<b>4 RESULTS.....</b>	<b>34</b>
4.1 Target-based virtual screening.....	34
4.2 Testing virtual hits experimentally .....	41
<b>5 DISCUSSION.....</b>	<b>49</b>
5.1 Target-based virtual screening.....	50
5.2 Testing virtual hits experimentally .....	54
5.3 Novel compounds with an effect on PAH.....	57
<b>6 FUTURE PERSPECTIVES .....</b>	<b>62</b>
<b>7 CONCLUDING REMARKS.....</b>	<b>63</b>
<b>8 REFERENCES .....</b>	<b>64</b>
<b>9 APPENDIX .....</b>	<b>72</b>

## ABSTRACT

Phenylalanine hydroxylase (PAH) is an enzyme that catalyses the hydroxylation of phenylalanine into tyrosine in the liver, which is the rate-limiting step in phenylalanine catabolism. PAH is a non-heme iron-dependent enzyme that also requires tetrahydrobiopterin (BH<sub>4</sub>) as cofactor to perform catalysis. PAH dysfunction results in phenylketonuria (PKU), characterized by neurotoxic accumulation of phenylalanine. PKU is caused by mutations in the *PAH* gene, usually resulting in a misfolding of PAH. A novel approach to treating misfolding diseases in general, is the use of pharmacological chaperones ó small molecules that can stabilize the native form of a protein and thus prevent misfolding and rescue protein function.

The aim of this Master project was to combine virtual and experimental methods to identify compounds that could act as pharmacological chaperones for PAH. The docking program Glide was used to screen a virtual library, and the compounds with the highest binding affinities were tested experimentally to validate their interactions with PAH, by testing thermostability and enzymatic activity of PAH in the presence of the compounds.

Target-based virtual screening and subsequent experimental validation identified a compound with a potential in pharmacological chaperone therapy for PKU. This compound had a significant protective effect on PAH activity, and its specificity and thermodynamic binding properties should be further investigated.

## ABBREVIATIONS

4 -OH-BH <sub>4</sub> :	Pterin-4 -carbinolamine
5-OH-Trp:	5-hydroxytryptophan
AAAH:	Aromatic amino acid hydroxylase
AAPA:	(2S, 3R)-3-amino-2-hydroxy-4-phenylbutyric acid
BH <sub>4</sub> :	(6R)-L-erythro-5,6,7,8-tetrahydrobiopterin
BSA:	Bovine serum albumin
CNS:	Central nervous system
DHFR:	Dihydrofolate reductase
DHPR:	Dihydropteridin reductase
DSF:	Differential scanning fluorimetry
DMSO:	Dimethyl sulfoxide
FDR:	False discovery rate
GIT:	Gastrointestinal tract
HEPES:	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HPLC:	High performance liquid chromatography
eHTS:	Experimental high throughput screening
HPA:	Hyperphenylalaninemia
hPAH:	Human phenylalanine hydroxylase
ITC:	Isothermal titration calorimetry
L-DOPA:	(S)-3,4-dihydroxyphenylalanine
LNAAs:	Large neutral amino acids
L-Phe:	L-phenylalanine
L-Tha:	L-thienylalanine
L-Trp:	L-tryptophan
L-Tyr:	L-tyrosine
L-Nle:	L-norleucine
LSDs:	Lysosomal storage diseases
PAH:	Phenylalanine hydroxylase
PAL:	Phenylalanine ammonia lyase
PCD:	Pterin carbinolamine dehydratase
PCR:	Polymerase chain reaction

PDB:	Protein Data Bank
PKU:	Phenylkentonuria
RMSD:	Root-square-mean deviation
SD:	Standard deviation
SEM:	Standard error of the mean
TH:	Tyrosine hydroxylase
THD:	Tyrosine hydroxylase deficiency
TPH:	Tryptophan hydroxylase
vHTS:	Virtual high throughput screening
wt-PAH:	Wild-type PAH

# 1 INTRODUCTION

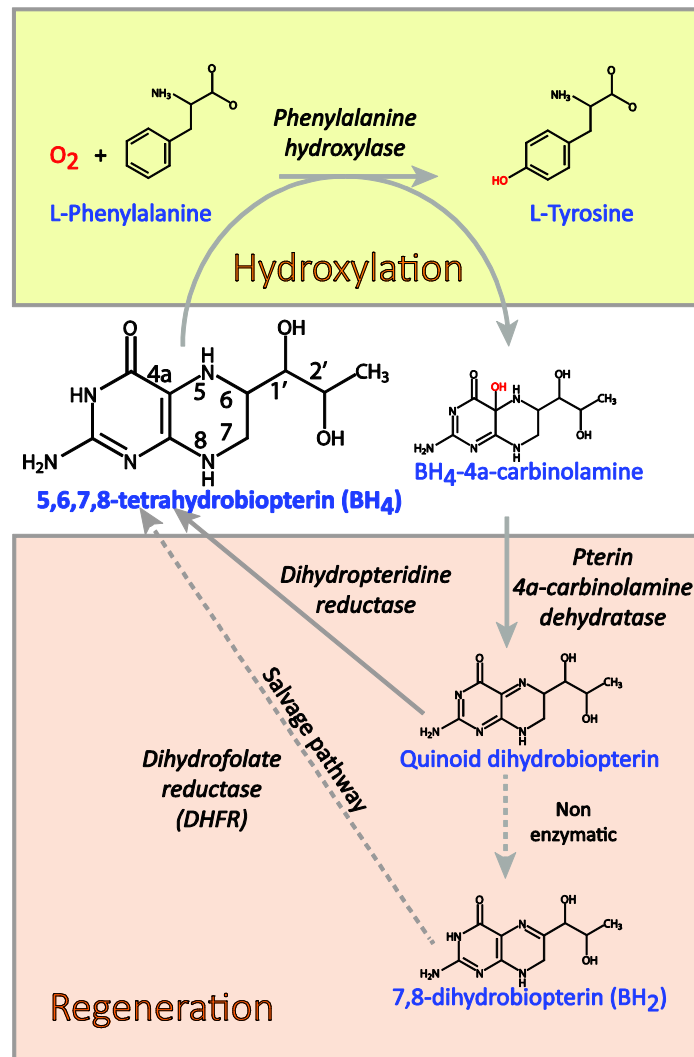
## 1.1 Phenylalanine hydroxylase

Phenylalanine hydroxylase (PAH) is an enzyme that catalyzes the hydroxylation of phenylalanine (L-Phe) into tyrosine (L-Tyr), which is the rate limiting step in the degradation of excess L-Phe in the liver. PAH dysfunction is caused by mutations in the *PAH* gene, and results in phenylketonuria (PKU), characterized by neurotoxic accumulation of phenylalanine (1, 2).

### 1.1.1 Catalytic mechanism

PAH is a non-heme iron- and (6R)-L-erythro-5,6,7,8-tetrahydrobiopterin (BH<sub>4</sub>)-dependent enzyme, and the catalysis also requires O<sub>2</sub> as additional substrate. The cofactor BH<sub>4</sub> is the source of electrons in the catalytic mechanism, and the complete catalytic cycle, including the regeneration of the cofactor, is illustrated in Figure 1-1. Oxygen binds to Fe<sup>2+</sup>, which is activated by BH<sub>4</sub>, forming a Fe<sup>2+</sup>-O-O-BH<sub>4</sub> complex. The O-O bond is cleaved, and a Fe(IV)=O intermediate is formed, hydroxylating phenylalanine to tyrosine. During catalysis, BH<sub>4</sub> is hydroxylated to pterin-4 -carbinolamine (4 -OH-BH<sub>4</sub>), and after each catalytic cycle BH<sub>4</sub> has to be regenerated. The enzyme pterin carbinolamine dehydratase (PCD) converts 4 -OH-BH<sub>4</sub> to q-BH<sub>2</sub>, which is converted to BH<sub>4</sub> by either dihydropteridin reductase (DHPR) or through BH<sub>2</sub> by dihydrofolate reductase (DHFR) (see reviews (3-5)).

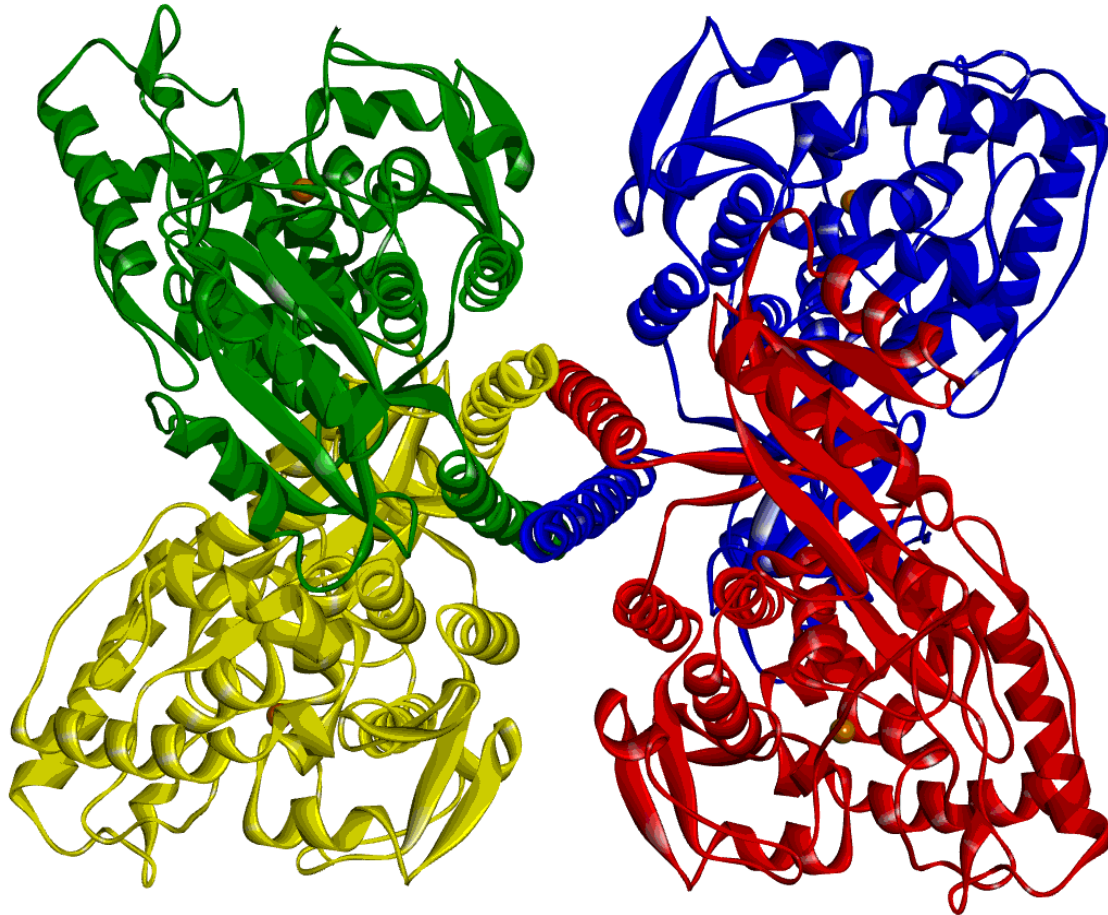




**Figure 1-1:** The phenylalanine hydroxylating system, showing hydroxylation of L-Phe into L-Tyr by PAH, and regeneration cycle for the cofactor BH<sub>4</sub>.

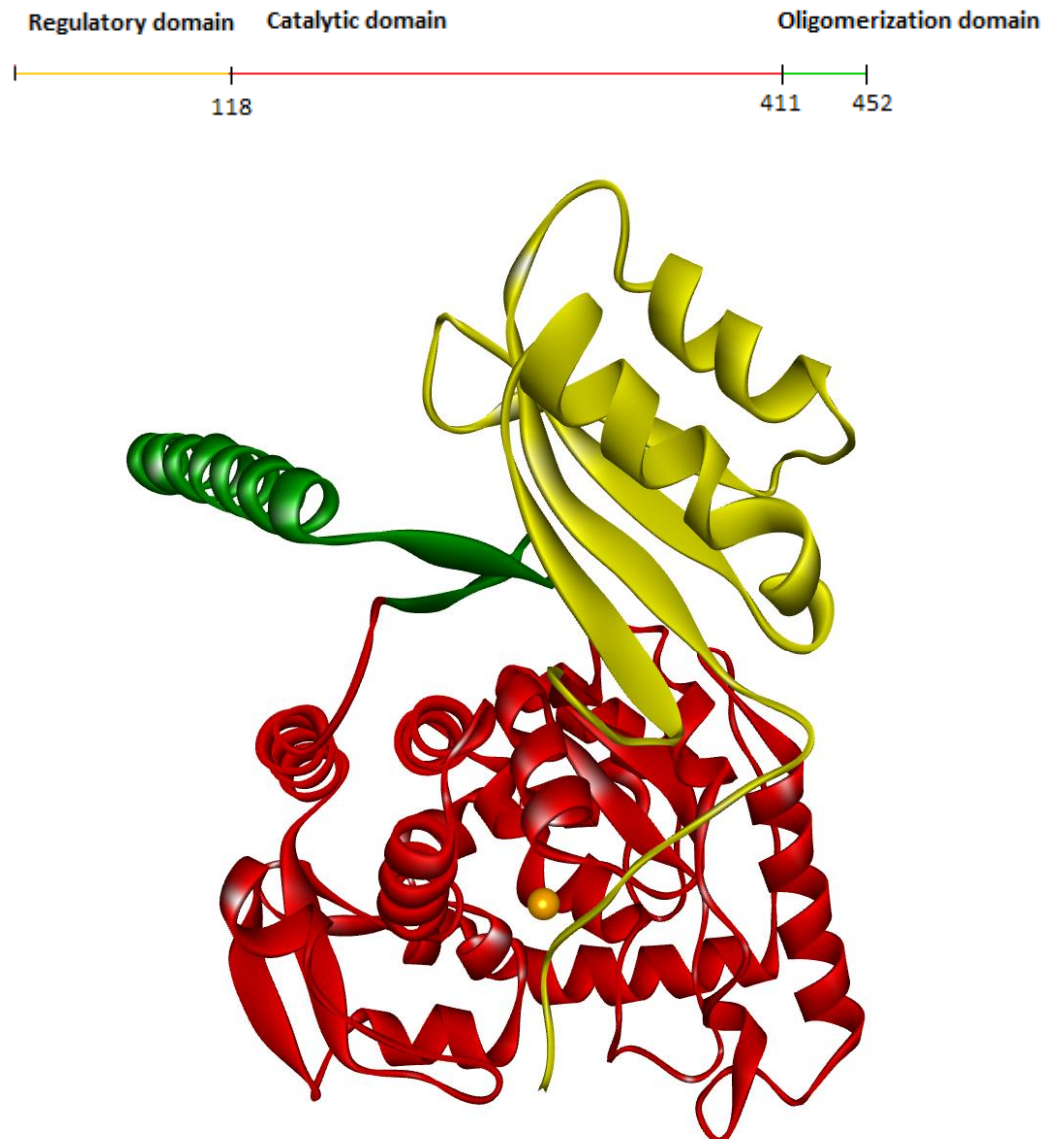
### 1.1.2 Structure and regulation

PAH is a tetramer consisting of four monomers arranged as two dimers bound together (Figure 1-2). Each monomer has three domains: an N-terminal regulatory domain, a catalytic domain, and a C-terminal oligomerization domain (Figure 1-3).



**Figure 1-2:** The PAH tetramer. Each monomer has a distinct colour and all domains are included (PDB entries 2pah and 1phz). The iron atom is shown as an orange sphere in each of the four active sites.

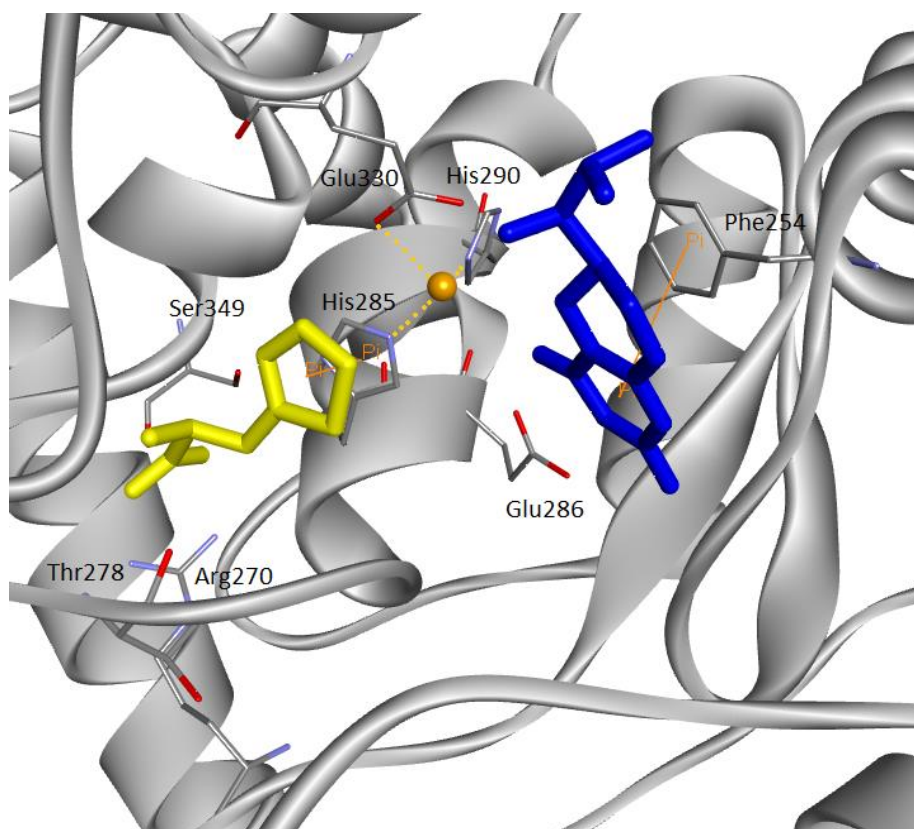
The crystal structures have been solved for dimeric human PAH (catalytic domain) (6), tetrameric human PAH (oligomerization domain plus catalytic domain) (7), and for dimeric rat PAH (regulatory domain plus catalytic domain) (8).



**Figure 1-3:** The PAH monomer showing the domain organization. The regulatory N-terminal is shown in yellow (residues 1-117, only 19-117 shown), the catalytic domain in red (residues 118-410) with the iron atom as an orange sphere, and the oligomerization C-terminal in green (residues 411-452) (PDB entries 2pah and 1phz).

The regulatory domain is located at the N-terminal. PAH is activated by phosphorylation of Ser16 by a protein kinase (9, 10), and by binding of phenylalanine (11). These regulatory mechanisms act synergistically, enhancing each other (10). In addition to its role as cofactor in the catalytic mechanism,  $\text{BH}_4$  is also a negative regulator of PAH activity at high concentrations (12). The oligomerization domain is located at the C-terminal and it is responsible for dimerization and tetramerization.

The catalytic domain consists of an active site with an iron atom coordinated to three amino acids, called a 2-His-1-carboxylate facial triad-motif (13), and two binding pockets; one cofactor binding site and one substrate binding site. Figure 1-4 shows the amino acids important for coordination of the iron atom (His285, His290 and Glu330), binding of the cofactor (especially pi-pi stacking through Phe254 and hydrogen bonding to Glu286), and binding of substrate, in particular His285 (pi-pi stacking), Arg270, Ser349 and Thr278 (14, 15).

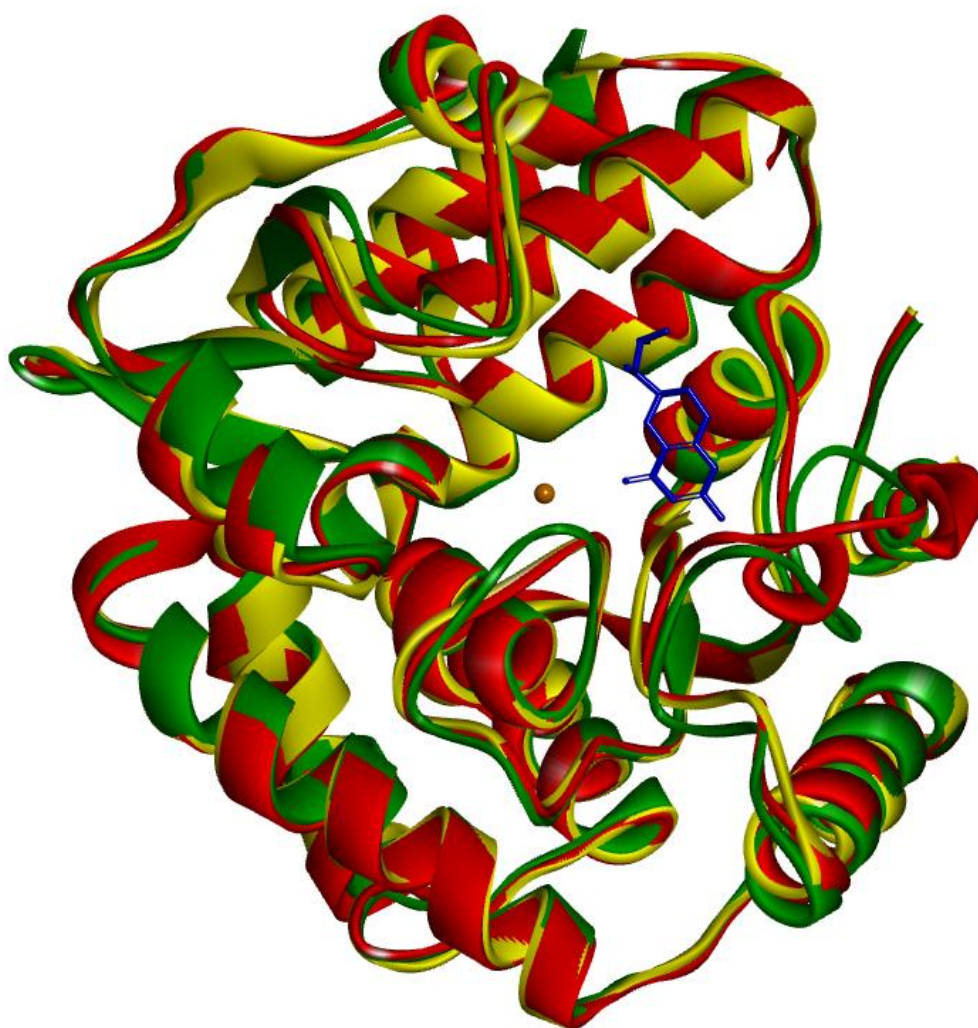


**Figure 1-4:** The active site with the cofactor and substrate binding sites, showing the most important amino acids involved in cofactor and substrate binding and coordination of the iron atom (PDB 1mmk). The amino acids are shown as sticks and colored by atom type. BH<sub>4</sub> is shown in blue in the cofactor binding site, and the substrate analog thienylalanine (L-Tha) is shown in yellow in the substrate binding site. Iron is shown as an orange sphere coordinating to three amino acids.

When the substrate binds to PAH, it is believed that large conformational changes occur. Only binding of the substrate analogous thienylalanine (L-Tha) and norleucine (L-Nle) have been solved by x-ray crystallography, but it is reasonable to assume that binding of phenylalanine is similar. The crystal structures of the substrate analogous in complex with the catalytic domain of PAH, were solved by Andersen et al in 2003 (16), and demonstrate that conformational changes occur in the catalytic domain upon substrate binding; most importantly, Tyr138 moves from a surface position to a position inside of the active site. The structural changes in the full length enzyme with intact regulatory domain have not yet been solved experimentally. However, there is experimental evidence of an increased exposure of the regulatory domain upon substrate binding (17), and it has been postulated that the regulatory domains rearrange to form an allosteric L-Phe binding site (18). In an alternative model, L-Phe is proposed to bind exclusively at the four catalytic sites in the tetramer, without any allosteric L-Phe binding site in the regulatory domain (19, 20). Direct experimental evidence for the proposed model of L-Phe substrate activation of the full-length PAH enzyme, represents a challenge for the future and require further structural information on the full-length enzyme.

## 1.2 The aromatic amino acid hydroxylases

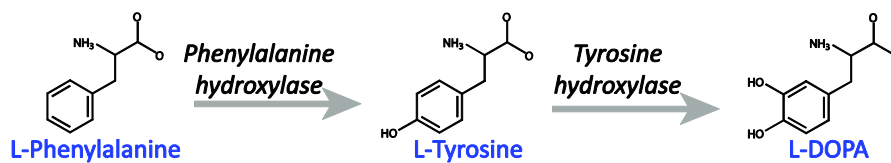
PAH is part of an enzyme family called the aromatic amino acid hydroxylases (AAAHs), which also include tyrosine hydroxylase (TH) and the tryptophan hydroxylases (TPH1 and TPH2). They all catalyse the hydroxylation of an aromatic amino acid substrate requiring  $\text{BH}_4$ ,  $\text{O}_2$  and  $\text{Fe}^{2+}$ , and are structurally very similar (Figure 1-5). Dysfunction caused by mutations in TH and the TPHs have been associated with neurological and psychiatric disorders (21).



**Figure 1-5:** Superimposition of the catalytic domains of PAH (red), TH (yellow) and TPH1 (green) (PDB entries 1j8u, 1toh and 1mlw). The cofactor in complex with PAH is shown as sticks (blue), and iron as an orange sphere.

### 1.2.1 Tyrosine hydroxylase

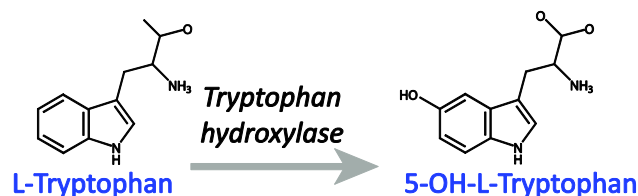
TH is found in the central nervous system (CNS), the peripheral sympathetic neurons and in the adrenal medulla (22), where it catalyses the hydroxylation of L-Tyr into L-DOPA, which is the rate-limiting step in the synthesis of the catecholamines dopamine, noradrenalin and adrenaline (Figure 1-6). Dysfunctions in TH are associated with the genetic neurological disorder tyrosine hydroxylase deficiency (THD), characterized by autosomal recessive DOPA-responsive dystonia and juvenile Parkinsonism, and TH levels are also severely reduced in Parkinson's disease (23-25).



**Figure 1-6:** PAH catalysis of hydroxylation of L-Phe into L-Tyr in the liver, and subsequent TH catalysis of hydroxylation of L-Tyr into L-DOPA in neuroendocrine tissues.

### 1.2.2 Tryptophan hydroxylase

TPH catalyses the hydroxylation of tryptophan (L-Trp) to 5-hydroxytryptophan (5-OH-Trp), which is the rate-limiting step in the synthesis of serotonin (Figure 1-7). Serotonin is further converted to melatonin. TPH exists as two isoforms, encoded by two genes: TPH1 is mainly found in the gastrointestinal tract (GIT) and in the pineal gland, where melatonin is produced, while TPH2 seems to be found exclusively in the brain and it is the main TPH isoform in the CNS (26). Serotonin has an important role in regulating sleep, mood, appetite and sexual behaviour, and mutations in the TPHs have been associated with various psychiatric disorders, such as depression, suicidal behaviour and hyperactivity (27-29).



**Figure 1-7:** TPH catalysis of hydroxylation of L-Trp into 5-OH-Trp in GIT and CNS.

### 1.3 Phenylketonuria

Phenylketonuria (PKU), first described by Fölling in 1934 (30), and probably better known as Föllings disease in Norway, is an inborn error of L-Phe metabolism caused by mutations in the gene encoding PAH. Defects in PAH lead to raised plasma levels and accumulation of phenylalanine (hyperphenylalaninemia; HPA), which impair brain function and development, if left untreated (1). In Norway, about 1:13000 newborns are diagnosed with PKU (31).

PKU is inherited as an autosomal recessive condition, i.e. both alleles have to be mutated. There are over 500 different mutations in the *PAH* gene, most of them causing a misfolding of PAH and subsequent decreased conformational stability (32, 33) (<http://www.pahdb.mcgill.ca>). It is distinguished between classical PKU, mild PKU and even non-PKU HPA, depending on the degree of phenylalanine accumulation and loss-of-PAH-function. Classical PKU has little or no PAH activity (plasma L-Phe > 1200  $\mu$ M), whereas mild PKU and non-PKU HPA has some PAH activity (plasma L-Phe 360-1200  $\mu$ M and 120-600  $\mu$ M, respectively) (34). A fourth type, generally known as  $\delta$ malignant PKU is not caused by PAH-mutations, but by mutations in the genes encoding for enzymes involved in BH<sub>4</sub> synthesis or regeneration, and in addition to HPA, this form is also characterized by severe deficiency in monoamine neurotransmitters (catecholamines and serotonin) (5).

Newborn screening tests and life-long dietary treatment has been successful in preventing brain damage resulting from PKU. However, a lower IQ than the average is observed for individuals with PKU, even with a phenylalanine-restricted diet from birth. Dietary treatment is also very demanding, often leading to malnutrition and psychosocial complications (35). Therefore, alternative treatments have appeared over the recent years, including treatment with synthetic BH<sub>4</sub> (Kuvan®), which has been shown to reduce plasma L-Phe levels and increase L-Phe tolerance in mild forms of PKU (36, 37). Also, gene therapy (38, 39), supplementation with large neutral amino acids (LNAA) (40) and enzyme replacement therapy with phenylalanine ammonia lyase (PAL) (41), are alternative approaches which are being investigated for treating PKU.



## 1.4 Protein folding and misfolding diseases

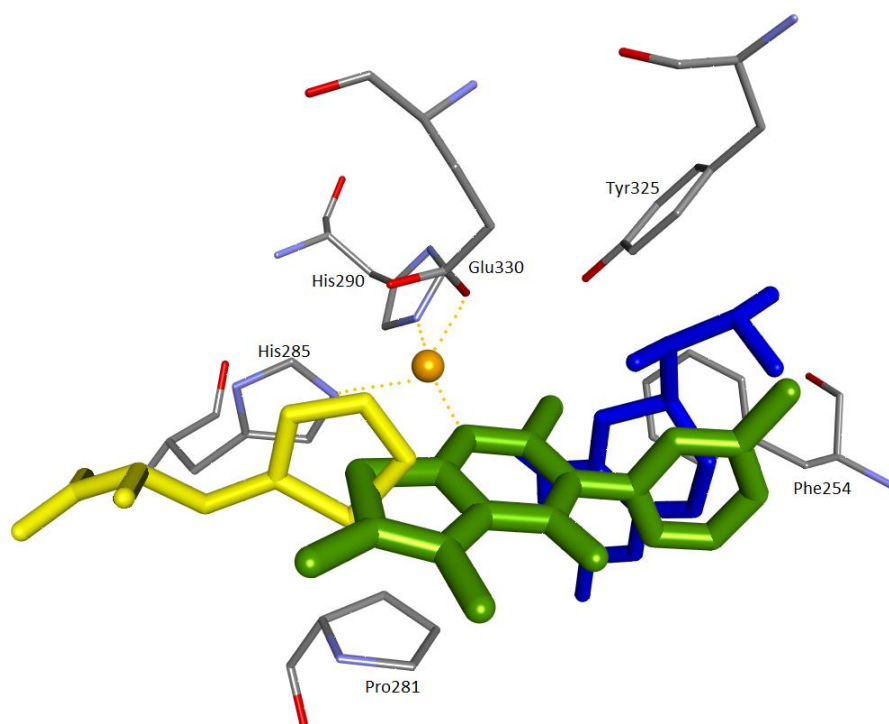
Folding of proteins occurs because of interactions between the amino acids in the primary structure, and correct folding is required for a functional protein. Proteins fold into the conformation that has the lowest free energy, and the free energy difference ( $\Delta G$ ) between the folded state (F) and the unfolded state (U) determines the conformational stability of the protein (42, 43).

Mutations in the amino acid sequence can cause misfolding or denaturation of the protein. Misfolded proteins are usually degraded, which leads to deficiency in the protein function, as seen for PKU, which is designated a loss-of-function misfolding disease. Gain-of-function misfolding diseases also exist, in which the mutant protein gains a toxic function where it is resistant to degradation and aggregates, forming amyloid deposits, as observed in Alzheimer's disease and Parkinson's disease (44).

## 1.5 Pharmacological chaperones

The concept of pharmacological chaperones comes from the term molecular chaperones. Molecular chaperones are large proteins that facilitate and assist protein folding, being part of the cellular quality control system that either assists misfolded proteins in refolding, or degrades them (45). Pharmacological chaperones are small molecules that stabilize the native folded conformation of a protein by interacting with residues in the active site, preventing or rescuing misfolding (46-48). These molecules therefore have a potential in the treatment of various misfolding diseases, such as PKU. Successful therapeutic use of pharmacological chaperones has been described for several of the lysosomal storage disorders (LSDs) and cystic fibrosis (see review (49)).

In fact, after being on the market for several years, synthetic BH<sub>4</sub> was discovered to act like a natural pharmacological chaperone for PAH. It has been shown to stimulate PAH activity partly by correcting mutant-PAH misfolding in some PKU patients (36, 37). Additional small molecule PAH chaperones have recently been identified by Pey et al. (50), and Santos-Sierra et al. (51). The crystal structure of one of these pharmacological chaperones, usually referred to as compound IV (5,6-dimethyl-3-(4-methyl-2-pyridinyl)-2-thioxo-2,3-dihydrothieno[2,3-d]pyrimidin-4(1H)-one), in complex with hPAH was solved by Torreblanca et al (52) in 2012, and demonstrated that compound IV bound in the active site, coordinating to iron through a nitrogen atom (N1), as shown in Figure 1-8. It is a weak inhibitor of PAH activity, but it still has the ability to stabilize PAH (both wild type and mutant forms) in vitro and in vivo (50).



**Figure 1-8:** Binding of compound IV in the active site of PAH (PDB 4anp and 1mmk). Compound IV is represented as green sticks, and coordination between N1 and iron (orange sphere) is shown. BH<sub>4</sub> is displayed in blue, and L-Tha in yellow. The protein is omitted for clarity, but important amino acids are shown as sticks.

Selectivity towards PAH represents a challenge, as the AAAs are structurally very similar, and chaperoning (or inhibitory) effect on TH and the TPHs could result in unacceptable side-effects. Some of the chaperones are tested on the other AAAs, but the selectivity is not fully established (53).

## 1.6 Computational drug design

Computer-aided drug design is increasingly being utilized both in academic research and in the pharmaceutical industry. Traditionally, experimental high throughput screening (eHTS) has been a common way of identifying new leads in the field of drug discovery. It requires a set of known ligands that are experimentally screened against a target. With virtual high throughput screening (vHTS), larger chemical diversity can be explored, and to a reduced cost compared to eHTS. In the following lead optimization process where hits/leads obtained in the drug discovery phase are being further developed into drugs, computational approaches provide the opportunity to analyze the interactions of the potential leads with their target in an earlier stage, and optimizing the lead to increase the probability of it becoming an actual drug (54-56).

Target-based virtual screening, or molecular docking, can be defined as the comprehensive process of searching a virtual library of ligands by positioning them in the binding site in a three dimensional structure of the target protein, and scoring the binding affinity. The main principles in docking is a search algorithm, responsible for positioning the ligands in the active site of the protein in various orientations, and a scoring function, that determines whether these orientations are the most energetically favourable (57). With the increasing amount of potential protein targets available, target-based virtual screening provides a powerful tool in the early stages of drug discovery, and it has been successfully used to identify therapeutically active compounds, such as several anti-viral agents (58).

## **2 AIMS OF THE PROJECT**

The overall aim of this Master project was to find compounds that could bind to and possibly stabilize PAH and thus act as pharmacological chaperones. Towards this aim we chose to screen a virtual database of ligands by molecular docking and to experimentally test the most promising ligands by measuring their effect on the thermal stability of PAH and on the preservation of enzymatic activity. The compounds that stabilize PAH can be considered potential pharmacological chaperones for PAH, which later might be derivatized and optimized for future treatment of PKU.

## 3 MATERIALS AND METHODS

### 3.1 Materials

#### SOFTWARE

Program	Version	Provider
Maestro	9.5	Schrödinger LLC
LigPrep	2.7	Schrödinger LLC
Protein Preparation Wizard	-	Schrödinger LLC
Glide	6.0	Schrödinger LLC
Epik	2.5	Schrödinger LLC
Discovery Studio	3.5	Accelrys Inc.
Accelrys Draw	4.1	Accelrys Inc.
GraphPad Prism	6	GraphPad Software
Canvas	1.7	Schrödinger LLC

#### INSTRUMENTS

Method	Instrument	Provider
Concentration measurements	Nanodrop ND-1000	Saveen Werner
Centrifugation	Centrifuge 5810R	Eppendorf
Weighing	Melter Toledo AB104-S	Bergman
pH measurements	691 pH Meter	Metrohm
DSF	LightCycler 480	Roche Applied Science
HPLC	1200 Infinity series	Agilent Technologies

#### CHEMICALS

Name	Provider
5000X SYPRO orange	Sigma Aldrich
Acetic acid (CH <sub>3</sub> COOH)	Sigma Aldrich
Ammonium iron (II) sulphate hexahydrate (NH <sub>4</sub> ) <sub>2</sub> Fe(SO <sub>4</sub> ) <sub>2</sub> · H <sub>2</sub> O	Sigma Aldrich
Bovine serum albumin (BSA)	Sigma Aldrich
Catalase	Sigma Aldrich
Compound IV (5,6-dimethyl-3-(4-methyl-2-pyridinyl)-2-thioxo-2,3-dihydrothieno[2,3-d]pyrimidin-4(1H)-one)	Maybridge LTD
Compound 1 (5-(3-hydroxybenzyl)hydantoin)	Sigma Aldrich
Compound 2 (N-(4-hydroxy-6-quinazoliny)acetamide)	Sigma Aldrich
Compound 3 (4,4-Diamino[1,1-biphenyl]-3,3-diol)	Sigma Aldrich
Compound 4 (3-(2,3-Dihydro-1,4-benzodioxin-6-yl)-1H-1,2,4-triazol-5-amine)	Sigma Aldrich
Compound 5 (Tyr-Phe)	Sigma Aldrich
Compound 6 ((2S,3R)-3-Amino-2-hydroxy-4-phenylbutyric acid)	Sigma Aldrich

hydrochloride)	
Compound 7 (Met-Leu-Phe acetate salt)	Sigma Aldrich
Compound 8 (( <i>R,S</i> )-4-Fmoc-3-carboxymethyl-piperazin-2-on)	Sigma Aldrich
Compound 9 (7-Hydroxycoumarinyl-4-acetic acid)	Sigma Aldrich
Compound 10 (L-Glutamic acid -benzyl ester)	Sigma Aldrich
Compound 11 (2- <i>N</i> -Fmoc-amino-3-(2- <i>N</i> -Boc-amino-pyrrolidiny) propionic acid)	Sigma Aldrich
Compound 12 (Phe-Arg -naphthylamide dihydrochloride)	Sigma Aldrich
Compound 13 (2-Acetyl-1,3-indanedione)	Sigma Aldrich
Compound 14 (Diminazene aceturate)	Sigma Aldrich
Compound 15 (fmoc-D-2-aminomethylphe(boc))	Sigma Aldrich
Compound 16 (1,3-Diiminoisindoline)	Sigma Aldrich
Compound 17 ( <i>N</i> -Benzoyl-Asn-Gly-Thr amide trifluoroacetate salt)	Sigma Aldrich
Compound 18 (( <i>S</i> )-(-)-2- <i>t</i> -butyl-2-piperazinecarboxamide)	Sigma Aldrich
Dimethyl sulfoxide (DMSO)	Sigma Aldrich
Distilled water	Milli-Q
Dithiothreitol (DDT)	Sigma Aldrich
Ethanol	Sigma Aldrich
HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid)	Sigma Aldrich
Hydrogen chloride (HCl)	VWR International
Propan-1-ol	Sigma Aldrich
Sodium hydroxide (NaOH)	Sigma Aldrich
Tetrahydrobiopterin (BH <sub>4</sub> )	Schircks Laboratories
L-Tyrosine	Sigma Aldrich
L-Phenylalanine	Sigma Aldrich

## BUFFERS AND SOLUTIONS

FPLC-buffer:

Concentration	Chemical	Mw
20 mM	HEPES	238,3 g/mol
200 mM	NaCl	58,44 g/mol

pH is adjusted to 7 with NaOH.  
Kept at -20 °C.

HEPES-buffer:

Concentration	Chemical	Mw
250 mM	HEPES	238,3 g/mol

pH is adjusted to 7 with NaOH  
Kept at -20 °C.

Stop solution:

Concentration	Chemical	Initial concentration
98 % (v/v)	Ethanol	Absolute
2 % (v/v)	Acetic acid	< 99,8 %

Kept at -20 °C.

HPLC-buffer:

Concentration	Chemical	Initial concentration
2 % (v/v)	1-Propanol	Absolute
0,1 % (v/v)	Acetic acid	< 99,8 %

## ENZYME

Recombinant human wild-type PAH was expressed and purified to homogeneity according to the protocol described by Martinez et al. (59) , and was provided by Ali Javier Sepulveda, technician in the Biorecognition group.



## 3.2 Target-based virtual screening

The Sigma library, which contains 66374 compounds, and is provided by Sigma-Aldrich and downloaded from Zinc (60, 61), was screened to identify compounds that fit with specific binding pockets in the correctly folded PAH structure, and therefore could be potential pharmacological chaperones for the enzyme. In order to test and validate the method, we performed a cross-docking of ligands that we know the binding mode of, including BH<sub>4</sub>, compound IV and two substrate analogues for PAH (L-Tha and L-Nle)

The docking program used in this project is Glide, which is part of a software package provided by Schrödinger (62, 63).

### 3.2.1 The Glide docking procedure

Before the docking job can be started, the protein and ligand structures need to be prepared, and the receptor grid, which is the area on the protein that is searched when attempting to dock the ligands, must be defined. Protein preparation involves assigning atom types, bond orders, partial charges and protonation states for the PDB-structure serving the role as the target protein. Ligand preparation ensures that the ligands are appropriate for docking, by assigning atom types and adding hydrogens and protonation states. Protein Preparation Wizard (64, 65) and LigPrep (65, 66) are provided from Schrödinger for these purposes.

Glide can perform flexible docking or rigid docking. In flexible docking, different ligand conformations are generated internally during the docking process, whereas for rigid docking only one ligand conformation is allowed, which is translated and rotated relative to the protein receptor. The protein receptor is treated by Glide as a rigid structure both in rigid and flexible docking. Glide has three different modes of precision; high throughput virtual screening (HTVS), standard precision (SP) and extra precision (XP). HTVS provides a rapid screen of large databases, but the conformational sampling is restricted. SP docking is the default, and it is suitable for screening large number of ligands of unknown quality. XP can be used on ligand poses that get a high score from the SP docking, and is more extensive (67).

Glide uses multiple search algorithms in a hierarchical manner where the ligands pass through a series of filters that evaluate the ligand's interactions with the protein receptor. The binding affinity is predicted by the scoring function GlideScore:

$$\text{GScore} = 0,05 \cdot \text{vdW} + 0,15 \cdot \text{Coul} + \text{Lipo} + \text{Hbond} + \text{Metal} + \text{Rewards} + \text{RotB} + \text{Site}$$

Van der Waals (vdW) and Coulomb energy (Coul) are calculated with reduced net ionic charges on groups with formal charges. The lipophilic term (Lipo) rewards favourable hydrophobic interactions. The hydrogen-bonding term (HBond) is separated into different weighted components depending on the charge on the hydrogen donors and acceptors. The metal binding includes only the interactions with anionic or highly polar atoms (if the net metal charge is positive). The rewards-term rewards or penalizes various features that are not mentioned separately. RotB adds a penalty for freezing rotatable bonds. The site-term rewards polar but non-hydrogen-bonding atoms in a hydrophobic region.

The protein-and ligand preparation procedures, and the search and scoring functions are thoroughly explained in the Glide User Manual (67).

### 3.2.2 Choice of PDB structures

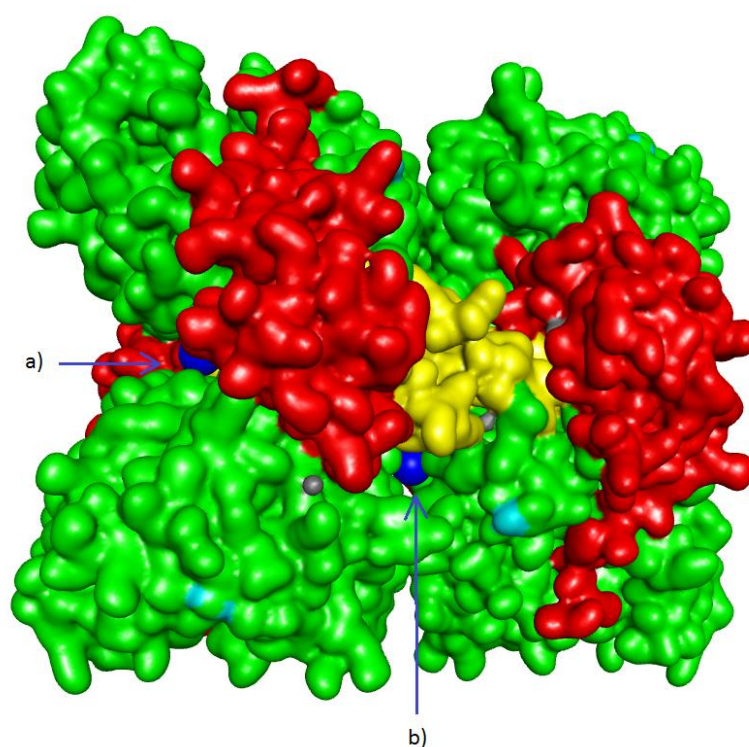
There are several crystal structures of PAH available in the Protein Data Bank (PDB), and which ones to use as target structures, was decided based on completeness of the structure, charge on the iron atom and the presence of cofactor or substrate analogues in the active site.

For the cross-docking we ended up using PDB 1j8u (representing the catalytic domain) as the protein receptor (68).

When docking the Sigma library, we have also considered alternative binding sites on the surface of the protein, in addition to the active site. Two potential binding sites were identified using LIGSITE (69); one in the interface between two monomers in a dimer, and the other one in the interface between two dimers in the tetramer (Figure 3-1). 1j8u only represents the catalytic domain. There is one PDB file (2pah) of tetrameric PAH (7), but this structure lacks the regulatory domain. Therefore, a hybrid structure was made out of 1j8u, 2pah and 1phz (8), the last one is a rat structure of PAH with an intact regulatory domain. To prepare a hybrid structure of tetrameric hPAH, the amino acids of the regulatory domain (from rat) were mutated to the corresponding human residues.

**Table 3-1:** Details of the PDB files used. Hetero-groups are the non-amino acid content of the structure.

<b>PDB entry</b>	<b>Organism</b>	<b>Resolution (Å)</b>	<b>Residues</b>	<b>Domain</b>	<b>Hetero-groups</b>
1j8u	<i>Homo sapiens</i>	1.5	118-424	Catalytic domain	Fe <sup>2+</sup> , BH <sub>4</sub>
2pah	<i>Homo sapiens</i>	3.1	118-452	Catalytic domain and oligomerization domain	Fe <sup>3+</sup>
1phz	<i>Rattus norvegicus</i>	2.2	19-427	Catalytic domain and regulatory domain	Fe <sup>3+</sup>



**Figure 3-1:** Binding pockets on the protein surface, illustrated as blue spheres (PDB 1j8u, 2pah and 1phz). The domains are shown in different colours. a) interface between monomers b) interface between dimers

### 3.2.3 Cross-docking of selected compounds

Cross-docking refers to docking of a ligand from one crystal structure of a protein into a different crystal structure of the same protein, to see whether the ligand is placed correctly. If the ligand is docked into the same crystal structure as it is taken from, the process is called re-docking. To address potential problems, validation by cross-docking or re-docking should always be performed, so that necessary adjustments can be made before starting the docking job. If having several crystal structures of the protein target, cross-docking is the validation method of choice because it is more challenging for the docking program than to just reproduce the binding mode of the ligand, and it is more similar to the real docking situation (70).

#### **Protein preparation and grid generation**

Two protein structures were prepared; one with all water molecules deleted and one keeping the water molecules located 3Å from *hetero-groups* (non protein chemical groups, i.e. active site iron and substrate/cofactor), in order to see if this made any difference in the docking of the cofactor. From the prepared protein structures, two grids were made. We chose a ligand diameter midpoint box of  $x=y=z=22$  Å, in order to test and get familiar with the program, allowing unusual or asymmetric binding modes using a large grid. All other settings were default.

#### **Ligand preparation**

Metal binding states were added, which generates high pH charges (enabling binding to iron), and determine chirality from 3D structure was chosen. All other settings were kept as default. When docking rigidly, we do not want the program to generate low energy ligand conformations, and this step was omitted in the standard LigPrep procedure. All tautomerization and ionization states at pH 5-9 were generated and docked.

#### **Docking**

BH<sub>4</sub>, compound IV, L-Tha and L-Nle, were docked using both rigid and flexible docking, and in all three modes of precision (HTVS, SP and XP), and the results were compared. They were all docked into the grid with no water molecules, and BH<sub>4</sub> was also docked flexibly into the grid containing some water molecules. Number of poses to keep per ligand was set to 5, all other settings were default.

To determine how well the ligands docked, we compared the results with reference compounds from PDB entries 1j8u (BH<sub>4</sub>), 4anp (compound IV) (52), 1mmk (Tha) (16) and 1mmt (Nle) (16), and calculated the root mean square deviation (RMSD), using Discovery Studio (71). The references were superimposed onto 1j8u (target protein).

### 3.2.4 Docking of the Sigma library

In addition to neutral pH, we included high pH protonation states for the Sigma ligands, to facilitate metal binding.

Based on the results from the cross-docking, we chose to dock the library in SP-mode. We also chose to include water molecules in some of the grids.

#### Protein preparation and grid generation

Five prepared protein structures were generated. Three of them are of the catalytic domain and are based on the structure of hPAH in complex with BH<sub>4</sub> (1j8u), after removal of the cofactor; the first one has all water molecules within 5 Å from *hetero*-groups included, the second one has no water molecules included, and the third one has three water molecules that coordinate to iron, included. The two last structures are of the alternative binding sites that were detected in LIGSITE (see above); both were generated with no water molecules included. All settings were kept default. From these prepared structures, five corresponding grids were made (Table 3-2): grid1-1, grid1-2 and grid 1-3 are the grids for the active site. Grid2 is the grid for the interface between the dimers, and grid3 is the grid for the interface between the monomers.

**Table 3-2:** Grids used when docking the Sigma-library.

Name	Description	Water molecules	PDB entry
Grid1-1	Active site	All	1j8u
Grid1-2	Active site	None	1j8u
Grid1-3	Active site	Three	1j8u
Grid2	Interface between dimers	None	2pah
Grid3	Interface between monomers	None	1j8u/2pah/1phz (hybrid)

#### Ligand preparation

The Sigma-ligands are downloaded as 3D structures, and partial charges were automatically assigned when the library was loaded to Glide and converted to the correct format for docking. Further preparation of the ligands was therefore not necessary.

## **Docking**

Docking was performed using flexible docking in SP-mode. Number of poses to keep per ligand was set to 1, and the total number of poses was limited to 500. All other settings were default.

Based on their Glide score and structures, 18 compounds were selected to be experimentally tested. All 2500 structures were analyzed with Canvas (72-74) using the substructure query-tool.

### 3.3 Experimental methods

#### 3.3.1 Thermostability assay

Differential scanning fluorimetry (DSF) was used to measure the thermal stability of PAH in the presence of the hit compounds from the virtual screening. The principle of DSF is to measure unfolding of the protein as the temperature is increased. A fluorescent dye (SYPRO orange) that interacts with the hydrophobic parts of the protein is added, and fluorescence is measured as the protein unfolds upon heating in a PCR machine (75, 76). When the protein unfolds, its hydrophobic parts are exposed to the dye, giving a fluorescent signal. The fluorescence ranges from 0 to 1, where 1 is the value at which all of the protein is unfolded. When the fluorescence is 0.5, 50 % of the protein is unfolded, termed the melting point of the protein ( $T_m$ ). Here, the shift in  $T_m$  ( $\Delta T_m$ ) when the hit compounds were added to the protein, were measured. A positive value indicates that the compound has a stabilizing effect on the protein, and a negative value indicates a destabilizing effect.

The experiment was performed with a 96-well microplate. The enzyme was diluted to 0.05 mg/mL in FPLC-buffer, and the fluorescent dye, SYPRO orange, was added. The compounds were dissolved in 100% DMSO in a concentration of 2 mg/mL, and added to the wells at a final concentration of 0.08 mg/mL (2% DMSO). Three controls with 2% DMSO were performed. The plate was heated from 30°C to 95°C with a temperature increase of 0.04 °C per second (4 measurements per degree). The fluorescent signals were detected as the protein was unfolded during temperature increase, giving values for  $T_m$  and  $\Delta T_m$ .

### 3.3.2 Enzymatic activity assay

In order to validate if the hit compounds were inhibitors or had a protective effect on PAH over time, we assayed their effect on PAH activity. Activity experiments were done using a standard PAH activity assay, as described by Knappskog et al (77), but with modifications of some of the conditions (Table 3-3 and Table 3-4). The amount of product formed was analysed using high performance liquid chromatography (HPLC) equipped with a fluorescence detector, after removing the protein by centrifugation.

#### **Standard PAH activity assay**

The enzyme was diluted to 0.05 mg/mL in FPLC-buffer (20 mM) and BSA (5 mg/mL) and kept on ice. BH<sub>4</sub> was added to a solution of acidic 2 mM DTT and kept on ice. The activity assay was performed at 37 °C. 5 µL of the enzyme was added (final concentration 0.005 mg/mL) to a mixture of 40 mM Hepes-buffer (pH 7), 5 mg/mL catalase and 1 mM phenylalanine. The incubation mixture was pre-incubated for 5 minutes. After 4 minutes 10 µM ferrous ammonium sulphate was added. At 5 minutes the reaction was started by adding 5 µL of BH<sub>4</sub> to a final concentration of 200 µM. The reaction was stopped after 1 minute by adding 50 µL stop solution containing 98% ethanol and 2% acidic acid. The reaction mix was precipitated at -20 °C for at least 30 minutes, and then centrifuged for 10 minutes on 14000 rpm.

The amount of product formed was analysed by HPLC. The substrate (L-Phe) and the product (L-Tyr) were separated in a column based on their affinity to the stationary phase compared to the mobile phase. The stationary phase was a cation exchanger, and the mobile phase was a 0.1 % acetic acid solution containing 2 % 1-propanol. L-Tyr is more polar than L-Phe because of the OH-group that is added during catalysis, and it therefore has a higher affinity to the mobile phase and will exit the column first. To determine the amount of produced L-Tyr, a fluorescence detector was used, measuring excitation at 274 nm and emission at 304 nm (78). The samples were compared to a standard with known concentration, and the specific activity was calculated [specific activity<sub>PAH</sub> = nmol<sub>Tyr</sub>/(mg<sub>PAH</sub> x min)].



**Table 3-3:** PAH activity assay. The upper part of the table shows the different components of the assay mixture and their concentration. The lower part represents the assay itself, starting at t=0 with the addition of PAH to the assay mixture. Pre-incubation time was 5 minutes and reaction time was 1 minute.

		<b>Original concentration</b>	<b>Volume (per sample)</b>	<b>Assay concentration</b>
MIX	Hepes-buffer	250 mM	8. $\mu$ L	40 mM
	Catalase	10 mg/mL	0.25 $\mu$ L	0.05 mg/mL
	L-Phe	50 mM	1 $\mu$ L	1 mM
	H <sub>2</sub> O		29.75 $\mu$ L	
	<b>SUM</b>		<b>39 <math>\mu</math>L</b>	
t=0	PAH	0.05 mg/mL	5 $\mu$ L	0.005 mg/mL (0.25 $\mu$ g)
t=4	Fe <sup>2+</sup>	0.5 mM	1 $\mu$ L	10 $\mu$ M
t=5	BH <sub>4</sub>	2 mM	5 $\mu$ L	0.2 mM
	<b>SUM</b>		<b>50 <math>\mu</math>L</b>	
t=6	Stop solution		50 $\mu$ L	
	<b>SUM</b>		<b>100 <math>\mu</math>L</b>	

### Optimization of activity assay

The assay was optimized by testing the effects of iron and BSA (bovine serum albumin) on the specific activity of PAH, by measuring activity in their presence and absence. The reason for testing this was to destabilize the enzyme, but still maintain good activity, which was preferable when we later wanted to obtain a destabilization curve for the time dependent activity loss of PAH. The purified PAH preparations have amounts of iron varying from 0.45 to 0.52 iron atoms per subunit PAH (77) (our preparation is assumed to be in the upper range), and iron was also being added in the assay to saturate the iron sites and increase the activity. BSA is a serum albumine protein that stabilizes the enzyme by acting as a crowding agent (for review on molecular crowding and effect on protein stability, see (79)).

Based on the results, we came up with an assay where additional iron is not added (as in the standard protocol for PAH activity), and where BSA is added in the reaction mixture instead of being added to the enzyme dilution, as shown in Table 3-4.

**Table 3-4:** Optimized PAH activity assay, where iron is removed and BSA added in the mixture. Pre-incubation time was 5 minutes and reaction time was 1 minute.

		<b>Original concentration</b>	<b>Volume (per sample)</b>	<b>Assay concentration</b>
MIX	Hepes-buffer	250 mM	8 $\mu$ L	40 mM
	Catalase	10 mg/mL	0.25 $\mu$ L	0.05 mg/mL
	L-Phe	50 mM	1 $\mu$ L	1 mM
	BSA	100 mg/mL	2.5 $\mu$ L	5 mg/mL
	H <sub>2</sub> O		28.25 $\mu$ L	
	<b>SUM</b>			<b>40 <math>\mu</math>L</b>
t=0	PAH	0.05 mg/mL	5 $\mu$ L	0.005 mg/mL (0.25 $\mu$ g)
t=5	BH <sub>4</sub> /DTT	2 mM	5 $\mu$ L	0.2 mM
	<b>SUM</b>		<b>50 <math>\mu</math>L</b>	
t=6	Stop solution		50 $\mu$ L	
	<b>SUM</b>		<b>100 <math>\mu</math>L</b>	

### Linearity of the specific activity of PAH

By varying the reaction time from 1 to 10 minutes, we wanted to see for how long the activity remained linear (constant specific activity). The measurements have to be performed in the linear range of product formation, so that the effects of possible pharmacological chaperones are visible.

### **Activity loss as a function of time**

The enzyme was kept on 37 °C for 1 to 40 minutes (pre-pre-incubation time) before starting the assay, in order for us to observe an activity loss over time.

### **Testing of the hit compounds**

We tested the effect of individual hit compounds on PAH activity and stability. Compound IV was also tested, because we know that it has a stabilizing effect and it is therefore interesting to compare with. Two experiments were carried out – the first one with no pre-pre-incubation in order to see how the compounds interfere with PAH activity, and the second one with a pre-pre-incubation time of 10 minutes to see if any of the compounds can preserve the activity and thus stabilize PAH and act as a pharmacological chaperone.

To make sure that the fluorescent signals detected in HPLC are caused by the produced L-Tyr, all the compounds were tested to see if they had any natural fluorescence.

The compounds were dissolved at a concentration of 2 mg/mL in 100% DMSO, and tested at a final concentration of 0.04 mg/mL in 2% DMSO, using 96-well micro plates, with one control with 2% DMSO on each row. The assays were designed so that each row was assayed at the same time using a multi pipette, and the rows were assayed in order from the first one to the last one. To get the pre-incubation and reaction time right for each row, the last one was done approximately 30 minutes after the first one. The assays were performed at 37 °C.

A varying number of parallels were performed for the controls and for each compound, and standard deviation (SD) and standard error (SEM) were calculated, to reflect the variation between the parallels (SD) and the precision of the mean value (SEM). A p-value was calculated for each compound compared to the control, using the multiple t-tests-function in Prism (unpaired t-test). This function controls the false discovery rate (FDR) when having multiple tests, by performing a Benjamini-Hochberg procedure to determine which p-values are small enough to be considered true discoveries (80).

### Effect on PAH activity

The mix was added to the wells, and one compound was added to each well. The assay was then started.

**Table 3-5:** Testing of hit compounds ó initial activity assay. The compound was added in the mixture before the assay was started at t=0. Pre-incubation time was 5 minutes and reaction time was 1 minute.

		<b>Original concentration</b>	<b>Volume (per sample)</b>	<b>Assay concentration</b>
MIX	Hepes-buffer	250 mM	8 µL	40 mM
	Catalase	10 mg/mL	0.25 µL	0.05 mg/mL
	L-Phe	50 mM	1 µL	1 mM
	BSA	100 mg/mL	2.5 µL	5 mg/mL
	H <sub>2</sub> O		27.25 µL	
	Compound	2 mg/mL	1 µL	0.04 mg/mL
	<b>SUM</b>		<b>40 µL</b>	
t =0	PAH	0.05 mg/mL	5 µL	0.005 mg/mL (0.25 µg)
t =5	BH <sub>4</sub> /DTT	2 mM	5 µL	0.2 mM
	<b>SUM</b>		<b>50 µL</b>	
t =6	Stop solution		50 µL	
	<b>SUM</b>		<b>100 µL</b>	

### Effect on PAH stability

The mix was added to the wells. The enzyme (0.05 mg/mL) was pre-pre-incubated with each compound (0.4 mg/mL) for 10 minutes at 37 °C, before the assay was started by adding 5 µL of the enzyme plus compound to the mix.

**Table 3-6:** Testing of hit compounds ó activity stability assay. The compound was pre-pre-incubated with PAH for 10 minutes before the assay was started by adding PAH plus compound to the mix at t=0. Pre-incubation time was 5 minutes and reaction time was 1 minute.

		<b>Original concentration</b>	<b>Volume (per sample)</b>	<b>Assay concentration</b>
MIX	Hepes-buffer	250 mM	8 µL	40 mM
	Catalase	10 mg/mL	0.25 µL	0.05 mg/mL
	L-Phe	50 mM	1 µL	1 mM
	BSA	100 mg/mL	2.5 µL	5 mg/mL
	H <sub>2</sub> O		28.25 µL	
	<b>SUM</b>		<b>40 µL</b>	
	t=0	PAH + Compound	0.05 mg/mL 0.4 mg/mL	5 µL
t=5	BH <sub>4</sub> /DTT	2 mM	5 µL	0.2 mM
	<b>SUM</b>		<b>50 µL</b>	
t=6	Stop solution		50 µL	
	<b>SUM</b>		<b>100 µL</b>	

## 4 RESULTS

### 4.1 Target-based virtual screening

#### 4.1.1 Cross-docking of selected compounds

The results from the cross-docking are presented in Table 4-1, Table 4-2 and Table 4-3, as RMSD values and Glide gscores. The RMSD between the docked ligand and the crystal structure (PDB 1j8u) should be as low as possible, indicating that the orientations match. We have chosen to define docking of the ligand in the active site with RMSD lower than 2 Å with respect to the crystal structure, as successful.

**Table 4-1:** Rigid cross-docking of BH<sub>4</sub>, compound IV, L-Tha and L-Nle to PAH, using HTVS, SP and XP. RMSD (compared to PDB 1j8u) and Glide gscore for each compound are displayed.

Compound	HTVS		SP		XP	
	RMSD (Å)	Glide gscore (kcal/mol)	RMSD (Å)	Glide gscore (kcal/mol)	RMSD (Å)	Glide gscore (kcal/mol)
BH <sub>4</sub>	1.324	-6.435	1.315	-6.504	1.319	-5.640
Comp.IV	0.8251	-5.659	0.8245	-5.750	0.8451	-3.859
L-Tha	8.862	-7.887	9.014	-8.047	9.003	-5.737
L-Nle	8.998	-7.150	9.178	-7.735	9.432	-5.604

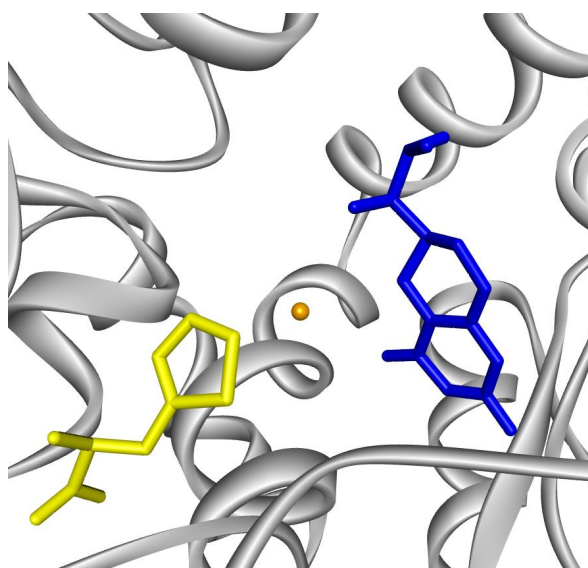
**Table 4-2:** Flexible cross-docking of BH<sub>4</sub>, compound IV, L-Tha and L-Nle to PAH, using HTVS, SP and XP. RMSD (compared to PDB 1j8u) and Glide gscore for each compound are displayed.

Compound	HTVS		SP		XP	
	RMSD (Å)	Glide gscore (kcal/mol)	RMSD (Å)	Glide gscore (kcal/mol)	RMSD (Å)	Glide gscore (kcal/mol)
BH <sub>4</sub>	1.750	-5.501	1.492	-5.886	1.816	-6.267
Comp.IV	2.502	-5.148	1.658	-6.562	5.503	-5.405
L-Tha	8.615	-8.578	8.436	-8.754	8.872	-10.55
L-Nle	8.889	-7.486	9.095	-7.753	9.957	-10.26

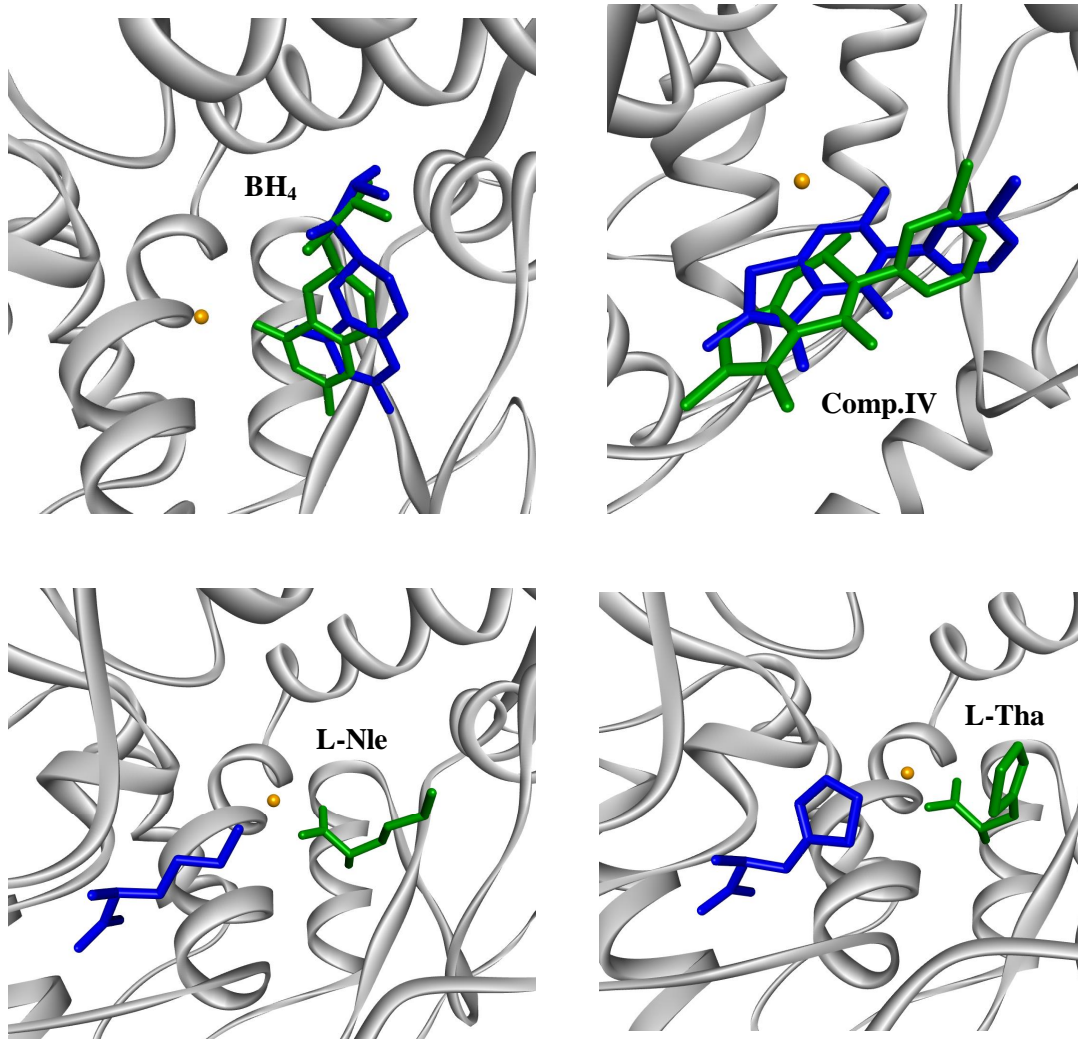
When comparing rigid and flexible docking, we see that rigid docking gives the best (lowest) RMSD values. This is expected, as rigid docking only docks one conformation (the correct one), whereas Glide generates ligand conformations internally during the flexible docking process, as mentioned earlier. In general, looking at both rigid and flexible docking, we see that XP performs poorer than HTVS and SP. According to Schrödinger (personal communication with Schrödinger Support), XP seems to be the least accurate when metals are involved.

Both for BH<sub>4</sub> and compound IV, the RMSDs are under 2 Å (1.49 Å and 1.66 Å in SP flexible mode). This is particularly good when considering compound IV's covalent bond to iron (metal coordination is a challenge to most docking programs, including Glide). The substrate analogues actually got much larger deviations (8.44 Å and 9.10 Å in SP flexible mode).

The binding modes of the cross-docked compounds are shown in Figure 4-2, and the binding mode of BH<sub>4</sub> and L-Tha as they appear in their crystal structures are also shown for comparison in Figure 4-1.



**Figure 4-1:** BH<sub>4</sub> (blue) in the cofactor binding pocket and Tha (yellow) in the substrate binding pocket (PDB 1mmk). Iron is shown as a sphere with three coordinating amino acids.



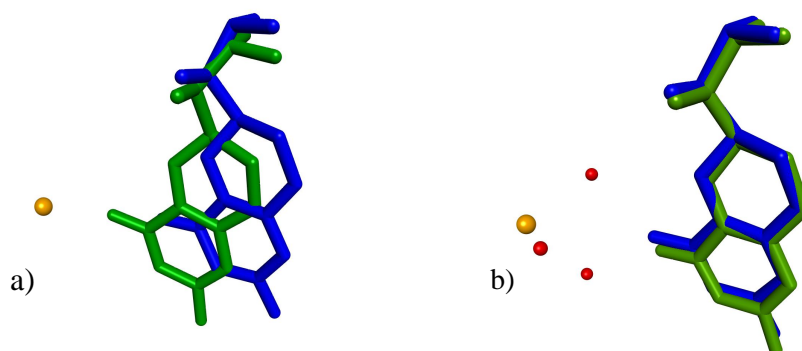
**Figure 4-2:** BH<sub>4</sub>, compound IV, Nle and Tha (all shown in green) docked to the cofactor binding pocket. Reference structures in blue. Iron is shown as a sphere.

We see that the four compounds were all docked into the cofactor binding pocket, including the substrate analogues (that we know bind in the substrate binding site, which explains their poor RMSD-values).

BH<sub>4</sub> was successfully re-docked to the cofactor binding site, with an RMSD-value of 1.49 Å (SP flexible mode) between the re-docked structure and the crystal structure with all water molecules removed. A lower RMSD value was obtained when including water molecules 3 Å from BH<sub>4</sub> in the grid (0.30 Å), as shown in Table 4-3 and Figure 4-3.

**Table 4-3:** SP flexible docking of BH<sub>4</sub> using two different grids; one with no waters present and one with water molecules within 3 Å from hetero-groups included.

BH <sub>4</sub> grid details	RMSD (Å)	Glide score (kcal/mol)
No waters present	1.492	-5.886
Waters present	0.3028	-11.11



**Figure 4-3:** SP-flexible re-docking of BH<sub>4</sub>. The re-docked BH<sub>4</sub> is shown in green, the original BH<sub>4</sub> in blue. a) waters excluded from grid b) waters within 3Å from hetero-groups included in grid. Three water molecules coordinating to iron are shown in red.

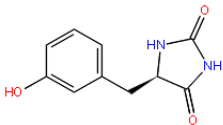
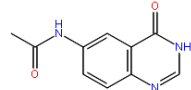
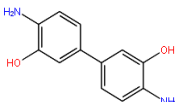
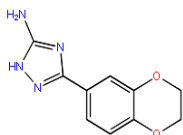
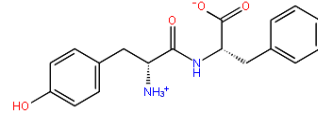
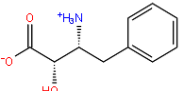
We see that the re-docked BH<sub>4</sub> and BH<sub>4</sub> as it appear in the crystal structure, have almost identical binding modes when some water molecules are included in the grid, because of the structural water molecules not being replaced by the cofactor upon binding.

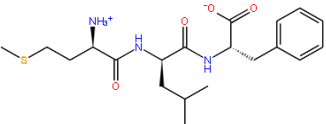
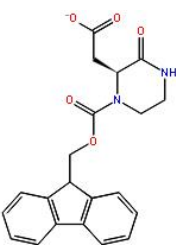
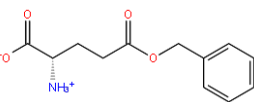
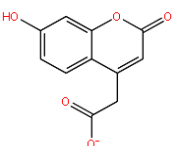
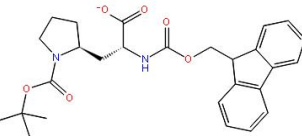
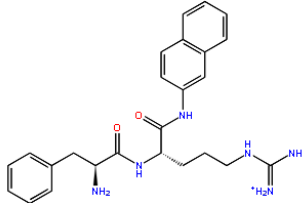
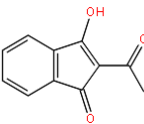
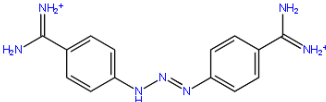


## 4.1.2 Docking of the Sigma library

Table 4-4 shows the selected top hits from the flexible docking of the Sigma library (the complete list of all structures is provided upon request). These 18 compounds were bought in order to test their binding experimentally. Grid1-1, grid1-2 and grid1-3 are the grids for the active site; the first one has all water molecules within 5 Å from hetero-groups included, the second one has no water molecules included and the third one has three water molecules included. Grid2 is the grid for the interface between the dimers, and grid3 is the grid for the interface between the monomers.

**Table 4-4:** Selected hits from flexible SP-docking of the Sigma library, with assigned compound numbers, zinc ID number, name and 2D structure, Glide gscore and pose number (i.e. ranking).

Compound number	Zinc ID	Structure/name	Pose number	Glide gscore
Grid1-1				
1	ZINC02540779	 5-(3-hydroxybenzyl)hydantoin	6	-10.71
2	ZINC17111121	 N-(4-hydroxy-6-quinazoliny)acetamide	25	-9.328
3	ZINC02038670	 4,4-Diamino[1,1-biphenyl]-3,3-diol	27	-9.313
4	ZINC26514552	 3-(2,3-Dihydro-1,4-benzodioxin-6-yl)-1H-1,2,4-triazol-5-amine	28	-9.225
Grid1-2				
5	ZINC05273977	 Tyrosylphenylalanine	5	-11.491
6	ZINC06096602	 (2S,3R)-3-Amino-2-hydroxy-4-phenylbutyric acid hydrochloride	8	-11.309

7	ZINC03874236	 Met-Leu-Phe acetate salt	14	-11.063
8	ZINC14632708	 ( <i>R,S</i> )-4-Fmoc-3-carboxymethyl-piperazin-2-on	19	-10.943
Grid1-3				
9	ZINC01700281	 L-Glutamic acid -benzyl ester	2	-10.753
10	ZINC00153920	 7-Hydroxycoumarinyl-4-acetic acid	5	-10.591
11	ZINC04284379	 2- <i>N</i> -Fmoc-amino-3-(2- <i>N</i> -Boc-amino-pyrrolidinyl) propionic acid	19	-10.158
Grid2				
12	ZINC27738213	 Phe-Arg -naphthylamide dihydrochloride	18	-9.475
13	ZINC05129922	 2-Acetyl-1,3-indanedione	12	-8.377
14	ZINC03830706	 Diminazene aceturate	16	-8.251

15	ZINC04241974	<p>Chemical structure of fmoc-D-2-aminomethylphe(boc) showing a fluorenylmethyl carbamate group attached to a phenyl ring, which is further substituted with a tert-butyl carbamate group and a 2-aminomethyl group.</p>	31	-7.913
<hr/>				
Grid3				
16	ZINC00388558	<p>Chemical structure of 1,3-Diiminoisoindoline, a five-membered heterocyclic ring with two imino groups (=NH<sub>2</sub><sup>+</sup>) and one imino group (=NH).</p>	3	-8.308
		1,3-Diiminoisoindoline		
17	ZINC15721778	<p>Chemical structure of N-Benzoyl-Asn-Gly-Thr amide trifluoroacetate salt, a peptide chain consisting of Asparagine, Glycine, and Threonine residues, with a benzoyl group attached to the Asparagine residue.</p>	8	-8.092
		<i>N</i> -Benzoyl-Asn-Gly-Thr amide trifluoroacetate salt		
18	ZINC26893390	<p>Chemical structure of (S)-(-)-2-t-butyl-2-piperazinecarboxamide, a piperazine ring substituted with a tert-butyl group and a carboxamide group.</p>	13	-7.7
		(S)-(-)-2-t-butyl-2-piperazinecarboxamide		

Many of the compounds obtained a good score (i.e. large negative Glide gscore number), and interact favourably with PAH. Most of the compounds that docked to the active site appear to bind in the cofactor binding pocket and are mainly cofactor- or substrate analogues or small peptides. Some of the larger compounds seem to occupy both the cofactor binding site and the substrate binding site. See Appendix 1 for binding modes.

## 4.2 Testing virtual hits experimentally

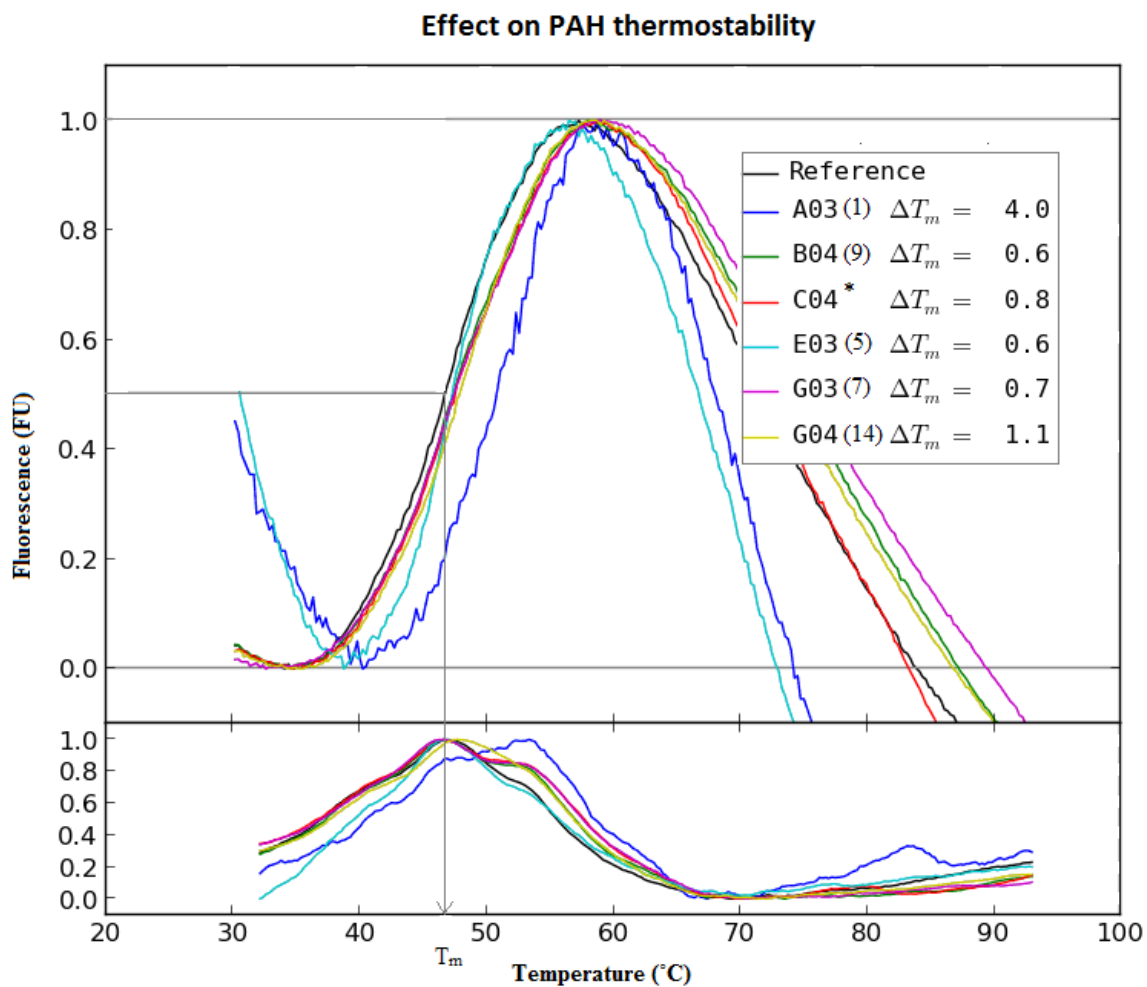
### 4.2.1 Effect of the compounds on the thermostability of PAH

The shift in  $T_m$  in the presence of each compound is given in Table 4-5.

**Table 4-5:**  $T_m$  for PAH in the presence of the hit compounds.  $T_m$  for the control was  $46.7 \pm 0.2$  °C (triplicate). Compounds giving a positive shift in  $T_m$  is highlighted.

Compound	$T_m$ (°C)	Compound	$T_m$ (°C)
1	4.0	10	0
2	0	11	0
3	-5.8	12	0
4	0	13	0
5	0.6	14	1.1
6	0	15	-4.1
7	0.7	16	0
8	0	17	0
9	0.6	18	0

We see that 10 compounds show no effect on thermal stability ( $T_m$  is zero), 2 compounds show a destabilizing effect (negative value of  $T_m$ ), and 6 compounds show a stabilizing effect (positive value of  $T_m$ ). The stabilizing effects of compounds 1, 5, 7, 9 and 14 are presented in Figure 4-4.



**Figure 4-4:** Fluorescence (given in fluorescence units, FU) plotted against temperature (one measurement per compound). The lower panel shows the derivative, for easier visualization of the shifts in  $T_m$ . Compound number is given in parentheses. The reference is the black line (2% DMSO).

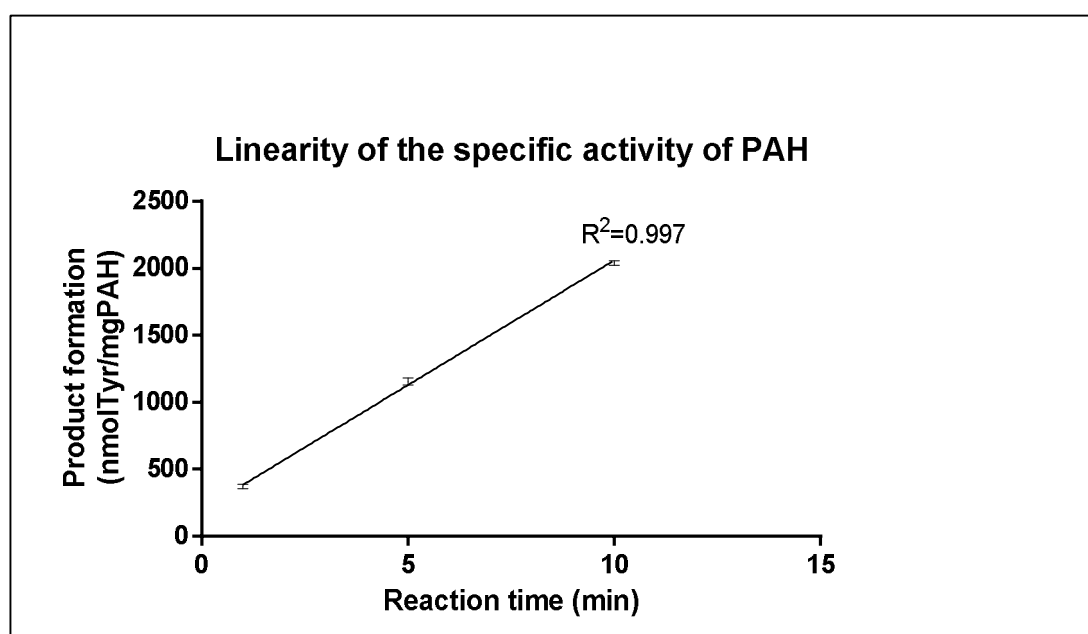
\*C04 is just enzyme (without DMSO). This means that DMSO destabilizes with 0.8 °C (i.e. not adding DMSO increase  $T_m$  with 0.8 °C).

We see that the lines are shifted to the right of the reference, increasing  $T_m$ . The best compound (number 1) increases  $T_m$  by 4 °C.

## 4.2.2 Effect of the compounds on activity of PAH

### Linearity of the specific activity of PAH

Figure 4-5 shows how much of tyrosine that is produced when the reaction time is varied from 1 to 10 minutes (measurements were done at 1, 5 and 10 minutes). The graph shows means of three parallels for each time, and the standard error is displayed as error bars (very small).

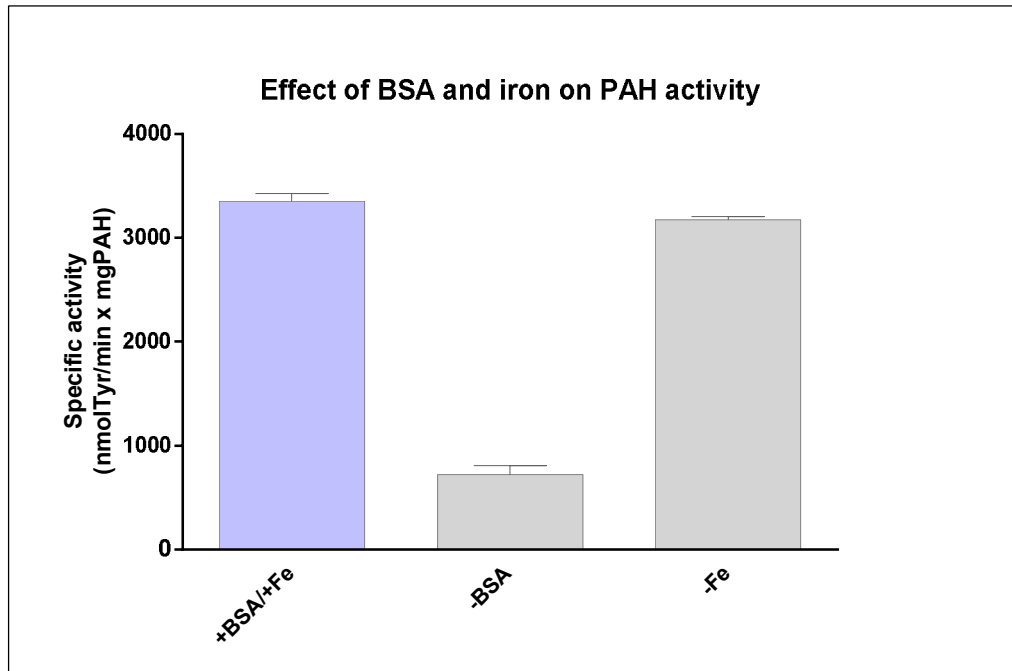


**Figure 4-5:** The product formation plotted against time (mean of triplicate measurements  $\pm$  SEM).

We see that the product formation and the time of reaction have an almost linear relationship. Being in the linear part of the curve is important (if we continued increasing the reaction time, the curve would stop being linear at some point because of increased product concentration in the reaction and eventually partial inactivation of PAH), as it allows us to see the effects of possible pharmacological chaperones. Also, we see that a reaction time of 1 minute when adding 0.25  $\mu\text{g}$  of PAH is sufficient to produce a detectable amount of L-Tyr.

## Optimization of activity assay

Figure 4-6 shows how removing BSA and removing iron affect PAH activity. Each column is the mean of three parallels, and the standard error is displayed.



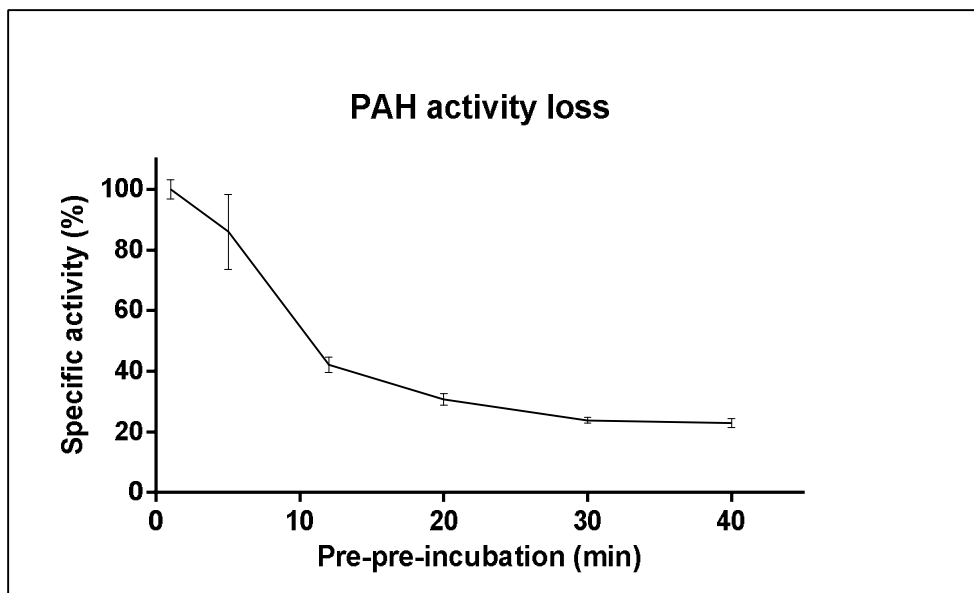
**Figure 4-6:** The effect of BSA and iron on PAH activity (mean of triplicate measurements  $\pm$  SEM)

We see that removing BSA clearly has a large impact on the activity, so we decided to keep it in the assay, but instead of having BSA in the enzyme solution, it was added in the assay mixture in future experiments.

In the original assay described by Knappskog et al (77), iron was added the last minute of pre-incubation. Here, adding iron does not seem to affect the activity, and so we have found that when iron is added in the growth medium as described earlier, it is not necessary to add it in the assay. This makes it much easier to coordinate the assay times.

### Activity loss as a function of time

Figure 4-7 shows the observed activity loss when PAH is kept on 37 °C for 1 to 40 minutes before the assay is started (pre-pre-incubation). Measurements were done at 1, 5, 12, 20, 30 and 40 minutes, six parallels each time.



**Figure 4-7:** Relative specific activity plotted against time of pre-pre-incubation (mean of six parallels  $\pm$  SEM). Activity is given as percentages of the activity at 1 minute, which was set to 100 %.

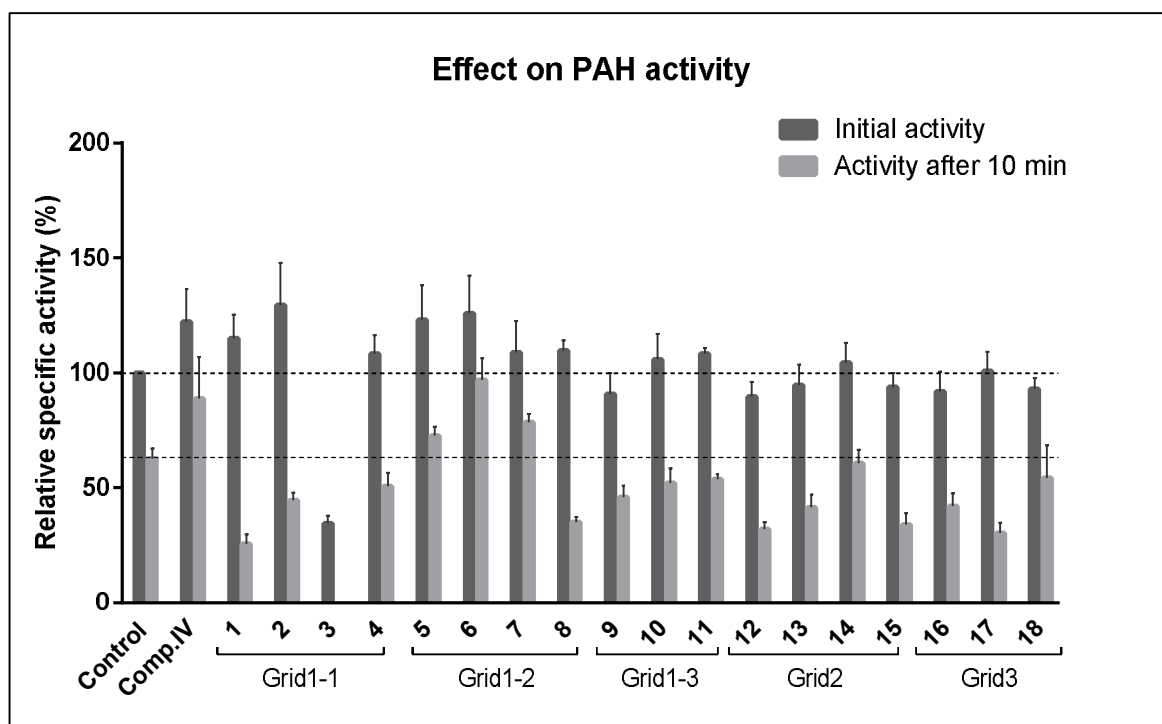
We measured a reduction of the activity to about 40 % after 12 minutes. This appears to be a good starting point for testing how the compounds interfere with PAH activity. If they can prevent or reduce the activity loss observed, they can be considered stabilizers for PAH.



## Testing of the hit compounds

To make sure that the fluorescent signals detected in HPLC were caused by the produced L-Tyr, all the compounds were tested to see if they had any natural fluorescence. As none of them had fluorescence at the same wavelength as L-Tyr, the signals should not be influenced by the compounds.

The results from the testing of the compounds are presented as means of four to eight parallels (varying number because of errors in readout of fluorescence, mistakes, or shortage of compounds), and given as percentages of the control at 0 min, which was set to 100%. Statistical data tables are also presented, including standard deviation (SD), standard error of the mean (SEM) and p-values. P-values were calculated using Prism's multiple t-tests function. All parallels and statistical data are shown in Appendix 2 and 3.



**Figure 4-8:** The compound's effect on PAH activity. Initial activity (0 min pre-pre-incubation) is shown next to the activity after 10 min pre-pre-incubation. The grids are specified. Compound IV (Comp.IV) is also included. The dotted lines represent 100 % activity (control at 0 min) and 63 % activity (control at 10 min). Columns are mean of 4-8 parallels  $\pm$  SEM.

The control has 63% remaining activity after 10 minutes of pre-pre-incubation (with 2% DMSO). 10 minutes were chosen as pre-pre-incubation time based on the results from the destabilizing curve in figure 4-7. The compounds being able to preserve an activity greater than 63% after 10 minutes are considered to be stabilizers.

**Table 4-6:** Statistical data for the initial activity assay (no pre-pre-incubation) showing mean, number of parallels, SD, SEM and p-value for each compound. \* marks a true discovery, and can be considered a significant difference from the control when correcting for multiple testing using FDR. The control was set to 100 %.

<b>Compound</b>	<b>Mean (%)</b>	<b>Number of parallels</b>	<b>SD</b>	<b>SEM</b>	<b>p-value</b>
Control 0 min	100.0	20	27.49	6.147	
Comp.IV	122.3	4	28.55	14.28	0.1318
1	115.1	6	25.34	10.34	0.2296
2	129.5	6	45.17	18.44	0.0192
3	34.58	6	8.272	3.377	< 0.0001*
4	108.5	6	19.80	8.085	0.4986
5	123.3	6	36.59	14.94	0.0640
6	125.8	6	40.49	16.53	0.0404
7	109.0	6	33.26	13.58	0.4737
8	109,7	7	12.16	4.598	0.4132
9	90.79	4	18.53	9.264	0.5332
10	105.9	7	29.1	11.00	0.6186
11	108.3	6	6.541	2.670	0.5088
12	89.78	7	16.87	6.376	0.3886
13	94.84	7	22.94	8.671	0.6632
14	104.5	6	21.19	8.649	0.7201
15	93.83	7	16.88	6.378	0.6026
16	92.05	7	22.75	8.601	0.5024
17	100.8	7	22.48	8.498	0.9462
18	93.12	7	12.52	4.732	0.5616

Looking at the results from the initial assay, we see that very few of the compounds tested seem to have a (significant) effect on the initial PAH activity. This is expected, as stabilizing effects will not be seen initially, but over time. Compound number 3 appears to be an inhibitor. Compound IV, which has previously been shown to be a weak inhibitor (50), does not have any significant effect on the initial activity in our experiment (if anything, it increases the activity).

**Table 4-7:** Statistical data for the activity stabilization assay (10 min pre-pre-incubation), showing mean, number of parallels, SD, SEM and p-value for each compound. The compounds having a higher activity than the control (63 %) are highlighted. \* marks a true discovery, and can be considered a significant difference from the control when correcting for multiple testing using FDR.

<b>Compound</b>	<b>Mean (%)</b>	<b>Number of parallels</b>	<b>SD</b>	<b>SEM</b>	<b>p-value</b>
Control 0 min	100.0				
Control 10 min	62.76	16	17.7	4.424	
Comp.IV	88.82	4	36.56	18.28	0.0080
1	25.74	7	11.06	4.181	< 0.0001*
2	44.63	7	8.965	3.388	0.0227
3	0	7	0	0	< 0.0001*
4	50.75	6	14.46	5.902	0.1525
5	72.62	6	9.587	3.914	0.2399
6	96.95	6	23.36	9.537	< 0.0001*
7	78.46	6	9.549	3.899	0.0617
8	35.11	6	5.567	2.273	0.0011*
9	46.03	6	12.09	4.934	0.0465
10	52.25	7	17.15	6.482	0.1858
11	53.73	7	6.452	2.439	0.2555
12	32.04	7	8.425	3.184	0.0001*
13	41.61	7	14.34	5.421	0.0080
14	60.71	7	15.4	5.82	0.7962
15	34.01	7	13.6	5.139	0.0003*
16	42.00	8	16.00	5.658	0.0064
17	30.42	8	12.39	4.379	< 0.0001*
18	54.34	8	40.04	14.16	0.2672

Looking at the results from the activity stabilization assay, we see that compound 5, 6 and 7, in addition to compound IV, seem to stabilize PAH. Only compound 6 has a statistically significant stabilizing effect. The sample with compound 3 apparently has no remaining activity after 10 minutes, as expected from the inhibitory effects seen in the initial activity assay.

## 5 DISCUSSION

Although dietary treatment has proven successful to prevent brain damage in PKU patients, there is still a great need for treatments that are easier to comply with (1, 35). As PKU is a loss-of-function misfolding disease usually caused by mutations in the gene encoding PAH, one approach is to attempt to correct the misfolding of the mutant PAH. The potential of pharmacological chaperones for this purpose represents a promising approach to treating PKU and other misfolding diseases (72). The natural cofactor BH<sub>4</sub> has been shown to act as a pharmacological chaperone for PAH in high concentrations, but only some patients are responsive (36). The pharmacological chaperones compounds III and IV (50), show promising effects in cells and normal mice, but the effects are low when tested in mouse models of PKU (Aurora Martinez, personal communication). No further development has been reported for the compounds discovered by Santos-Sierra et al. (51). Thus, better pharmacological chaperones for PAH are necessary.

In this project, we have screened a virtual library of drug-like compounds by docking them to the active site of PAH, and also to alternative binding pockets on the protein surface, with the aim of identifying compounds with a pharmacological chaperoning potential for PAH. This potential was investigated by evaluating the hit compounds in two in vitro assays, a thermostability assay and an activity assay. The use of molecular docking in screening of compounds is an established ó but also disputed ó approach in drug discovery (54-56). This thesis contributes to investigate the validity of docking and virtual screening as tools in the early phases of drug discovery, particularly in the development of pharmacological chaperones.

The screening identified 18 compounds that underwent thermostability and activity assays. One of the compounds has a significant stabilizing effect on PAH activity, and two other compounds might also be stabilizers (although not significant). One of the compounds inhibits PAH.

Here follows a discussion of the methods used and the results obtained.

## 5.1 Target-based virtual screening

There are several docking algorithms and software $\text{\AA}$ s available, with various features and different search and scoring functions. A number of papers evaluating the different docking performances have been published  $\acute{o}$  however, how to measure docking performance is not generally established, and it is therefore difficult to compare these papers. Glide, which is the program we have used, is generally appreciated as being one of the most accurate docking programs (57, 81).

An advantage of using virtual screening over traditional approaches in drug discovery, like experimental HTS, is the possibility of screening larger databases, containing novel ligands that have not yet been synthesized. The ligands can be filtered based on molecular properties, which is obviously very useful. Although the screening of large databases is computationally demanding, the cost is much lower and the effectiveness much higher, compared to eHTS (54-56)).

A prerequisite for docking is the availability of an x-ray structure of the target protein. However, with the sequencing of the human genome and the proteomic research increasing its scope, the number of protein targets is also increasing.

### **Choice of PDB structures**

There are several crystal structures of PAH available in the Protein Data Bank, and we chose PDB 1j8u as target structure for docking to the active site (see Table 3-1). This structure has the highest resolution out of the available structures (1.5  $\text{\AA}$ ), and the active site iron is in its reduced ferrous state, which is desirable, as this is the catalytically relevant state, as opposed to the earlier structures of PAH where the active site iron is in its inactive ferric state. Additionally, this structure was solved in complex with the natural cofactor of the enzyme,  $\text{BH}_4$ , which has been shown to act as a pharmacological chaperone for PAH (36, 37). There are no significant conformational changes observed upon cofactor binding to the catalytic domain of PAH (68), and we found the cofactor bound structure to be the most relevant for screening compounds towards the active site of PAH. It is important to keep in mind that Glide treats the protein receptor as a rigid structure, and the conformational changes that occur upon substrate binding, obviously will not happen virtually. We know that Tyr138

moves from a position on the outside of the receptor, to a position inside the active site (16) when the substrate binds, so it is sensible to use a structure without a bound substrate.

### **Docking to binding sites on the protein surface**

Docking to the active site of the target protein is the common procedure. We have also docked to two possible binding sites on the protein surface. We believe that investigating other sites is important, as there is no good reason to refrain from doing this. Interactions with residues on the protein surface could stabilize the protein, just as interactions in the active site could do so. Ligands that bind in the active site are likely to be inhibitors, and weak inhibition seems to be a common feature for pharmacological chaperones (82), but if a ligand binds too tightly and exerts considerable inhibition, that would of course not be desirable. It is therefore interesting to search for non-inhibitory chaperones that could stabilize the enzyme by binding to an alternative binding site. The program used to identify these sites, LIGSITE, is shown to have high precision (69), and the sites identified looked sensible when visualizing the pockets in Discovery Studio (Figure 3-1). The four active sites in the tetramer were correctly identified as binding sites by the program, which is reassuring. There is no crystal structure of full-length PAH, and so several structures were combined to make a hybrid structure of the tetramer to use as protein target when docking to the surface binding sites (Table 3-1).

#### **5.1.1 Cross-docking of selected compounds**

BH<sub>4</sub> and compound IV were both successfully re-docked/cross-docked, whereas the substrate analogues seem to dock to the cofactor binding site instead of the substrate binding site (Figure 4-2). It could be that the substrate binds to the cofactor binding site in the absence of cofactor, which possibly could explain the observed substrate inhibition of PAH at high L-Phe concentrations (83). However, this would have to be investigated in more detail to make any conclusions. Docking to PAH in complex with its cofactor could have guided the substrate analogues into the substrate binding site. However, due to time constraints we could not investigate the cross-docking of the substrate analogues any further.

Metal coordination and water molecules in the active site present a challenge for docking. Standard scoring functions are poor at estimating the contribution from possible interactions with metals in the receptor. The removal of structural water molecules from the active site would allow ligands that are unable to replace these to dock. The inclusion of non-structural

water molecules would exclude accessible areas of the active site for ligand to bind. Given that  $\text{BH}_4$  interacts with active site-residues through nearby water molecules, and that compound IV coordinates to iron, it is reassuring that both ligands seem to be re-docked/cross-docked satisfactory (Figure 4-2). Furthermore, we obtained a lower RMSD for  $\text{BH}_4$  when we included water molecules 3 Å from  $\text{BH}_4$  (0.30 Å) (Figure 4-3), and so deleting these seem to have an effect. Based on this we chose to keep water molecules in some of the grids when docking the Sigma library.

Taking a closer look at the results for compound IV, we see that the lowest RMSD-value (0.85 Å, SP rigid mode), is accompanied by a less favourable docking score (Table 4-1). This suggests that the orientation is correct, but that this orientation is considered unfavourable by Glide. The nitrogen atom (N1) of compound IV is non-charged, forming a bond to the iron atom, and it seems that Glide would rather prefer a charge on this nitrogen, thus giving the non-charged pose a high penalty. This addresses the importance of being aware of the limitations of docking when working with structures where metal coordination plays a role. If compound IV had been part of a large library that was screened, it would probably not have been a hit.

As both re-docking of  $\text{BH}_4$  and cross-docking of compound IV produced low RMSD values, we concluded that PDB 1j8u was a suitable target protein and that keeping most settings as default, as described in section 3.2.3, works satisfactory. Validation of docking to the alternative binding sites was not possible to perform, for obvious reasons.

### 5.1.2 Docking of the Sigma-library

The Sigma-library was chosen because it is easily accessible, and like the other catalogues provided by Zinc, it is filtered for drug discovery relevance and considered suitable for docking (60, 61).

The number of hits was limited to 500 compounds per grid. For the three active site-grids, many of the compounds are cofactor analogues, phenylalanine analogues or small peptides. For grid2 and 3 the compounds are generally larger, and many of them are peptides. In general, many of the hit compounds are L-Phe analogues or L-Phe-containing peptides. This could be thought to be due to an overrepresentation of this residue in the Sigma-library. A substructure search of the Sigma-library identified 1.3% L-Phe-analogues, whereas for our hits (all 2500 compounds) the corresponding number was 7.4%. Thus, the original library does not have an overrepresentation of L-Phe, suggesting that it provides favourable interactions in the binding pockets.

The 18 hits shown in Table 4-4, were chosen for experimental testing based on their Glide score, price and availability, and the fact that we wanted them to represent different types of compounds in all grids. Compound 1 is almost identical to one of the pharmacological chaperones previously mentioned, identified by Santos-Sierra et al. (51) (they only differ by a hydroxyl substituent on the ring for compound 1).



## 5.2 Testing virtual hits experimentally

By combining DSF with activity experiments, we were able to detect compounds that thermally stabilize PAH, and compounds that stabilize by protecting PAH from losing activity.

### 5.2.1 Thermostability

DSF is a quick and easy method for assaying the thermal stability of PAH in the presence of the hit compounds. It is time efficient and easy to perform, and it is in fact used as a high throughput experimental screening of compounds in the lab (75, 76). Here, it is used to complement the activity data.

Six compounds increase the melting point of PAH (Figure 4-4), compound number 1 giving the largest increase (4°C), suggesting that it has a stabilizing effect on the PAH structure. This compound (hydroxybenzylhydantoin) is a cofactor analogue (see structure in Table 4-4) that, as previously mentioned, is very similar to benzylhydantoin, which is the pharmacological chaperone identified by Santos-Sierra et al. (51). They used ligand-based virtual screening in the search for compounds with similar properties as BH<sub>4</sub>. Compound 1 docked to one of the water molecule-containing grids (grid1-1), that were generated based on the observation that this grid was optimal for re-docking of the cofactor (Table 4-3). A closer investigation of the compounds giving high docking score with grid1-1 reveals that they have similar interactions with PAH as the cofactor (Appendix 1). This suggests that inclusion of water molecules when preparing the grid for docking is the optimal approach to identify compounds that have the same binding properties as the cofactor. In general, this target-based approach of course requires that the compound for which analogues are being pursued, is available in complex with its receptor, with water molecules included.

Two compounds decrease the melting point (compound 3 with -5.8 °C and compound 15 with -4.1 °C). These two compounds presumably destabilize the folded state, which is not a desirable feature of a pharmacological chaperone. We chose to keep also these compounds for further validation by activity measurements.

## 5.2.2 Enzymatic activity

Several activity assays were performed, as described in section 3.3.2. The standardized PAH activity assay has a pre-incubation with L-Phe for five minutes, where iron is present the last minute of pre-incubation, before the reaction is started. This procedure ensures that PAH is fully activated by the time the reaction starts (77).

### **Optimization of the activity assay**

We found that adding iron to the reaction mixture did not increase the activity of the enzyme (Figure 4-6). This is contrary to what was found in the original paper where the procedure for activity measurements is presented (77). However, we added iron to the media when PAH was expressed by the bacteria, which presumably have saturated the purified enzyme with iron. BSA is traditionally used as a crowding agent to stabilize the enzyme in diluted solutions. When removing BSA from the pre-pre-incubation, we can observe a reduction in activity over a relatively short time frame (60 % reduction in 12 minutes of pre-pre-incubation) (Figure 4-7), which is reasonable for testing the effect of our potential pharmacological chaperones. Additionally, the presence of BSA during pre-pre-incubation with hit compounds could potentially diminish the effect of our compounds.

### **Testing of the compounds**

As the assays were done manually with pipettes, some differences between and within rows on the plate are unavoidable. There could also be time dependent differences within one plate, as the last row was done about 30 minutes after the first one. We do have controls on each row, and so we could have presented activities as percentages of the control on the same row to account for this. However, the variance between controls seems to be random. By presenting activities as percentages of the activity of the mean of the controls at 0 min, we are able to plot the results from the initial assay and the pre-pre-incubation assay together, providing more satisfactory comparisons.

Looking at Figure 4-8, we see that out of the compounds tested, compound 3 had a significant inhibitory effect on PAH activity, and compound 6 showed a significant protective effect on PAH activity. Also compounds 5 and 7 seem to be stabilizers, but these effects were not statistically significant. There is a large standard deviation for some of the compounds, reflecting high variability between the parallels. This could be due to different conditions

(concentration of enzyme, concentration of BH<sub>4</sub>, concentrations of components in the mix solution etc.), inaccurate timing of addition of enzyme, BH<sub>4</sub> and stop solution, aggregation, poor mixing, and errors concerning the separation on the HPLC column and readout of fluorescence.

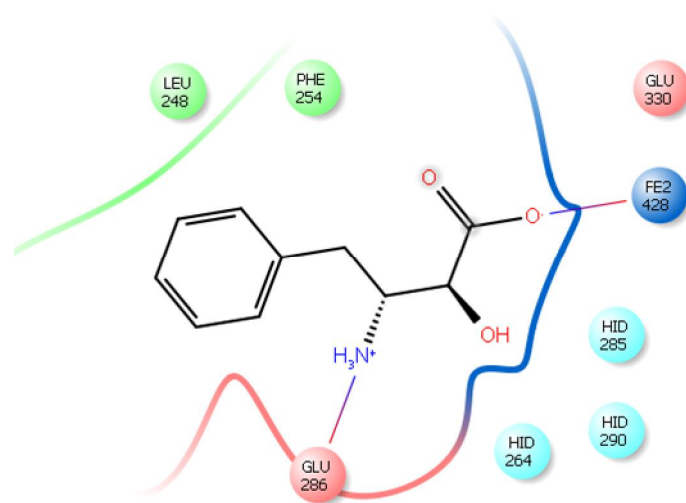
Compound IV has been shown to inhibit PAH to some extent (50). Our results suggest the opposite. This could be due to a stabilizing effect already seen in the initial activity assay ó even though there is no pre-pre-incubation, the compounds are present together with PAH in the 5 minute pre-incubation with L-Phe. This applies to all of the compounds, and we can not rule out an initial stabilizing effect.

The controls lost on average 40 % of their activity during 10 minutes of pre-pre-incubation with DMSO, compared to 60 % activity loss during 12 minutes of pre-pre-incubation without DMSO (as seen in the destabilization curve in Figure 4-7). The differences observed are likely due to effects of adding DMSO and the 2 minutes that differ in the pre-pre-incubation times.

### 5.3 Novel compounds with an effect on PAH

Looking at the results all together, the screening did identify several compounds that affect PAH. There is little consistency between the results from DSF and the activity experiments. The process of measuring protein stability is not straight forward, and we can not know the exact mechanisms of stabilization. Enzymatic activity is crucial for the protein function, and based on this we have considered it to be the most reliable measure to monitor protein stabilization in this project. Thus, compound 6 in particular, but also 5 and 7, appear to stabilize PAH. Compound 6 showed no effect on the melting point of the protein, and compounds 5 and 7 increased  $T_m$  with 0.6 and 0.7 °C, respectively, but as these values are very small and based on single measurements, they cannot be considered significant. Compound 3 stands out as a definite inhibitor, and it is conflicting that it appears to destabilize the protein structure (decrease  $T_m$ ).

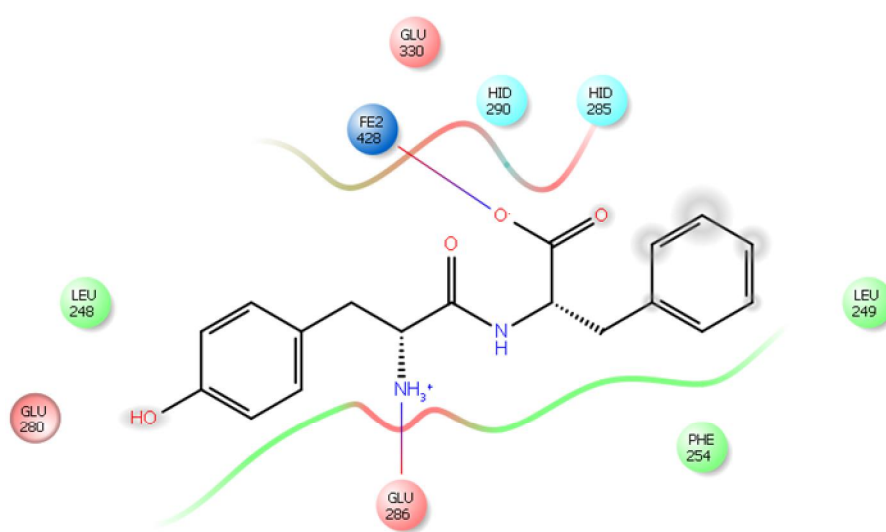
#### PAH stabilizers



**Figure 5-1:** Ligand interaction diagram for compound 6 ((2S, 3R)-3-amino-2-hydroxy-4-phenylbutyric acid). Amino acids are coloured by residue type (red: acidic, green: hydrophobic, blue: polar/metal, grey: other).

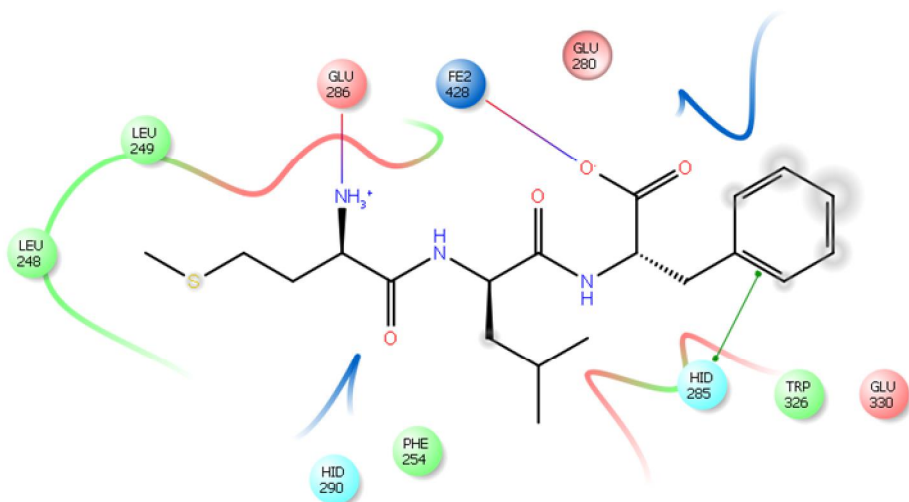
Compound 6 ((2S, 3R)-3-amino-2-hydroxy-4-phenylbutyric acid (AHPA)) is a substrate analogue that binds in the cofactor binding site (Appendix 1). It showed a significant stabilizing effect on PAH (97% remaining activity after 10 minutes compared to 63% for the control). It also appear to be a better stabilizer than compound IV (89% remaining activity) in our experiment, and it therefore has a great potential as a pharmacological chaperone for

PAH, and a possible future drug candidate for PKU. The AHPAs are shown to be enkephalinase inhibitors that potentiate the analgesic effect of morphine (84, 85). This implicates that compound 6 is not specific for PAH, and because it apparently penetrates the blood-brain-barrier (BBB), it is likely to affect the other AAAHs as well. However, the fact that this compound already has been studied makes it easier to access information about its properties and uses, and the structure could be further modified to be more specific to PAH. It meets the criteria of Lipinski's rule of five, which is a way of evaluating the "druglikeness" of a compound, i.e. whether a compound is likely to be successfully orally administered in humans (86).



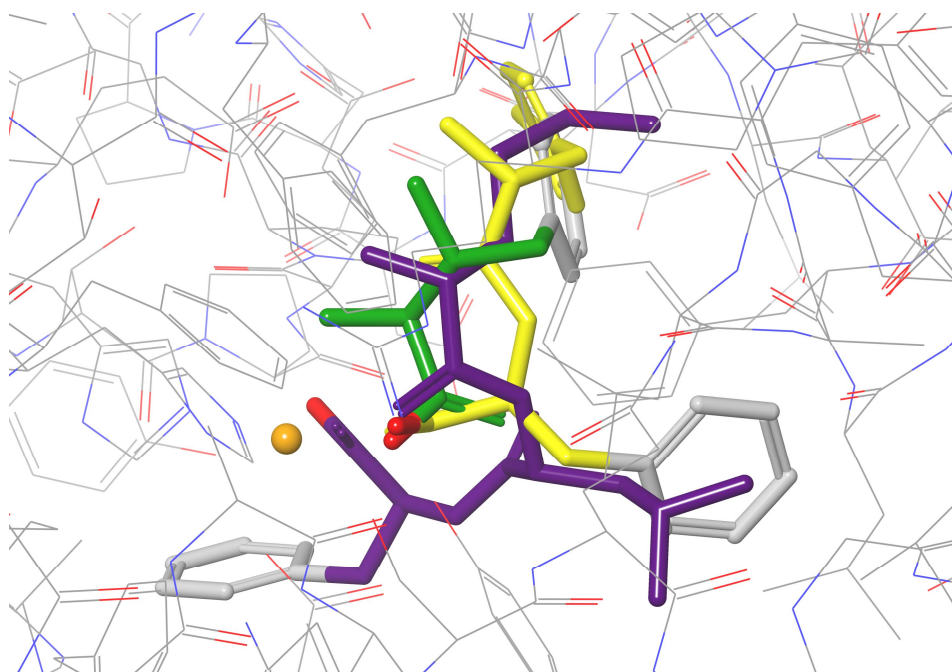
**Figure 5-2:** Ligand interaction diagram for compound 5 (tyrosylphenylalanine). Amino acids are coloured by residue type (red: acidic, green: hydrophobic, blue: polar/metal, grey: other)

Compound 5 (tyrosylphenylalanine) is a dipeptide that binds in the cofactor binding site. The presence of this compound resulted in a remaining PAH activity of 73% after 10 minutes pre-pre-incubation (compared to 63% for the control). This effect is not statistically significant, and more testing would be necessary to draw any conclusions. Peptides are generally difficult drug candidates, because of their susceptibility to degradation. However, modification of peptide structures into synthetic, more stable compounds that mimic the natural peptides (i.e. peptidomimetics) represents a promising field in drug design (see review (87)).



**Figure 5-3:** Ligand interaction diagram for compound 7 (methionyl-leucylphenylalanine). Amino acids are coloured by residue type (red: acidic, green: hydrophobic, blue: polar/metal, grey: other)

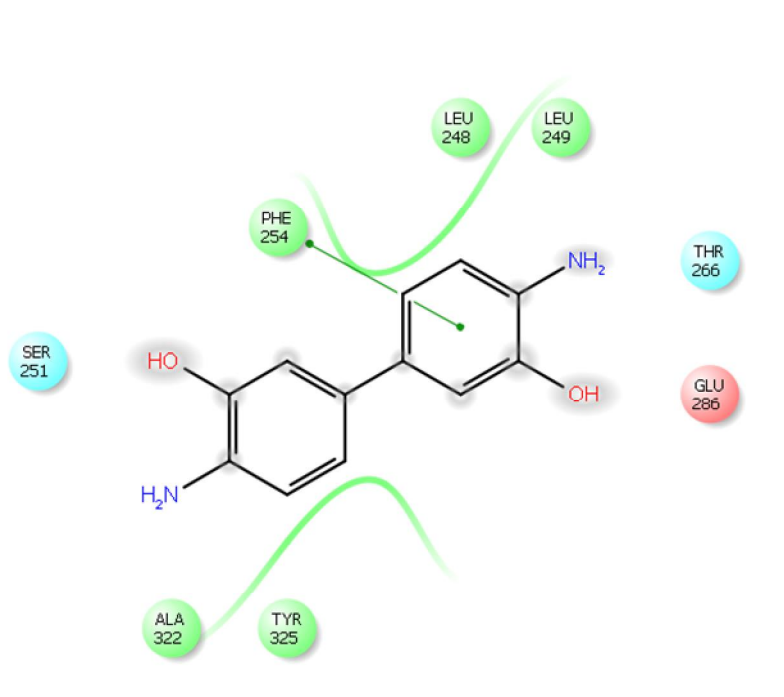
Compound 7 (methionyl-leucylphenylalanine acetate) is a tripeptide that binds in the active site pocket (extends through the cofactor binding site and the substrate binding site, see Appendix 1). The presence of this compound resulted in a remaining PAH activity of 78% after 10 minutes pre-pre-incubation (compared to 63% for the control). This effect is not statistically significant, and more testing would be necessary to draw any conclusions. Compound 7 has weak chemotactic properties, which has been utilized to design a peptide-grafted liposomal delivery system targeted to macrophages (88), and Primoquine (antiprotozoal drug) can be delivered this way. The same peptide issues concerning drug potential apply to compound 7.



**Figure 5-4:** Binding modes of compound 6 (green), compound 5 (yellow) and compound 7 (purple) in the active site. Iron is also shown as a sphere. Oxygen atom near iron is coloured red, and represents a possibility of coordination to iron. Phenylalanine rings, or phenylalanine-like rings in the case of compound 6, are coloured grey for comparison of their positions.

All three stabilizers have in common that they are either L-Phe analogues or peptides containing L-Phe. The phenylalanine-like ring are located in different positions, as illustrated in figure 5-4: for compound 5 the ring is located in the opposite end of the cofactor binding site compared to compound 6, and for compound 7 it is located in the substrate binding site. All compounds show a possible coordination to iron through the oxygen atom.

## PAH inhibitor



**Figure 5-5:** Ligand interaction diagram for compound 3 (4,4 -Diamino[1,1 -biphenyl]-3,3 -diol). Amino acids are coloured by residue type (red: acidic, green: hydrophobic, blue: polar/metal, grey: other)

Because of the strong inhibitory effects seen with compound 3 (4,4 -Diamino[1,1 -biphenyl]-3,3 -diol), it is worth mentioning, even though inhibition of PAH is not desirable. However, it could be of interest in *in vitro* assays where inducing PKU would be relevant. PAH activity goes from 34% initially in the presence of compound 3, to no activity after 10 minutes. Several of the compounds show inhibitory effects, but none as pronounced as compound 3.



## 6 FUTURE PERSPECTIVES

Out of the stabilizers, two of the compounds do not show a strictly significant stabilizing effect, probably because of variations between parallels. For these, the experiments should be repeated in order to validate our findings. Further optimization of affinity of the compounds towards PAH and screening for compound analogues could lead to increased stabilizing effects. Insight in all thermodynamic binding properties could for example be obtained by using isothermal titration calorimetry (ITC).

As selectivity for PAH is desirable, the compounds need to be tested on TH and TPH. Implications of stabilizing or inhibitory effects on the other AAAHs should be considered carefully, as well as the possibility of modifying the compounds to be more selective to PAH, if selectivity is an issue.

We have only tested the compounds on wt-PAH, and testing on specific mutants would be relevant and could open up to individualized treatment of different PKU genotypes.

For the compounds to be applicable as drug candidates for treating PKU, they would have to be tested in cells and further in animal models, to explore activity and toxicity, and eventually pharmacokinetic properties.

## **7 CONCLUDING REMARKS**

By combining virtual and experimental methods, we have identified a potential pharmacological chaperone that is able to protect PAH from losing activity, and therefore could be considered a possible candidate to treat PKU. We have also identified two peptides that show non-significant stabilizing effects and require further testing for them to be considered as potential pharmacological chaperones for PAH. A PAH inhibitor which might show interesting biotechnological approaches was also discovered.

The approach taken here, with target-based virtual screening and experimental validation of hits could be applied to other targets to find possible stabilizers or inhibitors.

## 8 REFERENCES

1. Scriver C R KS. Hyperphenylalaninemia phenylalanine hydroxylase deficiency. In: Scriver C R, Beaudet A L, Valle D, Sly W S, editors. *The Metabolic and Molecular bases of Inherited Disease*. 8th. New York: McGraw-Hill; 2001. pp. 1667-1724.
2. Flydal MI, Martinez A. Phenylalanine hydroxylase: function, structure, and regulation. *IUBMB life*. 2013;65(4):341-9.
3. Fitzpatrick PF. Mechanism of aromatic amino acid hydroxylation. *Biochemistry*. 2003;42(48):14083-91.
4. Olsson E TK, Martinez A, Jensen V R. . The Aromatic Amino Acid Hydroxylase Mechanism A Perspective from Computational Chemistry. *Adv. Inorg. Chem* . 2010;62:437-500.
5. Werner ER, Blau N, Thony B. Tetrahydrobiopterin: biochemistry and pathophysiology. *The Biochemical journal*. 2011;438(3):397-414.
6. Erlandsen H, Fusetti F, Martinez A, Hough E, Flatmark T, Stevens RC. Crystal structure of the catalytic domain of human phenylalanine hydroxylase reveals the structural basis for phenylketonuria. *Nature structural biology*. 1997;4(12):995-1000.
7. Fusetti F, Erlandsen H, Flatmark T, Stevens RC. Structure of tetrameric human phenylalanine hydroxylase and its implications for phenylketonuria. *The Journal of biological chemistry*. 1998;273(27):16962-7.
8. Kobe B, Jennings IG, House CM, Michell BJ, Goodwill KE, Santarsiero BD, et al. Structural basis of autoregulation of phenylalanine hydroxylase. *Nature structural biology*. 1999;6(5):442-8.
9. Abita JP, Milstien S, Chang N, Kaufman S. In vitro activation of rat liver phenylalanine hydroxylase by phosphorylation. *The Journal of biological chemistry*. 1976;251(17):5310-4.
10. Fitzpatrick PF. Allosteric regulation of phenylalanine hydroxylase. *Archives of biochemistry and biophysics*. 2012;519(2):194-201.
11. Nielsen KH. Rat liver phenylalanine hydroxylase. A method for the measurement of activity, with particular reference to the distinctive features of the enzyme and the pteridine cofactor. *European journal of biochemistry / FEBS*. 1969;7(3):360-9.

12. Bailey SW, Ayling JE. Separation and properties of the 6-diastereoisomers of l-erythro-tetrahydrobiopterin and their reactivities with phenylalanine hydroxylase. *The Journal of biological chemistry*. 1978;253(5):1598-605.
13. Koehntop KD, Emerson JP, Que L, Jr. The 2-His-1-carboxylate facial triad: a versatile platform for dioxygen activation by mononuclear non-heme iron(II) enzymes. *Journal of biological inorganic chemistry : JBIC : a publication of the Society of Biological Inorganic Chemistry*. 2005;10(2):87-93.
14. Teigen K, McKinney JA, Haavik J, Martinez A. Selectivity and affinity determinants for ligand binding to the aromatic amino acid hydroxylases. *Current medicinal chemistry*. 2007;14(4):455-67.
15. Teigen K, Froystein NA, Martinez A. The structural basis of the recognition of phenylalanine and pterin cofactors by phenylalanine hydroxylase: implications for the catalytic mechanism. *Journal of molecular biology*. 1999;294(3):807-23.
16. Andersen OA, Stokka AJ, Flatmark T, Hough E. 2.0Å resolution crystal structures of the ternary complexes of human phenylalanine hydroxylase catalytic domain with tetrahydrobiopterin and 3-(2-thienyl)-L-alanine or L-norleucine: substrate specificity and molecular motions related to substrate binding. *Journal of molecular biology*. 2003;333(4):747-57.
17. Bjorgo E, de Carvalho RM, Flatmark T. A comparison of kinetic and regulatory properties of the tetrameric and dimeric forms of wild-type and Thr427-->Pro mutant human phenylalanine hydroxylase: contribution of the flexible hinge region Asp425-Gln429 to the tetramerization and cooperative substrate binding. *European journal of biochemistry / FEBS*. 2001;268(4):997-1005.
18. Jaffe EK, Stith L, Lawrence SH, Andrade M, Dunbrack RL, Jr. A new model for allosteric regulation of phenylalanine hydroxylase: implications for disease and therapeutics. *Archives of biochemistry and biophysics*. 2013;530(2):73-82.
19. Thorolfsson M, Teigen K, Martinez A. Activation of phenylalanine hydroxylase: effect of substitutions at Arg68 and Cys237. *Biochemistry*. 2003;42(12):3419-28.
20. Stokka AJ, Carvalho, R.N., Barroso, J.F., Flatmark, T. Probing the role of crystallographically defined/predicted hinge-bending regions in the substrate-induced global conformational transition and catalytic activation of human phenylalanine hydroxylase by single-site mutagenesis. *J Biol Chem* 2004;279:26571-80.
21. Haavik J, Blau N, Thony B. Mutations in human monoamine-related neurotransmitter pathway genes. *Human mutation*. 2008;29(7):891-902.

22. Kaufman S. Tyrosine hydroxylase. *Advances in enzymology and related areas of molecular biology*. 1995;70:103-220.
23. Ludecke B, Knappskog PM, Clayton PT, Surtees RA, Clelland JD, Heales SJ, et al. Recessively inherited L-DOPA-responsive parkinsonism in infancy caused by a point mutation (L205P) in the tyrosine hydroxylase gene. *Human molecular genetics*. 1996;5(7):1023-8.
24. Swaans RJ, Rondot P, Renier WO, Van Den Heuvel LP, Steenbergen-Spanjers GC, Wevers RA. Four novel mutations in the tyrosine hydroxylase gene in patients with infantile parkinsonism. *Annals of human genetics*. 2000;64(Pt 1):25-31.
25. Haavik J, Toska K. Tyrosine hydroxylase and Parkinson's disease. *Molecular neurobiology*. 1998;16(3):285-309.
26. Walther DJ, Peter JU, Bashammakh S, Hortnagl H, Voits M, Fink H, et al. Synthesis of serotonin by a second tryptophan hydroxylase isoform. *Science (New York, NY)*. 2003;299(5603):76.
27. Zhang X, Gainetdinov RR, Beaulieu JM, Sotnikova TD, Burch LH, Williams RB, et al. Loss-of-function mutation in tryptophan hydroxylase-2 identified in unipolar major depression. *Neuron*. 2005;45(1):11-6.
28. Li D, He L. Further clarification of the contribution of the tryptophan hydroxylase (TPH) gene to suicidal behavior using systematic allelic and genotypic meta-analyses. *Human genetics*. 2006;119(3):233-40.
29. McKinney J, Johansson S, Halmoy A, Dramsdahl M, Winge I, Knappskog PM, et al. A loss-of-function mutation in tryptophan hydroxylase 2 segregating with attention-deficit/hyperactivity disorder. *Molecular psychiatry*. 2008;13(4):365-7.
30. Fölling A. Über ausscheidung von phenylbrenztraubensäure in den harn als stoffwechselanomalie in verbindung mit imbezillitat. *Hoppe-Seyl*. 1934;227:1696176.
31. Hofmann B. Nyfødtscreening - mer skjult tvang? *Tidsskr Nor Legeforen* 2010; 130: 291-3.
32. Scriver CR, Hurtubise M, Konecki D, Phommavanh M, Prevost L, Erlandsen H, et al. PAHdb 2003: what a locus-specific knowledgebase can do. *Human mutation*. 2003;21(4):333-44.
33. Scriver CR. The PAH gene, phenylketonuria, and a paradigm shift. *Human mutation*. 2007;28(9):831-45.
34. Lindner M. Treatment of phenylketonuria variants: European recommendations. In: Blau N, ed. *PKU and BH4: advances in*

phenylketonuria and tetrahydrobiopterin. Heilbronn:

SPS Verlagsgesellschaft mbH, 2006: 180687.

35. Enns GM, Koch R, Brumm V, Blakely E, Suter R, Jurecki E. Suboptimal outcomes in patients with PKU treated early with diet alone: revisiting the evidence. *Molecular genetics and metabolism*. 2010;101(2-3):99-109.
36. Erlandsen H, Pey AL, Gamez A, Perez B, Desviat LR, Aguado C, et al. Correction of kinetic and stability defects by tetrahydrobiopterin in phenylketonuria patients with certain phenylalanine hydroxylase mutations. *Proceedings of the National Academy of Sciences of the United States of America*. 2004;101(48):16903-8.
37. Pey AL, Perez B, Desviat LR, Martinez MA, Aguado C, Erlandsen H, et al. Mechanisms underlying responsiveness to tetrahydrobiopterin in mild phenylketonuria mutations. *Human mutation*. 2004;24(5):388-99.
38. Ding Z, Georgiev P, Thony B. Administration-route and gender-independent long-term therapeutic correction of phenylketonuria (PKU) in a mouse model by recombinant adeno-associated virus 8 pseudotyped vector-mediated gene transfer. *Gene therapy*. 2006;13(7):587-93.
39. Ding Z, Harding CO, Rebuffat A, Elzaouk L, Wolff JA, Thony B. Correction of murine PKU following AAV-mediated intramuscular expression of a complete phenylalanine hydroxylating system. *Molecular therapy : the journal of the American Society of Gene Therapy*. 2008;16(4):673-81.
40. Matalon R, Michals-Matalon K, Bhatia G, Burlina AB, Burlina AP, Braga C, et al. Double blind placebo control trial of large neutral amino acids in treatment of PKU: effect on blood phenylalanine. *Journal of inherited metabolic disease*. 2007;30(2):153-8.
41. Sarkissian CN, Gamez A, Wang L, Charbonneau M, Fitzpatrick P, Lemontt JF, et al. Preclinical evaluation of multiple species of PEGylated recombinant phenylalanine ammonia lyase for the treatment of phenylketonuria. *Proceedings of the National Academy of Sciences of the United States of America*. 2008;105(52):20894-9.
42. Anfinsen CB. The formation and stabilization of protein structure. *The Biochemical journal*. 1972;128(4):737-49.
43. Anfinsen CB. Principles that govern the folding of protein chains. *Science (New York, NY)*. 1973;181(4096):223-30.
44. Gregersen N, Bross P, Vang S, Christensen JH. Protein misfolding and human disease. *Annual review of genomics and human genetics*. 2006;7:103-24.
45. Ellis RJ. The molecular chaperone concept. *Seminars in cell biology*. 1990;1(1):1-9.

46. Morello JP, Petaja-Repo UE, Bichet DG, Bouvier M. Pharmacological chaperones: a new twist on receptor folding. *Trends in pharmacological sciences*. 2000;21(12):466-9.
47. Conn PM, Leanos-Miranda A, Janovick JA. Protein origami: therapeutic rescue of misfolded gene products. *Molecular interventions*. 2002;2(5):308-16.
48. Cohen FE, Kelly JW. Therapeutic approaches to protein-misfolding diseases. *Nature*. 2003;426(6968):905-9.
49. Aymami J, Barril X, Rodriguez-Pascau L, Martinell M. Pharmacological chaperones for enzyme enhancement therapy in genetic diseases. *Pharmaceutical patent analyst*. 2013;2(1):109-24.
50. Pey AL, Ying M, Cremades N, Velazquez-Campoy A, Scherer T, Thony B, et al. Identification of pharmacological chaperones as potential therapeutic agents to treat phenylketonuria. *The Journal of clinical investigation*. 2008;118(8):2858-67.
51. Santos-Sierra S, Kirchmair J, Perna AM, Reiss D, Kemter K, Roschinger W, et al. Novel pharmacological chaperones that correct phenylketonuria in mice. *Human molecular genetics*. 2012;21(8):1877-87.
52. Torreblanca R, Lira-Navarrete E, Sancho J, Hurtado-Guerrero R. Structural and mechanistic basis of the interaction between a pharmacological chaperone and human phenylalanine hydroxylase. *Chembiochem : a European journal of chemical biology*. 2012;13(9):1266-9.
53. Calvo AC, Scherer T, Pey AL, Ying M, Winge I, McKinney J, et al. Effect of pharmacological chaperones on brain tyrosine hydroxylase and tryptophan hydroxylase 2. *Journal of neurochemistry*. 2010;114(3):853-63.
54. Klebe G. Virtual ligand screening: strategies, perspectives and limitations. *Drug discovery today*. 2006;11(13-14):580-94.
55. Bleicher KH, Bohm HJ, Muller K, Alanine AI. Hit and lead generation: beyond high-throughput screening. *Nature reviews Drug discovery*. 2003;2(5):369-78.
56. MA M, Sperandio O, Villoutreix BO. Virtual Ligand Screening for Structure-based Drug Design: Approaches and Progress. *Bioautomation* 2007;7:104-21.
57. Young DC. *Computational Drug Design*: John Wiley & Sons, Inc.; 2009.
58. Kirchmair J, Distinto S, Liedl KR, Markt P, Rollinger JM, Schuster D, et al. Development of Anti-Viral Agents Using Molecular Modeling and Virtual Screening Techniques. *Infectious disorders drug targets*. 2010.
59. Martinez A, Knappskog PM, Olafsdottir S, Doskeland AP, Eiken HG, Svebak RM, et al. Expression of recombinant human phenylalanine hydroxylase as fusion protein in

- Escherichia coli circumvents proteolytic degradation by host cell proteases. Isolation and characterization of the wild-type enzyme. *The Biochemical journal*. 1995;306 ( Pt 2):589-97.
60. Irwin JJ, Sterling T, Mysinger MM, Bolstad ES, Coleman RG. ZINC: A Free Tool to Discover Chemistry for Biology. *Journal of chemical information and modeling*. 2012;52(7):1757-68.
  61. Irwin JJ, Shoichet BK. ZINC--a free database of commercially available compounds for virtual screening. *Journal of chemical information and modeling*. 2005;45(1):177-82.
  62. Schrödinger Release 2013-2: Maestro, version 9.5, Schrödinger, LLC. New York, NY.
  63. Small-Molecule Drug Discovery Suite 2013-2: Glide, version 6.0, Schrödinger, LLC. New York, NY 2013.
  64. Schrödinger Release 2013-2: Schrödinger Suite 2013 Protein Preparation Wizard; Epik version 2.5, Schrödinger, LLC, New York, NY, 2013; Impact version 6.0, Schrödinger, LLC, New York, NY, 2013; Prime version 3.3, Schrödinger, LLC, New York, NY, 2013. 2013.
  65. Madhavi Sastry G, Adzhigirey M, Day T, Annabhimoju R, Sherman W. Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *J Comput Aided Mol Des*. 2013;27(3):221-34.
  66. Schrödinger Release 2013-2: LigPrep, version 2.7, Schrödinger, LLC, New York, NY, 2013.
  67. Glide, version 6.0., User Manual, Schrödinger, LLC, New York, NY. 2013.
  68. Andersen OA, Flatmark T, Hough E. High resolution crystal structures of the catalytic domain of human phenylalanine hydroxylase in its catalytically active Fe(II) form and binary complex with tetrahydrobiopterin. *Journal of molecular biology*. 2001;314(2):279-91.
  69. Hendlich M, Rippmann F, Barnickel G. LIGSITE: automatic and efficient detection of potential small molecule-binding sites in proteins. *Journal of molecular graphics & modelling*. 1997;15(6):359-63, 89.
  70. Sutherland JJ, Nandigam RK, Erickson JA, Vieth M. Lessons in molecular recognition. 2. Assessing and improving cross-docking accuracy. *Journal of chemical information and modeling*. 2007;47(6):2293-302.
  71. Accelrys Software Inc., Discovery Studio Modeling Environment, Release 4.0, San Diego: Accelrys Software Inc., 2013.
  72. Schrödinger Release 2013-2: Canvas, version 1.7, Schrödinger, LLC, New York, NY, 2013.



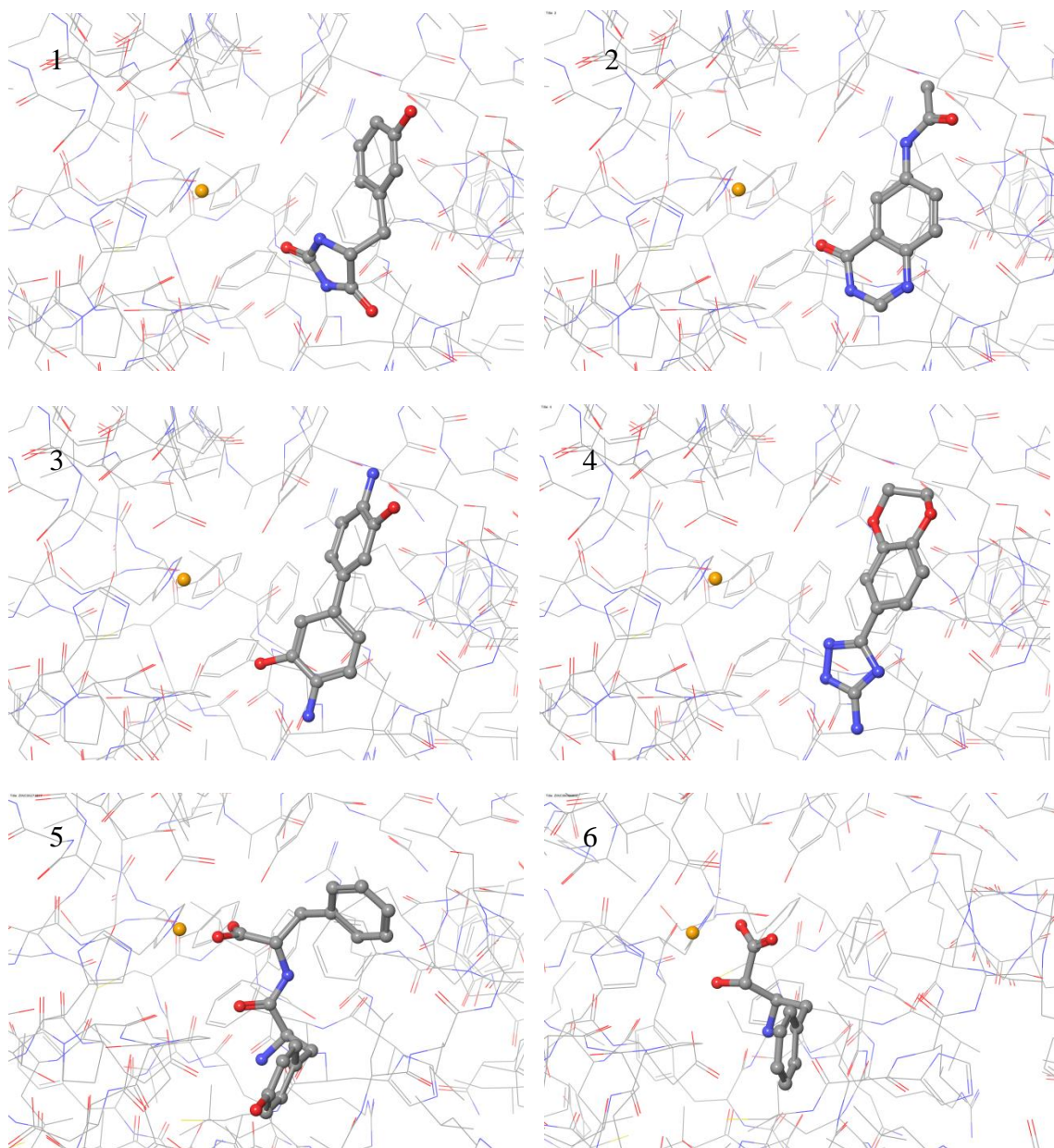
73. Duan J, Dixon SL, Lowrie JF, Sherman W. Analysis and comparison of 2D fingerprints: Insights into database screening performance using eight fingerprint methods. *Journal of Molecular Graphics and Modelling*. 2010;29(2):157-70.
74. Sastry M, Lowrie JF, Dixon SL, Sherman W. Large-Scale Systematic Analysis of 2D Fingerprint Methods and Parameters to Improve Virtual Screening Enrichments. *Journal of chemical information and modeling*. 2010;50(5):771-84.
75. Pantoliano MW, Petrella EC, Kwasnoski JD, Lobanov VS, Myslik J, Graf E, et al. High-density miniaturized thermal shift assays as a general strategy for drug discovery. *Journal of biomolecular screening*. 2001;6(6):429-40.
76. Lo MC, Aulabaugh A, Jin G, Cowling R, Bard J, Malamas M, et al. Evaluation of fluorescence-based thermal shift assays for hit identification in drug discovery. *Analytical biochemistry*. 2004;332(1):153-9.
77. Knappskog PM, Flatmark T, Aarden JM, Haavik J, Martinez A. Structure/function relationships in human phenylalanine hydroxylase. Effect of terminal deletions on the oligomerization, activation and cooperativity of substrate binding to the enzyme. *European journal of biochemistry / FEBS*. 1996;242(3):813-21.
78. Doskeland AP, Doskeland SO, Ogreid D, Flatmark T. The effect of ligands of phenylalanine 4-monooxygenase on the cAMP-dependent phosphorylation of the enzyme. *The Journal of biological chemistry*. 1984;259(18):11242-8.
79. Despa F, Orgill DP, Lee RC. Molecular crowding effects on protein stability. *Annals of the New York Academy of Sciences*. 2005;1066:54-66.
80. Benjamini YH, Yosef. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B* 57 (1): 289-300 1995.
81. Kellenberger E, Rodrigo J, Muller P, Rognan D. Comparative evaluation of eight docking tools for docking and virtual screening accuracy. *Proteins*. 2004;57(2):225-42.
82. Ringe D, Petsko GA. What are pharmacological chaperones and why are they interesting? *Journal of biology*. 2009;8(9):80.
83. Pey AL, Martinez A. The activity of wild-type and mutant phenylalanine hydroxylase and its regulation by phenylalanine and tetrahydrobiopterin at physiological and pathological concentrations: an isothermal titration calorimetry study. *Molecular genetics and metabolism*. 2005;86 Suppl 1:S43-53.
84. Hachisu M, Nakamura T, Kawashima H, Shitoh K, Fukatsu S, Koeda T, et al. Relationship between enhancement of morphine analgesia and inhibition of enkephalinase by

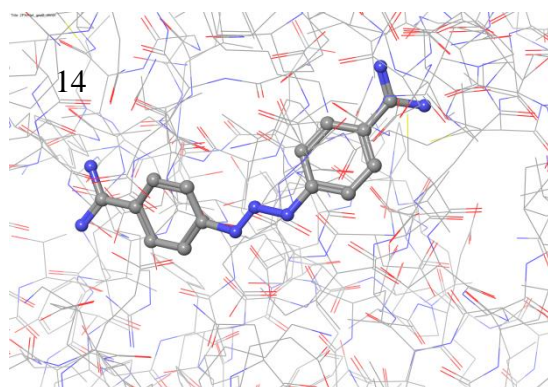
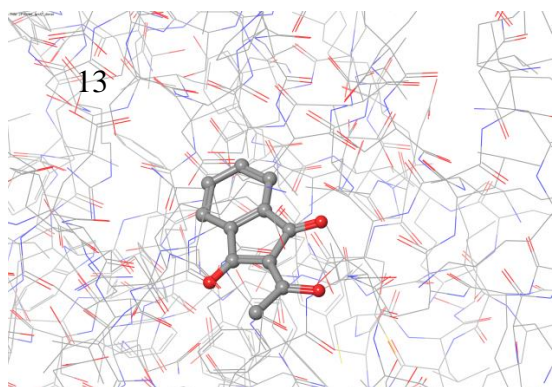
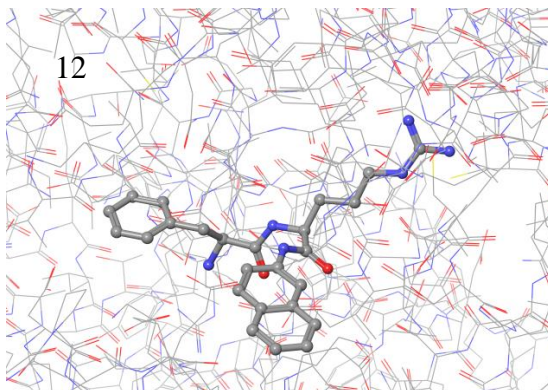
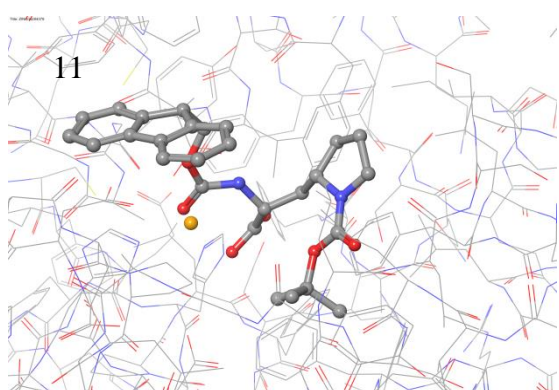
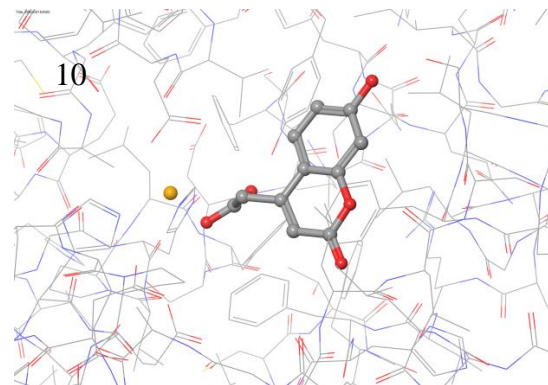
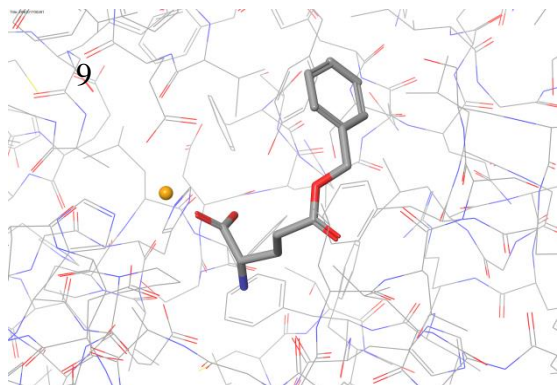
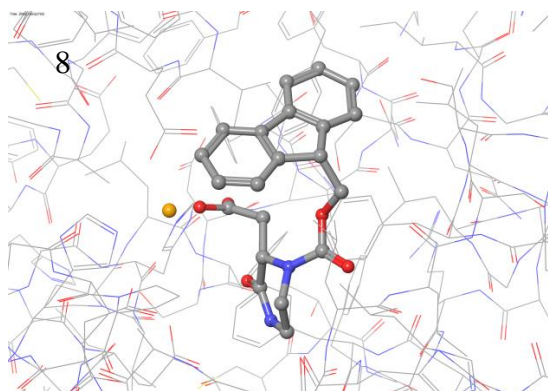
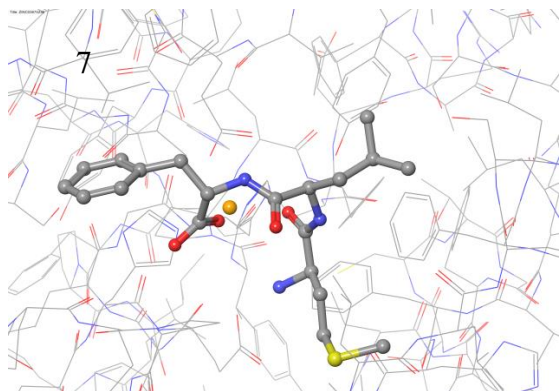
- 2S, 3R 3-amino-2-hydroxy-4-phenylbutanoic acid derivatives. *Life sciences*. 1982;30(20):1739-46.
85. Matsuoka Y, Satoh S, Uruno T, Kubota K. 2S,3R 3-amino-2-hydroxy-4-phenylbutanoic acid derivatives, enkephalinase inhibitors, augment met5-enkephalin-induced antinociception. *Japanese journal of pharmacology*. 1988;46(3):205-10.
86. Lipinski CA. Drug-like properties and the causes of poor solubility and poor permeability. *Journal of pharmacological and toxicological methods*. 2000;44(1):235-49.
87. Vagner J, Qu H, Hruby VJ. Peptidomimetics, a synthetic tool of drug discovery. *Current opinion in chemical biology*. 2008;12(3):292-6.
88. Banerjee G, Medda S, Basu MK. A novel peptide-grafted liposomal delivery system targeted to macrophages. *Antimicrobial agents and chemotherapy*. 1998;42(2):348-51.

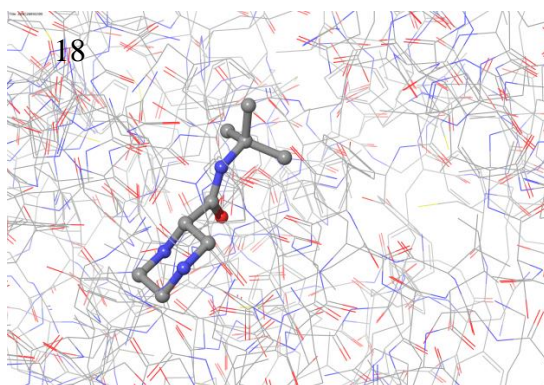
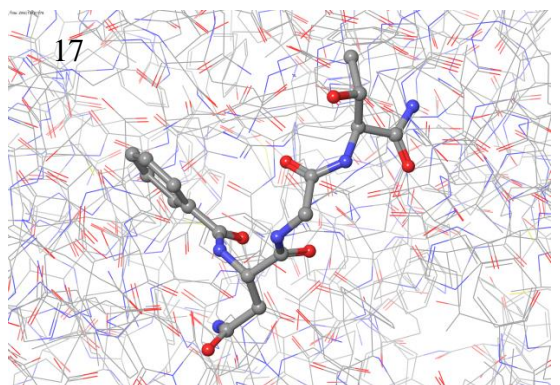
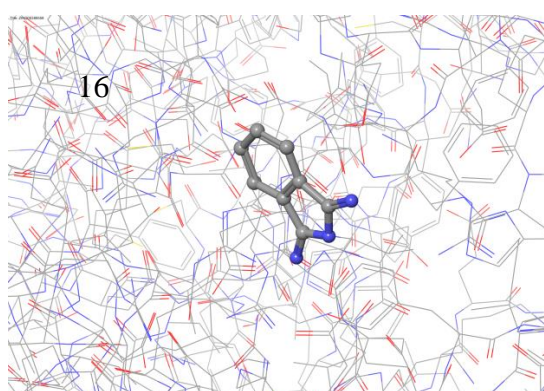
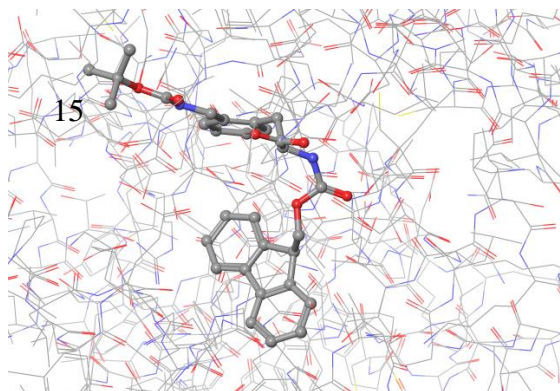
## 9 APPENDIX

### Appendix 1

Docking of compounds 1-18. Iron is shown as an orange sphere, and the compounds are coloured by element.







## Appendix 2

All parallels (designated P) and statistical calculations for the initial activity assay, printed from GraphPad Prism. Activity is given in percentage of the mean of the controls.

Table format: Grouped		Group A	Group B	Group C	Group D	Group E	Group F	Group G	Group H
		Control 0 min	Comp.IV	1	2	3	4	5	6
		Y	Y	Y	Y	Y	Y	Y	Y
1	P1	68.66497	132.864500	157.223000	199.296800	50.931400	132.864500	179.367100	183.795900
2	P2	85.83122	137.293400	132.864500	161.651800	33.216130	124.006900	146.151000	152.794200
3	P3	79.39387	139.507800	112.934800	137.293400	33.216130	117.363700	135.078900	139.507800
4	P4	92.26856	79.718710	94.414340	81.539660	27.895150	79.393880	83.685430	72.956540
5	P5	60.08185		96.560120	102.997500	32.186710	96.560120	98.705890	109.434800
6	P6	79.39387		96.560120	94.414340	30.040930	100.851700	96.560120	96.560120
7	P7	68.66497							
8	P8	66.51919							
9	P9	150.57980							
10	P10	152.79420							
11	P11	112.93480							
12	P12	79.39387							
13	P13	104.07720							
14	P14	106.29160							
15	P15	143.93660							
16	P16	119.57810							
17	P17	112.93480							
18	P18	106.29160							
19	P19	106.29160							
20	P20	104.07720							

		Group I	Group J	Group K	Group L	Group M	Group N	Group O	Group P	Group Q
		7	8	9	10	11	12	13	14	15
		Y	Y	Y	Y	Y	Y	Y	Y	Y
1		157.223000	128.435700	93.005170	141.722200	115.149300	108.506000	104.077200	121.792500	117.363700
2		128.435700	104.077200	99.648390	110.720400	112.934800	93.005170	95.219570	117.363700	101.862800
3		124.006900	121.792500	106.291600	126.221300	101.862800	93.005170	128.435700	124.006900	110.720400
4		75.102310	115.149300	64.217860	130.650100	111.580600	112.934800	115.149300	68.664970	93.005170
5		72.956540	96.560120		64.373410	109.434800	72.956540	83.685430	94.414340	83.685430
6		96.560120	100.851700		87.977000	98.705890	77.248090	64.373410	100.851700	75.102310
7			100.851700		79.393880		70.810750	72.956540		75.102310
8										
9										
10										
11										
12										
13										
14										
15										
16										
17										
18										
19										
20										

	Group R	Group S	Group T	Group U
	16	17	18	Data Set-U
	Y	Y	Y	Y
1	119.578100	119.578100	106.291600	
2	112.934800	121.792500	108.506000	
3	110.720400	121.792500	99.648390	
4	93.005170	110.720400	86.361940	
5	68.664970	81.539660	94.414340	
6	72.956540	72.956540	79.393880	
7	66.519200	77.248090	77.248090	
8				
9				
10				
11				
12				
13				
14				
15				
16				
17				
18				
19				
20				

Col. stats		A	B	C	D	E	F	G
		Control 0 min	Comp.IV	1	2	3	4	5
		Y	Y	Y	Y	Y	Y	Y
1	Number of values	20	4	6	6	6	6	6
2								
3	Minimum	60.08	79.72	94.41	81.54	27.90	79.39	83.69
4	25% Percentile	79.39	93.01	96.02	91.20	29.50	92.27	93.34
5	Median	104.1	135.1	104.7	120.1	32.70	109.1	116.9
6	75% Percentile	112.9	139.0	139.0	171.1	37.64	126.2	154.5
7	Maximum	152.8	139.5	157.2	199.3	50.93	132.9	179.4
8								
9	Mean	100.0	122.3	115.1	129.5	34.58	108.5	123.3
10	Std. Deviation	27.49	28.55	25.34	45.17	8.272	19.80	36.59
11	Std. Error of Mean	6.147	14.28	10.34	18.44	3.377	8.085	14.94
12								
13	Lower 95% CI of mear	87.13	76.91	88.50	82.13	25.90	87.72	84.86
14	Upper 95% CI of mear	112.9	167.8	141.7	176.9	43.26	129.3	161.7

		H	I	J	K	L	M	N	O	P
		6	7	8	9	10	11	12	13	14
		Y	Y	Y	Y	Y	Y	Y	Y	Y
1	6	6	7	4	7	6	7	7	6	
2										
3	72.96	72.96	96.56	64.22	64.37	98.71	70.81	64.37	68.66	
4	90.66	74.57	100.9	71.41	79.39	101.1	72.96	72.96	87.98	
5	124.5	110.3	104.1	96.33	110.7	110.5	93.01	95.22	109.1	
6	160.5	135.6	121.8	104.6	130.7	113.5	108.5	115.1	122.3	
7	183.8	157.2	128.4	106.3	141.7	115.1	112.9	128.4	124.0	
8										
9	125.8	109.0	109.7	90.79	105.9	108.3	89.78	94.84	104.5	
10	40.49	33.26	12.16	18.53	29.10	6.541	16.87	22.94	21.19	
11	16.53	13.58	4.598	9.264	11.00	2.670	6.376	8.671	8.649	
12										
13	83.35	74.14	98.42	61.31	78.95	101.4	74.18	73.62	82.28	
14	168.3	144.0	120.9	120.3	132.8	115.1	105.4	116.1	126.7	

		Q	R	S	T
		15	16	17	18
		Y	Y	Y	Y
1	7	7	7	7	
2					
3	75.10	66.52	72.96	77.25	
4	75.10	68.66	77.25	79.39	
5	93.01	93.01	110.7	94.41	
6	110.7	112.9	121.8	106.3	
7	117.4	119.6	121.8	108.5	
8					
9	93.83	92.05	100.8	93.12	
10	16.88	22.75	22.48	12.52	
11	6.378	8.601	8.498	4.732	
12					
13	78.23	71.01	80.01	81.54	
14	109.4	113.1	121.6	104.7	



Multiple t tests t tests		Discovery?	P value	Mean1	Mean2	Difference	SE of difference	t ratio	df
1	Comp.IV		0.131781	122.3	100.0	22.3	14.7702	1.5098	460.0
2	1		0.229605	115.1	100.0	15.1	12.5522	1.20297	460.0
3	2		0.0191859	129.5	100.0	29.5	12.5522	2.35018	460.0
4	3	*	2.829362e-007	34.58	100.0	-65.42	12.5522	5.21182	460.0
5	4		0.498638	108.5	100.0	8.5	12.5522	0.677171	460.0
6	5		0.064058	123.3	100.0	23.3	12.5522	1.85624	460.0
7	6		0.0404032	125.8	100.0	25.8	12.5522	2.05541	460.0
8	7		0.473735	109.0	100.0	9.0	12.5522	0.717004	460.0
9	8		0.413162	109.7	100.0	9.7	11.8425	0.819085	460.0
10	9		0.53323	90.79	100.0	-9.21	14.7702	0.623554	460.0
11	10		0.618577	105.9	100.0	5.9	11.8425	0.498207	460.0
12	11		0.508791	108.3	100.0	8.3	12.5522	0.661237	460.0
13	12		0.38859	89.78	100.0	-10.22	11.8425	0.862995	460.0
14	13		0.663244	94.84	100.0	-5.16	11.8425	0.43572	460.0
15	14		0.720132	104.5	100.0	4.5	12.5522	0.358502	460.0
16	15		0.602614	93.83	100.0	-6.17	11.8425	0.521006	460.0
17	16		0.502359	92.05	100.0	-7.95	11.8425	0.671312	460.0
18	17		0.94617	100.8	100.0	0.800003	11.8425	0.0675537	460.0
19	18		0.561553	93.12	100.0	-6.88	11.8425	0.580959	460.0

## Appendix 3

All parallels (designated P) and statistical calculations for the stability activity assay, printed from GraphPad Prism. Activity is given in percentage of the mean of the controls in the initial assay.

Table format: Grouped		Group A	Group B	Group C	Group D	Group E	Group F	Group G	Group H
		Control 10 min	Comp.IV	1	2	3	4	5	6
		Y	Y	Y	Y	Y	Y	Y	Y
1	P1	66.51919	95.219570	33.216130	37.644950	0	53.145810	66.432270	110.720400
2	P2	49.35295	104.077200	24.358500	46.502580	0	24.358500	68.646680	57.574630
3	P3	66.51919	119.833600	33.216130	50.931400	0	55.360220	86.361940	95.219570
4	P4	51.49873	36.140290	37.644950	44.288170	0	68.646680	59.789040	88.576350
5	P5	53.64451		18.668290	30.040930	0	51.498730	79.393880	102.997500
6	P6	27.89515		5.149873	45.061390	0	51.498730	75.102310	126.601000
7	P7	68.64667		27.895150	57.936070	0			
8	P8	66.43226							
9	P9	79.71872							
10	P10	79.71872							
11	P11	95.21958							
12	P12	68.64667							
13	P13	55.36022							
14	P14	64.21785							
15	P15	31.00172							
16	P16	79.71872							

	Group I	Group J	Group K	Group L	Group M	Group N	Group O	Group P	Group Q
	7	8	9	10	11	12	13	14	15
	Y	Y	Y	Y	Y	Y	Y	Y	Y
1	70.861080	33.216130	44.288170	86.361940	55.360220	37.644950	48.716990	73.075490	26.572910
2	64.217860	44.288170	26.572910	55.360220	62.003450	33.216130	62.003450	84.147530	42.073770
3	88.576350	37.644950	46.502580	44.288170	48.716990	39.859360	46.502580	66.432270	60.867850
4	79.718710	35.430540	64.373410	53.145810	62.003450	35.430540	33.216130	57.574630	32.186710
5	79.393880	27.895150	45.061390	30.040930	51.498730	36.478270	23.603580	38.624050	20.594000
6	87.977000	32.186710	49.352950	49.352950	45.061390	25.749370	51.498730	57.936070	25.749370
7				47.207170	51.498730	15.878780	25.749370	47.207170	30.040930
8									
9									
10									
11									
12									
13									
14									
15									
16									

	Group R	Group S	Group T	Group U	Group V
	16	17	18	Data Set-U	Data Set-V
	Y	Y	Y	Y	Y
1	28.787310	37.644950	64.217860		
2	66.432270	36.478270	57.574630		
3	60.867850	51.357250	138.854800		
4	49.352950	23.603580	40.769830		
5	31.001720	26.572910	50.931400		
6	20.151120	17.936710	13.507890		
7	38.624050	13.303840	8.797700		
8	40.769830	36.478270	60.081850		
9					
10					
11					
12					
13					
14					
15					
16					

Col. stats		A	B	C	D	E	F	G
		Control 10 min	Comp.IV	1	2	3	4	5
		Y	Y	Y	Y	Y	Y	Y
1	Number of values	16	4	7	7	7	6	6
2								
3	Minimum	27.90	36.14	5.150	30.04	0.0	24.36	59.79
4	25% Percentile	52.04	50.91	18.67	37.64	0.0	44.71	64.77
5	Median	66.48	99.65	27.90	45.06	0.0	52.32	71.87
6	75% Percentile	76.95	115.9	33.22	50.93	0.0	58.68	81.14
7	Maximum	95.22	119.8	37.64	57.94	0.0	68.65	86.36
8								
9	Mean	62.76	88.82	25.74	44.63	0.0	50.75	72.62
10	Std. Deviation	17.70	36.56	11.06	8.965	0.0	14.46	9.587
11	Std. Error of Mean	4.424	18.28	4.181	3.388	0.0	5.902	3.914
12								
13	Lower 95% CI of mean	53.33	30.64	15.51	36.34	0.0	35.58	62.56
14	Upper 95% CI of mean	72.19	147.0	35.96	52.92	0.0	65.92	82.68

		H	I	J	K	L	M	N	O	P
		6	7	8	9	10	11	12	13	14
		Y	Y	Y	Y	Y	Y	Y	Y	Y
1	6	6	6	6	6	7	7	7	7	7
2										
3	57.57	64.22	27.90	26.57	30.04	45.06	15.88	23.60	38.62	
4	80.83	69.20	31.11	39.86	44.29	48.72	25.75	25.75	47.21	
5	99.11	79.56	34.32	45.78	49.35	51.50	35.43	46.50	57.94	
6	114.7	88.13	39.31	53.11	55.36	62.00	37.64	51.50	73.08	
7	126.6	88.58	44.29	64.37	86.36	62.00	39.86	62.00	84.15	
8										
9	96.95	78.46	35.11	46.03	52.25	53.73	32.04	41.61	60.71	
10	23.36	9.549	5.567	12.09	17.15	6.452	8.425	14.34	15.40	
11	9.537	3.899	2.273	4.934	6.482	2.439	3.184	5.421	5.820	
12										
13	72.43	68.44	29.27	33.34	36.39	47.77	24.24	28.35	46.47	
14	121.5	88.48	40.95	58.71	68.11	59.70	39.83	54.88	74.95	

		Q	R	S	T
		15	16	17	18
		Y	Y	Y	Y
1	7	8	8	8	
2					
3	20.59	20.15	13.30	8.798	
4	25.75	29.34	19.35	20.32	
5	30.04	39.70	31.53	54.25	
6	42.07	57.99	37.35	63.18	
7	60.87	66.43	51.36	138.9	
8					
9	34.01	42.00	30.42	54.34	
10	13.60	16.00	12.39	40.04	
11	5.139	5.658	4.379	14.16	
12					
13	21.44	28.62	20.07	20.87	
14	46.59	55.38	40.78	87.81	

Multiple t tests t tests		Discovery?	P value	Mean1	Mean2	Difference	SE of difference	t ratio	df
1	Comp.IV		0.00804103	88.82	62.76	26.06	9.78252	2.66394	393.0
2	1	*	4.174283e-006	25.74	62.76	-37.02	7.93013	4.66827	393.0
3	2		0.022773	44.63	62.76	-18.13	7.93013	2.28622	393.0
4	3	*	2.564144e-014	0.0	62.76	-62.76	7.93013	7.91411	393.0
5	4		0.152469	50.75	62.76	-12.01	8.37725	1.43364	393.0
6	5		0.239909	72.62	62.76	9.86	8.37725	1.177	393.0
7	6	*	5.426551e-005	96.95	62.76	34.19	8.37725	4.08129	393.0
8	7		0.0616551	78.46	62.76	15.7	8.37725	1.87412	393.0
9	8	*	0.00105292	35.11	62.76	-27.65	8.37725	3.30061	393.0
10	9		0.0465061	46.03	62.76	-16.73	8.37725	1.99708	393.0
11	10		0.185834	52.25	62.76	-10.51	7.93013	1.32532	393.0
12	11		0.255524	53.73	62.76	-9.03	7.93013	1.13869	393.0
13	12	*	0.000125481	32.04	62.76	-30.72	7.93013	3.87383	393.0
14	13		0.00796844	41.61	62.76	-21.15	7.93013	2.66704	393.0
15	14		0.796151	60.71	62.76	-2.05	7.93013	0.258507	393.0
16	15	*	0.000326447	34.01	62.76	-28.75	7.93013	3.62541	393.0
17	16		0.00642998	42.0	62.76	-20.76	7.57751	2.73969	393.0
18	17	*	2.476037e-005	30.42	62.76	-32.34	7.57751	4.26789	393.0
19	18		0.267169	54.34	62.76	-8.42	7.57751	1.11118	393.0

