



Libraries and Learning Services

# University of Auckland Research Repository, ResearchSpace

## Copyright Statement

The digital copy of this thesis is protected by the Copyright Act 1994 (New Zealand).

This thesis may be consulted by you, provided you comply with the provisions of the Act and the following conditions of use:

- Any use you make of these documents or images must be for research or private study purposes only, and you may not make them available to any other person.
- Authors control the copyright of their thesis. You will recognize the author's right to be identified as the author of this thesis, and due acknowledgement will be made to the author where appropriate.
- You will obtain the author's permission before publishing any material from their thesis.

## General copyright and disclaimer

In addition to the above conditions, authors give their consent for the digital copy of their work to be used subject to the conditions specified on the [Library Thesis Consent Form](#) and [Deposit Licence](#).

---

*In silico* approaches in Drug  
Discovery and Development of  
Anticancer Drug Candidates

---

Chatchakorn Eurtivong

A thesis submitted in fulfilment of the requirements for the  
degree of Doctor of Philosophy in Chemistry

School of Chemical Sciences

University of Auckland, 2017

# Abstract

The late 20th century marked a phase shift in drug discovery and development by employing *in silico* tools that offer advantages for medicinal chemists to visualise molecular interactions through simulations of drug candidates and targets that coined the term 'molecular modelling'. Together with advances in development of software, increase in computational power and a wide range of well-characterised protein crystal structures available, it was possible for chemists to virtually screen for active drugs. Moreover, the accessibility of compound libraries gathered saw similarities in drugs and families of compounds that led to the growth of chemoinformatics that utilises the database information available, which can be applied to drug discovery and development projects. This thesis covers the bulk of today's drug discovery and development using *in silico* methods and its applicability mainly to anticancer drug candidates.

Phosphatidylinositol - specific phospholipase C (PI-PLC) is a protein that mediates cell signalling pathways initiated by the binding of an extracellular source/signal to the cell surface. Previous work provided evidences linking it to cancer progression. The 1*H*-pyrazole and 3-amino-thieno[2,3-*b*]pyridine (AThPs) class of compounds were the products of a virtual screen conducted earlier that displayed anticancer activities. Similarity approaches were conducted based on the chemical scaffolds for the two classes which rapidly expanded the structure activity relationship (SAR) resulted in an additional 717 hits. These were docked into the PI-PLC binding pocket, the putative target of the compounds, to further focus the selection. Thirteen derivatives of the AThPs were identified and tested against the NCI60 panel of human tumour cell lines. The most active derivative was potent against the MDA-MB-435 melanoma cell line with 50% growth inhibition (GI<sub>50</sub>) at 30 nM. Also, it was found that a piperidine moiety is tolerated on the AThP scaffold with GI<sub>50</sub> = 296 nM (MDA-MB-

435) considerably expanding the SAR for the series. For the 1*H*-pyrazoles, four derivatives were identified using *in silico* similarity approach and an additionally ten were synthesised with various substituents on the phenyl moiety to extend the SAR but only modest anticancer activity was found. Furthermore, preliminary *in silico* drug design of the AThPs were conducted to selectively target the  $\gamma$ 2-isoform showed dual *o*-methyl and *m*-chloro, and single substituted *m*-methoxy on the naphthalene moiety as reasonable candidates.

A series of AThPs were also prepared and tested in a phenotypic sea urchin embryo assay to identify potent and specific molecules that affect tubulin dynamics. The most active compounds featured a tricyclic core ring system with a fused cycloheptyl or cyclohexyl substituent and unsubstituted or alkyl-substituted phenyl moiety tethered via a carboxamide. Low nano-molar potency was observed in the sea urchin embryos for the most active compounds and was suggestive of a microtubule-destabilising effect. The molecular modelling studies indicated that the tubulin colchicine site is inhibited, which often leads to microtubule-destabilisation in line with the sea urchin embryo results. Finally, the identified hits displayed a robust growth inhibition (GI<sub>50</sub> of 50–250 nM) on multidrug-resistant melanoma MDA-MB-435 and breast MDA-MB-468 human cancer cell lines. This work demonstrated that for the AThPs, the most effective mechanism of action is microtubule-destabilisation initiated by binding to the colchicine pocket.

The next project demonstrated the use of an *in silico* docking software to search for active hit compounds against a drug target. A virtual screen was conducted against phosphatidylcholine - specific phospholipase C derived from *Bacillus cereus* (PC-PLC<sub>Bc</sub>) using the ChemBridge diversity collection of  $5 \times 10^4$  entities. Literature reports have established a relationship between PC-PLC activity and cancer progression. The virtual screen was employed in conjunction with the Amplex Red biochemical assay which identified four different classes of

novel PC-PLC inhibitors: *N*-phenylbenzenesulphonamide, 2,3,4,9-tetrahydro-1*H*-pyrido[3,4-*b*]indole, 2-morpholinobenzoic acids and benzamidobenzoic acids. The 2-morpholinobenzoic acids and 2,3,4,9-tetrahydro-1*H*-pyrido[3,4-*b*]indoles were confirmed for their antiproliferative effects; 50% concentration inhibition (IC<sub>50</sub>) ~ 1 - 2  $\mu$ M in MCF7 and MDA-MB-231 breast cancer cell lines. The *N*-phenylbenzenesulphonamide and benzamidobenzoic acids were shown to display more modest activities.

The last project addressed the applicability of known drugs in chemoinformatics. The Gaussian function was employed to calculate an index that can predict the quality of drug candidates. 1880 known drugs were collected and analysed for their mainstream molecular descriptors: molecular weight (MW), octanol-water partition coefficient (log P), hydrogen bond acceptor (HA) and donor (HD), rotatable bond (RB) and polar surface area (PSA). The statistical distributions were fitted to Gaussian functions for each of the descriptors and normalised to 1. This gave a mathematical tool to calculate a score, or an Index, for each descriptor. Known Drug Indexes (KDI) were derived by summation and multiplication giving one number for each molecule calculated. The KDI summation gives a theoretical maximum of 6 whereas the multiplication method results in 1. Both KDIs are advantageous methods in deriving optimally balanced drug candidates based on the molecular descriptors used with methysergide, amsacrine and fluorometholone being the best clinically used drugs according to both methods.

# Publications

1. **Eurtivong, C.**; Reynisdottir, I.; Kuczma, S.; Furkert, D. P.; Brimble, M. A.; Reynisson, J., Identification of anticancer agents based on the thieno[2,3-*b*]pyridine and 1*H*-pyrazole molecular scaffolds. *Bioorg. Med. Chem.* **2016**, 24 (16), 3521-6.
2. **Eurtivong, C.**; Semenov, V.; Semenova, M.; Konyushkin, L.; Atamanenko, O.; Reynisson, J.; Kiselyov, A., 3-Amino-thieno[2,3-*b*]pyridines as microtubule-destabilising agents: Molecular modelling and biological evaluation in the sea urchin embryo and human cancer cells. *Bioorg. Med. Chem.* **2017**, 25 (2), 658-664.
3. van Rensburg, M.; Leung, E.; Haverkate, N. A.; **Eurtivong, C.**; Pilkington, L. I.; Reynisson, J.; Barker, D., Synthesis and antiproliferative activity of 2-chlorophenyl carboxamide thienopyridines. *Bioorg. Med. Chem. Lett.* **2017**, 27 (2), 135-138.
4. **Eurtivong, C.**; Reynisson, J., The Development of an Index to Identify Quality Candidates for Drug Discovery. *J. Comput. Aided Mol. Des.* (Submitted May **2017**).
5. Bishit, R.; Jaiswal, J. K.; Oliver, V. F.; **Eurtivong, C.**; Reynisson, J.; Rupenthal, I. D., Preparation and evaluation of PLGA nanoparticle-loaded biodegradable light-responsive injectable implants as a promising platform for intravitreal drug delivery. *J. Drug Deliv. Sci. Technol.* **2017**, 40, 142-156.
6. Plažuk, D.; Wieczorek, A.; Ciszewski, W. M.; Kowalczyk, K.; Błaż, A.; Pawłędzio, S.; Makal, A.; **Eurtivong, C.**; Arabshahi, H. J.; Reynisson, J.; Hartinger, C. G.; Rychlik, B., A decoration of the side-chain of paclitaxel by ferrocenyl moiety and its influence on the cytotoxic activity of taxanes. *J. Med. Chem.* (Submitted May **2017**).
7. Zakharenko, A.; Luzina, O.; Sokolov, D.; Zakharova, O.; Dyrkheeva, N.; Chepanova, A.; Klabenkova, K.; Chelobanov, B.; Stetsenko, D.; **Eurtivong, C.**; Reynisson, J.; Volcho, K.; Salakhutdinov, N.; Lavrik, O., Synthesis and evaluation of benzylidenfuranone derivatives of usnic acid as novel Tdp1 inhibitors. *Bioorg. Med. Chem.* (Submitted May **2017**).

# Acknowledgements

First and foremost, I would like to thank my main supervisor, Dr. Jóhannes Reynisson for his efforts to train, advise and support me over the past four years. He often gives advices not just in the projects but also in our careers and life. Personally, I feel these are very important elements to absorb along the way to become successful. I was very fortunate to have him as my supervisor, a very patient, respectful, funny and considerate person.

Secondly, I would like to thank my co-supervisors, Dr. Euphemia Leung and Dr. Vijayalekshmi Sarojini for giving me support and advice. I was very fortunate to have Dr. Euphemia Leung who conducted various experimental tests for our compounds against PC-PLC and giving useful advice for my future career. Also, credit to Dr. Chris S. Xu for his involvement in screening compounds for PC-PLC activity. I would like to thank Distinguished Prof. Margaret Brimble and her group for their synthetic studies on the 1*H*-pyrazole derivatives and Prof. Inga Reynisdóttir for the MTT assay support on the PI-PLC project. I would like to thank our Russian collaborators; Prof. Alex Kiselyov, Prof. Victor Semenov, Dr. Marina Semenova, Leonid Konyushkin and Olga Atamanenko for their biological studies in sea urchin embryos on the AThP actives.

To my colleagues, John, Daniel, Ayesha and Krunal thanks for all the support in my work and the food you guys always provide in the office. A big thanks to John, who always helped me with the technical issues in software and computers. Will definitely miss the jokes and good times I spent together with you guys. We all know Krunal will be on the Forbes magazine one day and spend his riches, and the thing that I would definitely miss is the smell and orange colour of Daniel's chicken stew on rice that he brought in everyday, but deeply I hope he would seriously consider changing his recipe one day.

To my family in Thailand and friends, they have been the base of my support. Especially to my lovely parents, a big warm thank you to them, whom over the years supported me financially and gave me the strength when I needed. I am forever grateful. I love you very much.



# Table of Contents

Abstract.....	ii
Publications.....	v
Acknowledgements.....	vi
Table of Contents.....	viii
Abbreviations.....	xi
<i>Chapter 1</i> .....	1
1.1 Ancient to Modern Drug Discovery.....	2
1.2 <i>In vitro</i> drug screening .....	3
1.3 Structure-based Virtual High Throughput Screening.....	4
1.4 Genetic Optimisation for Ligand Docking.....	7
1.4.1 GoldScore .....	8
1.4.2 ChemScore.....	8
1.4.3 ChemPLP.....	9
1.4.4 Astex Statistical Potential.....	10
1.5 Similarity searching.....	11
1.6 Drug-likeness models.....	12
1.7 Chemical Space .....	13
<i>Chapter 2</i> .....	14
2.1 Introduction .....	15
2.1.1 Overview of phosphatidylinositol - specific phospholipase C .....	15
2.1.2 PLC- $\gamma$ .....	17
2.1.3 Catalytic site of PI-PLC- $\delta$ 1 and mechanism of cleavage .....	18
2.1.3 PLC inhibitors.....	20
2.1.4 Research Prologue .....	20
2.2 Results and Discussion.....	22
2.2.1 SAR of 3-amino-thieno[2,3- <i>b</i> ]pyridine (AThP).....	23
2.2.2 Docking AThPs against PLC- $\delta$ 1 .....	27
2.2.3 Molecular Modelling for PLC- $\gamma$ 2 selectivity.....	29
2.2.4 SAR of 1 <i>H</i> -pyrazoles.....	32
2.2.5 Docking 1 <i>H</i> -pyrazoles against PLC- $\delta$ 1 .....	34
2.3 Methodology .....	36
2.3.1 Similarity Search, Molecular modelling and CADD.....	36

2.3.1 Synthetic studies, NCI60 assay and MTT assay.....	37
2.4 Future Work and Conclusions.....	37
<i>Chapter 3</i> .....	39
3.1 Introduction .....	40
3.1.1 Microtubule structure and function .....	40
3.1.2 Microtubule dynamics .....	40
3.1.3 Microtubule-stabilising agents .....	43
3.1.4 Microtubule-destabilising agents.....	44
3.1.5 Sea urchin embryos as model organisms.....	46
3.2 Results and Discussion.....	46
3.2.2 Biological studies .....	47
3.2.3 Molecular Modelling .....	54
3.3 Methodology .....	63
3.3.1 Molecular Modelling .....	63
3.3.2 NCI60 assay and Phenotypic sea urchin embryo assay.....	64
3.4 Future Work and Conclusions.....	64
<i>Chapter 4</i> .....	65
4.1 Introduction .....	66
4.1.1 Overview of phosphatidylcholine – specific phospholipase C.....	66
4.1.2 PC-PLC in cancer progression .....	67
4.1.3 Binding site of PC-PLC <sub>Bc</sub> .....	68
4.1.4 Catalytic cleavage mechanism of PC-PLC <sub>Bc</sub> .....	69
4.2 Results and Discussion.....	70
4.2.1 Pilot Screen.....	70
4.2.2 Main Screen.....	85
4.3 Methodology .....	103
4.3.1 Molecular Modelling and Virtual Screening.....	103
4.3.2 Verification of hits.....	103
4.4 Future Works and Conclusion.....	104
<i>Chapter 5</i> .....	105
5.1 Introduction .....	106
5.1.1 Known Drug Index .....	106
5.1.2 Molecular Descriptors .....	107
5.1.3 Gaussian Distribution .....	107
5.1.4 Qikprop.....	108

5.2 Results .....	109
5.2.1 Distribution of Physicochemical Properties .....	109
5.2.2 Calculation of Known Drug Index .....	115
5.2.3 Robustness of the KDI <sub>7a/b</sub> indexes .....	119
5.3 Discussion .....	121
5.4 Methodology .....	123
5.5 Future Works and Conclusion.....	123
<i>Chapter 6</i> .....	125
Appendix.....	128
Bibliography .....	133

# Abbreviations

<b>ACD</b>	Available Chemical Directory
<b>Ala</b>	Alanine
<b>AMH</b>	Australian Medicines Handbook
<b>Arg</b>	Arginine
<b>Asn</b>	Asparagine
<b>ASP</b>	Astex Statistical Potential
<b>Asp</b>	Aspartic acid
<b>AThP</b>	3-amino-thieno[2,3- <i>b</i> ]pyridine
<b>ATOX1</b>	Antioxidant Protein 1
<b>CADD</b>	Computer-aided Drug Design
<b>CFDA</b>	China Food and Drug Administration
<b>CCS</b>	Copper chaperone for superoxide dismutase
<b>ChemDiv</b>	Chemical Diversity
<b>CMC</b>	Comprehensive Medicinal Chemistry
<b>CS</b>	ChemScore
<b>DAG</b>	Diacylglycerol
<b>DFT</b>	Density functional theory
<b>EC</b>	Effective threshold concentration
<b>FDA</b>	US Food and Drug Administration
<b>GDP</b>	Guanosine diphosphate
<b>GI<sub>50</sub></b>	50% Growth Inhibition
<b>Glu</b>	Glutamic acid
<b>Gly</b>	Glycine
<b>GOLD</b>	Genetic optimisation for Ligand Docking
<b>GPCR</b>	G protein-coupled receptor
<b>GS</b>	GoldScore
<b>GTP</b>	Guanosine-5'-triphosphate
<b>HA</b>	Hydrogen bond acceptor
<b>HB</b>	Hydrogen bond
<b>HD</b>	Hydrogen bond donor

<b>His</b>	Histidine
<b>HTS</b>	High Throughput Screen
<b>IBS</b>	InterBioscreen
<b>IC<sub>50</sub></b>	50% Concentration Inhibition
<b>IP<sub>3</sub></b>	Inositol triphosphate
<b>KDI</b>	Known Drug Index
<b>KDS</b>	Known Drug Space
<b>Leu</b>	Leucine
<b>Log P</b>	Octanol-water partition coefficient
<b>Lys</b>	Lysine
<b>MDDR</b>	MACCS-II Drug Data Report
<b>MHRA</b>	UK Medicines and Healthcare products Regulatory Agency
<b>MTT</b>	3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide
<b>MW</b>	Molecular weight
<b>NADPH</b>	Nicotinamide adenine dinucleotide phosphate hydrogen
<b>NCI60</b>	National Cancer Institute 60 cell panel
<b>NMR</b>	Nuclear magnetic resonance
<b>OPLS-AA</b>	Optimised Potentials for Liquid Simulation force field – all atom
<b>PC</b>	Phosphatidylcholine
<b>PC-PLC</b>	Phosphatidylcholine - specific phospholipase C
<b>PC-PLC<sub>Bc</sub></b>	Phosphatidylcholine - specific phospholipase C derived from <i>Bacillus cereus</i>
<b>PDB</b>	Protein Data Bank
<b>Phe</b>	Phenylalanine
<b>PLP</b>	Piecewise Linear Potential
<b>PI-PLC</b>	Phosphatidylinositol - specific phospholipase C
<b>PIP<sub>2</sub></b>	Phosphatidylinositol 4,5-bisphosphate
<b>PKC</b>	Protein kinase C
<b>PMDA</b>	Japan's Pharmaceuticals and Medical Devices Agency
<b>PSA</b>	Polar surface area
<b>RB</b>	Rotatable bond
<b>RMSD</b>	Root-mean squared deviation
<b>RSS</b>	Residual sum of squares
<b>RTK</b>	Receptor tyrosine kinase
<b>Ser</b>	Serine
<b>SAR</b>	Structure Activity Relationship

<b>Trp</b>	Tryptophan
<b>Tyr</b>	Tyrosine
<b>Val</b>	Valine
<b>vHTS</b>	Virtual High Throughput Screen

## Co-Authorship Form

This form is to accompany the submission of any PhD that contains published or unpublished co-authored work. **Please include one copy of this form for each co-authored work.** Completed forms should be included in all copies of your thesis submitted for examination and library deposit (including digital deposit), following your thesis Acknowledgements. Co-authored works may be included in a thesis if the candidate has written all or the majority of the text and had their contribution confirmed by all co-authors as not less than 65%.

Please indicate the chapter/section/pages of this thesis that are extracted from a co-authored work and give the title and publication details or details of submission of the co-authored work.

Chapter 2

Nature of contribution  
by PhD candidate

Similarity search, molecular modelling and writing up manuscript

Extent of contribution  
by PhD candidate (%)

70%

### CO-AUTHORS

Name	Nature of Contribution
Margaret A. Brimble	Synthetic studies
Daniel P. Furkert	Synthetic studies
Stephanie Kuczma	Synthetic studies
Inga Reynisdóttir	MTT assay
Jóhannes Reynisson	Supervision of PhD candidate

### Certification by Co-Authors

The undersigned hereby certify that:

- ❖ the above statement correctly reflects the nature and extent of the PhD candidate's contribution to this work, and the nature of the contribution of each of the co-authors; and
- ❖ that the candidate wrote all or the majority of the text.

Name	Signature	Date
Margaret A. Brimble		20/4/17
Daniel P. Furkert		28/4/17
Stephanie Kuczma		28/4/2017
Inga Reynisdóttir		21/5/2017
Jóhannes Reynisson		19/4 2017



## Co-Authorship Form

This form is to accompany the submission of any PhD that contains published or unpublished co-authored work. **Please include one copy of this form for each co-authored work.** Completed forms should be included in all copies of your thesis submitted for examination and library deposit (including digital deposit), following your thesis Acknowledgements. Co-authored works may be included in a thesis if the candidate has written all or the majority of the text and had their contribution confirmed by all co-authors as not less than 65%.

Please indicate the chapter/section/pages of this thesis that are extracted from a co-authored work and give the title and publication details or details of submission of the co-authored work.

Chapter 3

Nature of contribution  
by PhD candidate

Molecular modelling and writing up manuscript

Extent of contribution  
by PhD candidate (%)

70

### CO-AUTHORS

Name	Nature of Contribution
Victor Semenov	Biological studies
Marina Semenova	Biological studies
Leonid Konyushkin	Biological studies
Olga Atamanenko	Biological studies
Alex Kiselyov	Final reviewer for manuscript
Jóhannes Reynisson	Supervision of PhD Candidate

### Certification by Co-Authors

The undersigned hereby certify that:

- ❖ the above statement correctly reflects the nature and extent of the PhD candidate's contribution to this work, and the nature of the contribution of each of the co-authors; and
- ❖ that the candidate wrote all or the majority of the text.

Name	Signature	Date
Victor Semenov		19/4/2017
Marina Semenova		19/4/2017
Leonid Konyushkin		19/4/2017
Olga Atamanenko		19/4/2017
Alex Kiselyov		19/4/2017
Jóhannes Reynisson		19/4 2017



## Co-Authorship Form

Graduate Centre  
 The ClockTower – East Wing  
 22 Princes Street, Auckland  
 Phone: +64 9 373 7599 ext 81321  
 Fax: +64 9 373 7610  
 Email: [postgraduate@auckland.ac.nz](mailto:postgraduate@auckland.ac.nz)  
[www.postgrad.auckland.ac.nz](http://www.postgrad.auckland.ac.nz)

This form is to accompany the submission of any PhD that contains published or unpublished co-authored work. **Please include one copy of this form for each co-authored work.** Completed forms should be included in all copies of your thesis submitted for examination and library deposit (including digital deposit), following your thesis Acknowledgements. Co-authored works may be included in a thesis if the candidate has written all or the majority of the text and had their contribution confirmed by all co-authors as not less than 65%.

Please indicate the chapter/section/pages of this thesis that are extracted from a co-authored work and give the title and publication details or details of submission of the co-authored work.

Chapter 4

Nature of contribution by PhD candidate	Virtual Screening, Molecular modelling, similarity search, write up manuscript
Extent of contribution by PhD candidate (%)	70


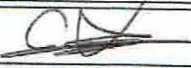
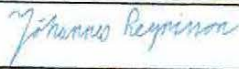
### CO-AUTHORS

Name	Nature of Contribution
Euphemia Leung	Biochemical assay and Biological studies, co-supervisor
Chris Sun Xu	Biological screening
Jóhannes Reynisson	Supervision of PhD Candidate

### Certification by Co-Authors

The undersigned hereby certify that:

- ❖ the above statement correctly reflects the nature and extent of the PhD candidate's contribution to this work, and the nature of the contribution of each of the co-authors; and
- ❖ that the candidate wrote all or the majority of the text.

Name	Signature	Date
Euphemia Leung		19/04/2017
Chris Sun Xu		20/4/2017
Jóhannes Reynisson		19/4 /2017

## Co-Authorship Form

Graduate Centre  
The Clock Tower – East Wing  
22 Princes Street, Auckland  
Phone: +64 9 373 7599 ext 81321  
Fax: +64 9 373 7610  
Email: [postgraduate@auckland.ac.nz](mailto:postgraduate@auckland.ac.nz)  
[www.postgrad.auckland.ac.nz](http://www.postgrad.auckland.ac.nz)

This form is to accompany the submission of any PhD that contains published or unpublished co-authored work. **Please include one copy of this form for each co-authored work.** Completed forms should be included in all copies of your thesis submitted for examination and library deposit (including digital deposit), following your thesis Acknowledgements. Co-authored works may be included in a thesis if the candidate has written all or the majority of the text and had their contribution confirmed by all co-authors as not less than 65%.

Please indicate the chapter/section/pages of this thesis that are extracted from a co-authored work and give the title and publication details or details of submission of the co-authored work.

Chapter 5

Nature of contribution  
by PhD candidate

Data collection and evaluation, write up manuscript

Extent of contribution  
by PhD candidate (%)

80


### CO-AUTHORS

Name	Nature of Contribution
Jóhannes Reynisson	Supervision of PhD Candidate and final review of manuscript

### Certification by Co-Authors

The undersigned hereby certify that:

- ❖ the above statement correctly reflects the nature and extent of the PhD candidate's contribution to this work, and the nature of the contribution of each of the co-authors; and
- ❖ that the candidate wrote all or the majority of the text.

Name	Signature	Date
Jóhannes Reynisson		19/4 2017

# *Chapter 1*

The Evolution of Drug Discovery: From  
Ancient times to *In silico* approaches

# 1.1 Ancient to Modern Drug Discovery

Since ancient times, the concept of alleviating disease symptoms was accomplished by consumption or applying substances to the body, mainly from plants, plant-derived mixtures or extracts.<sup>1</sup> The methods in discovering these ‘forms’ that have various medicinal properties were unclear but are safe to assume serendipity, empirical observations and trial and error were the main methods considering the lack of technology.<sup>1</sup> For example, the discovery of cinchona plant as an effective treatment for malaria was discovered by observing the Quechua people living in Peru chewing the bark portion to alleviate malarial symptoms.<sup>1-2</sup> The cinchona bark was soon introduced to Europe in 1638 followed by the isolation of the bioactive quinine alkaloid in 1820 which became the standard treatment for malaria until 2006.<sup>1-2</sup>

It was not until the late 19<sup>th</sup> and early 20<sup>th</sup> century that the basis of chemotherapy was established, many regarded to be the birth of modern drug discovery.<sup>1, 3</sup> Paul Ehrlich, a German physician and scientist, postulated the concept of chemical-receptor interactions that “chemoreceptors” influences the interactions between chemicals with living cells, *i.e.*, chemoreceptors of foreign invading organisms and cancer cells have stronger affinities for certain chemicals than normal cells.<sup>1, 3</sup> This concept stemmed from Ehrlich’s observation that different dyes stained various biological tissues at different intensities.<sup>1</sup> Ehrlich further postulated that chemicals can selectively target a disease-related organism that if applied could provide a “magic bullet” killing only that particular organism.<sup>1, 3</sup> One of the first evidence of the “magic bullet” theory was when Ehrlich cured two malaria-infected patients with methylene blue.<sup>4</sup> The fact that the patients stayed alive demonstrated selective properties of the dye against malarial parasites. The first ever recorded drug screen soon followed when

Ehrlich and colleagues screened hundreds of commercially available dyes against mice infected with *Trypanosoma equinum*, the parasite responsible for sleeping sickness that leads to the identification of Trypan red as a potent candidate.<sup>1, 4</sup> Unfortunately, resistant strains emerged as hindrance to its clinical success.<sup>1, 4</sup> Nevertheless, the approach was acclaimed to be the first attempt to discern the relationship between molecular structures and biological activity.<sup>1</sup> Ehrlich soon turned his interests to screen arsenic-based drugs against *Treponema pallidum*, a spirochete bacterium responsible for syphilis.<sup>1, 4</sup> The approach stemmed from the knowledge that atoxyl, an arsenic-based drug used to treat sleeping sickness in Africa is also active against spirochetes.<sup>1, 4</sup> Taking the advantage of chemical modifications as a way to expand the SAR, Ehrlich and coworkers conducted synthetic studies to provide derivatives which were screened against rabbits infected with syphilis.<sup>4</sup> As a result, the potent candidate, salvarsan entered clinical trials in 1909 and was later introduced to the clinic as an indication for syphilis.<sup>4</sup>

## 1.2 *In vitro* drug screening

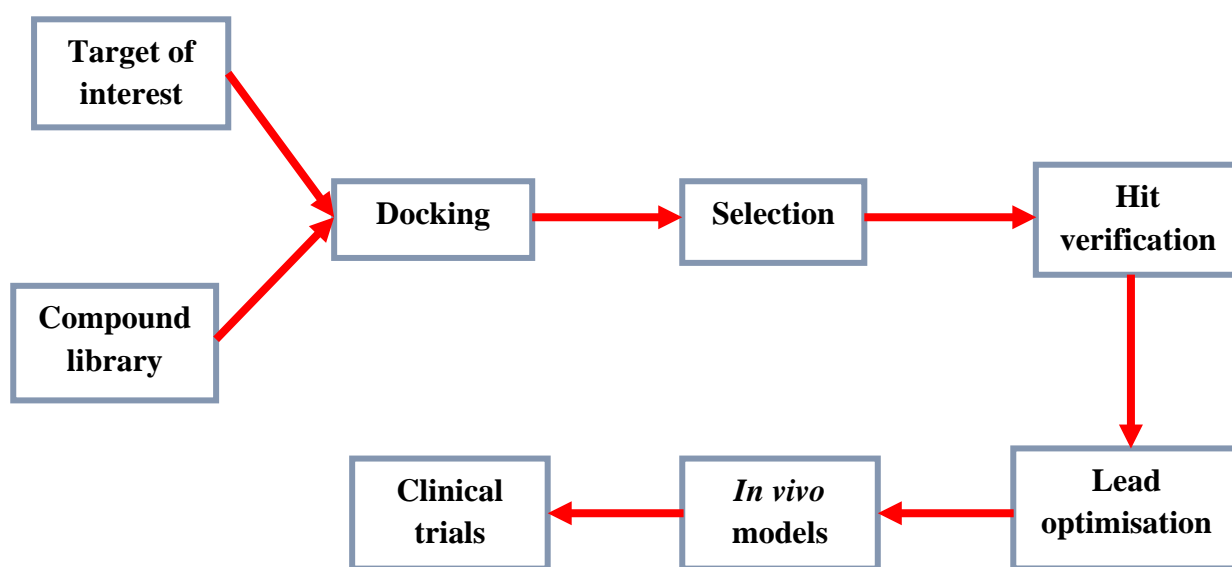
Today, the basic foundation in modern drug screening introduced by Ehrlich is retained with the only difference being the use of advanced technology and strategies employed. The 20<sup>th</sup> century marked a period that phenotypic screening was used for screening drugs in which changes in phenotypic expressions were studied for biological effects under both *in vivo* and *in vitro* settings. Many drugs were discovered through these approaches including various anticancer drugs without prior knowledge of their biomolecular targets, *e.g.*, vorinostat, pomalidomide, romidepsin and lenalidomide.<sup>5</sup> The setback of using animals or *in vivo* models as test subjects became increasingly controversial with issues relating to ethical practices<sup>6-7</sup> leading to the development of *in vitro* cell-based models as the preferential choice prior to

animal testing, *e.g.*, the NCI60 human tumour cell line that was developed in the late 1980s which consists of 60 human-derived tumour cell lines used for screening potential anticancer drug candidates.<sup>8</sup> However, a disadvantage of employing cell-based assays is that it does not provide the understandings at the molecular level that underlies the mechanisms of action, *i.e.*, it is unable to identify specific drug targets. It was not until the late 20<sup>th</sup> century that clearer understandings of molecular drug interactions with proteins were more accessible owing to advances in computational chemistry and biotechnology, *e.g.*, proteins became much easier to isolate and purify with improved assay techniques. Additionally, drug screening processes became much more feasible through advances in robotics that allow large numbers of pre-plated compounds to be screened for biological activity, a process known today as high throughput screening (HTS).

## 1.3 Structure-based Virtual High Throughput Screening

Over the past 30 years, HTS was regarded to be the keystone in drug discovery. Given the rapid progression of automation and robotics technology, the ease in screening compounds for biological activity was achieved. However, HTS is limited by its high cost and inability to unveil information involving mechanistic interactions between the active compound and its target. Molecular docking offers a solution that allows visualisation of predictive molecular interactions between compounds and protein targets that can elucidate binding mechanisms. The boom of the information technology industry in the late 1990s foresaw promising advances in development of computational power, sophisticated docking algorithms and software that saw through *in silico* tools as contributors to the next paradigm shift in drug discovery. In conjunction with collections of chemical and genetic databases of various drug

targets such as the Protein Databank (PDB),<sup>9-10</sup> it is feasible to simulate molecular interactions of drug targets with millions of compounds within short period of time, and to predict for the bioactivities of compounds against specific drug targets, a process known as structure-based virtual High Throughput Screening (vHTS). The basic drug discovery and development protocol using vHTS are shown in Fig. 1.1.



**Figure 1.1.** The basic flowchart of drug discovery and development protocol using the structure-based virtual High Throughput Screen.

In a vHTS, well-defined structures of drug targets are essential. Usually, these pre-defined structures were determined by x-ray crystallography,<sup>11</sup> NMR studies<sup>12</sup> and homology modelling,<sup>13</sup> which are publicly accessible, *e.g.*, the PDB.<sup>9-10</sup> Secondly, ligand collections are required, usually available from compound databases pre-designed based on its synthetic accessibility and their drug-like properties.<sup>14</sup> Together with a computer and well-established molecular docking software packages, a vHTS can take effect. One of the main objectives of a molecular docking software is to predict the binding modes and free binding energies ( $\Delta G_{\text{binding}}$ ) of ligands to a target. The software employs different search algorithms which



perform extensive pose searches that defines complementary fittings between ligands and the target.

Often, predictive binding energies are represented as scores which are ranked, *i.e.*, top-ranking scores represent compounds with most favourable affinities for the protein target. The top-scoring compounds are checked for their chemical stability, toxicity and plausible binding modes before a final selection of compounds or ‘virtual hits’ are obtained for experimental testing. Some of the ‘undesirable’ moieties are shown in Fig. 1.2. At this stage, the “virtual hits” are hypothesised to possess bioactivities. Furthermore, non-specific or promiscuous molecules should also be considered that typically form aggregates in assays. These aggregators tend to have higher log P values, lower solubility and are more sensitive to detergents than deaggregate compounds.<sup>15</sup> Thus, detergent-based assays have been suggested as a way to test for compounds with promiscuous properties.<sup>16</sup>

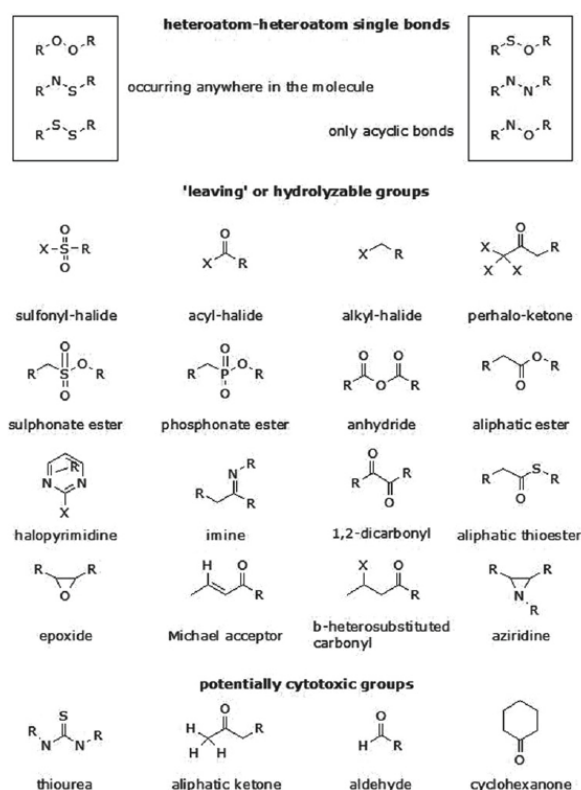


Figure 1.2. Chemical substructures that may cause interference with biochemical assays. (Diagram obtained from Bologa *et al.*<sup>17</sup>)



To identify “real hits”, *in vitro* testings are required, initially through biochemical assays to verify affinities against specific protein targets followed by cell-based testing. Potent hits are taken forward as lead compounds for further development. The early part of development depends heavily on synthetic and *in silico* studies to expand the SAR. Quite often, molecular modelling and organic synthesis are used in conjoint to rationalise drug-design. Toxicity and selectivity issues must also be addressed, *e.g.*, testing drug candidates using *in vivo* models and against a panel of putative biomolecular targets. This is especially important for cancer chemotherapy in which its success is much hindered by undesirable cytotoxic side effects. Drug candidates are then further proceeded towards clinical trials.

## 1.4 Genetic Optimisation for Ligand Docking

The Genetic Optimisation for Ligand Docking (GOLD) is an automated docking program developed in the late 1990s that is based on a genetic algorithm, a set of code to solve an optimisation problem based on genetic information passed on by natural selection.<sup>18-19</sup> GOLD is the docking software package used in the projects described in this thesis. It performs automated docking with full ligand flexibility in combination with partial protein flexibility within the protein binding site.<sup>19</sup> Partial flexibility refers to the rotatable bonds of polar atoms in certain side chains, *i.e.*, only OH on serine, threonine, and tyrosine and NH<sub>3</sub><sup>+</sup> on lysine.<sup>20</sup> This offered an advantage over softwares such as DOCK that only permitted rigid protein-ligand docking previously before its transition into flexible protein-ligand docking.<sup>21</sup> Furthermore, the docking duration is short as only the binding site of the protein is flexible, *i.e.*, much longer duration is required for the whole protein to search for all the possible conformations. The software package comes with four different scoring functions to assess

the fitness of each binding pose: GoldScore (GS),<sup>18-19</sup> ChemScore (CS),<sup>22-23</sup> Piecewise Linear Potential ChemPLP<sup>22, 24</sup> and Astex Statistical Potential (ASP).<sup>25</sup>

### 1.4.1 GoldScore

GS comprises of four main terms: protein-ligand hydrogen bonding energy, protein-ligand van der Waal's energy, ligand internal van der Waal's energy and ligand torsional straining.<sup>18-19</sup> The four terms are calculated based on molecular mechanics.<sup>18-19</sup> The fitness score is determined by summing all of the energy components.<sup>19</sup> GS operates by allowing poor intermolecular interactions at the beginning of the docking run to allow for better poses to evolve in the latter.<sup>20</sup>

### 1.4.2 ChemScore

CS is an empirical-based scoring function designed to approximate  $\Delta G_{\text{binding}}$  that is estimated using the following equation:

$$\Delta G_{\text{binding}} = \Delta G_0 + \Delta G_{\text{hbond}} + \Delta G_{\text{metal}} + \Delta G_{\text{lipo}} + \Delta G_{\text{rot}} \quad (1)$$

The scoring function was derived by fitting a training set of 82 ligand-protein complexes with known binding affinities against their predicted counterparts.<sup>22</sup> The function includes terms that express the predicted binding affinity at the y-intercept ( $\Delta G_0$ ), hydrogen bonding ( $\Delta G_{\text{hbond}}$ ), metal-ligand interactions ( $\Delta G_{\text{metal}}$ ), lipophilicity ( $\Delta G_{\text{lipo}}$ ), and bond flexibility ( $\Delta G_{\text{rot}}$ ).<sup>22-23</sup> Equation 1 can be rewritten as:

$$\Delta G_{\text{binding}} = \gamma_0 + \gamma_{\text{hbond}} P_{\text{hbond}} + \gamma_{\text{metal}} P_{\text{metal}} + \gamma_{\text{lipo}} P_{\text{lipo}} + \gamma_{\text{rot}} P_{\text{rot}} \quad (2)$$

Included in these terms are regression coefficients ( $\gamma$ ). The  $P$  terms are functions that estimate the magnitudes of each type of interaction for a given pose.<sup>20</sup> The  $\gamma$  coefficients are

multiplied by  $P$  terms to give the free energy contribution for a particular binding pose.<sup>20</sup> Additionally to  $\Delta G_{\text{binding}}$ , clash interactions and torsional strains are taken into account to compute the final fitness score.<sup>20, 23</sup>

### 1.4.3 ChemPLP

The ChemPLP stemmed from incorporating angle-dependent hydrogen bonding and ligand-metal interaction terms from ChemScore into the piecewise linear potential (PLP) scoring function.<sup>22, 26</sup> ChemPLP is an empirical-based scoring function used in GOLD and measures the molecular attractions and repulsions between heavy atoms (non-hydrogens) in proteins and ligands.<sup>22</sup> The types of molecular attractions accounted in PLP are steric and hydrogen bond interactions.<sup>20, 22</sup> A few examples of the atom types are given in Table 1.1: Secondary amines as hydrogen bond donors, oxygen on ethoxy as acceptors, hydroxyl groups function as both donors and acceptors and carbon are non-polar atoms.<sup>26</sup>

**Table 1.1.** Pairwise atomic interaction types used in the molecular recognition model.

Ligand atom type	Protein atom type			
-	Donor	Acceptor	Donor and Acceptor	Non-polar
Donor	Steric	Hydrogen bond	Hydrogen bond	Steric
Acceptor	Hydrogen bond	Steric	Hydrogen bond	Steric
Donor and Acceptor	Hydrogen bond	Hydrogen bond	Hydrogen bond	Steric
Non-polar	Steric	Steric	Steric	Steric

Internal ligand torsion and clash terms are included to account for any intramolecular interactions for each pose.

### 1.4.4 Astex Statistical Potential

The Astex Statistical Potential (ASP) is a knowledge-based scoring function that incorporated statistical information relating to the frequencies of various atom-atom interactions between proteins and ligands from the PDB which contain numerous structures of protein-ligand complexes.<sup>25</sup> Statistical potentials are calculated based on protein-ligand complexes in the PDB that includes 9209 ligands in 5839 proteins.<sup>25</sup> Calculating for the ASP fitness is determined by equation 3 shown below.

$$ASP\ Fitness = -C_s \sum_p \sum_l StatScore(p, l, r_{pl}) - E_{int} - E_{clash} \quad (3)$$

The ‘StatScore’ is the summation over all combinations of protein atoms ( $p$ ) and ligand atoms ( $l$ ) within 6.0 Å from PDB database and the distance between  $p$  and  $l$  denoted by  $r_{pl}$ .<sup>25</sup>  $C_s$  is the scaling factor, and the clash and internal energy terms from ChemScore are also included.<sup>25</sup> The clash term prevents proteins and ligands from clashing and the internal energy term account for intramolecular clashes and torsion forces for reasonable ligand geometries during docking.<sup>25</sup> It is worth noting that there are absences of chemical interaction terms in ASP such as hydrogen bonding due to the nature of knowledge-based scoring functions that measures statistical potentials rather than binding affinities, *i.e.*, ASP is more dependent on the quality of the protein-ligand complex collection in PDB than the other scoring functions.

## 1.5 Similarity searching

The concept of similarity search is based on the idea that structurally similar molecules tend to have similar properties and activities.<sup>27-28</sup> In medicinal chemistry, once an active hit is confirmed, the strategy can be implemented to search for structurally similar molecules.<sup>27-28</sup> The molecules resulting from the search are expected to exert a similar mode of action.<sup>27-28</sup> Given the luxury and ease of accessibility to chemical databases today, the SAR of an active hit can be rapidly derived using similarity methods. Perhaps the most common similarity method in use today is based on 2D fingerprinting.<sup>27-28</sup> The method employs sequences of bits with  $a$  bits set to “0” indicating a missing feature in molecule A or “1” to indicate the presence of a particular feature.<sup>27-28</sup> The same is applied to molecule B using  $b$  bits and any common feature between molecule A and B is represented by  $c$  bits.. Several similarity coefficients have been proposed for quantification, the most widely used is the Tanimoto coefficient (4).<sup>27</sup>

$$S_{AB} = \frac{c}{a + b - c} \quad (4)$$

The similarity ( $S_{AB}$ ) ranges between 0 and 1 in which  $S_{AB} = 0$  indicate no bits in common between two molecules, whereas  $S_{AB} = 1$  indicates identical fingerprinting representation but not necessarily the same molecule.

## 1.6 Drug-likeness models

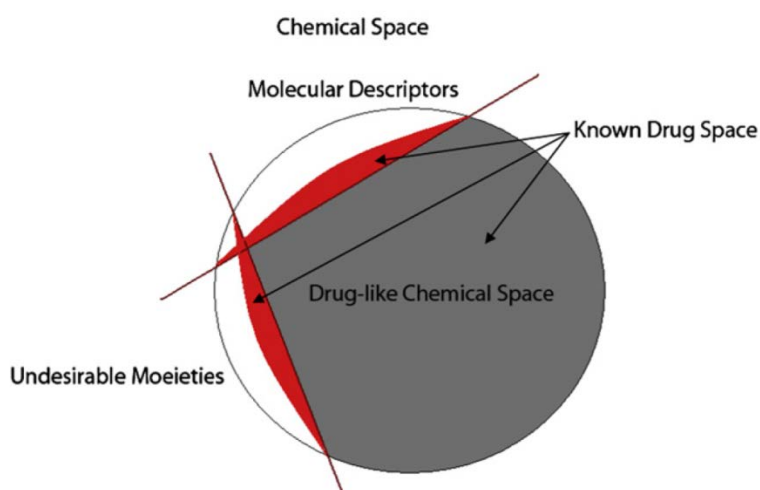
Combinatorial chemistry and the use of HTS became one of the key strategies to synthesise various compounds and to screen for their activities in a short period of time. Yet, significant small numbers of hits are generated and even fewer enter clinical trials compared to the number of compounds screened initially. Thus, the issue remained as how to distinguish drugs from non-drugs? Several approaches have been developed to address this issue. Perhaps the most famous is the “Lipinski’s Rule of Five” introduced in the late 1990s that described the distribution of physicochemical properties of orally active phase II drug candidates and was therefore assumed to be an indicator for good drug absorption.<sup>29</sup> The rule was widely implemented with numerous successes in the pharmaceutical industry,<sup>30</sup> which suggested a good permeation and absorption if a molecule has:

- $MW \leq 500 \text{ g mol}^{-1}$
- $\text{Log } P_{\text{octanol/water}} \leq 5$
- $\text{HDs} \leq 5$
- $\text{HAs} \leq 10$

Other attempts have also been made to define drugs. These included additional physicochemical properties introduced by Veber *et al.*<sup>31</sup> such as molecular flexibility ( $RB \leq 10$ ) and PSA that has  $140 \text{ \AA}^2$  or less,<sup>31</sup> density functional theory-derived descriptors,<sup>32</sup> development of molecular frameworks<sup>33</sup> and side chains,<sup>34</sup> integration of structural filters,<sup>35-36</sup> and development of drug-like indexes.<sup>37-39</sup> Evidently, efforts have been made to guide medicinal chemists more towards developing drug-like molecules. This left chemists with the notion to design improved drug candidates and compound libraries for more efficient screens, *e.g.*, removal of compounds with toxic moieties from compound libraries (see Fig. 1.2).

## 1.7 Chemical Space

It has been widely accepted that small molecule drugs exert their effect by binding to a biological target. The number of molecules that are capable of ‘triggering’ an effect is vastly innumerable raising the question of how many of these molecules are there and how many can be developed into drugs? This prompted the concept of chemical space to distinguish different properties of chemicals (see Fig. 1.3). Within the chemical space includes a drug-like chemical space which consists of all properties of drug-like molecules. These molecules have characteristics and features similar to drugs. However, they may contain features that hinder them to clinical success, *e.g.*, weak potency. Thus, other parameters are often compromised in known drugs or ‘*Known Drug Space*’ (KDS) such as toxic moieties and structures that complement their biological targets, *i.e.*, , the human genome codes approximately 19,000 to 20,000 different proteins.<sup>40</sup> Consequently, this resulted in a spectrum of drugs in the clinic ranging from very simple to large complex structures.



**Figure. 1.3.** Illustration of the concept of chemical space. Drug-like chemical space is depicted by the grey area that excludes moieties and molecular descriptors deemed undesirable in drugs. The red region depicts the undesirable properties and features that are included within KDS.

# *Chapter 2*

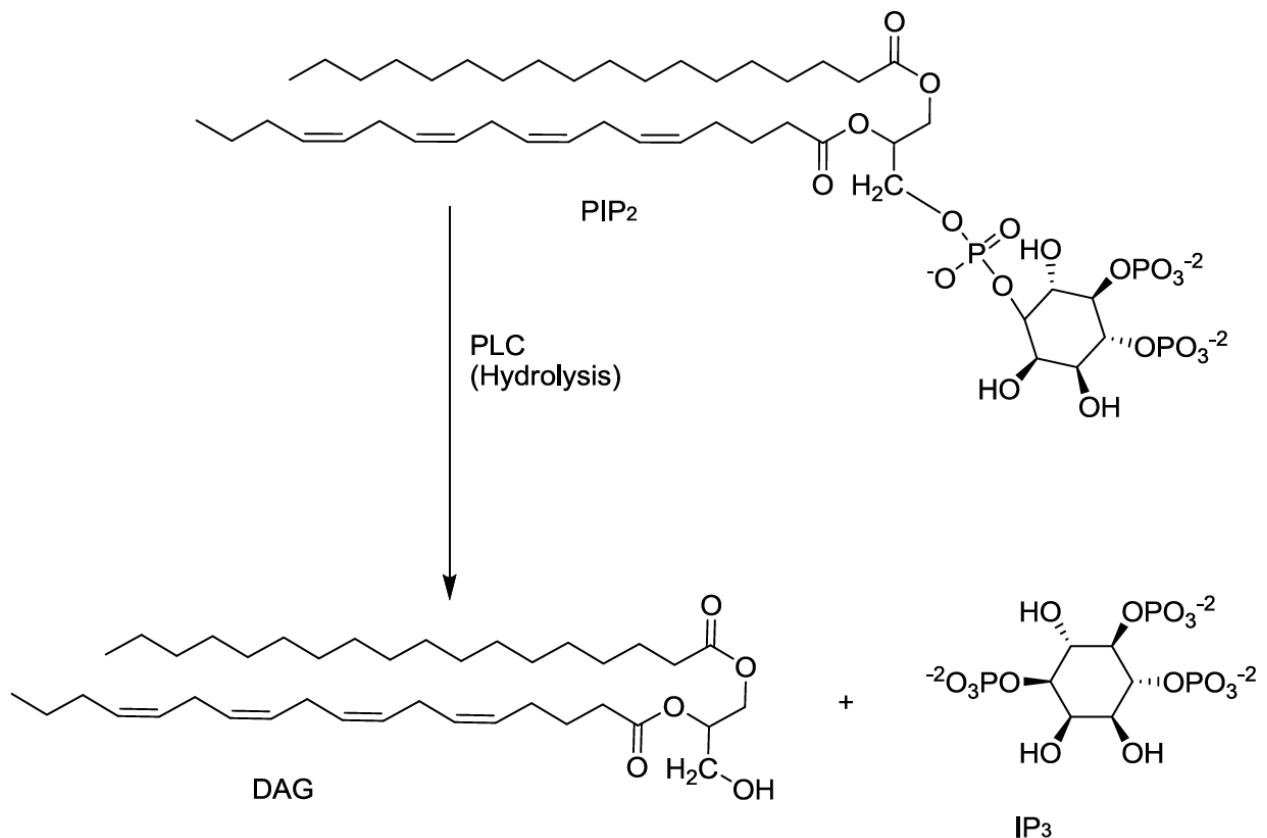
3-amino-thieno[2,3-*b*]pyridines and 1*H*-pyrazoles  
as phosphatidylinositol - specific phospholipase C  
inhibitors



## 2.1 Introduction

### 2.1.1 Overview of phosphatidylinositol - specific phospholipase C

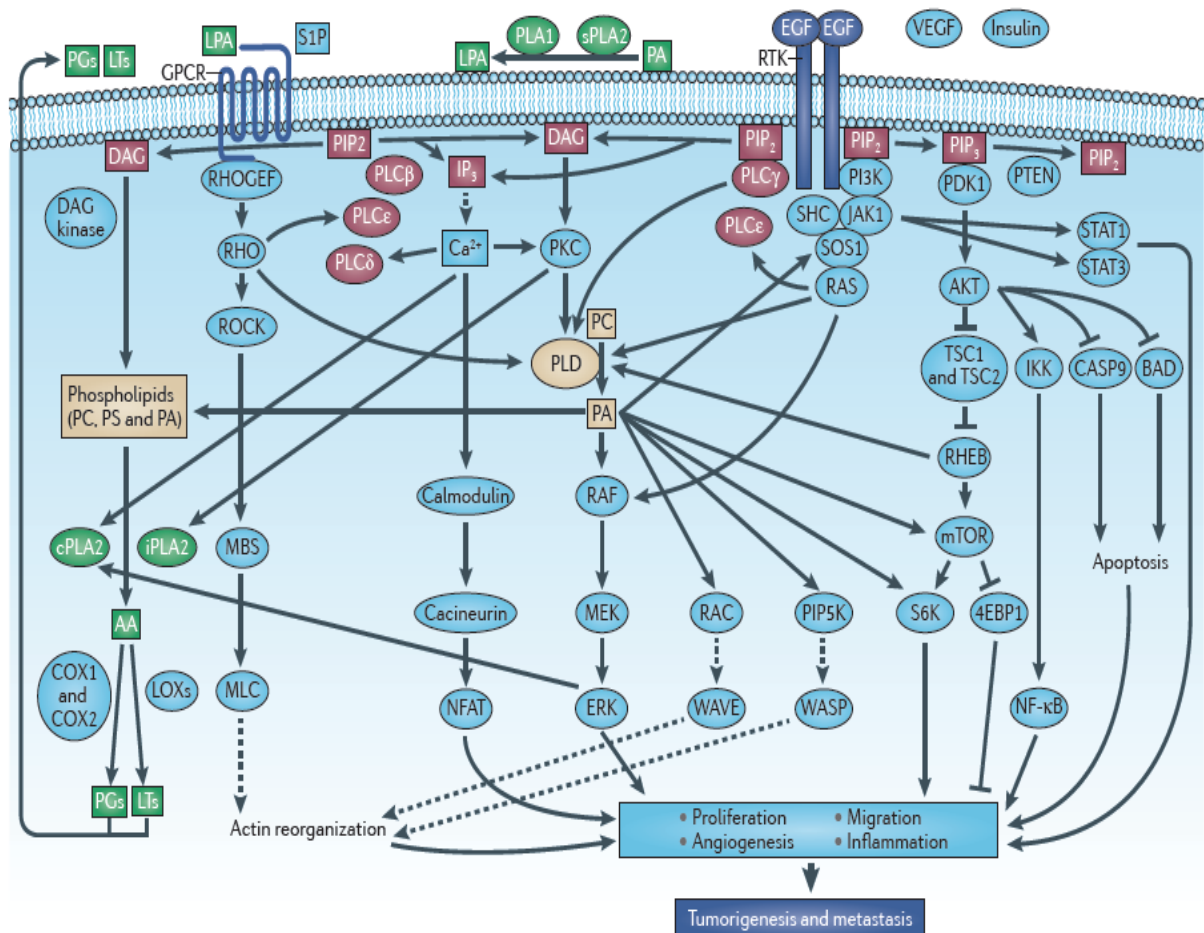
PI-PLC are essential mediators of intercellular and intracellular signalling that hydrolyse phosphatidylinositol 4,5-bisphosphate (PIP<sub>2</sub>) phospholipids into bioactive secondary messengers, inositol triphosphate (IP<sub>3</sub>) and diacylglycerol (DAG), which play a signalling role downstream within the signal transduction pathway.<sup>41-42</sup> The basic scheme for PIP<sub>2</sub> hydrolysis is shown in Fig. 2.1.



**Figure 2.1.** The hydrolysis of PIP<sub>2</sub> into secondary messengers DAG and IP<sub>3</sub>.

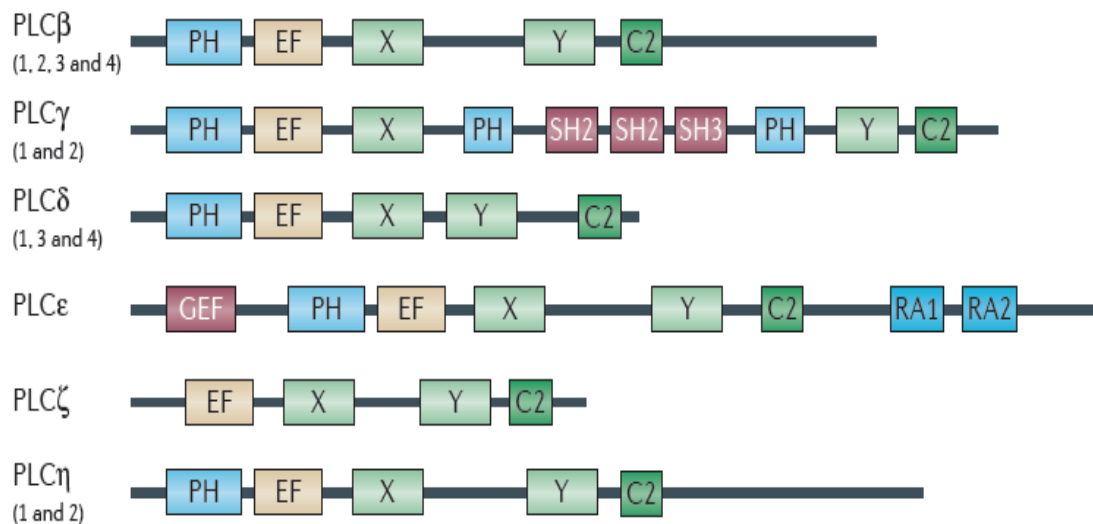
Specifically, IP<sub>3</sub> induces the release of Ca<sup>2+</sup> ions from intracellular stores within the endoplasmic reticulum and DAG together with Ca<sup>2+</sup> supplies activates the calcium-dependent enzyme, protein kinase C (PKC) to further propagate signal transduction.<sup>41-42</sup> Fig. 2.2 shows an overview

illustration of the phospholipase signalling network. PI-PLC can be activated by interaction of extracellular sources such as growth factors, hormones and lipids to receptor sites on the cell membrane surface, leading examples includes G-protein coupled receptors (GPCRs) and receptor tyrosine kinases (RTKs).<sup>42-43</sup> Other phospholipases A, B and D are not as prominently recognised for their roles in signalling pathways which are involved in stimulation of cancer-like behaviours.<sup>43</sup> Unregulated or overstimulation of the PI-PLC lead to a spur of abnormal cell proliferation, cell division, migration, increased cell motility and angiogenesis, typical features of cancer progression.<sup>43</sup>



**Figure 2.2.** A general overview of phospholipase family pathways and networks. Bold lines represent direct processes whilst dotted lines represent indirect processes. (Diagram obtained from Park *et al.*<sup>43</sup>)

Fig. 2.3 shows the distribution of domain types in six mammalian PLC subfamilies divided into thirteen different isoforms classified as PLC- $\beta$  ( $\beta$ 1- $\beta$ 4), PLC- $\gamma$  ( $\gamma$ 1- $\gamma$ 2), PLC- $\delta$  ( $\delta$ 1,  $\delta$ 3 and  $\delta$ 4), PLC- $\epsilon$ , PLC- $\xi$  and PLC- $\eta$  ( $\eta$ 1- $\eta$ 2).<sup>41-43</sup> So far only the PLC- $\delta$ 1, PLC- $\beta$ 2 and PLC- $\beta$ 3 have been crystallized and studied.<sup>42</sup> Several domains are known to have high structural similarities between the subfamilies and isoforms: the pleckstrin homology domains (PH), EF hand motifs (EF), X and Y catalytic domains and the calcium binding C2 domains have similarities greater than 40-50%.<sup>41-42</sup>



**Figure 2.3.** Structures of six different PI-PLC subfamilies in humans. (Diagram obtained from Park *et al.*<sup>43</sup>)

### 2.1.2 PLC- $\gamma$

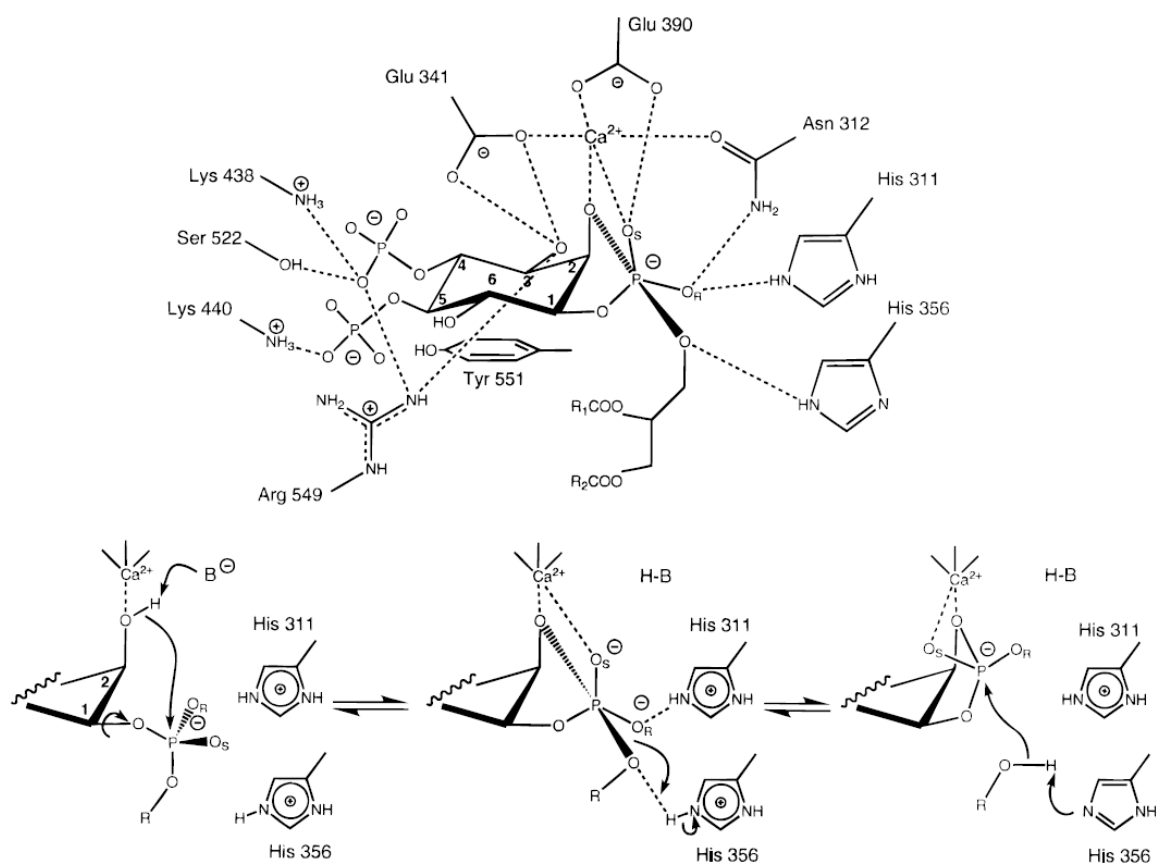
Given the large family of the PI-PLCs, the roles of each isoform differs in the signalling pathway resulting in various cellular expression.<sup>43</sup> For example, PLC- $\delta$ 1 activity had been identified as a tumour suppressor when functional studies showed that PLC- $\delta$ 1 expression suppresses both *in vitro* and *in vivo* tumourigenic ability in oesophageal squamous cell

carcinoma.<sup>44</sup> In contrast, PLC- $\gamma$  is more widely known for its role in cell proliferation and is activated by an array of oncogenic tyrosine kinase growth factor receptors.<sup>45</sup> Several studies have shown linkages between PLC- $\gamma$  activity and cancer. To list a few insightful examples: down-regulation of PLC- $\gamma$ 1 expression severely impaired activation of Rac protein and cell invasion of MBA-MB-231 breast and U87 brain cancer cell lines,<sup>46</sup> RNA knockdown of PLC- $\gamma$ 1 blocks cell mitogenesis in squamous cell carcinoma<sup>47</sup> and high levels of PLC- $\gamma$ 1 were found in human breast tissues when compared to normal breast tissues.<sup>48</sup> Two PLC- $\gamma$  isoforms have been identified: PLC- $\gamma$ 1 is ubiquitously expressed<sup>49</sup> and PLC- $\gamma$ 2 is found primarily in haematopoietic cells.<sup>50</sup> Due to the reasons mentioned, PLC- $\gamma$  is considered the most interesting PI-PLC. Furthermore, the lack of drug development targeting PI-PLCs spurs interest for this family of PLC as a viable anticancer drug target.

### 2.1.3 Catalytic site of PI-PLC- $\delta$ 1 and mechanism of cleavage

Unfortunately, the structure of the mammalian  $\gamma$ -isoform have been poorly characterised, *i.e.*, partial structures of the enzyme such as SH2 and SH3 domains are given in the PDB database. This prompted the use of the PLC- $\delta$  as a model for the  $\gamma$ -isoform.<sup>42-43, 49</sup> This approach is justified owing to shared genetic similarities amongst the isoforms (see section 2.1.1).

The catalytic site is occupied by a  $\text{Ca}^{2+}$  ion that is coordinated to the side chains of Glu341, Glu390, Asn312, Asp343 and a water molecule, part of the coordination complex is shown in Fig. 2.4.<sup>51</sup> Furthermore, the cyclic 1,4,5-IP<sub>3</sub> is stacked on the aromatic Tyr551 side chain, and equatorial diphosphates formed hydrogen bonds with polar Ser522 residue and positively charged amino acids Lys438, Lys440 and Arg549.

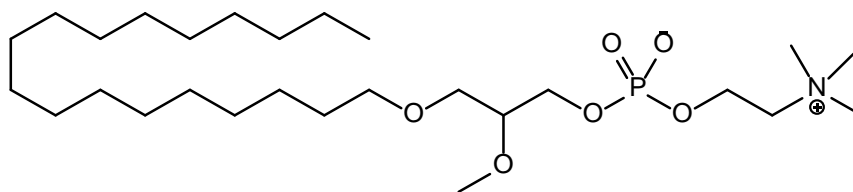


**Figure 2.4.** The catalytic site of PI-PLC- $\delta 1$  and mechanism of cleavage. (Diagram obtained from Essen *et al.*<sup>51</sup>)

As shown in Fig. 2.4, the role of the  $\text{Ca}^{2+}$  ion is to facilitate the binding of 2-hydroxyl group of PIP<sub>2</sub> from the axial position. This is followed by proton abstraction from the 2-hydroxyl group by a base. One of the roles played by the  $\text{Ca}^{2+}$  ion is probably to lower the  $\text{pK}_a$  of the 2-hydroxyl group for this step. Glu341 and Glu390 have been suggested as possible candidates for the base due to their favourable hydrogen bonding geometries.<sup>51-52</sup> His311 is considered to be too far away from accepting or donating a proton but instead stabilises the intermediate.<sup>51-52</sup> This follows a phosphorylase step, a nucleophilic attack by a water molecule releasing IP<sub>3</sub> and DAG-alkoxide that accepts a proton from His356.<sup>52</sup> It was reported from mutagenesis studies that replacing His311, His356 and Glu341 resulted in substantial decreases in PLC- $\delta 1$  catalytic activities.<sup>53</sup>

### 2.1.3 PLC inhibitors

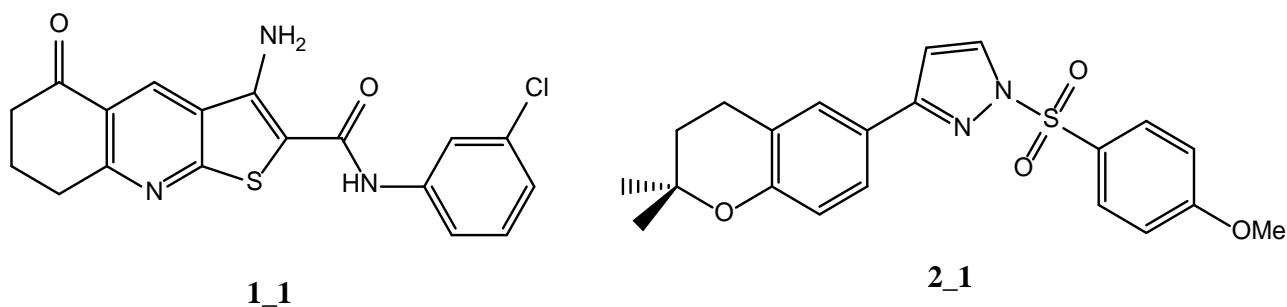
Several PLC inhibitors were discovered as steroids,<sup>54-55</sup> peptides, lipids and some were isolated from natural products and microbes.<sup>56</sup> Unfortunately, their development displayed mediocre micromolar potencies, many of which established inhibition against the  $\gamma$ -isoform. A few examples are hispidospermidine ( $IC_{50} = 16 \mu M$ ), licochalcone A ( $IC_{50} = 30 \mu M$ ), ariculatin ( $IC_{50} = 20 \mu M$ ), and myroridin K ( $IC_{50} = 6.7 \mu M$ ).<sup>56</sup> One of the prominent PI-PLC inhibitor is an antitumour lipid, edelfosine (see Fig. 2.5), a platelet-activating factor found in small concentrations in the body<sup>57-58</sup> that predominantly inhibit cytosolic fibroblast PLC- $\gamma$  ( $IC_{50} = 0.4 \mu M$ ).<sup>58</sup> Edelfosine was also evaluated in clinical trials for the treatment of advanced non-small-cell bronchogenic carcinoma,<sup>59</sup> acute leukaemia<sup>60</sup> and patented for the treatment of brain tumours.<sup>61</sup>



**Figure 2.5.** Molecular structure of PLC- $\gamma$  inhibitor, edelfosine ( $IC_{50} = 0.4 \mu M$ ).<sup>58</sup>

### 2.1.4 Research Prologue

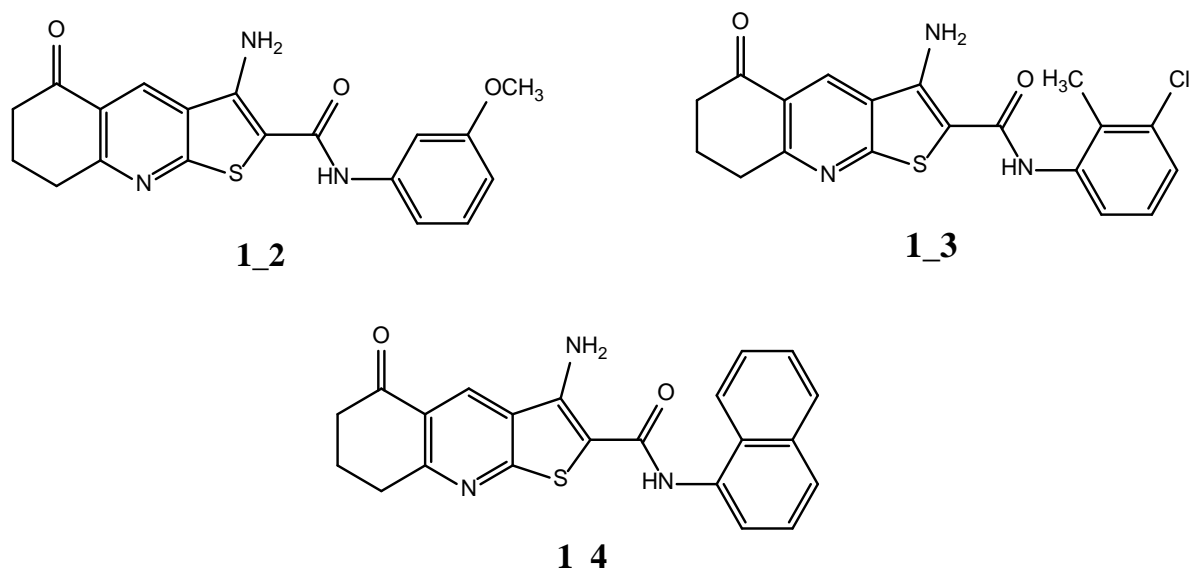
In a previous publication by Reynisson *et al.*<sup>62</sup> a library collection of commercially available compounds were downloaded from the ZINC<sup>63</sup> website and were docked to the PLC- $\delta 1$  crystal structure in the vHTS process.<sup>62</sup> Two of the hits were AThP **1\_1** and 1*H*-pyrazole **2\_1** (Fig. 2.6).



**Figure 2.6.** Two hit compounds **1\_1** and **2\_1** from vHTS by Reynisson *et al.*<sup>62</sup>

#### 2.1.4.1 3-amino-thieno[2,3-*b*]pyridine **1\_1**

**1\_1** was also tested in a  $^3\text{H}$ -PIP<sub>2</sub> flashplate biochemical assay and achieved 100% inhibition of PLC- $\gamma$ 2 at 50  $\mu\text{M}$ .<sup>62</sup> The assay relied on the ability for PLC- $\gamma$ 2 to hydrolyse  $^3\text{H}$ -PIP<sub>2</sub> into  $^3\text{H}$ -IP<sub>3</sub> and DAG resulting in loss of the radioactive signal.<sup>64</sup> It was further tested in leukaemia cell lines resulting in 63.8% growth arrest at 10  $\mu\text{M}$  single dose and LC<sub>50</sub>= 0.9  $\mu\text{M}$  (lethal concentration required to kill 50% of the cell population) in MDA-MB-435 cell line demonstrating the ability to not only inhibit growth but also kill cancer cells at a low micromolar concentration.<sup>65</sup> In order to expand the SAR, synthetic, *in silico* and biological studies were conducted yielding several derivatives.<sup>66-67</sup> One of the derivative synthesised, **1\_2** (Fig. 2.7) has *m*-methoxy substitution that showed high growth arrest particularly to leukaemia cell lines (98.4%) with 10  $\mu\text{M}$  single dose suggesting selectivity towards the PLC- $\gamma$ 2 isoform.<sup>66</sup> Two other derivatives, **1\_3** and **1\_4** with 3-chloro-2-methylphenyl and naphthalene substitutions, shown in Fig. 2.7 resulted from similarity search through collections of commercially available compounds.<sup>67</sup> They were tested in the NCI60 cell line panel using a 10  $\mu\text{M}$  single dose assay that resulted with the best NCI60 growth arrest so far; **1\_3** (83.6%) and **1\_4** (79.2%) with growth arrest for leukaemia cell lines at 95.5% and 91.1% respectively.<sup>67</sup>



**Figure 2.7.** Molecular structures of AThPs **1\_2**, **1\_3** and **1\_4** that showed 75.1%, 83.6% and 79.2% mean growth arrest respectively compared to 0% growth arrest in untreated cancer cells.

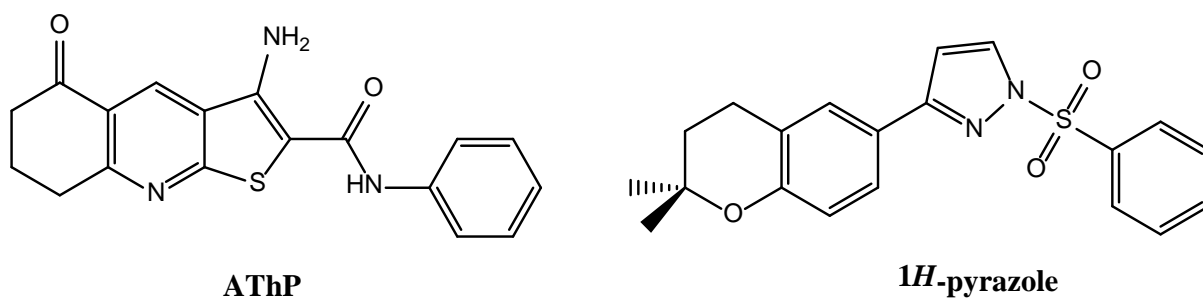
#### 2.1.4.2 1*H*-pyrazole **2\_1**

Hit **2\_1** was tested in a  $^3\text{H}$ -PIP<sub>2</sub> flashplate biochemical assay and showed evidence of PLC- $\gamma$ 2 inhibition (IC<sub>50</sub> ~7.5  $\mu\text{M}$ ).<sup>62</sup> **2\_1** was further tested in leukaemia and MDA-MB-468 (breast cancer) cell lines which resulted in 42.1% and 63.4% growth arrest compared to 0% growth arrest of untreated cancer cells.<sup>65</sup>

## 2.2 Results and Discussion

Clearly, anticancer activities were shown to be associated with AThP **1\_1** and 1*H*-pyrazole **2\_1**, and their analogues. To further explore the chemical space around the two chemical series (Fig. **2.8**) expanding their SAR profiles, synthetic, *in silico* and biological studies were conducted. The main focus of this chapter is based on *in silico* studies, in which a combination of similarity search, and molecular modelling strategies were employed.



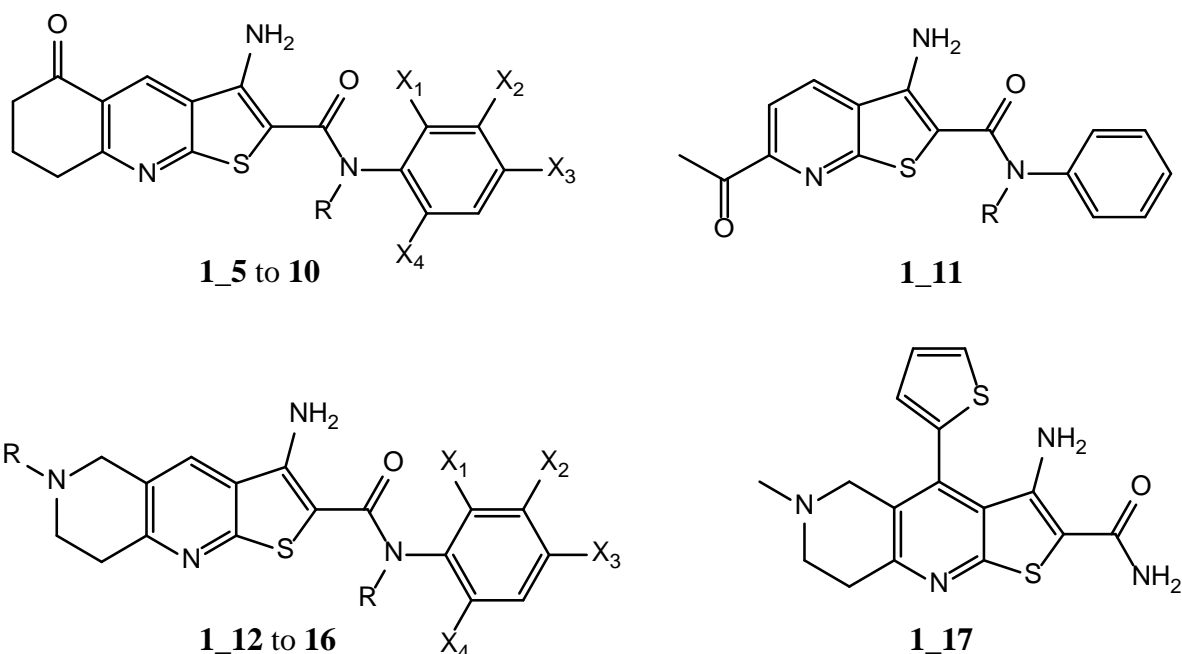


**Figure 2.8.** The basic molecular structures of the anticancer AThP (3-amino-5-oxo-N-phenyl-5,6,7,8-tetrahydrothieno[2,3-*b*]quinoline-2-carboxamide (left) and 1H-pyrazoles (3-(chroman-6-yl)-1-(phenylsulfonyl)-1H-pyrazoles), (right).

### 2.2.1 SAR of 3-amino-thieno[2,3-*b*]pyridine (AThP)

Structural similarity search was conducted on eMolecules,<sup>68</sup> a commercial compound database using the AThP structure shown in Fig. 2.8. Six hundred and fifty-four (654) compounds were found, using similarity coefficient of 0.7, and docked against the PLC- $\delta$ 1 scaffold (see Methodology). The structure of mammalian PLC- $\gamma$  has not been characterised prompting the use of the PLC- $\delta$  as a model for the  $\gamma$ -isoform.<sup>42-43, 49</sup> This approach is justified owing to the shared structural similarities between the isoforms (see section 2.1.1). Any compounds with no hydrogen bond formations and compounds satisfying the scoring criteria, GS < 60, CS < 26, ChemPLP < 64 and ASP < 30 were initially filtered. Derivatives **1\_9** (*m*-acetyl) and **17** (4-thiophen-2-yl) did not satisfy the criteria but were identified as compounds of ‘interests’ and were included in the selection. This resulted in thirteen virtual hits, which were tested using the NCI60 human tumour cell panel at 10  $\mu$ M concentration. Two structural classes were verified for the AThPs (**1\_5** to **10** and **12** to **16**) and two singletons (**1\_11** and **17**) with their molecular structures are shown in Table 2.1 with their mean growth of the NCI60 panel.

**Table 2.1** Antiproliferative activity of the AThPs and their molecular structures.



	R	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	Mean Growth Inhibition (%) <sup>a</sup>
<b>1_5</b>	H	H	Me	H	H	70.1
<b>1_6</b>	H	Et	H	H	H	74.6
<b>1_7</b>	H	Et	H	H	Me	74.8
<b>1_8</b>	H	H	Me	Me	H	73.7
<b>1_9</b>	H	H	Acetyl	H	H	2.1 <sup>b</sup>
<b>1_10</b>	Phenyl	H	H	H	H	1.0
<b>1_11</b>	-	-	-	-	-	-1.6
<b>1_12</b>	Benzyl	Me	H	Me	Me	72.7
<b>1_13</b>	Benzyl	H	Me	H	H	-0.6
<b>1_14</b>	Benzyl	H	OMe	H	H	47.3
<b>1_15</b>	Benzyl	H	H	H	H	40.6
<b>1_16</b>	Isopropyl	H	H	H	H	1.2
<b>1_17</b>	-	- <sup>c</sup>	- <sup>c</sup>	- <sup>c</sup>	- <sup>c</sup>	75.2

<sup>a</sup>Values represent relative growth (%) versus control as an average for 60 human tumour cell lines (NCI60). All compounds were tested at 10  $\mu$ M. <sup>b</sup>From Ref.<sup>67</sup> <sup>c</sup>No phenyl moiety.

It is clear that derivatives **1\_5** to **8**, **12** and **17** are the most active with mean growth inhibition of ~75% as compared to untreated cells (0%). When the substitution patterns of the phenyl rings are considered for derivatives **1\_5** to **10**, it can be concluded that small *o*- (X<sub>1</sub>) and *m*-alkyl (X<sub>2</sub>) substitutions are the most advantageous whilst electron withdrawing groups such as *m*-acetyl (**1\_9**) at X<sub>2</sub> is detrimental as compared to the un-substituted derivative that has a

relative growth inhibition of 69.8%.<sup>65</sup> It has been previously reported that *o*-substitution gives good results such as with fluoro (66.7%) and methyl (58.0%) groups but larger groups are not tolerated, *e.g.*, CF<sub>3</sub> (18.4%) and methoxy (38.7%).<sup>65, 67</sup> Furthermore, *m*-substitution is also favourable with fluoro at 68.6% and chloro at 63.7% relative growth inhibitions.<sup>65, 67</sup> This is in line with the results presented in Table 2.1.<sup>65, 67</sup> Interestingly, dual *m*- and *p*-methyl substitutions as for derivative **1\_8** has similar effect as derivatives **1\_5** to **7** suggesting the importance of *m*-alkyl substitutions at X<sub>2</sub> to anticancer activity since it has been previously seen that *p*-substitution is detrimental to the anticancer activity of the AThPs, *e.g.*, *p*-methoxy and *p*-chloro derivatives have no effect on the growth of the tumour cells.<sup>65, 67</sup> No anticancer activity was found for the *N,N*-diphenyl derivative **1\_10**.

Derivative **1\_11** is structurally unique in that it has a bicyclic instead of tricyclic ring system. It exhibited no activity, which can be explained by conserving the tricyclic ring system is important for anticancer activity of this class of compounds as has been previously observed.<sup>66</sup>

In general, the AThP derivatives **1\_13** to **17** exhibited more modest anticancer effects than the cyclohexanone-AThPs (**1\_5** to **10**) with the exception of derivatives **1\_12** and **17**, which had relative growth of 73% and 75%, respectively, which is similar to the most potent compounds in the **1\_5** to **10** series. In contrast to the cyclohexanone-AThPs (**1\_5** to **10**), *m*-methyl substitution is detrimental to the anticancer activity of the 3-amino-thieno[2,3-*b*][1,6]naphthyridines as seen for derivative **1\_13**. A possible explanation for this effect is that the two series are inhibiting different but related enzyme classes. Changing the R group from benzyl to isopropyl (**1\_16**) had a detrimental effect. Derivative **1\_17** is different in that it does not have a phenyl group but a thiophenyl on the pyridine moiety and can therefore be considered as a novel analogue in this chemical series. Nevertheless, it is potent in reducing the growth of the cancer cells.

Derivatives **1\_5** to **8** and **12**, **14**, **15** and **17** were selected for dose response testing based on their favourable growth inhibition at 10  $\mu$ M and the results are given in Table 2.2. The  $GI_{50}$  is the concentration for 50% of maximal inhibition of cell proliferation and therefore reflects the cytostatic concentration.<sup>8</sup> Five tumour cell lines were particularly affected: MDA-MB-435 (melanoma), MDA-MB-468 (breast), A498 (renal), SF-539 (central nervous system, CNS) and K-562 (leukaemia).

**Table 2.2.** The dose response for the most active compounds shown for five tumour cell lines in nano-molar (nM): MDA-MB-435 (melanoma), MDA-MB-468 (breast), A498 (renal), SF-539 (CNS) and K-562 (leukaemia).

	MDA-MB-435	MDA-MB-468	A498	SF-539	K-562
	$GI_{50}^a$	$GI_{50}^a$	$GI_{50}^a$	$GI_{50}^a$	$GI_{50}^a$
<b>1_5<sup>b</sup></b>	31/30	143/80	139/31	237/167	76/144
<b>1_6<sup>b</sup></b>	60/46	235/209	270/180	335/416	372/203
<b>1_7<sup>b</sup></b>	171/148	261/377	224/178	320/334	341/522
<b>1_8<sup>b</sup></b>	152/121	229/314	274/198	404/443	468/X
<b>1_12</b>	399	1780	1310	2440	903
<b>1_14</b>	296	2600	3820	1990	530
<b>1_15</b>	3430	17000	1780	2780	9100
<b>1_17</b>	3130	1750	2580	2350	4990

<sup>a</sup> $GI_{50}$ . <sup>b</sup>Tested twice and both values are given. X: no data.

It is clear that derivative **1\_5** is the most active compound particularly against MDA-MB-435 with an average  $GI_{50}$  of ~30 nM and inhibition in the 76 – 237 nM range for the other cell lines. The most active compound identified in this series has dual *o*-methyl and *m*-chloro substitution with  $GI_{50}$  of 18.5 nM for the MDA-MB-435 melanoma cell line and inhibition in the 23 – 38 nM range for the other cell lines.<sup>67</sup> Derivative **1\_6** is the second most active compound with activity in the 53 – 372 nM range followed by derivatives **1\_7** and **8**. Even though the 3-amino-thieno[2,3-*b*][1,6]naphthyridines are less active than their cyclohexanone-AThP counterparts identifying them using similarity search demonstrates the effectiveness of the technique to rapidly extend the SAR profile of the series.

## 2.2.2 Docking AThPs against PLC- $\delta$ 1

The AThPs were docked into the PLC- $\delta$ 1 substrate binding site in order to elucidate binding mechanisms of the compound series. Results showed some strongly potent derivatives have similar scores to modest derivatives. Interestingly, the second most active 3-amino-thieno[2,3-*b*][1,6]naphthyridine **1\_17** has the least scores across all four scoring functions (Table 2.3). This suggests the chemical class can target a wider range of proteins other than PLC- $\gamma$ . Literature review affirmed the AThPs with various mechanisms of action: antagonists adenosine A<sub>2A</sub> receptor (a G protein-coupled receptor),<sup>69</sup> as inhibitors of copper trafficking interfaces of ATOX1 and CCS proteins, for inhibition of tumour growth and Wilson's disease,<sup>70</sup> hepatocellular carcinoma-specific anti-tumour activity<sup>71-72</sup> and tubulin inhibitors.<sup>73-</sup>

74

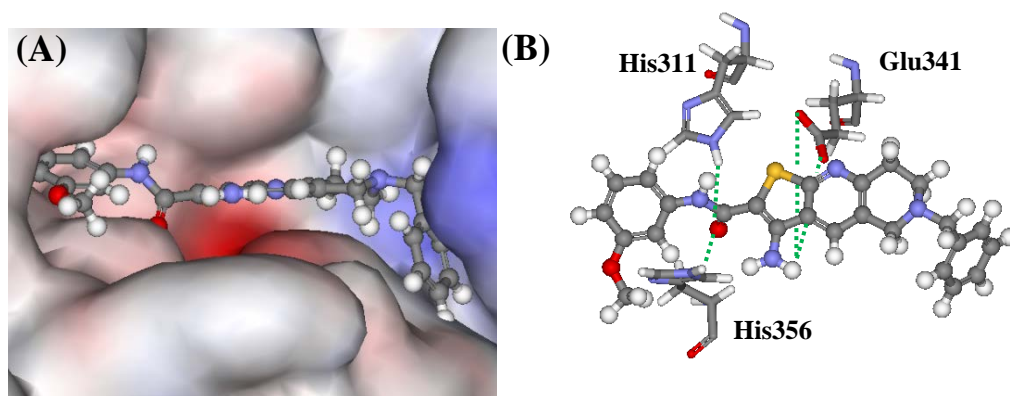
**Table 2.3.** Docking results and for AThP derivatives **1\_5** to **17** using the four scoring functions: GS, CS, ChemPLP and ASP. The antiproliferative activities are shown with the mean of the leukaemia tumour cell lines.

Derivative	GS	CS	ChemPLP	ASP	Mean Growth Inhibition (%)
<b>1_5</b>	62.0	32.3	72.3	36.1	70.1
<b>1_6</b>	66.0	27.6	65.7	34.6	74.6
<b>1_7</b>	63.7	34.9	85.0	37.0	74.8
<b>1_8</b>	64.1	31.6	66.9	31.7	73.7
<b>1_9</b>	59.6	31.2	71.4	38.8	2.1 <sup>a</sup>
<b>1_10</b>	65.2	35.5	83.0	39.7	1.0
<b>1_11</b>	60.7	30.3	73.6	33.0	-1.6
<b>1_12</b>	67.1	26.2	71.4	31.8	72.7
<b>1_13</b>	70.4	29.2	81.6	36.3	-0.6
<b>1_14</b>	69.9	29.5	81.7	37.1	47.3
<b>1_15</b>	68.2	28.6	77.8	40.2	40.6
<b>1_16</b>	60.6	26.5	64.9	33.7	1.2
<b>1_17</b>	46.4	22.4	49.4	26.7	75.2

<sup>a</sup>From Ref.<sup>67</sup>

Molecular modelling studies against PLC- $\delta$ 1 showed that the cyclohexanone-AThPs (**1\_5** to **10**) have hydrogen bonding to two amino acids histidine (His311) via the carboxamide and glutamic acid (Glu341) with the amine group. Also, arginine (Arg549) and lysine (Lys438)

interacts with the ketone group in the cyclohexanone moiety and the phenyl group is embedded into a lipophilic pocket. In general, the binding mode is the same as has been previously reported on this chemical series.<sup>65-67, 75</sup> The 3-amino-thieno[2,3-*b*][1,6]naphthyridine family had a similar predicted hydrogen bonding pattern, *i.e.*, His311 and His356 are involved in hydrogen bonding to the carboxamide moiety, Glu341 to the amino group and the phenyl moiety occupies the lipophilic pocket to the left hand side as shown in Fig 2.9 for derivative 1\_14. It was reported that alanine mutations at residues His311 and His356 caused ~20,000 and ~6,000 fold reduction in PLC- $\delta$ 1 activity, Glu341 involved with  $\text{Ca}^{2+}$  interactions and residues Lys438 and Arg549 are involved with interactions to phosphate groups which is associated with substrate selectivity.<sup>76</sup> The chemical series was consistently reported to interact with His311 and His356.<sup>65-67, 75</sup> which are involved in vital roles in protonation and deprotonation steps during catalysis.<sup>51, 76</sup> Moreover, a series of pyrimidinone and triazinone analogues were reported to be inactive,<sup>66</sup> the 3-amino group is incorporated into the ring system to eliminate the ability for Glu341 to act as a hydrogen bond acceptor suggesting the role of the AThP series in disrupting interactions with catalytic calcium, a cofactor essential for stabilising hydrogen bond networks during formation of the enzyme-substrate complex.<sup>76</sup> This led to speculation that Glu341 and dual histidine residues His311 and His356, are favourable residues for the AThP inhibition. The benzyl extension to the right hand side fills another lipophilic pocket, which could explain the reduction of activity for derivative 1\_16 due to its smaller isopropyl moiety.

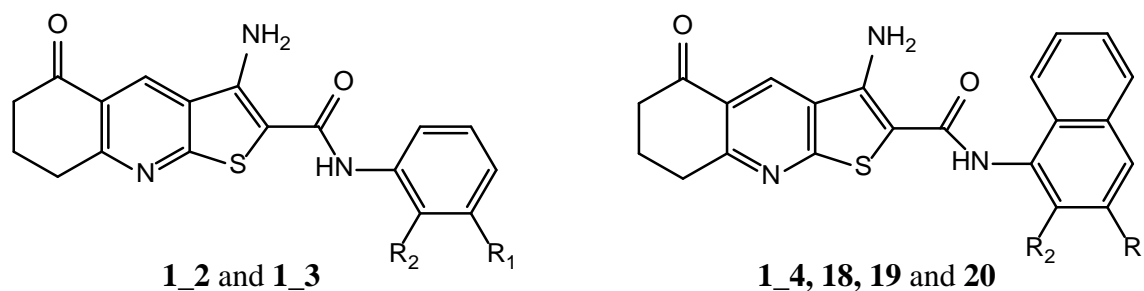


**Figure 2.9.** The docked configuration of derivative **1\_14** in the PLC- $\delta$ 1 binding site. **(A)** The phenyl group occupies the lipophilic cavity to the left hand side. Red depicts a negative partial charge on the surface, blue depicts positive partial charge and grey shows neutral/lipophilic regions. **(B)** Hydrogen bonds are depicted as green dotted lines between derivative **1\_14** and amino acids His311, Glu341 and His356.

### 2.2.3 Molecular Modelling for PLC- $\gamma$ 2 selectivity

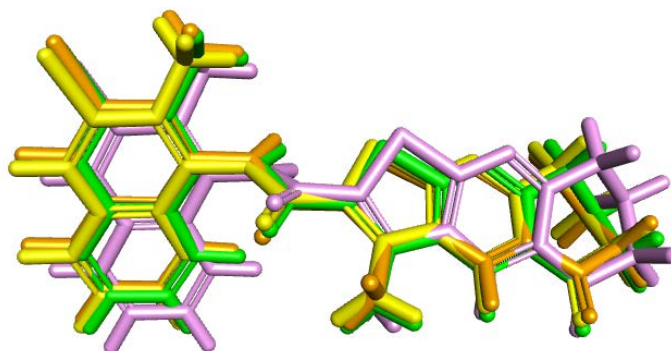
It is assumed that there is a degree of selectivity for the PLC- $\gamma$ 2 isoform based on the activity of the chemical series against the leukaemia cell lines particularly for **1\_2** (98.4%), **1\_3** (95.5%) and **1\_4** (91.1%). Computer-aided Drug Design (CADD) is used as a tool to modify the structure of the AThP series to potentially amplify selectivity against PLC- $\gamma$ 2 using the PLC- $\delta$ 1 structure as a model. The selectivity is presumed to be strongly linked to the naphthalene moiety and substitution patterns of the three actives. Thus, it is sensible to modify the structure based on these results.

**Table 2.4.** The molecular structures of three derivatives, **1\_18**, **19** and **20**, designed using CADD strategy. The scores were obtained using the GS, CS, ChemPLP and ASP algorithms. Modifications were made at R<sub>1</sub> and R<sub>2</sub>.



ID	R <sub>1</sub>	R <sub>2</sub>	GS	CS	ChemPLP	ASP
<b>1_2</b>	OCH <sub>3</sub>	H	60.3	28.5	70.3	35.1
<b>1_3</b>	Cl	CH <sub>3</sub>	66.5	38.1	82.4	38.5
<b>1_4</b>	H	H	71.0	35.9	83.1	41.1
<b>1_18</b>	Cl	CH <sub>3</sub>	72.8	42.9	90.9	42.4
<b>1_19</b>	OCH <sub>3</sub>	H	66.3	38.1	84.4	40.2
<b>1_20</b>	OCH <sub>3</sub>	CH <sub>3</sub>	72.8	39.2	90.6	42.4

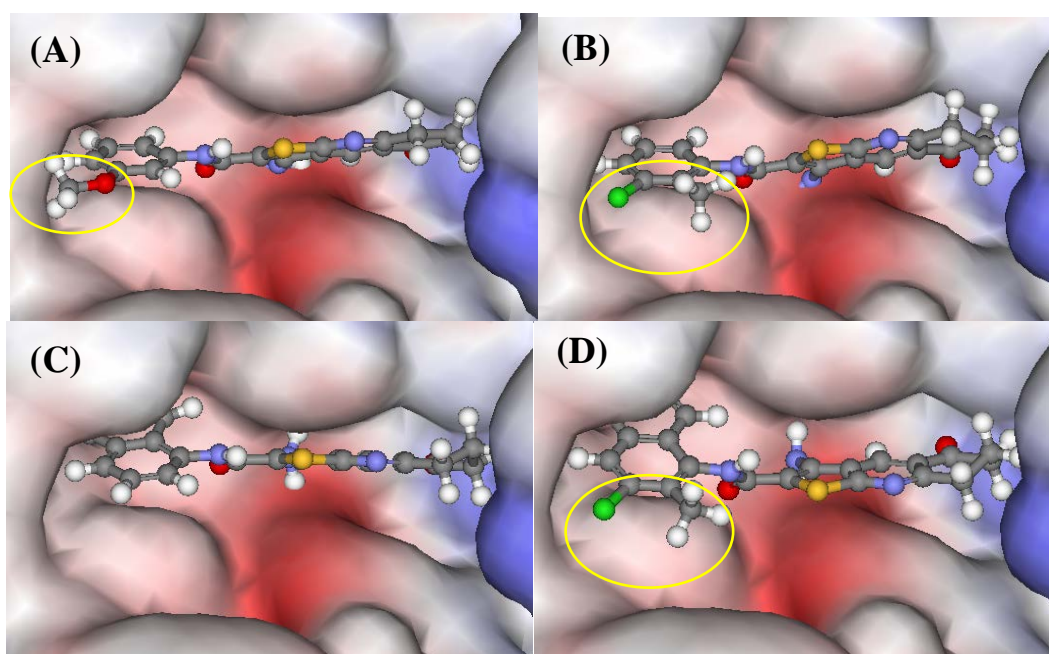
The naphthalene-based design at R<sub>1</sub> and R<sub>2</sub> prompted increases in the scores (Table 2.4) relative to the cyclohexanone-AThP series, a positive indication of increased binding affinity. The three hypothetical derivatives, **1\_18**, **19** and **20**, were further overlaid to investigate the reliability of the docking protocol. The average root-mean square deviations (RMSDs) were calculated to be near zero, for all three derivatives, *e.g.*, **1\_18** shown in Fig. 2.10 illustrates good overlays between all four scoring functions giving an agreeable consensus.



**Figure 2.10.** Overlaid docked configurations of **1\_18**. The ChemPLP configuration (green) was used as the reference for calculating the RMSDs; GS = 0.56Å (pink), CS = 0.20Å (yellow) and ASP = 0.39Å (orange).



The binding modes of the bioactive AThPs **1\_2**, **1\_3** and **1\_4** are shown in Fig. 2.11A to C, in which the substituents on the aromatic rings are circled in yellow facing the aqueous environment. The binding mode of these derivatives are in line with previous modelling studies on the AThP series having similar orientations, scores and hydrogen bonding patterns.<sup>66-67, 75</sup> The substituents on the phenyl rings face the water environment for **1\_2** and **1\_3**, whereas for derivative **1\_4**, the lipophilic naphthalene moiety is occupied in the deeper hydrophobic pocket. It was revealed after docking **1\_18**, **19** and **20** into the PLC- $\delta$ 1 binding site that they also have the same binding mode as the bioactive compounds, *i.e.*, the lipophilic naphthalene moiety occupies the deeper region of the hydrophobic pocket and the substituents on the naphthalene faces the water environment. The models presented are agreeably similar in binding modes with enhanced predicted binding affinities, thus the three hypothetical derivatives are recommended for synthesis and screening against the PLC- $\gamma$ 2 isoform.

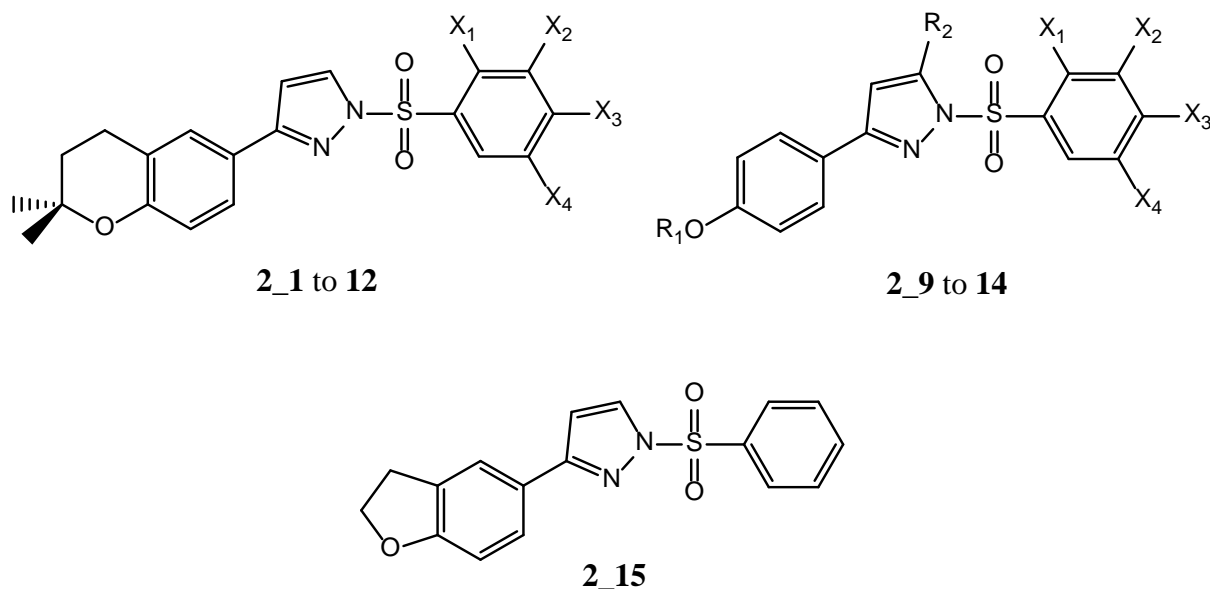


**Figure 2.11.** Docked configurations of **1\_2** (A), **1\_3** (B), **1\_4** (C) and **1\_18** (D). Red depicts a negative partial charge on the surface, blue depicts positive partial charge and grey shows neutral/lipophilic regions. The substituents on the aromatic systems are circled in yellow.

## 2.2.4 SAR of 1*H*-pyrazoles

The 1*H*-pyrazole **2\_1** (see Table 2.5) has a measured IC<sub>50</sub> ~ 7.5 μM for PLC-γ2.<sup>62</sup> Similarity coefficient value of 0.7 was used to conduct a structural similarity search based on **2\_1** resulting in 63 compounds. These were docked to the PLC-δ1 binding pocket and four virtual hits were identified (**2\_9** to **2\_11** and **2\_15**, Table 2.5). Unfortunately, no compounds with the tetrahydro-2*H*-pyran moiety were found as in **2\_1**, severely limiting the SAR study. In order to address this shortcoming seven derivatives with the tetrahydro-2*H*-pyran moiety were synthesised (**2\_2** to **2\_8**) with various substitutions on the phenyl ring as shown in Table 2.5. In particular, *p*-substitutions (X<sub>3</sub>) on the phenyl ring were made since the hit (**2\_1**) has a methoxy group there, indicating that this position is important. Also, a derivative without any substitutions on the phenyl ring was made as a reference. Furthermore, for the hits **2\_9** to **2\_11**, three close structural analogues were synthesised to extend the SAR study (**2\_12** to **2\_14**). The molecular structures are shown in Table 2.5 with the total growth of the NCI60 panel at 10 μM as compared to untreated cells (0%).

**Table 2.5.** Antiproliferative activity of the 1*H*-pyrazole structural derivatives. **2\_1** is the lead compound for the 1*H*-pyrazole series.



	R <sub>1</sub>	R <sub>2</sub>	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	Mean Growth Inhibition (%) <sup>a</sup>
<b>2_1</b>	-	-	H	H	OMe	H	22.7 <sup>c</sup>
<b>2_2</b>	-	-	H	H	H	H	23.5
<b>2_3</b>	-	-	H	H	CF <sub>3</sub>	H	10.2
<b>2_4</b>	-	-	H	H	<i>t</i> -Butyl	H	16.1
<b>2_5</b>	-	-	Cl	H	Cl	H	12.1
<b>2_6</b>	-	-	H	H	NHCOMe	H	35.5
<b>2_7</b>	-	-	H	OMe	OMe	H	27.6
<b>2_8</b>	-	-	H	Me	H	Me	14.5
<b>2_9<sup>b</sup></b>	(CH <sub>2</sub> ) <sub>3</sub> F	H	H	H	Me	H	15.0
<b>2_10<sup>b</sup></b>	Et	NH <sub>2</sub>	H	H	OMe	H	-2.2
<b>2_11<sup>b</sup></b>	(CH <sub>2</sub> ) <sub>3</sub> F	H	H	H	Br	H	12.5
<b>2_12</b>	Me	H	H	H	Me	H	15.8
<b>2_13</b>	Me	H	H	H	OMe	H	18.2
<b>2_14</b>	Isopropyl	H	H	H	OMe	H	13.0
<b>2_15<sup>b</sup></b>	2,3-dihydrofuran	-	H	H	H	H	18.1

<sup>a</sup>Values represent relative growth (%) versus control as an average for 60 human tumour cell lines (NCI60). All compounds were tested at 10 μM.

<sup>b</sup>Compounds resulted from structural similarity search.

<sup>c</sup>From Ref.<sup>65</sup>

As can be seen in Table 2.5, modest reduction of growth is achieved by the 1*H*-pyrazoles with only *p*-acetamide (2\_6) and dual *m*- and *p*-methoxy (2\_7) electron-donating groups showed marginally better results than 2\_1. Absence of the tetrahydro-2*H*-pyran ring (2\_9 to 14) and ring contraction from a six membered ring to five membered tetrahydrofuran (2\_15) did not result in improved anticancer activity. To further test the potency of the 1*H*-pyrazoles, 2\_1 was tested with aid from our collaborator using the MTT colorimetric assay, which measures the activity of NADPH-dependent redox enzymes in the cytosol giving information on the viability of cells.<sup>16</sup> The MTT assay relies on sources of NADPH reductase (viable cells) to reduce colourless MTT to detectable purple formazan product (see appendix Scheme A1).<sup>77</sup> The breast cancer cell line MDA-MB-468 was incubated with 10  $\mu$ M concentration of 2\_1 and were compared to the activity of the anticancer drug camptothecin (topoisomerase I poison). Minimal effect was observed (data not shown) confirming the modest anticancer effect for this chemical class.

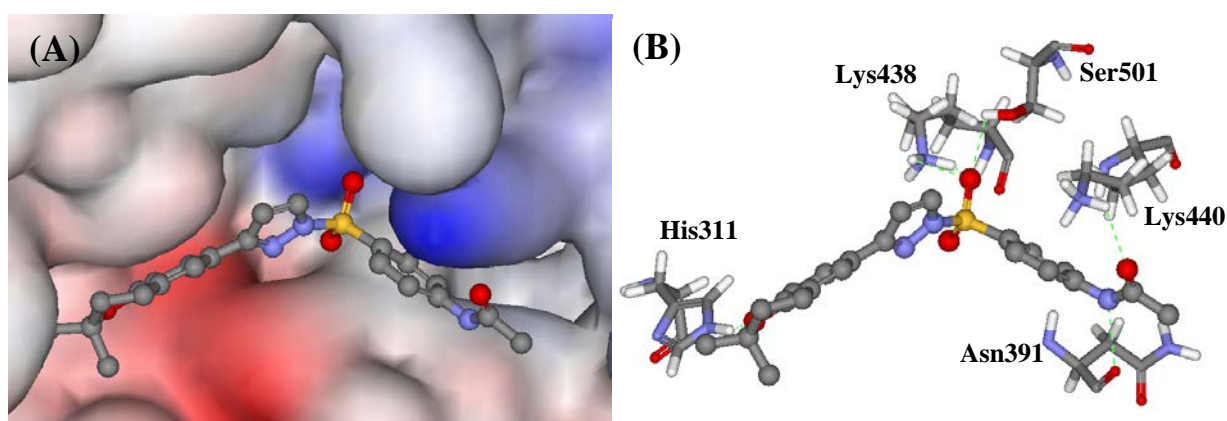
### 2.2.5 Docking 1*H*-pyrazoles against PLC- $\delta$ 1

The 1*H*-pyrazoles were docked into the PLC- $\delta$ 1 substrate binding site in order to elucidate their binding mechanisms. Results showed similar scores between the derivatives, the highest scoring derivative 2\_6 showed most prominent activity against the NCI60 panel with the strongest activity against leukaemia (mean = 44.2%) shown in Table 2.6. This suggests the 1*H*-pyrazoles inhibit a range of PLCs, more selectively for the  $\gamma$ 2-isoform.

**Table 2.6.** Docking results and for 1*H*-pyrazole derivatives **2\_1** to **15** using the four scoring functions: GS, CS, ChemPLP and ASP. The antiproliferative activities are shown with the mean of the leukaemia tumour cell lines.

Derivative	GS	CS	ChemPLP	ASP	Mean Growth inhibition (%)	Mean Leukaemia growth inhibition (%)
<b>2_1</b>	66.3	32.4	76.5	30.8	22.7	42.1
<b>2_2</b>	63.3	32.5	75.2	31.2	23.5	42.9
<b>2_3</b>	63.5	32.5	71.0	29.6	10.2	16.4
<b>2_4</b>	66.7	31.6	74.8	29.3	16.1	16.8
<b>2_5</b>	66.8	29.1	76.1	29.5	12.1	17.4
<b>2_6</b>	70.7	35.3	82.8	33.4	35.5	44.2
<b>2_7</b>	67.5	29.1	77.7	31.4	27.6	31.2
<b>2_8</b>	65.3	32.4	71.9	29.5	14.5	12.8
<b>2_9</b>	68.5	29.7	79.7	32.0	15.0	24.9
<b>2_10</b>	63.1	28.8	74.9	30.6	-2.2	-2.2
<b>2_11</b>	68.9	28.3	83.4	32.0	12.5	21.4
<b>2_12</b>	60.9	25.1	69.9	27.6	15.8	13.6
<b>2_13</b>	64.3	30.1	63.8	27.4	18.2	30.4
<b>2_14</b>	66.9	28.2	77.2	29.7	13.0	25.0
<b>2_15</b>	59.5	30.8	70.8	29.5	18.1	22.8

Molecular modelling revealed a complementary fit to the substrate binding site exemplified by derivative **2\_6** (Fig. **2.12A**), which shows occupation of 2,2-dimethylchroman and phenylacetamide to lipophilic pockets on the left and right hand sides respectively. Hydrogen bonding interactions were accommodated by residues His311, Lys438, Ser501, Lys440 and Asn391 (Fig. **2.12B**).



**Figure 2.12.** (A) The docked configuration of derivative **2\_6** in the PLC- $\delta$ 1 binding site. The hydrogens on the ligand was suppressed for clarity. Red depicts a negative partial charge on the surface, blue depicts positive partial charge and grey shows neutral/lipophilic regions. (B) Hydrogen bonds are depicted as green dotted lines between derivative **2\_6** and amino acids His311, Lys438, Ser501, Lys440 and Asn391.

The inhibitory action of the 1*H*-pyrazole series is speculated to disrupt hydrogen bonding and key catalytic steps by residue His311, and interference with phosphate binding to Asn391, Lys438, Lys440 and Ser501 residues shown in Fig. **2.12**. It was seen that unsubstituted and *p*-substituted phenyl with electron-donating groups seemed to display the best anti-leukaemia effects. Modelling *o*- and *m*-methoxy substituted phenyl derivatives did not show hydrogen bonding with Asn391 or Lys440, which suggested electro-donating groups at *p*-substituted position to be most selective for PLC- $\gamma$ 2 inhibition.

## 2.3 Methodology

### 2.3.1 Similarity Search, Molecular modelling and CADD

The AThP and 1*H*-pyrazole derivatives, were acquired from commercial sources using the eMolecules<sup>68</sup> web-based compound library. Substructure and Tanimoto similarity search methods were used to identify plausible inhibitors.<sup>78</sup> The structures shown in Fig. **2.8** were used as initial search scaffolds with similarity coefficient given at 0.7 and supplemented with substructure searches.

The compounds were docked to the crystal structure of PLC- $\delta$ 1 (PDB ID: 1DJX, resolution 2.3 Å),<sup>51</sup> which was obtained from the Protein Data Bank (PDB).<sup>9-10</sup> The Scigress v2.6<sup>79</sup> was used to prepare the crystal structure for docking, *i.e.*, hydrogen atoms were added, the co-crystallised ligand (D-Myo-Inositol-1,4,5-Triphosphate, IP<sub>3</sub>) was removed as well as crystallographic water molecules. The Scigress software suite was also used to build the inhibitors and the MM2<sup>80</sup> force field was used to optimise the structures. The centre of the binding pocket was defined as the position of the Ca<sup>2+</sup> ion (x = 126.257, y = 38.394, z = 22.370) with 10 Å radius. Fifty docking runs were allowed for each ligand with default

search efficiency (100%). The basic amino acids lysine and arginine were defined as protonated. Furthermore, aspartic and glutamic acids were assumed to be deprotonated. The GoldScore (GS),<sup>18-19</sup> ChemScore (CS),<sup>22-23</sup> ChemPLP<sup>22, 24</sup> and ASP<sup>25</sup> scoring functions were implemented to validate the predicted binding modes and relative energies of the ligands using the GOLD v5.4.0 software suite. The results were inspected visually and seventeen virtual hits selected.

### 2.3.1 Synthetic studies, NCI60 assay and MTT assay

Synthetic studies were conducted to expand the SAR of the 1*H*-pyrazoles. For the preparation of derivatives **2\_2** to **8** and **2\_12** to **14**, see Ref.<sup>81</sup> The derivatives were prepared by Distinguished Professor Margaret Brimble and colleagues from the synthetic laboratory at the University of Auckland.

All twenty-eight compounds were tested for their anticancer activity in the NCI60 assay by the National Cancer Institute, Maryland in the United States. For general protocol, see Ref.<sup>8</sup>

The MTT assay was conducted to verify for the anticancer activities of the 1*H*-pyrazoles in MDA-MB-468 breast cancer cells. For the protocol of the MTT assay, see Ref.<sup>77</sup> The assay was conducted by Prof. Inga Reynisdóttir from the Landspítali University Hospital, Reykjavík, Iceland.

## 2.4 Future Work and Conclusions

The anticancer activity of the AThP and 1*H*-pyrazole derivatives was investigated. Tests against the NCI60 cancer cell panel revealed that the AThPs were the most active with derivative **1\_5** giving GI<sub>50</sub> = 30 nM against the MDA-MB-435 melanoma cell line emphasising the importance of *m*-substituted phenyl ring moiety for this chemical series,

which is in line with previous findings. By applying a combination of similarity searching and molecular modelling, the 3-amino-thieno[2,3-*b*][1,6]naphthyridines were found to have anticancer activity extending the SAR. The 3-amino-thieno[2,3-*b*][1,6]naphthyridines is likely to inhibit a related target to PLC, *i.e.*, *m*-methyl (**1\_13**) showed no growth inhibition effect in contrast to derivative **1\_5**.

On the other hand, the 1*H*-pyrazoles only showed moderate growth inhibition for the NCI60 panel, which was verified using the MTT assay. Particularly, only leukaemia cell lines were affected by *p*-substituted electron-donating groups.

Additionally, molecular modelling showed the two chemical series are able to recognise key residues with the potential to inhibit catalytic mechanisms that were responsible for their anticancer effects.

The AThPs are known to be linked to various protein targets (see section **2.2.2**). Biochemical assays are required to identify their inhibitory activities against the known targets for an accurate evaluation on their selectivity. Furthermore, the chemical series is known to particularly affect MDA-MB-435 (melanoma) and MDA-MB-468 (breast) cell lines. It would be interesting to see proteins highly associated with the mentioned cell lines be experimentally tested against the AThPs. Additionally, the compounds can be docked and find correlations with the experimental data. A similar approach can be adopted for the 1*H*-pyrazoles that particularly affect leukaemia cell lines.



# *Chapter 3*

3-Amino-thieno[2,3-*b*]pyridines as microtubule-destabilising agents

## 3.1 Introduction

### 3.1.1 Microtubule structure and function

Microtubules are hollow cylindrical chain polymers constituting of repeating  $\alpha\beta$ -tubulin heterodimers that bind head to tail into protofilaments (Fig. 3.1).<sup>82</sup> Parallel contacts of usually 13 protofilaments give rise to one microtubule chain with dimensions approximately 25 nm in diameter, about 5 nm thickness and may stretch to 5 - 10  $\mu\text{m}$ .<sup>82-83</sup>

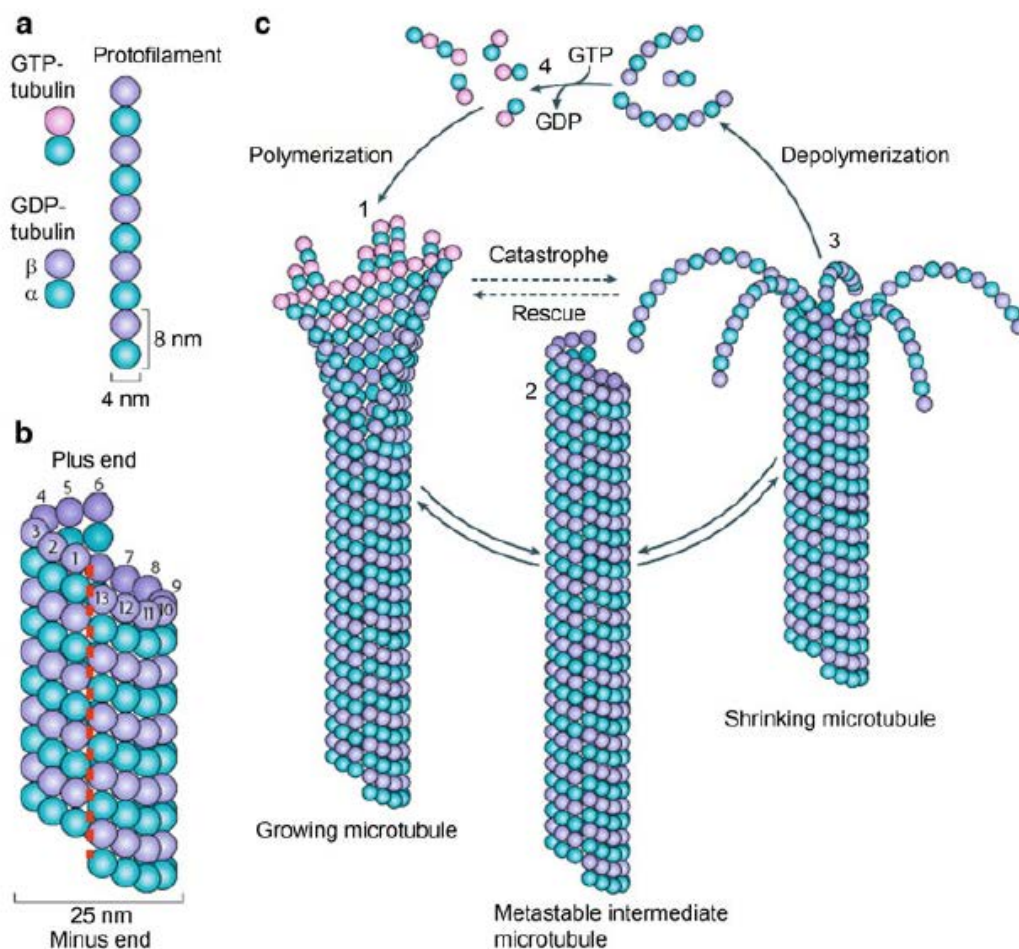
Microtubules play crucial roles in numerous cellular processes that include mitosis, intracellular trafficking of organelles and vesicles, architectural distribution of cell organelles, cell shape maintenance and migration.<sup>82-83</sup> Their strands are organised into more complex networks like centrioles, mitotic spindles, cytoplasmic networks for vesicle and organelle transport, flagellum and cilium as powerhouse structures for movement.<sup>82</sup>

### 3.1.2 Microtubule dynamics

Microtubules are highly dynamic structures which refer to the coexistence between growth and shortening phases of microtubules driven by two processes; polymerisation/assembly and depolymerisation/disassembly of heterodimer  $\alpha\beta$ -tubulin monomers.<sup>82, 84</sup> A main factor that governs the two processes during steady-state is the 'critical concentration' of free tubulin in the solution.<sup>84</sup> This is the concentration of free tubulin when there is no net change in elongation or shortening of microtubules, the concentration of free tubulin needs to exceed the critical concentration for a corresponding assembly of  $\alpha\beta$ -tubulin monomers, the reverse applies for disassembly of protofilaments.<sup>84</sup>

Polymerisation and depolymerisation utilises energy mediated by guanosine-5'-triphosphate (GTP) binding and hydrolysis.<sup>82-84</sup> It was hypothesised that GTP-bounded  $\alpha\beta$ -tubulin

monomers become assembled to the protofilaments more favourably at the (+)-end, and disassembly is more favourable at the (-)-end capped by  $\alpha$ -tubulin, *i.e.*, rate of growth is greater at the (+)-end than the (-)-end.<sup>82</sup> The GTP-bound  $\alpha\beta$ -tubulin polymerise to the protofilaments in a straight conformation and almost immediately GTP is hydrolysed to GDP and inorganic phosphate which stimulates a conformational change from a straight to curved chain.<sup>82, 85</sup> The curved protofilaments are forced to remain straight to further accommodate assembly of the filament chain giving rise to strained force.<sup>82, 85</sup> Continuous capping by the GTP-bound  $\alpha\beta$ -tubulins are required to stabilise the straightened conformation as shown in Fig. **3.1c**, the end where there is polymerisation has the  $\beta$ -subunit of GTP-bound  $\alpha\beta$ -tubulin coloured pink. On the other hand, depolymerisation ensued from build ups of GDP-bound  $\alpha\beta$ -tubulin capped ends illustrated as purple in Fig. **3.1c**. By far, this GTP/GDP-dependent model is the most acceptable in explaining the mechanisms underlying microtubule polymerisation and depolymerisation.<sup>85</sup> Depolymerisation is characterised by a 'flower-like' shape (Fig. **3.1c**) followed by fragmentation of protofilaments. Shifting from slow microtubule assembly to rapid assembly is known as 'rescue' whereas rapid disassembly is called 'catastrophe'.<sup>82-83</sup> The phase interchange between 'rescue' and 'catastrophe' has been accepted as a property of microtubule dynamics known as 'dynamic instability'.<sup>82-83</sup> A second property of microtubule dynamics is 'treadmilling', a process in which the assembly of  $\alpha\beta$ -tubulin heterodimers at the (+)-end is balanced by the disassembly of  $\alpha\beta$ -tubulin heterodimers at the (-)-end.<sup>82-83</sup> The two properties are important in maintaining microtubule dynamics during steady-state.

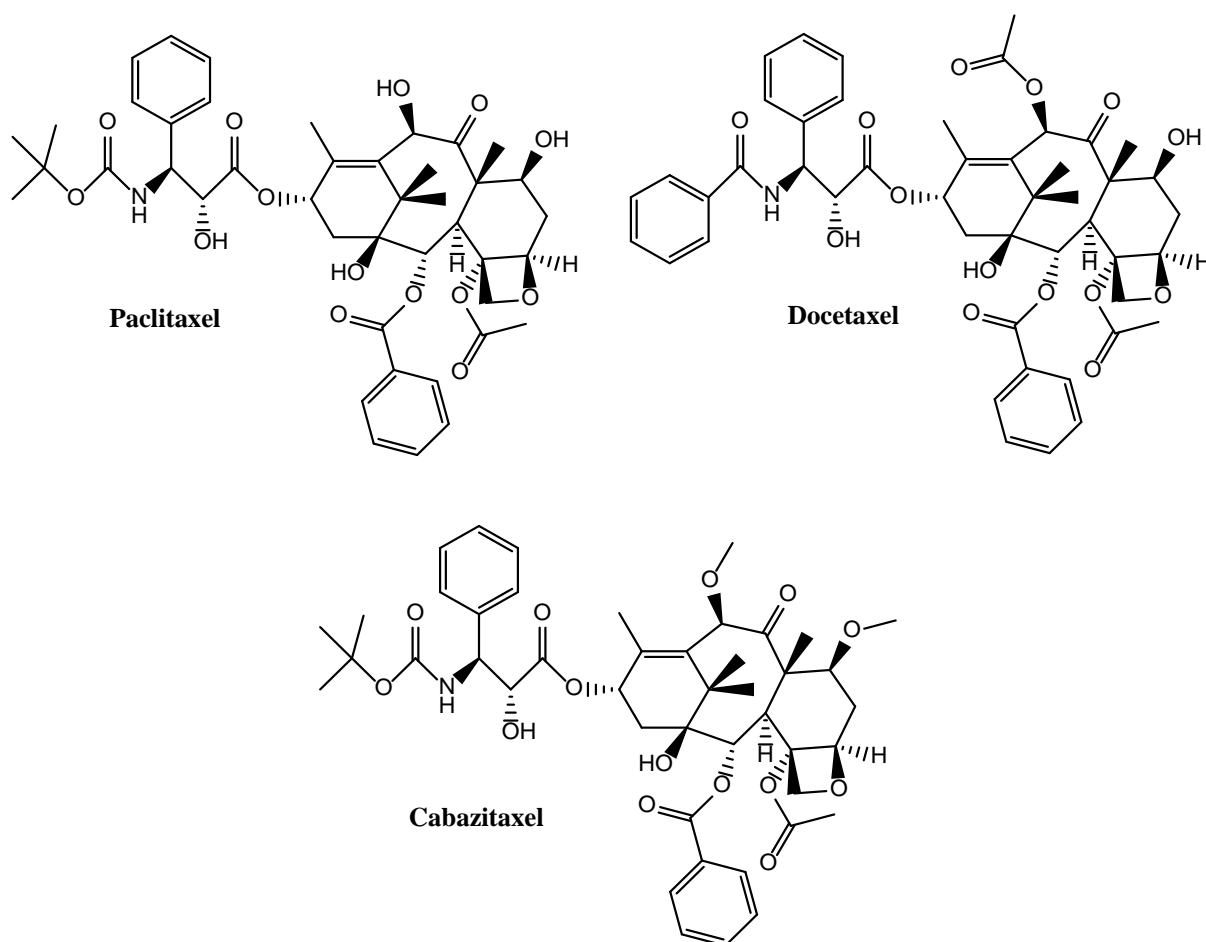


**Figure 3.1.** A basic representation of microtubule structures and dynamic instability. (Diagram was obtained from Draber *et al.*<sup>82</sup>) (a) Representation of GDP- and GTP-bound  $\alpha\beta$ -tubulin heterodimers and a protofilament. The  $\alpha$ -subunit is depicted as green.  $\beta$ -subunit bound to GDP and GTP is depicted as purple and pink, respectively. (b) Helical representation of a microtubule constituting of 13 protofilaments adjacently attached. The red-dashed line is to terminate the number of protofilaments at 13. (c) A simplified illustration of dynamic instability, a property of microtubule dynamics.

Microtubule dynamics are features that microtubule-targeted drugs are designed to perturb. These drugs are commonly used in treatment of cancer, particularly because mitosis is uncontrolled in cancer cells. A diverse range of small molecules and natural products were isolated and identified as microtubule-targeted drugs.<sup>83, 86</sup> Generally, these drugs are divided into two groups as either microtubule-stabilising or destabilising agents.

### 3.1.3 Microtubule-stabilising agents

Microtubule-stabilising agents are chemicals that stabilises the microtubule structure that polymerisation is more favourable than depolymerisation.<sup>83, 87</sup> Microtubule-stabilising agents are largely isolated and derived from natural products including plants, microorganisms and marine organisms.<sup>87-88</sup> Notably, the taxanes isolated from yew plants exhibited good potency with acceptable side effects that by far branded the taxanes as the most developed microtubule-stabilising agents.<sup>87, 89</sup> Fig 3.2 show examples of marketed drugs from the taxane family; paclitaxel (Taxol), docetaxel (Taxotere) and cabazitaxel (Jevtana).



**Figure 3.2.** The molecular structures of microtubule-stabilising taxanes; paclitaxel, docetaxel and cabazitaxel.

Taxanes bind more favourably to polymerised tubulin in which the binding site is located in the  $\beta$ -tubulin subunit commonly known as the 'taxol site' and is presumed as a common binding site for most microtubule-stabilising agents.<sup>83, 87, 89</sup> The location of the binding site was confirmed by crystallography investigation.<sup>90</sup> The exact molecular mechanism of microtubule-stabilising agents is unknown but is postulated to induce a conformational change that stabilise the structure of microtubules and enhance polymerisation.<sup>89</sup>

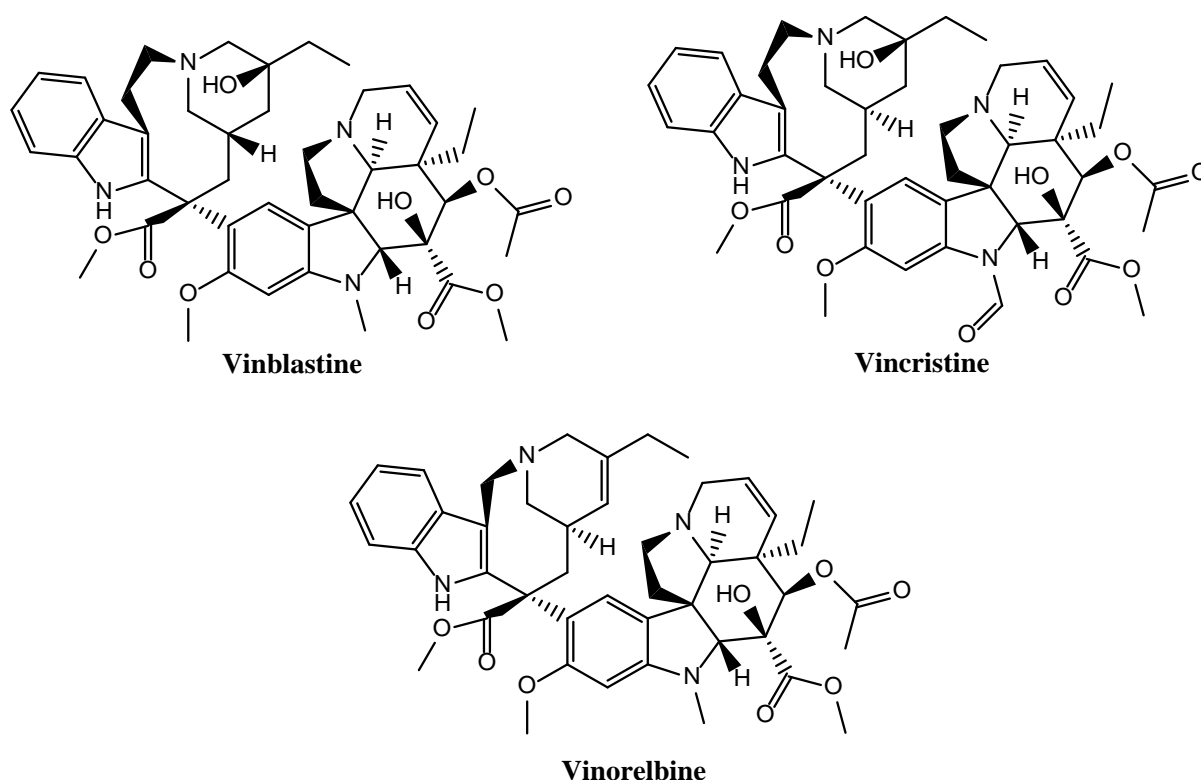
Under the microscope, the most recognised feature of microtubule stabilisation is the formation of defined stable microtubule bundles.<sup>87</sup> This feature has been widely approved for diagnosing tubulin stabilisation effects. A second feature of microtubule stabilisation is the formation of multipolar spindles resulting from the segregation of chromosomes in multiple directions.<sup>87</sup> Ultimately, as a consequence of failed mitotic separation lead to apoptosis.

### 3.1.4 Microtubule-destabilising agents

These agents are chemicals that destabilise the microtubule structure and favours depolymerisation over polymerisation.<sup>83, 86</sup> A diverse range of microtubule-destabilising agents exists, many of which were developed into marketed drugs. These include small molecules, natural products and peptide derivatives.<sup>83, 86</sup> The most prominently acclaimed microtubule-destabilising agents are derivatives of colchicine and *Vinca* alkaloids which displays similar destabilisation effects, *i.e.*, disintegration of microtubules.<sup>83, 86</sup> However, they bind at different sites in tubulin, namely the colchicine site and *Vinca* site.<sup>83, 86</sup> So far, only destabilising agents that target the *Vinca* site have been approved to the market, examples are shown in Fig. 3.3.

The colchicine site is located at the interface of  $\alpha\beta$ -tubulin heterodimer between the  $\alpha$ - and  $\beta$ -tubulin subunits.<sup>91-92</sup> A property of the colchicine site is its structural simplicity.<sup>92</sup> Consequently, an array of various small molecules were discovered and developed with good

affinity for the colchicine site whereas those that inhibit the taxol or *Vinca* domains are larger and more complex.<sup>92</sup> Colchicine binds more favourably to free tubulins in solution to form tubulin-colchicine complex followed by a slow conformational change in the complex which is incorporated into the growing microtubule chain.<sup>93</sup> A defect in the growing chain leads to a perturbation in microtubule dynamics resulting in depolymerisation.



**Figure 3.3.** Examples of three marketed microtubule-destabilising agents from the *Vinca* alkaloid family.

Similar to colchicine, inhibition at the *Vinca* site by the *Vinca* alkaloids result in destabilised microtubule dynamics and depolymerisation. The *Vinca* alkaloids bind to the  $\alpha\beta$ -interface between two tubulin molecules.<sup>86, 91-92</sup> Preferentially, they bind to (+)-end of microtubules at a high affinity that significantly results in suppression of microtubule dynamics and depolymerisation.<sup>94-95</sup>

### 3.1.5 Sea urchin embryos as model organisms

Sea urchin embryos have long been used for studies in developmental biology. The merits of using sea urchin embryos have attracted their applicability in drug testing, *e.g.*, such as assessing for features in microtubule-targeted drugs.<sup>96</sup> These merits include cost-effectiveness, ease of maintenance, avoids ethical-related issues, morphological changes are easily monitored and rapid cell division.<sup>96-97</sup> The phenotypic changes in morphology have been validated for both microtubule-stabilising and destabilising agents. Presence of paclitaxel, a microtubule-stabilising agent, show bundles of microtubule, multipolar spindles and cleavage arrest *i.e.*, lack of cell division during the embryonic phase.<sup>96-97</sup> On the other hand, presence of microtubule-destabilising agents altered cell division and at higher concentrations permitted embryonic spinning and formation of tuberculous-shaped eggs due to destruction of cilia and intracellular cytoskeleton networks, respectively.<sup>96-97</sup>

## 3.2 Results and Discussion

Initially, interests in screening the AThPs for microtubule-modulating activities stemmed from observations that phenstatin and nocodazole actives (see chemical structures Table 3.1) have structural similarities to the AThPs. Shown in Table 3.1 are the molecular structures of the AThP derivatives, **1\_21** to **25** and **1\_26** to **60** from series **A** and **B** featuring the carboxamide and ketone linkers, respectively. The test results in the phenotypic sea urchin embryo assay and the NCI60 human tumour cell line panel were also included. All forty compounds tested were purchased from Chemical Block Ltd. Derivatives **1\_21** to **1\_25** have a tricyclic AThP ring system linked to a phenyl ring *via* a carboxamide moiety. Derivatives **1\_26** to **1\_60** include cycloalkyl (**1\_26** to **39**), bicyclic (**1\_40** to **56**), quinuclidine (**1\_57** to **59**), and piperidine (**1\_60**) cores with a carbonyl linker. Specifically, derivatives **1\_40** to **48**



contain small substitutions at position groups R<sub>1</sub> - R<sub>3</sub> in the AThP ring system, whereas molecules **1\_49** to **56** have aromatic or heterocyclic substituents at R<sub>1</sub>.

### 3.2.2 Biological studies

#### 3.2.2.1 Sea urchin embryo assay results

Analysis of the data presented in Table **3.1** suggests that many of the AThPs displayed antiproliferative cytotoxic properties. In the *in vivo* phenotypic sea urchin embryo assay, the AThPs **1\_21** to **24** demonstrated the strongest effect, *i.e.*, caused pronounced cleavage alteration in the 10 – 50 nM range, cleavage arrest was observed at 40 – 500 nM concentrations, and embryo spinning (0.2 – 5 µM), suggesting antimitotic microtubule-destabilising activity comparable with the activity of phenstatin, a known destabilising agent. In this series, both cycloheptane and cyclohexane rings were well tolerated, since derivatives **1\_22** and **24** exhibited similar potencies. Introduction of methyl groups at the R<sub>6</sub> and R<sub>7</sub> positions resulted in derivative **1\_23** with similar activity, whereas the replacement of the cyclohexane ring in **1\_23** with cyclopentane (**1\_21**) reduced antimitotic effect. Interestingly, derivative **1\_25** substituted with the *m*-CF<sub>3</sub> group at R<sub>4</sub> did not affect sea urchin embryo development up to 4 µM concentration.

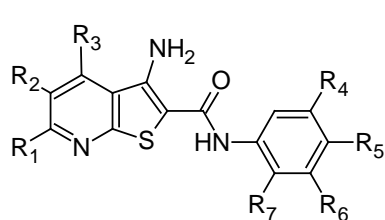
Tricyclic AThPs **1\_26** to **60** with carbonyl linker displayed weaker antimitotic effects upon linker replacement (ketone *vs.* amide) compared to the **1\_21** to **24** series. For the **1\_26** to **60** family, the unsubstituted derivative **1\_28** showed the strongest microtubule-destabilising effect (embryo spinning EC = 0.5 µM). In addition, compounds **1\_29**, **34**, **35**, and **38** can be considered as microtubule-destabilising agents, since they induced formation of tuberculate eggs typical for microtubule-destabilising agents.<sup>97</sup> The *p*-substituted methyl (**1\_26**), methoxy (**1\_30**), and halogen (**1\_27** and **1\_31** to **33**) derivatives were less active, unable to induce cleavage arrest and embryo spinning up to 4 µM concentration. Similar to series **1\_21** to **25**,

the replacement of the cyclohexane ring in **1\_29** with cyclopentane yielded less active molecule **1\_26**. Introduction of phenyl, 4-pyridine, and CF<sub>3</sub> groups at R<sub>3</sub> resulted in the loss of effect (**1\_35** vs. **36** and **37**; **38** vs. **39**).

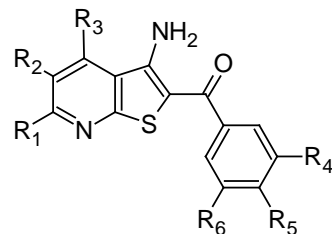
In general, bicyclic AThPs **1\_40** to **56** as well as quinuclidine (**1\_57** to **59**), and piperidine (**1\_60**) derivatives exhibited minimal or no antimitotic effect. Only *p*-methyl substituted derivative **1\_42** showed weak microtubule-destabilising activity, inducing the formation of tuberculate arrested eggs. Furthermore, AThPs **1\_41**, **45**, **46**, **48**, **51**, and **55** displayed weak antiproliferative activity.

The sea urchin embryo tests revealed that AThPs can target not only tubulin/microtubules, but also produce other tubulin-unrelated effects. Namely, derivatives **1\_36**, **37**, **39**, **47**, **50**, **52**, and **58** to **60** were unable to alter cleavage, however caused developmental retardation and alteration at later stages of development. Interestingly, thienyl derivatives **1\_53** to **55** exhibited specific embryotoxicity, inducing embryo disintegration and death at early pluteus, independently of the stage of exposure and treatment duration. These are qualitative observations and the mechanisms of these effects remain unclear and need further investigation in order to be established.

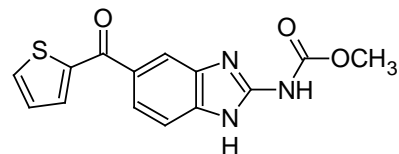
**Table 3.1.** The effects of AThPs on the sea urchin embryos and human cancer cells. Nocodazole and phenstatin are positive controls.



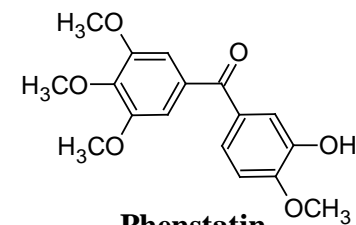
**Series A (1\_21 to 25)**



**Series B (1\_26 to 60)**



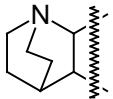
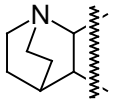
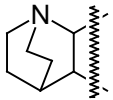
**Nocodazole**

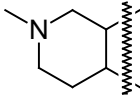


**Phenstatin**

Compound	R <sub>1</sub>	R <sub>2</sub>	R <sub>3</sub>	R <sub>4</sub>	R <sub>5</sub>	R <sub>6</sub>	R <sub>7</sub>	Sea urchin embryo effects, EC (μM) <sup>a</sup>			NCI60 screen	
								Cleavage alteration	Cleavage arrest	Embryo spinning	Mean GI, (%) <sup>b</sup>	Mean GI <sub>50</sub> , (μM) <sup>c</sup>
Nocodazole	-	-	-	-	-	-	-	0.005	0.001	0.1	-	0.0389 <sup>d</sup>
Phenstatin	-	-	-	-	-	-	-	0.01	0.05	0.5	-	0.06 <sup>e</sup>
1_21	-(CH <sub>2</sub> ) <sub>3</sub> -		H	H	H	Me	Me	0.05	0.5	5	76.4 <sup>f</sup>	0.42 <sup>f</sup>
1_22	-(CH <sub>2</sub> ) <sub>4</sub> -		H	H	H	H	H	0.01	0.05	0.2	69.7 <sup>f</sup>	1.00 <sup>f</sup>
1_23	-(CH <sub>2</sub> ) <sub>4</sub> -		H	H	H	Me	Me	0.01	0.1	0.5	ND <sup>g</sup>	
1_24	-(CH <sub>2</sub> ) <sub>5</sub> -		H	H	H	H	H	0.01	0.04	0.2	85.3 <sup>f</sup>	0.69 <sup>f</sup>
1_25	-(CH <sub>2</sub> ) <sub>5</sub> -		H	CF <sub>3</sub>	H	H	H	>4	>4	>4	90.0 <sup>f</sup>	3.09 <sup>f</sup>
1_26	-(CH <sub>2</sub> ) <sub>3</sub> -		H	H	Me	H	-	0.2	>4	>4	ND <sup>g</sup>	
1_27	-(CH <sub>2</sub> ) <sub>3</sub> -		H	H	Br	H	-	1	>4	>4	19.3	-
1_28	-(CH <sub>2</sub> ) <sub>4</sub> -		H	H	H	H	-	0.02	0.2	0.5	ND <sup>g</sup>	

1_29	-(CH <sub>2</sub> ) <sub>4</sub> -		H	H	Me	H	-	0.02	2 (TE) <sup>h</sup>	>4	16.4	-
1_30	-(CH <sub>2</sub> ) <sub>4</sub> -		H	H	OMe	H	-	0.5	>4	>5	ND <sup>g</sup>	
1_31	-(CH <sub>2</sub> ) <sub>4</sub> -		H	H	F	H	-	0.2	>4	>4	26.7	-
1_32	-(CH <sub>2</sub> ) <sub>4</sub> -		H	H	Cl	H	-	0.5	>4	>4	7.8	-
1_33	-(CH <sub>2</sub> ) <sub>4</sub> -		H	H	Br	H	-	0.5	>4	>4	8.1	-
1_34	-(CH <sub>2</sub> ) <sub>4</sub> -		H	Me	H	Me	-	0.05	0.2 (TE) <sup>h</sup>	>5	78.9	0.28
1_35	-(CH <sub>2</sub> ) <sub>4</sub> -		H	OMe	OMe	H	-	0.5	4 (TE) <sup>h</sup>	>5	60.5	3.24
1_36	-(CH <sub>2</sub> ) <sub>4</sub> -		Ph	OMe	OMe	H	-	>4	>4	>4	ND <sup>g</sup>	
1_37	-(CH <sub>2</sub> ) <sub>4</sub> -		4-Py	OMe	OMe	H	-	>4	>4	>4	ND <sup>g</sup>	
1_38	-(CH <sub>2</sub> ) <sub>4</sub> -		H	-OCH <sub>2</sub> O-		H	-	0.025	0.5 (TE) <sup>h</sup>	>5	ND <sup>g</sup>	
1_39	-(CH <sub>2</sub> ) <sub>4</sub> -		CF <sub>3</sub>	-OCH <sub>2</sub> O-		H	-	>4	>4	>4	ND <sup>g</sup>	
1_40	H	H	H	H	Br	H	-	1	>4	>4	ND <sup>g</sup>	
1_41	Me	H	H	H	Cl	H	-	0.5	>4	>4	18.4	-
1_42	Me	H	H	H	Me	H	-	1	4 (TE) <sup>h</sup>	>10	ND <sup>g</sup>	
1_43	Me	H	Me	H	Br	H	-	>4	>4	>4	ND <sup>g</sup>	
1_44	Me	H	Me	OMe	OMe	H	-	>4	>4	>4	14.2	-
1_45	Me	H	Me	-OCH <sub>2</sub> CH <sub>2</sub> O-		H	-	1	>4	>4	2.0	-
1_46	NH <sub>2</sub>	CN	H	H	F	H	-	4	>4	>4	ND <sup>g</sup>	

1_47	NH <sub>2</sub>	CN	H	H	Cl	H	-	>4	>4	>4	0.5	-
1_48	NH <sub>2</sub>	CN	H	OMe	OMe	H	-	4	>4	>4	4.6	-
1_49	2-MeO-C <sub>6</sub> H <sub>4</sub>	H	H	OMe	OMe	H	-	NA <sup>i</sup>			43.5	7.94
1_50	2-MeO-C <sub>6</sub> H <sub>4</sub>	H	H	-OCH <sub>2</sub> O-		H	-	>4	>4	>4	37.4	5.13
1_51	4-Py	H	H	OMe	OMe	H	-	4	>4	>4	17.1	-
1_52	4-Py	H	H	-OCH <sub>2</sub> O-		H	-	>4	>4	>4	ND <sup>g</sup>	
1_53	2-Thienyl	H	H	H	OMe	H	-	>4	>4	>4	ND <sup>g</sup>	
1_54	2-Thienyl	H	H	H	Cl	H	-	>4	>4	>4	ND <sup>g</sup>	
1_55	2-Thienyl	H	H	OMe	OMe	H	-	4	>4	>4	52.6	6.31
1_56	3,4-Methylene-dioxy-C <sub>6</sub> H <sub>3</sub>	H	CF <sub>3</sub>	OMe	OMe	H	-	NA <sup>i</sup>			ND <sup>g</sup>	
1_57			H	H	H	H	-	>4	>4	>4	7.3	-
1_58			H	H	OMe	H	-	>4	>4	>4	11.5	-
1_59			H	OMe	OMe	H	-	>4	>4	>4	11.7	-

1_60		H	H	H	H	-	4	>4	>4	13.3	
------	---	---	---	---	---	---	---	----	----	------	--

<sup>a</sup> The sea urchin embryo assay was conducted as described previously.<sup>97</sup> Fertilised eggs and hatched blastulae were exposed to 2-fold decreasing concentrations of compounds. Duplicate measurements showed no differences in effective threshold concentration (EC) values.

<sup>b</sup> GI %: single dose inhibition of cell growth at 10  $\mu$ M concentration.

<sup>c</sup> GI<sub>50</sub>: concentration required for 50% cell growth inhibition. This is the mean for the 60 cell lines tested in the NCI60 panel.

<sup>d</sup> NCI ID 238159.

<sup>e</sup> Data from Ref.<sup>98</sup>

<sup>f</sup> Data from Ref.<sup>99</sup>

<sup>g</sup> ND: not determined.

<sup>h</sup> TE: tuberculate eggs typical for microtubule destabilisers.

<sup>i</sup> NA: not tested, insoluble in 96% ethanol.

### 3.2.2.2 Cytotoxicity in NCI60 human cancer cell line panel

In NCI60 screen derivatives **1\_21**, **22**, **24**, and **25** at 10  $\mu\text{M}$  concentration inhibited cancer cell growth by  $\geq 70\%$  (Table **3.1**). Microtubule-destabilising derivatives **1\_34** and **35** also demonstrated pronounced cytotoxicity ( $\text{GI}_{50} = 0.28$  and  $3.24 \mu\text{M}$ , respectively), whereas *p*-substituted derivatives **1\_27**, **29**, **31** to **33** showed weaker cell growth inhibition as previously reported.<sup>5</sup> Bicyclic AThPs **1\_40** to **56** exhibited minimal microtubule modulation *in vivo* along with barely detectable cytotoxic properties. Independent studies of **1\_28** using *in vitro* (bovine) tubulin polymerisation inhibition assay yielded an  $\text{IC}_{50}$  value of  $13.28 \mu\text{M}$ , however the molecule did not cause cell cycle arrest in 3T3 cells.

In general, sea urchin embryo results for **1\_21**, **22**, **24** and **25** were in good correlation with the NCI60 data reported by Arabshahi *et al.*<sup>99</sup>, *i.e.*, human MDA-MB-435 melanoma and MDA-MB-468 breast cancer cell lines were found to be particularly sensitive to the active series and the pertinent  $\text{GI}_{50}$  values are given in Table **3.2**. Increased conformational flexibility by cycloalkyl ring expansion showed apparent increase in microtubule-destabilising activity and cell growth inhibition (molecules **1\_21** *vs.* **23** and **1\_22** *vs.* **24**). However, derivative **1\_21** showed the strongest effect against melanoma MDA-MB-435 cell line with  $\text{GI}_{50} = 52 \text{ nM}$  (Table **3.2**) but a relatively modest activity in the sea urchin embryo assay compared to other derivatives (*e.g.*, **1\_22** to **24**). This effect was more pronounced for derivative **1\_25**, which displayed the strongest general cancer growth inhibition with no potency in the sea urchin assay suggesting an alternative non-tubulin mediated mechanism of action. A similar outcome was observed for derivative **1\_50** ( $\text{R}_1 = 2\text{-methoxyphenyl}$ ,  $\text{R}_2 = 3, 4\text{-dimethoxyphenyl}$ ), it demonstrated a robust cytotoxic effect ( $\text{GI}_{50} = 5.13 \mu\text{M}$ ) with no activity seen in the sea urchin embryo assay. These data suggests that specific anti-tubulin

effect for the AThPs series is dependent on the substitution patterns for both the core and the phenyl pharmacophore.

**Table 3.2.** Effects of AThPs **1\_21**, **22**, **24** and **25** on MDA-MB-435 human melanoma and MDA-MB-468 breast cancer cell lines.

Compound	Mean GI, % <sup>a</sup>	Mean GI <sub>50</sub> (nM) <sup>b</sup>	
		MDA-MB-435 <sup>c</sup>	MDA-MB-468 <sup>c</sup>
<b>1_21</b>	23.6	52	197
<b>1_22</b>	30.3	172	332
<b>1_24</b>	14.7	205	254
<b>1_25</b>	10.0	2530	394

<sup>a</sup> GI %: single dose inhibition of cell growth at 10  $\mu$ M concentration.

<sup>b</sup> GI<sub>50</sub>: concentration required for 50% cell growth inhibition.

<sup>c</sup> Data from Ref.<sup>99</sup>

### 3.2.3 Molecular Modelling

Results from the sea urchin embryo assay are clear implications of microtubule destabilisation that displayed morphological changes similar to other microtubule-targeted destabilising agents *e.g.*, phenstatin and nocodazole from Table **3.1**, colchicine derivatives, and *Vinca* alkaloids. To elucidate the most likely binding sites on tubulin, molecular modelling was employed.

#### 3.2.3.1 Reliability and Reproducibility test

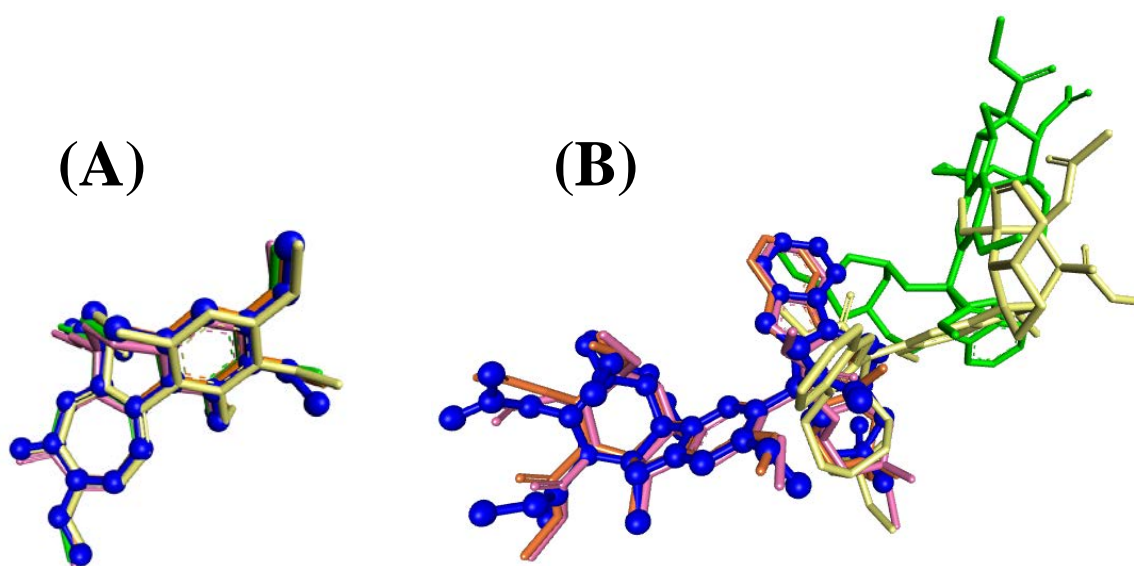
The software package GOLD was first tested for its reliability in order to reproduce the experimental binding modes. Four scoring functions were used, abbreviated as GS, CS, ChemPLP and ASP to strengthen consensual relationships. The co-crystallised ligands were removed and re-docked into their respective binding sites. The RMSDs between the co-crystallised ligand and the docked configurations were calculated (Table **3.3**). The docking

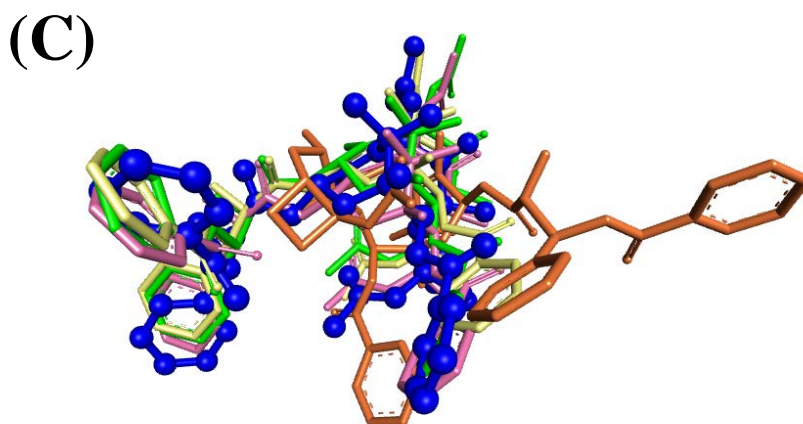


overlays are shown in Fig 3.5. Good RMSDs resulted from colchicine (0.394 Å) and paclitaxel (1.625 Å). However, only GS showed adequate RMSD = 1.307 Å for vinblastine (Fig. 3.5C). This may be associated with the fact that it is more unlikely for more complex and flexible structures to have the same pose for every docking run. Furthermore, the adequacy of the GS scoring function suggested it is the most reliable scoring function when docking into the *Vinca* site. Following this, all forty derivatives were docked into the three binding sites namely colchicine, *Vinca*, and taxol sites.

**Table 3.3.** RMSD values in Å calculated between the co-crystallised ligands and its re-docked configurations. Four scoring functions were used: GS, CS, ChemPLP and ASP.

Co-crystallised ligand	GS	CS	ChemPLP	ASP	Average RMSD (Å)
Colchicine	0.481	0.362	0.403	0.330	0.394
Vinblastine	1.307	12.000	12.547	6.020	7.969
Paclitaxel	1.090	1.667	1.684	2.060	1.625





**Figure 3.5.** Overlays of best scoring re-docked configurations. The co-crystallised ligands; colchicine (**A**), vinblastine (**B**) and paclitaxel (**C**) are depicted using ball and stick display style coloured blue. The other colours are representatives of each of the scoring functions; pink (GS), yellow (CS), green (ChemPLP) and orange (ASP). The hydrogen atoms of the ligands were suppressed.

### 3.2.3.2 Docking to colchicine, *Vinca*, and taxol sites

As mentioned before, the colchicine is regarded to be the most simple binding domain compared to the complexity of the *Vinca* and taxol sites. This is important for understanding the link behind the development of several colchicine site-targeted small molecules with destabilising effects. The structures of the *Vinca* and taxol sites are complicated with multiple sub-pockets and large cavity size favourable for large complex inhibitors. It should be noted that the taxol site is known to exert stabilising effects when inhibited, making it the most unlikely site for AThP inhibition. So, it can be assumed that the AThPs are most likely targeting the colchicine site. Derivatives **1\_21** to **60** were docked to all three sites and the scores are shown in Table **3.4**.

The *Vinca* and taxol sites showed similar scores across all four scoring functions. The colchicine site on the other hand clearly showed that GS and ChemPLP scored much higher than both the *Vinca* and taxol sites, whereas CS and ASP have similar scores. The scores are roughly 55 to 73 (GS) and 56 to 78 (ChemPLP) for the colchicine site with mostly concentrated in the mid and higher 60's. The *Vinca* site scored roughly within the range of 29

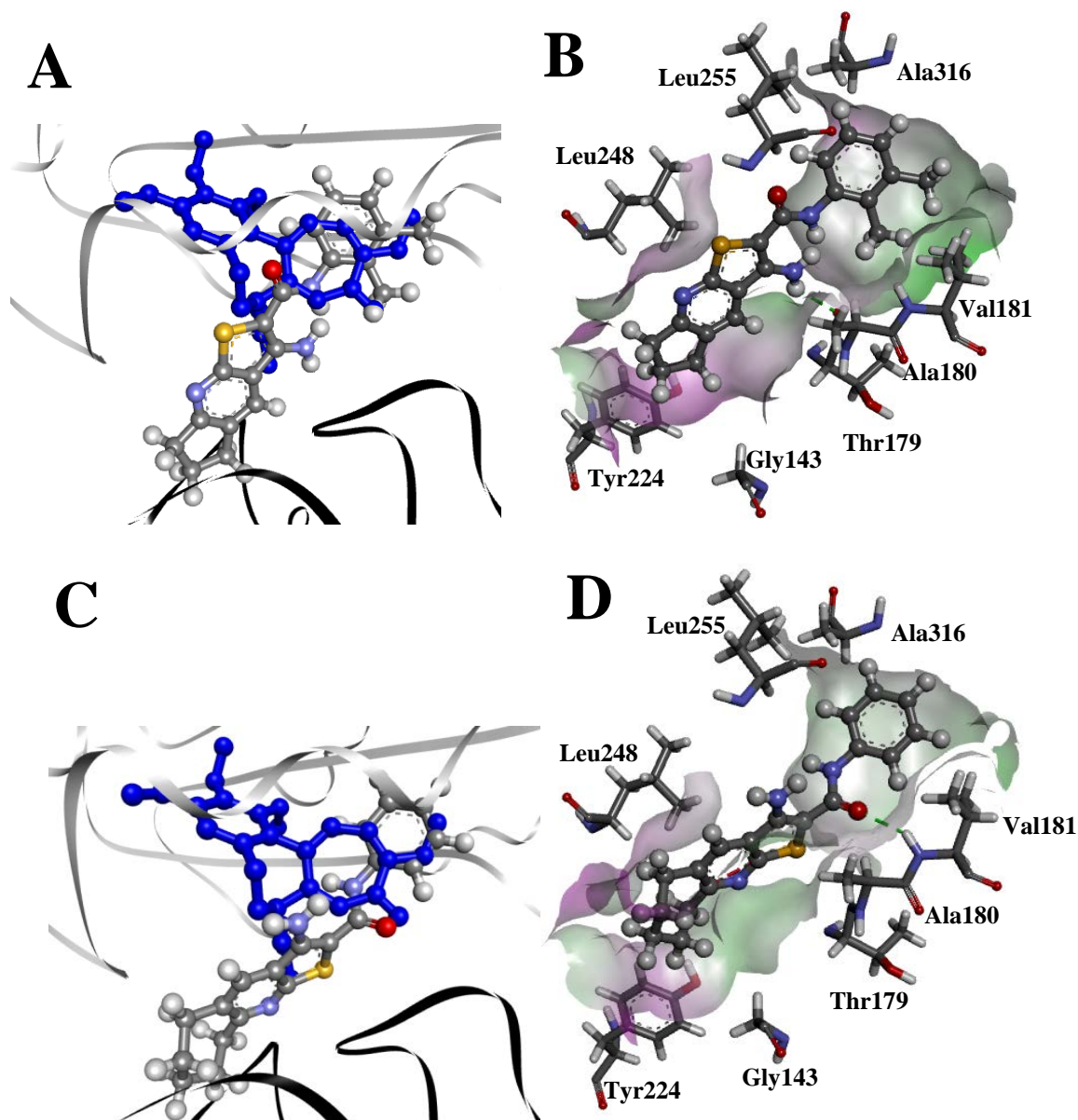
to 60 (GS) and 46 to 62 (ChemPLP) with most scoring in the lower and mid 50's. The stabilising taxol site scored around 47 to 65 (GS) and 46 to 70 (ChemPLP) with most scoring closely to the *Vinca* site around the lower and mid 50's. Thus, the score results indicated a close consensual relationship between GS and ChemPLP scoring functions suggesting the colchicine site as the most probable site for AThP inhibition.

In order to understand the binding mechanisms of the AThPs, the docked configurations to the colchicine site were studied for series **A**. Modelling studies were exemplified using compounds **1\_21** and **24** (Fig. 3.6). Hydrogen bonding with amino acid residues Val181 and Thr179 were observed. Additionally, hydrophobic contacts between the phenyl ring and lipophilic regions inside the  $\beta$ -subunit cavity where the side chains of hydrophobic amino acid residues Ala316, Leu255 and Val181 form a cavity. The cavity is known to be occupied by the heptalen-7-yl ring on the colchicine (Fig. 3.6A and C). The cycloalkyl moiety is predicted to be oriented towards a large hydrophobic area within the  $\alpha/\beta$  interface facilitating lipophilic contacts with amino acid residues Leu248, Tyr224, Gly143 and Ala180. This is an important observation as it established a relationship between ring size and biological activity, *i.e.*, a five-membered (**1\_21**) to a seven-membered ring (**1\_24**) size is likely to facilitate increased van der Waal's interactions within the hydrophobic pocket.

**Table 3.4.** Docked scores of colchicine, vinblastine, paclitaxel and AThP derivatives to the main three binding sites in tubulin monomer.

Compound	Colchicine				Vinca				Taxol			
	Goldscore	ChemScore	ChemPLP	ASP	Goldscore	ChemScore	ChemPLP	ASP	Goldscore	ChemScore	ChemPLP	ASP
Colchicine	83.6	31.1	83.0	35.1	-	-	-	-	-	-	-	-
Vinblastine	-	-	-	-	70.0	25.4	57.6	28.1	-	-	-	-
Paclitaxel	-	-	-	-	-	-	-	-	81.6	27.4	86.8	33.6
1_21	61.7	27.4	63.8	24.4	51.5	25.4	54.8	26.8	55.2	29.0	51.8	23.0
1_22	60.6	29.8	67.1	27.6	51.2	25.9	51.8	27.9	49.8	27.9	52.7	23.5
1_23	61.0	28.3	62.8	22.7	53.6	26.1	56.3	27.2	55.4	31.0	52.1	23.0
1_24	62.5	25.4	70.8	25.7	49.6	26.4	54.1	26.7	50.8	29.1	49.1	23.5
1_25	55.5	26.7	62.6	25.8	51.4	26.2	62.7	31.6	52.7	30.3	60.5	29.2
1_26	66.0	31.5	76.3	25.9	49.3	28.1	49.6	26.3	51.9	31.0	56.8	25.3
1_27	67.3	31.4	73.7	25.7	52.9	28.3	54.4	27.6	51.0	32.2	53.8	25.3
1_28	64.3	31.2	73.1	26.1	48.7	28.3	52.4	26.3	48.5	30.0	56.1	24.5
1_29	66.9	33.4	76.7	26.0	29.2	30.1	50.5	24.0	52.0	32.1	50.9	25.0
1_30	65.5	26.0	62.1	28.7	51.3	27.0	51.6	28.2	53.7	27.8	55.8	26.3
1_31	60.3	31.3	77.8	31.1	51.1	27.1	50.6	25.6	52.1	31.3	48.9	24.4
1_32	67.2	31.7	62.1	30.3	47.7	29.2	49.3	25.3	48.8	28.2	50.7	24.5
1_33	67.4	33.0	62.4	25.8	48.3	27.2	50.3	26.1	53.5	30.4	45.9	23.5
1_34	60.2	31.1	70.5	25.0	49.9	29.9	51.0	25.2	51.9	32.4	51.5	26.0
1_35	60.4	23.1	56.3	23.2	53.8	25.2	52.4	28.6	55.0	30.3	54.6	26.5
1_36	71.4	30.9	65.9	29.2	60.2	28.4	63.3	29.9	65.3	35.6	70.2	36.3
1_37	71.1	27.5	63.8	30.1	58.7	27.6	67.4	30.3	63.3	34.4	67.2	36.5
1_38	65.7	30.0	63.1	29.0	51.0	28.5	52.4	27.7	53.0	29.8	55.6	25.9
1_39	65.5	28.8	74.8	34.0	50.0	26.4	57.1	30.4	53.7	27.6	57.9	28.5
1_40	62.2	27.8	63.9	26.4	45.9	25.3	50.0	25.1	47.7	28.0	48.3	25.9
1_41	62.2	27.8	67.1	28.0	48.9	25.6	47.6	24.7	48.5	28.3	50.0	25.3
1_42	66.6	25.7	67.5	27.3	46.5	24.9	46.0	24.1	49.1	28.2	50.8	26.7
1_43	67.8	29.6	56.3	24.6	49.4	26.0	51.3	24.2	49.4	29.4	52.8	24.9
1_44	69.0	25.0	65.2	24.7	52.1	21.5	49.8	27.0	55.4	26.1	51.8	26.0
1_45	73.0	31.5	68.9	28.1	50.2	28.2	50.8	27.6	50.8	29.0	49.1	26.2
1_46	57.7	24.0	56.6	28.5	48.5	25.2	50.1	28.1	48.2	28.2	52.1	28.2

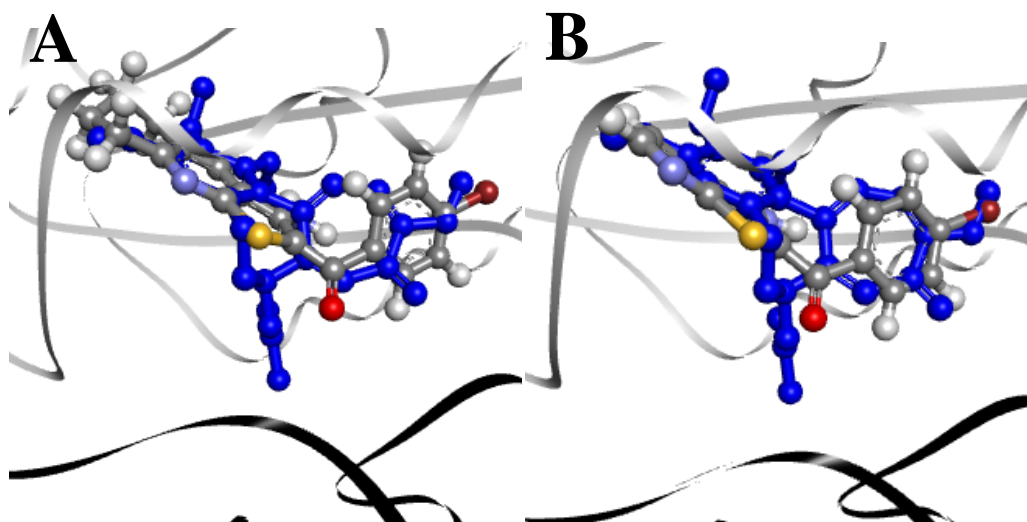
<b>1_47</b>	61.7	24.4	67.6	30.6	53.5	25.1	51.7	26.2	49.7	27.3	51.4	26.5
<b>1_48</b>	59.1	20.5	50.6	28.4	58.5	21.9	53.7	28.1	50.9	25.5	50.4	26.9
<b>1_49</b>	68.7	26.3	60.0	29.5	59.1	25.9	59.1	30.7	61.3	29.6	64.6	32.9
<b>1_50</b>	68.0	30.1	61.4	30.2	58.3	27.4	62.1	31.3	64.6	28.5	59.9	30.4
<b>1_51</b>	61.6	25.5	50.6	27.3	53.3	25.0	51.9	30.1	68.4	27.5	59.9	32.0
<b>1_52</b>	65.7	28.6	52.1	29.1	55.6	28.8	54.8	30.6	58.7	30.8	60.0	31.6
<b>1_53</b>	70.6	26.1	57.0	27.3	59.4	26.7	54.8	28.0	59.6	27.4	59.2	29.0
<b>1_54</b>	73.0	28.3	59.6	27.3	58.2	27.9	57.6	26.7	57.6	30.1	56.2	30.1
<b>1_55</b>	68.5	24.8	62.4	28.8	58.6	25.3	56.8	29.2	61.2	27.4	59.3	29.1
<b>1_56</b>	76.6	25.9	64.4	34.3	67.1	25.4	71.3	35.8	66.0	29.2	71.3	38.5
<b>1_57</b>	52.7	30.9	66.9	23.5	48.9	29.6	48.7	27.8	46.9	30.1	49.7	25.8
<b>1_58</b>	59.1	25.9	65.1	23.0	52.9	25.5	52.9	25.3	53.6	28.7	50.8	25.8
<b>1_59</b>	55.6	24.6	51.0	21.8	53.2	24.3	53.2	27.4	50.7	27.2	54.0	26.2
<b>1_60</b>	61.3	29.7	69.1	28.7	51.4	26.8	48.1	27.2	49.8	29.8	54.6	25.5



**Figure 3.6.** Binding poses of compounds **1\_21** (**A** and **B**) and **1\_24** (**C** and **D**). Colchicine (blue) is overlain with the ligands (**A** and **B**). The grey and black polypeptide ribbons belong to the  $\beta$ -chain and  $\alpha$ -chain respectively. Green dotted lines represent hydrogen bonding interactions. The binding site of the protein is rendered where grey is neutral and green and pink represent hydrogen bond acceptor and donor groups.

Additionally, the series **B** tricyclic and bicyclic AThPs with ketone linkers were modelled. Results revealed that there is substantially more overlapping with the colchicine ligand and less interactions with the  $\alpha/\beta$ -interface and more occupation in the  $\beta$ -subunit cavity for both series **B** tricyclic and bicyclic AThPs exemplified by derivatives **1\_29** and **44** respectively (Fig. 3.7). Thus, it is suggested that cycloalkyl interactions with the  $\beta$ -subunit cavity are

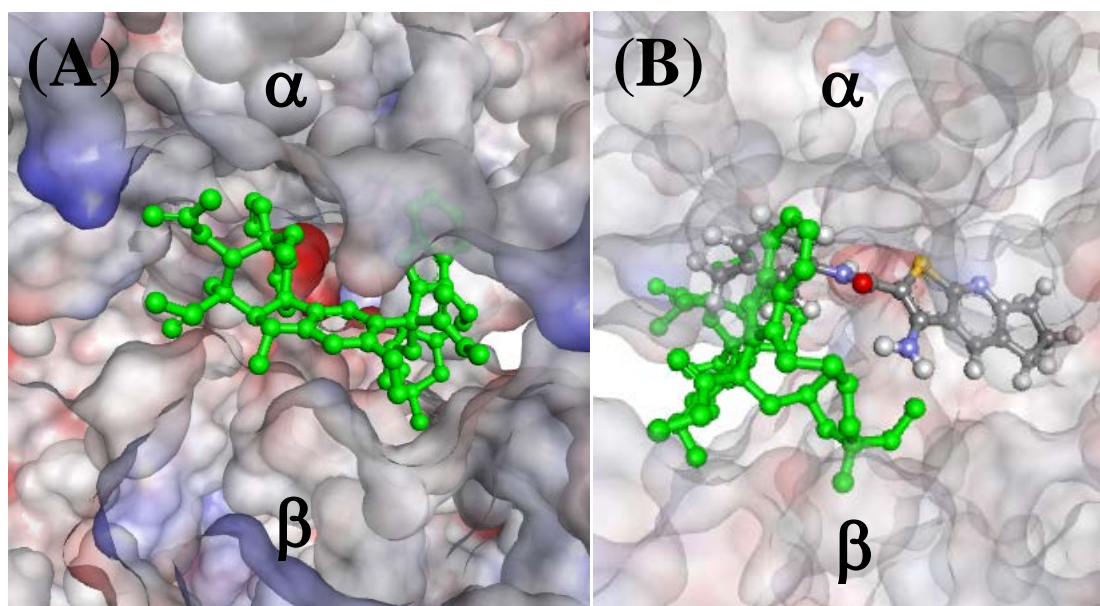
unfavourable to increases in destabilising effects. This model is supported by the biological data presented in Table 3.1.



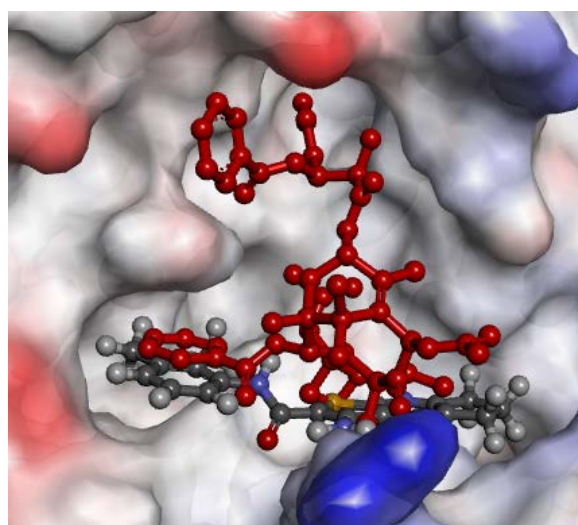
**Figure 3.7.** Binding poses of series **B** AThP derivatives **1\_29** (**A**) and **44** (**B**). Colchicine (blue) is overlain with the compounds. The grey and black polypeptide ribbons belong to the  $\beta$ -chain and  $\alpha$ -chain respectively.

The second binding site on tubulin is the *Vinca* site located at the  $\alpha\beta$ -interface between two tubulin molecules. Inhibition of the site is known to cause destabilising effects, thus leaving speculation as a possible location for AThP inhibition. Structural visualisation showed a large binding site with multiple grooves and pockets suitable for larger flexible molecules. To further clarify, shown in Fig. 3.8 are illustrations of vinblastine, a known *Vinca* site inhibitor and derivative **1\_21** docked to the *Vinca* site. The tricyclic core clearly lacked any form of overlays with vinblastine which strongly suggests that the AThPs are unlikely to interact with the *Vinca* site in a similar fashion to vinblastine and other *Vinca* site inhibitors.





**Figure 3.8.** Binding poses of vinblastine co-crystallised ligand (green) (A) and **1\_21** (B) in the *Vinca* site. The protein surface is rendered and  $\alpha$ - and  $\beta$ -subunits labeled. Partially positive, negative and neutral regions are depicted as blue, red and grey areas.



**Figure 3.9.** Binding poses of paclitaxel co-crystallised ligand (red) and derivative **1\_21**. The protein surface is rendered. Partially positive, negative and neutral regions are depicted as blue, red and grey areas.

A similar case appeared for the taxol site, the binding site is relatively large with multiple pockets and clefts suitable for large flexible complex molecules (Fig. 3.9). The site is known for their stabilisation effects when inhibited, which did not correspond to the destabilising



effects demonstrated by the AThPs. As shown in Fig. 3.9, derivative **1\_21** cannot accommodate all the binding pockets and molecular interactions as thoroughly as paclitaxel. This explained the reason for a consistent drop in scores for both the GS and ChemPLP (Table 3.4) when comparing the colchicine and taxol sites, a similar case can be attributed to the *Vinca* site.

## 3.3 Methodology

### 3.3.1 Molecular Modelling

The compounds were docked to the crystal structures of tubulin with their respective PDB IDs: 4O2B (resolution 2.3 Å),<sup>100</sup> 4EB6 (resolution 3.47 Å)<sup>101</sup> and 1JFF (resolution 3.5 Å).<sup>90</sup> The Scigress v2.6<sup>79</sup> was used to prepare the crystal structure for docking, *i.e.* hydrogen atoms were added, the co-crystallised ligands were removed as well as crystallographic water molecules. The Scigress software suite was also used to build the chemical structures and were optimised using the MM2<sup>80</sup> force field. The centre of the binding were defined in crystal structures: 4O2B as C3 on colchicine (x = 13.222, y = 8.371, z = -23.331), 4EB6 as C58 on vinblastine (x = 13.391, y = 90.610, z = 103.739) and 1JFF as O11 on paclitaxel (x = 1.403, y = -16.979, z = 16.391) with 10 Å radius. Fifty docking runs were allowed for each ligand with default search efficiency (100%). The basic amino acids lysine and arginine were defined as protonated. Furthermore, aspartic and glutamic acids were assumed to be deprotonated. The GoldScore (GS),<sup>18-19</sup> ChemScore (CS),<sup>22-23</sup> ChemPLP<sup>22, 24</sup> and ASP<sup>25</sup> scoring functions were implemented to validate the predicted binding modes and relative energies of the ligands using the GOLD v5.4.0 software suite.

### 3.3.2 NCI60 assay and Phenotypic sea urchin embryo assay

Compounds were verified for their anticancer activity in the NCI60 assay by the National Cancer Institute, Maryland in the United States. For general method see Ref.<sup>8</sup>

The phenotypic sea urchin embryo assay was conducted by our collaborators, Prof. Victor Semenov, Dr. Marina Semenova, Leonid Konyushkin, Olga Atamanenko and Prof. Alex Kiselyov at the N. D. Zelinsky Institute of Organic Chemistry and N. K. Kol'tsov Institute of Developmental Biology in Moscow, Russian Federation. The protocol is outlined in Ref.<sup>97</sup>

## 3.4 Future Work and Conclusions

In this work, it was shown that tubulin destabilisation is one of the main biological mechanisms of action for the AThPs based on the sea urchin embryo data. Molecular modelling determined the colchicine site to be the most likely inhibition site for anti-tubulin activity. Inevitably, to confirm this would need to employ biochemical assays that can show evidences of AThP inhibition, *e.g.*, using radioactive [<sup>3</sup>H]-colchicine as competing substrates has been suggested.<sup>73</sup> Derivatives **1\_21** to **24** have low nano-molar activities but the rest of the compounds are less active. A carboxamide linker between the thienopyridine core and the phenyl group is crucial for the activity, which can be rationalised with enhanced lipophilic contact at the  $\alpha/\beta$ -interface in tubulin compared to the derivatives containing a ketone linker according to molecular modelling to the colchicine site. The potency in the sea urchin assay is reflected in the NCI60 human tumour panel with some notable exceptions that suggested that this class of compounds has a number of modes of action, *e.g.*, derivative **1\_25**. It is clear that more work is required in order to identify all of the biological targets of the AThPs but the results presented here substantially advances our insight into their anticancer mechanisms of action.

# *Chapter 4*

Identification of novel phosphatidylcholine – specific phospholipase C inhibitors using virtual high throughput screening

## 4.1 Introduction

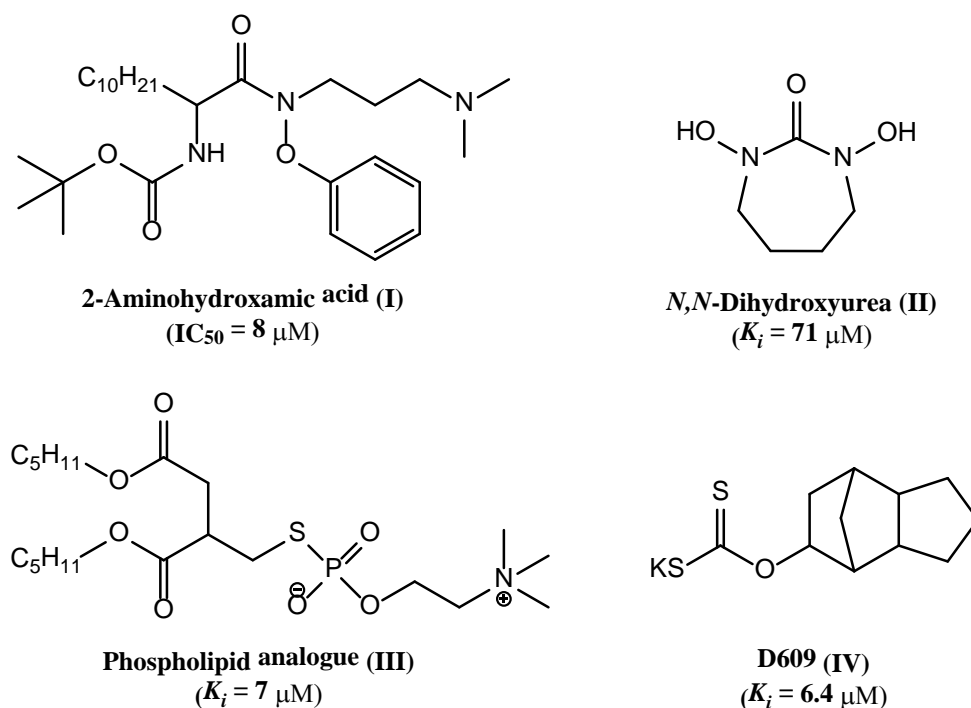
### 4.1.1 Overview of phosphatidylcholine – specific phospholipase C

PC-PLC is a type of phospholipase C mediator that preferentially hydrolyses phosphatidylcholine (PC) phospholipids into secondary messengers DAG and phosphocholine.<sup>102-103</sup> DAG is known for its involvement in activating PKC, a kinase that phosphorylates downstream proteins in the signal transduction cascade.<sup>42, 104</sup> The role of phosphocholine in the signaling pathway is not entirely clear other than it is required for the synthesis of essential lipids for signal transduction.<sup>105</sup> However, there is a strong relationship between high phosphocholine levels and cell proliferation activity, particularly in cancer tumours making it a useful biomarker for monitoring cell proliferation.<sup>106-107</sup>

The progression in understanding the mechanisms, structures and functions of mammalian PI-PLCs have been well-developed. Unfortunately, this is not the same for mammalian PC-PLCs. Only the prokaryotic PC-PLCs have been purified and its structure characterised in detail, the most well-known was extracted from *Bacillus cereus* (PC-PLC<sub>Bc</sub>).<sup>108</sup> Initially, the idea of using PC-PLC<sub>Bc</sub> emerged from using bacteria from the *Bacillus* genus to screen for the hydrolysis of PCs that revealed *Bacillus cereus* to be the most responsive strain.<sup>109</sup> Thus, PC-PLC<sub>Bc</sub> is used as a model for studying the structure and functions of mammalian PC-PLCs. The approach is justifiable as both mammalian and prokaryotic PC-PLCs were reported to share some similar features, *e.g.*, antigenic similarities<sup>110</sup> and able to mimic similar responses such as enhancing prostaglandin biosynthesis.<sup>111</sup>

### 4.1.2 PC-PLC in cancer progression

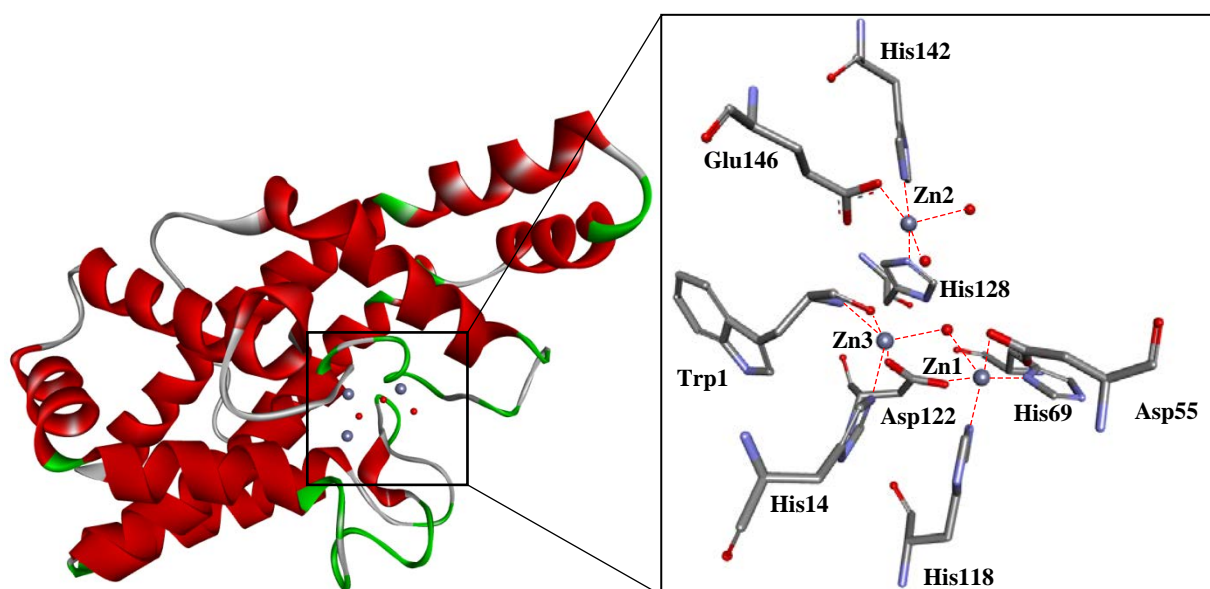
Similar to the PI-PLCs, PC-PLCs hydrolyse phospholipid substrates to form bioactive signalling intermediates that further activate an array of cellular events, which if unregulated can lead to cancer.<sup>104</sup> Due to a lack of results on mammalian PC-PLCs, specific information regarding their cellular roles, structures and isotype variations are limited. Nonetheless, a growing body of evidence linking PC-PLC with cancers have been established, *e.g.*, increased activation of PC-PLC in MDA-MB-231 breast<sup>112</sup> and epithelial ovarian cancers,<sup>113</sup> and it was implicated to be involved in oncogenic mitogen-activated protein kinase activation.<sup>114</sup> Only a handful of PC-PLC inhibitors are known which are not considered to be drug-like, *e.g.*, 2-aminohydroxamic acids,<sup>115</sup> univalent anions,<sup>116</sup> *N,N'*-dihydroxyureas,<sup>117</sup> phospholipid analogues<sup>118</sup> and the acclaimed xanthate, D609 (see molecular structures in Fig. 4.1).<sup>119-120</sup> Therefore, the data available suggests that PC-PLC is a novel target for anticancer drug discovery with novel drug-like structures pending to be identified.



**Figure 4.1.** Examples of known PC-PLC inhibitors (I - IV) and their measured inhibition activities.

### 4.1.3 Binding site of PC-PLC<sub>Bc</sub>

It is worth studying the binding site of PC-PLC<sub>Bc</sub> to gain the structural insights of the enzyme. The binding site is buried in a hydrophobic region surrounded by  $\alpha$ -helical chains and loops connecting the helices (see Fig. 4.2).<sup>108</sup> The active site is embedded with a tri-metal centre consisting of three Zn<sup>2+</sup> ions.<sup>103</sup> The tri-metal centre are penta-coordinated to amino acids and three water molecules that are proposed to play central roles in the catalytic cleavage of PC substrates.<sup>103, 108, 121</sup> Zn1 is bridged to Zn3 by a water molecule and ligates to residues Asp55, His69, His118 and Asp122.<sup>108</sup> Zn2 is ligated to two water molecules and amino acid residues His128, Glu146 and His142.<sup>108</sup> Zn3 is ligated to amino acids Trp1, His14 and Asp122.<sup>108</sup>



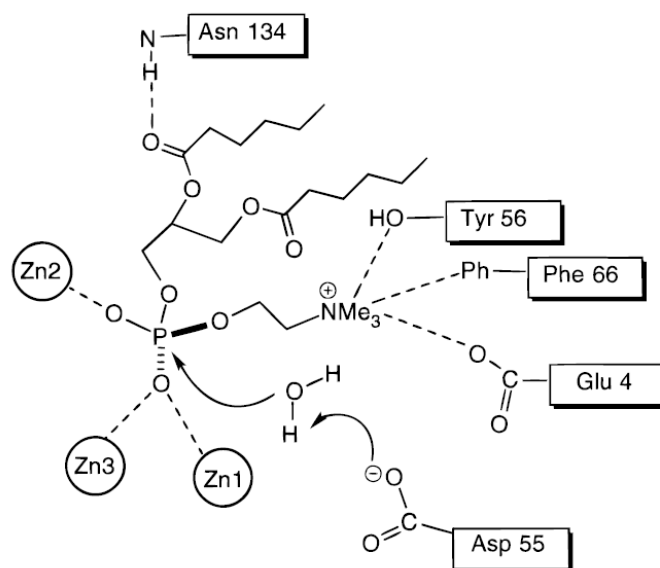
**Figure 4.2.** The structure of wild-type PC-PLC<sub>Bc</sub> enzyme (PDBid: 1AH7).<sup>108</sup> The phospholipase monomer is displayed as red ribbons embedded with three catalytic Zn<sup>2+</sup> ions and water molecules depicted as grey and red spheres, respectively (left). Coordinating water molecules and amino acids to the Zn<sup>2+</sup> ions represented as sticks are coloured by atom types (right).

The substrate binding site of PC-PLC<sub>Bc</sub> are accommodated by other residues that includes Ala3, Glu4, Tyr56, Ser64, Thr65, Phe66, Phe70 and Asn134.<sup>121</sup> Additionally to the tri-metal centre, these residues can accommodate different parts of the PC substrate; Glu4, Tyr56 and Phe66 have been implicated to accommodate the choline head group,<sup>103</sup> whereas it is difficult to

identify the residues involved for lipophilic contacts with the acyl chains.<sup>122</sup> This was partly due to the conformational flexibility and disorder of the long acyl chains which resulted in difficulties in measuring the intermolecular distances between the amino acid residues and the acyl chains.<sup>122</sup> Nevertheless, the basis of favourable hydrophobic – hydrophobic contacts can be assumed as a general guide between the PC substrate and the hydrophobic residues.

#### 4.1.4 Catalytic cleavage mechanism of PC-PLC<sub>Bc</sub>

It was confirmed that the phosphate group coordinate to the Zn<sup>2+</sup> ions that can result in the displacement of water molecules from the tri-metal centre.<sup>122</sup> Consequently, this leads to the phosphate group to become more electrophilic or ‘*activated*’.<sup>103</sup> An amino acid residue can act as a base and abstract a proton from a water molecule. Several amino acid residues have been proposed as bases owing to their short intermolecular distances from nucleophilic water: Glu4<sup>123</sup> and Asp55.<sup>123-124</sup> A mechanism proposed using Asp55 as the general base is shown in Fig. 4.3. Abstracting a proton from water activates a nucleophilic hydroxide ion to attack on the phosphodiester bond that resulted in two separate products: phosphocholine and DAG alkoxide. A proton donor is required to protonate DAG alkoxide. It was proposed that Asp55 can also act as the proton donor.<sup>124</sup> Although, the mentioned residues play important mechanistic roles in catalysis, the roles of the histidines should not be overlooked. The side chain of histidine is an imidazolium heterocycle that has an estimated pK<sub>a</sub> value ~7 near physiological pH, which suits histidine to function as both an acid and a base.<sup>125</sup>



**Figure 4.3.** A basic outline of the cleavage mechanism by the PC-PLC<sub>Bc</sub> enzyme. (Diagram obtained from Hergenrother *et al.*<sup>103</sup>)

## 4.2 Results and Discussion

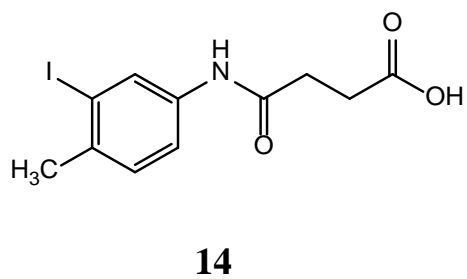
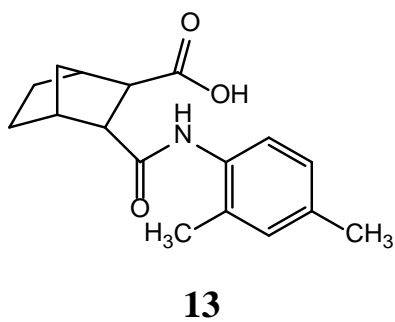
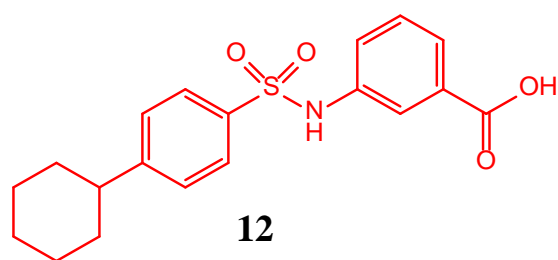
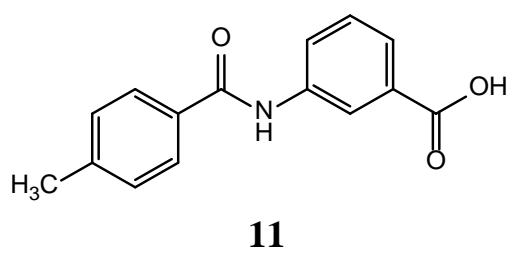
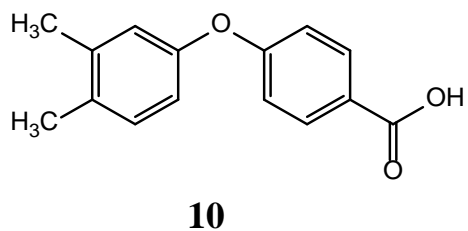
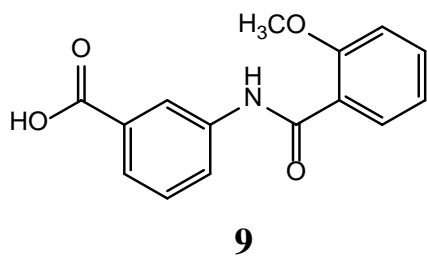
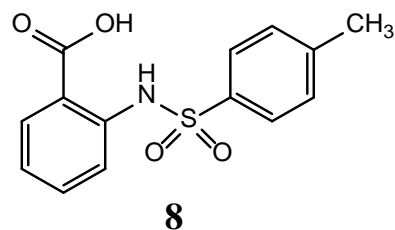
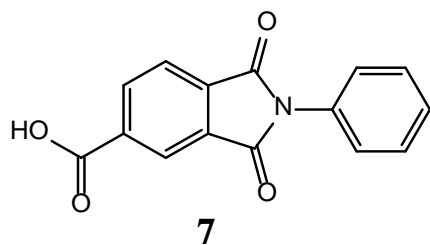
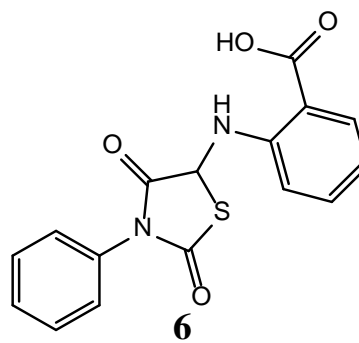
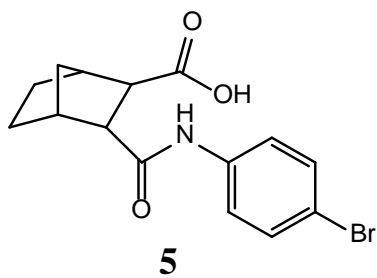
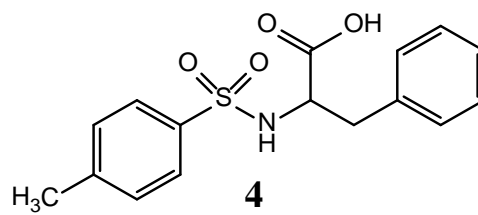
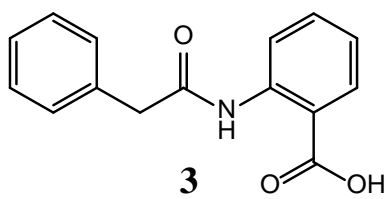
The aim of this project is to search for novel inhibitors of mammalian PC-PLC using the PC-PLC<sub>Bc</sub> crystal structure as a model. The vHTS methodology was used due to its ability to identify inhibitors of various drug targets (see introduction section 1.3).

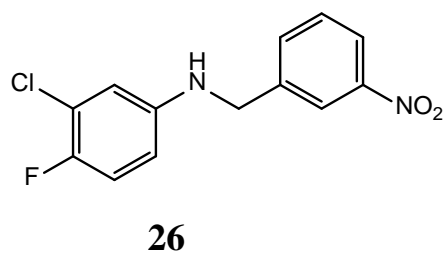
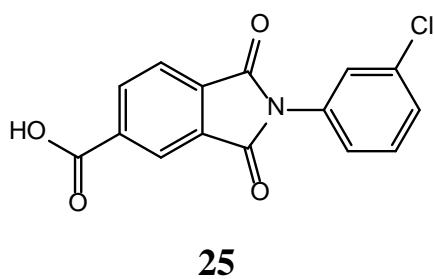
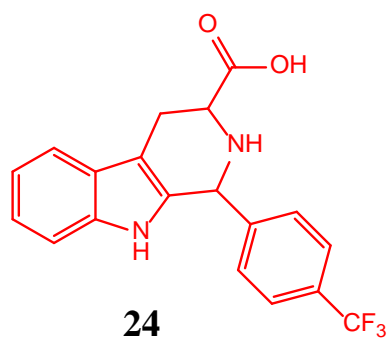
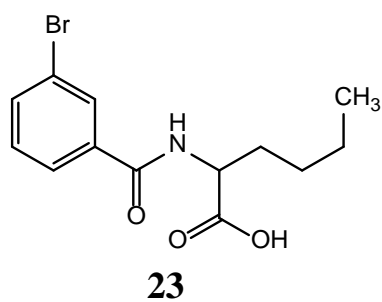
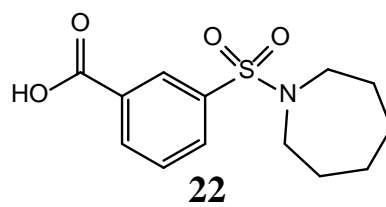
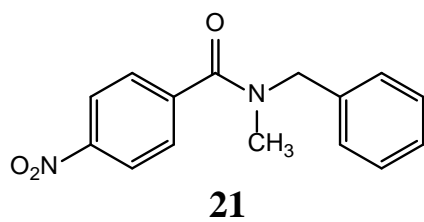
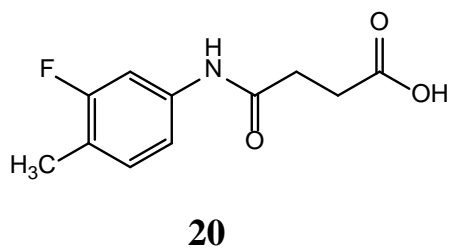
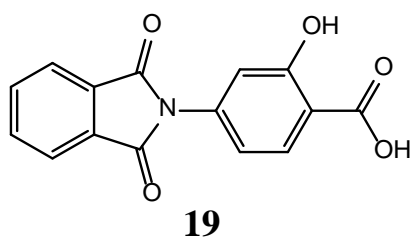
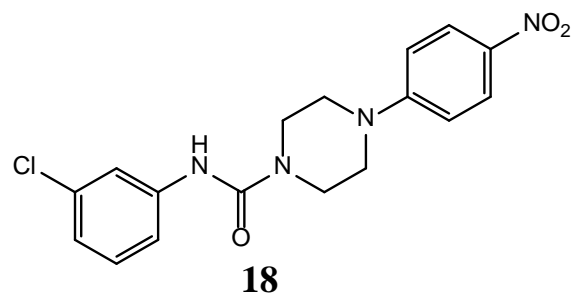
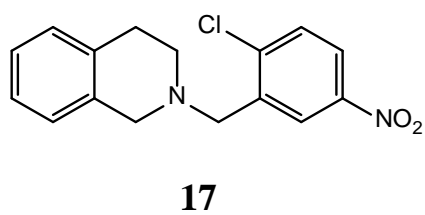
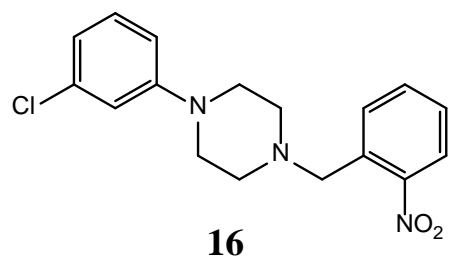
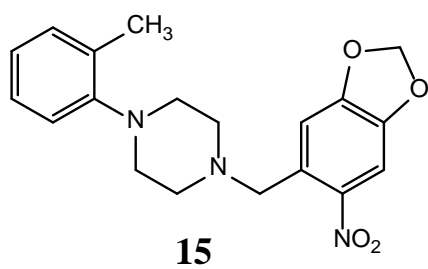
### 4.2.1 Pilot Screen

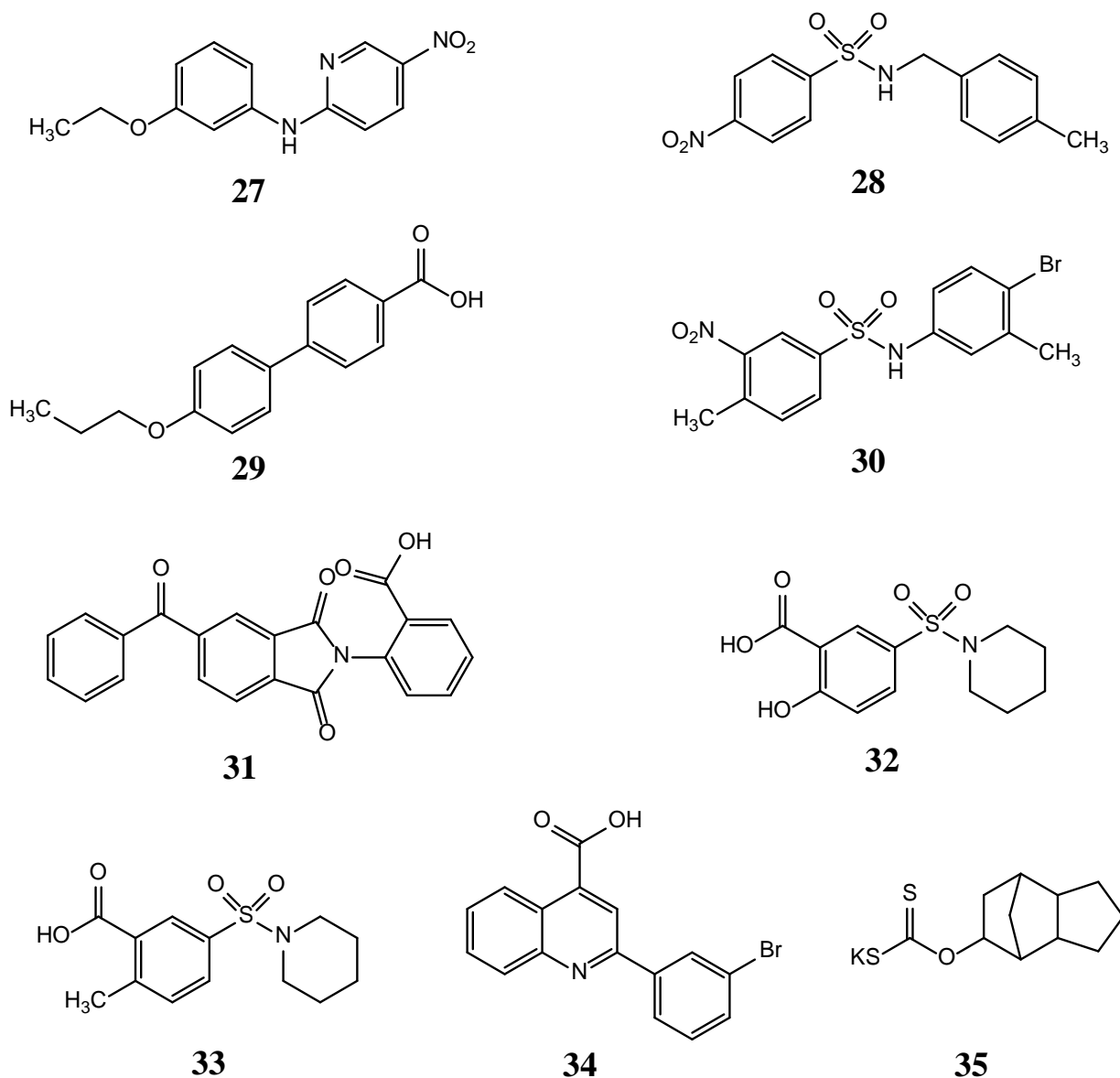
$1 \times 10^4$  molecular entities were downloaded from the ChemBridge diversity collection.<sup>126</sup> It was filtered based on Lipinski's rules<sup>29</sup> resulting in 8373 drug-like molecules. The docking scaffold used was the wild-type PC-PLC<sub>Bc</sub> crystal structure (structure shown in Fig. 4.2). Zn3 was selected as the centre of binding as it was reported that the Zn<sup>2+</sup> ions play important roles for catalytic cleavage by binding to charged phosphate groups on the phospholipid substrates.<sup>108, 117, 121</sup> Four scoring functions, GoldScore (GS),<sup>18-19</sup> ChemScore (CS),<sup>22-23</sup>



ChemPLP<sup>22, 24</sup> and ASP<sup>25</sup> were implemented at 30% efficiency. Molecules without hydrogen bonding (HB = 0) and those that have scores GS < 55, CS < 26, ChemPLP < 65 and ASP < 25 were eliminated. The initial elimination process resulted in 1721 molecules. These were re-docked at a higher efficiency (100%), *i.e.*, higher quality pose of the ligand is predicted. Molecules with HB = 0, GS < 60, CS < 34, ChemPLP < 65 and ASP < 28 were further eliminated. Additionally, molecules with undesirable reactive, cytotoxic and chemically unstable moieties suggested to produce false positives were removed (see introduction section **1.3**, Fig. **1.2**).<sup>17</sup> The remaining molecules were checked for the best predicted poses between the scoring functions and whether a plausible binding mode was predicted, *e.g.*, chemical groups directed towards Zn<sup>2+</sup> ions to form electrostatic interactions and no lipophilic moieties facing the aqueous environment. In total, thirty-two compounds (see molecular structures Fig. **4.4**) were purchased from the ChemBridge Corporation Online chemical store for experimental testing.<sup>127</sup>



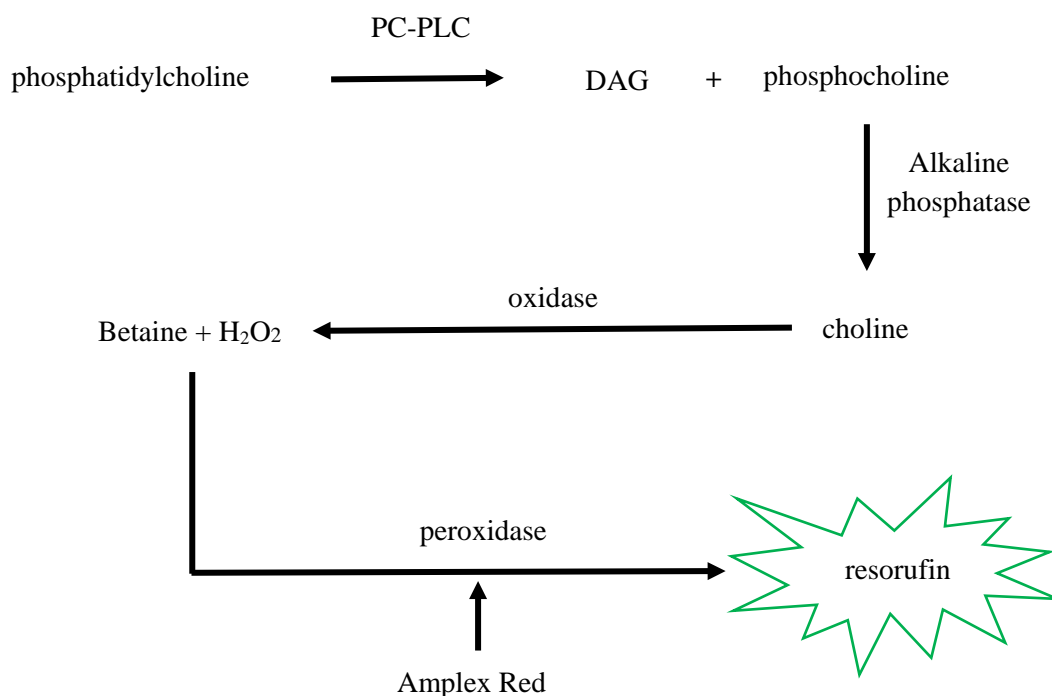




**Figure 4.4.** Molecular structures of thirty two compounds tested with the Amplex Red biochemical assay for activity against PC-PLC. The known inhibitor is D609 (**35**) was used as a positive control.<sup>119-120</sup> Active hits **12** and **24** are coloured red.

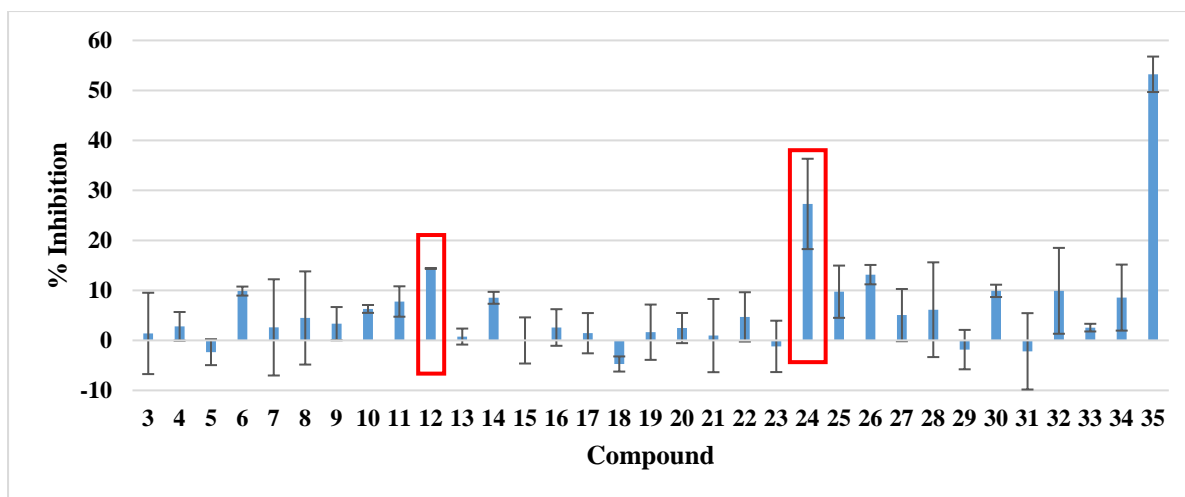
The compounds were tested with the Amplex Red assay that indirectly measures the amount of hydrogen peroxide produced with high sensitivity and specificity using the 10-acetyl-3,7-dihydroxyphenoxazine (Amplex Red reagent) fluorogenic probe (see Scheme 4.1).<sup>128</sup> The results are shown in Fig. 4.5. The assay indirectly measures the activity of PC-PLC<sub>Bc</sub> that determines the amount of hydrogen peroxide produced. Firstly, PC-PLC<sub>Bc</sub> cleaves its natural

substrate, PC into DAG and phosphocholine followed by the hydrolysis of phosphocholine by alkaline phosphatase into choline and inorganic phosphate. In the presence of choline oxidase, choline is oxidised to betaine and hydrogen peroxide. The highly fluorescent product, resorufin is formed when hydrogen peroxide reacts to the Amplex Red reagent in the presence of a peroxidase.



**Scheme 4.1.** Flow chart of the Amplex Red assay for indirect monitoring of PC-PLC<sub>Bc</sub> activity.

The compounds were considered as hits with inhibition >15% of the fluorescence emitted. Initially, two hits were identified: 3-[(4-cyclohexylphenyl)sulfonamido]benzoic acid (**12**) and 1-[4-(trifluoromethyl)phenyl]-2,3,4,9-tetrahydro-1*H*-pyrido[3,4-*b*]indole-3-carboxylic acid (**24**) having ~15% and ~27% inhibition at 10  $\mu$ M. Interestingly, both compounds have a carboxylic acid moiety that can coordinate to the catalytic Zn<sup>2+</sup> ions. This suggests coordination to the Zn<sup>2+</sup> ions to be important to the inhibition mechanism.



**Figure 4.5.** The average PC-PLC<sub>Bc</sub> activity and their standard deviations after screening the 32 selected compounds from the pilot screen at 10  $\mu$ M using the Amplex Red biochemical assay. The positive control used is D609 (35). Active hits were identified as having >15% inhibitory activity resulting in hits 12 and 24 as bracketed in red.

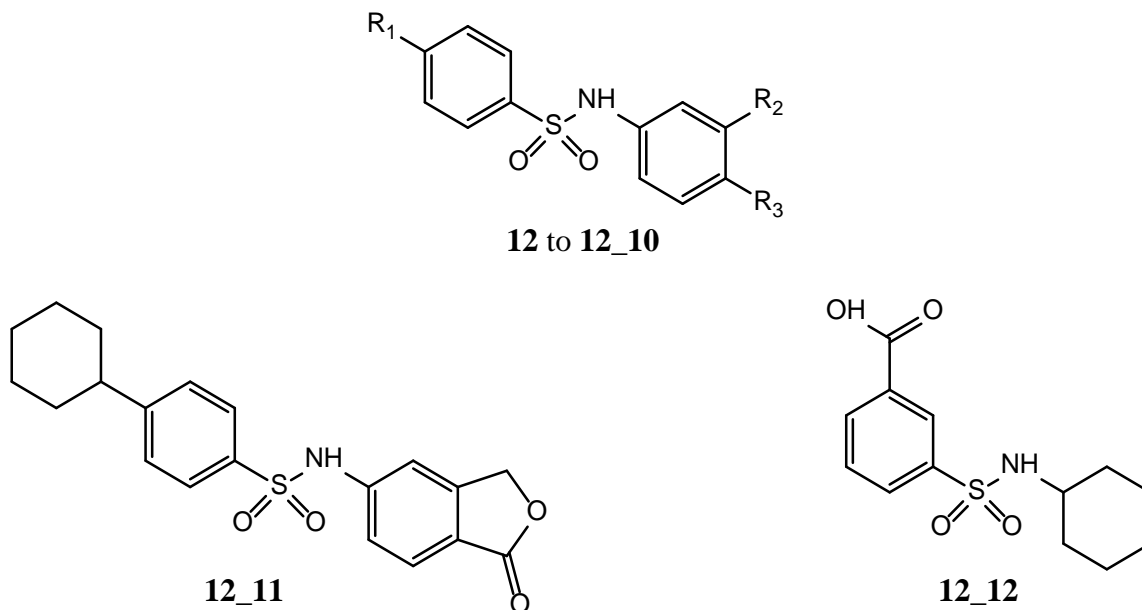
In order to explore the SARs of **12** and **24**, structurally similar analogues were found for experimental testing from eMolecules.com,<sup>68</sup> a website that hosts commercially available compounds, three compound providers were used: ChemBridge Corporation,<sup>127</sup> Chemical Diversity Inc (ChemDiv)<sup>129</sup> and InterBioScreen Ltd (IBS).<sup>130</sup> The chemical series resulted from hits **12** and **24** are classified as the *N*-phenylbenzenesulphonamides and 2,3,4,9-tetrahydro-1*H*-pyrido[3,4-*b*]indoles based on their core scaffolds.

#### 4.2.1.1 *N*-phenylbenzenesulphonamides

Hit **12** exhibited 15% inhibition in the biochemical assay and features *N*-phenylbenzenesulphonamide as the core structure, *m*-substituted carboxylic acid moiety to the sulphonamide bridge in *N*-phenyl and *p*-substituted cyclohexyl to the sulphonamide scaffold linked to benzene. A similarity coefficient of 0.7 was used to search for structurally similar compounds to **12** in eMolecules.com. Twelve *N*-phenylbenzenesulphonamides were purchased and experimentally tested and the results are given in Table 4.1. *N*-phenylbenzenesulphonamides **12\_1** to **12\_10** have mainly large cyclic structures at position R<sub>1</sub> with various substituents at positions R<sub>2</sub> and R<sub>3</sub> that includes carboxylic acids, halogens,

methyls, benzamide, *N,N*-dimethyl and nitro groups. Compound **12\_11** has an isobenzofuran-1(3*H*)-one substituted for a phenyl ring, whereas **12\_12** is the only compound with a non-aromatic cyclohexyl ring linked to the sulphonamide.

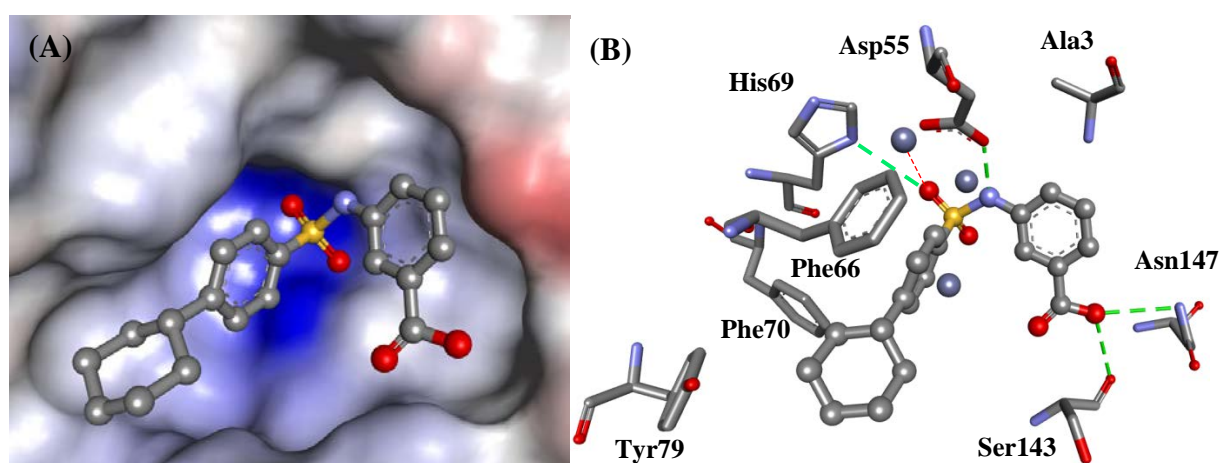
**Table 4.1.** The molecular structures and inhibition activities of the *N*-phenylbenzenesulphonamide derivatives. The inhibition was measured at 10  $\mu$ M using the Amplex Red biochemical assay. Compound **35** (D609) was used as the positive control.



Compound	R <sub>1</sub>	R <sub>2</sub>	R <sub>3</sub>	Inhibition (%)
<b>35</b>	-	-	-	53
<b>12<sup>a</sup></b>	cyclohexyl	CO <sub>2</sub> H	H	14
<b>12_1</b>	cyclohexyl	CO <sub>2</sub> H	N(CH <sub>3</sub> ) <sub>2</sub>	9
<b>12_2</b>	cyclohexyl	benzamide	H	5
<b>12_3</b>	cyclohexyl	H	CO <sub>2</sub> H	5
<b>12_4</b>	cyclohexyl	H	NO <sub>2</sub>	6
<b>12_5</b>	phenyl	CO <sub>2</sub> H	H	3
<b>12_6</b>	cyclohexyl	CH <sub>3</sub>	H	8
<b>12_7</b>	CO <sub>2</sub> H	F	H	5
<b>12_8</b>	cyclohexyl	H	Br	6
<b>12_9</b>	methyl	CO <sub>2</sub> H	H	0
<b>12_10</b>	cyclohexyl	H	CH <sub>3</sub>	3
<b>12_11</b>	-	-	-	11
<b>12_12</b>	-	-	-	6

<sup>a</sup>Hit compound from pilot screen.

As can be seen in Table 4.1, only modest inhibitory activities against PC-PLC<sub>Bc</sub> was achieved by the *N*-phenylbenzenesulphonamides and none of the derivatives demonstrated improved activities over **12**. The second best inhibitor was for compound **12\_11** showed ~11% inhibition. Removal of ring systems (**12\_7** and **12**) and substituting cyclohexyl for aromatic rings at R<sub>1</sub> (**12\_5** and **9**) did not improve potency. Substitutions by *N,N*-dimethyl (**12\_1**), benzamide (**12\_2**), carboxylic acid (**12\_3**), nitro (**12\_4**), halogen (**12\_7** and **8**) and methyl (**12\_6** and **10**) at positions R<sub>3</sub> and R<sub>4</sub> on the phenyl ring showed no improvement in inhibitory activities.



**Figure 4.6.** The docked configuration of *N*-phenylbenzenesulphonamide **12** in the PC-PLC<sub>Bc</sub> binding site. (A) The protein surface is rendered. Red and blue surfaces depicts negative and positive partial charge whereas grey shows neutral lipophilic areas. (B) Zn<sup>2+</sup> ions are depicted as grey unconnected spheres which form coordinated bonds shown as red dotted lines with the sulphonyl functional group, whereas hydrogen bonding are depicted as green dotted lines with amino acid residues Ser13, Asp55, His69 and Ser143. The hydrogen atoms were suppressed for clarity.

Molecular modelling studies for the most active *N*-phenylbenzenesulphonamide **12** predicted hydrogen bonding interactions with amino acid residues Asp55, His69, Ser143 and Asn147 (see Fig. 4.6). The benzoic acid is predicted to be oriented towards a hydrophobic area formed by Ala3 side chains and Ser2 backbone, whereas the cyclohexylbenzene is situated nearby residues Phe66 and Phe70 in which the two moieties are expected to form

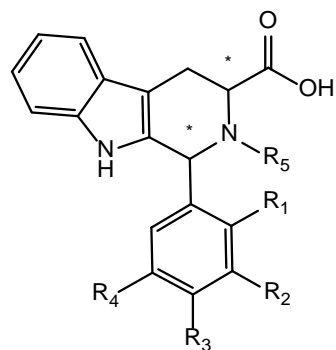


hydrophobic contacts. Replacing cyclohexyl with phenyl (**12\_5**) and small methyl group (**12\_9**) is detrimental to activity, suggesting the importance of conformational flexibility. Visual inspection of derivative **12\_9** binding mode showed lack of spatial occupation by methyl and is assumed to have less favourable hydrophobic contacts. This phenomenon was also seen for derivative **12\_12**. The sulphonamide functional group is facing towards the triple  $\text{Zn}^{2+}$  ions, where they are expected to coordinate with the metals which can explain its inhibitory effect on PC-PLC<sub>Bc</sub> activity. The *N*-phenylbenzenesulphonamides have similar binding poses and interactions with the exception of **12\_2** where *m*-benzamide inserted at R<sub>2</sub> gave a non-plausible binding mode, *e.g.*, phenyl ring positioned outside the hydrophobic site.

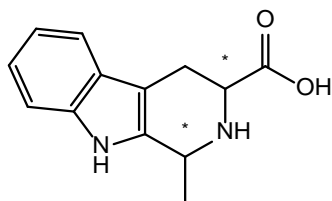
#### **4.2.1.2 2,3,4,9-tetrahydro-1*H*-pyrido[3,4-*b*]indoles**

Hit **24** has a measured inhibition of 27% (refer to Fig. **4.5**). It features a 2,3,4,9-tetrahydro-1*H*-pyrido[3,4-*b*]indole heterocyclic core, the pyridine moiety is linked to a phenyl ring and a carboxylic acid. Again, a similarity coefficient value of 0.7 was used to search for structurally similar compounds on eMolecules.com. Thirty-five compounds were purchased for experimental testing. Specifically, the compounds contain various substitutions around the phenyl ring (**24\_1** to **32**), methyl (**24\_33**), pyridine (**24\_34**) and naphthalene moieties (**24\_35**) replacing the phenyl ring were also obtained.

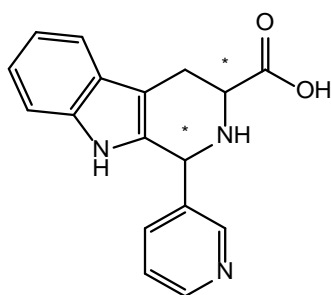
**Table 4.2.** The molecular structures and inhibition activities of the 2,3,4,9-tetrahydro-1*H*-pyrido[3,4-*b*]indole derivatives. The inhibition was measured at 10  $\mu$ M using the Amplex Red biochemical assay. Compound **35** (D609) was used the positive control. Stereocenters are marked with an asterisk (\*).



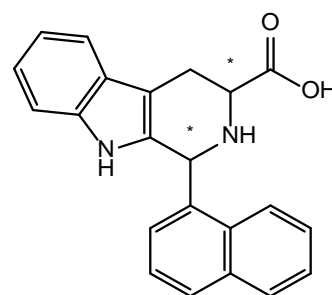
**24 to 24\_32**



**24\_33**



**24\_34**



**24\_35**

Compound	R <sub>1</sub>	R <sub>2</sub>	R <sub>3</sub>	R <sub>4</sub>	R <sub>5</sub>	Inhibition (%)
<b>35</b>	-	-	-	-	-	62
<b>24<sup>a</sup></b>	<b>H</b>	<b>H</b>	<b>CF<sub>3</sub></b>	<b>H</b>	<b>H</b>	<b>27</b>
<b>24_1</b>	H	H	H	H	H	17
<b>24_2</b>	OCH <sub>3</sub>	H	H	H	H	25
<b>24_3</b>	H	H	OCH <sub>3</sub>	H	H	9
<b>24_4</b>	H	H	OCH <sub>2</sub> CH <sub>3</sub>	H	H	34
<b>24_5</b>	OCH <sub>3</sub>	OCH <sub>3</sub>	H	H	H	22
<b>24_6</b>	H	OCH <sub>3</sub>	OCH <sub>3</sub>	H	H	38
<b>24_7</b>	H	OCH <sub>3</sub>	OCH <sub>2</sub> CH <sub>2</sub> CH <sub>3</sub>	H	H	20
<b>24_8</b>	H	OCH <sub>3</sub>	OCH <sub>3</sub>	OCH <sub>3</sub>	H	15
<b>24_9</b>	H	OH	H	H	H	18
<b>24_10</b>	<b>H</b>	<b>OH</b>	<b>OH</b>	<b>H</b>	<b>H</b>	<b>75</b>
<b>24_11</b>	H	OH	OCH <sub>3</sub>	H	H	47
<b>24_12</b>	H	OH	OCH <sub>2</sub> CH <sub>3</sub>	H	H	53

24_13	H	OCH <sub>3</sub>	OH	H	H	43
24_14	H	OCH <sub>3</sub>	OH	OCH <sub>3</sub>	H	60
24_15	H	H	N(CH <sub>3</sub> ) <sub>2</sub>	H	H	49
24_16	H	F	H	H	H	34
24_17	H	H	F	H	H	24
24_18	Cl	H	H	H	H	14
24_19	H	H	Cl	H	H	17
24_20	Cl	H	Cl	H	H	44
24_21	Br	H	H	H	H	27
24_22	H	Br	H	H	H	21
24_23	H	CF <sub>3</sub>	H	H	H	19
24_24	H	H	OCHF <sub>2</sub>	H	H	16
24_25	NO <sub>2</sub>	H	H	H	H	26
24_26	H	NO <sub>2</sub>	H	H	H	23
24_27	CH <sub>3</sub>	H	H	H	H	29
24_28	H	H	CH <sub>3</sub>	H	H	19
24_29	H	H	isopropyl	H	H	-10
24_30	CH <sub>3</sub>	H	H	CH <sub>3</sub>	H	21
24_31	H	H	H	H	COCH <sub>3</sub>	5
24_32	H	H	(CH <sub>2</sub> ) <sub>7</sub> CH <sub>3</sub>	OCH <sub>3</sub>	H	-28
24_33	-	-	-	-	-	4
24_34	-	-	-	-	-	24
24_35	-	-	-	-	-	37

<sup>a</sup>Hit compound from pilot screen.

As can be seen from Table 4.2, there is a clear indication of PC-PLC<sub>Bc</sub> inhibition for the 2,3,4,9-tetrahydro-1*H*-pyrido[3,4-*b*]indoles, several that showed improved inhibition (>27%). Generally, the best compounds are *m*-alkoxy and *p*-hydroxyl electron-donating groups on the phenyl ring. Dual *m*- and *p*-OH substitutions (24\_10) has the strongest inhibition effect (~75%), followed by dual *m*-methoxy substitutions on both sides of the ring together with *p*-hydroxyl substitution (24\_14) at ~60% and following closely by 24\_12, 15 (*p*-*N,N*-dimethyl),

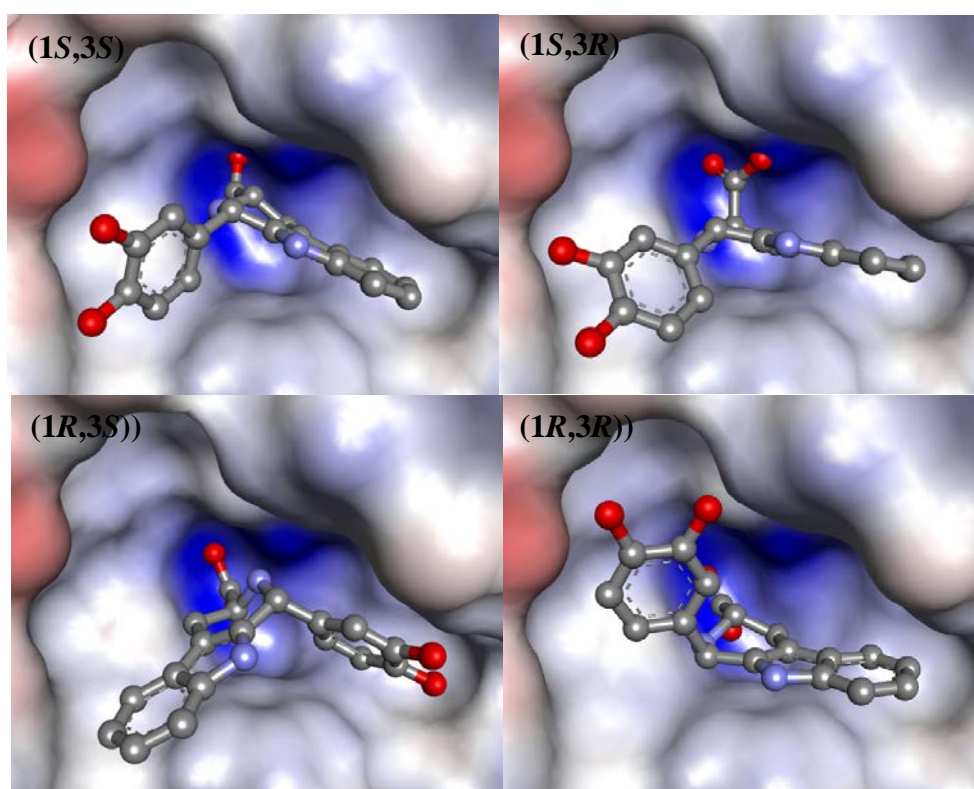
**11** and **13** ranging between 43 – 53% inhibition. However, alkoxy substitutions alone showed modest inhibitory effects (**24\_3**, **5**, **6**, and **8**). Single halogen substitutions did not improve efficacy with the exception of *m*-fluoro (**24\_16**), whereas dual *o*- and *p*-chloro (**24\_20**) demonstrated improved inhibition (~44%). The remaining substituents on the phenyl ring include nitro (**24\_26** and **25**) and *m*-CF<sub>3</sub> (**24\_23**), hydrogens (**24\_1**) and alkyls (**24\_28**, **27**, **29**, **30** and **32**) did not improve inhibition. Replacing phenyl with pyridine (**24\_34**) did not improve activity, whereas improved activity is seen with naphthalene (**24\_35**). Methyl (**24\_33**) substantially reduced inhibition suggesting the importance of bulky aromatic systems. Introducing acetyl on the piperidine ring at R<sub>5</sub> (**24\_31**) demonstrated detrimental effect for the inhibition.

**Table 4.3.** Docking scores of the four possible diastereoisomers of compound **24\_10** against the PC-PLC<sub>Bc</sub> crystal structure.

Compound	GS	CS	ChemPLP	ASP
<b>24_10 (1S,3S)</b>	75.6	45.0	84.3	43.8
<b>24_10 (1S,3R)</b>	69.5	44.1	83.5	41.5
<b>24_10 (1R,3R)</b>	64.1	43.0	77.6	42.4
<b>24_10 (1R,3S)</b>	73.8	46.4	95.3	45.5

Molecular modelling was conducted to understand the inhibition mechanisms behind the 2,3,4,9-tetrahydro-1*H*-pyrido[3,4-*b*]indoles. Since the 2,3,4,9-tetrahydro-1*H*-pyrido[3,4-*b*]indoles were tested as racemic mixtures, the diastereoisomer corresponding to the best activity was unclear. The problem was investigated by docking all four diastereoisomers (Table 4.3) using the most active compound, 1-(3,4-dihydroxyphenyl)-2,3,4,9-tetrahydro-1*H*-pyrido[3,4-*b*]indole-3-carboxylic acid (**24\_10**). The results revealed that the (**1R,3S**)-diastereoisomer scored best for the CS, ChemPLP and ASP functions. It also scored comparably well to the highest GS scoring (**1S,3S**)-diastereoisomer. The consensus concept

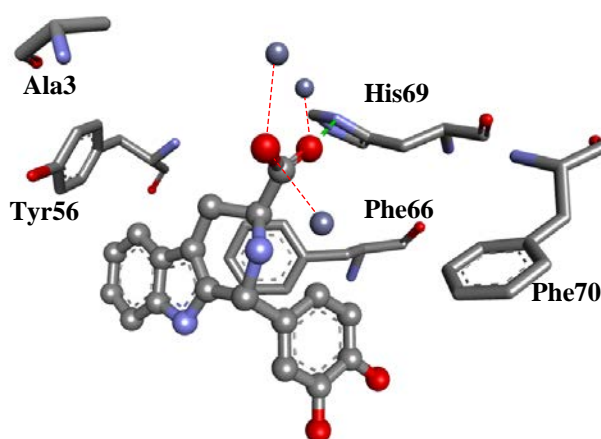
was used to show that the (**1R,3S**)-diastereoisomer is predicted to be most active, *i.e.*, three scoring functions (CS, ChemPLP and ASP) out of four favoured the (**1R,3S**)-diastereoisomer. Thus, it is reasonable to use the (**1R,3S**)-diastereoisomer to model for the 2,3,4,9-tetrahydro-1*H*-pyrido[3,4-*b*]indole series. Visual inspection showed plausible binding modes (see. Fig. 4.7) for all four diastereoisomers, *e.g.*, phenyl and indole moieties occupy lipophilic areas and the carboxylic acid moiety is facing the Zn<sup>2+</sup> ions.



**Figure 4.7.** The docked configurations of the four 1-(3,4-dihydroxyphenyl)-2,3,4,9-tetrahydro-1*H*-pyrido[3,4-*b*]indole-3-carboxylic acid (**24\_10**) diastereomers in the PC-PLC<sub>Bc</sub> binding site. The protein surface is rendered and hydrogen atoms are suppressed. Red and blue surfaces depicts negative and positive partial charge respectively whereas grey shows neutral lipophilic areas.

(**1R,3S**)-2,3,4,9-tetrahydro-1*H*-pyrido[3,4-*b*]indole **24\_1** to **32** are analogues with different substituents around the phenyl ring showed high similarity in binding modes and scores. The binding mode and interactions are represented by **24\_10** in Fig. 4.8. The phenyl group is situated close to lipophilic areas partly consisting of Phe66 and Phe70 residues where they are expected to form  $\pi$ - $\pi$  stacking, whereas the tricyclic 2,3,4,9-tetrahydro-1*H*-pyrido[3,4-

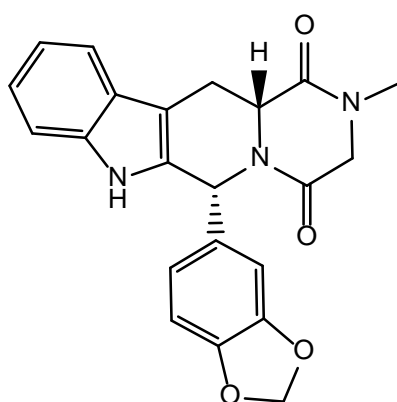
b]indole heterocycle occupies a lipophilic region accommodated by residues Ala3, Tyr56 and Phe66. The hydroxyl substituents at R<sub>2</sub> and R<sub>3</sub> faces the water environment, which explains the preference for polar groups at these positions over large alkyls, *e.g.*, **24\_29** and **32**. Hydrogen bonding interaction was predicted with His69 *via* the carboxylic acid functionality. In some cases, Glu146 formed hydrogen bond with pyridine, *e.g.*, modelling compound **24\_15**. The carboxylic acid functional group facing the deeply buried positively charged regions are expected to form coordination bonds with the Zn<sup>2+</sup> ions.



**Figure 4.8.** Binding interactions of (1*R*,3*S*)-**24\_10** in the PC-PLC<sub>Bc</sub> binding site has hydrogen bonding with His69 residue depicted as a green dotted line. Zn<sup>2+</sup> ions are depicted as unconnected grey spheres and are coordinated by the carboxylic acid functionality shown as red dotted lines. Hydrophobic amino acid residues Ala3, Tyr56, Phe66 and Phe70 are also shown. The hydrogen atoms are suppressed.

Evidence from the scores (see appendix) have limited weight in explaining the SAR of different substitution patterns on the phenyl ring, *i.e.*, they lacked correlations between activities and substitution patterns. For example, the more active alkoxy substituted compounds **24\_2** to **14** have similar scores to halogenated compounds **24\_16** to **24**. Replacing methyl for the phenyl ring (**24\_33**) resulted in score reduction but showed comparable activity to non-substituted phenyl (**24\_1**).

Interestingly, tadalafil (see Fig. 4.9), a marketed drug branded as Cialis, consist of the 2,3,4,9-tetrahydro-1*H*-pyrido[3,4-*b*]indole core as its molecular scaffold. Tadalafil is a phosphodiesterase type 5 (PDE5) inhibitor and is orally available as a treatment for erectile dysfunction.<sup>131</sup> This suggests the 2,3,4,9-tetrahydro-1*H*-pyrido[3,4-*b*]indole core is likely to be compatible for use in the clinic with acceptable pharmacokinetic profile. There is no record of the drug with anticancer properties which suggests selectivity is independent of the tricyclic 2,3,4,9-tetrahydro-1*H*-pyrido[3,4-*b*]indole core.



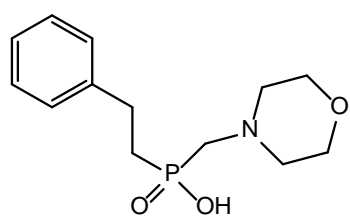
**Figure 4.9.** The molecular structure of tadalafil, a marketed phosphodiesterase type 5 inhibitor.

## 4.2.2 Main Screen

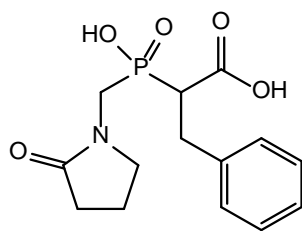
The success of the pilot screen in finding active hits encouraged further search of more inhibitors on a larger scale. An additional  $4 \times 10^4$  compounds were obtained from ChemBridge which was filtered using Lipinski's Rule of Five leaving  $\sim 3.3 \times 10^4$  drug-like molecules. All four scoring functions in the GOLD software suite were used: GS, CS, ChemPLP and ASP. The same consensus concept used as in the pilot screen was applied, *i.e.*, only ligands that scored well with all scoring functions and had hydrogen bonding activities were taken forward.<sup>132-133</sup> The first filter criteria was as follow: GS > 55, CS > 25, ChemPLP > 65, ASP > 25 and compounds with acceptable hydrogen bonding activity, that is HB > 0 (GS) and HB  $\geq$  1 for CS and ChemPLP. Imposing this filter criteria resulted in 5249

candidates which were re-docked at 100% search efficiency. The compound candidates were filtered through a second score criteria: GS > 55, CS > 25, ChemPLP > 70 and ASP > 25. Furthermore, compounds with good hydrogen bonding activities were selected: HB > 10 (GS) and HB  $\geq$  1 (CS and ChemPLP) resulting in 871 candidates. Compounds with carboxylic acid functionality and moieties that can ionise to form negative charges were taken forward. This approach assumed the strongest inhibitors have the ability to coordinate to the Zn<sup>2+</sup> ions. One hundred and fifty-five candidates were selected that had carboxylic and sulphonic acid moieties. Candidates with undesirable reactive, cytotoxic and chemically unstable moieties suggested to produce false positives were removed.<sup>17</sup> This was followed by visual inspections for plausible binding modes (see pilot screen section **5.2.1**) and they had to have an agreeable binding pose between the four scoring functions, *i.e.*, good overlays between the binding poses given by low RMSDs. In total, fifty four compounds were selected (see molecular structures in Fig. **4.10**) and were purchased for experimental testing.

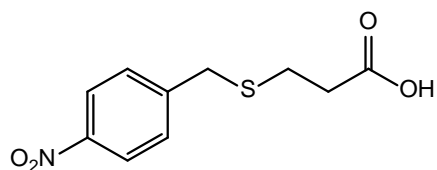




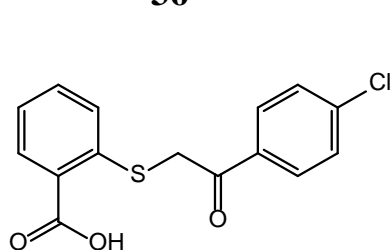
**36**



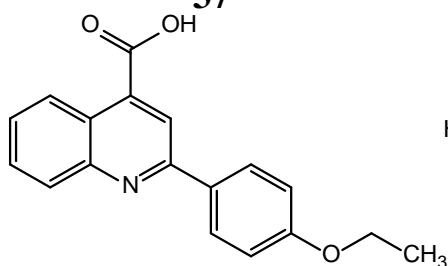
**37**



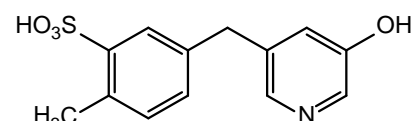
**38**



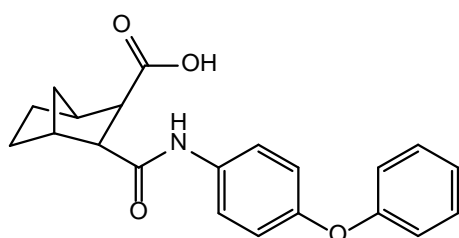
**39**



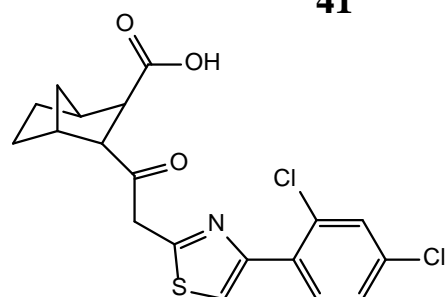
**40**



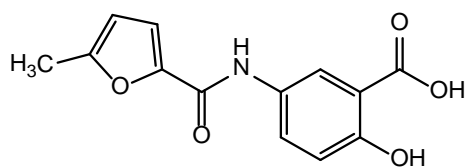
**41**



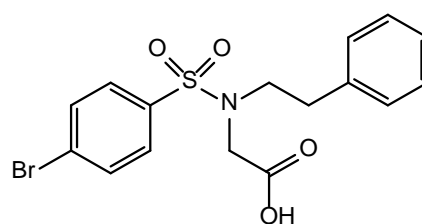
**42**



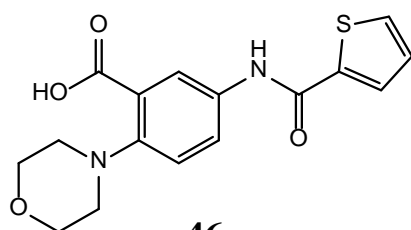
**43**



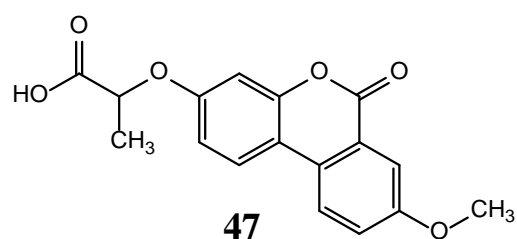
**44**



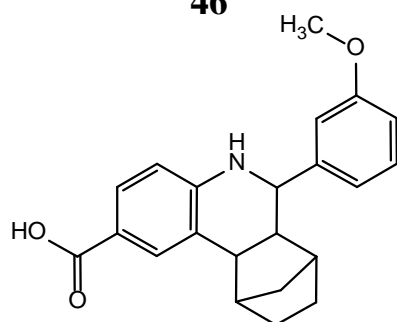
**45**



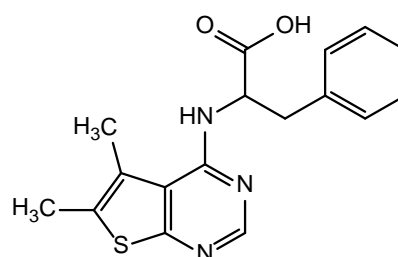
**46**



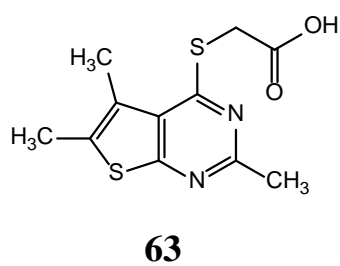
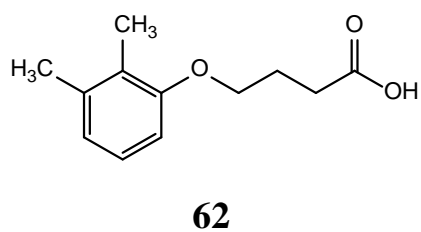
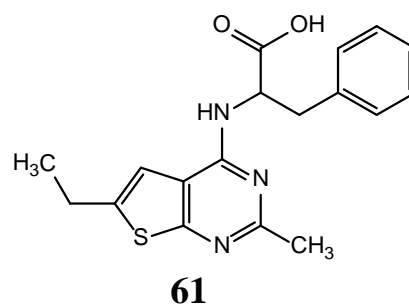
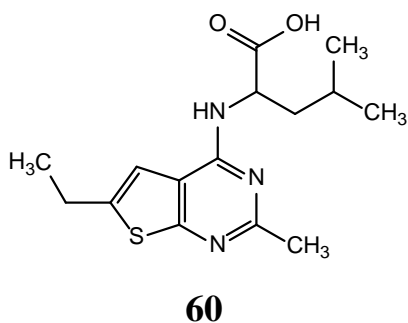
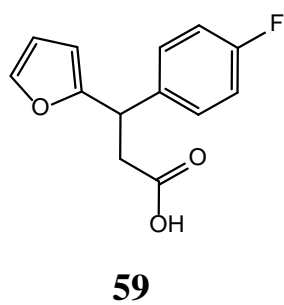
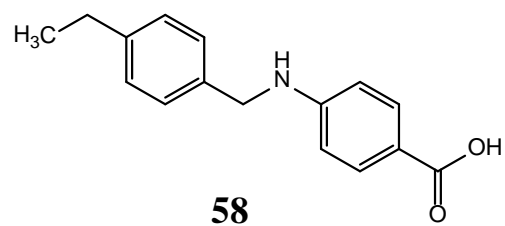
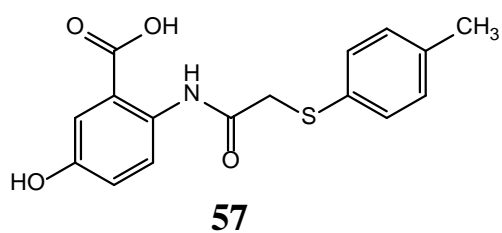
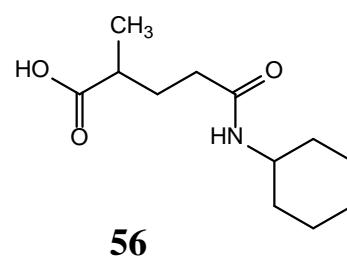
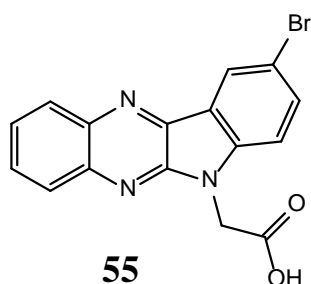
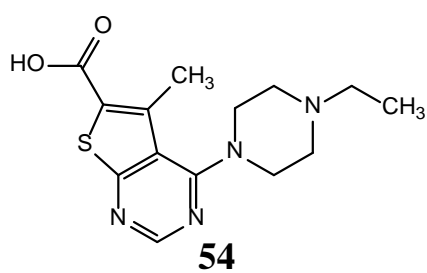
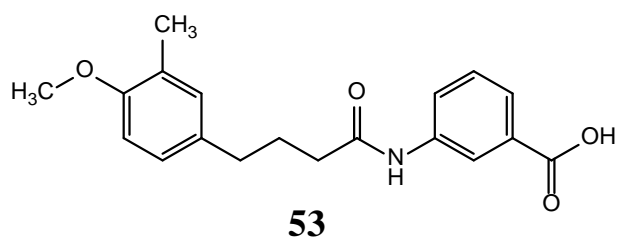
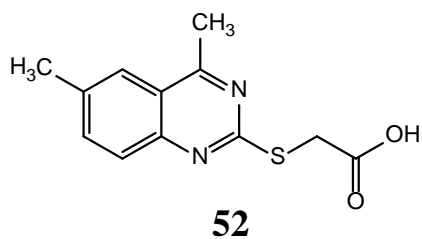
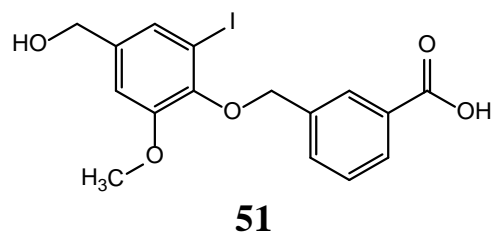
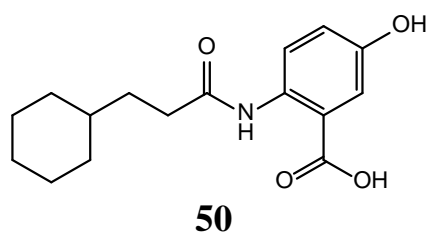
**47**

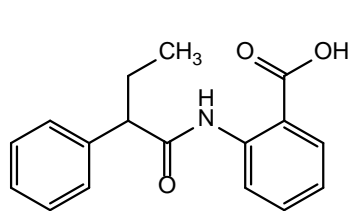


**48**

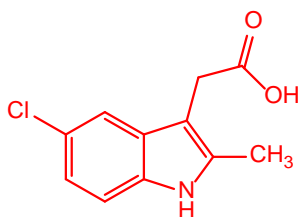


**49**

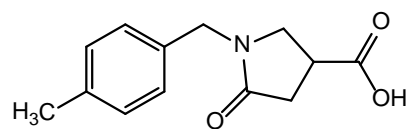




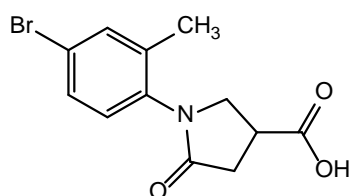
**64**



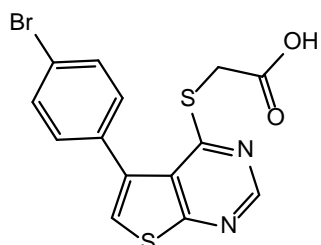
**65**



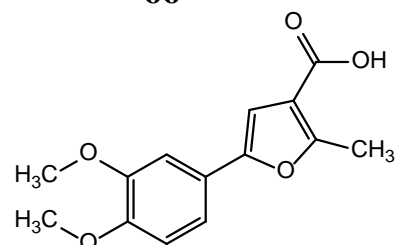
**66**



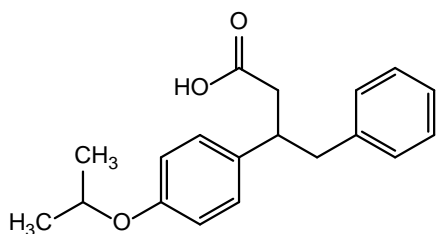
**67**



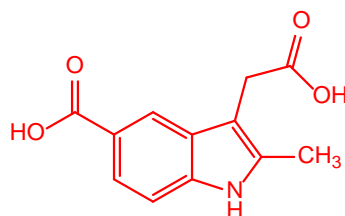
**68**



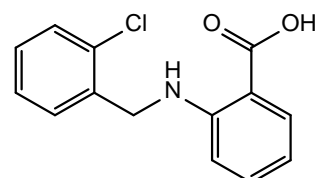
**69**



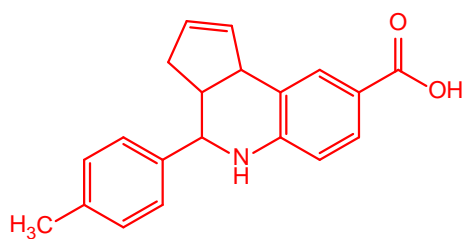
**70**



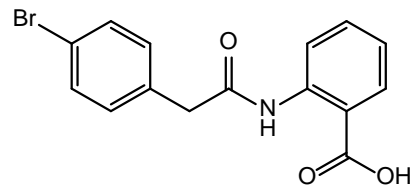
**71**



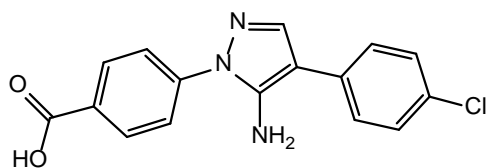
**72**



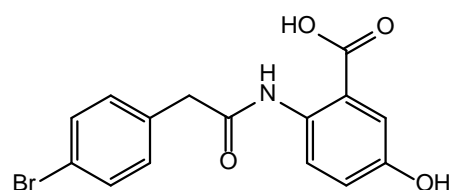
**73**



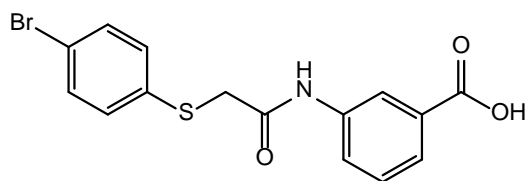
**74**



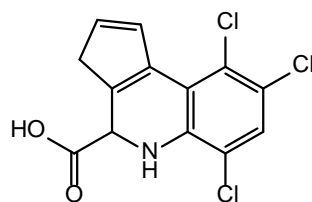
**75**



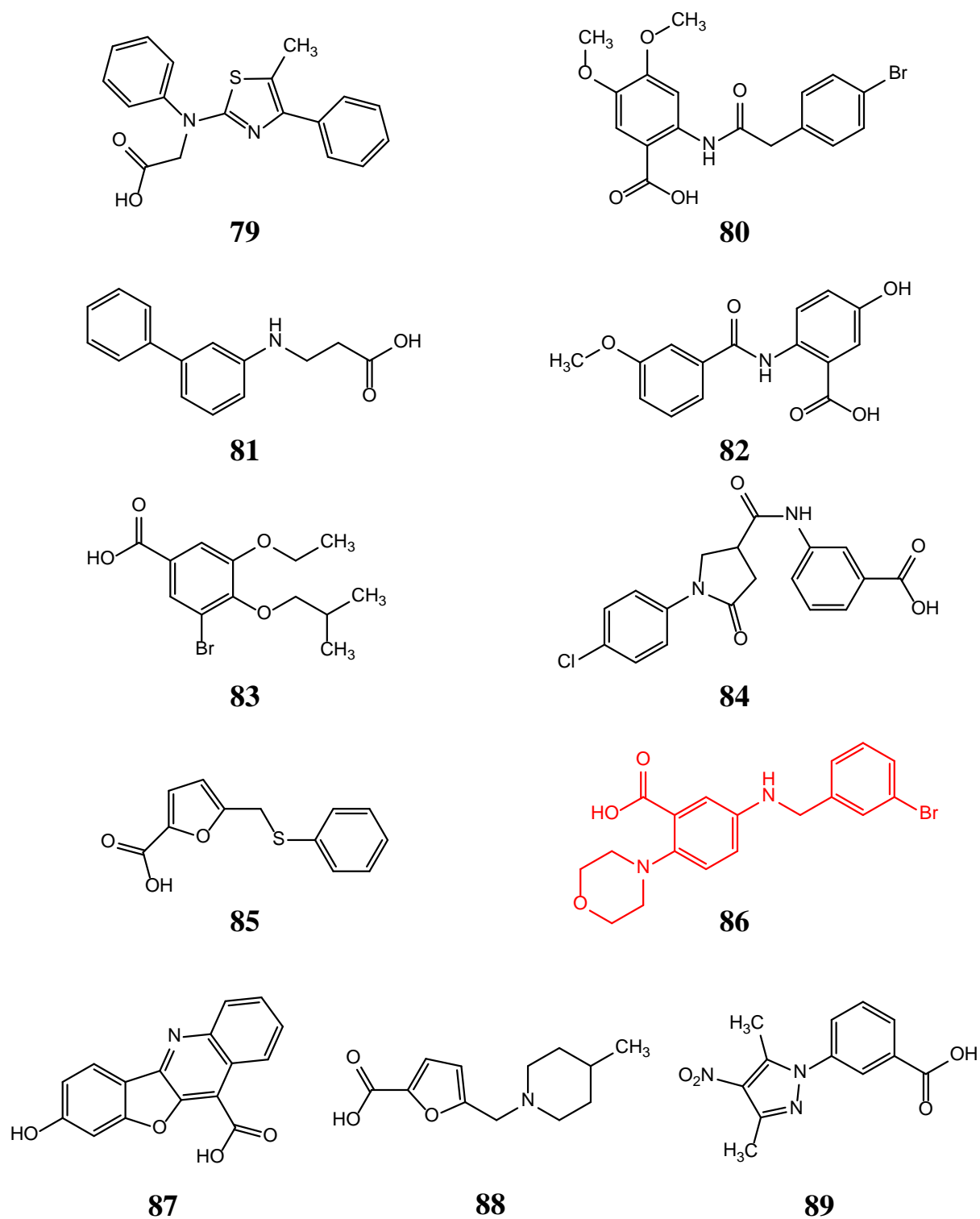
**76**



**77**



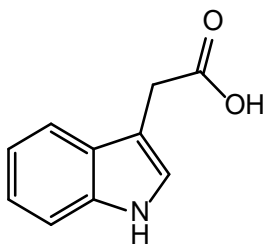
**78**



**Figure 4.10.** Molecular structures of fifty four compounds identified and tested with the Amplex Red biochemical assay for PC-PLC inhibition. The top four active hits **86**, **65**, **71** and **73** are coloured red.

Four compounds emerged as strong hits with inhibition >50%. Hit **86** showed the best inhibition effect (80%) followed by **65** (65%), **71** (60%) and **73** (54%). Hits emerged with medium inhibitory activities (25 – 50%) were **39** (25%), **50** (43%), **51** (28%), **75** (40%), **76** (29%) and **82** (47%). Interestingly, some compounds of the same family appeared as actives from the main screen: indole-3-acetic acids (**65** and **71**), and benzamidobenzoic acids (**50**, **76** and **82**). Four more benzamidobenzoic acid derivatives were also recognised with weak activities (<25%): **57**, **64**, **74** and **80**.

Similarity searches were conducted on the ChemBridge Online Chemical store based on hits **73** and **86**, and the more modest benzamidobenzoic acid **82**. Search results revealed no similar compounds to **73** at 0.7 threshold. Lowering the threshold to 0.5 resulted in appearance of several 2,3,4,9-tetrahydro-1*H*-pyrido[3,4-*b*]indoles. Forty-three derivatives of **86** and twenty derivatives of **82** were obtained using the search coefficient value of 0.7. Although the indole-3-acetic acids **65** and **71** displayed strong inhibition effects, they were opted from the search as they were speculated to be false positives. The compounds resemble the plant growth hormone, auxin (see molecular structure in Fig. **4.11**). Naturally, auxin is biosynthesised from its amino acid precursor, tryptophan,<sup>134</sup> and was reported to possess anticancer activities in the presence of horseradish peroxidase.<sup>135-137</sup> The mechanism suggested was through the generation of free radicals that resulted in apoptotic cell death.<sup>135-137</sup> Auxin alone did not induce cytotoxicity in human melanoma cells,<sup>138</sup> which confirmed their inactivity *in vitro*.

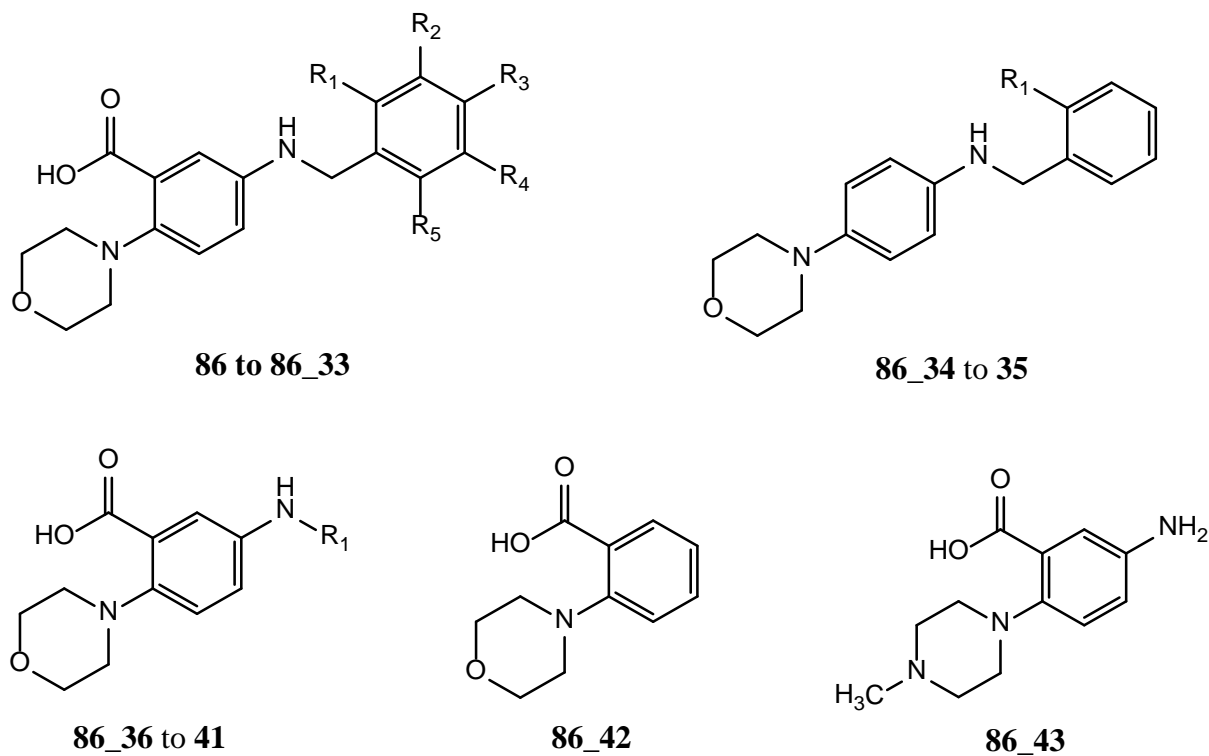


**Figure 4.11.** The molecular structure of the plant growth hormone, auxin.

#### 4.2.2.1 2-morpholinobenzoic acids

Hit **86** emerged as the strongest inhibitor from the main screen (80% inhibition). It features a 2-morpholinobenzoic acid scaffold linked to a benzyl group by an amine. The amine is positioned *para* to the morpholine ring. A total of forty-three analogues of **86**, classified as the 2-morpholinobenzoic acids, were purchased for experimental testing to expand the SAR. Compounds **86\_1** to **33** have various substitution groups on the phenyl ring. Two analogues, **86\_34** and **35**, replaced the natural position of the carboxylic acid group with a fluorine and methyl. Compounds **86\_36** to **41** have various heterocyclic and aromatic groups linked to the 5-amino functionality. The 5-amino group is replaced with a hydrogen for **86\_42**. Singleton **86\_43** contain a piperazine ring instead of a morpholine. The experimental results are presented in Table **4.4**, which clearly suggest that the 2-morpholinobenzoic acids display inhibition activities against PC-PLC<sub>Bc</sub>.

**Table 4.4.** The molecular structures and inhibition activities of the 2-morpholinobenzoic acid derivatives. The inhibition was measured at 10  $\mu$ M using the Amplex Red biochemical assay. Compound **35** (D609) was used the positive control.



Compound	R <sub>1</sub>	R <sub>2</sub>	R <sub>3</sub>	R <sub>4</sub>	R <sub>5</sub>	Inhibition (%)
<b>35</b>	-	-	-	-	-	35
<b>86<sup>a</sup></b>	<b>H</b>	<b>Br</b>	<b>H</b>	<b>H</b>	<b>H</b>	<b>80</b>
<b>86_1</b>	H	H	H	H	H	67
<b>86_2</b>	F	H	H	H	H	70
<b>86_3</b>	H	H	F	H	H	70
<b>86_4</b>	Cl	H	H	H	H	89
<b>86_5</b>	H	Cl	H	H	H	89
<b>86_6</b>	Cl	H	H	H	F	72
<b>86_7</b>	H	H	Br	H	H	71
<b>86_8</b>	Cl	H	Cl	H	H	65
<b>86_9</b>	CF <sub>3</sub>	H	H	H	H	66
<b>86_10</b>	H	H	SCH <sub>3</sub>	H	H	68
<b>86_11</b>	OCH <sub>3</sub>	H	H	H	H	67
<b>86_12</b>	H	OCH <sub>3</sub>	H	H	H	72
<b>86_13</b>	H	H	OCH <sub>3</sub>	H	H	68

86_14	OCH <sub>2</sub> CH <sub>3</sub>	H	H	H	H	71
86_15	H	OCH <sub>2</sub> CH <sub>3</sub>	H	H	H	73
86_16	H	H	OCH <sub>2</sub> CH <sub>3</sub>	H	H	64
86_17	OCH <sub>3</sub>	OCH <sub>3</sub>	H	H	H	68
86_18	OCH <sub>3</sub>	H	OCH <sub>3</sub>	H	H	70
86_19	H	OCH <sub>3</sub>	OCH <sub>3</sub>	H	H	70
86_20	OCH <sub>2</sub> CH <sub>3</sub>	OCH <sub>3</sub>	H	H	H	69
86_21	H	OCH <sub>3</sub>	OCH <sub>2</sub> CH <sub>3</sub>	H	H	69
86_22	H	OCH <sub>2</sub> CH <sub>3</sub>	OCH <sub>3</sub>	H	H	74
86_23	benzyloxy	OCH <sub>3</sub>	H	H	H	74
86_24	H	OCH <sub>3</sub>	benzyloxy	H	H	70
86_25	OCH <sub>3</sub>	H	H	Cl	H	67
86_26	OCH <sub>2</sub> CH <sub>3</sub>	H	H	Cl	H	69
86_27	OCH <sub>3</sub>	OCH <sub>3</sub>	Br	H	H	69
86_28	H	OCH <sub>3</sub>	OCH <sub>3</sub>	Cl	H	72
86_29	OCH <sub>3</sub>	Cl	H	Cl	H	67
86_30	CH <sub>3</sub>	H	H	H	H	65
86_31	H	H	CH <sub>3</sub>	H	H	67
86_32	H	H	CH <sub>2</sub> CH <sub>3</sub>	H	H	70
86_33	H	H	1-pyrrolidine	H	H	72
86_34	F	-	-	-	-	101
86_35	CH <sub>3</sub>	-	-	-	-	98
86_36	H	-	-	-	-	68
86_37	2-phenylacetamide	-	-	-	-	4
86_38	thiophen-2-ylmethyl	-	-	-	-	69
86_39	(2-methoxynaphthalen-1-yl)methyl	-	-	-	-	73
86_40	naphthalene-1-ylmethyl	-	-	-	-	70
86_41	cinnamyl	-	-	-	-	70
86_42	-	-	-	-	-	-25
86_43	-	-	-	-	-	32

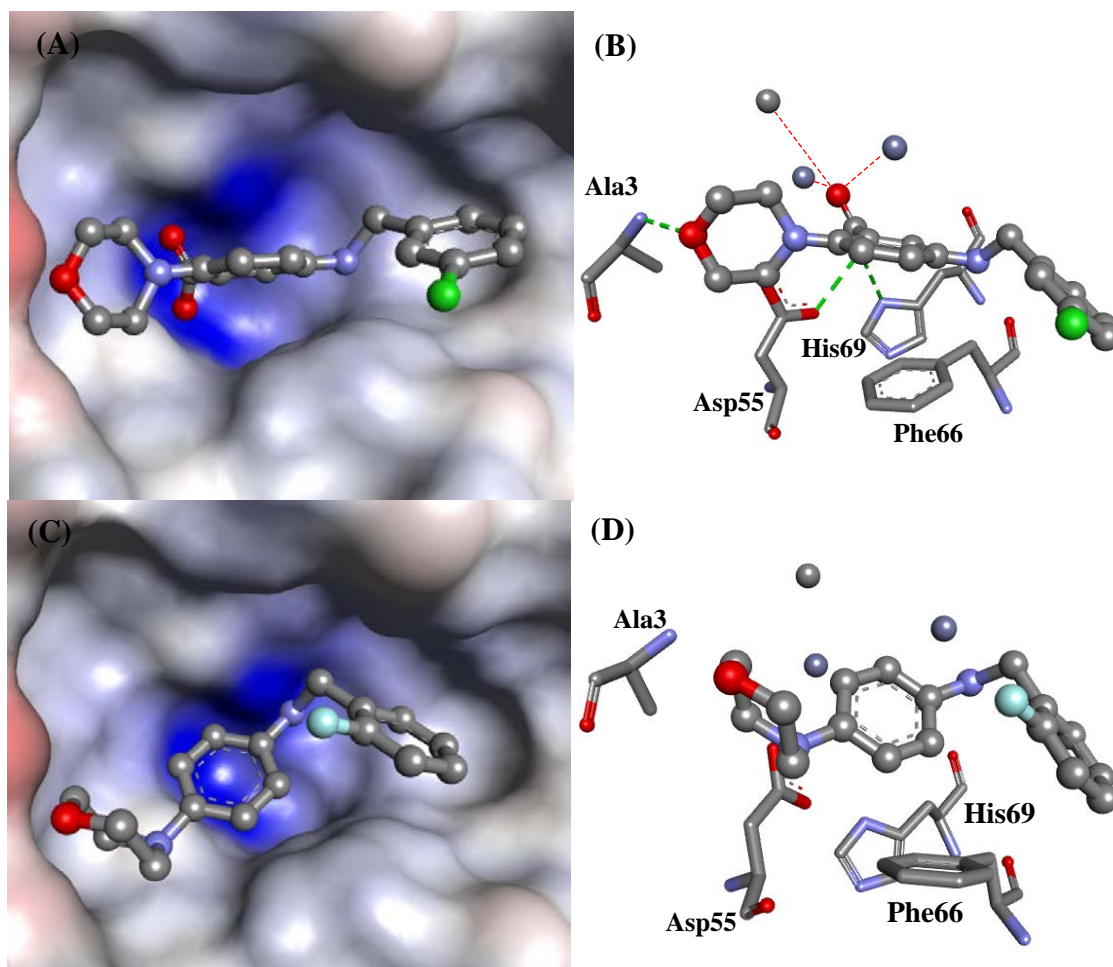
<sup>a</sup>Hit compound from main screen. The 2D molecular structures of derivatives **86\_37** to **41** are provided in Figure A1 in the appendix.



Generally, inhibition effects are tolerated with various substitution patterns around the phenyl ring, mainly ranging between 65 – 75%. Unsubstituted phenyl (**86\_1**) displayed 67% inhibition effect. Most single halogen substitutions (**86** and **86\_2** to **7**) have inhibition >70% with *o*- (**86\_4**) and *m*-chloro (**86\_5**) showed pronounced inhibitory effect at 89%, whereas dual *o*- and *p*-chloro substitution (**86\_8**) have reduced activity at 65%. Introducing *o*-CF<sub>3</sub> (**86\_9**), *p*-SCH<sub>3</sub> (**86\_10**), single alkoxys (**86\_11** to **16**), double alkoxys (**86\_17** to **24**), halogen and alkoxy combinations (**86\_25** to **29**), single alkyls (**86\_30** to **32**) and 1-pyrrolidine (**86\_33**) did not exhibit pronounced inhibition effects.

Unexpectedly, compounds without carboxylic acid functional groups (**86\_34** and **35**) displayed the strongest inhibition effect (~100%). This finding contradicted with the proposed mechanism of inhibition as it had been suggested that a negatively charged functional group, in this case a carboxylate, can inhibit the catalytic action of Zn<sup>2+</sup> ions.<sup>116-117, 139</sup> On the other hand, it was reported that zinc coordinating functionalities are not required for inhibition, and potency is more dependent on selective recognition of surrounding amino acids.<sup>115</sup>

Compounds **86\_36** to **41** showed tolerance in the absence of benzyl groups at R<sub>1</sub> and exhibited inhibition with more than 68% with the exception of 2-phenylacetamide, which displayed significant loss in activity suggesting detrimental effects of amides to the SAR. The absence of 5-amino group (**86\_42**) and replacement of morpholine for 4-methylpiperazine-1-yl (**86\_43**) also displayed significant losses to inhibitory effects.



**Figure 4.12.** The docked configuration of 2-morpholinobenzoic acid **86\_5** (A) and **34** (C) in the PC-PLC<sub>Bc</sub> binding site. The protein surface is rendered. Red and blue surfaces depicts negative and positive partial charge whereas grey shows neutral lipophilic areas. (B) Binding interactions with **86\_5** show Zn<sup>2+</sup> ions as grey unconnected spheres form coordinate bonds depicted as red dotted lines, whereas hydrogen bonding are depicted as green dotted lines with residues Ala3 and Asp55. (D) Binding mode of **86\_34** showed lack of hydrogen bonding neither coordination. The hydrogen atoms were suppressed.

Modelling of the 2-morpholinobenzoic acids is **86\_5** and **34** in Fig. 4.12. It shows the phenyl group occupying a lipophilic space surrounded side chains of Phe70, Tyr79 and Thr133. The benzoic acid core is seen directly above a hydrophobic basin prominently formed by Phe66 that could exhibit  $\pi$ - $\pi$  stacking stabilising the enzyme-inhibitor complex. The morpholine to Ala3 hydrogen bond interaction can be rationalised as a selective interaction. The reason is that substituting a piperazine (**86\_43**) demonstrated significant reduction in activity and showed no hydrogen bond with Ala3. Another hydrogen bond formed is between the

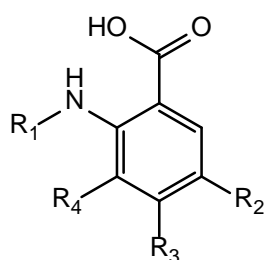
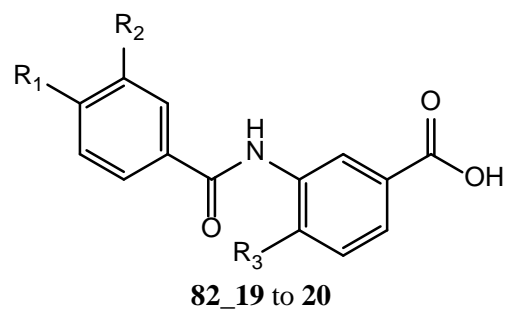
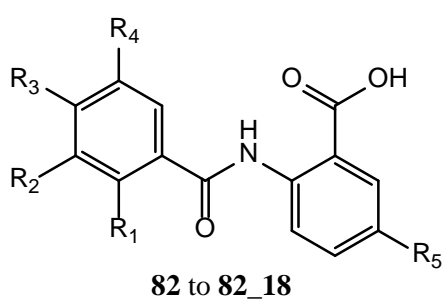
carboxylic acid and Asp55, which was seen for the much highly active compounds **86\_34** and **35** suggesting the importance of lipophilicity for improved inhibition.

#### 4.2.2.2 Benamidobenzoic acids

Hits **50**, **76** and **82** emerged from the main screen with medium inhibitory activities (25 – 50%) whilst weak hits of the same family (**57**, **64**, **74**, and **80**) were recognised. The strongest inhibitor is hit **82** (47% inhibition) has a 3-methoxyphenyl linked to a benzoic acid by an amide as the molecular scaffold. Hits **50** and **76** have non-aromatic ethylcyclohexane and benzyl substituted moieties which replaced the phenyl group. Weak inhibitors **57** has *p*-tolylsulphane replacing the phenyl and **64**, **74** and **80** have benzyl replacing phenyl. The compound family is classified as the benamidobenzoic acids.

A total of twenty structurally similar derivatives resulted from similarity search based on hit **82** and were purchased from the ChemBridge Online Chemical store for experimental testing. Three derivatives (**82\_1** to **3**) are halogen substituted compounds and seven (**82\_4** to **10**) have dual halo-alkoxy substitutions. Compounds **82\_11** to **18** have alkoxy, methyl and hydroxyl – based substitutions. Two compounds (**82\_19** to **20**) have *m*-substituted carboxylic acids, a feature distinctly different to the normal *o*-carboxylic acids.

**Table 4.5.** The molecular structures and inhibition activities of the benzamidobenzoic acid derivatives. The inhibition was measured at 10  $\mu$ M using the Amplex Red biochemical assay. Compound **35** (D609) was used as the positive control.

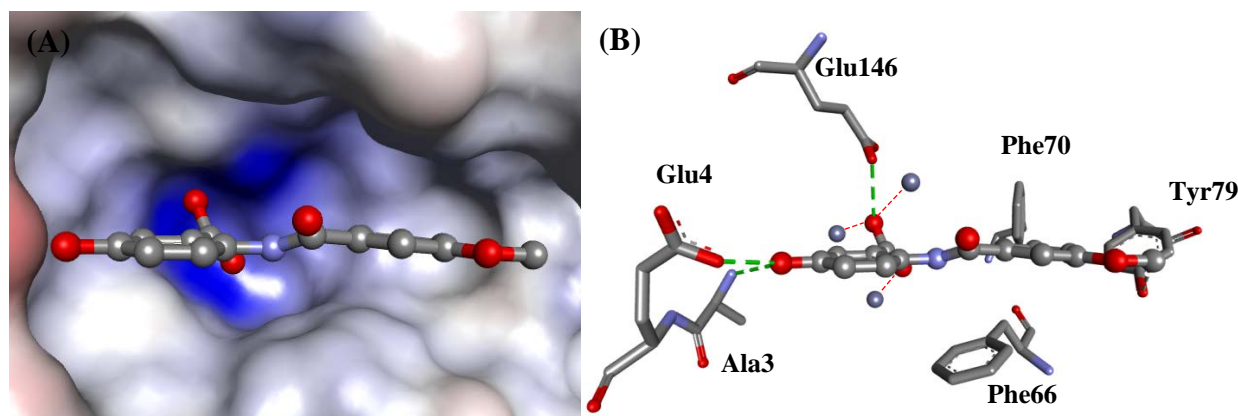


Compound	R <sub>1</sub>	R <sub>2</sub>	R <sub>3</sub>	R <sub>4</sub>	R <sub>5</sub>	Inhibition (%)
<b>35</b>	-	-	-	-	-	67
<b>82<sup>a</sup></b>	H	OCH <sub>3</sub>	H	H	OH	47
<b>82_1</b>	H	Br	H	H	H	27
<b>82_2</b>	H	Br	H	H	Cl	13
<b>82_3</b>	H	H	Cl	H	Br	11
<b>82_4</b>	H	Br	H	H	OH	23
<b>82_5</b>	H	I	H	H	OH	36
<b>82_6</b>	OCH <sub>3</sub>	H	H	Cl	Cl	7
<b>82_7</b>	OCH <sub>3</sub>	H	H	Br	Cl	9
<b>82_8</b>	H	OCH <sub>3</sub>	H	H	Cl	5
<b>82_9</b>	OCH <sub>2</sub> CH <sub>3</sub>	H	H	Cl	OH	18
<b>82_10</b>	OCH <sub>3</sub>	H	H	Br	OH	19
<b>82_11</b>	CH <sub>3</sub>	H	H	H	OH	16
<b>82_12</b>	H	CH <sub>3</sub>	H	H	OH	19
<b>82_13</b>	H	CH <sub>3</sub>	CH <sub>3</sub>	H	OH	31

<b>82_14</b>	H	H	CH <sub>3</sub>	H	OH	26
<b>82_15</b>	H	H	OCH <sub>3</sub>	H	OH	35
<b>82_16</b>	H	OCH <sub>3</sub>	H	OCH <sub>3</sub>	H	5
<b>82_17</b>	OCH <sub>3</sub>	H	H	H	OH	30
<b>82_18</b>	OCH <sub>3</sub>	H	OCH <sub>3</sub>	H	OH	28
<b>82_19</b>	I	H	CH <sub>3</sub>	-	-	0
<b>82_20</b>	H	OCH <sub>3</sub>	OH	-	-	4
<b>50<sup>a</sup></b>	3-cyclohexylpropanamide	OH	H	H	-	43
<b>57<sup>a</sup></b>	(3-methylthio)acetamide	OH	H	H	-	20
<b>64<sup>a</sup></b>	2-phenylbutamide	H	H	H	-	11
<b>74<sup>a</sup></b>	(4-bromophenyl)acetamide	H	H	H	-	17
<b>76<sup>a</sup></b>	(4-bromophenyl)acetamide	OH	H	H	-	29
<b>80<sup>a</sup></b>	(4-bromophenyl)acetamide	OCH <sub>3</sub>	OCH <sub>3</sub>	H	-	8

<sup>a</sup>Hit compounds from main screen. The 2D molecular structures of the benzamidobenzoic acids **50**, **57**, **64**, **74**, **76** and **80** are provided in Figure A2 in the appendix.

In general, the benzamidobenzoic acids displayed weak activities, the strongest is hit **82**. Halogen-substituted phenyl (**82\_1** to **3**) displayed relatively weak activities (11 - 27% inhibition). An improvement in inhibition was seen for hydroxyl substitution at R<sub>5</sub> (**82\_9** to **10**). Dual alkoxy – halogen substituted phenyl did not improve the activity (**82\_6** to **12**). Alkyl and alkoxy substituted phenyl (**82\_10** to **15**, **17** and **18**) displayed inhibition ranging from 16 – 35% with hydroxyl at R<sub>5</sub>. It can be seen clearly that hydroxyl substitution at R<sub>5</sub> is crucial for improved SAR, *e.g.*, substantial loss of activity for methoxy substituted phenyl (**82\_16**) with hydrogen at R<sub>5</sub>. Shifting the carboxylic acid from an *o*- to *m*-substituted position is detrimental to activity.



**Figure 4.13.** The docked configuration of hit **82** in the PC-PLC<sub>Bc</sub> binding site. (A) The protein surface is rendered where red and blue areas depicts negative and positive partial charges whereas grey areas are neutral. (B) Hydrogen bonding interactions depicted as green dotted lines of hit **82** with surrounding amino acid residues Ala3, Glu4 and Glu146. Coordination to Zn<sup>2+</sup> ions are depicted as red dotted lines. The positions of hydrophobic amino acids Phe66, Phe70 and Tyr79 are shown. The hydrogen atoms were suppressed.

Molecular docking revealed similar scores and plausible binding mode for the benzamidobenzoic acids. The binding mode is represented by the most active hit **82** shown in Fig. **4.13**. The carboxylic acid moiety is oriented towards the positive region where the three Zn<sup>2+</sup> ions are situated forming coordinate bond and aromatic moieties occupied lipophilic grey regions. The methoxybenzene is positioned near hydrophobic amino acid residues Phe66, Phe70 and Tyr79, where they are expected to be stabilised by hydrophobic contacts and  $\pi$ - $\pi$  interactions. The hydroxyl on the benzoic acid moiety is accommodated by hydrogen bonds with Ala3 and Glu4, a finding that explains improved activity when R<sub>5</sub> = hydroxyl; interaction with Ala3 is frequently seen with the morpholine moiety of the 2-morpholinobenzoic acid series, most likely a favourable interaction point that stabilises the enzyme-inhibitor complex and interaction with Glu4 can perturb the formation of base in the cleavage mechanism.<sup>123</sup> An additional hydrogen bond was seen with Glu146, an amino acid that ligates to Zn2 (see Fig. **4.2**). It is postulated that this interaction resulted in an additional mechanism that perturbed Zn2 – Glu146 ligation.

#### 4.2.2.3 Experimental IC<sub>50</sub> values

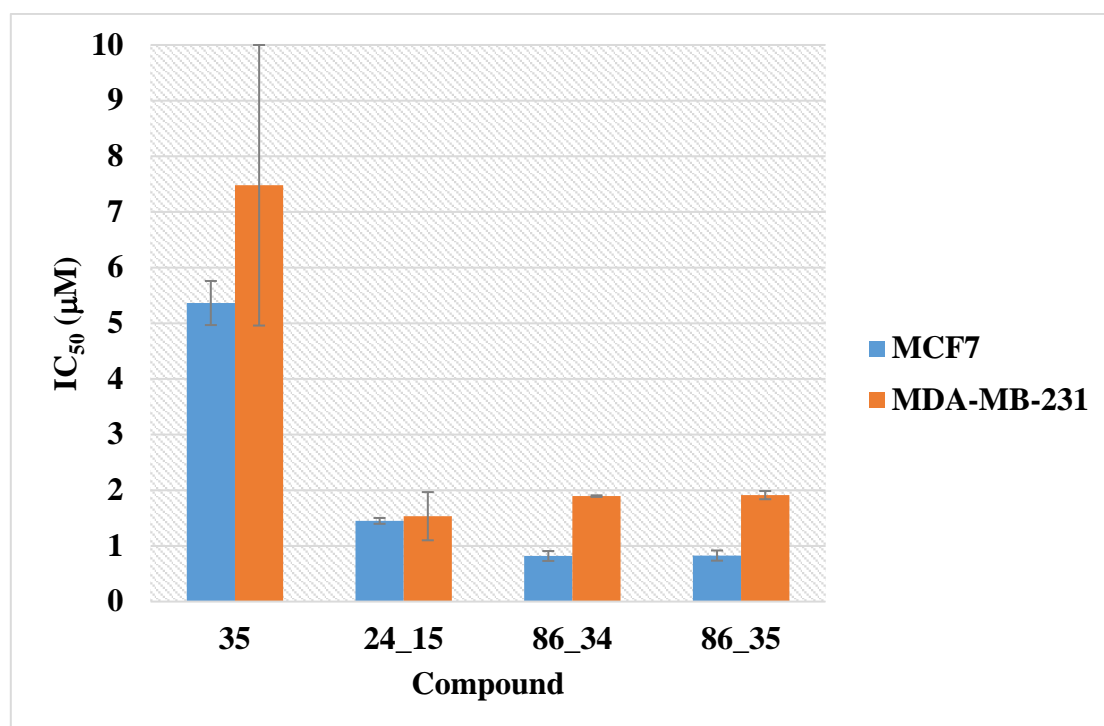
The pilot and main screens resulted in the 2,3,4,9-tetrahydro-1*H*-pyrido[3,4-*b*]indoles and 2-morpholinobenzoic acids that displayed the most promising inhibition for PC-PLC, *i.e.*, they demonstrated >50% inhibition. Two compounds, **24\_10** and **15**, from the 2,3,4,9-tetrahydro-1*H*-pyrido[3,4-*b*]indoles and five compounds, **86** and derivatives **86\_4**, **5**, **34** and **35**, from the 2-morpholinobenzoic acid series were chosen and their IC<sub>50</sub> values were measured (Table 4.6). It is shown that only **24\_15** displayed weaker activity than D609, the reference compound, whereas the most active 2,3,4,9-tetrahydro-1*H*-pyrido[3,4-*b*]indole **24\_10** showed comparable potencies to the 2-morpholinobenzoic acids. Derivatives **86\_34** and **35** displayed the strongest IC<sub>50</sub> ~ 1 μM, 2 - 3 times more potent than analogues **86**, **86\_4** and **5**.

**Table 4.6.** Inhibition (IC<sub>50</sub>) of PC-PLC by the 2,3,4,9-tetrahydro-1*H*-pyrido[3,4-*b*]indole and 2-morpholinobenzoic acid derivatives. The IC<sub>50</sub> values are the averages of five replicates using the Amplex Red biochemical assay. Six different concentrations were used to derive the IC<sub>50</sub> values: 0, 0.1, 0.4, 1.1, 3.3 and 10.0 μM. Compound **35** (D609) is the positive control.

Compound	IC <sub>50</sub> (μM)
<b>35</b>	7.8
<b>24_10</b>	1.8
<b>24_15</b>	9.1
<b>86</b>	2.7
<b>86_4</b>	2.0
<b>86_5</b>	3.1
<b>86_34</b>	1.0
<b>86_35</b>	1.1

#### 4.2.2.4 Cell proliferation assay

The 2,3,4,9-tetrahydro-1*H*-pyrido[3,4-*b*]indole derivative **24\_15** and the two most active 2-morpholinobenzoic acid derivatives, **86\_34** and **35**, were tested for the effect on cell proliferation in two breast cancer cell lines: MCF7 and MDA-MB-231. The thymidine incorporation assay<sup>140</sup> was employed. The assay relies on the uptake of radioactive-labelled [<sup>3</sup>H]-thymidine to be incorporated into DNA during mitosis followed by measurement of radioactivity levels that gives an indication of cell proliferative activities.<sup>140</sup> The results are shown in Figure 4.15. Results from the MCF7 cell lines reflects the IC<sub>50</sub> (μM) values from the Amplex Red biochemical assay, *i.e.*, 2-morpholinobenzoic acid derivatives **86\_34** and **86\_35** displayed the strongest effects followed by **24\_15** and **35**. On the other hand, it was seen **24\_15** displayed the strongest effect in MDA-MB-231 cell lines followed closely by derivatives **86\_34** and **86\_35**, whereas the reference compound **35** demonstrated the least antiproliferative activity.



**Figure 4.15.** Cell proliferation activities (IC<sub>50</sub> in μM) for compounds **D609**, **24\_15**, **86\_34** and **86\_35** on the MCF7 (blue) and MDA-MB-231 (orange) breast cancer cell lines.



## 4.3 Methodology

### 4.3.1 Molecular Modelling and Virtual Screening

The crystal structure was obtained from Protein Data Bank (PDB)<sup>10</sup> ID: 1AH7<sup>108</sup> with resolution 1.50 Å. The Scigress v2.6 program<sup>79</sup> was used to prepare the crystal structure for docking, *i.e.*, hydrogen atoms were added and crystallographic water molecules removed. The centre of PC-PLC<sub>Bc</sub> binding pocket was defined as the position of Zn<sup>2+</sup> ion (x = 42.482, y = 22.996, z = 8.556) with 10 Å radius. The basic amino acids lysine and arginine were defined as protonated. Furthermore, aspartic and glutamic acids were assumed to be deprotonated. The GoldScore (GS),<sup>19</sup> ChemScore (CS),<sup>22-23</sup> ChemPLP<sup>22, 24</sup> and ASP<sup>25</sup> scoring functions were implemented to validate the predicted binding modes and relative energies of the ligands using the Genetic Optimisation for Ligand Docking software package (GOLD) v5.4.0. Its capabilities to identify bioactive compounds in virtual screening projects are well-documented in the literature.<sup>19, 141-142</sup> The vHTS was conducted with the ChemBridge diversity collection of  $5 \times 10^4$  entities. Substructure and Tanimoto similarity methods<sup>78</sup> were used to identify structurally similar compounds through commercially available databases (details in results and discussion section 4.2).

### 4.3.2 Verification of hits

The hits were verified using the Amplex Red biochemical assay<sup>128</sup> and cell proliferation activities of the actives were confirmed following the thymidine incorporation assay protocol.<sup>140</sup> The work was carried out by Dr. Euphemia Leung and Dr. Chris Sun Xu at the Auckland Cancer Research Centre, Auckland.

## 4.4 Future Works and Conclusion

The aim to find novel inhibitors of mammalian PC-PLC was successful based on a wealth of active chemical series from the virtual screen. The 2-morpholinobenzoic acids was identified to be the most potent series with derivatives without the carboxylic acid functional group (**86\_34** and **35**) displayed the strongest inhibition effect ( $IC_{50} \sim 1 \mu M$ ) based on tests from Amplex Red assay. This shows that zinc coordinating groups are not required for inhibition activity and the interactions with the surrounding amino acids seemed to be more important. The 2,3,4,9-tetrahydro-1*H*-pyrido[3,4-*b*]indole series displayed the second best inhibition with hydroxyl and alkoxy substitutions on the phenyl ring being most active. Molecular modelling predicted the 1*R*,3*S*-diastereomer as the most active diastereomer, however, this require further synthetic studies followed by experimental testing for verification. Both the 2-morpholinobenzoic acid and 2,3,4,9-tetrahydro-1*H*-pyrido[3,4-*b*]indole derivatives displayed stronger antiproliferative effects than the known inhibitor, D609, which were verified in the MCF7 and MBA-MB-231 breast cancer cell lines. Experimental testing and molecular modelling studies on the more modest benzamidobenzoic acid series revealed that Ala3 and Glu4 interactions with *p*-hydroxyl on the benzoic acid moiety is a key interaction for activity. Their antiproliferative activities are pending to be tested.

These preliminary studies provided positive outcomes of the project. Future works heavily rely on exploring the SAR through synthesis and biological testing aided by molecular modelling to optimise their potencies. The main focus is to develop the two most active families: 2-morpholinobenzoic acid and 2,3,4,9-tetrahydro-1*H*-pyrido[3,4-*b*]indoles. Following their optimisation, hopefully their potencies qualify for *in vivo* tests to investigate toxicity and therapeutic effects prior to clinical trials.

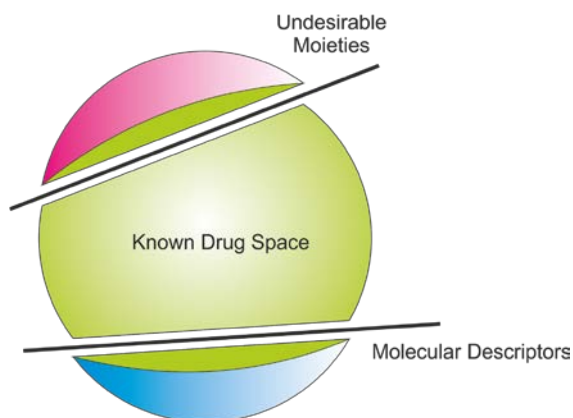
# *Chapter 5*

The Development of an Index to Identify Quality  
Candidates for Drug Discovery Projects

# 5.1 Introduction

## 5.1.1 Known Drug Index

Over the years, an open question that left medicinal chemists to ponder is how drugs differ from non-drugs? In order to distinguish the two categories, attempts have been made to try and distinguish properties of drugs and non-drugs. Several approaches have been developed to distinguish drugs from non-drugs by quantification methods (see introduction section 1.6). This chapter is based on the development of a simple method to calculate for *Known Drug Index* (KDI), a value that quantitatively measure the quality of a drug candidate based on a collection of clinically approved drugs. KDI was derived from the concept of KDS which includes all drugs that have passed clinical trials and remained in medical use (see graphical representation Fig. 5.1).<sup>143</sup> By simply using KDI as a guide, medicinal chemists can identify benign areas in chemical space to develop improved lead compounds for drug development. It has been shown that drugs have larger parameters both for the molecular descriptors and unwanted molecular moieties than non-drugs.<sup>144-146</sup> This is attributable to their need for selective interactions with protein targets and certain requirements to penetrate pharmacokinetic barriers.<sup>147</sup>



**Figure 5.1.** A graphical representation of KDS with undesirable moieties and molecular descriptors.

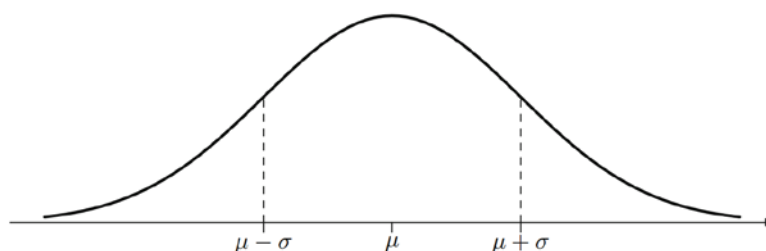
### 5.1.2 Molecular Descriptors

The analysis of chemical structural information can be made through molecular descriptors.<sup>148-149</sup> These are numerical values that characterises properties of a molecule.<sup>148</sup> The idea of using molecular descriptors to describe drugs is not new. To list a few insightful examples; Lipinski have established a criteria in selecting drugs that favour oral bioavailability,<sup>147</sup> molecular flexibility and polar surface area (PSA) have been shown to be important for drug distribution and target binding,<sup>31</sup> and density functional theory-derived (DFT) descriptors used to define KDS.<sup>32</sup> Additionally, molecular frameworks of known drugs,<sup>33-34</sup> and certain molecular moieties were suggested as guides to potentially improve pharmacokinetic profiles of drugs and eliminate toxic and promiscuous moieties.<sup>17, 144</sup> Perhaps, the most widely accepted descriptors used to define drug-like molecules are the six mainstream descriptors: MW, Log P, HD, HA, RB and PSA. Together, they are capable in capturing the bulk properties of molecules that allows rapid identification of a molecule's '*drug-likeness*'.

### 5.1.3 Gaussian Distribution

The human body is a complex system with a vast range of biomolecular targets and with a spectrum of diseases in the human population, one can expect drugs to at the very least differ slightly in their properties. Nonetheless, drugs that were accepted for clinical use share two similar features: they have an acceptable pharmacokinetic profile and their side effects are tolerable. Hence, the statistical distribution profiles for molecular descriptors can be mapped for known drug collections to observe for similarities in drug properties and differences when compared to non-drug collections. The Gaussian/normal distribution (Fig. 5.2) is suitable for mapping the statistical variations in populations. The peak of the distribution is the most

frequently occurring variable ( $\mu$ ) and the horizons are broadened depending on the variations ( $\sigma$ ) in the population.



**Figure 5.2.** A Gaussian distribution curve. Parameter  $\mu$  denotes the most commonly occurring variable and  $\sigma$  denotes the variation in the population (Diagram obtained from Heumann *et al.*<sup>150</sup>)

$$f(x) = ae^{\frac{-(x-b)^2}{2c^2}} \quad (5)$$

The distribution is described by the Gaussian function (5) where parameters  $a$ ,  $b$ , and  $c$  denotes the peak height, centre of peak and standard deviation respectively.

### 5.1.4 Qikprop

The Qikprop software was developed with a goal to accurately predict molecular descriptors that could evaluate the pharmacokinetic characteristics of a molecule. The software utilised Monte Carlo simulations in the presence of Optimised Potentials for Liquid Simulation force field (OPLS-AA)<sup>151-152</sup> that includes all atoms of a solute surrounded by water molecules. The force field was parameterised to reproduce experimental thermodynamic and structural data in fluids, making it suitable to simulate fluidic systems such as the human body.<sup>151-152</sup> The descriptors were calculated based on different configurations of the molecules and the average calculated as the output, *e.g.*, the HD count was averaged from different configurations of a solute interacting to the water molecules, thus integers may not be shown as the output. The Qikprop software was employed in this project as a tool for calculating molecular descriptors. Its reliability is well-established.<sup>153</sup>

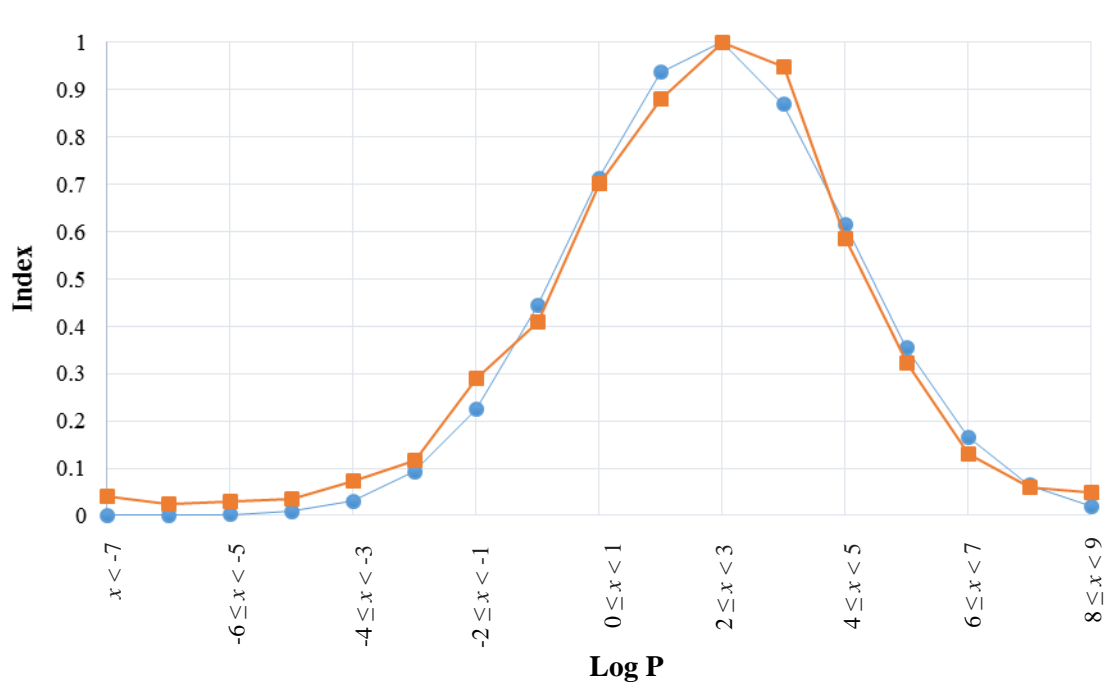
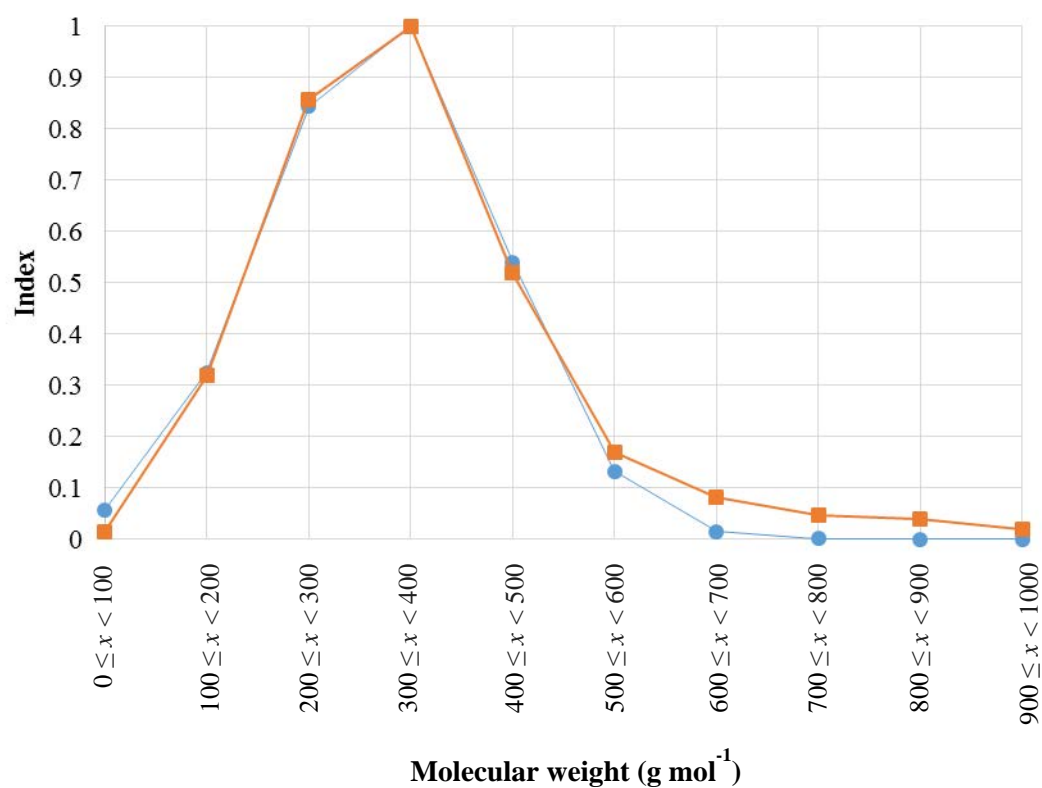
## 5.2 Results

In this study, we explored a simple approach to calculate the drug-like quality of molecules based on the statistical distributions of known drugs in clinical use. The known drug collection was obtained from the following websites and resources: US Food and Drug Administration (FDA),<sup>154</sup> Health Canada,<sup>155</sup> the UK Medicines and Healthcare products Regulatory Agency (MHRA),<sup>156</sup> China Food and Drug Administration (CFDA),<sup>157</sup> Australian Medicines Handbook (AMH)<sup>158</sup> and Japan's Pharmaceuticals and Medical Devices Agency (PMDA)<sup>159</sup> (see methodology for more details on data collection). The statistical distribution profiles for MW, log P, HD, HA, RB and PSA were normalised and fitted to the Gaussian function (5). The derived parameters  $a$ ,  $b$  and  $c$  were used to calculate a score for each descriptor, which were summed up into one number giving a quantitative index for the *quality* of the molecule under consideration.

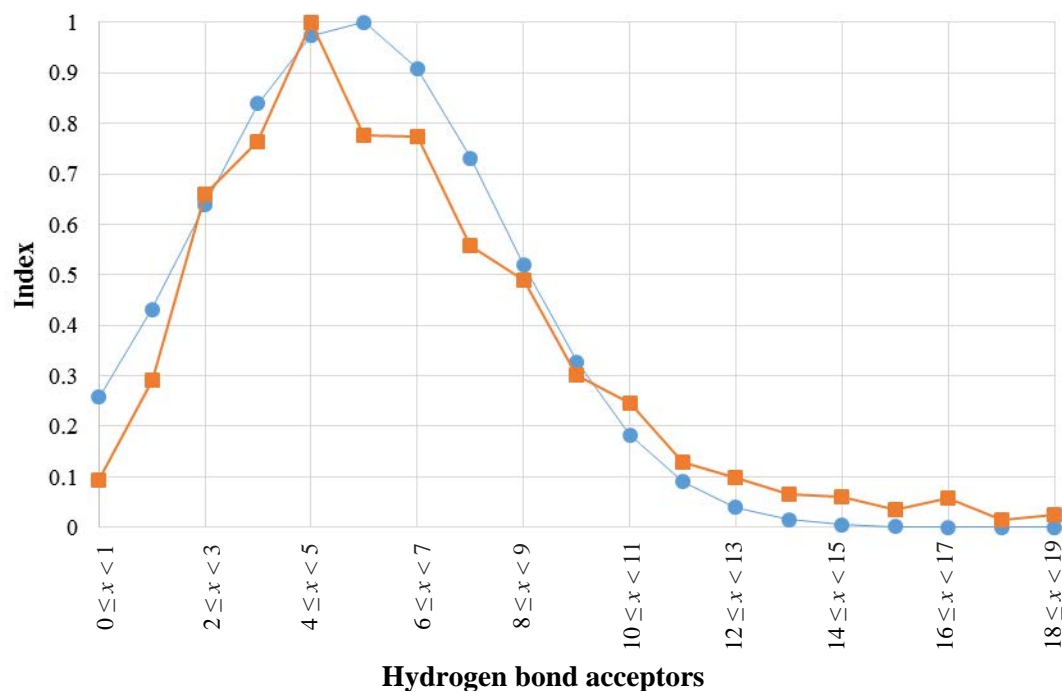
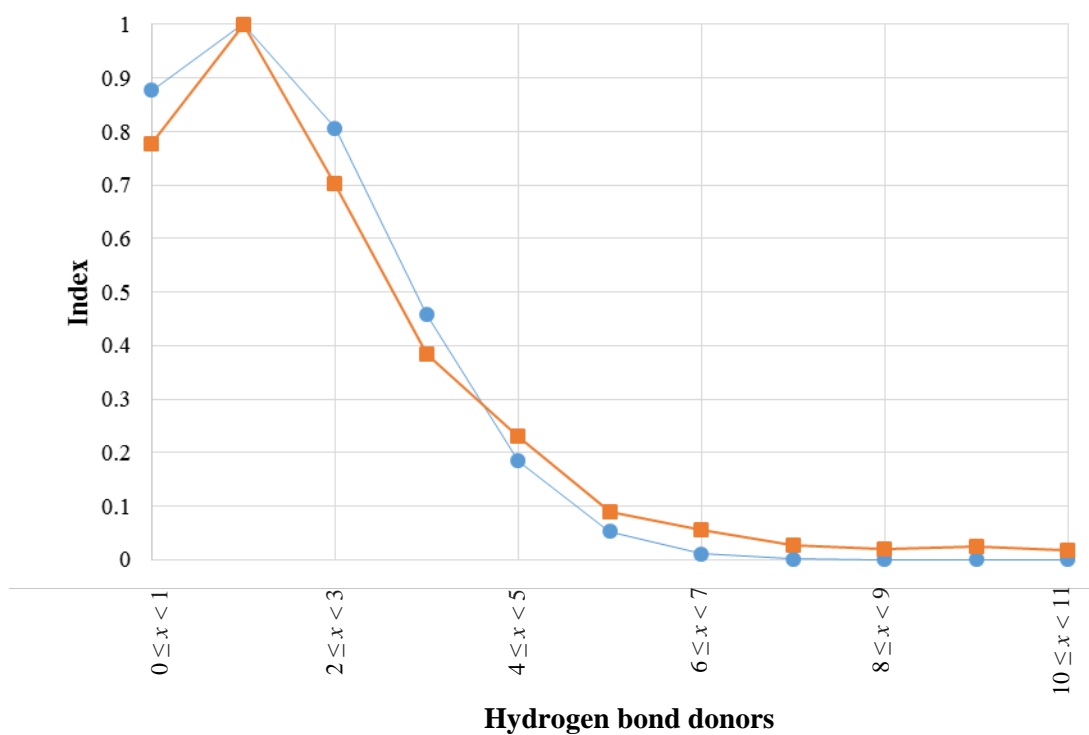
### 5.2.1 Distribution of Physicochemical Properties

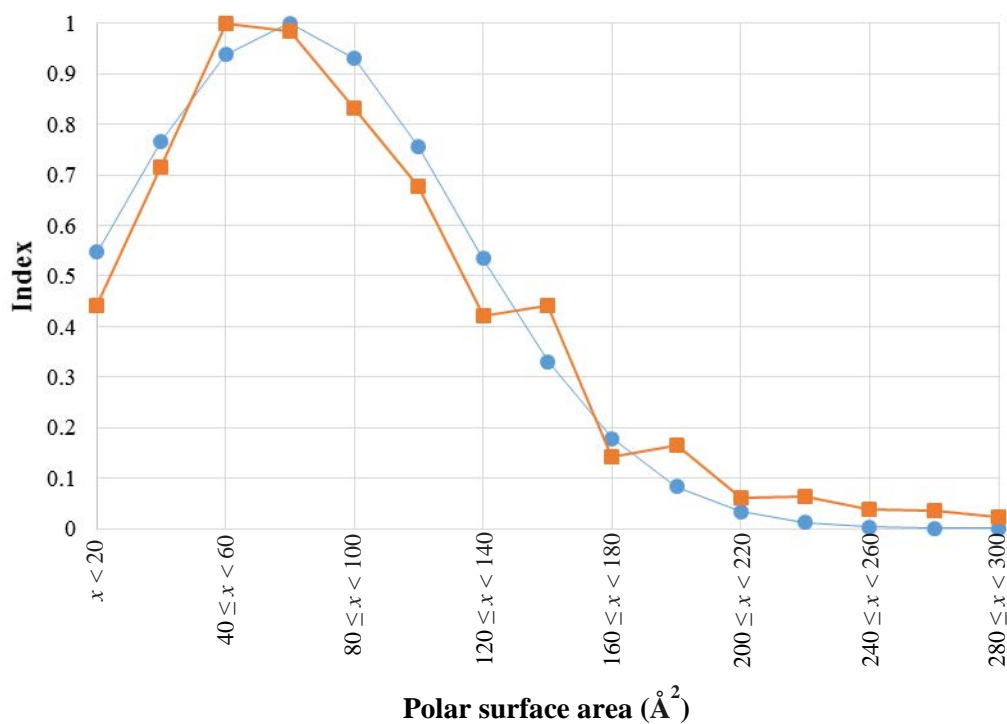
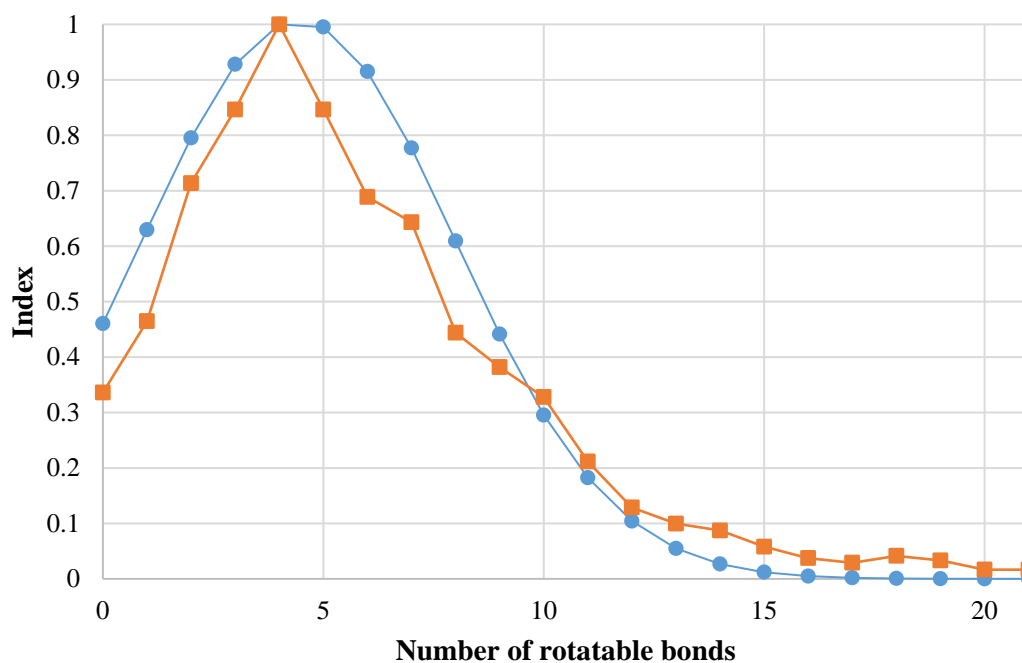
The statistical distributions of each descriptor for the collection of 1880 marketed drugs are shown in Fig. 5.3. The molecular descriptors were plotted as domain boundaries ( $x$ ) against percentages of drugs,  $f(x)$ . The input  $x$ -value for equation 5 is the next whole number greater than the domain, *e.g.*, for the molecular weight domain of  $300 \leq x < 400$ , the input  $x$ -value is 400. The percentages of drugs,  $f(x)$ , were normalised to give an 'Index' value for each molecular descriptor, *i.e.*, the height of the distribution's apex was defined as 1 ( $a = 1.00$  and Index = 1.00). The basic assumption was made that the most frequent descriptor value ( $x$ ) gives the best drug-like quality. The tails of the distributions with few data points were excluded from the graphs, thus the indexes for these points were estimated to be zero. The derived non-linear parameters  $b$  and  $c$  (Table 5.1) can be used to calculate the Index of any

molecule by applying the Gaussian function (5). In this equation  $b$  is the most occurring value in the statistical distribution and  $c$  is the standard deviation









**Figure 5.3.** The statistical distributions of physicochemical properties of the 1880 marketed drugs. The Gaussian fitted plot is coloured blue (circles) and the raw data is represented in orange (squares). The y-axis is given as an Index normalised to 1.00 at the apex of the plot.

The spread of the distribution curves indicated substantial variation of the molecular descriptors in marketed drugs, which can be explained in terms of differences in administration routes, doses, formulations and biochemical targets. Excellent fits of the Gaussian function to the distributions were obtained and is reflected by the low residual sum of squares (*RSS*) as shown in Table 5.1. The *RSS* is expressed by equation 6 and is used to measure the variance between the predicted Gaussian model  $f(x)$  and the raw data ( $y$ ). Moreover, the shapes and patterns observed are in line with previously reported distributions for the molecular descriptors for drugs and drug-like compounds.<sup>37, 39, 153, 160</sup>

$$\text{Residual sum of squares} = \sum_{i=1}^n (y_i - f(x_i))^2 \quad (6)$$

**Table 5.1.** The parameters,  $b$  and  $c$ , calculated by non-linear least squares curve fitting according to equation 5. The apex,  $a$ , was normalised to 1.00 for all six descriptors. The *RSS* (residual sum of squares) values indicate quality of the fits, low values indicating good fits. Parameter  $b$  is the mode or most occurring in the distribution

	MW	Log P	HD	HA	RB	PSA
<b>b</b>	371.76	2.82	1.88	5.72	4.44	79.40
<b>c</b>	112.76	2.21	1.70	2.86	3.55	54.16
<b>RSS</b>	0.04	0.04	0.10	0.16	0.12	0.11

The largest proportion of drugs have MW within  $300 \leq x < 400 \text{ g mol}^{-1}$  range accounting for 31.7%. Short chain drugs with MWs below  $100 \text{ g mol}^{-1}$  accounted for only 0.5%, *e.g.*, acetohydroxamic acid (MW =  $75.1 \text{ g mol}^{-1}$ ) used in treating urinal tract infections. A gradual decline in KDI was seen after  $400 \text{ g mol}^{-1}$  and a tail occurring around  $\sim 700 \text{ g mol}^{-1}$  indicative of a non-favourable region in chemical space. Larger drugs tend to fall into categories of natural products, peptides, glycosides and penicillin-derived antibiotics, some of which are orally available. It is undeniable that most drugs comply with Lipinski's  $500 \text{ g mol}^{-1}$  limit.<sup>29</sup> This is plausible as it is more difficult for large molecules to diffuse through biological

membranes.

Log P is a measure of molecular lipophilicity. As shown in Fig. 5.3, the most favourable log P region is between  $2 \leq x < 3$  accounting for 17.6% and tails emerge at  $\sim 7$  and  $\sim 9$ . This finding is in agreement with Oprea,<sup>160</sup> which used a compound collection consisting of applied, or studied medicinal agents, in humans from the Comprehensive Medicinal Chemistry (CMC)<sup>161</sup> database to compare property distributions with the Available Chemical Directory (ACD)<sup>162</sup> and MACCS-II Drug Data Report (MDDR)<sup>163</sup> databases. The ACD library consist of a set of starting materials used for drug synthesis whereas the MDDR contains small molecules under development, which were synthesised and screened for activity. Furthermore, Oprea<sup>160</sup> reported that the most frequent log P = 2.1 for the sample drug collection in the CMC database comparing to log P = 3.2 in ACD and log P = 3.4 in MDDR databases, which suggests that approved drugs have smaller log P than non-drugs. Furthermore, evaluation of a drug database by Biswas *et al.*<sup>37</sup> suggested log P = 2 has the best drug-like score.

HDs and HAs are associated with the polarity of molecules and are important for hydrogen bonding with biomolecular targets. The most frequent HDs are between  $1 \leq x < 2$  accounted for 29.4% of the total collection. A few reports are in line with this finding, *e.g.*, Xu and Stevenson<sup>39</sup> evaluated a collection of CMC drugs resulted HD = 1 to be most frequent and Oprea<sup>160</sup> showed that a collection of MDDR compounds, after removal of reactive species, resulted in HD = 1 being the most frequent. The distribution of HDs is different from the other descriptors since fewer HDs tend to be favourable for drugs. In comparison, the HA distribution is more similar in shape to the MW, RB and PSA distributions. The most prevalent HAs are between  $4 \leq x < 5$  with a tail occurring at  $14 \leq x < 15$ . Note that the Gaussian model estimated  $5 \leq x < 6$  to be more prevalent due to the positive skewness of the distribution. However, evaluations of CMC drugs by Xu and Stevenson<sup>39</sup> showed most

common HA = 3 whereas the same ACD and MDDR compounds from Oprea<sup>160</sup> have HA = 3 and 4, respectively. The HAs in these reports were relatively lower than those found here most likely due to the inclusion of large drugs that are in clinical use.

RBs are associated with molecular flexibility, influencing oral bioavailability,<sup>164</sup> and is important for reducing the entropic penalty for ligand-target binding.<sup>165</sup> It was found that RB = 4 is most frequent accounting for 12.7% followed by RB = 5 with 10.7% of the total. Clearly, this trend has similar distribution shapes as MW, HA and PSA. The number of favourable RBs in drugs is somewhat controversial with different values given in the literature ranging from 3 to 5 by Oprea,<sup>160</sup> 2 to 3 was suggested by Biswas *et al.*<sup>37</sup> and Xu and Stevenson<sup>39</sup> even proposed RB = 9 to be most favourable based on the CMC library.

PSA is defined as the surface sum over all polar atoms and plays an important role in the pharmacokinetic profile of drugs. According to Figure 5.3 PSA between  $40 \leq x < 60 \text{ \AA}^2$  is the most favourable accounting for 16.7% of the total and with a tail appearing at  $260 \leq x < 280 \text{ \AA}^2$ . The PSA range and trend is comparable to reported by Biswas *et al.*,<sup>37</sup> which suggested  $\sim 50 \text{ \AA}^2$  to have the best drug-like score.

### 5.2.2 Calculation of Known Drug Index

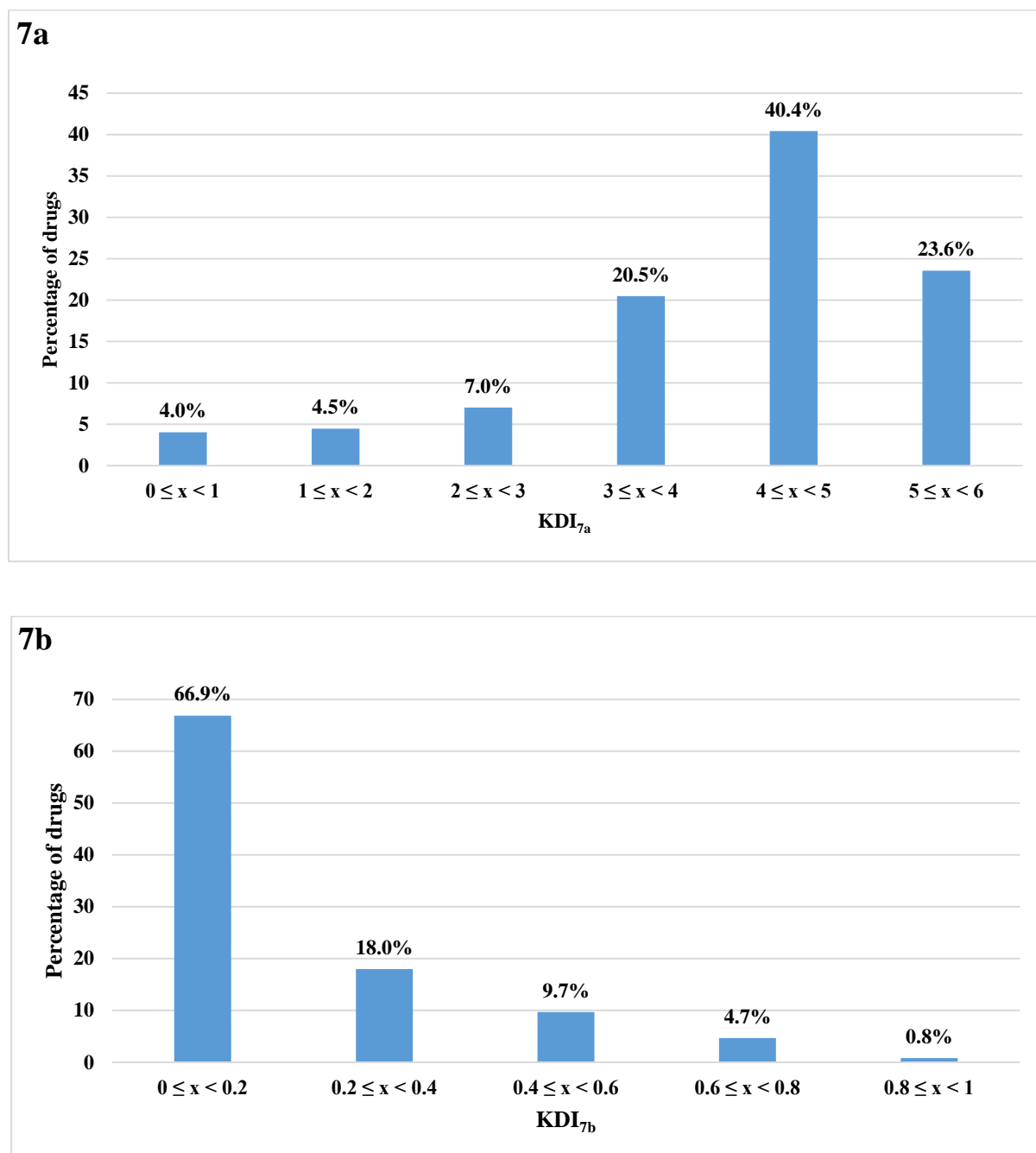
In order to generate an index for the overall quality of a compound, the Indexes for each of the six descriptors were combined to give the Known Drug Index (KDI). Two approaches were explored: summation as shown in equation 7a and multiplication in equation 7b.

$$\text{KDI}_{7a} = \text{I}_{\text{MW}} + \text{I}_{\log P} + \text{I}_{\text{HD}} + \text{I}_{\text{HA}} + \text{I}_{\text{RB}} + \text{I}_{\text{PSA}} \quad (7a)$$

$$\text{KDI}_{7b} = \text{I}_{\text{MW}} \times \text{I}_{\log P} \times \text{I}_{\text{HD}} \times \text{I}_{\text{HA}} \times \text{I}_{\text{RB}} \times \text{I}_{\text{PSA}} \quad (7b)$$

Equation 7a is based on the Index summation of the six mainstream descriptors, thus  $\text{KDI}_{\text{max}} = 6.00$  is assumed to have ideal qualities of a drug. The idea of multiplication in equation 7b has been applied in various scenarios such as for calculating the impact of population growth

on the environment<sup>166-167</sup> and estimating the number of civilisations in the Milky Way galaxy.<sup>168</sup> The ideal quality of a drug candidate using equation **7b** is 1.00. The results of deriving KDI using both equations **7a/b** for the 1880 drugs are shown in Fig. **5.4**.

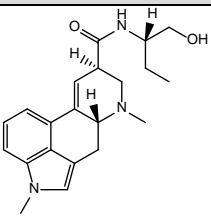
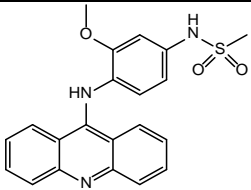


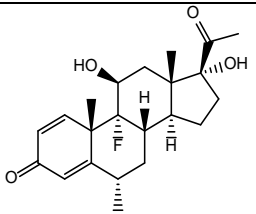
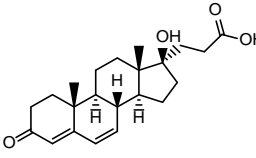
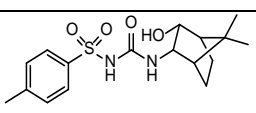
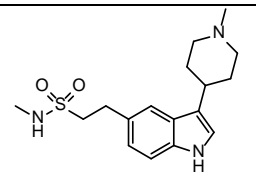
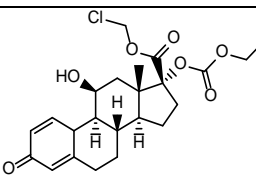
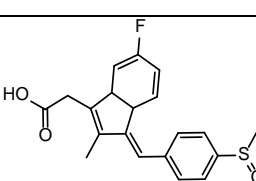
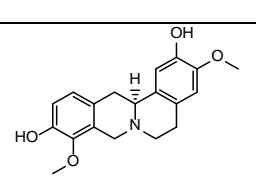
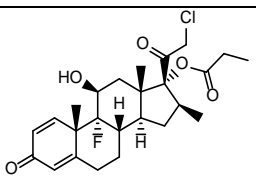
**Figure 5.4.** The Known Drug Index (KDI) distributions of the total drug collection using equation **7a** and **7b**. The former gives a maximum at 4 – 5 but the latter results in most of the drugs having < 0.2 score.

For equation **7a**, the average  $KDI_{7a}$  is  $4.08 \pm 1.27$  with most of the drugs between  $4 \leq x < 5$  (40.5%) followed by  $5 \leq x < 6$  (23.6%). Twenty-six drugs (1.4%) had  $KDI_{7a} \sim 0$  and these are peptides, glycosides, natural products and their derivatives. Using equation **7b** the average  $KDI_{7b}$  is only  $0.18 \pm 0.20$ . The largest percentage of drugs 66.9% is between  $0 \leq x < 0.2$  followed by  $0.2 \leq x < 0.4$  (18.0%). 36.4% of drugs have  $KDI_{7b} = 0$ , *i.e.*, no drug-quality at all according to this equation. It is apparent that equation **7b** is less suitable for calculating drug indexes than **7a** considering the large proportion of marketed drugs having no drug-like quality at all, *e.g.*, acclarubicin is a drug approved for treating cancer by the MHRA that resulted  $KDI_{7b} = 0$ , simply because the  $I_{HA}$  is nil (0). Nonetheless, equation **7b** has its merits in distinguishing the most balanced drugs from the collection and can be used a tool to select the best balanced compounds from large screening collections.

The top 10 highest scoring drugs are shown in Table **5.2** with their indications, drug classes, and administration routes. Both methods identify the same drugs with  $KDI_{7a/b} \leq 5.89$  and 0.90, respectively. A range of different biomolecular targets and indications is seen for the drugs, interestingly with four steroids and the top drug an ergot derivative. Half of the top 10 drugs are orally administered with various other administration routes present.

**Table 5.2.** The top 10 drugs with the highest  $KDI_{7a/b}$  scores. Their drug class, indication and route of administrations are shown.

Drug	Molecular Structure	$KDI_{7a}$	$KDI_{7b}$	Class	Indication	Administration
Methysergide		5.89	0.90	Serotonin 5-HT receptors	Migraine	Oral
Amsacrine		5.89	0.89	DNA intercalating agent	Leukaemia	Intravenous

<b>Fluorometholone</b>		5.88	0.88	Corticosteroid	Ophthalmic inflammation	Ophthalmic
<b>Canrenoic acid</b>		5.86	0.87	Aldosterone antagonist	Heart failure	Intravenous
<b>Glibornuride</b>		5.85	0.85	Hypoglycaemic agent	Type II diabetes	Oral
<b>Naratriptan</b>		5.84	0.85	5-HT receptor agonist	Migraine	Oral
<b>Loteprednol</b>		5.84	0.85	Corticosteroid	Ophthalmic inflammation	Ophthalmic
<b>Sulindac</b>		5.83	0.84	NSAIDs	Arthritis	Oral
<b>Stepholidine</b>		5.82	0.83	D1 receptor agonist and D2 antagonist	Migraine and hyperactivity disorder	Oral
<b>Clobetasol</b>		5.81	0.82	Corticosteroid	Skin inflammation and dermatosis	Topical

The effects of the descriptors were investigated where each one was systematically removed and the corresponding  $KDI_{7a/b}$  values calculated. By removing a descriptor results in  $KDI_{7a}$ ,



$\text{max} = 5.00$  whilst  $\text{KDI}_{7b, \text{max}} = 1.00$  remains unchanged. The average  $\text{KDI}_{7a/b}$  average values are shown in Table 5.3 and are relatively unchanged between 3.36 to 3.45 for **7a**, and 0.21 to 0.24 for **7b** suggesting that the molecular descriptors are equally important in determining the quality of drug candidates.

**Table 5.3.** The averages of  $\text{KDI}_{7a/b}$  with standard deviations calculated after systematic removal of each descriptor with their standard deviations. Only modest changes are observed indicating equal contribution of the descriptors.

	MW	Log P	HD	HA	RB	PSA
<b>KDI<sub>7a</sub></b>	$3.45 \pm 1.09$	$3.40 \pm 1.10$	$3.37 \pm 1.10$	$3.43 \pm 1.03$	$3.40 \pm 1.08$	$3.36 \pm 1.05$
<b>KDI<sub>7b</sub></b>	$0.24 \pm 0.23$	$0.22 \pm 0.22$	$0.22 \pm 0.23$	$0.22 \pm 0.22$	$0.23 \pm 0.23$	$0.21 \pm 0.21$

### 5.2.3 Robustness of the $\text{KDI}_{7a/b}$ indexes

In this study 3122 building blocks and 5231 molecules from the diversity library were obtained from the ChemBridge Corporation website<sup>126</sup> as well as 7883 natural products from the InterBioScreen Ltd (IBS).<sup>130</sup>

The average  $\text{KDI}_{7a/b}$  were calculated for each library and the results are shown in Table 5.4 as well as the averages of the descriptors. The ChemBridge diversity library has the highest  $\text{KDI}_{7a} = 5.14$  followed by the IBS library  $\text{KDI}_{7a} = 4.53$  compared to total drugs  $\text{KDI}_{7a} = 4.08$  with the Building Blocks collection with  $\text{KDI}_{7a}$  only at 3.26. It was seen that libraries with prior Lipinski's and Veber's optimisations have a higher  $\text{KDI}_{7a}$ , *i.e.*, the  $\text{KDI}_{7a}$  of the drug collection increased from 4.13 to 4.80 when these filters were applied as shown in Table 5.4. Ideally, the drug collection should score the highest  $\text{KDI}_{7a/b}$  since it is composed of drugs in clinical use. KDS includes drugs with different administration routes such as IV and topical, which explains to some extent the lower  $\text{KDI}_{7a/b}$  values of the drugs as compared to the collections optimised to adhere to drug-like chemical space. In general, drugs in clinical use are very varied, *e.g.*, the lithium ion is a useful drug for treating psychiatric conditions and

the serendipity phenomenon has played a major role in the discovery of the medication in use today.<sup>169</sup> When the KDI<sub>7b</sub> index is considered, the same trend is seen as for its KDI<sub>7a</sub> counterpart, *i.e.*, the filtered libraries have a better score and the building blocks generate the poorest result.

**Table 5.4.** The results of average KDI<sub>7a/b</sub> with standard deviations and the corresponding molecular descriptors for the compound libraries investigated.

	Total drugs	Total Drugs after drug-like filter	IBS natural products	ChemBridge drug-like molecules	ChemBridge building blocks
<i>No. of entries</i>	1880	1279	7883	5231	3122
<i>KDI<sub>7a</sub></i>	4.08 ± 1.27	4.64 ± 0.67	4.53 ± 0.89	5.14 ± 0.43	3.26 ± 0.68
<i>KDI<sub>7b</sub></i>	0.18 ± 0.20	0.24 ± 0.21	0.24 ± 0.20	0.41 ± 0.20	0.04 ± 0.07
<i>MW (g mol<sup>-1</sup>)</i>	374.7 ± 202.2	294.2 ± 85.0	351.8 ± 107.9	361.1 ± 55.0	186.4 ± 44.2
<i>Log P</i>	2.1 ± 2.6	2.0 ± 1.8	2.9 ± 2.0	3.0 ± 1.4	0.3 ± 1.3
<i>HD</i>	2.1 ± 2.6	1.4 ± 1.2	1.1 ± 1.3	0.9 ± 0.8	0.8 ± 1.0
<i>HA</i>	6.9 ± 5.6	5.0 ± 2.2	5.9 ± 3.2	6.5 ± 1.7	2.6 ± 1.6
<i>RB</i>	6.8 ± 6.9	4.2 ± 2.5	4.2 ± 3.5	4.6 ± 1.9	0.7 ± 1.4
<i>PSA (Å<sup>2</sup>)</i>	93.0 ± 77.4	67.7 ± 32.6	79.7 ± 42.1	67.5 ± 21.3	29.8 ± 18.8

In order to further check the robustness of the indexes they were correlated with the oral bioavailability of the drugs and the results are given in Table 5.5 with the averages of the molecular descriptors. Oral bioavailability of drugs was collated by Matuszek and Reynisson<sup>32</sup> for 708 marketed drugs with reported data; 455 were categorised as having a ‘high’ bioavailability ( $\geq 50\%$ ), 137 drugs as ‘medium’ (10 - 49%) and 116 drugs as ‘low’ ( $< 10\%$ ). The average KDI<sub>7a</sub> of the high category is 4.43, medium 4.08 and low is 2.74, giving a *very* clear trend for higher numbers for good oral bioavailability. The same trend is seen for the KDI<sub>7b</sub>, albeit less pronounced. This strong trend supports the use of the KDI<sub>7a/7b</sub> as a tool to gauge the quality of compound collections and individual molecules in drug design projects. A general trend is seen for the molecular descriptors that high oral bioavailability

corresponds to smaller values with the notable exception of log P, it remained unchanged for drugs with medium and high bioavailability.

**Table 5.5.** The relationship between oral bioavailability and the  $KDI_{7a/b}$  indexes. The values of the molecular descriptors are given as averages with standard deviations. Bioavailability is categorised into ‘high’ bioavailability ( $\geq 50\%$ ), ‘medium’ (10 - 49%) and ‘low’ ( $< 10\%$ ) based on 709 known drugs.

Descriptor	Bioavailability categories		
	<i>Low</i>	<i>Medium</i>	<i>High</i>
<b><i>KDI<sub>7a</sub></i></b>	$2.78 \pm 1.81$	$4.08 \pm 1.34$	$4.43 \pm 0.88$
<b><i>KDI<sub>7b</sub></i></b>	$0.09 \pm 0.19$	$0.15 \pm 0.18$	$0.21 \pm 0.20$
<b><i>MW (g mol<sup>-1</sup>)</i></b>	$536.6 \pm 343.4$	$371.5 \pm 148.8$	$320.2 \pm 115.3$
<b><i>Log P</i></b>	$0.1 \pm 3.8$	$2.2 \pm 2.3$	$2.2 \pm 2.0$
<b><i>HD</i></b>	$4.8 \pm 4.9$	$2.0 \pm 1.6$	$1.7 \pm 1.4$
<b><i>HA</i></b>	$12.4 \pm 9.8$	$7.1 \pm 4.4$	$5.6 \pm 3.3$
<b><i>RB</i></b>	$13.3 \pm 11.9$	$6.5 \pm 4.6$	$4.9 \pm 3.4$
<b><i>PSA (Å<sup>2</sup>)</i></b>	$180.0 \pm 139.2$	$91.4 \pm 49.0$	$75.1 \pm 44.1$

## 5.3 Discussion

The method presented here is an excellent mathematical tool in deriving well balanced compounds for drug development, each molecular descriptor is given an equal weight based on its statistical distribution. The two approaches to either sum up (**7a**) or multiply (**7b**) the individual Index values for each descriptor give interesting results as shown in Fig. 5.4. Equation **7a** is more benign since a poor result for one descriptor can be compensated by good results by the others. This is not the case for equation **7b** due to the nature of multiplying small numbers leads to even a smaller number and in the extreme a nil ‘0’ for one descriptor gives this result for  $KDI_{7b}$ . This explains why only 0.8% of the known drugs

have a score between 0.8 and 1.0 for  $KDI_{7b}$  and 23.6% of the drugs are in the highest category, 5 – 6, for  $KDI_{7a}$ .

The commercially available collections from ChemBridge (diversity) and IBS both get better scores based on the  $KDI_{7a/b}$  than the drug collection. It is exceedingly unlikely that any of these thousands of commercially available molecules will ever be tested in humans, let alone being a useful drug. Nevertheless, a superior  $KDI_{7a/b}$  is achieved and the question emerges why? The most plausible answer is that the molecular descriptors that are used do not fully capture the physicochemical properties required for a drug molecule to negotiate the human physiological milieu. There are two main perpendicular dimensions to be considered: in/stability and lipo/hydrophilicity. For the former it is estimated that 75% of all drugs are metabolised in the liver by the cytochrome P450 class of enzymes<sup>170</sup> and none of the six descriptors used have any bearing on the redox properties of the molecules. Interestingly, it is now known that the statistical distributions for the ionisation potentials (one-electron oxidation) and electron affinities (one-electron reduction) calculated using the DFT of known drugs are relatively narrow<sup>32</sup> indicating that a defined volume within chemical space is unique for drugs regarding their redox properties. Furthermore, it has been shown that the hydration energies ( $\Delta G_{hyd}$ ), again calculated using DFT, are clearly more exothermic for known drugs compared to collections of small organic molecules and drug-like compounds.<sup>171</sup> These DFT-derived descriptors correlated well with their experimental counterparts, firmly linking them to empirical physical properties but their application is limited by the substantial CPU required for their calculation. We are currently working on expanding the drug and reference collections in order to gauge the applicability of the DFT descriptors in locating favourable regions in chemical space with the aim of adding them to the  $KDI_{7a/b}$  equations to improve their ability to identify quality hits, leads and drug candidates.

## 5.4 Methodology

The lists of approved drugs were collected (see section 5.2). All non-biologically active molecules such as contrasting agents and excipients were deleted from the lists. Furthermore, large compounds with molecular weight greater than  $2000 \text{ g mol}^{-1}$ , elements such as zinc and iron, organometallic complexes, nutrient supplements, food additives, cosmetic substances and experimental drug candidates were excluded from the drug pool. The resulting 1880 drugs were matched on a Microsoft Excel 2013 spreadsheet. Duplicates were removed based on the generic names and molecular weights. The structures of the drugs were compiled from the DrugBank,<sup>172-174</sup> PubChem<sup>175</sup> and ChemSpider<sup>176</sup> databases in 3D format. Counter ions were removed and the molecules structurally optimised using the MM2<sup>80</sup> force-field in the modelling software Scigress v2.6.<sup>79</sup> The QikProp v4.6<sup>177</sup> software package was used to generate the descriptors: MW, log P, HD, HA, RB and PSA. The reliability of QikProp has been previously established for the descriptors used.<sup>153</sup> The results are given as additional information provided in the CD-ROM. The descriptors were plotted using the Gaussian function (see equation 5) in Microsoft Excel 2013. The non-linear parameters  $a$ ,  $b$ , and  $c$  were estimated through non-linear least squares curve fitting using the solver add-in function.

## 5.5 Future Works and Conclusion

In this work the idea is explored whether one number, an index, can be developed to reflect the quality of a drug candidate. This was achieved by fitting the Gaussian functions to the statistical distributions of six mainstream molecular descriptors of a collection of 1880

clinically approved drugs. In this way a number can be derived for a molecule reflecting its adherence to the drug distributions for each descriptors. Two methods were investigated in treating these numbers, *i.e.*, addition (equation **7a**) and multiplication (equation **7b**). The addition method gave useful results as it identified 84.5% of the known drugs with a score >3 and almost two thirds (64.0%) >4 out of 6 being the maximum. Due to the nature of multiplying small numbers the latter approach (**7b**) only gave 0.8% of the known drugs as >0.8 and 5.5% with >0.6 with 1 being the maximum. Both approaches are excellent to derive an optimally balanced drug molecule with methysergide, amsacrine and fluorometholone identified as the best. It was confirmed that drugs with good oral bioavailability are linked to the KDIs.

One of the main problem with developing a useful index is the lack of descriptors, which reliably reflect the redox stability of the compounds and hence their metabolic rate. Ionisation potentials and electron affinities are descriptors that can be included into the KDI equations to address the problem. An important future work is the need to develop drug indexes for different routes of administration based on the collated known drugs. From the results, a large portion of the drug collection with low KDIs were not orally available, however they were clinically approved for treatment.

# *Chapter 6*

## Summary and Conclusions

This thesis have shown the applicability of, *in silico* methods to be effective for anticancer drug discovery and development projects. Chapters **2** to **4** involve development of small molecules as inhibitors of cancer-related protein targets: similarity methods to search for structurally similar inhibitors of PI-PLC and using CADD to predict for improved PLC- $\gamma$ 2 selectivity, molecular docking technique to search for favourable AThP inhibition sites on tubulin, and vHTS that identified several novel inhibitors of PC-PLC. Lastly, Chapter **5** proposed a simple method to quantify for the quality of drug candidates.

Chapter **2** showed the ability to expand SARs of the AThP and 1*H*-pyrazole chemical series based on similarity searches on the eMolecules compound database. The key finding is a related series of AThPs, 3-amino-thieno[2,3-*b*][1,6]naphthyridine series, which substituted the cyclohexanone for a piperidine ring in the tricyclic core. Unfortunately, most of the 3-amino-thieno[2,3-*b*][1,6]naphthyridines demonstrated weaker anticancer effects than the AThPs, whereas the 1*H*-pyrazole series remained relatively modest.

Chapter **3** established AThPs as microtubule-destabilising agents seen in test results of the phenotypic sea urchin embryo assays. The assay results revealed greater cytotoxic effects with carboxamide linkers in comparison to ketone linkers. Molecular docking was employed to derive the binding modes of these agents and predict the most favourable AThP inhibition site on tubulin. The results suggested the colchicine site as most favourable for AThP inhibition.

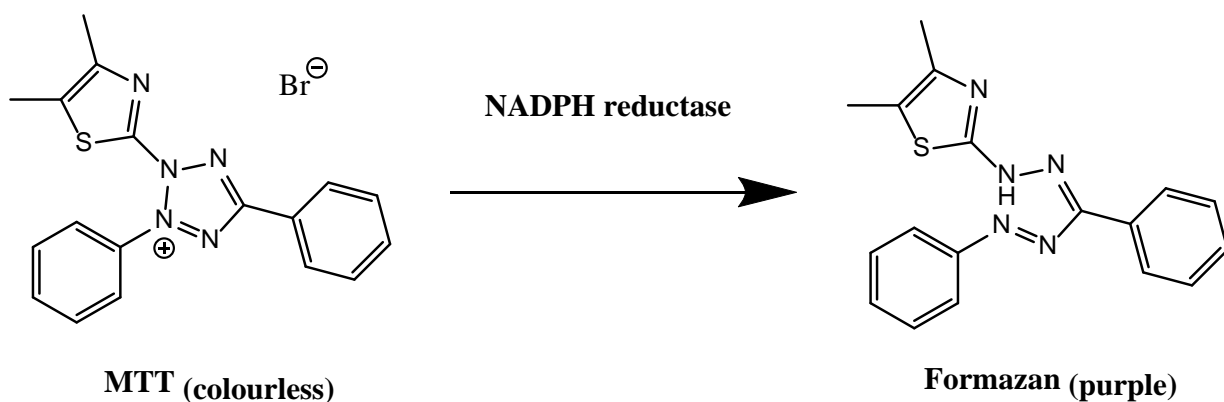
Chapter **4** demonstrated vHTS as a cost effective way for identifying novel drug candidates. Four different chemical series were identified as inhibitors of PC-PLC in which the 2,3,4,9-tetrahydro-1*H*-pyrido[3,4-*b*]indole and 2-morpholinobenzoic acid series were the most active. Cell-based *in vitro* assays confirmed the two series as antiproliferative agents. Interestingly,



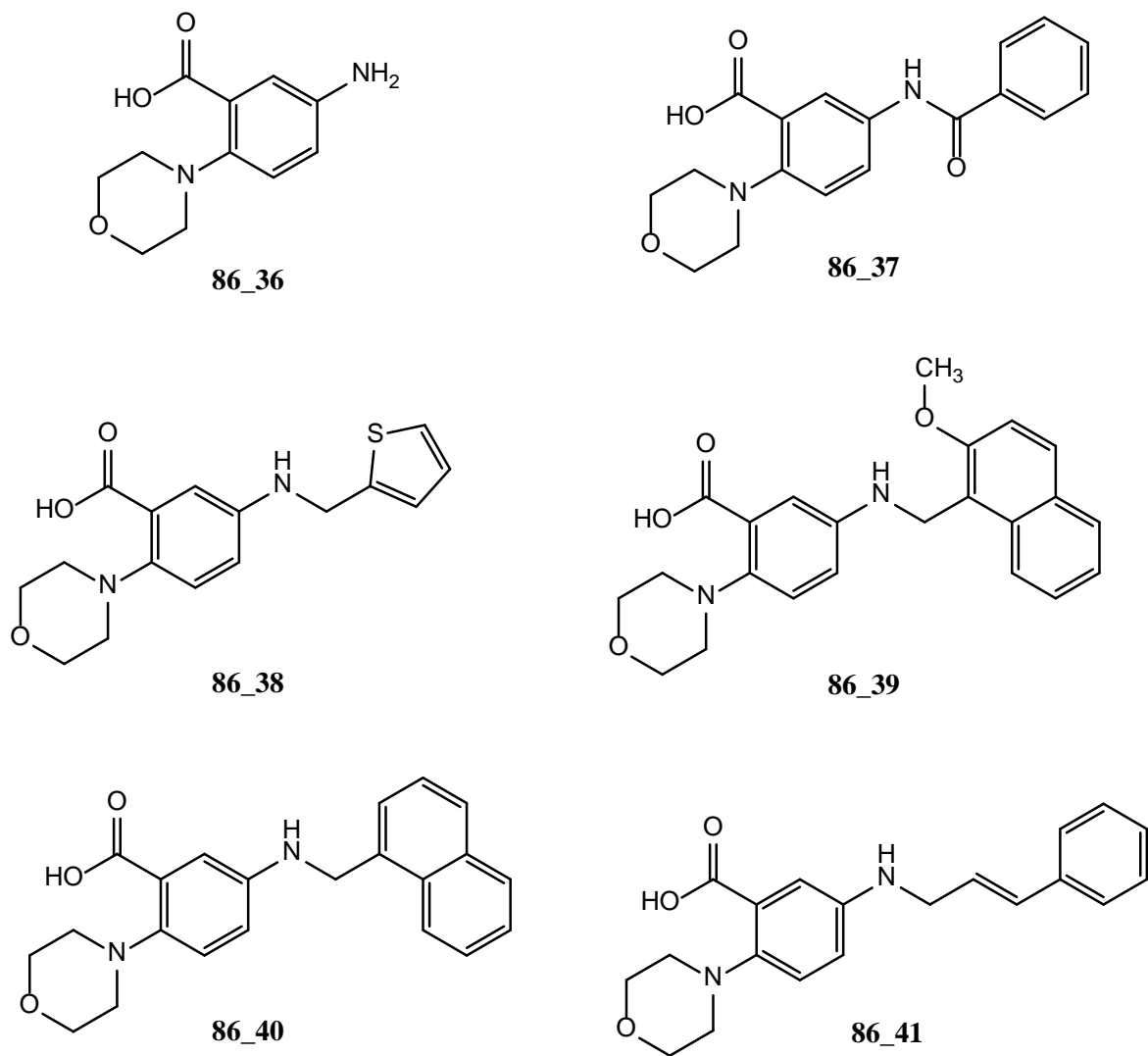
results suggested the mode of inhibition is more dependent on the interactions with the surrounding amino acids rather to the catalytic  $\text{Zn}^{2+}$  ions.

Chapter 5 outlined the use of a drug collection to derive an index for quantifying the quality of drug candidates. The study failed to clearly distinguish drugs from drug-like collections that showed further needs of including descriptors that could further discriminate KDS from drug-like chemical space. Nevertheless, the approach is robust and simplistic to quantify drug qualities based on the parameters derived.

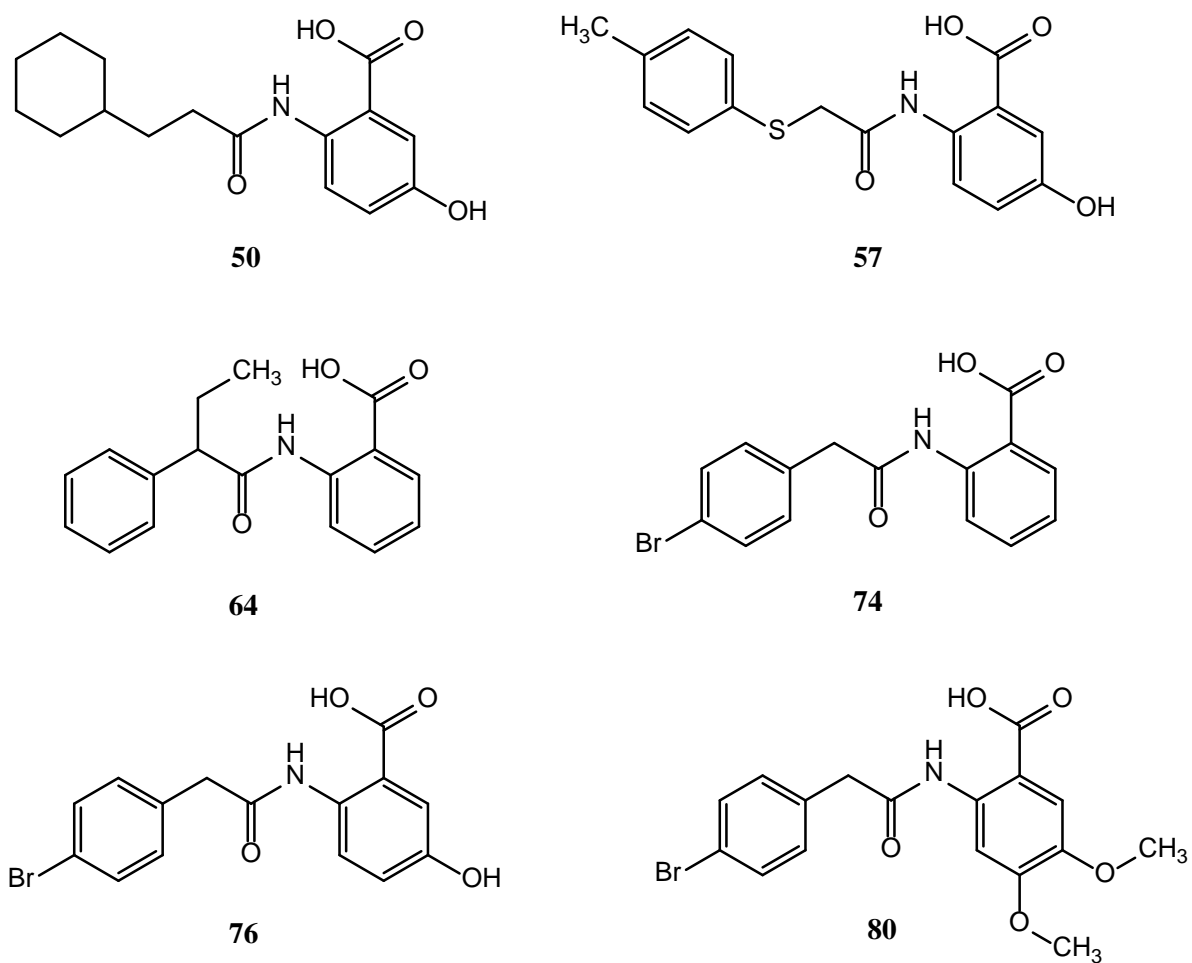
# Appendix



**Scheme A1.** Reduction of colourless MTT to purple formazan in an MTT colorimetric assay.



**Figure A1.** The 2D molecular structures of 2-morpholinobenzoic acid derivatives **86\_37** to **41**.



**Figure A2.** The 2D molecular structures of benzamidobenzoic acid derivatives **50**, **57**, **64**, **74**, **76** and **80**.

**Table A1.** The results of the scoring functions for the *N*-phenylbenzenesulphonamide derivatives.

Compound	GS	CS	ChemPLP	ASP
<b>12</b>	81.1	36.4	80.6	36.9
<b>12_1</b>	84.7	44.0	85.0	33.9
<b>12_2</b>	83.7	43.0	89.2	40.8
<b>12_3</b>	83.7	45.8	93.2	33.2
<b>12_4</b>	82.8	47.0	90.4	33.1
<b>12_5</b>	80.1	43.5	78.6	34.4
<b>12_6</b>	77.4	43.2	77.3	31.0
<b>12_7</b>	77.2	42.6	80.9	34.7
<b>12_8</b>	82.4	45.3	84.9	31.4
<b>12_9</b>	82.8	43.3	77.7	35.1
<b>12_10</b>	82.4	43.7	83.7	30.7
<b>12_11</b>	87.0	44.4	85.3	34.1
<b>12_12</b>	80.2	41.5	79.5	32.4

**Table A2.** The results of the scoring functions for the 2,3,4,9-tetrahydro-1*H*-pyrido[3,4-*b*]indole derivatives.

Compound	GS	CS	ChemPLP	ASP
24	69.7	49.3	91.4	50.9
24_1	69.9	54.6	98.2	42.9
24_2	75.5	45.6	92.6	45.0
24_3	71.0	50.2	93.0	44.8
24_4	72.2	51.8	96.4	46.2
24_5	77.3	49.4	100.6	43.3
24_6	72.4	50.9	88.0	41.2
24_7	77.2	44.0	84.7	42.1
24_8	71.5	41.7	80.2	39.4
24_9	72.3	53.6	92.7	41.3
24_10	73.8	46.4	95.3	45.5
24_11	75.2	46.6	88.8	41.3
24_12	76.0	45.1	86.7	44.4
24_13	73.9	41.6	89.1	45.6
24_14	72.2	42.6	79.4	39.7
24_15	70.1	52.8	96.1	44.1
24_16	69.7	50.4	96.9	43.0
24_17	69.1	49.8	100.2	45.4
24_18	73.1	49.0	97.8	40.6
24_19	72.6	52.5	100.0	43.8
24_20	76.9	52.2	93.8	44.5
24_21	74.4	48.8	89.5	40.3
24_22	71.7	49.4	95.7	41.7
24_23	73.4	50.0	106.0	48.7
24_24	72.4	44.5	94.1	46.8
24_25	78.6	48.8	93.4	43.9
24_26	72.2	48.8	97.9	45.3
24_27	72.6	50.7	98.9	41.4
24_28	72.1	52.7	101.7	43.7
24_29	74.8	53.6	98.2	41.6
24_30	75.9	49.3	86.0	40.0
24_31	46.4	29.6	69.7	35.7
24_32	76.6	50.7	91.5	44.0
24_33	64.8	44.7	81.2	37.1
24_34	69.8	52.0	100.0	43.8
24_35	69.1	52.1	99.5	41.1

**Table A3.** The results of the scoring functions for the 2-morpholinobenzoic acid derivatives.

<b>Compound</b>	<b>GS</b>	<b>CS</b>	<b>ChemPLP</b>	<b>ASP</b>
<b>86</b>	56.5	34.3	67.7	38.2
<b>86_1</b>	67.3	43.6	92.9	41.1
<b>86_2</b>	65.1	42.7	88.1	44.1
<b>86_3</b>	67.7	45.0	95.1	42.5
<b>86_4</b>	69.3	44.2	86.9	41.2
<b>86_5</b>	65.2	47.6	95.8	45.0
<b>86_6</b>	71.2	43.9	84.5	42.2
<b>86_7</b>	69.2	45.6	94.7	42.9
<b>86_8</b>	68.3	45.7	93.0	41.3
<b>86_9</b>	69.4	44.2	93.5	43.0
<b>86_10</b>	72.2	46.6	95.6	41.5
<b>86_11</b>	69.5	42.3	77.6	41.3
<b>86_12</b>	69.3	45.2	90.5	44.7
<b>86_13</b>	69.9	45.9	92.3	44.5
<b>86_14</b>	72.7	42.0	90.7	42.5
<b>86_15</b>	70.8	44.6	96.9	45.0
<b>86_16</b>	72.4	46.5	93.0	45.7
<b>86_17</b>	70.2	40.4	90.3	44.1
<b>86_18</b>	70.8	44.4	89.6	46.2
<b>86_19</b>	70.8	46.1	90.5	44.2
<b>86_20</b>	77.5	45.4	80.8	42.3
<b>86_21</b>	70.5	43.6	95.0	44.3
<b>86_22</b>	69.9	41.6	95.4	43.8
<b>86_23</b>	75.0	38.5	88.5	47.5
<b>86_24</b>	76.6	44.9	100.1	48.5
<b>86_25</b>	69.0	44.2	82.7	42.7
<b>86_26</b>	72.1	40.9	89.9	43.7
<b>86_27</b>	72.5	46.0	90.7	41.7
<b>86_28</b>	72.5	43.7	91.1	43.7
<b>86_29</b>	73.6	43.8	80.8	42.8
<b>86_30</b>	67.3	43.7	90.2	41.1
<b>86_31</b>	68.7	45.4	95.2	44.2
<b>86_32</b>	69.4	45.5	93.3	41.2
<b>86_33</b>	71.1	46.4	95.2	43.4
<b>86_34</b>	51.1	31.1	68.9	39.9
<b>86_35</b>	71.1	34.0	66.3	38.8
<b>86_36</b>	58.1	34.8	67.1	30.5
<b>86_37</b>	55.9	39.4	83.9	38.5
<b>86_38</b>	72.6	40.4	88.6	39.9
<b>86_39</b>	69.0	36.4	85.7	42.2
<b>86_40</b>	75.3	46.6	90.7	43.5
<b>86_41</b>	69.0	44.0	90.1	42.3
<b>86_42</b>	69.7	36.6	70.0	29.3
<b>86_43</b>	55.9	37.5	70.7	28.6

**Table A4.** The results of the scoring functions for the benzamidobenzoic acid derivatives.

<b>Compound</b>	<b>GS</b>	<b>CS</b>	<b>ChemPLP</b>	<b>ASP</b>
<b>82</b>	59.3	40.5	78.3	34.5
<b>82_1</b>	60.6	38.4	77.5	34.2
<b>82_2</b>	59.6	37.6	73.8	30.4
<b>82_3</b>	57.3	35.6	66.7	27.8
<b>82_4</b>	57.6	38.0	78.3	34.1
<b>82_5</b>	61.5	40.8	76.5	35.0
<b>82_6</b>	53.7	31.1	64.6	31.0
<b>82_7</b>	54.4	33.0	64.9	30.3
<b>82_8</b>	54.2	36.0	72.1	32.9
<b>82_9</b>	61.3	36.0	80.1	35.5
<b>82_10</b>	59.9	36.9	78.8	35.5
<b>82_11</b>	59.1	38.6	77.5	32.9
<b>82_12</b>	59.7	38.5	78.8	33.0
<b>82_13</b>	58.0	40.0	79.9	34.4
<b>82_14</b>	56.4	38.8	77.7	33.5
<b>82_15</b>	56.5	36.0	79.0	35.5
<b>82_16</b>	54.7	33.1	63.8	34.1
<b>82_17</b>	58.4	36.1	77.9	34.5
<b>82_18</b>	56.7	35.2	76.8	36.5
<b>82_19</b>	60.3	40.4	70.8	31.6
<b>82_20</b>	63.9	43.7	85.3	37.4
<b>50</b>	70.9	41.8	82.8	34.6
<b>57</b>	70.3	46.1	87.3	37.3
<b>64</b>	60.7	40.2	73.0	35.5
<b>74</b>	65.0	44.5	83.3	33.5
<b>76</b>	71.2	46.7	88.9	35.7
<b>80</b>	64.8	41.7	76.7	39.0

# Bibliography

1. Blass, B. E., Chapter 2 - The Drug Discovery Process: From Ancient Times to the Present Day. In *Basic Principles of Drug Discovery and Development*, Academic Press: Boston, 2015; pp 35-86.
2. Achan, J.; Talisuna, A. O.; Erhart, A.; Yeka, A.; Tibenderana, J. K.; Baliraine, F. N.; Rosenthal, P. J.; D'Alessandro, U., Quinine, an old anti-malarial drug in a modern world: role in the treatment of malaria. *Malar. J.* **2011**, *10* (1), 144.
3. Drews, J., Drug discovery: a historical perspective. *Science* **2000**, *287* (5460), 1960-4.
4. Kaufmann, S. H. E., Paul Ehrlich: founder of chemotherapy. *Nat. Rev. Drug Discov.* **2008**, *7* (5), 373-373.
5. Moffat, J. G.; Rudolph, J.; Bailey, D., Phenotypic screening in cancer drug discovery - past, present and future. *Nat. Rev. Drug Discov.* **2014**, *13* (8), 588-602.
6. Franco, N. H., Animal Experiments in Biomedical Research: A Historical Perspective. *Animals : an Open Access Journal from MDPI* **2013**, *3* (1), 238-273.
7. Festing, S.; Wilkinson, R., The ethics of animal research. Talking Point on the use of animals in scientific research. *EMBO reports* **2007**, *8* (6), 526-30.
8. Shoemaker, R. H., The NCI60 Human Tumour Cell line Anticancer Drug Screen. *Nat. Rev. Drug Dis.* **2006**, *6*, 813-823.
9. Berman, H.; Henrick, K.; Nakamura, H., Announcing the worldwide Protein Data Bank. *Nat. Struct. Biol.* **2003**, *10*, 980.
10. Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E., The Protein Data Bank. *Nuc.Acids Res.* **2000**, *28*, 235-242.
11. Davis, A. M.; Teague, S. J.; Kleywegt, G. J., Application and limitations of X-ray crystallographic data in structure-based ligand and drug design. *Angew. Chem. Int. Ed. Engl.* **2003**, *42* (24), 2718-36.
12. Cavalli, A.; Salvatella, X.; Dobson, C. M.; Vendruscolo, M., Protein structure determination from NMR chemical shifts. *Proc. Natl. Acad. Sci.* **2007**, *104* (23), 9615-9620.
13. Oshiro, C.; Bradley, E. K.; Eksterowicz, J.; Evensen, E.; Lamb, M. L.; Lanctot, J. K.; Putta, S.; Stanton, R.; Grootenhuis, P. D., Performance of 3D-database molecular docking studies into homology models. *J. Med. Chem.* **2004**, *47* (3), 764-7.
14. Guido, R. V.; Oliva, G.; Andricopulo, A. D., Virtual screening and its integration with modern drug design technologies. *Curr. Med. Chem.* **2008**, *15* (1), 37-46.
15. Seidler, J.; McGovern, S. L.; Doman, T. N.; Shoichet, B. K., Identification and prediction of promiscuous aggregating inhibitors among known drugs. *J. Med. Chem.* **2003**, *46* (21), 4477-86.
16. Ryan, A. J.; Gray, N. M.; Lowe, P. N.; Chung, C.-w., Effect of Detergent on "Promiscuous" Inhibitors. *J. Med. Chem.* **2003**, *46* (16), 3448-3451.
17. Bologna, C. G.; Oprea, T. I., Compound Collection Preparation for Virtual Screening. In *Bioinformatics and Drug Discovery*, Larson, R. S., Ed. Humana Press: Totowa, NJ, 2012; pp 125-143.
18. Jones, G.; Willet, P.; Glen, R. C., Molecular Recognition of Receptor Sites using a Genetic Algorithm with a Description of Desolvation. *J.Mol.Biol.* **1995**, *245*, 43-45.
19. Jones, G.; Willet, P.; Glen, R. C.; Leach, A. R.; Taylor, R., Development and Validation of a Genetic Algorithm for Flexible Docking. *J.Mol.Biol.* **1997**, *267*, 727-748.

20. CCDC *GOLD User Guide version 5.4*, 2015.
21. Lang, P. T.; Moustakas, D.; Brozell, S.; Carrascal, N.; Mukherjee, S.; Balias, T.; Allen, W. J.; Holden, P.; Pegg, S.; Raha, K.; Shivakumar, D.; Rizzo, R.; Case, D.; Schoichet, B.; Kuntz, I. *DOCK 6.7 Users Manual*, Regents of the University of California.
22. Eldridge, M. D.; Murray, C.; Auton, T. R.; Paolini, G. V.; Mee, P. M., Empirical scoring functions: I. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. *J. Comp. Aided Mol. Des.* **1997**, *11*, 425-445.
23. Verdonk, M. L.; Cole, J. C.; Hartshorn, M. J.; Murray, C. W.; Taylor, R. D., Improved protein-ligand docking using GOLD. *Proteins.* **2003**, *52* (4), 609-623.
24. Korb, O.; Stützle, T.; Exner, T. E., Empirical Scoring Functions for Advanced Protein-Ligand Docking with PLANTS. *J. Chem. Inf. Model.* **2009**, *49*, 84-96.
25. Mooij, W. T. M.; Verdonk, M. L., General and targeted statistical potentials for protein-ligand interactions. *Proteins.* **2005**, *61* (2), 272-287.
26. Gehlhaar, D. K.; Verkhivker, G. M.; Rejto, P. A.; Sherman, C. J.; Fogel, D. R.; Fogel, L. J.; Freer, S. T., Molecular recognition of the inhibitor AG-1343 by HIV-1 protease: conformationally flexible docking by evolutionary programming. *Chem. Biol.* **1995**, *2* (5), 317-324.
27. Leach, A. R.; Gillet, V. J., Similarity Methods. In *An Introduction To Chemoinformatics*, Springer Netherlands: Dordrecht, 2007; pp 99-117.
28. Willett, P.; Barnard, J. M.; Downs, G. M., Chemical Similarity Searching. *J. Chem. Inform. Comput. Sci.* **1998**, *38* (6), 983-996.
29. Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J., Experimental and computational approaches to estimate solubility and permeability in drug discovery and development setting. *Adv. Drug Deliv. Rev.* **1997**, *23*, 3-25.
30. Lipinski, C. A., Lead- and drug-like compounds: the rule-of-five revolution. *Drug Discov Today Technol* **2004**, *1* (4), 337-341.
31. Veber, D. F.; Johnson, S. R.; Cheng, H.; Smith, B. R.; Ward, K. W.; Kopple, K. D., Molecular Properties that Influence the Oral Bioavailability of Drug Candidates. *J. Med. Chem.* **2002**, *45*, 2615-2623.
32. Matuszek, A. M.; Reynisson, J., Defining Known Drug Space Using DFT. *Mol. Inform.* **2016**, *35* (2), 46-53.
33. Bemis, G. W.; Murcko, M. A., The Properties of Known Drugs. 1. Molecular Frameworks. *J. Med. Chem.* **1996**, *39*, 2887-2893.
34. Bemis, G. W.; Murcko, M. A., Properties of Known Drugs. 2. Side Chains. *J. Med. Chem.* **1999**, *42*, 5095-5099.
35. Baell, J. B.; Holloway, G. A., New Substructure Filters for Removal of Pan Assay Interference Compounds (PAINS) from Screening Libraries and for Their Exclusion in Bioassays. *J. Med. Chem.* **2010**, *53*, 2719-2740.
36. Rishton, G. M., Reactive compounds and in vitro false positives in HTS. *DDT* **1997**, *2* (9), 382-384.
37. Biswas, D.; Roy, S.; Sen, S., A simple approach for indexing the oral druglikeness of a compound: discriminating druglike compounds from nondruglike ones. *J. Chem. Inf. Model.* **2006**, *46* (3), 1394-401.
38. Sadowski, J.; Kubinyi, H., A Scoring Scheme for Discriminating between Drugs and Nondrugs. *J. Med. Chem.* **1998**, *41* (18), 3325-3329.
39. Xu, J.; Stevenson, J., Drug-like Index: A New Approach To Measure Drug-like Compounds and Their Diversity. *J. Chem. Inform. Comput. Sci.* **2000**, *40* (5), 1177-1187.
40. Ezkurdia, I.; Juan, D.; Rodriguez, J. M.; Frankish, A.; Diekhans, M.; Harrow, J.; Vazquez, J.; Valencia, A.; Tress, M. L., Multiple evidence strands suggest that there may



- be as few as 19,000 human protein-coding genes. *Hum. Mol. Genet.* **2014**, 23 (22), 5866-78.
41. Fukami, K.; Inanobe, S.; Kanemaru, K.; Nakamura, Y., Phospholipase C is a key enzyme regulating intracellular calcium and modulating the phosphoinositide balance. *Prog. Lipid Res.* **2010**, 49 (4), 429-437.
  42. Vines, C. M., Phospholipase C. In *Calcium Signaling*, Islam, M. S., Ed. Springer Netherlands: Dordrecht, 2012; pp 235-254.
  43. Park, J. B.; Lee, C. S.; Jang, J. H.; Ghim, J.; Kim, Y. J.; You, S.; Hwang, D.; Suh, P. G.; Ryu, S. H., Phospholipase signalling networks in cancer. *Nat. Rev. Cancer.* **2012**, 12 (11), 782-92.
  44. Fu, L.; Qin, Y.-R.; Xie, D.; Hu, L.; Kwong, D. L.; Srivastava, G.; Tsao, S. W.; Guan, X.-Y., Characterization of a Novel Tumor-Suppressor Gene PLC $\delta$ 1 at 3p22 in Esophageal Squamous Cell Carcinoma. *Cancer Res.* **2007**, 67 (22), 10720-10726.
  45. Wells, A.; Grandis, J. R., Phospholipase C- $\gamma$ 1 in tumor progression. *Clin. Exp. Metastasis.* **2003**, 20 (4), 285.
  46. Sala, G.; Dituri, F.; Raimondi, C.; Previdi, S.; Maffucci, T.; Mazzeo, M.; Rossi, C.; Iezzi, M.; Lattanzio, R.; Piantelli, M.; Iacobelli, S.; Broggin, M.; Falasca, M., Phospholipase C $\gamma$ 1 Is Required for Metastasis Development and Progression. *Cancer Res.* **2008**, 68 (24), 10187-10196.
  47. Xie, Z.; Chen, Y.; Liao, E.-Y.; Jiang, Y.; Liu, F.-Y.; Pennypacker, S. D., Phospholipase C- $\gamma$ 1 is Required for the Epidermal Growth Factor Receptor-induced Squamous Cell Carcinoma Cell Mitogenesis. *Biochem. Biophys. Res. Commun.* **2010**, 397 (2), 296-300.
  48. Arteaga, C. L.; Johnson, M. D.; Todderud, G.; Coffey, R. J.; Carpenter, G.; Page, D. L., Elevated content of the tyrosine kinase substrate phospholipase C- $\gamma$ 1 in primary human breast carcinomas. *Proc. Natl. Acad. Sci. U. S. A.* **1991**, 88 (23), 10435-9.
  49. Rhee, S. G.; Kim, H.; Suh, P.-G.; Choi, W. C., Multiple forms of phosphoinositide-specific phospholipase C and different modes of activation. *Biochem. Soc. Trans.* **1991**, 19 (2), 337-341.
  50. Homma, Y.; Takenawa, T.; Emori, Y.; Sorimachi, H.; Suzuki, K., Tissue- and cell type-specific expression of mRNAs for four types of inositol phospholipid-specific phospholipase C. *Biochem. Biophys. Res. Commun.* **1989**, 164 (1), 406-12.
  51. Essen, L.-O.; Perisic, O.; Katan, M.; Wu, Y.; Roberts, M. F.; Williams, R. L., Structural Mapping of the Catalytic Mechanism for a Mammalian Phosphoinositide-Specific Phospholipase C. *Biochemistry.* **1997**, 36 (7), 1704-1718.
  52. Rebecchi, M. J.; Pentylä, S. N., Structure, function, and control of phosphoinositide-specific phospholipase C. *Physiol. Rev.* **2000**, 80 (4), 1291-335.
  53. Cheng, H.-F.; Jiang, M.-J.; Chen, C.-L.; Liu, S.-M.; Wong, L.-P.; Lomasney, J. W.; King, K., Cloning and Identification of Amino Acid Residues of Human Phospholipase C $\delta$ 1 Essential for Catalysis. *J. Biol. Chem.* **1995**, 270 (10), 5495-5505.
  54. Xie, W.; Peng, H.; Kim, D. I.; Kunkel, M.; Powis, G.; Zalkow, L. H., Structure - Activity relationship of Aza-steroids as PI-PLC inhibitors. *Bioorg. Med. Chem.* **2001**, 9 (5), 1073-1083.
  55. Xie, W.; Peng, H.; Zalkow, L. H.; Li, Y. H.; Zhu, C.; Powis, G.; Kunkel, M., 3 $\beta$ -hydroxy-6-aza-cholestane and related analogues as phosphatidylinositol specific phospholipase C (PI-PLC) inhibitors with antitumor activity. *Bioorg. Med. Chem.* **2000**, 8 (4), 699-706.
  56. Lee, J. S.; Kim, J., Phospholipase C $\gamma$  as a target for the development of new anticancer agents from natural sources. *Drugs Future.* **2001**, 26 (2), 163-173.

57. Mollinedo, F.; Gajate, C.; Martin-Santamaria, S.; Gago, F., ET-18-OCH<sub>3</sub> (edelfosine): a selective antitumour lipid targeting apoptosis through intracellular activation of Fas/CD95 death receptor. *Curr. Med. Chem.* **2004**, *11* (24), 3163-84.
58. Powis, G.; Seewald, M. J.; Gratas, C.; Melder, D.; Riebow, J.; Modest, E. J., Selective inhibition of phosphatidylinositol phospholipase C by cytotoxic ether lipid analogues. *Cancer Res.* **1992**, *52* (10), 2835-40.
59. Drings, P.; Günther, I.; Gatzemeier, U.; Ulbrich, F.; Khanavkar, B.; Schreml, W.; Lorenz, J.; Brugger, W.; Schick, H. D.; Pawel, J. v.; Nordström, R., Final Evaluation of a Phase II Study on the Effect of Edelfosine (an Ether Lipid) in Advanced Non-Small-Cell Bronchogenic Carcinoma. *Oncol. Res. Treat.* **1992**, *15* (5), 375-382.
60. Vogler, W. R.; Berdel, W. E.; Geller, R. B.; Brochstein, J. A.; Beveridge, R. A.; Dalton, W. S.; Miller, K. B.; Lazarus, H. M., A phase II trial of autologous bone marrow transplantation (ABMT) in acute leukemia with edelfosine purged bone marrow. *Adv. Exp. Med. Biol.* **1996**, *416*, 389-96.
61. Nagler, A., Edelfosin for the treatment of brain tumors. Google Patents: 2003.
62. Reynisson, J.; Court, W.; O'Neill, C.; Day, J.; Patterson, L.; McDonald, E.; Workman, P.; Katan, M.; Eccles, S. A., The identification of novel PLC-gamma inhibitors using virtual high throughput screening. *Bioorg. Med. Chem.* **2009**, *17* (8), 3169-76.
63. Irwin, J. J.; Shoichet, B. K., ZINC - A Free Database of Commercially Available COMpounds for Virtual Screening. *J. Chem. Inf. Model.* **2005**, *45*, 177-182.
64. Katan, M.; Rodriguez, R.; Matsuda, M.; Newbatt, Y. M.; Aherne, G. W., Structural and mechanistic aspects of phospholipase C $\gamma$  regulation. *Adv. Enzyme Regul.* **2003**, *43* (1), 77-85.
65. Feng, L.; Reynisdottir, I.; Reynisson, J., The effect of PLC-gamma2 inhibitors on the growth of human tumour cells. *Eur. J. Med. Chem.* **2012**, *54*, 463-9.
66. Hung, J. M.; Arabshahi, H. J.; Leung, E.; Reynisson, J.; Barker, D., Synthesis and cytotoxicity of thieno[2,3-*b*]pyridine and furo[2,3-*b*]pyridine derivatives. *Eur. J. Med. Chem.* **2014**, *86*, 420-37.
67. Arabshahi, H. J.; Leung, E.; Barker, D.; Reynisson, J., The development of thieno[2,3-*b*]pyridine analogues as anticancer agents applying in silico methods. *MedChemComm.* **2014**, *5* (2), 186-191.
68. eMolecules. [www.emolecules.com](http://www.emolecules.com).
69. Katritch, V.; Jaakola, V.-P.; Lane, J. R.; Lin, J.; Ijzerman, A. P.; Yeager, M.; Kufareva, I.; Stevens, R. C.; Abagyan, R., Structure-Based Discovery of Novel Chemotypes for Adenosine A<sub>2A</sub> Receptor Antagonists. *J. Med. Chem.* **2010**, *53* (4), 1799-1809.
70. Wang, J.; Luo, C.; Shan, C.; You, Q.; Lu, J.; Elf, S.; Zhou, Y.; Wen, Y.; Vinkenburg, J. L.; Fan, J.; Kang, H.; Lin, R.; Han, D.; Xie, Y.; Karpus, J.; Chen, S.; Ouyang, S.; Luan, C.; Zhang, N.; Ding, H.; Merkx, M.; Liu, H.; Chen, J.; Jiang, H.; He, C., Inhibition of human copper trafficking by a small molecule significantly attenuates cancer cell proliferation. *Nat Chem.* **2015**, *7* (12), 968-979.
71. Zhou, R.; Huang, W. J.; Guo, Z. Y.; Li, L.; Zeng, X. R.; Deng, Y. Q.; Hu, F. Y.; Tong, A. P.; Yang, L.; Yang, J. L., Molecular mechanism of hepatocellular carcinoma-specific antitumor activity of the novel thienopyridine derivative TP58. *Oncol. Rep.* **2012**, *28* (1), 225-31.
72. Zeng, X. X.; Zheng, R. L.; Zhou, T.; He, H. Y.; Liu, J. Y.; Zheng, Y.; Tong, A. P.; Xiang, M. L.; Song, X. R.; Yang, S. Y.; Yu, L. T.; Wei, Y. Q.; Zhao, Y. L.; Yang, L., Novel thienopyridine derivatives as specific anti-hepatocellular carcinoma (HCC) agents: synthesis, preliminary structure-activity relationships, and in vitro biological evaluation. *Bioorg. Med. Chem. Lett.* **2010**, *20* (21), 6282-5.

73. Romagnoli, R.; Baraldi, P. G.; Kimatrai Salvador, M.; Preti, D.; Aghazadeh Tabrizi, M.; Bassetto, M.; Brancale, A.; Hamel, E.; Castagliuolo, I.; Bortolozzi, R.; Basso, G.; Viola, G., Synthesis and biological evaluation of 2-(alkoxycarbonyl)-3-anilinobenzo[*b*]thiophenes and thieno[2,3-*b*]pyridines as new potent anticancer agents. *J. Med. Chem.* **2013**, *56* (6), 2606-18.
74. Eurtivong, C.; Semenov, V.; Semenova, M.; Konyushkin, L.; Atamanenko, O.; Reynisson, J.; Kiselyov, A., 3-Amino-thieno[2,3-*b*]pyridines as microtubule-destabilising agents: Molecular modelling and biological evaluation in the sea urchin embryo and human cancer cells. *Bioorg. Med. Chem.* **2017**, *25* (2), 658-664.
75. Leung, E.; Hung, J. M.; Barker, D.; Reynisson, J., The effect of a thieno[2,3-*b*]pyridine PLC- $\gamma$  inhibitor on the proliferation, morphology, migration and cell cycle of breast cancer cells. *MedChemComm.* **2014**, *5* (1), 99-106.
76. Ellis, M. V.; James, S. R.; Perisic, O.; Downes, C. P.; Williams, R. L.; Katan, M., Catalytic Domain of Phosphoinositide-specific Phospholipase C (PLC): Mutational analysis of residues within the active site and hydrophobic ridge of PLC $\delta$ 1. *J. Biol. Chem.* **1998**, *273* (19), 11650-11659.
77. Berridge, M. V.; Herst, P. M.; Tan, A. S., Tetrazolium dyes as tools in cell biology: New insights into their cellular reduction. In *Biotechnol. Ann. Rev.*, Elsevier: 2005; Vol. Volume 11, pp 127-152.
78. Leach, A. R.; Gillet, V. J., Representation And Manipulation Of 2D Molecular Structures. In *An Introduction To Chemoinformatics*, Springer Netherlands: Dordrecht, 2007; pp 1-25.
79. Fujitsu Scigress, v2.6; Fijitsu Limited: 2016.
80. Allinger, N. L., Conformational analysis. 130. MM2. A hydrocarbon force field utilizing V1 and V2 torsional terms. *J. Am. Chem. Soc.* **1977**, *99*, 8127-8134.
81. Eurtivong, C.; Reynisdóttir, I.; Kuczma, S.; Furkert, D. P.; Brimble, M. A.; Reynisson, J., Identification of anticancer agents based on the thieno[2,3-*b*]pyridine and 1*H*-pyrazole molecular scaffolds. *Bioorg. Med. Chem.* **2016**, *24* (16), 3521-6.
82. Dráber, P.; Dráberová, E., Microtubules. In *Cytoskeleton and Human Disease*, Kavallaris, M., Ed. Humana Press: Totowa, NJ, 2012; pp 29-53.
83. Jordan, M. A.; Wilson, L., Microtubules as a target for anticancer drugs. *Nat. Rev. Cancer.* **2004**, *4* (4), 253-265.
84. Valiron, O.; Caudron, N.; Job, D., Microtubule dynamics. *Cell. Mol. Life Sci.* **2001**, *58* (14), 2069-84.
85. Nogales, E., Structural insights into microtubule function. *Annu. Rev. Biochem.* **2000**, *69*, 277-302.
86. Avendaño, C.; Menéndez, J. C., Chapter 9 - Anticancer Drugs Targeting Tubulin and Microtubules. In *Medicinal Chemistry of Anticancer Drugs (Second Edition)*, Elsevier: Boston, 2015; pp 359-390.
87. Horwitz, S. B.; Fojo, T., Microtubule Stabilizing Agents. In *The Role of Microtubules in Cell Biology, Neurobiology, and Oncology*, Fojo, T., Ed. Humana Press: Totowa, NJ, 2008; pp 307-336.
88. Rohena, C. C.; Mooberry, S. L., Recent progress with microtubule stabilizers: new compounds, binding modes and cellular activities. *Nat. Prod. Rep.* **2014**, *31* (3), 335-55.
89. Arnal, I.; Wade, R. H., How does taxol stabilize microtubules? *Curr. Biol.* **1995**, *5* (8), 900-908.
90. Löwe, J.; Li, H.; Downing, K. H.; Nogales, E., Refined structure of  $\alpha\beta$ -tubulin at 3.5 Å resolution1. *J. Mol. Biol.* **2001**, *313* (5), 1045-1057.
91. Botta, M.; Forli, S.; Magnani, M.; Manetti, F., Molecular Modeling Approaches to Study the Binding Mode on Tubulin of Microtubule Destabilizing and Stabilizing Agents. In

- Tubulin-Binding Agents: Synthetic, Structural and Mechanistic Insights*, Carlomagno, T., Ed. Springer Berlin Heidelberg: Berlin, Heidelberg, 2009; pp 279-328.
92. Gigant, B.; Cormier, A.; Dorléans, A.; Ravelli, R. B. G.; Knossow, M., Microtubule-Destabilizing Agents: Structural and Mechanistic Insights from the Interaction of Colchicine and Vinblastine with Tubulin. In *Tubulin-Binding Agents: Synthetic, Structural and Mechanistic Insights*, Carlomagno, T., Ed. Springer Berlin Heidelberg: Berlin, Heidelberg, 2009; pp 259-278.
  93. Skoufias, D. A.; Wilson, L., Mechanism of inhibition of microtubule polymerization by colchicine: inhibitory potencies of unliganded colchicine and tubulin-colchicine complexes. *Biochemistry*. **1992**, *31* (3), 738-46.
  94. Jordan, M. A.; Wilson, L., Kinetic analysis of tubulin exchange at microtubule ends at low vinblastine concentrations. *Biochemistry*. **1990**, *29* (11), 2730-9.
  95. Wilson, L.; Jordan, M. A.; Morse, A.; Margolis, R. L., Interaction of vinblastine with steady-state microtubules in vitro. *J. Mol. Biol.* **1982**, *159* (1), 125-149.
  96. Nishioka, D.; Marcell, V.; Cunningham, M.; Khan, M.; Von Hoff, D. D.; Izbicka, E., The use of early sea urchin embryos in anticancer drug testing. *Methods Mol. Med.* **2003**, *85*, 265-76.
  97. Semenova, M. N.; Kiselyov, A.; Semenov, V. V., Sea urchin embryo as a model organism for the rapid functional screening of tubulin modulators. *BioTechniques*. **2006**, *40* (6), 765-74.
  98. Pettit, G. R.; Toki, B.; Herald, D. L.; Verdier-Pinard, P.; Boyd, M. R.; Hamel, E.; Pettit, R. K., Antineoplastic Agents. 379. Synthesis of Phenstatin Phosphate1a. *J. Med. Chem.* **1998**, *41* (10), 1688-1695.
  99. Arabshahi, H. J.; van Rensburg, M.; Pilkington, L. I.; Jeon, C. Y.; Song, M.; Gridel, L.-M.; Leung, E.; Barker, D.; Vuica-Ross, M.; Volcho, K. P.; Zakharenko, A. L.; Lavrik, O. I.; Reynisson, J., A synthesis, in silico, in vitro and in vivo study of thieno[2,3-b]pyridine anticancer analogues. *MedChemComm*. **2015**, *6* (11), 1987-1997.
  100. Prota, A. E.; Danel, F.; Bachmann, F.; Bargsten, K.; Buey, R. M.; Pohlmann, J.; Reinelt, S.; Lane, H.; Steinmetz, M. O., The Novel Microtubule-Destabilizing Drug BAL27862 Binds to the Colchicine Site of Tubulin with Distinct Effects on Microtubule Organization. *J. Mol. Biol.* **2014**, *426* (8), 1848-1860.
  101. Ranaivoson, F. M.; Gigant, B.; Berritt, S.; Joullie, M.; Knossow, M., Structural plasticity of tubulin assembly probed by vinca-domain ligands. *Acta Crys. D*. **2012**, *68* (8), 927-934.
  102. Glunde, K.; Bhujwalla, Z. M.; Ronen, S. M., Choline metabolism in malignant transformation. *Nat. Rev. Cancer*. **2011**, *11* (12), 835-48.
  103. Hergenrother, P. J.; Martin, S. F., Phosphatidylcholine-Preferring Phospholipase C from *B. cereus*. Function, Structure, and Mechanism. In *Bioorganic Chemistry of Biological Signal Transduction*, Waldmann, H., Ed. Springer Berlin Heidelberg: Berlin, Heidelberg, 2001; pp 131-167.
  104. Park, J. B.; Lee, C. S.; Jang, J. H.; Ghim, J.; Kim, Y. J.; You, S.; Hwang, D.; Suh, P. G.; Ryu, S. H., Phospholipase signalling networks in cancer. *Nat. Rev. Cancer*. **2012**, *12*, 782-792.
  105. Ridgway, N. D., The role of phosphatidylcholine and choline metabolites to cell proliferation and survival. *Crit. Rev. Biochem. Mol. Biol.* **2013**, *48* (1), 20-38.
  106. Daly, P. F.; Lyon, R. C.; Faustino, P. J.; Cohen, J. S., Phospholipid metabolism in cancer cells monitored by <sup>31</sup>P NMR spectroscopy. *J. Biol. Chem.* **1987**, *262* (31), 14875-8.
  107. Aboagye, E. O.; Bhujwalla, Z. M., Malignant transformation alters membrane choline phospholipid metabolism of human mammary epithelial cells. *Cancer Res.* **1999**, *59* (1), 80-4.

108. Hough, E.; Hansen, L. K.; Birknes, B.; Jynge, K.; Hansen, S.; Hordvik, A.; Little, C.; Dodson, E.; Derewenda, Z., High-resolution (1.5 Å) crystal structure of phospholipase C from *Bacillus cereus*. *Nature*. **1989**, 338 (6213), 357-360.
109. Mcgaughey, C. A.; Chu, H. P., The Egg-Yolk Reaction of Aerobic Sporing Bacilli. *Microbiology*. **1948**, 2 (3), 334-340.
110. Clark, M. A.; Shorr, R. G.; Bomalaski, J. S., Antibodies prepared to *Bacillus cereus* phospholipase C crossreact with a phosphatidylcholine preferring phospholipase C in mammalian cells. *Biochem. Biophys. Res. Commun.* **1986**, 140 (1), 114-9.
111. Levine, L.; Xiao, D. M.; Little, C., Increased arachidonic acid metabolites from cells in culture after treatment with the phosphatidylcholine-hydrolyzing phospholipase C from *Bacillus cereus*. *Prostaglandins*. **1987**, 34 (5), 633-42.
112. Abalsamo, L.; Spadaro, F.; Bozzuto, G.; Paris, L.; Cecchetti, S.; Lugini, L.; Iorio, E.; Molinari, A.; Ramoni, C.; Podo, F., Inhibition of phosphatidylcholine-specific phospholipase C results in loss of mesenchymal traits in metastatic breast cancer cells. *Breast Cancer Res.* **2012**, 14 (2), R50-R50.
113. Spadaro, F.; Ramoni, C.; Mezzanzanica, D.; Miotti, S.; Alberti, P.; Cecchetti, S.; Iorio, E.; Dolo, V.; Canevari, S.; Podo, F., Phosphatidylcholine-specific phospholipase C activation in epithelial ovarian cancer cells. *Cancer Res.* **2008**, 68 (16), 6541-9.
114. Plo, I.; Lautier, D.; Levade, T.; Sekouri, H.; Jaffrezou, J. P.; Laurent, G.; Bettaieb, A., Phosphatidylcholine-specific phospholipase C and phospholipase D are respectively implicated in mitogen-activated protein kinase and nuclear factor kappaB activation in tumour-necrosis-factor-alpha-treated immature acute-myeloid-leukaemia cells. *Biochem. J.* **2000**, 351 (Pt 2), 459-67.
115. González-Bulnes, P.; González-Roura, A.; Canals, D.; Delgado, A.; Casas, J.; Llebaria, A., 2-Aminohydroxamic acid derivatives as inhibitors of *Bacillus cereus* phosphatidylcholine preferred phospholipase C PC-PLCBc. *Bioorg. Med. Chem.* **2010**, 18 (24), 8549-8555.
116. Aakre, S. E.; Little, C., Inhibition of *Bacillus cereus* phospholipase C by univalent anions. *Biochem. J.* **1982**, 203 (3), 799-801.
117. Martin, S. F.; Follows, B. C.; Hergenrother, P. J.; Franklin, C. L., A novel class of zinc-binding inhibitors for the phosphatidylcholine-preferring phospholipase C from *Bacillus cereus*. *J. Org. Chem.* **2000**, 65 (15), 4509-14.
118. Martin, S. F.; Wong, Y.-L.; Wagman, A. S., Design, Synthesis, and Evaluation of Phospholipid Analogs as Inhibitors of the Bacterial Phospholipase C from *Bacillus cereus*. *J. Org. Chem.* **1994**, 59 (17), 4821-4831.
119. Adibhatla, R.; Hatcher, J. F.; Gusain, A., Tricyclodecan-9-yl-Xanthogenate (D609) Mechanism of Actions: A Mini-Review of Literature. *Neurochem. Res.* **2012**, 37 (4), 671-679.
120. Amtmann, E., The antiviral, antitumoural xanthate D609 is a competitive inhibitor of phosphatidylcholine-specific phospholipase C. *Drugs Exp. Clin. Res.* **1996**, 22 (6), 287-94.
121. Martin, S. F.; Hergenrother, P. J., Catalytic Cycle of the Phosphatidylcholine-Preferring Phospholipase C from *Bacillus cereus*. Solvent Viscosity, Deuterium Isotope Effects, and Proton Inventory Studies. *Biochemistry*. **1999**, 38 (14), 4403-4408.
122. Hansen, S.; Hough, E.; Svensson, L. A.; Wong, Y. L.; Martin, S. F., Crystal structure of phospholipase C from *Bacillus cereus* complexed with a substrate analog. *J. Mol. Biol.* **1993**, 234 (1), 179-87.
123. Martin, S. F.; Hergenrother, P. J., General base catalysis by the phosphatidylcholine-preferring phospholipase C from *Bacillus cereus*: the role of Glu4 and Asp55. *Biochemistry*. **1998**, 37 (16), 5755-60.

124. da Graca Thrige, D.; Buur, J. R.; Jorgensen, F. S., Substrate binding and catalytic mechanism in phospholipase C from *Bacillus cereus*: a molecular mechanics and molecular dynamics study. *Biopolymers*. **1997**, *42* (3), 319-36.
125. Rebek, J., On the structure of histidine and its role in enzyme active sites. *Struct. Chem.* **1990**, *1* (1), 129-131.
126. ChemBridge Cooperation. <http://www.chembridge.com/>.
127. Hit2Lead ChemBridge Chemical Store, ChemBridge Cooperation. [www.hit2lead.com](http://www.hit2lead.com).
128. Zhou, M.; Diwu, Z.; Panchuk-Voloshina, N.; Haugland, R. P., A Stable Nonfluorescent Derivative of Resorufin for the Fluorometric Determination of Trace Hydrogen Peroxide: Applications in Detecting the Activity of Phagocyte NADPH Oxidase and Other Oxidases. *Anal. Biochem.* **1997**, *253* (2), 162-168.
129. Chemical Diversity. [www.chemdiv.com](http://www.chemdiv.com).
130. InterBioScreen. <http://www.interbioscreen.com/>.
131. Coward, R. M.; Carson, C. C., Tadalafil in the treatment of erectile dysfunction. *Ther. Clin. Risk Manag.* **2008**, *4* (6), 1315-1330.
132. Wang, R.; Wang, S., How Does Consensus Scoring Work for Virtual Library Screening? An Idealized Computer Experiment. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 1422-1426.
133. Yang, J.-M.; Chen, Y.-F.; Shen, T.-W.; Kristal, B. S.; Hsu, D. F., Consensus Scoring Criteria for Improving Enrichment in Virtual Screening. *J. Chem. Inf. Model.* **2005**, *45* (4), 1134-1146.
134. Zhao, Y., Auxin Biosynthesis: A Simple Two-Step Pathway Converts Tryptophan to Indole-3-Acetic Acid in Plants. *Mol. Plant.* **2012**, *5* (2), 334-338.
135. Kim, S. Y.; Ryu, J. S.; Li, H.; Park, W. J.; Yun, H. Y.; Baek, K. J.; Kwon, N. S.; Sohn, U. D.; Kim, D. S., UVB-activated indole-3-acetic acid induces apoptosis of PC-3 prostate cancer cells. *Anticancer Res.* **2010**, *30* (11), 4607-12.
136. Wardman, P., Indole-3-acetic acids and horseradish peroxidase: a new prodrug/enzyme combination for targeted cancer therapy. *Curr. Pharm. Des.* **2002**, *8* (15), 1363-74.
137. Kim, D. S.; Jeon, S. E.; Jeong, Y. M.; Kim, S. Y.; Kwon, S. B.; Park, K. C., Hydrogen peroxide is a mediator of indole-3-acetic acid/horseradish peroxidase-induced apoptosis. *FEBS Lett.* **2006**, *580* (5), 1439-46.
138. Kim, D.-S.; Jeon, S.-E.; Park, K.-C., Oxidation of indole-3-acetic acid by horseradish peroxidase induces apoptosis in G361 human melanoma cells. *Cell. Signal.* **2004**, *16* (1), 81-88.
139. Benfield, A. P.; Goodey, N. M.; Phillips, L. T.; Martin, S. F., Structural Studies Examining the Substrate Specificity Profiles of PC-PLC(Bc) Protein Variants. *Arch. Biochem. Biophys.* **2007**, *460* (1), 41-47.
140. Goncharova, E. A.; Lim, P.; Goncharov, D. A.; Eszterhas, A.; Panettieri, R. A., Jr.; Krymskaya, V. P., Assays for in vitro monitoring of proliferation of human airway smooth muscle (ASM) and human pulmonary arterial vascular smooth muscle (VSM) cells. *Nat. Protoc.* **2006**, *1* (6), 2905-8.
141. Nissink, J. W. M.; Murray, C.; Hartshorn, M.; Verdonk, M. L.; Cole, J. C.; Taylor, R., A New Test Set for Validating Predictions of Protein-Ligand Interaction. *Proteins.* **2002**, *49*, 457-471.
142. Kontoyianni, M.; McClellan, L. M.; Sokol, G. S., Evaluation of Docking Performance: Comparative Data on Docking Algorithms. *J. Med. Chem.* **2004**, *47*, 558-565.
143. Mirza, A.; Desai, R.; Reynisson, J., Known drug space as a metric in exploring the boundaries of drug-like chemical space. *Eur. J. Med. Chem.* **2009**, *44* (12), 5006-5011.
144. Axerio-Cilies, P.; Castañeda, I. P.; Mirza, A.; Reynisson, J., Investigation of the incidence of “undesirable” molecular moieties for high-throughput screening compound libraries in marketed drug compounds. *Eur. J. Med. Chem.* **2009**, *44*, 1128-1134.

145. Bade, R.; Chan, H.-F.; Reynisson, J., Characteristics of known drug space. Natural products, their derivatives and synthetic drugs. *Eur. J. Med. Chem.* **2010**, *45* (12), 5646-5652.
146. Drew, K. L. M.; Baiman, H.; Khwaounjoo, P.; Yu, B.; Reynisson, J., Size estimation of chemical space: how big is it? *J. Pharm. Pharmacol.* **2012**, *64*, 490–495.
147. Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J., Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.* **2001**, *46* (1-3), 3-26.
148. Leach, A. R.; Gillet, V. J., Molecular Descriptors. In *An Introduction To Chemoinformatics*, Springer Netherlands: Dordrecht, 2007; pp 53-74.
149. Xue, L.; Bajorath, J., Molecular descriptors in chemoinformatics, computational combinatorial chemistry, and virtual screening. *Comb. Chem. High Throughput Screen.* **2000**, *3* (5), 363-72.
150. Heumann, C.; Schomaker, M.; Shalabh, Probability Distributions. In *Introduction to Statistics and Data Analysis : With Exercises, Solutions and Applications in R*, Springer International Publishing: Cham, 2016; pp 153-178.
151. Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J., Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *J. Am. Chem. Soc.* **1996**, *118* (45), 11225-11236.
152. Jorgensen, W. L.; Tirado-Rives, J., The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. *J. Am. Chem. Soc.* **1988**, *110* (6), 1657-1666.
153. Ioakimidis, L.; Thoukydidis, L.; Mirza, A.; Naeem, S.; Reynisson, J., Benchmarking the Reliability of QikProp. Correlation between Experimental and Predicted Values. *QSAR Comb. Sci.* **2008**, *27* (4), 445-456.
154. U.S. Food and Drug Administration. <http://www.fda.gov/default.htm> (accessed 15th Jun 2015).
155. Health Canada. <http://www.hc-sc.gc.ca/> (accessed 1st Jul 2015).
156. Medicines and Healthcare Products Regulatory Agency. <http://www.mhra.gov.uk> (accessed 21st Jun 2015).
157. China food and drug administration. [eng.sfda.gov.cn](http://eng.sfda.gov.cn) (accessed 25th Jun 2015).
158. *Australian Medicines Handbook*. 2014.
159. Japan's Pharmaceuticals and Medical Devices Agency. [www.pmda.go.jp](http://www.pmda.go.jp) (accessed 1st Jul 2015).
160. Oprea, T. I., Property distribution of drug-related chemical databases. *J. Comput. Aided Mol. Des.* **2000**, *14* (3), 251-64.
161. Hansch, C. H., *Comprehensive medicinal chemistry : the rational design, mechanistic study & therapeutic application of chemical compounds; [in 6 volumes]*. Pergamon Press: Oxford, 1991.
162. BIOVIA Available Chemical Directory. <http://accelrys.com/products/collaborative-science/databases/sourcing-databases/biovia-available-chemicals-directory.html> (accessed 14th Jan 2017).
163. BIOVIA MACCS-II Drug Data Report. <http://accelrys.com/products/collaborative-science/databases/bioactivity-databases/mddr.html>.
164. Veber, D. F.; Johnson, S. R.; Cheng, H. Y.; Smith, B. R.; Ward, K. W.; Kopple, K. D., Molecular properties that influence the oral bioavailability of drug candidates. *J. Med. Chem.* **2002**, *45* (12), 2615-23.
165. You, W.; Huang, Y. M.; Kizhake, S.; Natarajan, A.; Chang, C. E., Characterization of Promiscuous Binding of Phosphor Ligands to Breast-Cancer-Gene 1 (BRCA1) C-



- Terminal (BRCT): Molecular Dynamics, Free Energy, Entropy and Inhibitor Design. *PLoS Comput. Biol.* **2016**, *12* (8), e1005057.
166. Schneider, U. A.; Havlík, P.; Schmid, E.; Valin, H.; Mosnier, A.; Obersteiner, M.; Böttcher, H.; Skalský, R.; Balkovič, J.; Sauer, T.; Fritz, S., Impacts of population growth, economic development, and technical change on global food production and consumption. *Agric. Syst.* **2011**, *104* (2), 204-215.
  167. Ehrlich, P. R., Human impact: the ethics of I=PAT. *Ethics Sci. Environ. Polit.* **2014**, *14* (1), 11-18.
  168. Vakoch, D. A.; Dowd, M. F., *The Drake Equation: Estimating the Prevalence of Extraterrestrial Life Through the Ages*. 1st ed.; Cambridge University Press: United Kingdom, 2015.
  169. Hargrave-Thomas, E.; Yu, B.; Reynisson, J., Serendipity in anticancer drug discovery. *W. J. Clin. Oncol.* **2012**, *3* (1), 1-6.
  170. Rydberg, P.; Olsen, L.; Norrby, P.-O.; Ryde, U., General Transition-State Force Field for Cytochrome P450 Hydroxylation. *J. Chem. Theory Comput.* **2007**, *3*, 1765-1773.
  171. Zafar, A.; Reynisson, J., Hydration Free Energy as a Molecular Descriptor in Drug Design: A Feasibility Study. *Mol. Inf.* **2016**, *35*, 207-214.
  172. Knox, C.; Law, V.; Jewison, T.; Liu, P.; Ly, S.; Frolkis, A.; Pon, A.; Banco, K.; Mak, C.; Neveu, V.; Djoumbou, Y.; Eisner, R.; Guo, A. C.; Wishart, D. S., DrugBank 3.0: a comprehensive resource for 'omics' research on drugs. *Nucleic Acids Res.* **2011**, *39*, D1035-41.
  173. Wishart, D. S.; Knox, C.; Guo, A. C.; Cheng, D.; Shrivastava, S.; Tzur, D.; Gautam, B.; Hassanali, M., DrugBank: a knowledgebase for drugs, drug actions and drug targets. *Nuc. Acid. Res.* **2008**, *36*, D901-6.
  174. Wishart, D. S.; Knox, C.; Guo, A. C.; Shrivastava, S.; Hassanali, M.; Stothard, P.; Chang, Z.; Woolsey, J., DrugBank: a comprehensive resource for in silico drug discovery and exploration. *Nuc. Acids Res.* **2006**, *34*, D668-D672
  175. Pubchem. <https://pubchem.ncbi.nlm.nih.gov> (accessed 25th Jun 2015).
  176. ChemSpider. <http://www.chemspider.com/>.
  177. Schrödinger *QikProp*, v4.6; Schrödinger: 2016.