



THE UNIVERSITY OF
SYDNEY

COPYRIGHT AND USE OF THIS THESIS

This thesis must be used in accordance with the provisions of the Copyright Act 1968.

Reproduction of material protected by copyright may be an infringement of copyright and copyright owners may be entitled to take legal action against persons who infringe their copyright.

Section 51 (2) of the Copyright Act permits an authorized officer of a university library or archives to provide a copy (by communication or otherwise) of an unpublished thesis kept in the library or archives, to a person who satisfies the authorized officer that he or she requires the reproduction for the purposes of research or study.

The Copyright Act grants the creator of a work a number of moral rights, specifically the right of attribution, the right against false attribution and the right of integrity.

You may infringe the author's moral rights if you:

- fail to acknowledge the author of this thesis if you quote sections from the work
- attribute this thesis to another author
- subject this thesis to derogatory treatment which may prejudice the author's reputation

For further information contact the University's Director of Copyright Services

sydney.edu.au/copyright

**DESIGN, SYNTHESIS AND TESTING OF NOVEL
ANTI-CANCER AGENTS TARGETING SECRETORY
PATHWAY CALCIUM ATPASE**

Thesis submitted in fulfilment for the degree of

MASTER OF PHILOSOPHY

FACULTY OF PHARMACY

UNIVERSITY OF SYDNEY

2014



THE UNIVERSITY OF
SYDNEY

By

Jennifer Hsuan-Yu Lin

CONTENTS

Acknowledgement	6
Declaration.....	7
Abbreviations	8
Chapter 1: Introduction.....	11
1.1: Breast cancer.....	11
1.1.1: Definition and subtypes	11
1.1.2: Current chemotherapy for basal-like breast cancer	12
1.2: Calcium signalling mechanism and calcium homeostasis.....	13
1.2.1: Membrane proteins for calcium homeostasis	14
1.2.2: Store-operated calcium entry	14
1.3: Calcium ATPases.....	17
1.3.1: Sarcoplasmic/endoplasmic reticulum calcium ATPase (SERCA)	17
1.3.1.1: SERCA inhibitors	21
1.3.1.1.1: Cyclopiazonic acid.....	21
1.3.1.1.2: Thapsigargin	22
1.3.1.2: SERCA-related diseases	23
1.3.2: Plasma membrane calcium ATPase (PMCA).....	23
1.3.2.1: PMCA-related diseases.....	25
1.3.3: Secretory pathway calcium ATPase (SPCA).....	25
1.3.3.1: SPCA-related disease.....	27
1.3.3.2: Role of calcium signalling in cancer.....	27
1.4: Molecular modelling.....	29
1.4.1: Grid-based liganding docking with energetic (Glide)	29
1.4.2: Sitemap and receptor grid generation	32
1.4.3: Protein preparation.....	34
1.4.4: Glide scores and MM-GBSA methods	35
1.5: Aims of the project	37

Chapter 2: Methods.....	39
2.1: <i>in silico</i> molecular modelling	39
2.1.1: Software and software components used.....	39
2.1.2: Comparative or homology modelling	39
2.1.2.1: SPCA	40
2.1.2.2: SERCA.....	41
2.1.3: Identifying ligand binding sites	43
2.1.4: Docking of ligands.....	44
2.1.4.1: Ligand preparation.....	45
2.1.4.2: Ligand docking on SPCA models.....	45
2.1.4.3: Ligand docking on SERCA models.....	47
2.1.5: CombiGlide trial	48
2.2: Compound availability.....	50
2.3: Bioassay.....	50
2.3.1: MDA-MB-231 cell culture	50
2.3.2: Fluorometric imaging plate reader (FLIPR)	51
2.3.2.1: Dose-response measurement.....	53
2.4: Chemical synthesis	53
2.4.1: Molecular docking of compounds with pyrrolidone core.....	53
2.4.2: Experimental.....	58
2.4.2.1: Preparation of 1-(2,4-dihydroxy-3-methylphenyl)-2-phenylethanone	58
2.4.2.2: Preparation of diallyl bisphenol A	59
2.4.2.2.1: Preparation of diallyl ether of bisphenol A.....	59
2.4.2.2.2: Rearrangement of ether to 3,3'-diallyl bisphenol A.....	59
2.4.2.3: Preparation of 3,3'-diallyl- <i>o,o'</i> -bis(carboxymethyl)bisphenol A.....	60
2.4.2.4: Preparation of (1 <i>E</i> ,4 <i>Z</i> ,6 <i>E</i>)-5-hydroxy-1,7-bis(4-hydroxy-3-methoxyphenyl)hepta-1,4,6-trien-3-one	61
2.4.2.4.1: X-ray data collection and reduction.....	62
2.4.2.5: Preparation of 3,5-bis{2'-(4-hydroxy-3-methoxyphenyl)-(1' <i>E</i>)-ethen-1'-yl}1 <i>H</i> -pyrazole.....	64

Chapter 3: Results and discussion.....	65
3.1: Molecular modelling results	65
3.1.1: SPCA	65
3.1.1.1: SPCA2	65
3.1.1.2: SPCA1d	68
3.1.2: SERCA.....	71
3.1.2.1: SERCA1.....	71
3.1.2.2: SERCA2.....	71
3.1.2.3: SERCA3.....	72
3.1.3: CombiGlide results	72
3.1.4: Results from second cycle of molecular modelling.....	77
3.1.5.1: Results from first cycle of compound selection.....	81
3.1.5.2: Results from second cycle of compound selection.....	85
3.2: Dose-response curves from FLIPR.....	86
3.3: Intracellular calcium response curves from FLIPR	90
3.3.1: Results for compound 8 and compound 3	90
3.3.2: Results for compounds 10, 13, 15, 16 and 17.....	100
3.4: Discussion.....	106
3.4.1: Evaluation of molecular modelling results	106
3.4.1.1: SPCA	106
3.4.1.2: SERCA.....	109
3.4.1.3: Evaluation of CombiGlide results.....	110
3.4.2: Evaluation of results from tested compounds.....	112
3.4.2.1: Compound 8 and compound 3.....	112
3.4.2.1.1: Limitations from molecular modelling and FLIPR testing.....	115
3.4.2.2: Compounds 10, 13, 15, 16 and 17	119
3.4.2.2.1: Limitations from molecular modelling.....	127
3.4.2.2.2: Potency and structural aspects of tested compounds	131
3.4.2.3: FLIPR.....	133
3.4.3: Conclusions and future directions.....	135

References.....	139
Appendix A: Literature review	150
Appendix B: Cell culture	152
Appendix C: FLIPR protocol.....	154
Appendix D: Repeat FLIPR tests.....	161
Appendix E: Homology model sequence alignments.....	165

ACKNOWLEDGEMENT

I would like to sincerely thank my supervisor, Professor David (Dai) Hibbs and associate supervisor, Professor Paul Groundwater for their abundant and wonderful knowledge, expertise and advice towards every parts of this project. Our postdoc, Linda, has also diligently guided me through the chemistry experiments involved to make the compounds required for testing. Special thanks were also given to Vivian for the training and guidance on the synthesis of curcumin and its cyclic pyrazole analogue. Greatest appreciation should also be expressed for all my fellow colleagues from N213. Without them, life was not as interesting and stimulating as ever before throughout this MPhil project.

I would also like to thank Professor Gregory Monteith and his team at the School of Pharmacy, University of Queensland. Their postdocs, Amelia and Iman, have patiently trained me throughout all the biological experiments involved in this project when I spent time with them in June 2013 and also involved in the compound testings which followed after. Their fellow PhD students have also greatly assisted with every part of the procedures involved.

Finally, I have to thank my family and friends for all their supports both physically and mentally throughout this journey. Without everyone mentioned here, it would not be possible to complete this project. Thank you to all.

DECLARATION

I would like to declare that all the work involved within this project was true and original. All the experiments conducted were carried out by the author under the supervision, guidance and assistance from Professor David (Dai) Hibbs and Professor Paul Groundwater at Faculty of Pharmacy, University of Sydney and also Professor Gregory Monteith from School of Pharmacy, University of Queensland unless otherwise stated.

ABBREVIATIONS

ADP	Adenosine diphosphate
ATP	Adenosine triphosphate
BHQ	2,5-di-(<i>t</i> -butyl)-1,4-benzohydroquinone
BRCA	Tumour suppressor gene
Ca ²⁺	Calcium (ion)
¹³ C	Carbon
CPA	Cyclopiazonic acid
DCM	Dichloromethane
DMF	Dimethyl formamide
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
Dscore	Druggability score
DMEM	Dulbecco's Modified Eagle Medium
EDTA	Ethylenediaminetetraacetic acid
EtOAc	Ethyl acetate
ER	Endoplasmic reticulum; Oestrogen receptor
HER2	Human epidermal growth factor receptor
HHD	Hailey-Hailey disease
HTVS	High throughput virtual screening
HCl	Hydrochloric acid

IAM	Independent Atom Model
IR	Infrared
IGF1R	Insulin-like growth factor receptor
Ins(1,4,5)P ₃	Inositol-1,4,5-triphosphate
IP ₃ R	Inositol-1,4,5-triphosphate receptor
Mn ²⁺	Manganese (ion)
MW	Molecular weight
NCBI	National Centre for Biotechnology Information
NP	4- <i>n</i> -Nonylphenol
NMR	Nuclear magnetic resonance
ORTEP	Oak Ridge Thermal Ellipsoid Plot
PBS	Phosphate buffered saline
P-type	Phosphorylation-type
PB	Poisson-Boltzmann
PDB	Protein Data Bank
PMCA	Plasma-membrane calcium ATPase
PMR1	Plasma membrane ATPase-related 1
PMR2	Plasma membrane ATPase-related 2
PR	Progesterone receptor
¹ H	Proton
RCSB	Research Collaboratory for Structural Bioinformatics

NaCl	Sodium chloride
Na ₂ SO ₄	Sodium sulphate
SCREEN	Surface Cavity Recognition and Evaluation
SERCA	Sarcoplasmic or endoplasmic reticulum calcium ATPase
SOCE	Store-operated calcium entry
SP	Standard precision
SR	Sarcoplasmic reticulum
SPCA	Secretory pathway calcium ATPase
THF	Tetrahydrofuran
TG	Thapsigargin
TLC	Thin layer chromatography
XP	Extra precision
2D	Two-dimensional
3D	Three-dimensional

Poster presentation

ASCEPT 2013, Melbourne, Australia

“Design, synthesis and testing of novel anti-cancer agents targeting secretory pathway calcium ATPase”

CHAPTER 1: INTRODUCTION

1.1 Breast cancer

1.1.1 Definition and subtypes

The main subtypes of breast cancer were originally identified *via* DNA microarray and were categorised into luminal A, luminal B, human epidermal growth factor receptor (HER2) 2-positive, basal-like, claudin-low and normal-like subtypes.¹⁻³ Luminal subtypes were all known to be oestrogen receptor (ER) – positive. Luminal A tumours were low-grade with progesterone receptor (PR) – positive and HER2 – negative features. Luminal B tumours were of higher grade, PR-positive or PR-negative and either HER2-positive or HER2- negative.⁴ About 13 – 20% of invasive breast tumours had shown HER2-positivity which were treated with trastuzumab, a humanised anti-HER2-specific monoclonal antibody. Claudin-low subtypes usually had low to absent expression of luminal markers, abundant epithelial-to-mesenchymal transition markers, immune response genes and cancer stem cell-like features.⁵

Basal-like breast cancer, on the other hand, has drawn many controversial discussions over its proper pathological definition due to the heterogeneous nature of the disease. So far, it has been defined in both immunohistochemical and deoxyribonucleic acid (DNA) microarray-related perspectives and there was yet no internationally-accepted definition. Basal-like breast cancer possessed triple-negative markers which meant that these basal-like tumour cells were HER2-negative, ER-negative and PR-negative when examined in DNA microarray analysis. The immunohistochemical features for these tumour subtypes consisted of cytokeratin 5/6-positive

and/or epidermal growth factor receptor-positive. Because of the presence of the triple-negative markers for basal-like breast cancer, it could be ambiguously referred to as triple-negative breast cancer. However, after extensive studies were done it was shown that these two cancer subtypes were not synonymous as there were some overlaps in the disease features between these two subtypes of breast cancer.^{1,2,6} To further emphasise the heterogeneity of breast cancer, it was also found that about both basal-like breast cancer and BRCA1-mutated tumours shared many common features and about 75% of BRCA1-related tumours were discovered to be basal-like breast cancer.^{7,8}

1.1.2 Current chemotherapy for basal-like breast cancer

Basal-like breast cancer is the most aggressive breast cancer subtype with the poorest prognostic outcome.^{1,9} It is more likely to occur in pre-menopausal women who are younger than 40 years of age. Currently there are very minimal specific and effective treatment agents available to target basal-like breast cancer. It has a poor response to most of the conventional chemotherapies used to treat other breast cancer subtypes.^{1,10} From recent reviews on the treatment of basal-like breast cancer, it is confirmed that conventional chemotherapy and/or DNA damaging agents for the duration of at least four to six months is the best possible treatment regimen to date for basal-like breast cancer patients.^{9,11} Several other single agents such as bevacizumab, polyadenosine diphosphate-ribose polymerase inhibitors or agents targeting epidermal growth factor receptor are presented with conflicting results after trials in triple-negative breast cancer.

1.2 Calcium signalling mechanism and calcium homeostasis

Calcium (Ca^{2+}) is an important signalling agent coordinating a variety of physiological processes in the body.¹² The signalling role of calcium was first discovered by serendipity in 1883 where isolated rat hearts were found to contract in hard water rather than in distilled water.¹³ Further signalling roles were later investigated and numerous cellular processes such as cell fertilisation, proliferations and apoptosis were found to have calcium involved as the essential messenger.¹³ Ca^{2+} signalling are crucial for a vast amount of cellular functions such as short-term responses in contraction and secretion and longer-term control of transcription, cell division and apoptosis.¹⁴ In order to achieve balanced calcium signalling processes in our body, calcium homeostasis is required to be maintained at all time as it is also a critical step for cell survival.¹⁵

There are two main sources of signal calcium which are from the internal cellular stores and external medium.¹⁵ The internal stores are mainly kept in the sarcoplasmic reticulum (SR) or endoplasmic reticulum (ER) where the signal calcium are controlled either by calcium itself or a group of messengers such as inositol-1,4,5-triphosphate ($\text{Ins}(1,4,5)\text{P}_3$), cyclic adenosine diphosphate (ADP) ribose, nicotinic acid adenine dinucleotide phosphate or sphingosine-1-phosphate. The external medium consists of the plasma-membrane channels which will control the entry of calcium ion by responding to various stimuli, for example, membrane depolarisation, stretch, noxious stimuli, extracellular agonist action, intracellular messaging or depletion of the intracellular calcium stores.¹⁵

1.2.1 Membrane proteins for calcium homeostasis

During the “ON” reaction where Ca^{2+} is introduced into the cells, Ca^{2+} enters mainly from the external medium across a large electrochemical gradient through the plasma membrane and activates various entry channels during the process.¹⁵ Other sources of Ca^{2+} entry during the reaction are from the existing internal stores from the SR or ER where the calcium entry is regulated by $\text{Ins}(1,4,5)\text{P}_3$ and ryanodine receptors. To counteract the “ON” reaction, an “OFF” reaction for the calcium homeostasis is necessary to be in place in order to switch off the calcium signal within the cytoplasm to avoid excessive accumulations of Ca^{2+} . In order to achieve the “OFF” reaction, there are four main pumps and exchangers acting as the tools for carrying out this process. These are the sodium/calcium exchangers, plasma-membrane Ca^{2+} -ATPase (PMCA), sarcoplasmic or endoplasmic reticulum calcium ATPase (SERCA) and the mitochondrial uniporter.^{13,15} By maintaining a calcium homeostasis in the cells, the resting level of Ca^{2+} is at about 100nM in the cytoplasm.

1.2.2 Store-operated calcium entry

Store-operated calcium entry (SOCE) was one of the many important cellular functions involved in many non-excitabile and excitable cells.^{14,16} SOCE was vital in cell proliferation, apoptosis and differentiation. It was also involved in cellular shape changes and regulation of T cell motility. Other roles of SOCE in cellular functions also included regulation of cellular secretion and neurotransmission. It was a process involving the depletion of intracellular Ca^{2+} in the ER store which then triggered the influx of Ca^{2+} from the extracellular space through store-operated channels (SOCs). An early example was shown by using thapsigargin (TG) as a molecular probe for understanding the intracellular Ca^{2+} storage and release.¹⁷ It demonstrated that SERCA

inhibition took place when TG was present and blocked the reuptake of Ca^{2+} into the ER. This then caused a slow-emptying of the ER Ca^{2+} store which then induced SOCE.

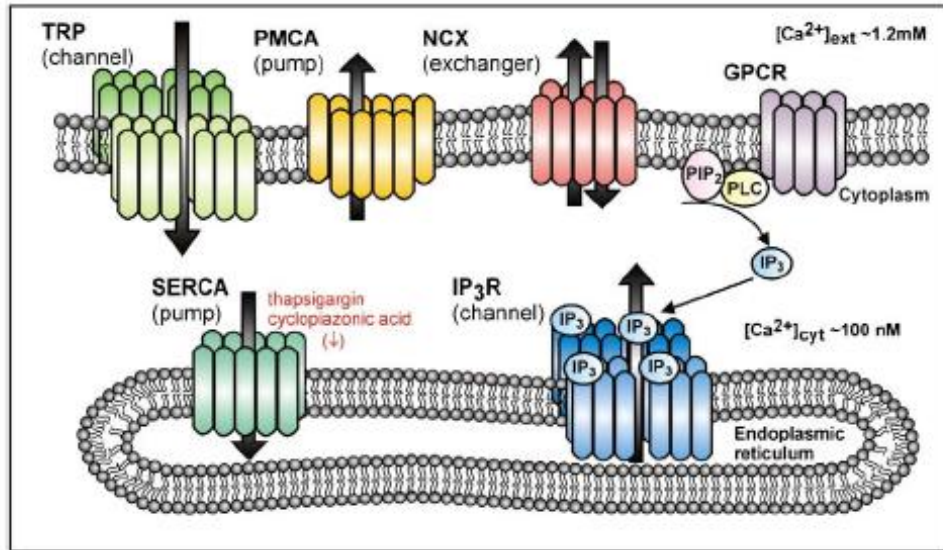


Figure 1 Different types of Ca^{2+} pumps involved in Ca^{2+} homeostasis (taken from ref. 16)

The ER Ca^{2+} store depletion takes place as the remaining Ca^{2+} in ER slowly leaks out from the IP_3R channel eventually emptying the Ca^{2+} store in the ER, while SERCA can be blocked by TG inhibiting the re-sequestration of Ca^{2+} (**Figure 1**)^{16,18}. Similar SOCE effect would also occur with other SERCA inhibitors such as cyclopiazonic acid (CPA).¹⁸ Another commonly known example of SOCE occurred when G protein-coupled receptors on the plasma membrane were activated which in turn would induce $\text{Ins}(1,4,5)\text{P}_3$ to activate the inositol-1,4,5-triphosphate receptor ($\text{Ins}(1,4,5)\text{P}_3\text{R}$) on the ER. $\text{Ins}(1,4,5)\text{P}_3\text{R}$ would then begin the release of the ER Ca^{2+}

store. The depleted Ca^{2+} store would trigger a message to be sent to plasma membrane allowing the entry of extracellular Ca^{2+} via SOCs. In a different scenario but within the same principle, SOCE would also be induced as the stromal interacting molecule (STIM1) detected the depletion. STIM1 would aggregate or translocate to ORAI1, which was a plasma membrane Ca^{2+} influx protein associated with SOCE. ORAI1 would then activate the entry of Ca^{2+} from extracellular space (**Figure 2**).

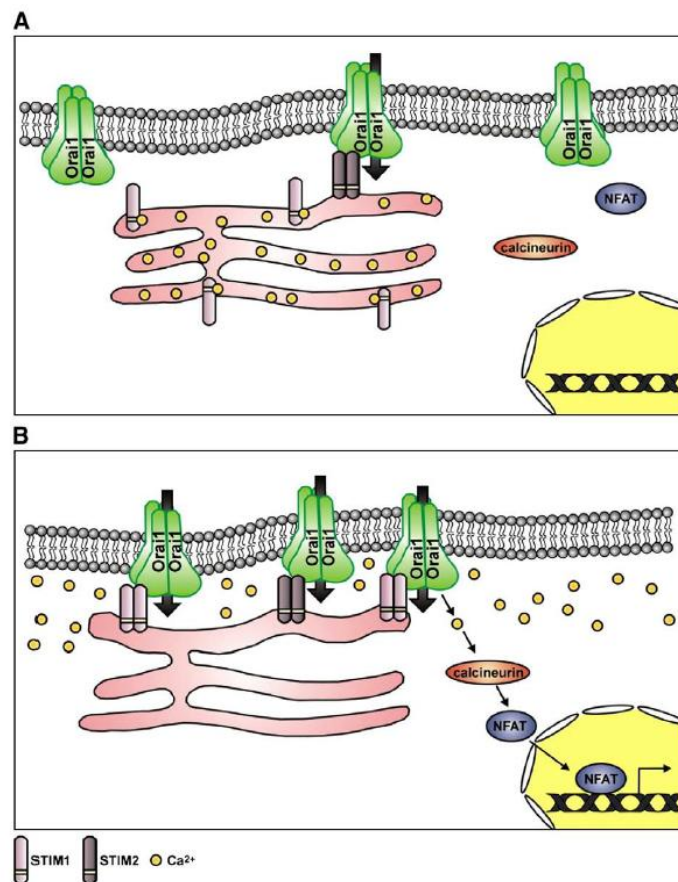


Figure 2 Store-operated Ca^{2+} entry involving ORAI1 and STIM1 (taken from ref. 16)

1.3 Calcium ATPases

1.3.1 Sarcoplasmic/endoplasmic reticulum calcium ATPase

The involvement of adenosine triphosphate (ATP) in the calcium signalling process was first identified in the 1960s where Ebashi *et. al.*^{19,20} and Hasselbach *et. al.*²¹ found sarcoplasmic reticulum vesicles accumulated Ca^{2+} by using ATP as their energy source.¹³ The transporters involved in actively transporting the cations across the cell membrane are named the phosphorylation-type (P-type) ATPases since they form a covalent aspartylphosphate group *via* the pyrophosphate bond of the ATP.^{22,23} They have originally evolved from two ancestral types where Type I P-type ATPases transport heavy metals such as Cu^+ , Cd^{2+} and Hg^{2+} and Type II transport lower atomic mass metals like Na^+ , K^+ , Ca^{2+} and H^+ .²⁴ Calcium ATPase is a member of this P-type ATPase family.

Calcium ATPases are comprised of SERCA, PMCA and Secretory Pathway Ca^{2+} -ATPase (SPCA). SERCA remains the most extensively studied calcium ATPase, and the majority of the structural studies on SERCA structure has been done by Toyoshima²⁵⁻³⁴ who has prepared and analysed the crystal structures of SERCA sourced from rabbit muscles. This has greatly assisted in understanding the relationship between its structure and function and also the mechanism of this calcium pump.²⁹ SERCA is found to be embedded in the cell membrane of sarcoplasmic or endoplasmic reticulum with a relative molecular mass of 110,000. Its structure consists of three domains in the cytoplasmic part of the SERCA and ten transmembrane α -helices named M1 to M10 (**Figure 3**).^{29,30} The anchor or actuator domain (domain A) is the smallest domain out of all three domains. It has about 110 amino acid residues located between M2 and M3 with long loops

connecting to the transmembrane region. The nucleotide domain (domain N) is the largest domain headpiece present. It has seven-stranded antiparallel β -sheet with the critical ATP-binding residue, Phe487, found within this region. Phosphorylation domain (domain P) is situated between domain A and N and is important as it processes the ATP phosphorylation site. Another important residue involved in the phosphorylation process, Asp351, is situated at the C-terminal end of the central β -strand close to M5.

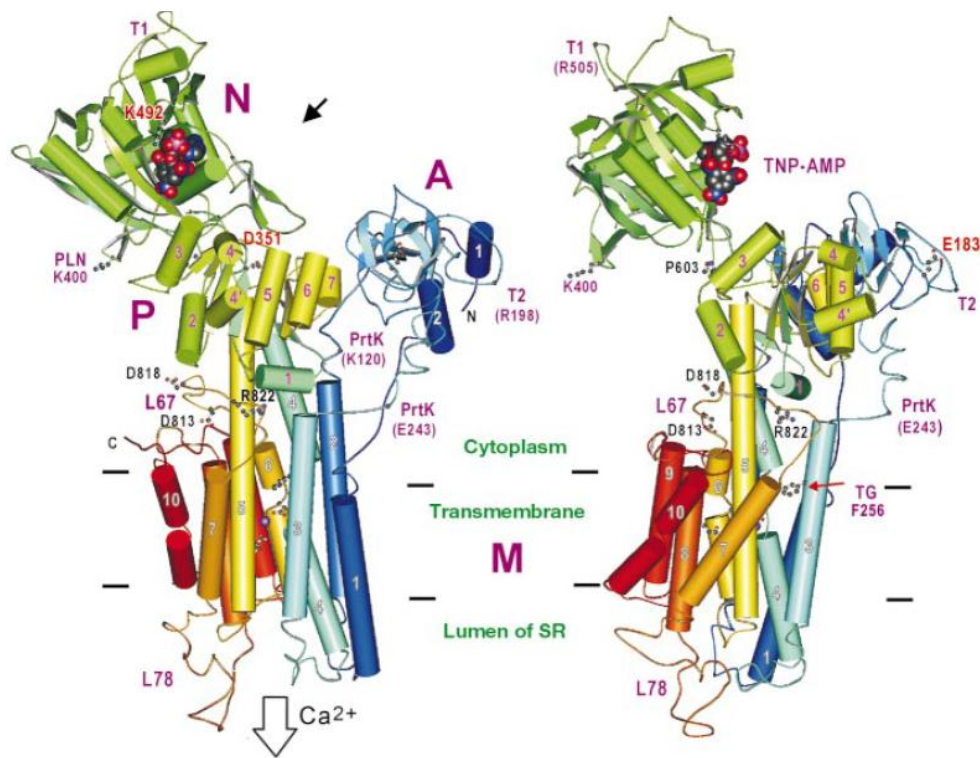


Figure 3 Structure of SERCA displaying three domains (A, N and P) with 10 transmembrane helices (M1 – M10) (taken from ref. 27)

There are two Ca^{2+} binding sites present in the M4, M5, M6 and M8 region.^{29,30} Ca^{2+} binding site one is contributed by the side-chain oxygen atoms of five residues (Asn768, Glu771, Thr799,

Asp800 and Glu908) on M5, M6 and M8. Site two is almost on the M4 helix where six residues (Val304, Ala305, Ile307, Asn796, Asp800 and Glu309) are involved in forming the binding site. Asp800 is a critical residue involved in the binding of both Ca^{2+} atoms. The binding sites are stabilised by extensive hydrogen-bond networks amongst the participating residues and also the transmembrane helices.

During the muscle relaxation phase, SERCA pumps Ca^{2+} from the cytoplasm into the reticulum which is known as the E1 phase of the E1-E2 model. The E1-E2 model shown in **Figure 4** is also known as the Post-Albers cycle.^{30,35-37} It involves the phosphorylation of one ATP with two Ca^{2+} transferred across the membrane. Initially the two Ca^{2+} or ATP from the cytoplasm bind to the E1 conformation of the SERCA. This transforms the calcium ATPase into a high-energy phosphointermediate (2Ca^{2+} -E1-ATP). Phosphorylation then follows to produce an ADP sensitive and another high energy intermediate (2Ca^{2+} -E1-P). The next conformational change takes place to reach the E2 phase and forms the ADP insensitive and low energy intermediate (2Ca^{2+} -E2-P). At this point the calcium-binding sites switch to face the luminal side and the two Ca^{2+} and phosphate are released to bring the calcium ATPase back to the E1 phase. During the whole process two or three H^+ are also counter-transported.²³ This phosphorylation cycle is a classic characteristic of the P-type ATPases.

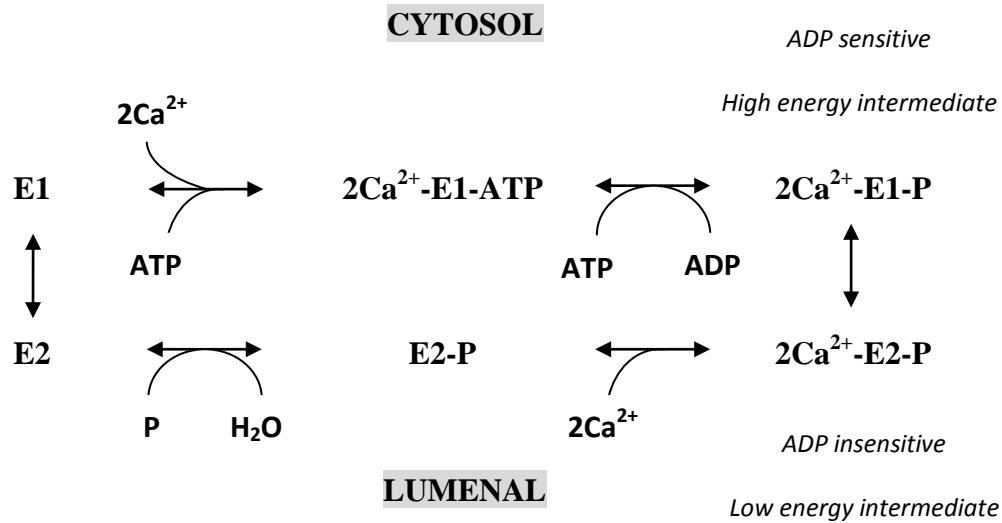


Figure 4 Post-Albers cycle showing the transition between E1 to E2 states of P-type ATPases

SERCA is coded by genes ATP2A1, ATP2A2 and ATP2A3.¹² There are three main SERCA isoforms present from alternative gene splicing and this gives rise to SERCA1a and 1b, SERCA2a, 2b and 2c and SERCA3a, 3b, 3c, 3d, 3e and 3f subtypes. Each SERCA isoform are expressed in different types of tissues, for example, SERCA1a and 1b are present in fast-twitch skeletal muscle, whereas SERCA2a are mostly found in the brain, cardiac, slow-twitch and smooth muscle cells.^{12,24} SERCA2b is expressed in all tissue types and 2c are found in epithelial, mesenchymal and haematopoietic cells. SERCA3a, 3b, 3c and 3e are in most tissues and some can be discovered in haematopoietic cells, salivary glands, trachea, lung, pancreas, kidney and colon.

1.3.1.1 SERCA inhibitors

1.3.1.1.1 Cyclopiazonic acid

Cyclopiazonic acid (CPA) **1** (Figure 5) is an indole tetramic acid metabolite of *Aspergillus* and *Penicillium*.^{38,39} It inhibits SERCA with high specificity and affinity. The mechanism of inhibition is implemented by blocking the calcium access channel and making the structures of the involved transmembranes rigid and hence immobilised.⁴⁰ The CPA binding site involves amino acid residues from M1, M2, M3 and M4 (Leu61 and Val62 from M1, Asn101 and Ala102 from M2, Gly257 from M3, Leu311 and Pro312 from M4). The structure of CPA consists of two main regions, the tetramic acid portion and a hydrophobic region containing an indole nucleus. This indole nucleus is important for exerting the inhibitory effect of CPA as it bridges from M1 to M4 helices. Removal of this indole nucleus eliminates the inhibiting effect of CPA.⁴¹

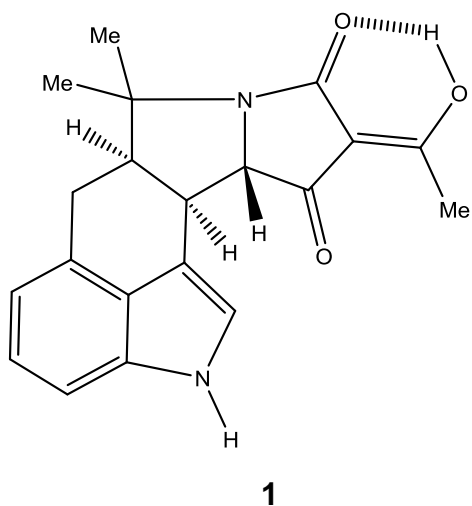


Figure 5 Structure of CPA

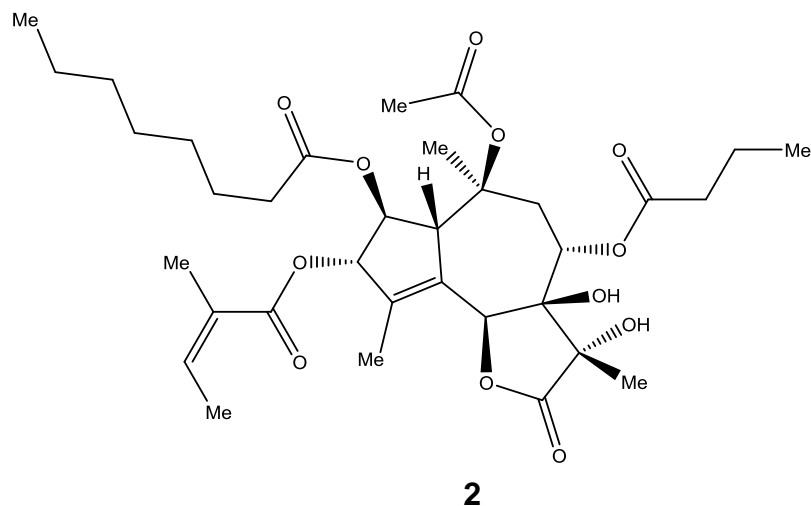


Figure 6 Structure of TG

1.3.1.1.2 Thapsigargin

Thapsigargin (TG) **2** (**Figure 6**) is a naturally occurring sesquiterpene lactone compound, normally derived from *Thapsia garganica*.^{42,43} It has been found to inhibit ER Ca²⁺ ATPase with high potency. This is largely observed when the intracellular Ca²⁺ store increases after TG is added to rat liver microsomes. TG also inhibits SR Ca²⁺ ATPase obtained from rabbit skeletal muscle.⁴⁴ This is achieved by using a subnanomolar concentration of thapsigargin to produce the total inhibition of SR Ca²⁺ ATPase. As a result, TG is the most potent and specific SERCA inhibitor since the same effect was not observed in PMCA in erythrocytes.

A crucial amino acid residue within SERCA, Phe256, is discovered to render TG ineffective when it is mutated.⁴⁵ The effect occurs when Phe256 mutates to Leu256 in a Chinese hamster lung fibroblast cell line after it is subjected to different TG concentration treatments.⁴⁵ The cell line developed resistance to the TG inhibition of the Ca²⁺ transport function within the cell. A similar TG-resistant effect also occurs in the Syrian hamster smooth muscle cell line within the same study where Phe256 mutates to Ser256.⁴⁵ These evidences suggest that Phe256 is a critical amino acid residue in the SERCA and TG interaction.

Another study has conducted F256V mutations in all three isoforms of SERCA (1b, 2b and 3a) expressed in COS-7 cells which are treated with four different SERCA inhibitors, TG, CPA, 2,5-di-(^tbutyl)-1,4-benzohydroquinone (BHQ) and 4-*n*-Nonylphenol (NP)⁴⁶. TG sensitivity is noted to vary in different SERCA isoforms and after the F256V mutation, TG sensitivity reduces across all three isoforms, but with variable results among them. Other SERCA inhibitors such as

CPA, BHQ and NP are only modestly affected by the F256V mutation. This indicates these inhibitors bind to a different hydrophobic site than TG. TG typically binds to M3 (Phe256), M5 (Ile765) and M7 (Tyr837) helices.⁴³

1.3.1.2 SERCA-related diseases - Darier disease and Brody's disease

One of the documented cases of SERCA-related diseases was in 1969 involving a reduced Ca^{2+} uptake by SERCA.⁴⁷ The symptoms in individuals with this defect included muscle stiffness during exercises and painless cramps, this became known as Brody's disease. Its main feature involved a slow muscular relaxation phase which caused the unusual muscle stiffness during exercise. Biopsy evidence was also found in quadriceps and cultured muscle cells in Brody's disease patients with reduced SERCA1 pump activity.^{48,49} The reduced SERCA1 pump activity was not only found in slow-twitch but also in fast-twitch muscles in some clinical studies.⁴⁹ Another SERCA-related disorder is Darier-White disease, also known as Darier's disease or keratosis follicularis⁵⁰. Darier's disease is known to be a rare autosomal dominant skin disorder and was discovered earlier than Brody's disease. It involved seborrheic papules in upper trunk, scalp, palmar pits and other skin flexures. The major histological observation was the loss of epidermal cell adhesions and abnormal keratinisation due to mutations in genes encoding SERCA pumps.

1.3.2 Plasma membrane calcium ATPase (PMCA)

The structure of PMCA was predicted to be similar to SERCA with ten transmembrane helices and also the three A, N and P domains.^{12,51} There are no PMCA protein structures solved to date

but comparative modelling has been used to build the tentative PMCA structure based on the known protein structures of SERCA and protein sequence of PMCA. This however is still insufficient to inform the actual mechanism of its long C-terminal tail which was known to be involved in two different conformational states of PMCA. The location of PMCA as suggested by its name was mainly on the plasma membrane of the cells and was first discovered in 1966 from erythrocytes.⁵² PMCA was coded by four different genes which have undergone alternative splicing at two sites of its protein sequence to produce four different PMCA isoforms, PMCA1, 2, 3 and 4.^{12,51} The location of each isoform differed where PMCA 1 and 4 were expressed ubiquitously and type 2 and 3 were mainly restricted to the brain. Some of the PMCA2 were also found in mammary gland tissues and some PMCA3 in skeletal muscles.

Calmodulin is known to activate and regulate PMCA activity as it increased the Ca^{2+} binding affinity to the pump.⁵¹ PMCA pumps are also known to be activated by acidic phospholipids such as phosphatidylserine, cardiolipin, phosphatidylinositol and phosphatidic acid but its physiological significance is yet unknown.⁵³ Another agent, calpain, was also found to activate PMCA pump irreversibly *via* cleaving the C-terminal tail.^{12,51} On the other hand, there were no known PMCA inhibitors comparable to TG for SERCA pumps. However, one suggested PMCA inhibitor was the caloxin peptides which were found to bind to the extracellular domain of PMCA and had selective inhibitory effect on different PMCA isoforms.⁵⁴

1.3.2.1 PMCA-related diseases

The four isoforms of PMCA also had their own individual impact on disease pathology. Since PMCA1 was distributed widely in the tissues, it has crucial effects on embryonic development and organogenesis. Other interesting findings also linked PMCA1 to human systolic blood pressure and coronary artery disease.^{55,56} PMCA 4, on the other hand, did not affect the embryonic development as far as PMCA1 and instead has more impact on male fertility.⁵⁷ Defects in genes encoding PMCA4 were likely to reduce sperm motility which led to sterility in males. PMCA4 was also found to regulate cardiac contractility in both *in vitro* and *in vivo* studies.^{58,59} As alluded to previously, PMCA2 were mainly concentrated in the central nervous system, especially in the cerebellum. It was thus expected to observe some central nervous system related problems when there were defects or mutations in the PMCA2 genes. One of such problems was discovered in PMCA2 mutant mice with ataxic symptoms found in standing and walking. PMCA2 gene mutations were also associated with deafness in some human cases.^{60,61} Similar to PMCA2, PMCA3 is mainly distributed in the central nervous system and has been found to be related to a familial X-linked congenital cerebellar ataxia in PMCA3 mutation.⁶²

1.3.3 Secretory pathway calcium ATPase

A new subset of calcium ATPase was first discovered in yeast, *Saccharomyces cerevisiae*.⁶³ Two genes, PMR1 (plasma membrane ATPase-related 1) and PMR2 (plasma membrane ATPase-related 2), isolated from the yeast were found to encode the calcium pump in the secretory pathway. The location of the yeast calcium pump is identified in the Golgi apparatus where ATP is required for calcium transport.⁶⁴ Similar finding has also been shown by Sorin where yeast lysate fractions were assayed and PMR1 in the fractions was found to contain Golgi marker,

guanosine diphosphatase.⁶⁵ Other supporting evidence includes when SERCA inhibitors such as TG, CPA and vanadates did not inhibit PMR1 Ca^{2+} ATPase with the same intensity as that in SERCA. Another study has also characterised a PMR1 homologue in *Caenorhabditis elegans*. The PMR1 proteins expressed in COS-1 cells have accumulated Ca^{2+} and manganese (Mn^{2+}) in the Golgi apparatus.⁶⁶ A subsequent study was conducted in mammalian cells and the presence of PMR1 was confirmed with similar inhibitor sensitivities as the previous study in *Caenorhabditis elegans*.⁶⁷

This new subgroup of Ca^{2+} ATPase was later named SPCA in humans.⁶⁸ This is in line with the nomenclature and phylogenetic tree for SERCA and PMCA. SPCA is localised to the Golgi which is responsible for the modification and transport of proteins and biosynthesis of lipids.⁶⁹ One of the major functions of SPCA is Ca^{2+} and Mn^{2+} transport into the secretory pathway.⁶⁸ Mn^{2+} transport is important for protein glycosylation, especially N-linked and O-linked glycosylations.⁷⁰ SPCA also helps with cytosolic Ca^{2+} signalling as demonstrated in HeLa cells with SPCA1 inhibition via RNA-mediated interference.⁷¹

There are two human genes, ATP2C1 and ATP2C2, encoding for SPCA which gives two SPCA subtypes; SPCA1 and SPCA2.⁷² ATP2C1 undergoes alternative splicing at the 3'-end of the primary mRNA and this gives rise to four different SPCA1 isoforms, SPCA1a, 1b, 1c and 1d. Newer SPCA isoforms such as SPCA1e and SPCA1f have also been identified recently.⁷³ SPCA2, on the other hand, only exists in one isoform as ATP2C2 does not undergo alternative splicing. SPCA2 has been investigated to possess the same ability as SPCA1 to transport Ca^{2+} and Mn^{2+} without being affected by thapsigargin.⁷⁴ The tissue expressions for human SPCA2 are

in the gastrointestinal and respiratory tract, prostate, thyroid, salivary and mammary glands. SPCA1 is mainly found in human epidermal keratinocytes and also in most of other tissues such as brain, heart, lungs and kidneys.⁷⁵

1.3.3.1 SPCA-related disease – Hailey-Hailey disease

A well-known example of an inherited autosomal dominant disease resulted from ATP2C1 mutation is Hailey-Hailey disease (HHD).^{75,76} HHD was characterised by skin blistering and erosions in carriers of the mutation. Human epidermal keratinocytes gathered from the carriers of ATP2C1 mutation have revealed inadequate intracellular Ca²⁺ regulation under both resting and stimulated states. This potentially affected the keratinocyte cell-to-cell adhesions leading to HHD's clinical symptoms. This was also one of the examples of a clinical disorder related to mutations in human Ca²⁺ ATPase. Another well-known disorder related to impaired Ca²⁺ ATPase was the aforementioned Darier's disease which also shared similar clinical symptoms. It was therefore justifiable to state that both SERCA and SPCA pumps were essential for regulating and maintaining Ca²⁺ homeostasis in keratinocytes. Disruption to these Ca²⁺ pumps and thus the Ca²⁺ homeostasis would likely lead to these diseases. This link was demonstrated when a loss of a functional allele of human gene, hSPCA1, was found to relate to HHD.⁶⁸

1.3.3.2 Role of calcium signalling in cancer

More recent findings have subsequently found an important link between calcium signalling and cancer. Calcium signalling can regulate tumourigenesis, enabling cancer cells to escape apoptosis and maintain angiogenesis to assist cancer metastases.^{77,78} Two studies have

demonstrated this link *via* producing heterozygous mutations of ATP2A2 encoding SERCA2 in mice. This has resulted in squamous cell tumours in the murine skin, oesophagus, oral mucosa and stomach.^{79,80} Another study looks at the effect of mutating ATP2C1 encoding SPCA1 in mice.⁸¹ Similar effect has resulted where adult mice with heterozygous ATP2C1 gene expressions developed squamous cell cancers without any evidence of HHD. Cases of calcium signalling related to some cancers were also identified in remodelled PMCA isoform expressions.^{78,82,83} It was found that altered Ca²⁺ efflux ability would reduce the sensitivity of the cells to apoptosis and increase the response of cancer cells to cellular proliferative stimuli. PMCA1 and 4 were found to be highly expressed in breast cancer cell lines such as MCF-7 and MDA-MB-231.⁸⁴⁻⁸⁶ Silencing of PMCA4 facilitated the apoptosis of MDA-MB-231 cells mediated by the Bcl-2 inhibitor ABT-263 which was being evaluated in clinical trials treating small-cell lung cancer and B cells-related tumours.^{82,87} Therefore it was suggested that blocking PMCA4 might enhance the effectiveness of some anti-cancer therapies.

A recent study by Grice *et. al.* has focussed on the connection between SPCA1 and different breast cancer subtypes. SPCA1 is found to be associated with basal-like breast cancers, which, as mentioned earlier, has the poorest prognosis with minimal therapeutic agents available.⁸⁸ Increased expression of insulin-like growth factor receptor (IGF1R) has been identified in this study, which has a strong involvement in cancer initiation, proliferation and resistance to anti-cancer therapy. This finding provides a window of opportunity to place SPCA1 as a new therapeutic target for basal-like breast cancer. In another words, if SPCA1 is inhibited, will this reduce the IGF1R production that will ultimately inhibit basal-like breast cancer tumour growth? Therefore the aims of this research project are to fully understand the role of SPCA1 in basal-like

breast cancer and to design, synthesise and test chemical compounds that will inhibit the growth of basal-like breast cancer cells *in vitro*.

1.4 Molecular modelling

1.4.1 Grid-based ligand docking with energetics (Glide)

The Glide program suite essentially searches for favourable interactions between a ligand and a macromolecular receptor, which is usually a protein molecule.⁸⁹ It carries out this function based on using a hierarchical filter to evaluate the interactions between selected ligands and the target receptor.⁹⁰⁻⁹² The filter tests the spatial fit of the ligand to the defined active binding site. Then it will assess the ligand-receptor interactions by using a grid-based method produced from an empirical ChemScore function.⁹³ Ligand poses that have passed the initial screening will enter into the final stage of the hierarchy.

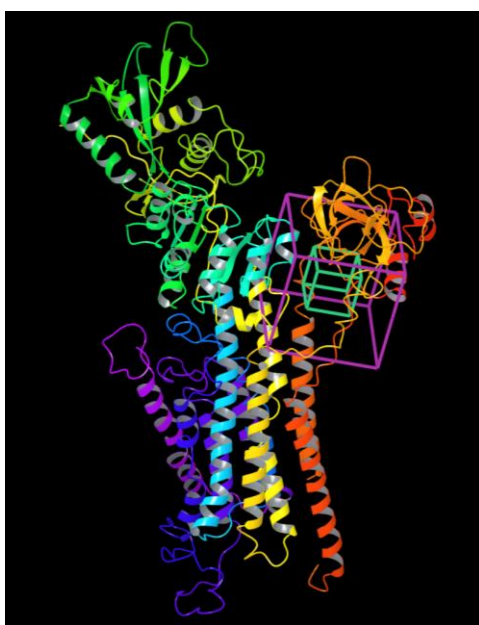


Figure 7. GLIDE generated receptor grid (shown as two boxes representing the grid boundary)

The first stage of the hierarchy filter is the Site-Point Search.⁸⁹ This initially involves looking for all possible ligand locations in the active binding site. Each ligand molecule is divided into a core region and a number of rotamer groups due to being attached to its core *via* rotatable bonds. Each ligand molecule will have about 500 conformations generated based on the number of rotatable bonds, any 5- and 6- membered rings, any presence of long-range internal hydrogen bonds and asymmetric pyramidal trigonal nitrogen centres in the core. Then, during the docking process, every possible location and orientation will be searched over the entire active site of the protein for the generated ligand conformations. The active site or ligand binding site is represented by two boxes which will be referred to as “grid” from hereafter in the Site-Point Search.⁸⁹ A 2Å grid represented by the purple box in **Figure 7** is the volume in which the grid potentials are computed and where all the ligand atoms are located. The purple box is formed based on the distances from the site point to the receptor surface. These distances are evaluated by the programme at a series of pre-specified directions and then sorted into different distance ranges covering the active site region. A second 1Å grid, represented by the green box in **Figure 7**, is then assigned based on the distances from ligand centre, which is the midpoint of the two most separated atoms, to the ligand surface. The green box represented the volume that the ligand centre explores during the Site-Point Search. Therefore, the whole Site-Point Search process runs from comparing the distance ranges from the site point to the receptor and the distance from the ligand centre to its surface (purple grid (box) versus green grid (box) as shown on **Figure 7**). The Glide programme will place the ligand centre at the site point for a good match within the Site-Point Search.

The second stage of the hierarchical filter is followed by Diameter Test which is to examine the atom placements within a specified distance of a line between the most separated atoms.⁸⁹ This is also known as the ligand diameter. The ligand-receptor orientation will be skipped if there are any steric clashes present within the ligand diameter. The rotation of the ligand around the ligand diameter is also tested. To evaluate this property, a Greedy Scoring system⁹³ is set up based on Schrödinger's own version of ChemScore empirical scoring function. Each subset of ligands or atoms that are capable of making hydrogen bonds or ligand-metal interactions with the receptor will be scored *via* this Greedy Scoring. Greedy Score also takes into account any favourable hydrophobic interactions, and unfavourable steric clashes and allows atoms to move $\pm 1\text{\AA}$ in x, y or z directions to compensate for large 2\AA jump in the site point or ligand centre position. The top Greedy Scores will then be re-scored *via* a refinement procedure where the ligands are allowed to move rigidly by $\pm 1\text{\AA}$ in the Cartesian directions.

Only about 100 to 400 ligand-receptor poses will reach the final Grid Minimisation stage. This stage mainly focuses on conducting energy minimisation on a set of van der Waals electrostatic grids which have been fine-tuned to reduce the large energy and gradient terms resulting from close inter-atomic interactions. Eventually, the energy-minimised poses will be re-scored again by using Glide score scoring function generated by Schrödinger.⁹² Glide (G) score is based on ChemScore as mentioned previously but also takes into account of steric-clash term, other rewards and penalties such as buried polar terms, amide twice penalties, hydrophobic enclosure terms and excluded volume penalties plus other modifications to other terms mentioned in the G score algorithm developed by Schrödinger (as shown below). The standard unit of G score was measured in kilocalories per mole (kcal/mol) (**Equation 1**).

G score = 0.05 x van der Waals interactions + 0.15 x Coulomb energy + Lipophilic term + Hydrogen bond + Metal-binding term + Rewards + Penalty for freezing rotatable bonds + Polar interactions in the active site (polar but non-hydrogen bonding atoms in a hydrophobic region are also rewarded)

Equation 1. Schrödinger Glide Score

1.4.2 Sitemap and receptor grid generation

In the field of structure-based drug design, there are times when the actual drug binding site is not known or alternative binding sites are required in order to know any new biological effects of a drug. A specific *in silico* visual tool is needed in order for researchers to ascertain the probable drug interactions prior to the experimental stage. Numerous computational models have thus been derived to search for optimal drug binding sites.⁹⁴⁻¹⁰⁷ A variety of different methods such as using binding site geometry, structures, physiochemical properties, fragment probes and combining the electrostatic potential, binding surface curvature with the amino acid residues have also been assessed for identifying desirable binding sites. Surface Cavity REcognition and EvaluationN (SCREEN) was developed and used to identify and characterise protein surface cavities.¹⁰⁸ SCREEN is a surface segmentation method attempting to define and characterise cavities on a set of 99 proteins co-crystallised with drugs. The most decisive factors found to identify a druggable binding site were the shape and size of these surface cavities.

Schrödinger's SiteMap is also a useful tool developed for identifying druggable sites on receptors.¹⁰⁹ A systematic approach is employed throughout the binding site identification process. It begins with a search on the protein surface to determine if one or more regions could be the potential ligand binding sites.¹¹⁰ Site points are linked throughout this stage to mark the sites that have the greatest potential to provide tight protein-ligand or protein-protein binding. Second stage involves forming maps showing regions of hydrophobicity or hydrophilicity. The hydrophilic map would be further divided into donor, acceptor and metal-binding regions. Altogether there are six types of maps available *via* colour-coded mesh systems to enable users to identify the desirable active or druggable site (**Figures 8**). The hydrophobic region would be represented by yellow and hydrophilic map would be displayed in green. The hydrogen-bonding donor region would be coloured violet blue and hydrogen-bonding acceptor region in red. Metal-binding map on the other hand would be in pink mesh form. The surface region was then shown in grey. The final stage involved evaluating all the energy property calculations at each site, which can then be inspected.

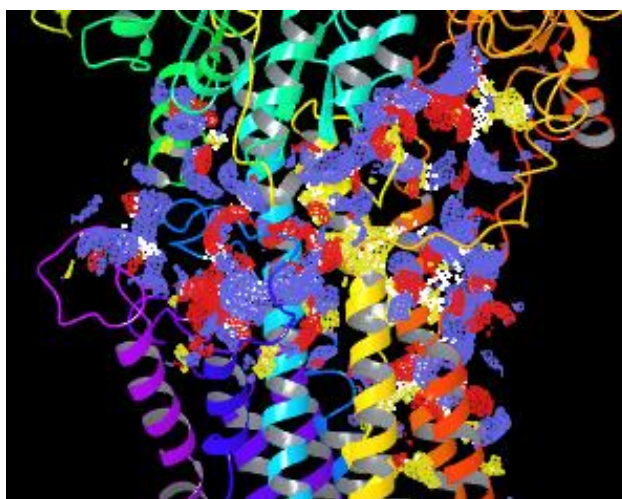


Figure 8 A close-up view of druggable sites (colour-coded mesh system) on SPCA1d-1SU4 model

SiteMap has also confirmed size and relative openness of the binding sites as significant factors in site identification when compared with SCREEN.¹¹¹ However, Sitemap has further developed a binding site predicting term, Site Score, to effectively distinguish sites that could bind ligands to sites that could not. Site Score was developed based on the number of site points involved in a binding site, degree of enclosure and hydrophilic scores for the site. The degree of site enclosure was measured based on the openness of the binding site. Typically a Site Score of 0.80 would aid in discriminating between a druggable and non-druggable site. Site Score of at least 1 and above indicated a reasonable chance for the site to be druggable. However, Site Score only aids in distinguishing a potential drug binding site and not showing the degree of druggability of the site. To address this, SiteMap has generated a similar scoring term, the Druggability Score (Dscore), to classify the druggability of the protein binding site to tightly bind passively absorbed small molecules. Dscore differs from Site Score where the hydrophilic score was not capped at 1 as in Site Score. This ensured Dscore would truly reflect a druggable binding site since it was found that difficult or undruggable binding site usually were more hydrophilic and less hydrophobic than the druggable sites.^{105,111}

1.4.3 Protein preparation

Protein preparation is a crucial step to ensure the protein crystal structures for the homology models were refined adequately and correctly to optimise docking results. A few examples of structure-based virtual screening in drug discovery have used protein preparation wizard with some promising results.¹¹²⁻¹¹⁴ Protein preparation can be done *via* the Protein Preparation Wizard¹¹⁵ or manually. It involved a series of steps to optimise hydrogen bonds, add any missing

hydrogen atoms and remove any steric clashes within the receptor model prior to the docking process.^{116,117} Other functions that Protein Preparation Wizard provided were to correct any metal ionisation state of the protein and remove water molecules if required.

1.4.4 Glide scores and molecular mechanics generalised born surface area (MM-GBSA) methods

Molecular dynamics have long been an interesting field for molecular modelling research involving biological macromolecules. Molecular dynamics research was separated into four different eras investigating the complex system of biological molecules *in silico*.¹¹⁸ The computer simulation in theory would largely reduce financial cost in conducting drug-discovery research and also to enable studies of the intricate biological systems. The first era began with the most primitive method of studying protein structure, where there were inadequate computational power to account for the surrounding environment of the biological macromolecules. This has evolved later to combine both free energy calculations from the second era and structure calculations from the third era into the final era plus what was already discovered from the first era to enable the understanding of macromolecular molecular dynamics.

This then lead to two major solvation models developed to assist our understanding of molecular dynamics¹¹⁹. Explicit solvent model initially focussed on treating solvent molecules at atomic levels which equally meant that it would use high computer power and cost to run modelling work. Then along came the Poisson-Boltzmann (PB) methods for implicit solvent model to calculate solvation free energies but it still required a significant amount of computer power.

Therefore a simpler implicit solvation model was developed and known as the generalised Born model which used less computational expense and reproduced similar promising results when compared to the PB methods ¹²⁰.

One of the studies comparing different methods for free energy prediction for protein-ligand complexes have shown that the molecular mechanics Poisson-Boltzmann surface area (MM-PBSA) and MM-GBSA methods had the best correlation between the estimated and experimental values ¹²¹. Both of these methods have different strengths and weaknesses. MM-PBSA was shown to be better at predicting absolute but not relative binding free energies than MM-GBSA ¹²². However, as mentioned before, due to MM-GBSA being more computationally efficient, it would be more useful in drug design.

Essentially, the Prime MM-GBSA module within the Schrödinger programme suite was developed to predict the binding affinities and strain energies of the ligands to the target protein.¹²³ The calculated values would not necessarily be the same as the experimental binding affinities but MM-GBSA would rank-order the ligands based on the calculated binding energies which is named MM-GBSA dG bind. The MM-GBSA dG bind is based on this equation from Schrödinger: $DG_{bind} = E_{complex}(minimised) - E_{ligand}(minimised) - E_{receptor}(minimised)$. The ligand strain energy can also be calculated from the difference between the calculated energy for the ligand extracted from the optimized complex (as if the ligand was optimised in the receptor binding pocket) and the calculated energy minimisation of the ligand outside the receptor. The ligand would be treated alone in solution for both calculations.

The Prime MM-GBSA module had originated from a surface-area-based version of the generalised Born (SGB) model.¹²⁴ The SGB model had empirical correction scheme trialled to correct the energy prediction errors in order to be comparable to the PB results. This SGB model was further refined to take into account the hydrophobic term and variable dielectric constant to improve protein loop and side chain predictions in the homology models.^{125,126} The newer version of Prime MM-GBSA 2011 and later versions added the ability to assign the protonation state of the receptor model which improved long protein loop prediction of up to 20 residues.¹²⁷ There have been numerous different studies using MM-GBSA to explore the energetics and structure of nucleic acid, protein-protein interactions where, for example, the binding free energies of filamin A subunits with bound peptides were estimated and in lead optimisation and discovery projects involving protein kinase inhibitors.¹²⁸⁻¹³¹

1.5 Aims of the project

The goal of this research project was to discover potential chemical compounds that could be further developed to become lead compounds to target SPCA1 and also SERCA pumps. The drug design process would need to be robust enough to ask the question; could a SERCA inhibitor be developed based on the drug design process involving molecular modelling, chemical synthesis and biological testing? If this first step was achieved then the next critical step was to design a SPCA1 inhibitor as SPCA1 was found to be highly involved in basal-like breast cancer. The potential lead compounds would then have the opportunity to become novel anti-cancer agents targeting basal-like breast cancer in this context. The ultimate aim was to

widen the current therapeutic agents available for patients with basal-like subtype of breast cancer in the hope to further improve their quality of life and life expectancy.

CHAPTER 2: METHODS

2.1 *in silico* molecular modelling

Structure-based drug design was the main technique adopted in this project. This technique involved modelling the target proteins, SERCA and SPCA, in their 3D forms and evaluating the binding affinities of potential ligand molecules to these protein targets. The Schrödinger suite of programs was used exclusively throughout the modelling process.^{123,132}

2.1.1 Software and software components used

All software and software components were run on the Maestro interface from Schrödinger unless otherwise stated. Comparative or homology modelling was carried out using Prime with two software components used, which were Protein Preparation Wizard and Structure Prediction Wizard. Ligand binding sites were found using SiteMap with Receptor Grid Generation used. Ligand preparation process used LigPrep. Ligand docking required Glide to carry out the docking process with details explained in section 2.1.4.2. Side-chain substitution on some of the ligands was trialled using CombiGlide with its embedded component BREED used. Details of the BREED component were further explained in section 3.1.4. The crystal structure of compound **16** was solved by using SHELXS refined by using SHELXL.¹³³

2.1.2 Comparative or homology modelling

2.1.2.1 SPCA

The human protein sequence of SPCA was found from National Centre for Biotechnology Information (NCBI) Protein and its FASTA sequence was used as the protein template for the SPCA homology models.^{123,134,135} The most suitable SERCA homolog was selected based on at least 30% similar identities between the homolog and SPCA protein sequence (**Table 1**). It was suggested that if the sequence identities were below 30%, the sequence alignment would fall into the twilight zone and become less accurate.¹³⁶ The crystal structure resolution was at least less than 3Å for the SERCA homolog to increase the clarity and accuracy for the SPCA model (**Table 1**). Two comparative models were built for SPCA1d and SPCA2 (**Appendix E**). For the SPCA2 PDB ID model, the SERCA homolog used was PDB ID 1IW0 (chain A). SPCA1d model used 1SU4 (chain A). Once the SPCA homology models were constructed (**Figures 9 and 10**), their 3D protein structures were prepared *via* Protein Preparation Wizard under Prime.¹¹⁵ The preparation process included assigning bond orders, adding hydrogen atoms, adjusting bond orders and formal charges, enhancing hydrogen bonds, deleting crystallographic waters, eliminating atomic clashes *via* protein minimisation with the OPLS_2005 force field.^{116,117}



Figure 9 Homology model of SPCA2-1IW0

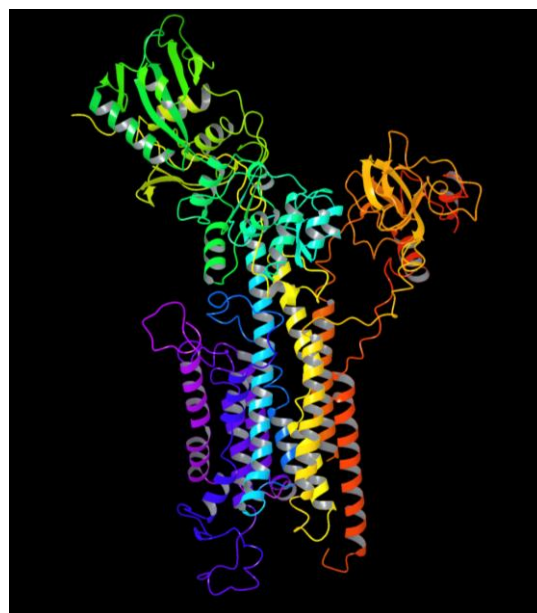


Figure 10 Homology model of SPCA1d-1SU4

2.1.2.2 SERCA

Similar procedures were undertaken to build the homology models for all three SERCA isoforms. Each of the PDB-derived protein crystal structures were selected from BLAST Homology Search under the Structure Prediction Wizard.¹²³ The SERCA homologues selected were based on the criteria mentioned previously for SPCA homology models (**Table 1**). Sequence alignments of each homology model were available in **Appendix E**.

Ca ²⁺ ATPase isoforms	PDB ID	Name	Identities	Positives	Source	Experiment	Resolution
SPCA1d	1SU4	Chain A of crystal structure of Ca ²⁺ -ATPase with two bound Ca ²⁺	32%	51%	<i>Oryctolagus cuniculus</i>	X-ray diffraction	2.4 Å
SPCA2	1IW0	Chain A of the SR Ca ²⁺ -ATPase in the absence of Ca ²⁺	32%	50%	<i>Oryctolagus cuniculus</i>	X-ray diffraction	3.1 Å
SERCA1	3AR5	Chain A of Ca ²⁺ pump crystal structure with bound TNP-AMP and TG	96%	98%	<i>Oryctolagus cuniculus</i>	X-ray diffraction	2.2 Å
SERCA 2	3AR4	Chain A of Ca ²⁺ pump crystal structure with bound ATP and TG in the absence of Ca ²⁺	84%	93%	<i>Oryctolagus cuniculus</i>	X-ray diffraction	2.15 Å
SERCA 3	3AR7	Chain A of Ca ²⁺ pump crystal structure with bound TNP-ATP and TG in the absence of Ca ²⁺	75%	88%	<i>Oryctolagus cuniculus</i>	X-ray diffraction	2.15 Å

Table 1 Details and statistics for all SPCA & SERCA homology models. **Identities %** corresponds to the percentage of residues that are identical between the sequences. **Positives %** is the percentage of residues that are positive matches according to the similarity matrix selected – default setting uses BLOSUM62 matrix.

2.1.3 Identifying ligand binding sites

SiteMap was used for identifying ligand binding sites for SPCA1d model under default settings. The only change made was 6 sites were chosen for the reporting of site-point groups by SiteMap instead of the default setting of 5 sites. This was done to increase the chance of finding more ligand binding sites, which were named as site-point groups in the SiteMap software component. Ligand binding sites could be set up *via* Receptor Grid Generation. All the grids generated in this project used the default setting. Grids were shown in **Figure 9** with both larger purple and smaller green 3D grids. Each grid represented the binding site for the ligand sets.

Additionally, other methods used to identify ligand binding sites on the receptor models was by compiling a Venn diagram of all the amino acid residues present within the ligand binding sites (**Figure 11**). Initially, the library of TimTec ligands¹³⁷ was docked on the SPCA2 model on the original TG binding site. The decision of using TimTec library was due to previous experience in the drug design area in our lab group. Then all the amino acid residues present within the ligand binding site were recorded for each docked ligand. The amino acid residues that have recurred most often for all the ligands docked were used to set up the trial grids. The numbered positions of at least three amino acid residues within the receptor were used to set up a receptor grid. Each specific position of the amino acid residue was represented by three digits after the shorthand of the amino acid name (**Figure 11**). The three amino acid residues were randomly chosen from the six most common amino acid residues from the Venn diagram to assist with the setting of ligand binding sites (grids).

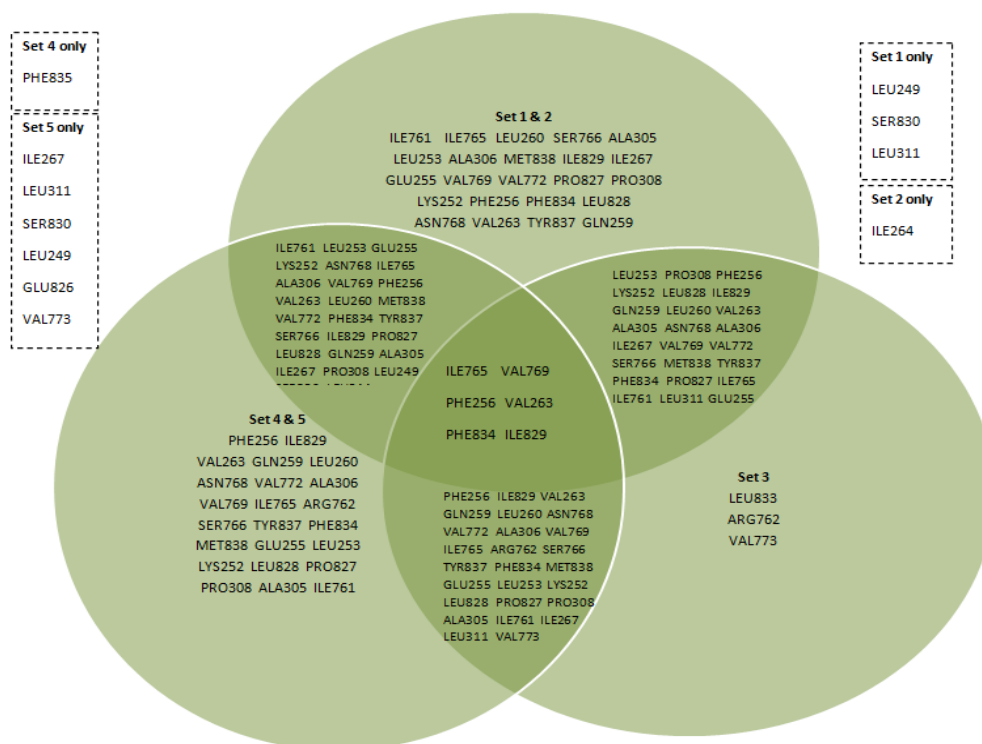


Figure 11 Venn diagram used for identifying common amino acid residues

2.1.4 Docking of ligands

The ligands used in this project were obtained from ZINC, a free database of 13 million commercially-available compounds, readily accessible by academic or pharmaceutical researchers involved in structural-based virtual screening research.¹³⁷ Ligands were downloaded as SDF files from the TimTec sub-database from ZINC, in reference pH of 7. The downloaded ligands were separated in sets of five with a total number of 271,965 used for high throughput virtual screening for both SERCA and SPCA models (set 1: 136,987 structures, set 2: 90,629 structures, set 3: 17,692 structures, set 4: 10,325 structures and set 5: 16,332 structures).

2.1.4.1 Ligand preparation

Schrödinger's LigPrep was used to ensure the desired ligand was in a low-energy state with correct stereochemistry for its structure.¹³⁸ The importance of stereochemistry in virtual screening was emphasised by Brooks *et. al.* where inclusion of all stereoisomers of all potential lead compounds was crucial to minimise any false negatives or to prevent losing a significant proportion of the potential leads.¹³⁹ LigPrep can also convert a 2D drawn structure to its 3D form which would be further processed *via* a series of steps to prepare the 3D ligand. The preparatory steps involved ensuring the ligands existed in appropriate ionisation states, tautomers, ring conformations, molecular weights and also the number and types of functional groups.

Numerous studies using structure-based virtual screening to search for potential lead compounds targeting specific diseases have used LigPrep.^{131,140-143} LigPrep was also used to prepare *in silico* enzyme inhibitors during a molecular docking study targeting a multifunctional kinase involved in a wide range of diseases.¹⁴⁴ Other studies have used LigPrep to find novel biosynthetic pathways for fuels such as 1-butanol, to build a 3D-quantitative structure-activity relationship model to predict the activity of desired candidate compound prior to synthesis and also for understanding the binding modes and activities of bromobenzimidazole casein kinase 2 inhibitors.¹⁴⁵⁻¹⁴⁷

Non-TimTec ligands used were prepared by using LigPrep prior to all docking protocols.¹³⁸ All other TimTec ligands were also prepared by LigPrep prior to the docking process.

2.1.4.2 Ligand docking on SPCA models

All TimTec ligands downloaded from ZINC were subjected directly to high throughput virtual screening (HTVS) docking initially, followed by standard precision (SP) and then extra precision (XP) docking.⁸⁹ These three docking modes were applied to all docking carried out in this project in this order. HTVS docking was the first docking mode used to reduce the number of intermediate conformations throughout the docking funnel. The SP mode was used to screen the ligands with reasonable HTVS Glide scores as chosen by the user. XP mode was designed for use on top-scoring ligand poses to screen false positives and active compounds that could bind to a specific conformation of the receptor. Each docking job of the selected ligands was processed and separated into a number of subjobs to be run by a limited number of central processing units (CPUs) in the computer, restricted by the number of Glide licences present.

To validate the method, ligand docking was first trialled on TG-binding site to the SPCA2-1IW0 model. This was done by using Glide *via* the Maestro interface.^{89-92,132} Prior to the ligand docking, TG was prepared by using LigPrep. Six prepared TG conformers generated from LigPrep were docked onto the SPCA2-1IW0 homology model. This step was done to assist in finding the six TG-binding sites which would then be trialled for subsequent ligand docking. The Glide docking process was then started with Set 1 of the TimTec ligands as mentioned before. In order to filter out the desirable ligands for the target receptor, the Glide docking process was carried out in the three docking modes in the order as mentioned previously.

After the docking process was completed for SPCA2-1IW0 model and the Venn diagram compiled, a tentative binding site was formed firstly for the SPCA1d-1SU4 model. The receptor grid, site 1, was set on LEU765, LEU256 and ILE834 which was originally based on the TG-binding site. All five sets of TimTec ligands were docked on this same site. Following on, four more receptor grids were set up where all five sets of TimTec ligands were docked in the three docking modes. Site 2 grid was set on ILE769, PHE262 and PHE854. Site 3 was based on VAL761, VAL274 and ILE828. For site 4, the grid was selected on LEU260, LEU765 and ILE769. Last grid, site 5, was generated on LEU260, LEU765 and ILE837. Once all of the five ligand binding sites were trialled for docking, SiteMap was used to generate better ligand binding sites. Altogether three SiteMap sites for SPCA1d-1SU4 were discovered, selected and produced *via* using Receptor Grid Generation. SiteMap site 4 was selected based on the location of ALA142, LEU141 and PHE236. SiteMap site 1 was on LEU319, GLY322 and LYS251. SiteMap site 3 was on ASN756, GLY808 and SER252.

2.1.4.3 Ligand docking on SERCA models

All three SERCA models were subjected to Glide docking. The same grid-setting protocol was followed with SERCA as that for the SPCA models for the ease of post-docking comparison. SiteMap site 4 for SERCA1-3AR5 was set exactly the same as that for SPCA1d-1SU4 model (ALA142, LYS141 and ARG236). SiteMap site 1 for SERCA1-3AR5 was also the same as that for the SPCA1d-1SU4 model (LEU319, GLY322 and GLN251). SiteMap site 3 for SERCA1-3AR5 was also the same (ASN756, GLY808 and LYS252). A TG-equivalent binding site was

also set up on ILE765, PHE256 and PHE834, according to the Venn diagram produced for SPCA2-1IW0 model. XP docking was then started for the same 26 ligands on these four set grids.

For SERCA2-3AR4 model, four equivalent receptor grids were set up as for the SERCA1-3AR5 model. SiteMap site 4 was on ALA142, LYS141 and ARG236. SiteMap site 1 had LEU319, GLY322 and GLN251. SiteMap site 3 was set on ASN756, GLY808 and LYS252. Lastly for the TG-equivalent binding site, this was generated on ILE765, PHE256 and PHE834. For SERCA3-3AR7 model, the SiteMap site 4 was placed on ALA142, ARG141 and ARG236. For SiteMap site 1, this was located on LEU319, GLY322 and ARG251 and for SiteMap site 3, this was on ASN756, GLY808 and LYS252. The TG-binding site was lastly set on ILE765, PHE256 and PHE834. Same XP docking for the same group of 26 ligands were done for these eight grids for both the SERCA2-3AR4 and SERCA3-3AR7 models.

2.1.5 CombiGlide trial

Side-chain substitutions were also trialled on the six compounds selected preliminarily *via* CombiGlide to see if better hit generation could be obtained prior to biological testing.¹⁴⁸ The side-chain substituents were taken from the Schrödinger website, where a total of 665 fragments were downloaded and separated into three combinatorial libraries to enable efficient docking processes. XP docking mode was used throughout the fragment docking for all six selected compounds. Interactive enumeration and docking function was used within CombiGlide. The point of side-chain substitution on the six selected compound structures was manually selected to indicate where the fragment would be attached (**Figure 12**).

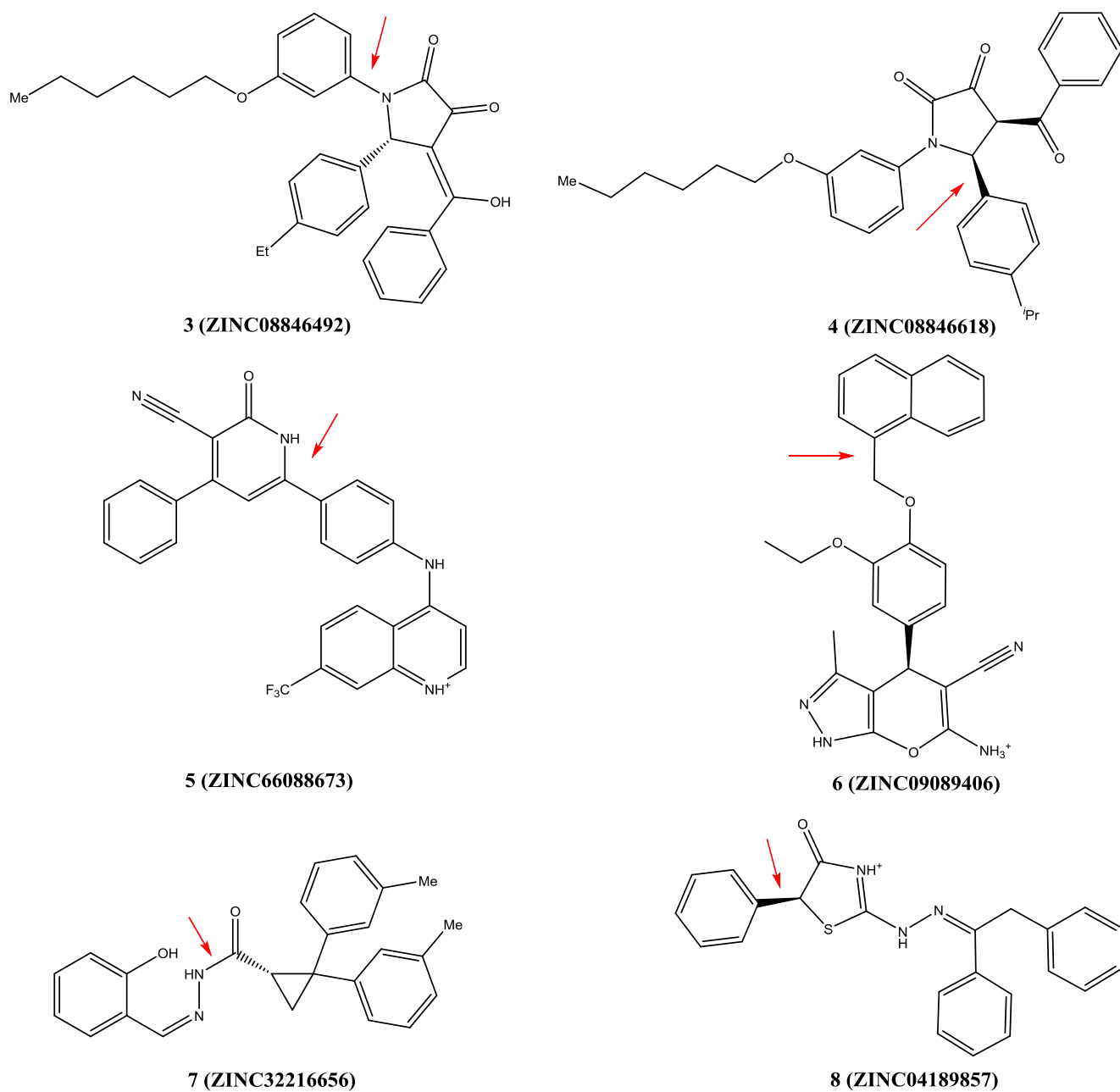


Figure 12 Six structures showing the site of side-chain substitutions. Red arrows indicate the broken bonds with new fragment attached. (ZINC08846492 and ZINC08846618 for SPCA1d-1SU4, ZINC09089406 and ZINC66088673 for SPCA2-1IW0, ZINC04189857 for SERCA1-3AR5 and SERCA2-3AR4 and ZINC32216656 for SERCA3-3AR7)

2.2 Compound availability

After completing all the docking and MM-GBSA analysis of the binding energies of the ligands, there were two options to obtain the target compounds. The compounds would either be synthesised in a chemistry laboratory or purchased from a suitable chemical supplying vendors if the stereochemistry of the target compounds were too complex to obtain a stable stereoisomer from laboratory synthesis. In the first round of structure-based virtual screening process, it was decided to purchase two compounds from TimTec due to complex stereochemistry of the selected compounds.

2.3 Bioassay – intracellular calcium measurement

2.3.1 MDA-MB-231 cell culture

The selected compounds were tested in MDA-MB-231 breast cancer cell line to observe the effect on the intracellular Ca^{2+} level. The cells were initially thawed after removal from liquid nitrogen store. The thawed cells were then transferred to a new falcon tube with 9mL of Dulbecco's Modified Eagle Medium (DMEM) added. The tube was then centrifuged to aid the removal of supernatant which would normally contain dimethyl sulfoxide (DMSO), the cryoprotectant agent. The cells would then be resuspended in 1mL of DMEM. The cell suspension was then transferred to a culture flask containing the DMEM and mixed thoroughly. The culture flask was stored in the incubator at 37°C with 5% carbon dioxide (CO_2) and observed for any contamination after 24 hours.

Cell feeding was carried out by replacing old media with fresh DMEM each time when the desired cell confluency was not reached. Cell splitting was carried out once every two to three days using trypsin after phosphate buffered saline (PBS)/ethylenediaminetetraacetic acid (EDTA) was added and separated into different flask for cell culture. The passage number and dilution factor were recorded after each cell splitting. Cells were plated when the cell growth was confluent with evenly-spread cells in the culture flask. Cell count was done to record the number of cells plated in each well. The final plate had about 15,000 cells in each well on a 96-well plate incubated at 37°C for 48 hours prior to running fluorometric imaging plate reader (FLIPR). Detailed cell culture methods were provided in **Appendix B**.

2.3.2 Fluorometric imaging plate reader (FLIPR)

The main instrument used for testing the Ca²⁺ signalling response in the MDA-MB-231 cell line was FLIPR^{®TETRA} High Throughput Cellular Screening System from Molecular Devices, LLC. It was intended to work as a high throughput optical screening tool for cell-based fluorescent assays.^{149,150} One of the major features of this instrument was the presence of simultaneous pipetting and reading capabilities to allow for rapid kinetic cellular assays. The number of pipette heads used in the experiment was 96 which could be increased to 382 or 1536 for other high throughput assays.

The machine of FLIPR consisted of three main parts (**Figure 13**). The optical detection of FLIPR used the cooled charge coupled device (CCD) technology, in which the CCD camera would take a picture of the bottom of the plate to record signals for each individual cell concurrently. The

signal would be derived from the emitted light coming from the fluorescent indicator dye, fluo-4 AM, which was excited by argon laser. During the running process, the FLIPR robot could aspirate the reagents from each of the two different fluid-addition plates and dispense them into the third one containing the cultured cells. The Ca^{2+} response outcomes within the cells could then be accessed from the software, FLIPR[®]TETRA system, operated on ScreenWorks[®] from Molecular Devices.

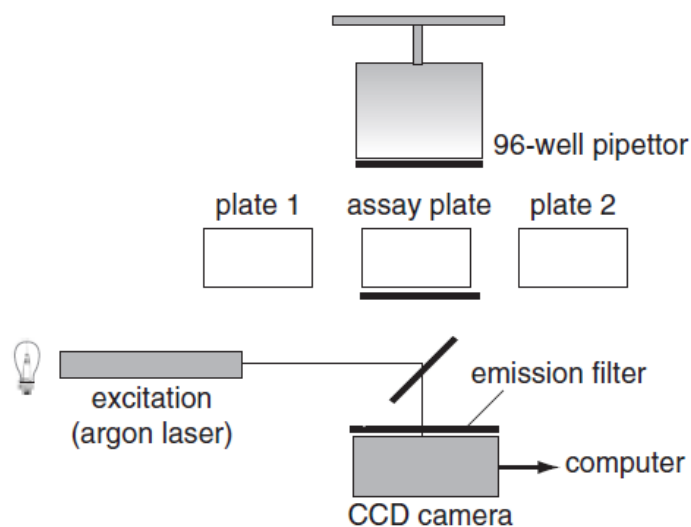


Figure 13 Main components of a FLIPR system (taken from ref. 139)

Prior to running FLIPR, two buffer solutions had to be made up. The first buffer solution was the 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) stock buffer that would maintain the pH of the solution. PSS buffer solution was the second buffer needed throughout the plating and preparation process for the similar purpose. The components within each buffer solution are given in **Appendix C**.

The FLIPR process was undertaken over two days (details in **Appendix C**). Day one involved plating cells in the 96-well plate in complete growth media. Day two involved preparing all three reagent plates and running the FLIPR assay. First reagent plate would have BAPTA added to each well. The second reagent plate would contain BAPTA, CPA, compounds 8 and 3 and DMSO. The third reagent plate would have Ca^{2+} , BAPTA, DMSO, both compounds 8 and 3 added to separate wells on the plate. FLIPR robot would then add the reagents from plate one, plate two then plate three to the cell plate in this order.

2.3.2.1 Dose-response measurement

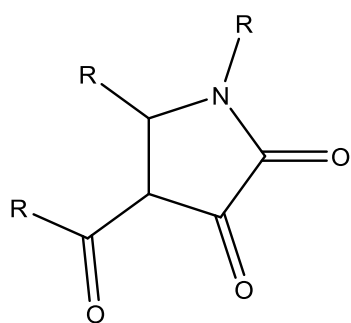
Dose-response measurement was carried out after the completion of FLIPR to ascertain the biological activity of the tested compounds. Dose-response graphs were constructed using XY table and dose vs. response graphs from GraphPad Prism version 5 for Windows, GraphPad Software, La Jolla California USA. Error bars shown on the dose-response curves were automatically generated by GraphPad Prism by using mean and error plot with standard error of the mean.

2.4 Chemical synthesis

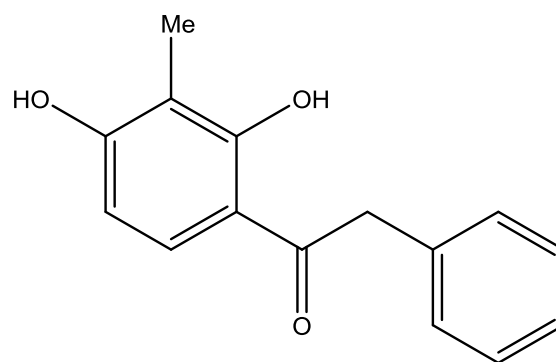
2.4.1 Molecular docking of compounds with pyrrolidone core

Further molecular docking were performed after the results from the first two selected compounds were finalised from the FLIPR assay. The aim of the docking was to search for compounds with a pyrrolidone core from SciFinder, ZINC databases and PubChem. This pyrrolidone core **9** was identified after comparing the structure of Compound **3** (ZINC08846492

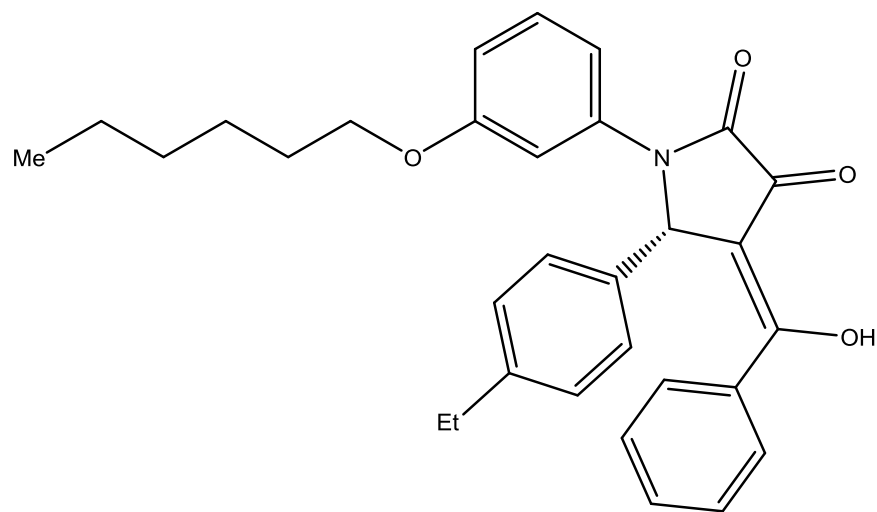
in **Figure 12**) with CPA **1** where both have similar structural motifs, incorporating a five-membered pyrrolidone ring incorporating two carbonyl groups and an acetyl group.¹⁵¹ Therefore, this particular core structure might confer some structural-activity effect towards the design of a SERCA inhibitor.



9

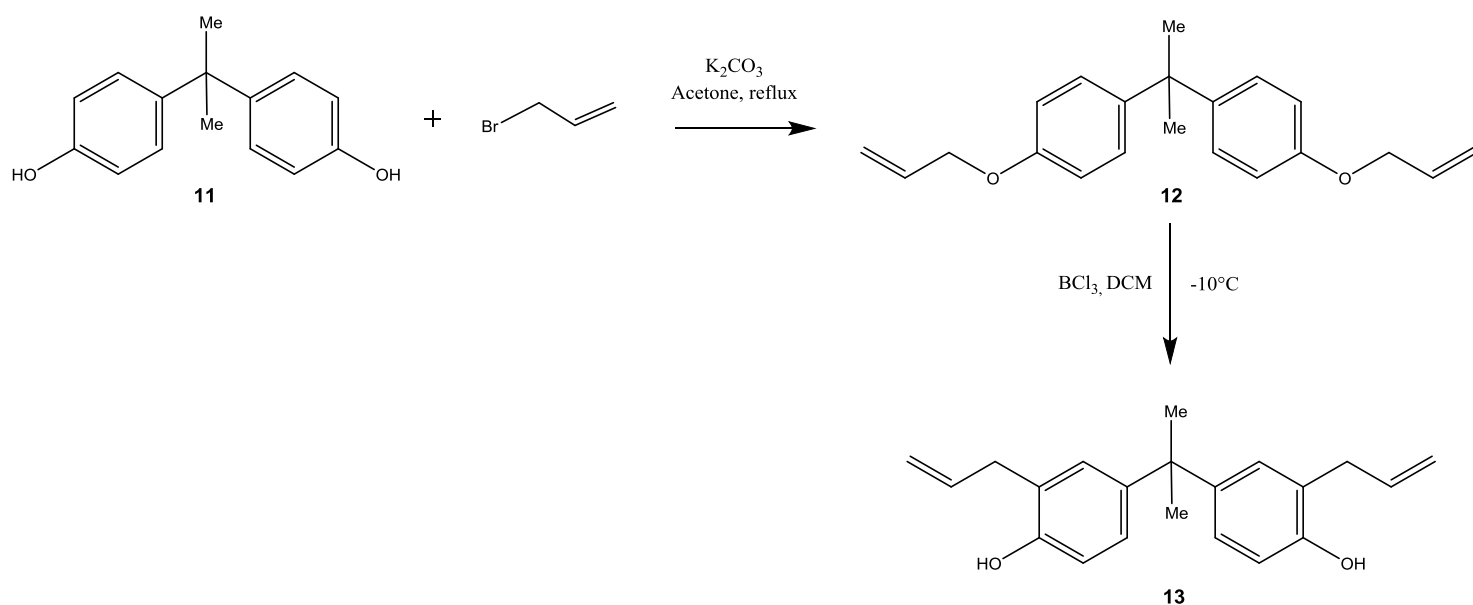


10



3 (ZINC08846492)

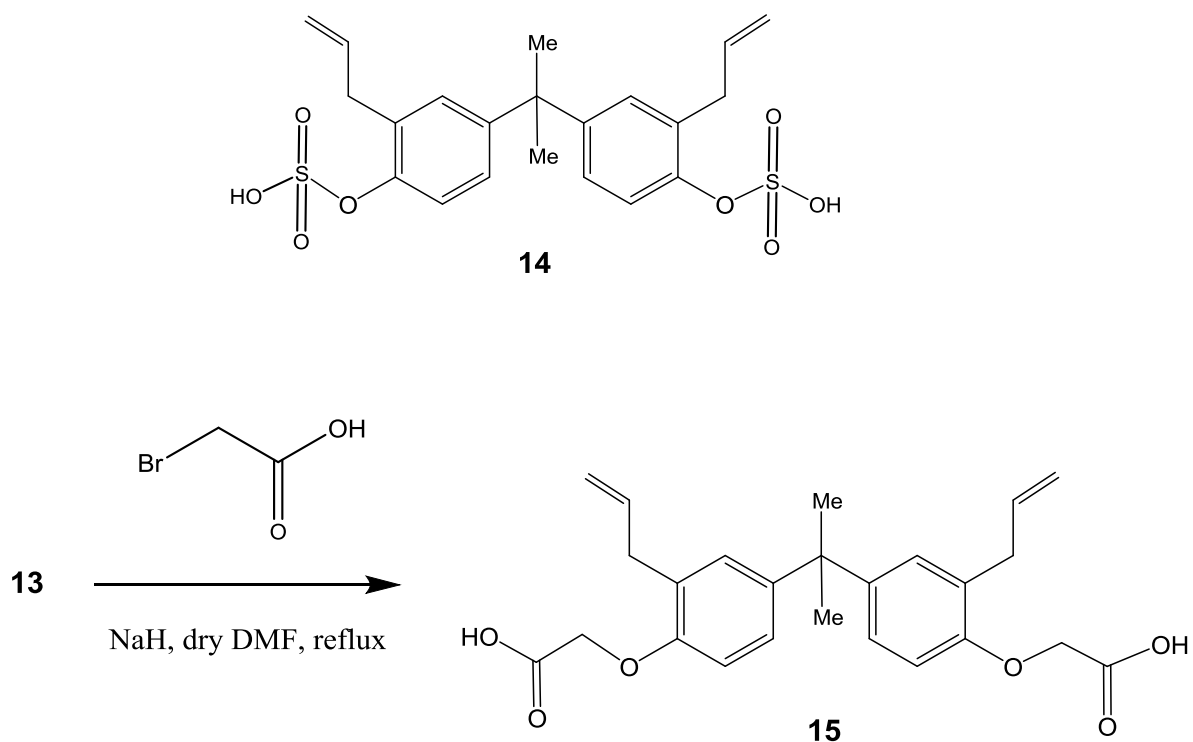
One compound, 1-(2,4-dihydroxy-3-methylphenyl)-2-phenylethanone **10** was synthesised within our lab group by J. Hanh/H. Gao following literature methods.¹⁵² Because of the possibility of keto-enol tautomerism within this structure, it was decided to add this compound to the biological assay as this keto-enol tautomerism theme was common to this structure, CPA and the desired pyrrolidone-like core structure, despite other differences in their ring structures. The structure of diarylethanone **10** was confirmed by both ¹H and ¹³C NMR spectroscopy, and the spectra correlated with the literature data.¹⁵²



Scheme 1

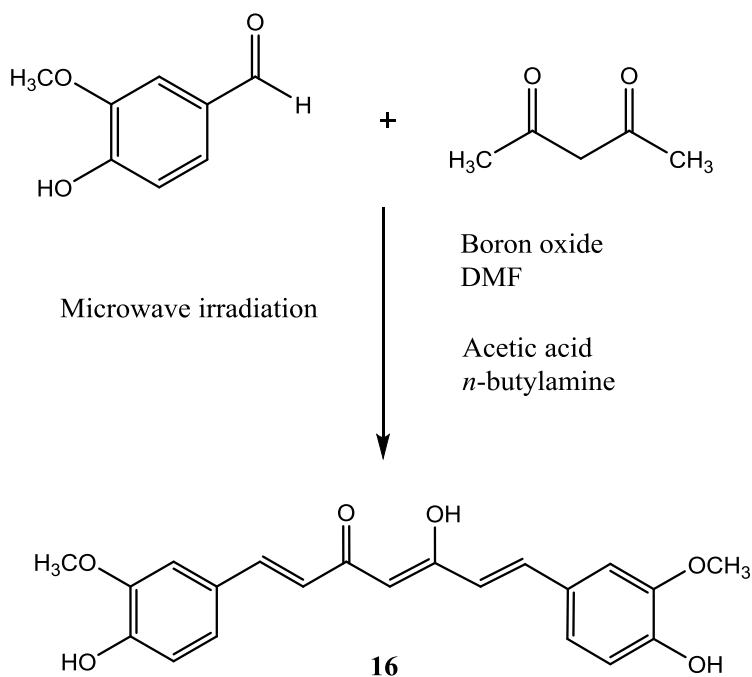
The diallyl bisphenol A **13**, Scheme 1, was considered when its structural activity relationships were investigated and found to increase the intracellular calcium levels in the human skin fibroblast cells, a phenomenon consistent to a SERCA inhibitor.¹⁵³ This compound was synthesised from bisphenol A **11** *via* its allyl ether **12** derivative using published methods.¹⁵⁴ However, due to the hydrophobic properties of this compound, as a result of the allyl groups

attached on both aromatic rings, it was not ideally soluble in DMSO or aqueous media for testing in FLIPR. It was therefore decided to synthesise more water-soluble derivatives of diallyl bisphenol A; diallyl bisphenol A sulphate **14** and acetic acid **15** derivatives. Preparation of **14** applying a modified literature method ¹⁵⁵ was not successful. Synthesis of 3,3'-diallyl-*O,O'*-bis(carboxymethyl)bisphenol A **15** was carried out using bromoacetic acid, scheme 2.



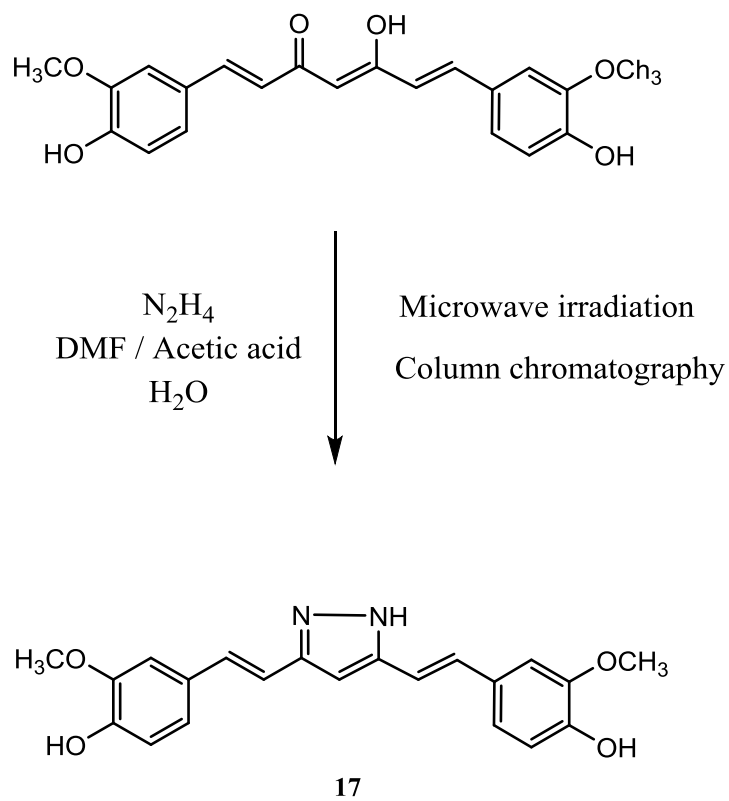
Scheme 2

Based on the known anticancer effect of curcumin and also the link between SPCA and basal-like breast cancer,¹⁵⁶ curcumin **16** was synthesised using the established procedure¹⁵⁷ reacting 2 equivalent of 4-hydroxy-3-methoxy-benzaldehyde with 1 equimolar of acetylacetone in the presence of boron oxide, as detailed in the Experimental Section, **Scheme 3**. The crystals were analysed and confirmed *via* X-ray diffraction and also ¹H and ¹³C NMR which had correlated with literature data.¹⁵⁸



Scheme 3

As pyrazole derivative **17** showed greater inhibitory potential on calcium/calmodulin dependent protein kinase II, it was considered as a suitable candidate to be subjected to biological testing. Heterocyclic analogue **17** was prepared by the published microwave assisted method¹⁵⁷ using the synthesised curcumin **16** and hydrazine, **Scheme 4**.



Scheme 4

2.4.2 Experimental section

2.4.2.1 Preparation of 1-(2,4-dihydroxy-3-methylphenyl)-2-phenylethanone 10

1-(2,4-Dihydroxy-3-methylphenyl)-2-phenylethanone **10** was prepared by a literature procedure¹⁵²; (3.3 g, 54 %); ¹H NMR (400 MHz, CDCl₃) δ_H 2.11 (3H, s, CH₃), 4.20 (2H, s, CH₂), 5.59 (1H, broad, OH), 6.34 (1H, d, *J* = 8.8 Hz, H-5'), 7.23-7.26 (3H, m, Ar-H), 7.30-7.34 (2H, m, Ar-H), 7.61 (1H, d, *J* = 8.8 Hz, H-6'), 12.96 (1H, s, OH); ¹³C NMR (75 MHz, CDCl₃) δ_C 7.2 (CH₃), 44.7 (CH₂), 106.9 (C_{Ar}), 111.2 (C_{Ar}), 112.9 (C_{Ar}), 126.9 (C_{Ar}), 128.6 (C_{Ar}), 129.2 (C_{Ar}), 129.4 (C_{Ar}), 134.4 (C_{Ar}), 160.2 (C_{Ar}), 163.4 (C_{Ar}), 202.1 (C=O, quat.).

2.4.2.2 Preparation of diallyl bisphenol A 13

2.4.2.2.1 Preparation of diallyl ether of bisphenol A 12¹⁵⁴

Bisphenol A **11** (2.968 g, 13 mmol), allyl bromide (4.718 g, 38.99 mmol), and potassium carbonate (5.5 g) were dissolved in acetone (30 mL) and refluxed overnight, under nitrogen. The reaction was monitored by thin layer chromatography (TLC). The solution was then cooled, filtered and the solvent evaporated under reduced pressure. The residue was dissolved in DCM (50 mL) and washed with dilute aq. sodium carbonate solution (2 × 25 mL) followed by water (25 mL). The remaining organic layer was dried over anhydrous sodium sulphate, filtered and evaporated under reduced pressure to give bisphenol A diallyl ether **12** a yellow-brown oil (2.49 g, 84 %); ¹H NMR (400 MHz, CDCl₃) δ_H 1.62 (6H, s, 2 × CH₃), 4.48 (4H, d, *J* = 6.4 Hz, 2 × CH₂), 5.24-5.27 (4H, m, 2 × CH₂=), 5.37 (1H, s, OH), 5.99-6.07 (2H, m, 2 × CH=), 6.80 (4H, m, Ar-H, H₃, H₅), 7.11 (4H, m, Ar-H, H₂, H₆).

2.4.2.2.2 Rearrangement of ether 12 to 3,3'-diallyl bisphenol A 13¹⁵⁴

The diallyl ether of bisphenol A **12** (2.49 g, 8.07 mmol) was dissolved in DCM (24 mL). The reaction was carried out on an ice bath containing a mixture of ice, acetone and sodium chloride (-10 °C). Boron trichloride (12 mL; 1M solution in DCM; 12 mmol) was carefully added and stirring was continued for 4 hours at -10 °C. The reaction was monitored by TLC and after it was complete, water (30 mL) was added and the mixture was stirred for a further 15 minutes. The DCM layer was then separated and washed with water (3 × 50 mL) then dried over anhydrous sodium sulphate, filtered and evaporated under reduced pressure to give 3,3'-diallylbisphenol A **13** as a brown oily liquid (1.87 g, 6.06 mmol, 75 %); ¹H NMR (400 MHz, CDCl₃) δ_H 1.63 (6H, s,

2 × CH₃), 3.38 (4H, d, *J* = 6.4 Hz, 2 × CH₂), 5.12-5.17 (4H, m, 2 × CH₂=), 5.30 (1H, s, OH), 5.96-6.06 (2H, m, 2 × CH=), 6.71 (2H, d, *J* = 8.8 Hz, H-5,5'), 6.96-6.99 (4H, m, Ar-H).

2.4.2.3 Preparation of 3,3'-diallyl-*o,o'*-bis(carboxymethyl)bisphenol A **15**¹⁵⁹

3,3'-Diallylbisphenol A **13** (1.87 g, 6.07 mmol) in THF (100 mL) was added to sodium hydride (0.87 g; 60% dispersion in mineral oil, 36.40 mmol) in dry DMF (25 mL) and the solution was stirred for 30 minutes at ambient temperature. Bromoacetic acid (2.01 g, 14.60 mmol) was added and the mixture was heated at 60°C under reflux, overnight. The solution was allowed to cool then 1M aq. HCl was added dropwise until a pH of 2 was obtained. The aqueous layer was extracted with ethyl acetate (3 × 40 mL). The collected organic layers were dried over Na₂SO₄ and the solvent was evaporated yielding a brown oil which was then triturated with petroleum ether to give **15** as a whitish-yellow solid (13.168 g, 100 %, 146 - 148°C); ¹H NMR (400 MHz, DMSO-*d*₆) δ_H 1.51 (6H, s, 2 × CH₃), 3.29 (4H, d, *J* = 6.4 Hz, 2 × CH₂), 4.60 (4H, s, 2 × OCH₂COOH), 4.94 (4H, m, 2 × CH₂=), 5.89 (2H, m, 2 × CH=), 6.69 (2H, d, *J* = 9.2 Hz, H-5,5'), 6.92-6.94 (4H, m, 4 × Ar-H), 12.91 (2H, m, 2 × COOH). ¹³C NMR (75 MHz, DMSO-*d*₆) δ_C 31.2 (CH₃), 34.5 (CH₂-CH=), 41.6 (CH₃), 65.2 (O-CH₂COOH), 111.4 (CH=CH₂), 115.8 (CH=CH₂), 125.5 (C_{Ar}-5 or 6), 127.6 (C_{Ar}-2), 128.3 (C_{Ar}-5 or 6), 137.4 (C_{Ar}-3), 143.2 (C_{Ar}-1), 153.6 (C_{Ar}-4), 170.8 (O-CH₂COOH); IR (Nujol) ν_{max}/cm⁻¹ 1746 (COOH), 1246 (C-O), 913 (C-H_{Ar}); HRMS (ESI) Found: MNa⁺, 447.1779. Calc. for C₂₅H₂₈O₆Na: MNa⁺, 447.1784.

2.4.2.4 Preparation of (1*E*,4*Z*,6*E*)-5-hydroxy-1,7-bis(4-hydroxy-3-methoxyphenyl)hepta-1,4,6-trien-3-one **16**¹⁵⁷

Vanillin (304.3 g), acetyl acetone (100.13 g) and boron oxide (200 mg) was dissolved in DMF (1 mL). The reaction mixture was subjected to microwave irradiation at 55-60°C for 5 minutes. Acetic acid (60 μ L) and n-butyl amine (50 μ L) was added to the reaction vessel followed by repeated microwave irradiation for another 5 minutes. The resulting mixture was poured on 20% acetic acid in water (50 mL). The precipitate formed was filtered and recrystallised from ethanol to give **16** as red crystals (0.68 g, 100 %, 178 - 180°C) ¹H NMR (400 MHz, DMSO-*d*₆) δ _H 3.80 (6H, s, 2 \times OCH₃), 6.02 (1H, s, CH=), 6.72 (2H, d, *J* = 16.0 Hz, 2 \times CH=), 6.78 (2H, d, *J* = 8.0 Hz, CH_{Ar}, H-5), 7.11 (2H, dd, *J* = 8.0 and 1.6 Hz, 2 \times CH_{Ar}, H-6), 7.28 (2H, d, *J* = 1.6 Hz, 2 \times CH_{Ar}, H-2), 7.51 (2H, d, *J* = 15.6 Hz, CH=); ¹³C NMR (75 MHz, DMSO-*d*₆) δ _C 56.5 (CH₃), 101.3, 111.7, 116.1, 121.5, 123.6, 126.7, 141.1, 148.4, 149.8, 183.6 (C=O); IR (Nujol) ν _{max}/cm⁻¹ 1620 (C=O), 1600 – 1550 (C=C), 1283-1153 (C-O), 1026(C-H_{Ar}).

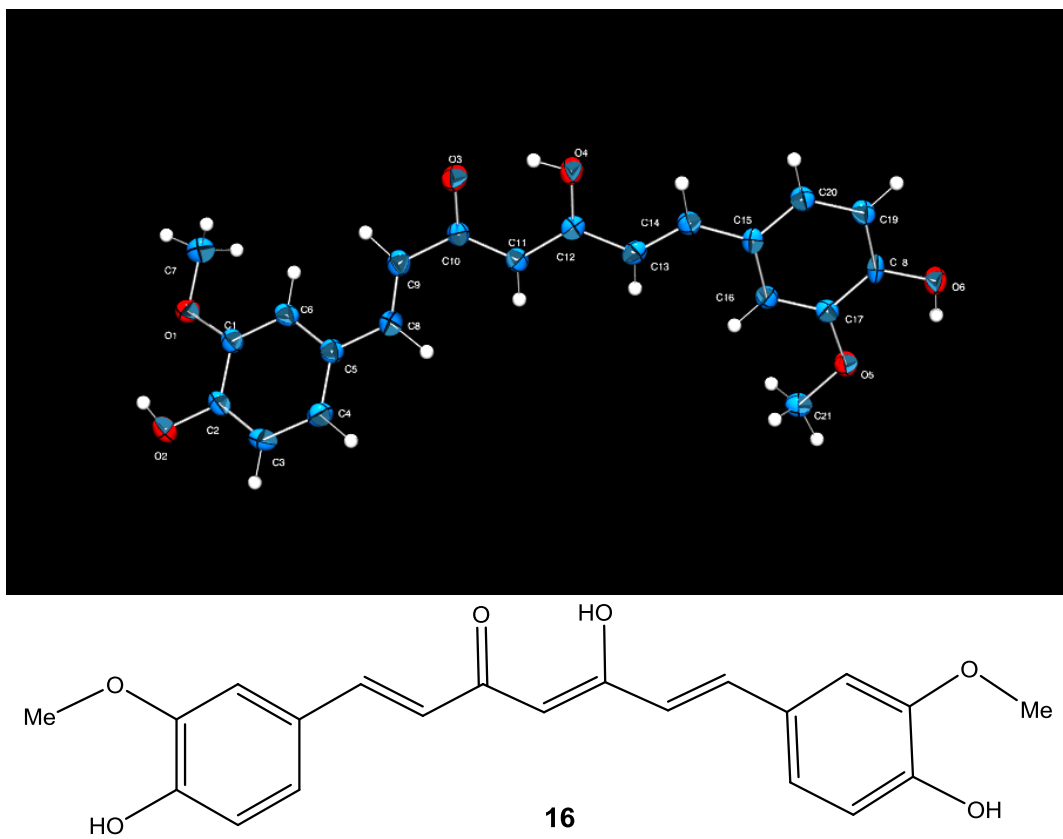


Figure 14 An X-ray refinement and Oak Ridge Thermal Ellipsoid Plot (ORTEP¹⁶⁰) drawing for compound **16** (ellipsoids are at the 30% probability level)

2.4.2.4.1 X-ray data collection and reduction

The single-crystal X-ray diffraction experiments were carried out at the University of Sydney using a Bruker APEX2 CCD-based diffractometer with an X-ray wavelength of 0.7107 Å (Mo K α) and at an experimental temperature of 150 K. The red single crystal of **16** (0.25 × 0.15 × 0.15 mm) was mounted on the tip of a thin glass fibre with a minimum amount of Paratone N oil and inserted in the cold N₂ stream of an Oxford Cryosystem COBRA instrument. X-ray diffraction data were collected using 0.3° $\Delta\omega$ scans, maintaining the crystal-to-detector distance

at 6.01 cm. The diffraction data were integrated using SAINT+,¹⁶¹ and the unit cell parameters for **16** at 150 K were refined from 500 reflections in the monoclinic space group $P2_1/n$ with $Z = 4$, $F(000) = 776$, and $\mu = 0.101 \text{ mm}^{-1}$ (**Table 2**).

Table 2 Independent Atom Model (IAM) Refinement of compound 16

Empirical formula	$C_{21}H_{20}O_6$
Molecular mass	368.37
Crystal size/ mm^3	$0.25 \times 0.15 \times 0.15$
Temperature/K	150
Crystal system	monoclinic
Space group	$P2_1/n$
a (Å)	12.614
b (Å)	7.086
c (Å)	20.009
α (°)	90.00
β (°)	94.83
γ (°)	90.00
Volume (Å ³)	1782.1
Z	4
$\rho_{\text{calc}}/\text{mg}/\text{mm}^3$	1.373
μ/mm^{-1}	0.101
F(000)	776.0
Radiation	MoK α ($\lambda = 0.71073$)
2 Θ range for data collection	3.68 to 42.5°
Index ranges	$-12 \leq h \leq 12, -7 \leq k \leq 6, -20 \leq l \leq 20$
Reflections collected	7628
Independent reflections	1974[R(int) = 0.0646]
Data/restraints/parameters	1974/0/257
Goodness-of-fit on F^2	1.009
Final R indexes [$I \geq 2\sigma(I)$]	$R_1 = 0.0420, wR_2 = 0.0931$
Final R indexes [all data]	$R_1 = 0.0816, wR_2 = 0.1112$
Largest diff. peak/hole / e Å ⁻³	0.21/-0.18

The structure of **16** was solved using direct method SHELXL-97,¹³³ and full-matrix least-squares refinement on F^2 was carried out using SHELXL-97. All non-hydrogen atoms were refined anisotropically. Hydrogen atoms were included in idealised position, with U_{150} tied to 1.5 times that of the parent atom. An ORTEP drawing of compound **16** could be seen in **Figure 14**. Given the low values of R_1 and wR_2 , we could conclude that this was a structure of reasonable quality.

2.4.2.5 Preparation of 3,5-bis{2'-(4-hydroxy-3-methoxyphenyl)-(1'E)-ethen-1'-yl}1H-pyrazole **17**¹⁵⁷

Curcumin **16** (0.3 g), acetic acid (1 mL) and hydrazine hydrate (80 μ L) was dissolved in DMF (1 mL) and subjected to microwave irradiation at 55-65 °C for 5 minutes. After addition of hydrazine hydrate (80 μ L) the irradiation cycle was repeated. The resulting mixture was poured into water (40 mL) and the solid formed upon standing was filtered to give crude product **17**, which was then purified by gradient column chromatography (ethyl acetate-hexane 1-1 to 9-1) to yield an off-yellow solid **17** (0.16 g, 53 %, 217 - 219°C); ¹H NMR (400 MHz, DMSO-*d*₆) δ_H 3.79 (6H, s, 2 \times CH₃), 6.59 (1H, s, CH, H-4₁), 6.74 (2H, d, J = 8.0 Hz, 2 \times CH_{Ar}, H-5), 6.87-6.91 (4H, m, 2 \times CH_{Ar} and 2 \times CH=), 7.01 (2H, d, J = 16.8 Hz, 2 \times CH=), 7.11 (2H, s, 2 \times CH_{Ar}, H-2), 9.14 (1H, broad, OH), 12.8 (1H, broad, OH); ¹³C NMR (75 MHz, acetone-*d*₆) δ_C 55.3 (CH₃), 99.1 (C=C_{pyrazole}), 108.9 (C_{Ar}-2), 115.1 (C_{Ar}-5), 120.3 (C=C), 129.2 (C=C), 129.8 (C_{Ar}-6), 146.8 (C_{Ar}-3 or 4 or C=N), 147.7 (C_{Ar}-3 or 4 or C=N); IR (Nujol) $\nu_{\max}/\text{cm}^{-1}$ 3478(OH_{phenol}), 3314(N-H), 1594(C=N), 1280(C-O); HRMS (ESI) Found: MH⁺, 365.1000. Calc. for C₂₁H₂₁N₂O₆: MH⁺, 365.1501.

CHAPTER 3: RESULTS AND DISCUSSION

3.1 Molecular modelling results

3.1.1 SPCA

3.1.1.1 SPCA2

As mentioned earlier in section 2.1.4, there were a total of five sets of TimTec ligands downloaded for docking. Set 1 contained 136,987 TimTec ligand structures, set 2 had 90,629 TimTec structures, set 3 had 17,692 structures, set 4 had 10,325 structures and lastly set 5 had 16,332 structures. HTVS docking mode was used as the first docking step and was repeated three times in total for set 1 and 2 TimTec ligands. Each time the subjob number was reduced for docking to see if this would increase the number of ligands with G scores of -7 or below. TimTec ligands from set 3, 4 and 5 were repeated twice in the three docking modes for the same purpose of increasing the number of hits for virtual screening. The binding site for all the TimTec ligands from all dockings methods was based on the TG binding site.

Altogether there were two important criteria used in this project for choosing compounds with the highest potential for biological activity. G scores were one of the factors and the other component was the results from MM-GBSA. G scores were a useful indicator of how well the compounds have performed in the *in silico* docking based on a rank-scoring system as mentioned previously. According to Schrödinger, G scores around -8 and -9 kcal/mol were considered to be very good for shallow binding sites or sites with mainly hydrophobic interactions.⁸⁹ For Glide

XP mode docking, G scores were usually -12 kcal/mol or below. For Glide HTVS or SP mode docking, G scores of -10 kcal/mol or lower usually indicated good binding of the ligand.

Docking for set 1 ligands was initially run by 16 subjobs in the HTVS mode. The result from this gave 49 ligands with G scores of at least -8 or below. For the second round, 8 subjobs were used to run the HTVS mode docking for the same set 1 ligands at the same site and this provided 66 ligands with G scores of at least -8 or below. For the third round of HTVS docking, 6 subjobs were used and this produced 140 ligands with G scores of at least -8 or below. The number of subjobs was reduced from 16 to 8 then lastly to 6 for each round. This applied to all the subsequent docking mentioned hereafter unless otherwise stated. The HTVS docking results for set 1, 2, 3, 4 and 5 ligands were shown in **Table 3**.

	Number of Tim Tec ligands with G scores of at least -8 or less				
HTVS docking mode	Set 1	Set 2	Set 3	Set 4	Set 5
16 subjobs used	49	54	N/A	N/A	N/A
8 subjobs used	66	74	30	3	38
6 subjobs used	140	100	56	4	77

Table 3 HTVS docking mode results for Tim Tec ligands on SPCA2 model

The set 1 ligand SP docking has produced a total of 131 ligands with G scores of at least -8 or below by using 8 subjobs. For set 2 ligands, this has produced 219 ligands of G scores of about -8 or below. For set 3, the top 300 ligands with G scores of about -8 or below were chosen to run the XP mode docking. This has led to a total of 19 ligands with top G scores of at least -10 to about -11.6. Set 4 SP docking has given 100 ligands with G scores from about -7.7 to -8.9. These ligands were subjected to XP docking and produced two ligands with G scores of about -10.3. From SP docking for set 5 ligands, this has generated a total of 400 ligands with G scores of about -8 to -10.6. XP docking was then performed for these ligands and found a total of 11 ligands with G scores of -10 to -11.5. All the ligands generated from XP mode dockings from set 1 to 5 were saved in separate SDF files and 2-dimensional (2D) viewer PDF files.

The 131 ligands from set 1 SP docking has been trialled for XP docking but only generated two ligands with G scores of about -10.7 to -10.8. Therefore, the SP docking was repeated and the number of subjobs used to run the SP docking was reduced from 8 to 6 to see if this would provide better result. This gave altogether 100 ligands with G scores of about -8 to -9.8 from the SP docking. These ligands were then subjected to the XP docking, which produced 5 ligands with G scores of about -10 to -11. The SP docking was also repeated for set 2 ligands with the subjob number reduced from 8 to 6. The 100 ligands from the HTVS docking for set 2 were put through the SP docking mode. The result of this has provided 90 ligands with G scores of about -7.4 to -9.3. These ligands were sent through to XP docking and finally produced 3 ligands with G scores of about -10.2 for two of the ligands and the last one had about -10.9.

From the docking results of the Tim Tec ligands (set 1 to 5) on SPCA2-1IW0 model, there were altogether 40 ligands that had XP G scores of around -10. A Venn diagram was compiled to show all the amino acid residues involved at the ligand binding sites from the dockings completed previously. Six amino acid residues, ILE765, VAL769, PHE256, VAL263, PHE834 and ILE829, were discovered to be the most common amino acid residues for all 40 ligands. The grid was re-set by using three of the six common amino acid residues, ILE765, PHE256 and PHE834, to be the ligand binding site. The XP docking for these 40 ligands was repeated again to evaluate and confirm the previous G scores. This has given 26 out of the 40 ligands from the repeated XP docking with G scores of at least -10 or below.

3.1.1.2 SPCA1d

At the time of writing, there were no published studies investigating into the specific SPCA1 isoforms present in basal-like breast cancer cells. Therefore, SPCA1d protein sequence was adopted due to its commonly expressed nature. After the SPCA1d-1SU4 homology model was prepared for docking, the ligand binding site was set via forming grid on LEU765, LEU256 and ILE834. The same process of using the three docking modes was followed for this homology model as the SPCA2 model. For Tim Tec set 1 ligands, HTVS docking has identified top 500 ligands with G scores of about -6.4 and below. These 500 ligands were selected to be docked in SP mode by using 6 subjobs. This has generated 300 ligands with G scores of about -6.4 and below and subsequently put through for XP docking. However this has not produced any ligands with G scores of at least -10 or below. Similar XP docking results were also found for Tim Tec ligand set 2, 3, 4 and 5.

In fact, the XP docking for Tim Tec set 1 ligands have only produced 12 ligands with G scores between -8.0 to -9.1. For set 2 ligands, only 7 ligands had G scores between -8.1 to -8.7. For set 3, there were 9 ligands with G scores of about -8.1 to -9.0. Only one ligand with G score of -8.2 was present for set 4 and lastly for set 5, none of the ligands from XP docking had G scores above -8. Overall, the average highest G scores of -8 to -9 were found for all five set of Tim Tec ligands only.

Because of the above result, the receptor grid was re-set again to ILE769, PHE262 and PHE854 and named site 2, which was close to the grid for SPCA2-1IW0 model. All three docking modes were trialled at site 2 for all five sets of Tim Tec ligands, final results did not identify any ligands with XP G scores of -10 or below. Another receptor grid, site 3, was trialled again to see if this would make any difference. This grid was set to be on VAL761, VAL274 and ILE828 due to being closest to VAL769, VAL263 and ILE829 found from the SPCA2-1IW0 Venn diagram. Site 3 dockings had the same outcome as for site 2 docking; nil ligands had any XP G score of at least -10 or below. Site 4 was then generated on LEU260, LEU765 and ILE769 to trial all five set of ligands for docking. Similar outcome was reached except one ligand from set 3 had a XP G score of -10.2. Lastly, one more other grid was set up on LEU260, LEU765 and ILE837 to trial one more run of the three docking modes and the result for set 1 ligands was that none of the ligands had XP G scores of -10 or below.

The solution to the problem of not obtaining the expected desirable XP G scores, such as -10 or below, was to run SiteMap to identify any druggable sites on SPCA1d-1SU4 model. First site generated from SiteMap for SPCA1d-1SU4 based on residue ALA142, LEU141 and PHE236 (SiteMap site 4) was subjected to XP docking with the previous 26 Tim Tec ligands derived from the SPCA2-1IW0 model. The G score range obtained from this docking ranged from -3.367 to -6.750. The second site from SiteMap for the same homology model was set on residues LEU319, GLY322 and LYS251 (SiteMap site 1). This only brought G scores from -4.161 to -7.688 for the same set of 26 ligands. Lastly, SiteMap site 3 was selected based on ASN756, GLY808 and SER252 and XP docking trialled on these same set of ligands. The G score ranged from -4.235 to -7.973.

SPCA2-1IW0 model was also trialled with these three SiteMap-derived grids for SPCA1d-1SU4 model. SiteMap site 4 was based on ALA142, LYS141 and ARG236 (only difference was LYS and ARG being the different amino acid residues with exactly same residue positions as that for SPCA1d-1SU4 model). For SiteMap site 1, the grid was set on LEU319, GLY322 and GLN251 (exactly same residue positions but differs in amino acid residue GLN). SiteMap site 3 had ASN756, GLY808 and LYS252 (same residue positions with only residue LYS different). The same set of 26 ligands was trialled for XP docking on these three grids with G scores ranged from 3.459 to -3.959 for SiteMap site 4; -2.920 to -8.028 for SiteMap site 1; -0.641 to -11.180 for SiteMap site 3.

3.1.2 SERCA

3.1.2.1 SERCA1

Similar receptor grids were set to all the SERCA homology models *via* using SiteMap.¹⁰⁹ SiteMap site 4 was set on ALA142, LYS141 and ARG236. The G scores obtained after XP docking of the same set of 26 ligands ranged from 0.718 to -4.108. The same XP docking of the 26 ligands were then applied to the following three receptor binding sites. SiteMap site 1 was on LEU319, GLY322 and GLN251. The G scores ranged from -2.993 to -8.877. SiteMap site 3 was based on ASN756, GLY808 and LYS252 with G scores ranged from -3.584 to -11.423. Lastly, for the original TG-equivalent binding site on ILE765, PHE256 and PHE834, the G scores ranged from -8.812 to -11.745.

3.1.2.2 SERCA2

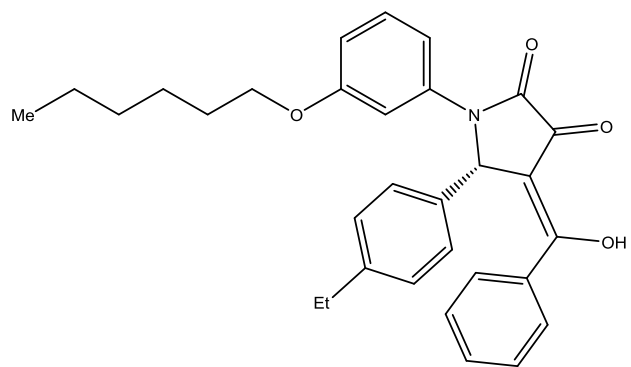
For SERCA2 homology model, the same four receptor grids were set up for XP docking of the same 26 ligands as before. For SiteMap site 4 on ALA142, LYS141 and ARG236, the G scores ranged from 3.624 to -5.082. For SiteMap site 1 on LEU319, GLY322 and GLN251, the G scores were found to range from -2.661 to -6.676. From SiteMap site 3 on ASN756, GLY808 and LYS252, the G scores ranged from -2.976 to -11.900. For the TG-equivalent binding site on ILE765, PHE256 and PHE834, the G scores were found to be between -8.225 to -11.659.

3.1.2.3 SERCA3

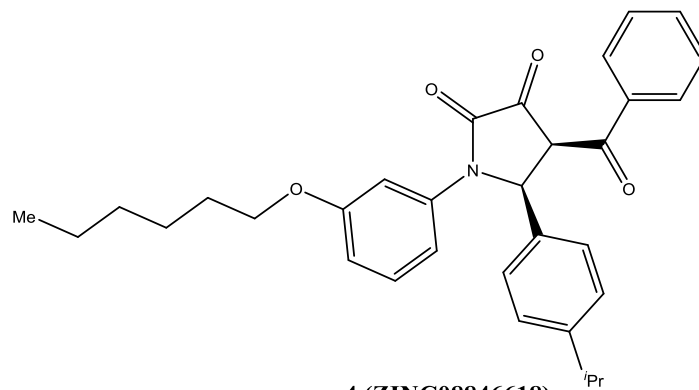
Same four receptor grids as for SERCA1 and SERCA2 models were also set up for SERCA3 homology model. SiteMap site 4 was based on ALA142, ARG141 and ARG236. The G scores obtained after the XP docking of the same set of 26 ligands ranged from 2.043 to -3.865. For SiteMap site 1 on LEU319, GLY322 and ARG251, the G scores were found to be between -1.169 to -8.279. The grid for SiteMap site 3 was set on ASN756, GLY808 and LYS252 and the G scores ranged from -1.560 to -9.397. For the TG-equivalent site on ILE765, PHE256 and PHE834, the G scores were between -5.597 to -11.565.

3.1.3 CombiGlide results

The details of applying CombiGlide to the six compounds selected preliminarily were elaborated in details on page 48 in Chapter two, the Methods section.¹⁴⁸ The side-chain substitutions were applied to all six selected compounds and the compounds with substituted side groups were subjected to XP docking on SPCA1d, SPCA2, SERCA1 and SERCA3 homology models. The overall average G scores achieved from the CombiGlide process did not exceed the average G scores from the previous dockings.

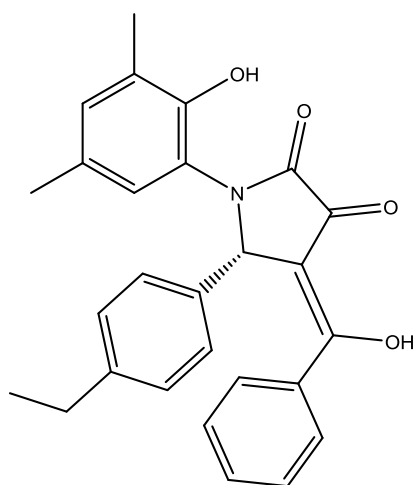


3 (ZINC08846492)

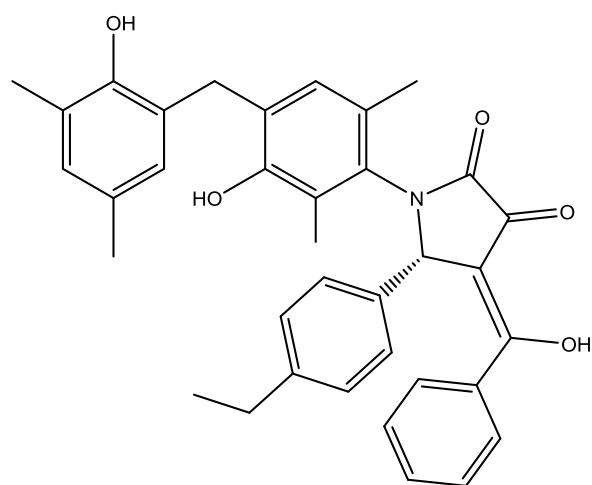


4 (ZINC08846618)

Figure 15

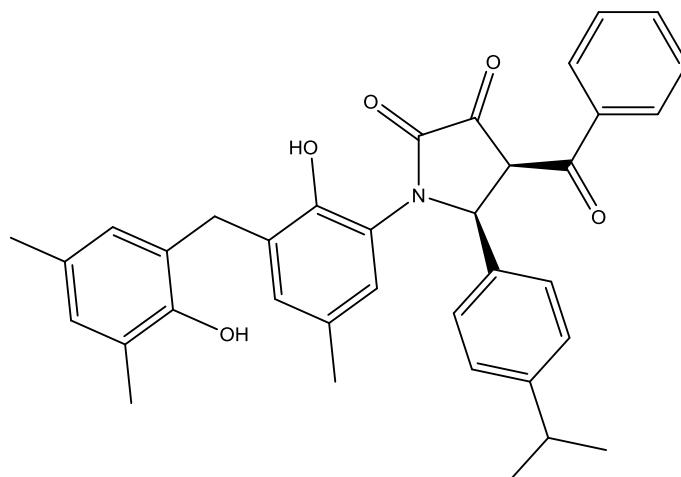


ZINC08846492 with half of bisphenol attachment (2,4-dimethylphenol)



ZINC08846492 with a bisphenol group attachment

Figure 16



**ZINC08846618 with a
bisphenol group attachment**

Figure 17

The two selected ligands for SPCA1d model were compounds **3** (ZINC08846492) and **4** (ZINC08846618) (**Figure 15**). From the first combinatorial library with imported Schrödinger fragment 1 to 50a, ZINC08846492 with fragment number 50 had the highest G score of -7.233. From the second combinatorial library of imported fragment 51 to 200a, ZINC08846492 with fragment number 154a had the highest G score of -7.049. On the same ligand with fragment 291 from the third combinatorial library with fragment 201 to 441, the highest G score obtained was -8.765. Several other different side-chain attachment points were also trialled and the G scores obtained ranged from -4.001 to -7.688. A trial was also done to attach bisphenol group onto ZINC08846492 and the G scores ranged from -7.027 to -7.711 (**Figure 16**). CombiGlide process was also applied to ZINC08846618. The bisphenol group was also trialled on ZINC08846618 but only a G score of -6.377 was obtained (**Figure 17**). The first combinatorial library included Schrödinger fragments 1 to 120, the highest G score was -8.162. For the second combinatorial

library containing fragments 121 to 250, the highest G score was -8.972. For the third combinatorial library with fragments 251 to 441, the highest G score was -8.103.

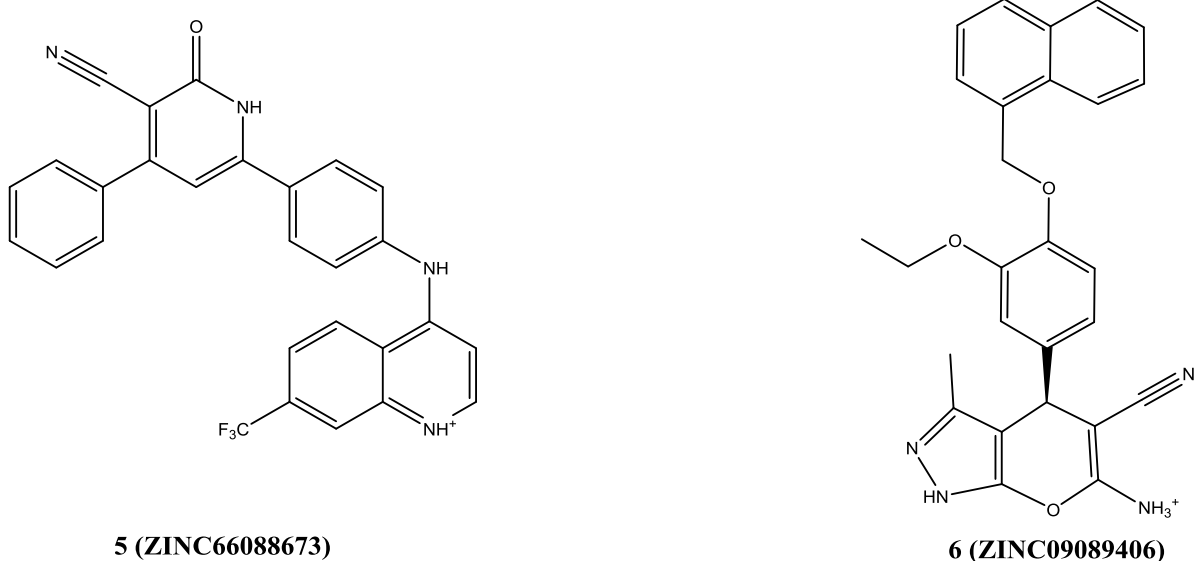


Figure 18

For the SPCA2 model, the selected Tim Tec ligands were compounds **6** (ZINC09089406) and compound **5** (ZINC66088673) (**Figure 18**). The first combinatorial library set up contained fragments 1 to 170. The highest G score for ZINC09089406 was -9.347 with fragment 169. For ZINC66088673, the highest G score was -10.399 with fragment 39. The second combinatorial library had fragments 171 to 310c. The highest G score for ZINC09089406 was -8.769 for fragment 306b and that for ZINC66088673 was -9.276 for fragment 279. For the third combinatorial library, it contained fragments 311 to 441. The highest G score for ZINC09089406 was -8.613 for fragment 394a and that for ZINC66088673 was -9.70 for fragment 429.

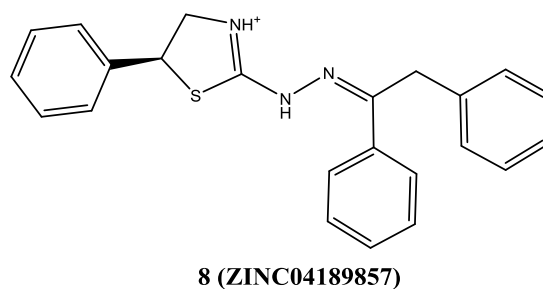
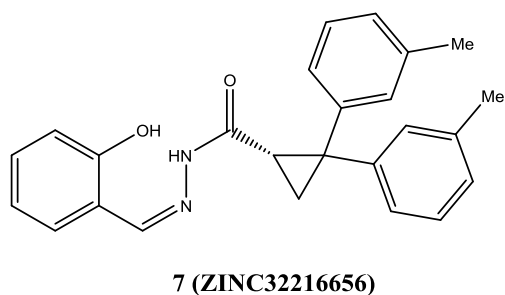
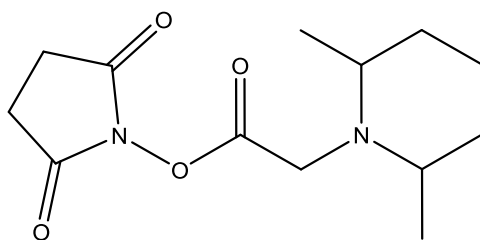


Figure 19

For the SERCA1 model, the ligands selected were compound **8** (ZINC04189857) (**Figure 19**). The first combinatorial library set up had fragments 1 to 170 and the highest G score was -10.889 for fragment 57. The second combinatorial library contained fragments 171 to 310c. The highest G score from the XP docking was -10.472. The third combinatorial library had fragments 311 to 441, the highest G score was -8.64. For SERCA3 model, the selected ligand was compound **7** (ZINC32216656) (**Figure 19**). The first combinatorial library had fragments 1 to 180a with the highest G score of -9.867. The second combinatorial library was set up with fragments 181 to 310c and the highest G score from the docking was -9.307. The third combinatorial library comprised of fragments 311 to 441. The highest G score from docking was -9.869.

3.1.4 Results from second cycle of molecular modelling



**Pyrrolidone-like structure
from PubChem**

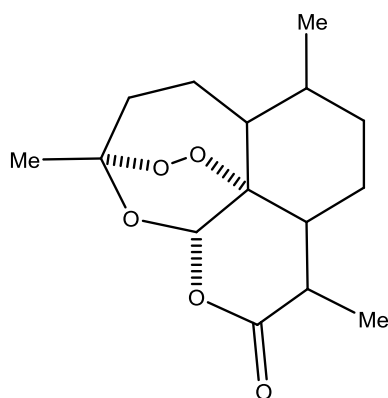
Figure 20

From SciFinder, one compound was found to bear the same structure as compound **9** (from section 2.4.1).¹⁵¹ A trial XP docking of this compound was done on SPCA1d model with only the highest G score of -4.03 obtained. From PubChem, one compound with a pyrrolidone-like core was discovered (**Figure 20**).¹⁶² The highest G score obtained from docking this compound onto the SPCA1d model was -5.106. Both compounds were prepared *via* LigPrep prior to docking¹³⁸.

BREED method was originally developed by Vertex Pharmaceuticals as an automated method to generate novel inhibitors from known structures of the ligands binding to target proteins.¹⁶³ Schrödinger has licensed and inherited this algorithm from Vertex Pharmaceuticals. The BREED method was accessed from the script centre under the fragment application of the Maestro interface. This method was applied to the structures found from the ZINC database where 22 pyrrolidone-like and also another 26 less pyrrolidone-like structures were discovered. The 22 pyrrolidone-like structures were subjected to the BREED algorithm and generated a total of 154 hybrid structures. The highest G score produced from XP docking onto the SPCA1d model was -

5.903 from these 154 hybrid structures. Also, there were altogether 182 hybrid compounds generated from the BREED algorithm for the 26 less pyrrolidone-like structures. The highest G score produced was -7.166 for the second batch of the hybrid compounds.

A trial of docking CPA on SERCA1 was done to see the actual G scores produced. A total of three binding sites were trialled; Sitemap-generated site 1, 3 and 4. The G score obtained from site 3 was the highest with -6.786 which was followed by G score of -4.949 from site 1 and lastly a G score of -2.614 from site 4. BREED method was again used to produce hybrid molecules from CPA and the 26 less pyrrolidone-like structures. A total of 202 hybrid molecules were formed and docked on SERCA1 model. The highest G score was found to be -7.650 on Sitemap-generated site 3. Comparing this G score with the docking of the 26 less pyrrolidone-like structures onto SERCA1, the highest G score of -6.081 was obtained on Sitemap site 3 and this score did not have a huge difference to the score for the hybrid molecule.



Artemisinin

Figure 21

The next compound trialled was artemisinin (**Figure 21**) as it was found to inhibit the SERCA orthologue of *Plasmodium falciparum* in *Xenopus* oocytes.¹⁶⁴ The structure of artemisinin was downloaded *via* ZINC database and prepared *via* LigPrep.¹³⁸ XP docking of artemisinin was done on SERCA1 model on the same Sitemap site 1, 3 and 4. The highest G score was on Sitemap site 1 with -3.248 only. Due to this low G score and after some considerations, a new receptor model of the SERCA orthologue, PfATP6, of *Plasmodium falciparum* was built to trial docking the artemisinin structure. This was done *via* importing the PDB structure of 1U5N into Maestro and using Protein Preparation Wizard to prepare the protein model. Sitemap was used specifically to identify any druggable sites.¹⁰⁹ Two receptor grids were set up with Sitemap site 4 based on ARG1153, LEU326 and ALA1136 and Sitemap site 1 on LEU125, VAL112 and ASP1038. Artemisinin was docked in XP mode on both sites with highest G score only at -2.139.

Curcumin in its keto-enol form was also trialled on the SERCA1 model. The structure was obtained from ZINC database and was prepared *via* LigPrep prior to docking.¹³⁸ The highest G score was found to be from Sitemap site 1 which had -5.626. Hybrid molecules were again produced by combining artemisinin and the 26 less pyrrolidone-like structures by using BREED algorithm. Altogether 189 hybrid compounds were generated. These molecules were docked onto the SERCA1 model in XP mode. Highest G score was -7.431 (ZINC32105134 + ZINC49646993; **Figure 22**) from Sitemap site 3. BREED method was applied to generate hybrid molecules by combining CPA, artemisinin, curcumin and the 26 less pyrrolidone-like structures. The highest G score was -7.650 which was the same hybrid molecule (ZINC04300723 + ZINC49647007; **Figure 23**) generated from the previous batch of hybrid compounds *via* merging CPA and the 26 less pyrrolidone-like structures.

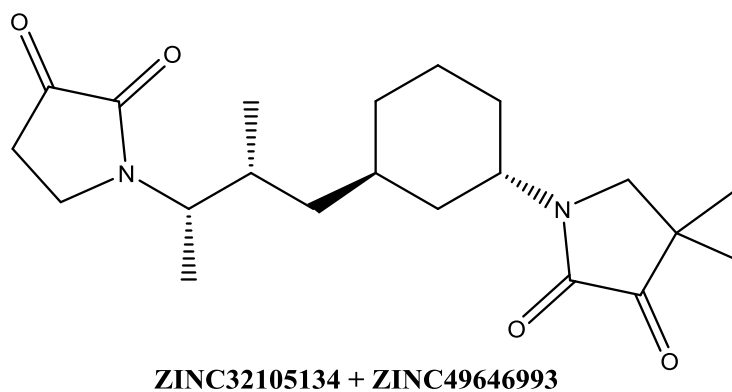


Figure 22

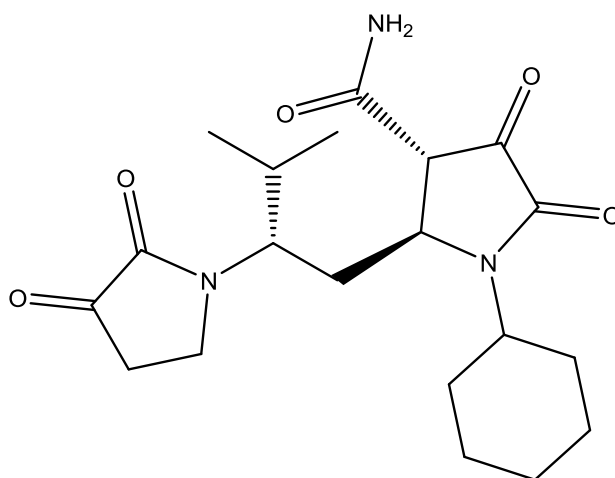


Figure 23

A cyclic pyrazole derivative of curcumin was drawn in 2D sketcher on the Maestro interface and converted to 3D structure. Same process was carried out for diallyl bisphenol A, which was tested to inhibit SERCA.¹⁵³ The cyclic pyrazole curcumin derivative was docked onto SERCA1 model on Sitemap site 1, 3 and 4. The highest G score was -5.968 on site 1. Same docking was also applied to diallyl bisphenol A on SERCA1 with the same three binding sites. The highest G score found was -6.348. LEU61, VAL62, LEU65, ILE97, LEU253, ILE307, PRO308, LEU311 and PRO312 were found to be the hydrophobic residues interacting with most of the non-polar

parts of the bisphenol inhibitors as suggested by Woeste *et. al.*¹⁵³ ASP254 was also found to be the critical amino acid residue interacting with the hydroxyl group of bisphenol inhibitors via hydrogen bonding. Therefore, a new grid was formed on SERCA1 on ASP254, ILE97 and PRO312. The two prepared diallyl bisphenol A were docked onto this new binding site with the highest G score of -4.803. The original CPA binding site on SERCA was also trialled with the grid set on GLN56, LEU61 and ASN101.⁴⁰ The G score obtained was -4.418.

3.1.5.1 Results from first cycle of compound selection

The G scores and MM-GBSA results facilitated the final compound selection for biological testing. From the 26 ligands identified from the docking process on the SPCA2 model, six compounds were preliminarily selected. These six compounds were chosen on the basis of possessing G scores of at least -7 or below with reasonable MM-GBSA results of binding energy between the ligand and receptor (**Table 4**). Even though CombiGlide was trialled on the six preliminarily selected compounds, the overall average G scores did not surpass the previous G scores obtained. Therefore two compounds, ZINC08846492 (compound **3**) for SPCA1d and ZINC04189857 (compound **8**) for SERCA1, were selected to be purchased from Tim Tec out of the six compounds selected originally.

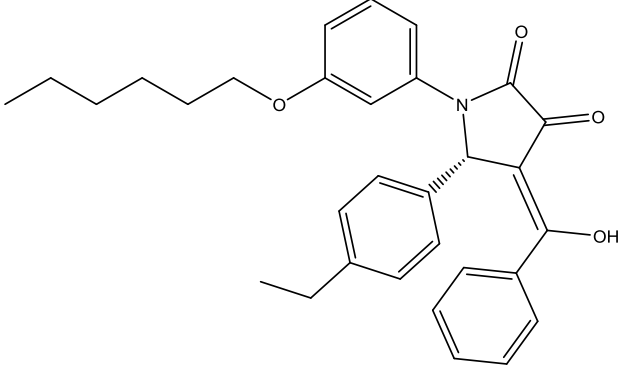
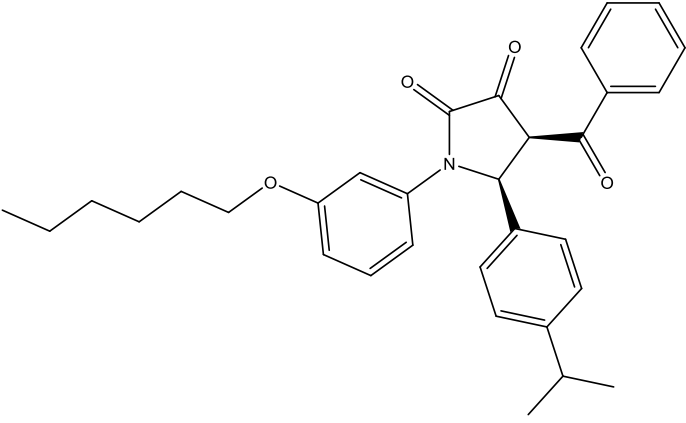
Screening compounds for 1SU4-SPCA1d	Properties	Scores
<p style="text-align: center;">ZINC08846492</p> 	XP G Score	-7.688
	MMGBSA dG bind energy	-90.593513
	Ligand strain energy	13.58
	Prime energy (of the complex)	45305.1
<p style="text-align: center;">ZINC08846618</p> 	XP G Score	-7.027
	MMGBSA dG Bind energy	-85.14
	Ligand strain energy	38.47
	Prime energy (of the complex)	45276.3

Table 4 Six pre-selected compounds with corresponding G scores and MM-GBSA results

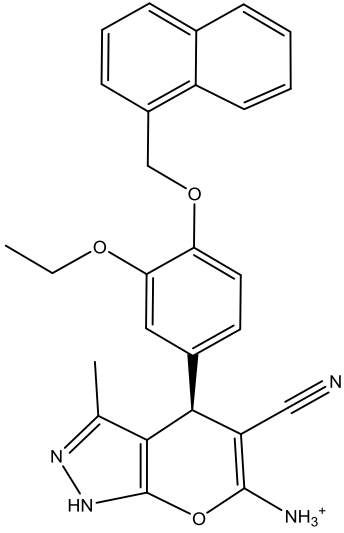
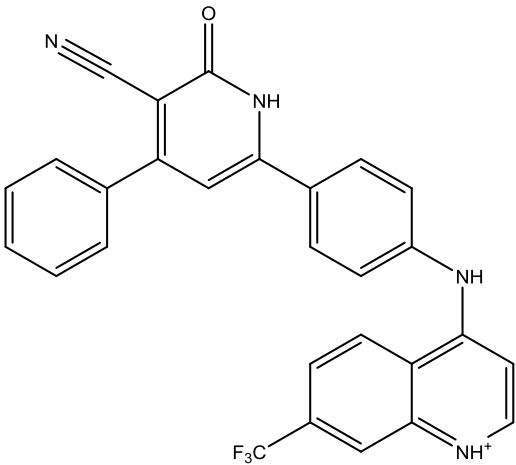
Screening compounds for 1IW0-SPCA2	Properties	Scores
ZINC09089406 	XP G Score	-8.028
	MMGBSA dG Bind energy	-71.092079
	Ligand strain energy	8.62
	Prime energy (of the complex)	-40069.1
ZINC66088673 	XP G Score	-11.116
	MMGBSA dG Bind energy	-70.420023
	Ligand strain energy	6.68
	Prime energy (of the complex)	-40065.8

Table 4 (continued) Six pre-selected compounds with corresponding G scores and MM-GBSA results

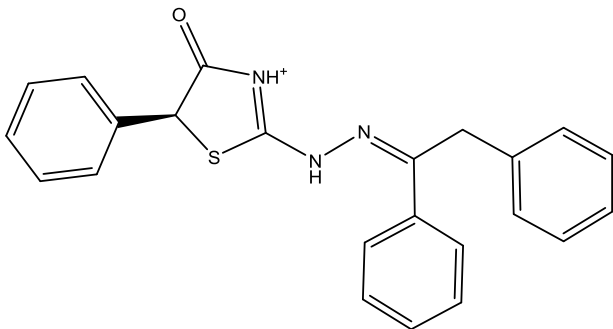
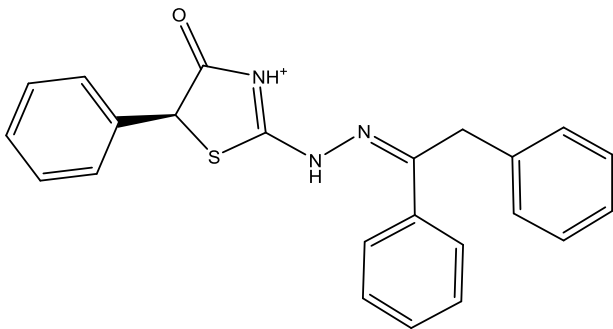
Screening compounds for 3AR5-SERCA1	Properties	Scores
<p style="text-align: center;">ZINC04189857</p> 	XP G Score	-10.917
	MMGBSA dG Bind energy	-78.466592
	Ligand strain energy	4.425
	Prime energy (of the complex)	-26763.2
Screening compounds for 3AR4-SERCA2 (same compound as 3AR5-SERCA1)	Properties	Scores
<p style="text-align: center;">ZINC04189857</p> 	XP G Score	-11.9
	MMGBSA dG Bind energy	-77.767271
	Ligand strain energy	7.721
	Prime energy (of the complex)	-41203.8

Table 4 (continued) Six pre-selected compounds with corresponding G scores and MM-GBSA result

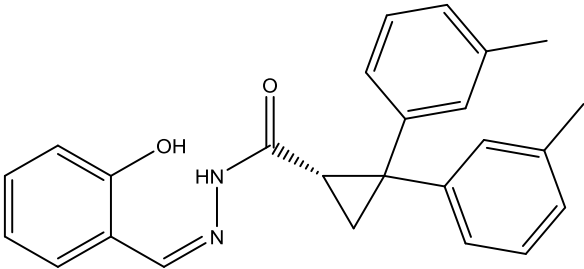
Screening compounds for 3AR7-SERCA3	Properties	Scores
<p style="text-align: center;">ZINC32216656</p> 	XP G Score	-9.17
	MMGBSA dG Bind energy	-85.252258
	Ligand strain energy	1.982
	Prime energy (of the complex)	-41040.2

Table 4 (continued) Six pre-selected compounds with corresponding G scores and MM-GBSA results

3.1.5.2 Results from second cycle of compound selection

The initial draft selection of compounds that had the most potential to be the testing compounds were two of the hybrid molecules generated from the BREED method. These hybrid molecules had the top two G scores which were ZINC04300723 + ZINC49647007 (**Figure 23**) with a G score of -7.650 and ZINC32105134 + ZINC49646993 (**Figure 22**) with a G score of -7.431. Other potential compounds considered were the diallyl bisphenol A, curcumin and cyclic pyrazole derivative of curcumin. All of these three compounds had G scores higher than that for CPA relatively. Therefore, it was decided to synthesise these compounds in the laboratory (compounds **12**, **13**, **15**, **16** and **17**). The two hybrid molecules however had complex stereochemistry and were not readily available commercially hence they were not chosen to be synthesised at this stage. Also in collaboration with the honours students, another compound with a similar structural core was also decided to be tested (compound **10**).

3.2 Dose - response curves from FLIPR

FLIPR testing was based on the principle of SOCE which was explained in details in Chapter one. The results were presented in the Ca^{2+} response or dose-response (for compounds **3** and **8** only) curves with two distinctive Ca^{2+} peaks shown. This was one of the standard experimental protocols used to examine and study the Ca^{2+} influx phase in details.¹⁶ A small Ca^{2+} increase occurred initially when the remaining Ca^{2+} in ER leaked out from IP_3R while ER was blocked by CPA or tested compounds. This was then followed by a slow reduction in Ca^{2+} response as the intracellular Ca^{2+} returned to its basal level. Overall this was represented by the first Ca^{2+} peak. After the addition of Ca^{2+} , there would be a larger Ca^{2+} influx, represented by the second Ca^{2+} peak. An example of graphical interpretation of these two Ca^{2+} peaks could be seen in **Figure 24**.

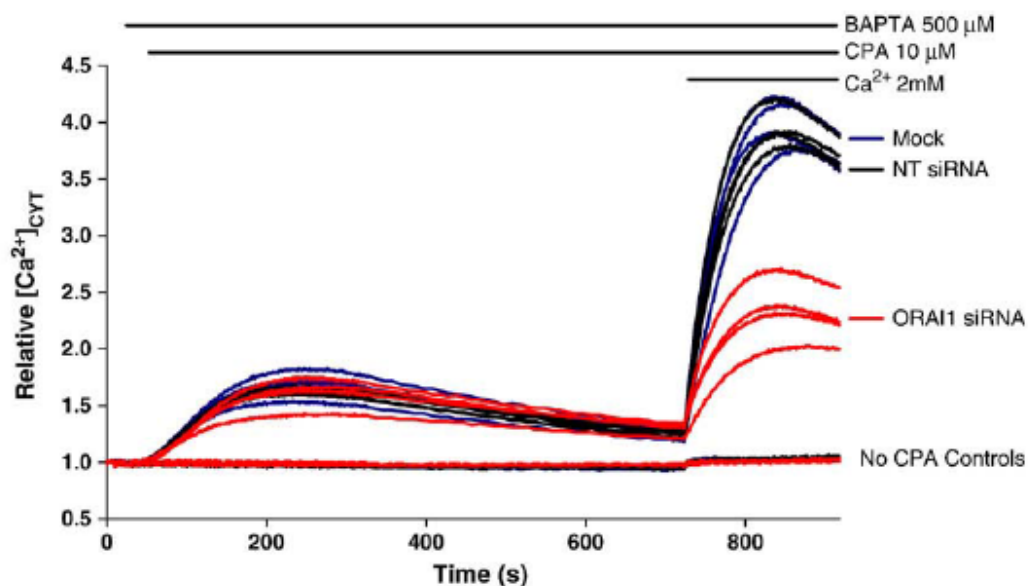


Figure 24 Example of a Ca^{2+} response profile (taken from ref. 16)

From the first two compounds tested (compound **8** – ZINC04189857 and compound **3** – ZINC08846492), dose–response curves were generated based on the first and second Ca^{2+} peaks produced after treating the MDA-MB-231 cells. These peaks were compared with the Ca^{2+} peaks from the control, CPA. Compound **8** was not adequate to generate any dose response for the first Ca^{2+} peak. The second peak from compound **8** only managed to produce a very minimal Ca^{2+} response when compared to that from CPA. On the other hand, compound **3** was able to generate a slightly better dose response on the first Ca^{2+} peak than that from compound **8**. The second Ca^{2+} peak after compound **3** treatment only produced a similar minimal dose response as from compound **8** treatment. Both compounds did not elicit significant Ca^{2+} dose response in the MDA-MB-231 cells as CPA. The comparison results could be seen in **Figure 25** with the individual dose-response curves shown in **Figure 26**.

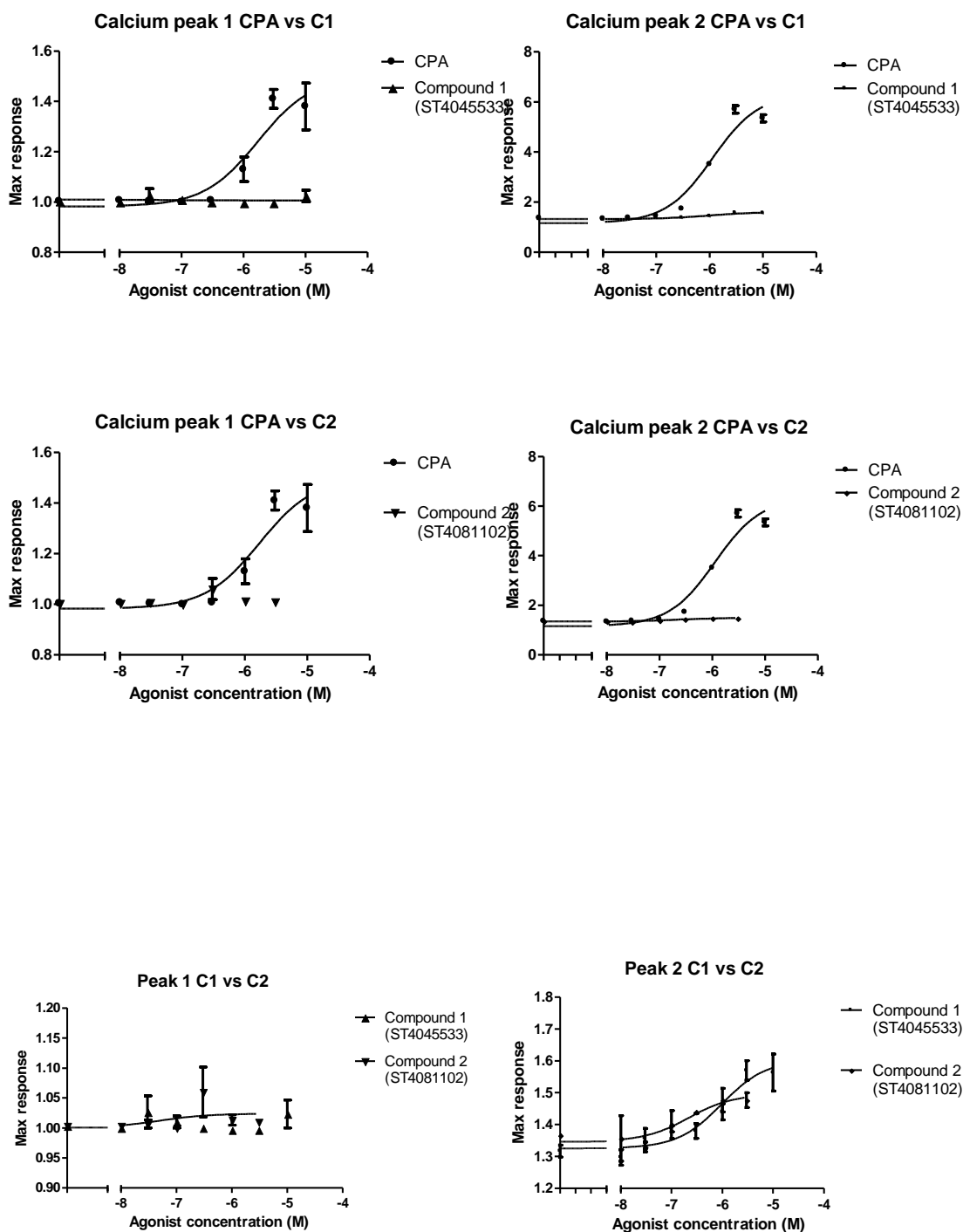


Figure 25 Comparison dose response curves for compound **8** or **C1** (ZINC04189857), compound **3** or **C2** (ZINC08846492) and CPA. Peak 1 represented the first Ca^{2+} peak and peak 2 represented the second Ca^{2+} peak from the FLIPR testing (M = Molar concentration in semi-logarithmic plot)

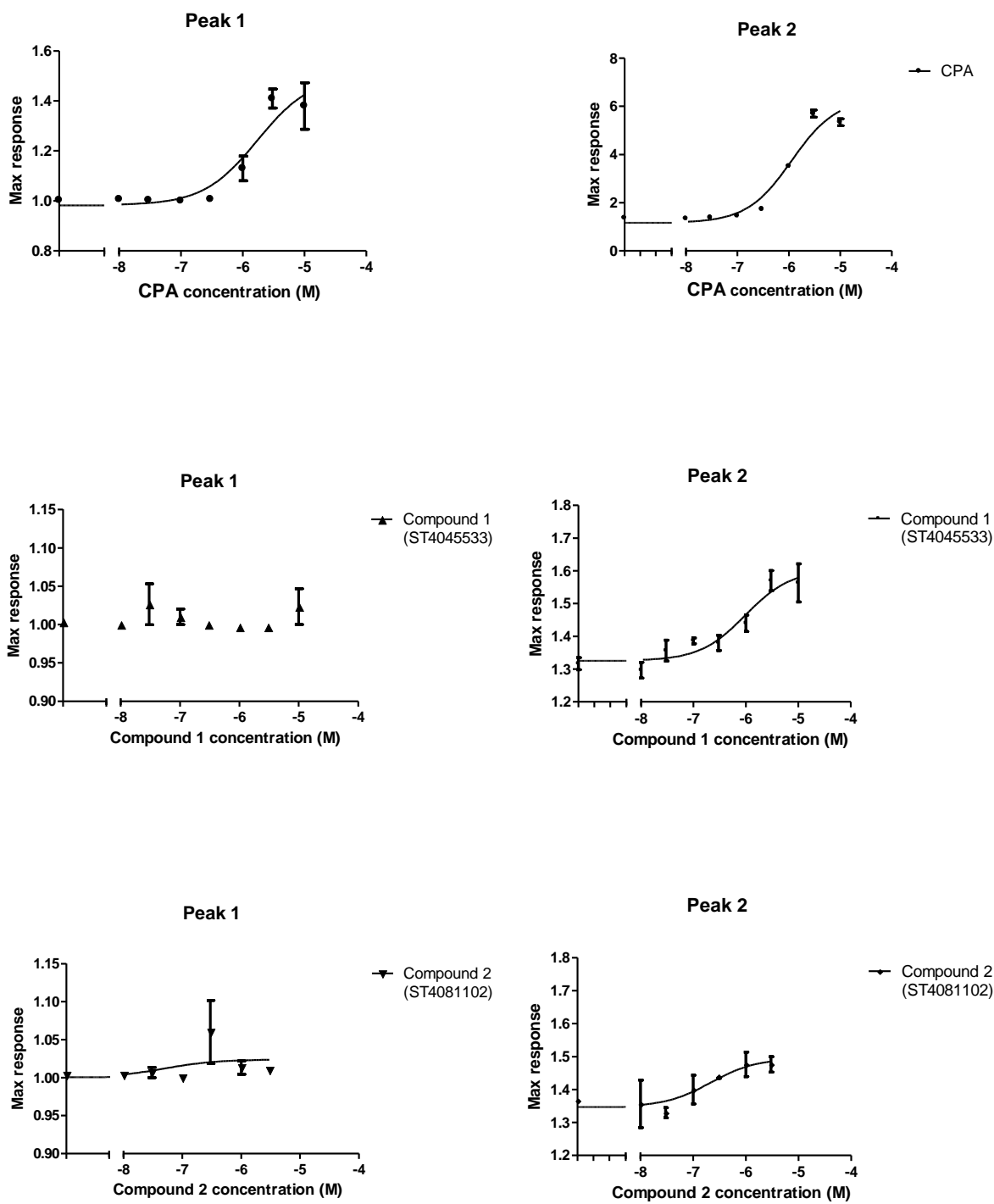


Figure 26 Individual dose response curves for compound **8** (Compound 1 or ZINC04189857), compound **3** (Compound 2 or ZINC08846492) and CPA. Peak 1 represented the first Ca²⁺ peak and peak 2 represented the second Ca²⁺ peak from the FLIPR testing (M = Molar concentration in semi-logarithmic plot)

3.3 Intracellular calcium response curves from FLIPR

3.3.1 Results for compound 8 (ZINC04189857) and compound 3 (ZINC08846492)

The Ca^{2+} response curves were obtained from the FLIPR and shown in the following figures. These figures demonstrated the details of the intracellular Ca^{2+} signalling over time within the MDA-MB-231 cells after these two compound treatments when compared to the control groups, CPA and DMSO. **Figure 27** was composed of the different compound concentrations applied to the MDA-MB-231 cell line. It demonstrated the first Ca^{2+} peak when BAPTA and compound **8** and **3** were added. The most prominent Ca^{2+} peaks that could be easily observed at the top were from CPA treatment with concentrations from 10 μM , 3 μM and 1 μM . The lower ranges of CPA concentrations were buried within the compound **8** and **3** groups. All of the compound **8** and **3** concentrations used only provided a Ca^{2+} response from 0.75 to 1.0 over the baseline Ca^{2+} response. In another words, there were no Ca^{2+} peak generated when compound **8** and **3** were applied to the cell line.

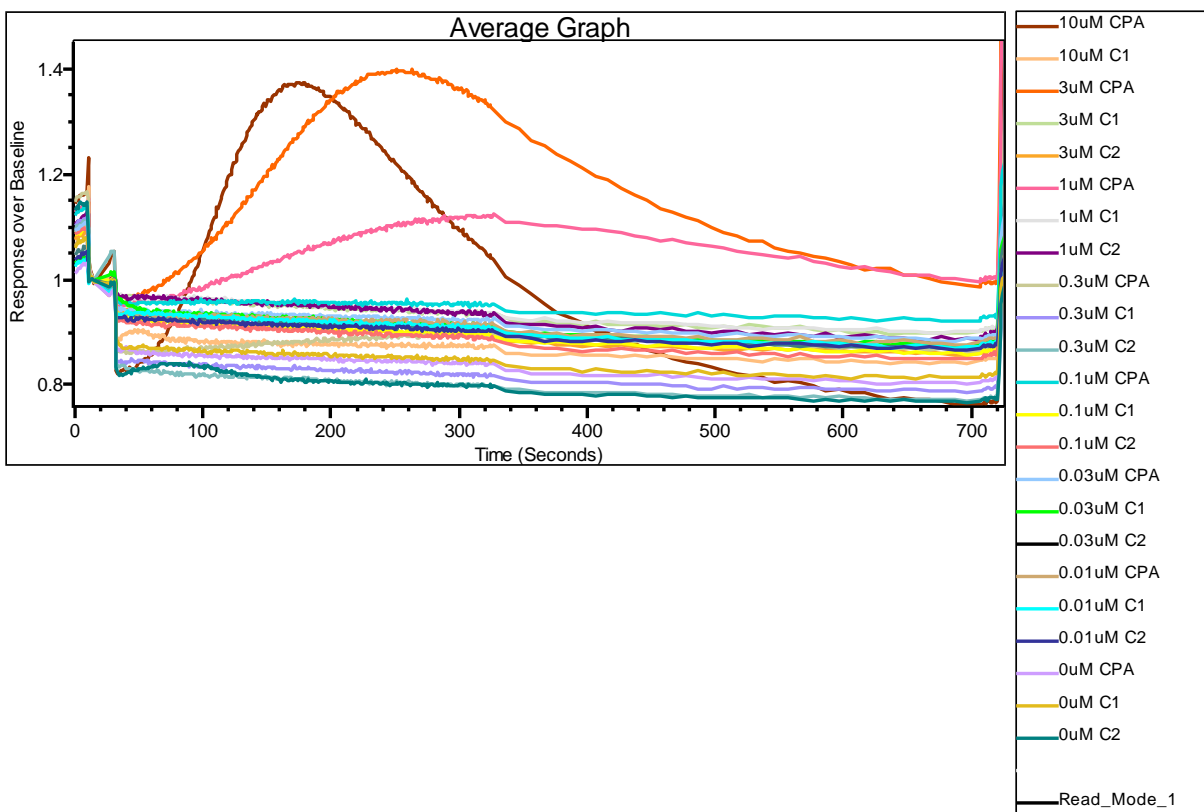


Figure 27 Agonist concentrations of CPA, Compound **8** (C1 or ZINC04189857) and Compound **3** (C2 or ZINC08846492) versus Ca^{2+} response over time (zoom-in for first Ca^{2+} peak)

For **Figure 28**, this showed the second Ca^{2+} peak when Ca^{2+} was added to the MDA-MB-231 cell line. Again the most obvious peaks were from the CPA treatment (positive control group). The easily seen CPA concentrations were also the same as that for the first Ca^{2+} peak, for 10 μM , 3 μM and 1 μM but also included the 0.3 μM . Compound **8** in 3 μM was the highest concentration to give a slightly higher second Ca^{2+} peak than the rest of the concentrations for both Compound **8** and **3**. Overall the second Ca^{2+} signals were very weak for both compound **8** and **3** when compared with our control, CPA group. **Figure 29** demonstrated the overall Ca^{2+} response for all concentrations of all three tested agonists, compounds **8**, **3** and CPA.

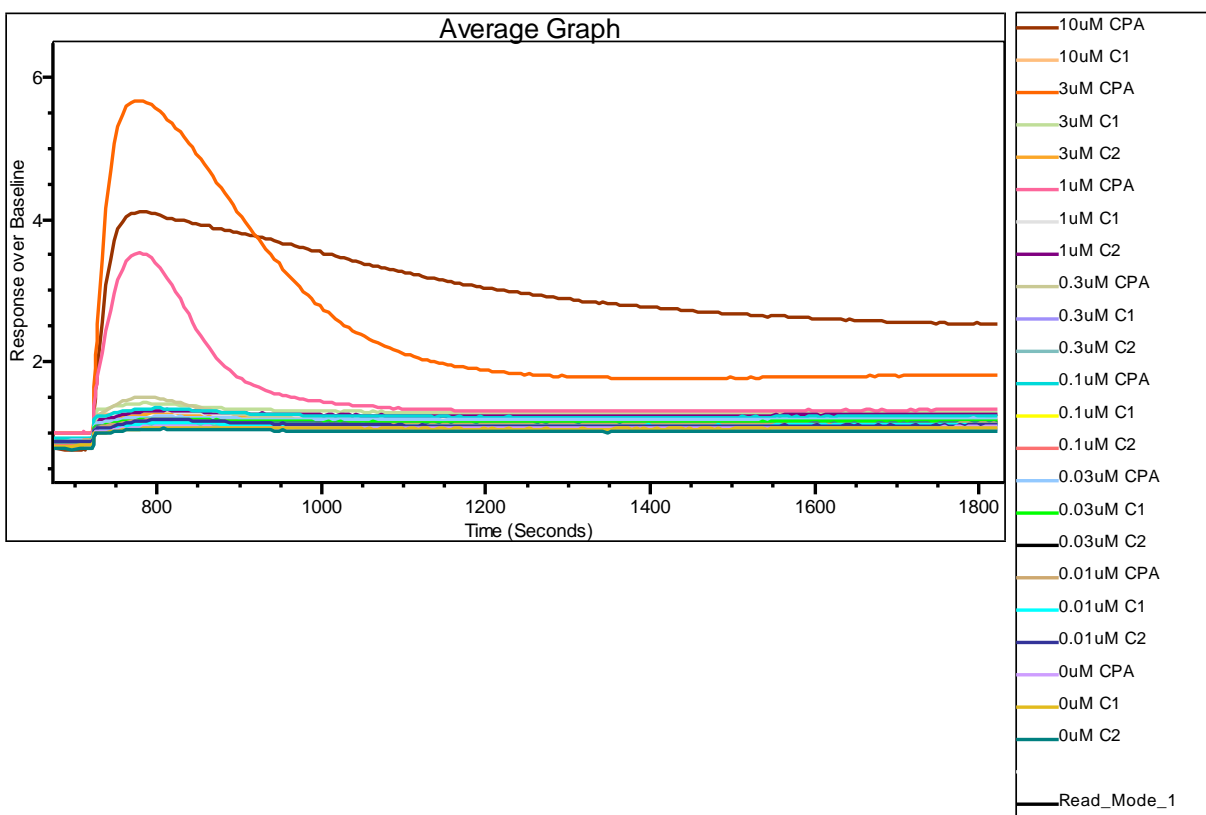


Figure 28 Agonist concentrations of CPA, Compound **8** (C1 or ZINC04189857) and Compound **3** (C2 or ZINC08846492) versus Ca^{2+} response over time (zoom in for second Ca^{2+} peak)

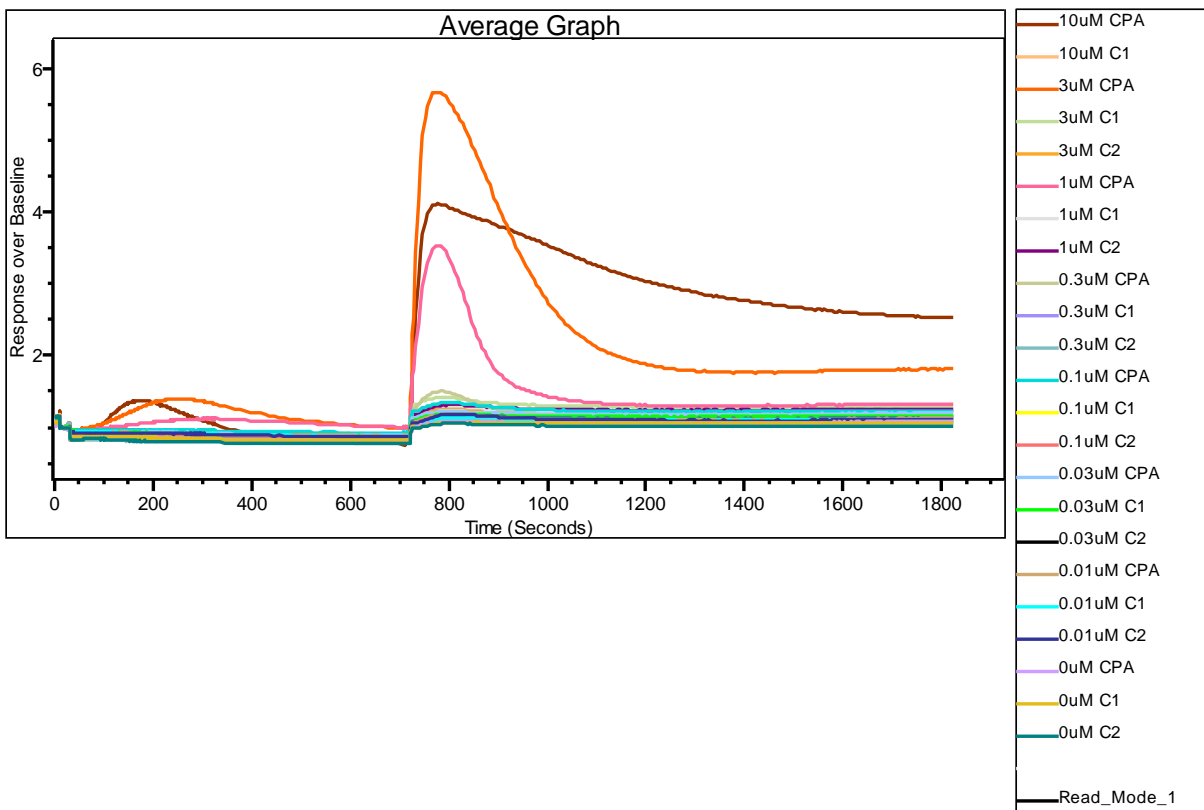


Figure 29 Agonist concentrations of CPA, Compound **8** (C1 or ZINC04189857) and Compound **3** (C2 or ZINC08846492) versus Ca^{2+} response over time

Figure 30 only included the top four concentrations for all three agonist treatment groups for the ease to observe the Ca^{2+} response clearly. The concentrations ranged from $0.1\mu\text{M}$ to $10\mu\text{M}$ for all three agonist groups. A similar Ca^{2+} signalling effect was seen as from the previous figures. Both compound **8** and **3** did not generate significant first and second Ca^{2+} peaks. This was especially clear in the enlarged zoom-in graphs for first and second peaks. The peaks generated from CPA again were the standard SERCA inhibitor Ca^{2+} signalling curves in the cell line as seen in the figure.

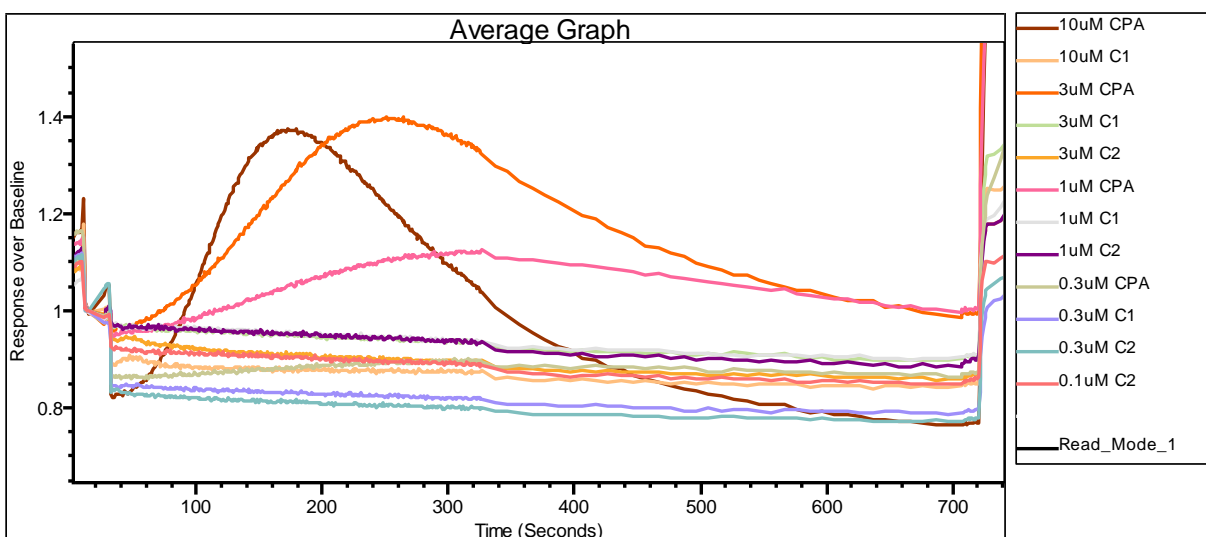


Figure 30 Top four agonist concentrations for three treatment groups (*Top*: zoom-in for first Ca^{2+} peak *Bottom*: zoom-in for second Ca^{2+} peak)(C1 = compound **8** C2 = compound **3**)

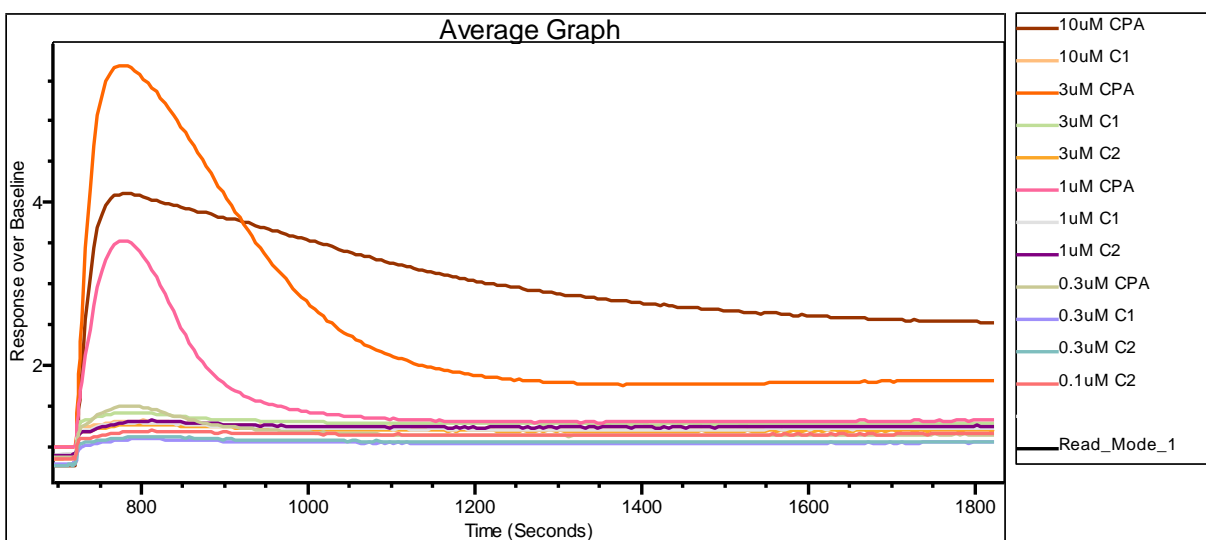


Figure 30 (continued) Top four agonist concentrations for three treatment groups (*Top*: zoom-in for first Ca^{2+} peak *Bottom*: zoom-in for second Ca^{2+} peak)(C1 = compound **8** C2 = compound **3**)

Figure 31 has focussed on just the top three agonist concentrations for all three treatment groups. First Ca^{2+} peaks for both compounds **8** and **3** at all concentrations only produced flat Ca^{2+} response when compared with our positive control group, CPA. However for the second Ca^{2+} peak, both compounds **8** and **3** produced a small increase in Ca^{2+} response after Ca^{2+} was added. Overall the small Ca^{2+} increase for second peak was not large enough to warrant a resemblance to the positive control. **Figure 32** again showed only the top two agonist concentrations for all three treatment groups with similar outcome as mentioned previously.

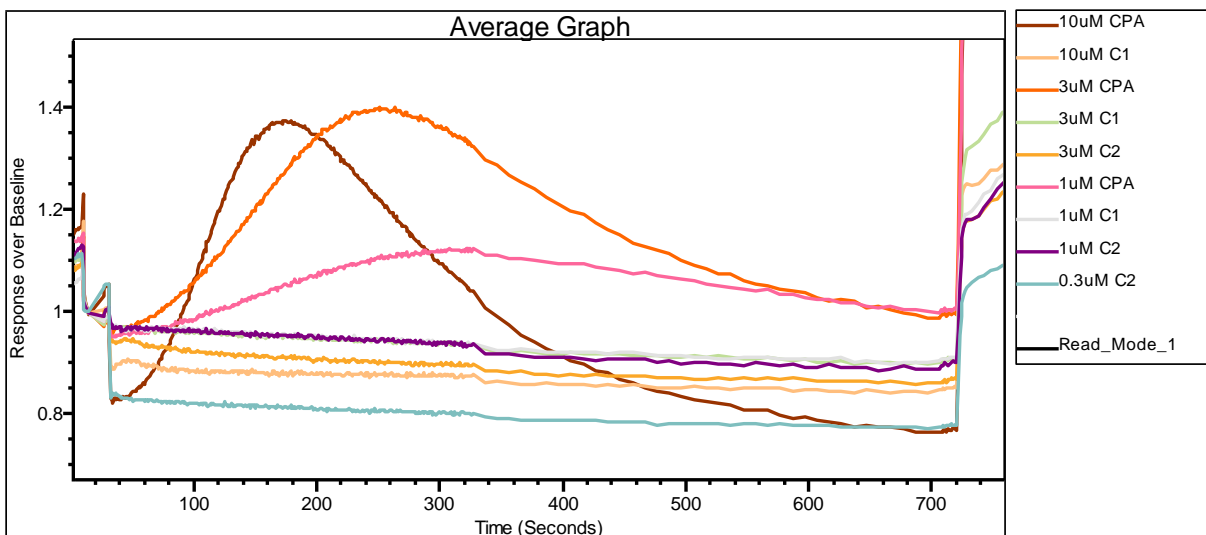


Figure 31 Top three agonist concentrations for three treatment groups (*Top*: zoom-in for first Ca^{2+} peak *Bottom*: zoom-in for second Ca^{2+} peak) (C1 = compound **8** C2 = compound **3**)

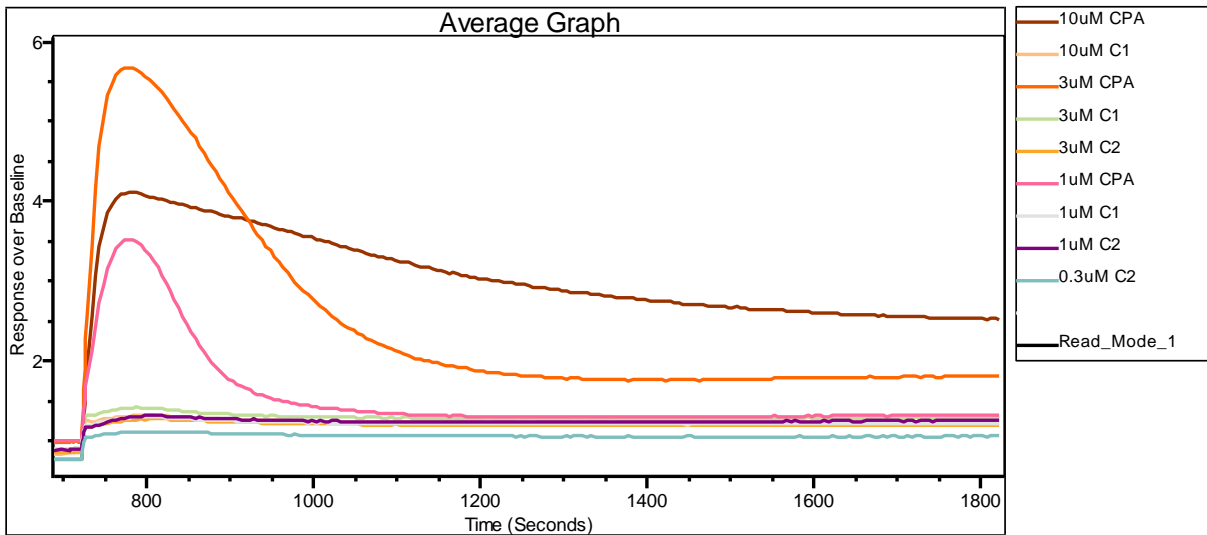


Figure 31 (continue) Top three agonist concentrations for three treatment groups (*Top*: zoom-in for first Ca^{2+} peak *Bottom*: zoom-in for second Ca^{2+} peak) (C1 = compound 8 C2 = compound 3)

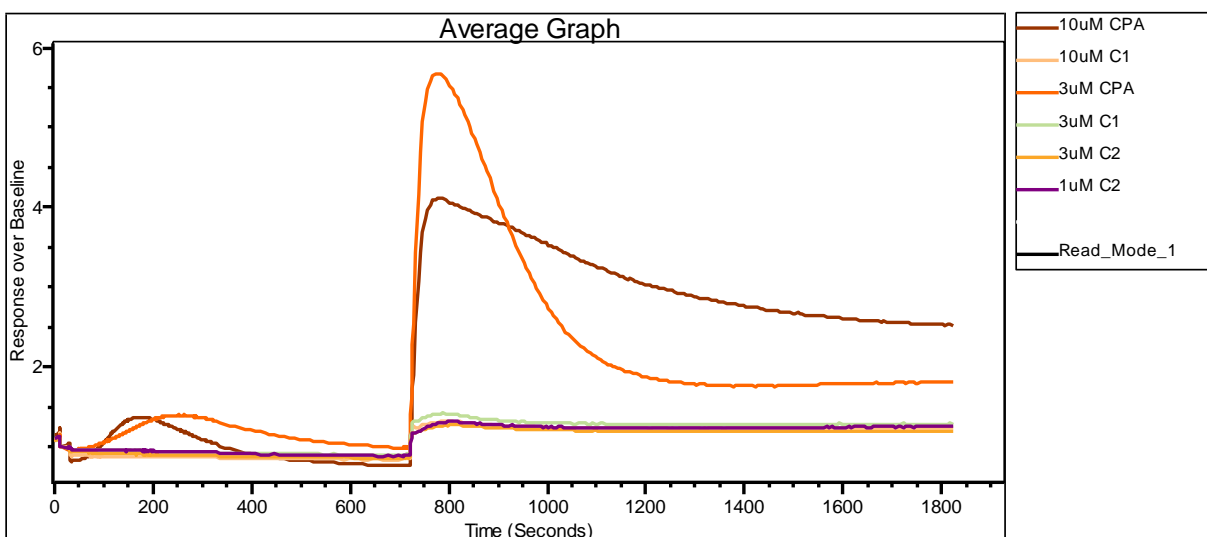


Figure 32 Top two agonist concentrations for three treatment groups (C1 = compound **8** C2 = compound **3**)

For **Figures 33 and 34**, the positive control group, CPA, was compared individually with both compounds **8** and **3**. Only the top two concentrations were shown in the figures. Both compounds **8** and **3** were very similar in their overall Ca^{2+} signalling profiles. In **Figure 35**, the three agonist treatment groups were compared with the negative control, DMSO. Clearly the top two concentrations of CPA demonstrated the typical SERCA inhibitor Ca^{2+} response curves. Both compounds **8** and **3**, on the other hand, were much less prominent but when compared with our negative control group, the Ca^{2+} responses were still slightly higher than that for DMSO (shown as 0 μM of CPA, compounds **8** and **3**). The DMSO concentration used in the experiment was set at 4 $\mu\text{L}/\text{mL}$. The direct comparisons between 10 μM of compound **8** and 3 μM of compound **3** against DMSO were shown in the figures. Here the slight increments in Ca^{2+} were present but smaller when compared with the figure consisting 10 μM of CPA against DMSO. **Figures 36, 37 and 38** have shown the top four concentrations of CPA, compounds **8** and **3** against DMSO to reveal the differences between each agonist.

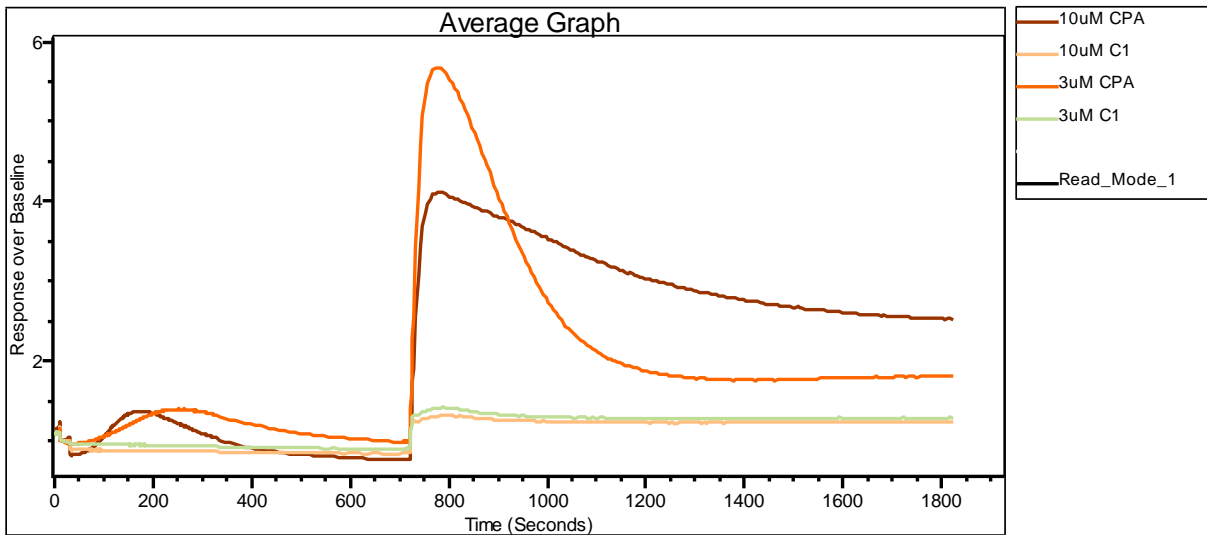


Figure 33 Top two concentrations of CPA versus Compound **8** (C1 or ZINC04189857)

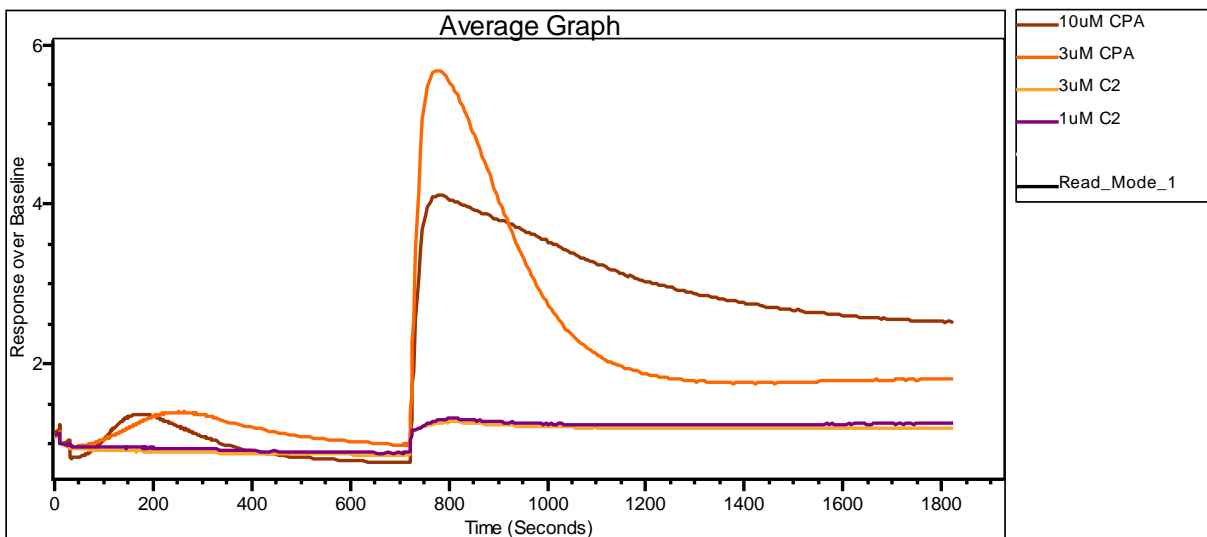


Figure 34 Top two concentrations of CPA versus Compound **3** (C2 or ZINC08846492)

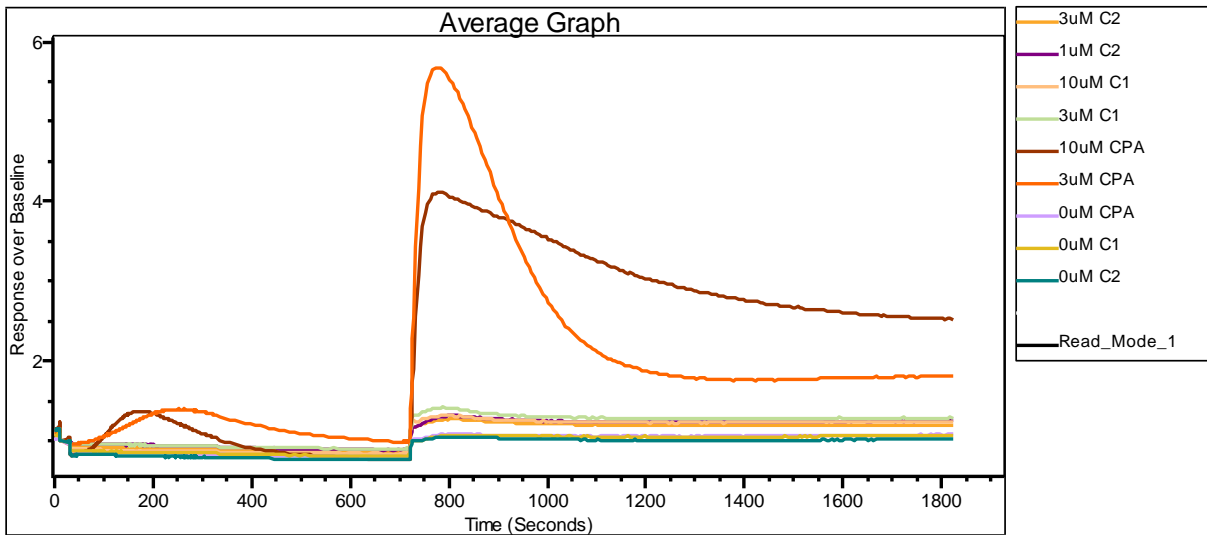


Figure 35 Top two concentrations of three agonist treatments versus DMSO (4 μ L/mL)
(C1 = compound 8 C2 = compound 3)

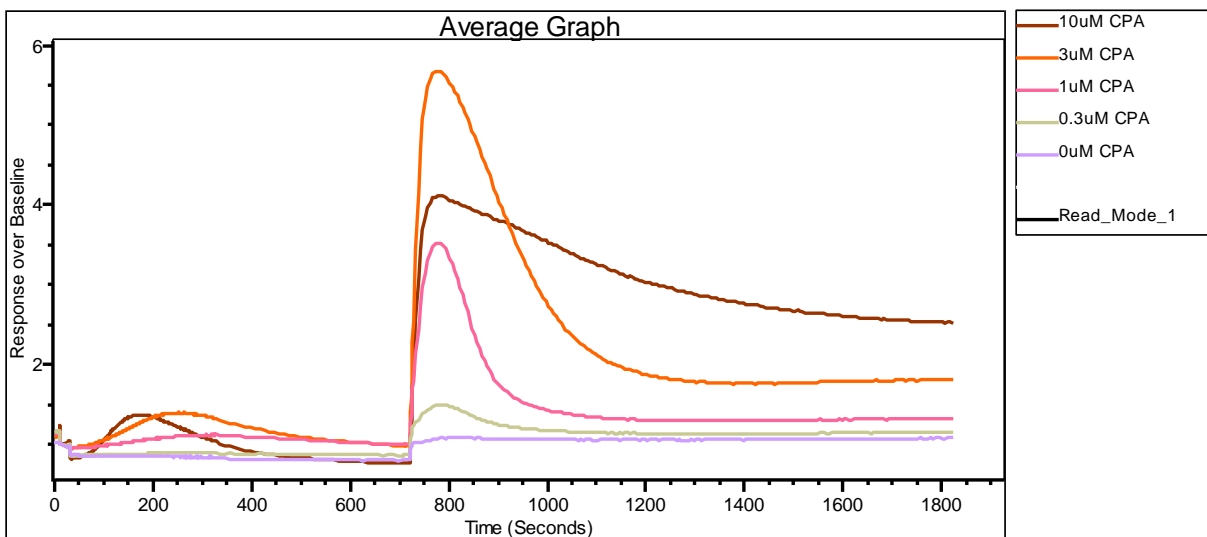


Figure 36 Top four concentrations of CPA versus DMSO (4 μ L/mL)

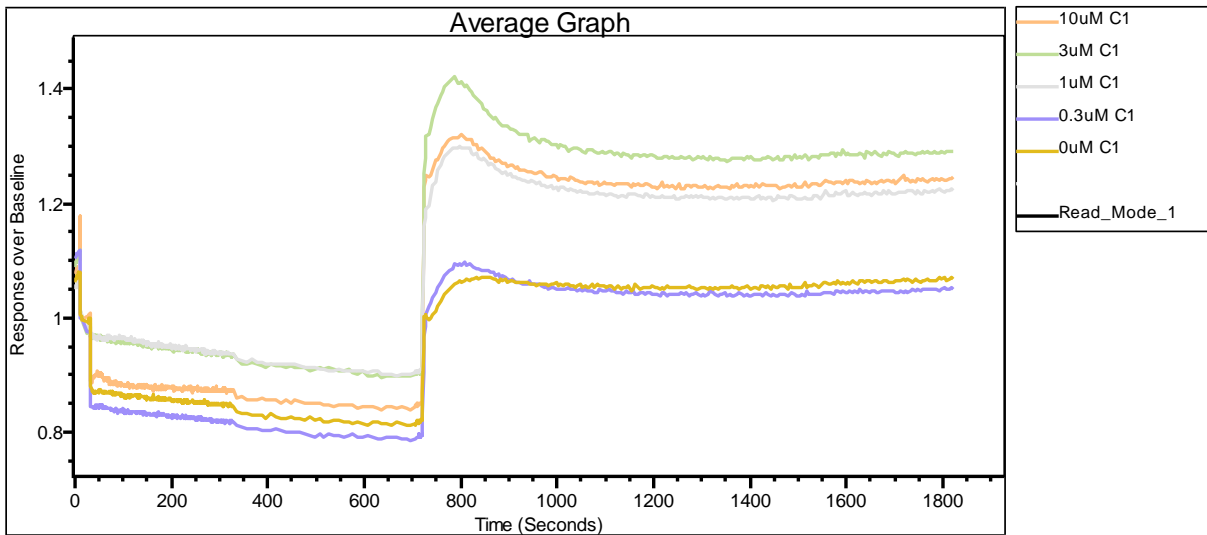


Figure 37 Top four concentrations of Compound **8** (C1 or ZINC04189857) versus DMSO (4 μ L/mL)

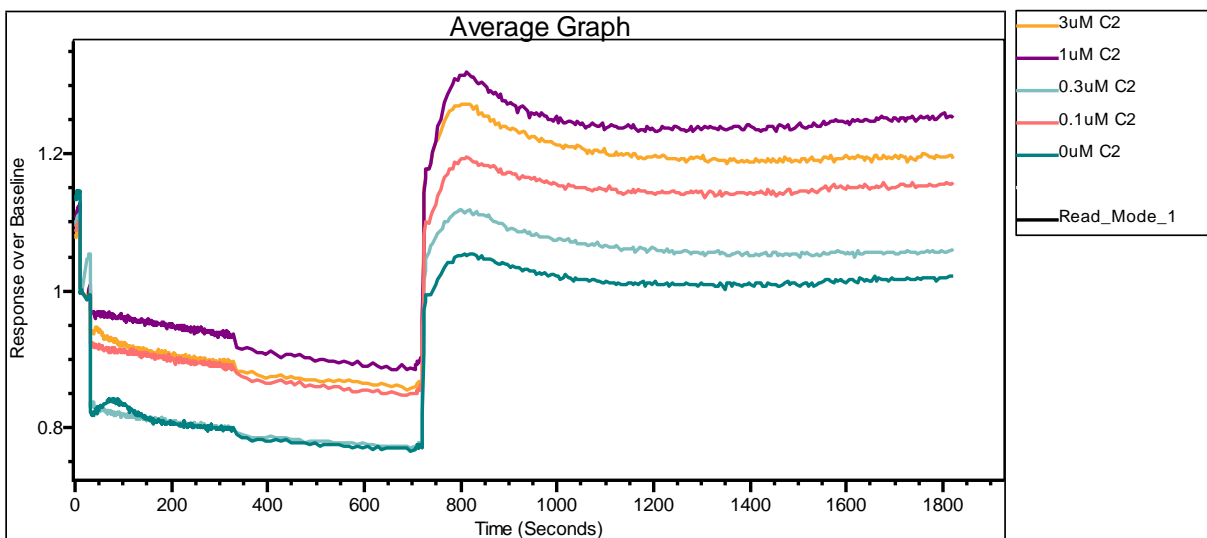


Figure 38 Top four concentrations of Compound **3** (C2 or ZINC08846492) versus DMSO (4 μ L/mL)

3.3.2 Results for compounds 10, 13, 15, 16 and 17

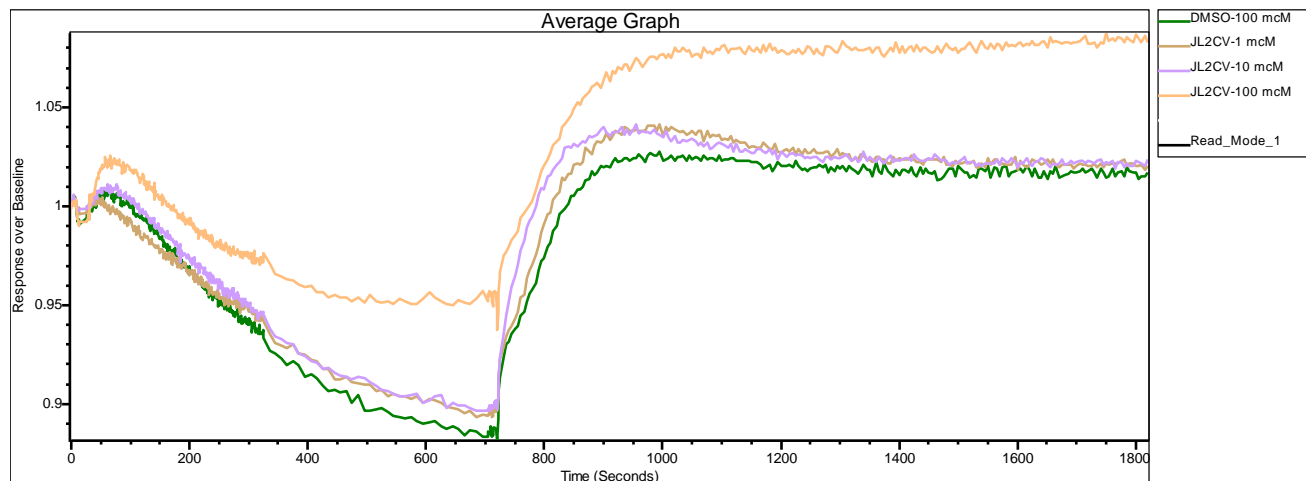


Figure 39 Ca²⁺ response for compound **17** (JL-2CV)

Figure 39 provided the Ca²⁺ response for compound **17**. At the beginning, there seemed to be an increase of the Ca²⁺ signalling but in fact this increase was due to compound addition artefact from the FLIPR. The dark green line represented the negative control, DMSO. The rest of the three lines represented three concentrations of compound **17** in 1 μ M, 10 μ M and 100 μ M used in the FLIPR testing. Overall, there were no significant changes or increase in Ca²⁺ signalling of compound **17** when compared with DMSO.

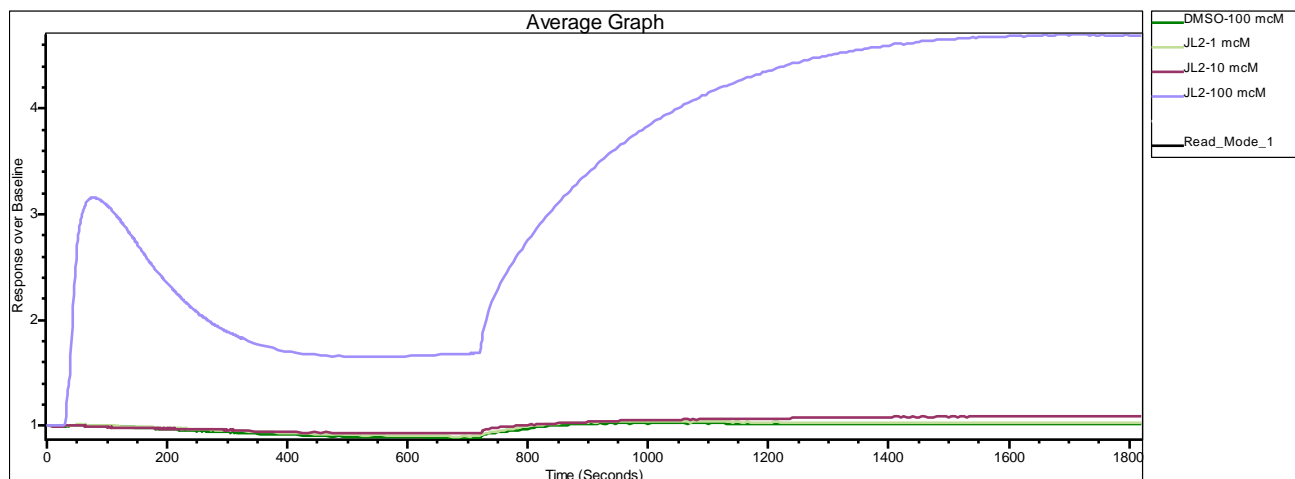


Figure 40 Ca^{2+} response for compound **13** (JL-2)

The **Figure 40** showed the Ca^{2+} signalling profile of compound **13** compared with DMSO at 100 μM . Three different concentrations of compound **13** were used, starting with the lowest concentration of 1 μM and then increasing to 10 μM and 100 μM . DMSO was represented by the dark green line. Both 1 μM and 10 μM of compound **13** did not generate any increased Ca^{2+} signalling when compared with DMSO. At 100 μM of compound **13**, a significantly higher Ca^{2+} was observed by a first Ca^{2+} peak and also a second Ca^{2+} peak which was not recovered but sustained throughout the process. The sustained second Ca^{2+} peak was different from the Ca^{2+} signalling profile from CPA, where the second Ca^{2+} peak would recover back to baseline response. On the other hand, a contrast in the Ca^{2+} response was seen in compound **15** at 100 μM from compound **13** in **Figure 41**. All three concentrations of compound **15** did not produce prominent Ca^{2+} response as that from CPA at 10 μM . These concentrations were in fact more in line with the effect from DMSO.

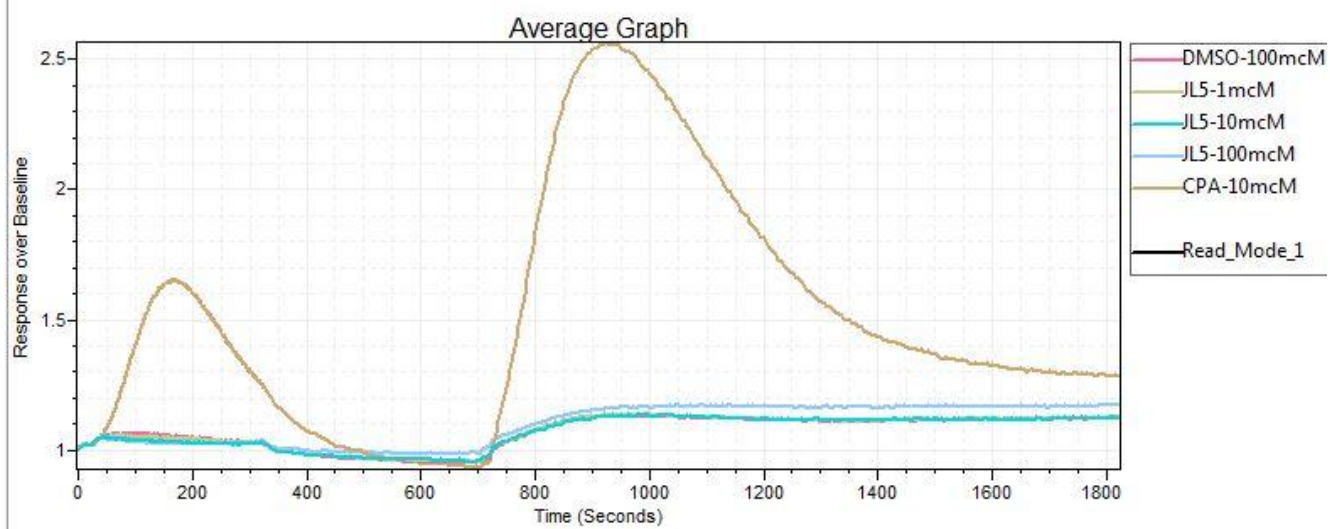


Figure 41 Ca^{2+} response for compound **15** (JL-5)

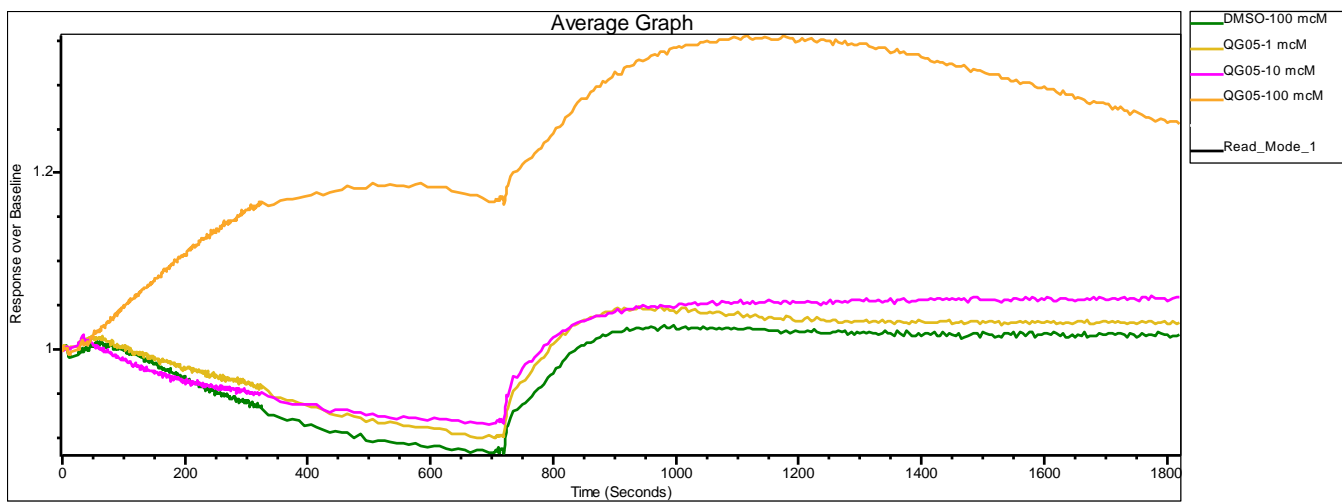


Figure 42 Ca^{2+} response for compound **10** (QG-05)

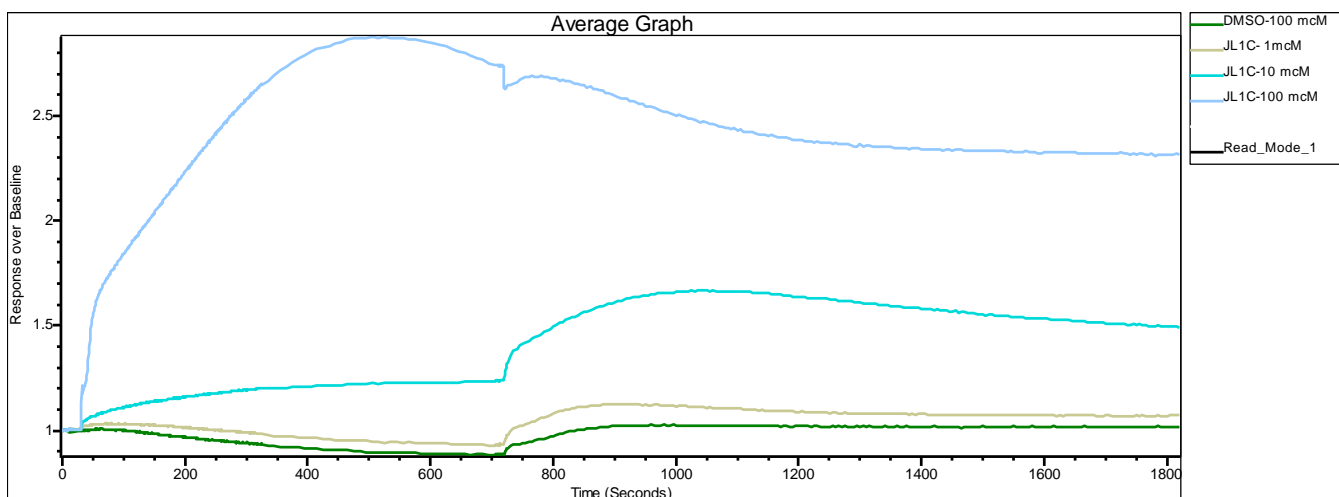


Figure 43 Ca^{2+} response for compound **16** (JL-1C)

In **Figure 42**, DMSO was shown in dark green line and compared with three other concentrations of compound **10**. The Ca^{2+} signalling profiles for both 1 μM and 10 μM of compound **10** were very similar to the Ca^{2+} response from DMSO. Only the 100 μM of compound **10** produced a higher Ca^{2+} response. There was a higher first Ca^{2+} peak and the second peak had a slight recovery towards the end of 1800 seconds. For **Figure 43**, compound **16** was tested in three concentrations, 1 μM , 10 μM and 100 μM . Compound **16** in 1 μM had not much difference to DMSO but with 10 μM of compound **16**, a higher first and second Ca^{2+} peaks were shown. At 100 μM of compound **16**, an even higher Ca^{2+} response was shown than at 10 μM .

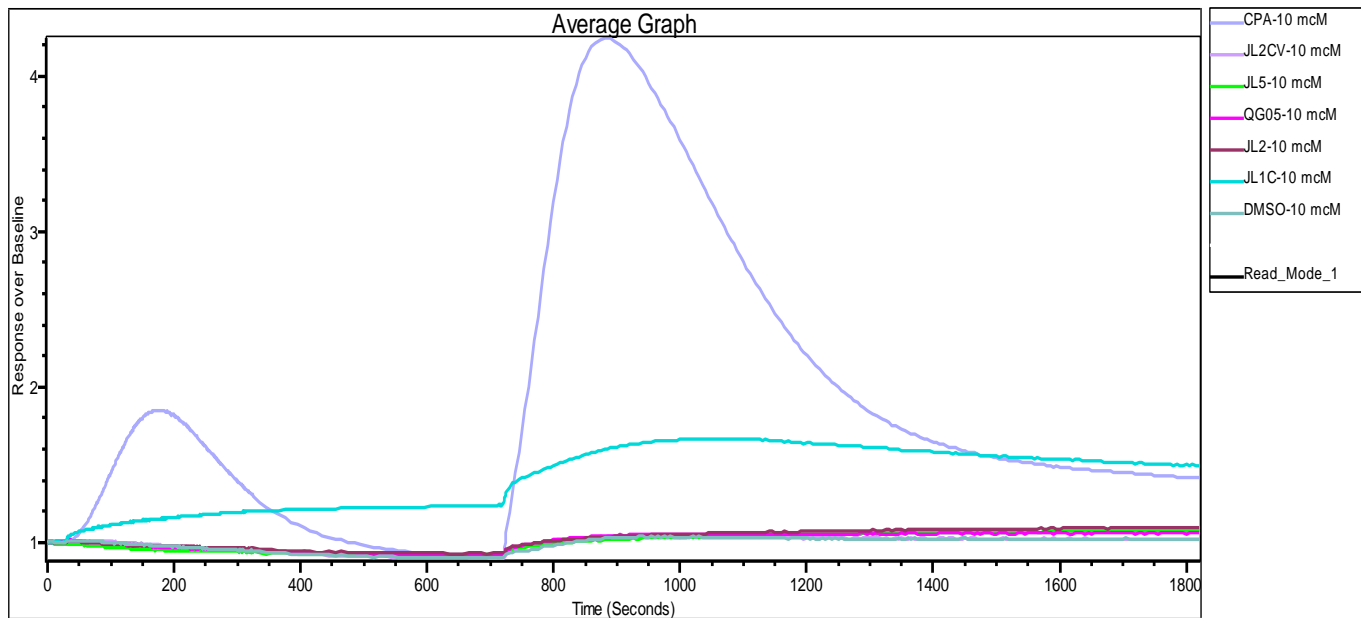


Figure 44 Ca²⁺ response for all five tested compounds at 10 μM

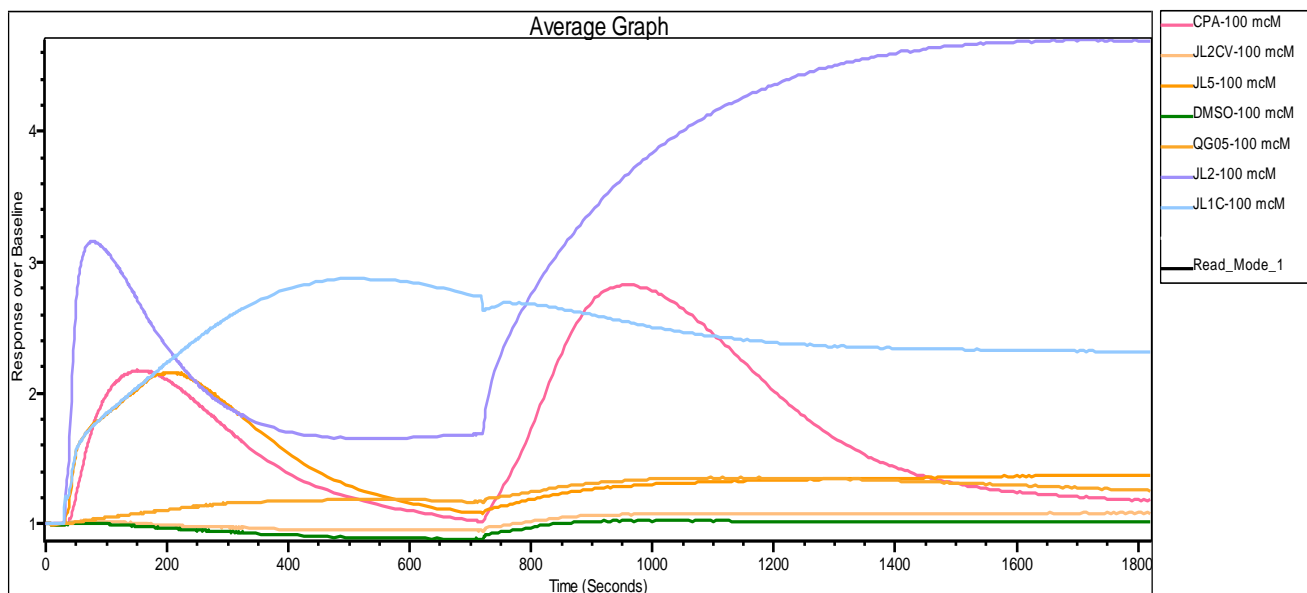


Figure 45 Ca²⁺ response for all five tested compounds at 100 μM

In **Figure 44** comparing all five compounds with CPA and DMSO at 10 μM , it was clearly shown that compound **16** had a higher Ca^{2+} response than the rest of the tested compounds including DMSO. CPA itself gave the standard Ca^{2+} signalling profile with the highest response out of all compounds tested. The comparison of all five compounds with CPA and DMSO at 100 μM was given in **Figure 45**. There were four out of five tested compounds that had higher Ca^{2+} responses than compounds **17**, **10** and DMSO. Compound **13** had higher Ca^{2+} response than CPA whereas the first Ca^{2+} peak response for compound **16** was higher than CPA but its second Ca^{2+} peak did not exceed that from CPA.

3.4 DISCUSSION

3.4.1 Evaluation of molecular modelling results

3.4.1.1 SPCA

The molecular modelling process began with ascertaining the ligand binding site of a well-known SERCA inhibitor, TG, on the first Ca^{2+} ATPase model, SPCA2. Therefore the TG binding site was used on this model for the initial virtual screening of the TimTec compounds derived from the ZINC database.

The molecular docking began from screening all downloaded Tim Tec compounds on the SPCA2 model. The number of subjobs was reduced from 16 to 6 in the HTVS docking because there were only limited numbers of Glide licence available for the project. If the number of

subjobs exceeded the number of Glide licence present, the running of the subjobs would be terminated which meant no results would be available after the Glide docking. Consequently this might reduce the number of ligands with desirable G scores for further investigation. The number of subjobs used in the entire docking process was kept at 6 unless otherwise stated to avoid variations in the docking results between different Ca²⁺ ATPase models.

A Venn diagram was compiled after completing all three stages of docking modes, from HTVS to SP and finally the XP docking mode. The purpose of the Venn diagram was to differentiate the ligand-interacting amino acid residues on the SPCA2 model. It was essential to identify the exact ligand binding site by knowing which amino acid residues were involved. Altogether there were six amino acid residues found after analysing the ligand-interacting residues from 40 Tim Tec ligands. This would assist in finding the desirable ligand binding sites for further compound docking on the SPCA1d model and subsequently on all three SERCA homology models. Three of the six amino acid residues (ILE765, PHE256 and PHE834) were randomly chosen and set as the ligand binding site and XP mode docking trialled for the second time on the 40 Tim Tec ligands. This was to confirm this binding site would give at least similar G scores of -10 or below. At the end, 26 out of the 40 compounds gave this G score.

For the SPCA1d model, the first ligand binding grid was set on LEU765, LEU256 and ILE834. The only difference here was that the amino acid residues were different from the one used for the second trial for the SPCA2 model, which had isoleucine (ILE) and two phenylalanines (PHE) on position 765, 256 and 834 respectively. Four other different ligand binding grids were also set

up as described in details in the Results Section. A combination of three amino acid residues with either same amino acid residues or same residue position number as the previously found six residues from the Venn diagram was used. None of the docking results reached the G score of at least -10 as before. This could be attributed to the difference in the ligand binding site set with variable amino acid residues and positions present. The molecular binding and interaction would change drastically from the previous ligand-receptor bindings on the SPCA2 model. Subsequently this could affect the G scores obtained in the selected Tim Tec ligands. This outcome was reflected in the docking from the SPCA1d model.

In an attempt to resolve the issues with the lower G scores obtained from SPCA1d docking, SiteMap was employed to identify new druggable sites on the SPCA1d model. Three new binding sites were found and the 26 Tim Tec ligands previously found to have relatively high G scores were trialled for docking on these three sites. The G scores reached for all three binding sites did not actually reach -10 with the highest of only about -7.9 for SiteMap-derived site 3. Oppositely, another trial was also carried out on SPCA2 model to test these three SiteMap-derived sites. The specific grids were described in details in the Results section. The highest G score obtained here was about -11 from SiteMap site 3. The rest of the ligands had lower G scores than -11 which were also the same for the result from the docking on SPCA1d model. The possibilities of not acquiring the expected high G scores could be due to two main reasons. Firstly, only the 26 found Tim Tec ligands were used for ligand docking on the three SiteMap-derived sites for both SPCA2 and SPCA1d models. The outcome could have been different if the original whole five sets of Tim Tec ligands were subjected to HTVS, SP and XP mode dockings. Secondly, the accuracy of the homology models might not have reflected the real-life protein

models of SPCA2 and SPCA1d. At the time of constructing these homology models, there were no actual crystal structures of SPCA2 and SPCA1d available from human source. This was therefore taken into account to the best possible extent by building these SPCA2 and SPCA1d models based on human-sourced protein sequences.

3.4.1.2 SERCA

Homology models for all three isoforms of SERCA were constructed based on human-sourced SERCA protein sequences. The same three SiteMap-derived sites were used for SERCA1, 2 and 3 models. Again a combination of amino acid residues at the same positions were used for setting up the ligand binding grids but the amino acid residues might not be exactly the same as the initial ones found from SPCA1d model. The highest G score found from XP docking of SERCA1 model was about -11.7 from the original TG-based site derived from SPCA2 docking.

The same four binding sites were also applied to SERCA2 model, which included the original TG-equivalent site. XP docking was done for all four sites with the same set of 26 Tim Tec ligands. The highest G score was about -11.9 from SiteMap-derived site 3. For SERCA3 model, the same docking process was applied on the four binding sites. The highest G score was about -11.5 from the TG-equivalent site. Overall, all three SERCA models produced similar docking results.

3.4.1.3 Evaluation of CombiGlide results

The main reason of adopting and trialling CombiGlide in the molecular docking process was to attempt maximising the number of compound hits with G scores of about -10 or below. This programme suite was designed to allow substitutions of side-chain groups or so-called “fragments” of the chemical compounds in the *in silico* context. This allowed more time and flexibility without putting in high expenditure into the chemical synthesis in order to pre-empt the first-hand effect of substituted compound on the desired receptor target. However, downfalls of this technique could be that it was nonetheless only an *in silico* simulation and not a real-life representation of the compound-receptor binding so the result from this application might not lead to the expected experimental outcome.

CombiGlide was applied to the six compounds selected after the completion of all the docking processes for SPCA2, SPCA1d and all three SERCA models. There were in total 667 fragments downloaded from the Schrödinger website. Two compounds (ZINC08846492 and ZINC08846618; **Figure 15**) were selected from the docking process for SPCA1d model and subjected to fragment substitutions in CombiGlide. An average G score of -7.78 was obtained from the CombiGlide trial on SPCA1d model. The highest G score obtained was -8.972 from ZINC08846618. For SPCA2 model, there were also two Tim Tec ligands selected (ZINC09089406 and ZINC66088673; **Figure 18**) for fragment substitutions. A higher average G score of -9.35 was reached for the compounds chosen from SPCA2 docking than from SPCA1d model. The highest G score was -10.399 from ZINC66088673.

The compound selected after evaluation of the docking results from SERCA1 and SERCA2 models was ZINC04189857 (**Figure 19**), which was subjected to fragment substitutions. The average G score resulted was -10.0. The highest G score found was -10.889. For SERCA3 model, the selected ligand was ZINC32216656 (**Figure 19**). The highest G score was -9.869 and an average G score of -9.68 were achieved. Overall, after close evaluation, none of the G scores obtained from CombiGlide were higher than the original G scores reached from previous dockings. Therefore, it was decided not to change any of the six compounds selected previously for experimental testing.

One critical factor influencing the docking result of the substituted compounds on the targeted receptor model was that the fragment substitution point was randomly chosen. This meant that there was no systematic or consistent point of substitution of each fragment onto the selected ligand. This undoubtedly would affect the docking results to a certain degree, such as variability in G scores from XP mode docking. Potentially a change in the ligand-receptor binding affinities and energies due to changes in molecular interactions would also be reflected if this was specifically investigated. However, due to the time constraint and project aims, this was not particularly looked into. Therefore it would be more valuable to investigate the impact of substituting fragments at set points or randomly selected points on a chosen compound structure and evaluate the outcomes both *in silico* and experimentally.

3.4.2 Evaluation of the results from tested compounds

3.4.2.1 Compound 8 and compound 3

Since CombiGlide trial did not produce any superior G scores than previously obtained, it was decided to test two of the compounds out of the chosen six which aimed to target SPCA1d and SERCA1 models. The final chosen compounds were ZINC04189857 (compound **8**) and ZINC08846492 (compound **3**). Compound **8** was selected to target SERCA1 and Compound **3** was to target SPCA1d. There were no compounds chosen to target SPCA2 since this was not the exact target SPCA isoform for basal-like breast cancer. Hence, one compound was chosen to target SPCA1d and another compound was selected to target SERCA1.

At the initial testing stage and also due to time constraint of the project, both compounds **8** and **3** were purchased from the commercial chemical supplier, Tim Tec. Approximately 20mg of both compound **8** and **3** were purchased and samples of about 1 mg of each compounds were examined by NMR spectroscopy to ensure the compounds purity. The NMR spectrums for both compounds **8** and **3** confirmed their purity.

The first biological testing of these compounds involved using FLIPR to measure the Ca^{2+} signalling activities within the MDA-MB-231 cells after compound treatment. Our positive control in this project was CPA, a well-known SERCA inhibitor in the literature.^{38,40} The negative control used in the experiment was DMSO which was used at a concentration of 4 $\mu\text{L}/\text{mL}$ for each well of the 96-well plate. Ca^{2+} ions were added before 800 seconds in order to

show whether the calcium pumps in the sarcoplasmic reticulum were functioning as expected theoretically in the breast cancer cells when Ca^{2+} ions were present intracellularly. Similar fluorescent-based intracellular calcium level testings were also adopted in other previous studies.^{153,165} The result would provide first-hand comparison between the tested agents and the control agonist, CPA.

The result outputs from FLIPR were shown as a series of Ca^{2+} response curves over time (see all figures in Results Section). All the data obtained from FLIPR have been normalised at baseline to ensure all treatment groups started at the same time point. CPA itself showed a standard two- Ca^{2+} peak curves. The first Ca^{2+} peak came from the inhibition of the SERCA which prevented the Ca^{2+} to be sequestered into the sarcoplasmic reticulum. Hence these Ca^{2+} were present within the cells and would be measured and presented as the peak on the Ca^{2+} response curves. However other Ca^{2+} channels, such as PMCAs, would still be functioning at the time when SERCA was inhibited by CPA. Therefore, some Ca^{2+} would be able to gain entry into the organelles present within the cells such as sarcoplasmic reticulum or Golgi network. In this case, the level of Ca^{2+} would then slowly drop. This process was represented by the first peak in the Ca^{2+} response curves. The second Ca^{2+} peak arrived after more Ca^{2+} ions were added into the cells to demonstrate a higher intracellular Ca^{2+} level than the first peak to help differentiating the different Ca^{2+} effects inside the breast cancer cells.

When compound **8** was added to the MDA-MB-231 cells, there was no significant first Ca^{2+} peak effect when compared to that from CPA. **Figure 27**, a zoom-in shot for first calcium peak,

showed relatively flat Ca^{2+} responses for compound **8**. This finding also applied to effect from compound **3** where there was no obvious Ca^{2+} peaks observed. Only the 10 μM , 3 μM and 1 μM CPA have demonstrated the first Ca^{2+} peak. It could also be noticed that before 350 seconds, the Ca^{2+} response readings were in zigzag lines for all three treatment groups and not as smooth as the readings after 350 seconds. This was due to the FLIPR reading the Ca^{2+} signals once every second before 350 seconds. After this time point, FLIPR was reading the Ca^{2+} signals once every ten seconds which was the default setting of the FLIPR.

Figure 28 had the enlargement of the second Ca^{2+} peaks for all concentrations tested for the three agonists mentioned. A prominent difference could be seen between the Ca^{2+} peaks from CPA and compounds **8** and **3**. Again the same three Ca^{2+} concentrations for CPA provided the significant higher Ca^{2+} peaks which gradually recovered as time passed. The fourth highest Ca^{2+} peak seen for CPA was the 0.3 μM which could just be seen slightly above the rest of the flat Ca^{2+} response curves from compounds **8** and **3**. It could also be observed that there were a subtle increase in Ca^{2+} signals for both compounds **8** and **3** treatment groups. This effect could be confirmed from **Figure 35** which compared the Ca^{2+} signals between the three agonist groups with DMSO. Only the top two concentrations for all three agonist treatment groups were compared with DMSO, which were represented by 0 μM CPA, 0 μM C1 and 0 μM C2 in the figure.

Figure 30 demonstrated the top four concentrations of all three agonist treatment groups. It compared between CPA, compounds **8** and **3** in concentrations of 10 μM , 3 μM , 1 μM , 0.3 μM

and 0.1 μM . Compound **3** had poor solubility in DMSO and hence the highest concentration reachable was at 3 μM . Here the zoom-in graphs of the first and second Ca^{2+} peaks clearly showed a much weaker Ca^{2+} response for both compounds **8** and **3** as discussed above. This again could be confirmed in **Figures 33** and **34** where both compounds **8** and **3** only produced a very low and weak Ca^{2+} response when compared individually with CPA.

3.4.2.1.1 Limitations from molecular modelling and FLIPR testing

One interesting aspect of the Ca^{2+} signalling effects shown in **Figure 28** was that the Ca^{2+} response for 3 μM CPA was higher than 10 μM CPA. In the expected outcome, the Ca^{2+} response would have been higher for the higher CPA concentration group than a lower concentration treatment. This could be due to the inevitable variations in the processes and handling of cell culture and plating, reagent plate preparations and fluorescent dye treatment. Throughout the process of cell culture and plating, it could be possible to have plated the cells unevenly in some of the wells in the 96-well cell plate. Uneven cells in the well could lead to variations in the FLIPR fluorescent signal readings. Another probable reason could be due to human error which might involve mixing up the 10 μM and 3 μM CPA well additions during the reagent plate preparation process. It was also possible that there were remnants of the Fluo-4-AM dye in the 3 μM CPA wells even though the plate was washed at least twice throughout the process. In this case this could lead to a higher reading of the Ca^{2+} signals in the 3 μM CPA wells.

From the calcium response curves obtained for CPA, compounds **8** and **3**, it was evident to suggest that both compound **8** and **3** did not perform as expected. This could be attributed to two major factors where the accuracy of molecular modelling and the assay method used in this project might have contributed to this outcome. The compound selection for biological testing was influenced heavily from the molecular modelling. It was noticed that choosing compounds based on G scores and MM-GBSA results might not be adequate as docking scores were not considered as one of the factors during the compound selection process. It was also observed that in some cases of the ligand-receptor docking, the docking scores and G scores were exactly the same but in some other cases they differed to a large degree. Docking scores were almost like G scores, except they also took into account of the Epik penalty of the binding complex. For this reason, to further improve the accuracy of choosing the compound with highest potential to target the desirable molecular target, it would be necessary to also take into account the docking scores for each ligand-receptor bindings during the ligand selection process.

Because of the unavailability of SPCA1d crystal structure at the beginning of the project, *in silico* homology model of SPCA1d had to be built and generated. The accuracy of the SPCA1d homology model to the real-life SPCA1d protein crystal structure was not possible to be confirmed at the time of the project. However it would be fair to point out that this SPCA1d homology model was likely not a true reflection of real SPCA1d protein structure since it was only an *in silico* simulation of the protein structure. This would also have an impact on identifying the correct target binding sites on the homology model since the identified binding sites for the screened ligands might not be the actual ligand binding sites on the model.¹⁶⁶ As a result, there was a high degree of inaccuracy in the attempt to design effective SPCA1d

inhibitors. The only factor that might give the project results with some more credibility was that human-sourced SPCA1d protein sequence was used to produce the SPCA1d homology model. Then a SERCA orthologue was chosen as the structure base for the SPCA1d homology model. The accuracy of the SPCA1d receptor model would be greatly enhanced and improved if there were human-sourced crystal structures for SPCA1d available in the future.

Another important aspect of molecular modelling limitation was the LigPrep process for the ligands selected for docking. All the Tim Tec ligands downloaded from the ZINC database have not been subjected to LigPrep prior to docking which began from HTVS to SP and lastly to XP mode as discussed previously. This might have potentially affected the virtual screening result which might have reduced the number of potential compounds for biological testing. However, the LigPrep step was not undertaken previously since it was also suggested by Schrödinger that ZINC database usually provided the prepared 3D databases which used Epik for state generation therefore those structures could be used directly for virtual screening such as in Glide. Nevertheless, it was also later understood that in order to achieve the most optimal results, it would be wiser to prepare all ligands with LigPrep prior to any docking process.

It was also worth noting that even if LigPrep was applied to each of the downloaded ligand, there would still be some limitations present to the LigPrep application itself. LigPrep has not been subjected to extensive testing for preparing ligands other than largely organic and drug-like molecules¹³⁸. Some degrees of uncertainty would exist after the LigPrep process was applied to the ligands outside this molecule class. It was possible that some of the downloaded Tim Tec

ligands might be outside the tested molecule class and hence would lose the chance of being submitted into the docking process. Another limitation would be that some of the input structures might be lost during the LigPrep process due to difficult *in silico* structure preparation. Problematic structures would thus not be reflected in the output structure file and this would reduce the number of ligands for the docking process. The file size of input ligand structures could also limit the number of structures for the computers to process. The larger the number of input structures such as 150,000 ligands or above might place a considerable strain on the computer operating system and also the hardware. In this project, about 271,965 compounds were screened. However, not all of the screened compounds were subjected to LigPrep all at once but splitted in sets of five.

Another area for improvement involved refining and checking the compound stereochemistry. Even though LigPrep process would take into account the stereochemical factors but as mentioned above, the ZINC database-derived Tim Tec compounds were not put through the LigPrep process. This also meant that the stereochemistry of all Tim Tec ligands was overlooked during the molecular modelling step. As suggested by Brooks *et. al.*, it was recommended to check the stereochemistry of the ligand-like compounds prior to virtual screening process.¹³⁹ If the aspect of checking compound stereochemistry was neglected, there was a potential risk of losing a significant proportion of potential lead compounds. This would thus deeply impact the outcome of the biological assays. It was also noteworthy to consider another scenario where even if LigPrep was used for all the downloaded ligands, it would also be useful to check the stereochemistry of the ligands to see if the application has corrected the compound stereochemistry.

3.4.2.2 Compounds 10, 13, 15, 16 and 17

The second cycle of virtual screening began after the first two compounds were selected and tested in FLIPR. The structures of compounds **8** and **3** and CPA were reviewed and compared. A common structural theme to both CPA and compound **3** were deduced and found. The structural motif was shown as compound **9**. It was a pyrrolidone-like core structure that could undergo keto-enol tautomerism. For this reason, during the initial attempt to conduct the second cycle of virtual screening, the aim was to search for presence of this particular core structure in chemical compounds from databases such as SciFinder, PubChem and ZINC. All the structures found from all the databases mentioned previously were submitted to Glide XP docking on SPCA1d model. All ligand-like compounds derived from the databases were also prepared *via* LigPrep prior to any docking activity.

An exact structure of the pyrrolidone-like core structure was searched and found from SciFinder. Another compound that had the pyrrolidone-like core structure was also found from Pubmed. However, both compounds were subjected to XP mode docking and did not generate satisfactory G scores as expected and also due to G scores were ranked in a relative term, it would be necessary to carry out other compound dockings in order to make more comparisons. While more compounds were searched from ZINC database containing the pyrrolidone structure-like core, BREED algorithm was trialled in an effort to generate more *in silico* compounds. BREED algorithm did this *via* producing hybrid molecules from selected compounds submitted by the user. One of the limitations with this method was that the actual chemical synthesis might be

difficult for the hybrid molecule such as restriction from complex compound stereochemistry and thus might affect the purity and yield of the compound. This was also the major reason for not testing one of the hybrid molecules in the FLIPR even though a reasonable G score was obtained at -7.166. Also the number of hybrid molecule formations might also be limited by whether if the ligands had swappable fragments, substituents or bonds that could be mixed and matched to ensure chemically and structurally-fit hybrid compounds.¹⁶³

At this stage the focus of the second cycle of molecular docking has shifted to developing a SERCA inhibitor due to unsatisfactory G scores obtained from above mentioned structures with pyrrolidone-like core rather than attempting to design a SPCA1d inhibitor. Another important reason of changing the focus was also to ensure a viable and effective SERCA inhibitor could be designed and developed where the same drug design process could be applied to design a SPCA1d inhibitor without doubts. This would also ensure an effective drug design process was established on the SERCA model which was already well-studied in the past and current literature than the more newly found SPCA.

Since G scores were comparative terms, it was necessary to dock CPA on the SERCA1 model to distinguish the G score generated from CPA with the other trialled compounds. In an *in silico* environment, the G score produced from the docking of CPA on SERCA1 model would also behave as a control to be compared with the rest of the trialling compounds. As pointed out in the results section, the highest G score reached by CPA was on Sitemap-generated site 3 at -6.786. A total of 202 hybrid molecules derived from CPA and the 26 chemical structures with

less pyrrolidone-like core downloaded from ZINC database were also generated. The docking of these hybrid molecules had a higher G score of -7.650 than both the highest G score of docking CPA alone and also the highest G score from docking the 26 less pyrrolidone-like structures on the SERCA1 model.

Artemisinin was also trialled for XP docking on the SERCA1 model with the same three binding sites as mentioned before. The major reason for this trial was due to the discovery of artemisinin inhibiting the SERCA orthologue of *Plasmodium falciparum* in *Xenopus oocytes*.¹⁶⁴ Subsequently it would be useful to investigate the effect of artemisinin on human SERCA. However after trialling XP docking on three Sitemap-generated sites, the G score reached was not as high as all the compounds trialled so far in the second round of virtual screening. Later it was considered to generate a receptor model based on the original SERCA orthologue of *Plasmodium falciparum* which was also known as PfATP6. This time a crystal structure actually existed for PfATP6 which was found from PDB database with an ID of 1U5N. This crystal structure was directly imported into the Maestro interface and used. The main purpose of this was to observe any distinguished drug selectivity between the protozoal and human SERCAs. However, the result of docking artemisinin onto 1U5N actually produced a slightly lower G score than the one obtained from above. Both G scores obtained were below -3.5 which were much lower than all the previously obtained G scores in the second cycle virtual screening. The difference in the two highest G scores from both human and protozoal SERCA models in the artemisinin trial was only by a score of -1.109 which did not demonstrate observable drug selectivity between the two models.

One of the few earliest studies tested and found the anti-tumour activity of curcumin *in vitro* using Chinese Hamster Ovary cells and *in vivo* in mice model with Dalton's lymphoma.¹⁶⁷ Another study also discovered the ability of curcumin to reduce the production of nitric oxide which was known to be involved in inflammatory and carcinogenic processes in cells.¹⁶⁸ Curcumin was also known to have a similar effect on the Ca^{2+} -uptake activity in ER as the SERCA inhibitors, TG and CPA, which related to its anti-cancer property.¹⁶⁹ It was proposed that curcumin was likely to inhibit the proprotein convertase activity by inhibiting the Ca^{2+} activity mediated by SERCA and thus exhibited the anti-cancer property to the three colon cancer cell lines. Due to this evidence, curcumin in a keto-enol form was submitted for docking trial. On the other hand, a cyclic pyrazole analogue of curcumin was also prepared for XP docking. The differences between the compounds could be seen in schemes 3 and 4. The highest G scores between the two structures were quite close in that the G score of -5.968 for the cyclic pyrazole derivative of curcumin was slightly higher than the G score of -5.626 for curcumin itself.

BREED algorithm was again applied to the keto-enol form of curcumin in combination with artemisinin and the 26 less pyrrolidone-like structures found from ZINC database. A higher G scores was observed from the hybrid molecule than the highest G scores obtained from docking of curcumin in the keto-enol form. The combination of compounds used for BREED algorithm was then expanded to include CPA and cyclic pyrazole curcumin analogue. However this did not generate any new hybrid molecule as the highest G score obtained was from the same hybrid molecule generated from the previous BREED combination of CPA and the 26 less pyrrolidone-

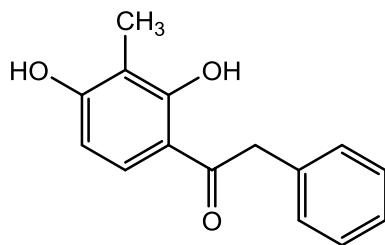
like structures. Nevertheless, as mentioned above, none of the hybrid molecules were selected for FLIPR testing due to difficulties discussed in the chemical synthesis process.

Diallyl bisphenol A was also sent for docking trial due to the known ability of bisphenols to inhibit SERCA.^{153,170} One of the studies aimed to find out the structural-activity relationship of the inhibitory effect for 27 commercially-available bisphenols on SERCA. The major finding was that the presence of a variety of non-polar substituents on the two phenyl rings in bisphenols provided the most potent inhibitory activity of these compounds on SERCA. Examples of potent non-polar substituents were small alkyl groups such as allyl side chain used in the docking trial or halide groups. It was also found that methyl or cyclohexyl groups at the central carbon atom connecting the two phenyl rings would also enhance the potency of the compound. This was however not trialled in the docking trial in this round due to time constraint of the project. Another interesting study also tested the effect of bisphenol or also known as bis(2-hydroxy-3-tert-butyl-5-methyl-phenyl)-methane and 2-aminoethoxydiphenylborate on SPCA1d and SERCA2b.¹⁶⁵ It found that these two compounds did selectively inhibit SPCA1d more than SERCA2b. Bisphenol A was also tested in a similar fashion but it was found to be inhibiting both SPCA1d and SERCA2b at a similar level without distinguished selectivity over one or another. Since the current aim in this round of virtual screening was to focus on finding and developing a SERCA inhibitor, therefore bisphenol A with two allyl side chains was selected to be docked onto the SERCA1 model. As suggested by Woeste *et. al.*, the hydrophobic interactions provided from the hydrophobic residues present on SERCA plus the hydrogen bonding from ASP254 were the two key factors in the chemical interaction between the bisphenol inhibitors and SERCA. Hence, the bisphenol A with two allyl groups was docked onto

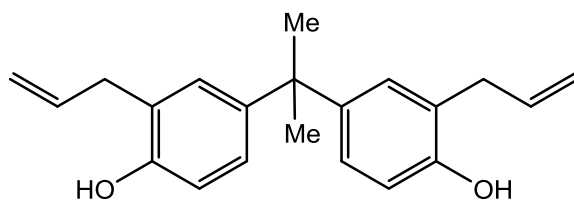
a set grid based on these two factors as mentioned in the results section. Another binding site was also trialled which was the original CPA-binding site and also the three Sitemap-generated binding sites were also trialled. The highest G score of -6.348 reached for this bisphenol A compound was based on the Sitemap-generated site rather than the CPA binding site and also the set binding site on ASP254, ILE97 and PRO312.

In the second cycle of compound selection for FLIPR testing, G scores were one of the deciding factors as past literature evidence on the inhibitory effect of bisphenols was also used in the selection process. The real outcome still lied within the actual biological experiments therefore it was decided to test the bisphenol A with the two allyl side groups, curcumin plus its cyclic pyrazole derivative. It was also worth mentioning that the G scores obtained for these three compounds were in fact higher than that from the CPA docking trial. It was also decided that soluble salt forms of bisphenol A should be considered since at this stage, chemical synthesis was the next step to acquire these compounds rather than obtaining from commercial source. The main reason to produce the soluble salt forms of the bisphenol A compound was due to FLIPR testing normally used DMSO as the solvent and also as the negative control, therefore the testing compound should at least be soluble in DMSO to aid the smooth running of FLIPR. Therefore, compound **15** was synthesised in the laboratory which was also useful to demonstrate any result difference from compound **13**. Compound **10** had a similar pyrrolidone-like core present in its structure which would be another potential candidate for FLIPR testing. Therefore it was decided that five compounds in total would be synthesised and tested in FLIPR to see their effect on Ca^{2+} signalling within the MDA-MB-231 cells.

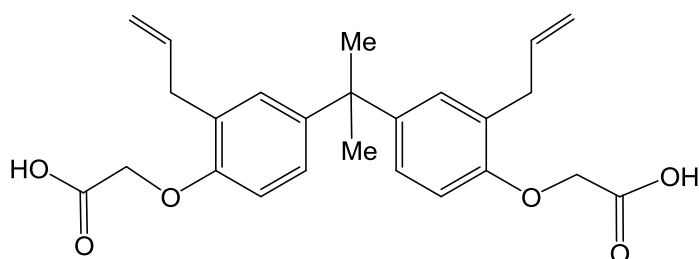
The chemical synthesis process has taken similar amount of time as the molecular modelling stage. The production of compounds **16** and **17** only involved a short amount of time than the making of compound **15** due to the ability to conduct the synthesis by microwave irradiation. The manual column chromatography required more time input for separating the desired cyclic pyrazole analogue of curcumin from the unwanted chemicals. Throughout the process of making all five compounds, constant TLC monitoring were taken place to ensure mid-reaction compounds were actually correct and wanted for the end product. Final products for each compound were also submitted for NMR spectroscopy, mass spectroscopy (for compounds **15** and **17**), IR (for compound **15**, **16** and **17** only) and X-ray diffraction for compound **16** only to ensure purity.



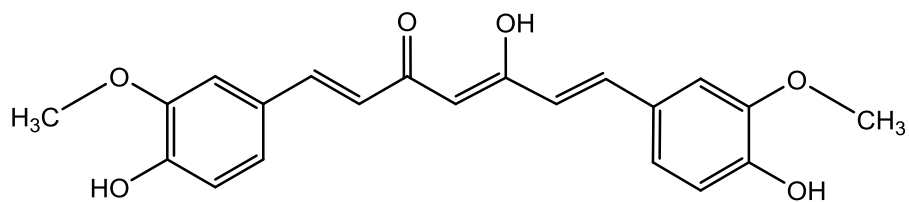
Compound 10 QG-05



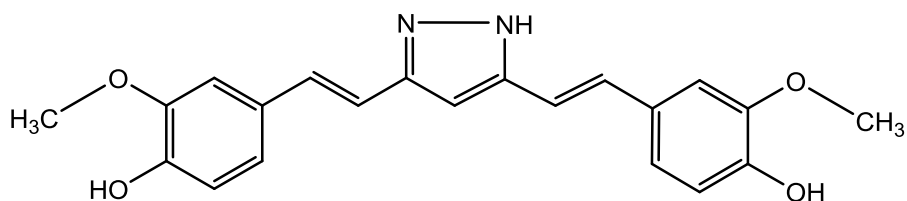
Compound 13 JL-2



Compound 15 JL-5



Compound 16 JL-1C



Compound 17 JL-2CV

Figure 46 Chemical structures of compounds **10**, **13**, **15**, **16** and **17**

These five compounds, shown in **Figure 46**, have each demonstrated different Ca^{2+} response effects in the second batch of compound testing. Each compound was tested in three incremental concentrations from 1 μM , 10 μM and 100 μM . As pointed out previously, **Figure 39** representing the Ca^{2+} response from compound **17** showed an initial addition artefact where there was a small increase in Ca^{2+} signals. Overall, it could be confirmed that this compound did not affect the Ca^{2+} response in the breast cancer cells at all. On the other hand, compound **16** acted differently to compound **17** even though compound **16** at 1 μM also did not change the Ca^{2+} response and had a similar effect to DMSO. However, for 10 μM of compound **16**, a higher Ca^{2+} response than DMSO was observed and again at the highest concentration of 100 μM , a significant increase of Ca^{2+} response was seen in the **Figure 43**. It was interesting to note that at 10 μM , the first Ca^{2+} peak was not as prominent as the one for 100 μM . The second Ca^{2+} peak however at 10 μM showed an increase in Ca^{2+} signal after addition of Ca^{2+} took place and the peak slowly recovered but not completely. At 100 μM , the first Ca^{2+} peak also did not recover completely and quite a sharp reduction in Ca^{2+} could be noticed at around the time when more Ca^{2+} were added to create the second Ca^{2+} peak. Again, the same incomplete Ca^{2+} recovery curve was observed. This phenomenon could implicate that compound **16** acted more like a PMCA inhibitor rather than a SERCA inhibitor. This could also be explained by the Ca^{2+} response profile produced here by compound **16** at 100 μM which did not resemble exactly as the profile for CPA, one of the well-studied SERCA inhibitors. Compound **16** was likely a moderately potent PMCA inhibitor since it required a concentration at 100 μM to elicit the higher Ca^{2+} response rather than at a smaller concentration.

3.4.2.2.1 Limitations from molecular modelling

Between compounds **16** and **17**, the result from the FLIPR test did not reflect the modelling result attained earlier where the highest G score for compound **17** was in fact higher than that for compound **16**. This showed that the outcome from structure-based virtual screening might not always be reflected in the actual biological assays. Virtual screening was highly dependent on two essential factors; the quantity and quality of the available data used and predictive or discriminatory ability of the underlying screening algorithm.¹⁷¹ One of the virtual screening pitfalls was unrealistic assumptions and expectations. It was common to assume that the binding pose produced on the software screen after docking was the correct binding pose for the docked compound on target receptor. However, it was also highly possible that even though high G score was assigned to a particular ligand-receptor binding pose, the actual binding pose shown was in fact false. It was even mentioned that some drug discovery successes using structure-based virtual screening might be due to pure serendipity. Another limitation within structure-based screening was compound selection. This process was entirely based on subjectivity, even though G scores and MM-GBSA results were used to help with selection process. This process potentially could introduce bias as unselected compounds were not sent for testing and might produce unexpected results. This scenario could likely apply to the first round of compound selection where only two out of six trialled compounds were sent for FLIPR testing, the other four compounds missed out the opportunity to be validated due to financial reason and time constraints.

Another pitfall related to the size and diversity of the chemical compound libraries used.¹⁷¹ As this was only a small-scale project, only one compound library was used with the total number of compounds screened less than 300,000. Therefore it could also explain why the virtual screening results were not reflected highly in the bioassay since only one small-size and less diversified library of compounds was screened. In an ideal situation, more than one chemical compound library should be used at least with special attention paid on selecting a wide range of different types or classes of compounds to increase better chance of finding the hit compounds.

Another likely problem in the virtual screening used in this project could come from the vendor libraries which provided free-access to a variety of compound libraries. These libraries sometimes might also contain molecules with chemically active groups or other unwanted functional groups that would cause interference to the high throughput screening technique used in the virtual screening software.¹⁷¹ This situation could not be totally ruled out for the compound library used in this project and could contribute to the result discrepancy between the virtual screening and FLIPR testing.

As mentioned previously, this project was run on a small-scale basis therefore the only virtual screening method adopted was structure-based rather than ligand-based. Ideally, it would produce more potential compound candidates and likely more accurate results if more than one screening method was used and then to use the results collectively from all the methods used.¹⁷¹ It was also worth to point out another downside within the virtual screening process was that

most of the dockings done on each receptor model were only done once and not repeated. If each docking was at least repeated once, this could potentially produce different docking outcomes.

The other pair of compounds tested was compound **13** and **15**. Compound **13** was the original diallyl bisphenol A and compound **15** was its analogue with two acetic acid groups attached to the benzene rings. From the Ca^{2+} response curve shown for compound **13**, a noticeable incomplete Ca^{2+} recovery could be seen for the second Ca^{2+} peak at 100 μM . A similar case could also be seen for the first Ca^{2+} peak where there was also a slight incomplete Ca^{2+} recovery. The Ca^{2+} responses for the other two concentrations unfortunately did not differ hugely from DMSO, set as the negative control in the second round of FLIPR testing. Consequently since compound **13** only began to show more prominent Ca^{2+} signals at 100 μM rather than at 10 μM or below, it was probably also a moderately potent Ca^{2+} pump inhibitor. The full Ca^{2+} response profile indicated that compound **13** might inhibit both the SERCA and PMCA pumps. By analysing both Ca^{2+} peaks, compound **13** might have inhibited PMCA more than SERCA since its first Ca^{2+} peak almost resembled that from CPA but had an incomplete Ca^{2+} recovery. This could probably mean that some of the PMCA pumps were inhibited by compound **13** and were not available for the transit of Ca^{2+} into the ER. This could be interpreted by the incomplete Ca^{2+} recovery to indicate some Ca^{2+} still remained in the intracellular space. The second peak had a sustained Ca^{2+} response which was different from CPA's complete recovery in its second Ca^{2+} peak. Again, the features observed would likely lead to the compound behaving more like a PMCA pump inhibitor than SERCA inhibitor. Nevertheless, compound **15** showed an opposite Ca^{2+} response at 100 μM when compared to compound **13**. At 1 μM and 10 μM , similar Ca^{2+} response effects were observed in both compounds **13** and **15**. At these two concentrations, both

compounds **13** and **15** did not induce any Ca^{2+} response in the MDA-MB-231 cells. However, a difference in Ca^{2+} effect was observed at 100 μM where compound **15** did not elicit any Ca^{2+} response at all.

Compound **10** also only showed an increased Ca^{2+} response at 100 μM . At 1 μM and 10 μM , the Ca^{2+} response were about the same as DMSO. However, overall the Ca^{2+} signals for compound **10** at 100 μM were still lower than CPA at 100 μM . There was a similar incomplete Ca^{2+} recovery for the first peak which was also the same for second Ca^{2+} peak. Since the compound was only showing effect at 100 μM , it could also be concluded that it was not a very potent Ca^{2+} pump inhibitor. From the overall Ca^{2+} signalling profile, compound **10** might also act more like a PMCA inhibitor than SERCA inhibitor. This was based on a sustained second Ca^{2+} peak which might mean that some of the Ca^{2+} added were not allowed transit into the SR via PMCA since it was likely inhibited by compound **10**.

3.4.2.2.2 Potency and structural aspects of tested compounds

Compound **16** was the most potent out of all five compounds as it had elicited Ca^{2+} signals at 10 μM where as other compounds only had prominent Ca^{2+} signals at 100 μM . This was shown in **Figures 44** and **45**. At 100 μM (**Figure 45**), only two out of the five compounds had explicit higher Ca^{2+} response on top of CPA acting as the positive control. These compounds were compound **16** and **13**. However, for compounds at 10 μM or below (**Figure 44**), only compound **16** managed to produce a higher intracellular Ca^{2+} response than the rest of the tested compounds.

One of the probable explanations for the dichotomous effect between compound **13** and **15** could be their structural differences at the phenolic hydroxyl position. Compound **13** possessed the original hydroxyl group whereas compound **15** had the acetic acid group substituted at the hydroxyl position. Compounds **13** and **16** all had the phenol group present within their structures. Compound **10** had one extra hydroxyl group attached to the phenol group already present. Therefore the presence of phenol might be important to provide some probable inhibitory effect of these compounds on SERCA or PMCA. The acetic acid group within compound **15** would theoretically increase the solubility of the diallyl bisphenol A which would be considered as a useful trait during drug design process but in this case, its presence might have hindered the possible Ca^{2+} -eliciting response in these basal-like breast cancer cells.

The results from compound **16** and **17** also gave an interesting opposite effect which could possibly be explained by their structural differences. Compound **17** did not induce any Ca^{2+} response where compound **16** was the most potent compound out of the five compounds tested. The possible inactivity from compound **17** might be due to the presence of the cyclic pyrazole structure instead of the keto-enol group in compound **16** since this was the only difference between the two compounds. Therefore from the results shown it might be possible that the presence of the phenol and keto-enol group within compound **16** provided the highest possible potency out of all tested compounds. However in order to determine the exact structure-activity relationship, more compound analogues would need to be synthesised and tested to confirm this theory thoroughly.

Repeat FLIPR testings were also conducted for compounds **13**, **16**, **17** and **10**. Compound **15** was currently under the process of repeat testing. Consistent outcomes were seen for compounds **16**, **17** and **10** except **13** (see appendix D). The Ca^{2+} effect for both first and second Ca^{2+} peaks were much lower than the first test for compound **13** as shown from the FLIPR graphs. The most probable reason for this inconsistent repeat result for compound **13** could be due to compound hydrolysis which would have caused a loss in compound effect.

Thus in terms of selecting the most optimal compound candidate for further drug development at this stage, it could be concluded that compound **16** had the most probable potential to be further developed to become lead compounds targeting SERCA. Both compounds **17** and **15** did not have any effect on the Ca^{2+} response in the breast cancer cells therefore would not be suitable candidates for lead compound development. Compound **10** only had the higher Ca^{2+} response at 100 μM but when comparing its overall Ca^{2+} response profile; the Ca^{2+} signals were lower than both compounds **13** and **15**. Therefore, compound **10** might not be one of the most ideal compounds to be further refined for lead development.

3.4.2.3 FLIPR

To decide if an unknown compound would be either a SERCA or SPCA inhibitor in the cells, a control such as CPA was used in the experiment to compare the effect of the tested compound on the intracellular Ca^{2+} response. Finding out the likelihood of what the tested compounds could be

was the major purpose of conducting the FLIPR tests in this project. From the results, it would then be possible to decide if the unknown compound was acting like a SERCA, SPCA or PMCA inhibitor. SERCA inhibitor should in theory produce a first Ca^{2+} peak resembling to that from CPA. SPCA inhibitor on the other hand should have a smaller first Ca^{2+} peak since there was no other main Ca^{2+} channels present on the Golgi apparatus apart from SPCA therefore the Ca^{2+} leakage will be smaller than that of ER. For PMCA inhibitors, the intracellular Ca^{2+} response would be more sustained and not recovered since ER Ca^{2+} store could still release Ca^{2+} into the cytoplasm maintaining the Ca^{2+} level.

The whole process of FLIPR began with the addition of BAPTA, an intracellular Ca^{2+} -chelating agent. This was to ensure there was no intracellular Ca^{2+} present prior to the addition of the chemical compounds in the MDA-MB-231 cells. Then CPA or one of the tested compounds was added. The addition of Ca^{2+} then took place at around 700 seconds. The purpose of the FLIPR test was to see if the compound selected for testing resembled the Ca^{2+} signalling profile CPA had on the MDA-MB-231 cells. Once this was confirmed, it would be a sound evidence for using the same drug designing process to design an agent to inhibit SPCA1d.

The principle of the Ca^{2+} response seen in CPA could be explained by knowing that SERCA was inhibited by CPA as it was added into the cells. This meant no more Ca^{2+} could be sequestered into the ER. This also meant that the normal Ca^{2+} leak from ER could also not be re-sequestered. Hence this would cause an increase in the intracellular Ca^{2+} level. A measure of this Ca^{2+} level was done by FLIPR to provide the first Ca^{2+} peak shown in the graph. However since there were

also PMCA and other Ca^{2+} channels present on ER, this would lead to a slow entrance of the intracellular Ca^{2+} into the ER and thus a reduction in the Ca^{2+} response over the baseline. As mentioned previously, more Ca^{2+} were added at around 700 seconds by the FLIPR robot. This would lead to an increase in intracellular Ca^{2+} level but since the effect of CPA would wear off a little therefore small amount of the added Ca^{2+} would be sequestered by SERCA into the ER with also the access through PMCA and other Ca^{2+} channels, therefore the second Ca^{2+} peak would recover at the end.

The downside of the FLIPR test itself was that it was not designed to be a specific test to differentiate between different calcium pumps such as SERCA, SPCA and PMCA. Therefore, the results of the FLIPR test were only preliminary at this stage to inform the likely possibility of which compounds might possess the potential to inhibit either SERCA or SPCA1d. Since the assay method adopted in this project was understandably inadequate consequently further testings such as doing ATPase assay would likely affirm more promising results. The only probable argument of using FLIPR as the only assay method in the project was again due to project time constraint and also there were yet not many studies focussing on developing and designing a SPCA1 inhibitor in the context of targeting basal-like breast cancer.

3.4.3 Conclusions and future directions

A combination of using well-prepared chemical database, careful set-up of parameter choices along with any key benchmarking studies to compare screening methods and results would likely lead to better virtual screening outcome. A high-quality virtual screening technique, thorough

literature background search related to the specific drug target and presence of drug design expertise were some key factors required to produce viable lead compounds in drug discovery projects. Therefore, to advance and improve the virtual screening process used in this study, it would be necessary to address all the above mentioned limitations or issues as discussed in order to broaden the first batch of hit compounds. Essentially, virtual screening was used to help reducing costs and time involved in chemical synthesis and biological testings in order to maximise hit compounds in an efficient manner.

The initiative for this project was to find and design a candidate compound that might have the potential to be further refined to inhibit SPCA1d. However towards the halfway of the project after testing the first two virtually screened compounds, this goal has proved to be more difficult to be achieved therefore the aim was then adjusted to focus on finding and designing a compound that would inhibit SERCA since there were known SERCA inhibitors present such as TG and CPA, to be the reference compounds. From the second FLIPR testing of the five compounds, at least three out of the five tested compounds had an effect on the intracellular Ca^{2+} signals than the two tested previously. However, both the potency and selectivity of these three compounds would need to be further improved to become better SERCA inhibitors. Thus increasing the selectivity and potency of these compounds would be the next critical step for future experimental work.

Diallyl Bisphenol A and its derivatives, compounds **13** and **15** respectively, were known to be ubiquitous chemical substances used in numerous industrial polycarbonate productions.¹⁷²

However, since both compounds **13** and **15** were known to disrupt the endocrine system and posed with adverse effects to pregnant women, they might not be the best candidate compounds to enter the next stage of compound refinement.^{172,173} Therefore even though diallyl bisphenol A and its derivative were not suitable lead compound candidates, they might be useful compounds to study the calcium signalling and relevant cellular processes involved within SERCA.

The FLIPR result from testing compound **16**, which was also known as curcumin, might provide a new direction for this particular project due to the improved Ca^{2+} response shown in the breast cancer cells. As discussed previously, curcumin had similar effect as TG and CPA on Ca^{2+} signalling in colon cancer cells, it was also now widely known for its chemo-preventative and chemo-therapeutic effect in addition to numerous effects such as anti-inflammatory, antioxidant, immunomodulatory and anti-angiogenic properties.^{156,169} Hence it would be a suitable compound candidate for further drug development in light of inhibiting SERCA or SPCA1d to target basal-like breast cancer. However, it was also known that curcumin had low *in vivo* bioavailability which might prevent it from being a useful and effective chemotherapeutic agent.¹⁷⁴ Thus improving its pharmacokinetic profile, particularly oral bioavailability, might be another interesting aspect to be investigated in the future.

It was also peculiar to see nil Ca^{2+} effect from compounds **17** and also **15**. This could possibly be explained from a structure-activity relationship point of view as discussed previously. This would be another interesting aspect to find out if further work were to be carried out for this project. For compound **16**, it might be worth submitting the compound to do Ca^{2+} ATPase assay

and cell viability assays to observe its effect on MDA-MB-231 cell proliferation since curcumin was known to have an anticancer effect.

One of the other interesting aspects from the results obtained from testing the five compounds was that some of these compounds appeared to act more like a PMCA inhibitor. As introduced earlier, remodelling of PMCA isoform expressions was also highly linked to some breast or colon cancers. The preliminary result here might open up other different research opportunities to investigate the possibility of some of these compounds to target PMCA isoforms in breast or colon cancer cells.

REFERENCES

- (1) Badve, S.; Dabbs, D. J.; Schnitt, S. J.; Baehner, F. L.; Decker, T.; Eusebi, V.; Fox, S. B.; Ichihara, S.; Jacquemier, J.; Lakhani, S. R.; Palacios, J.; Rakha, E. A.; Richardson, A. L.; Schmitt, F. C.; Tan, P. H.; Tse, G. M.; Weigelt, B.; Ellis, I. O.; Reis-Filho, J. S.: Basal-like and triple-negative breast cancers: a critical review with an emphasis on the implications for pathologists and oncologists. *Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc* **2011**, *24*, 157-67.
- (2) De Summa, S.; Pinto, R.; Sambiasi, D.; Petriella, D.; Paradiso, V.; Paradiso, A.; Tommasi, S.: BRCAness: a deeper insight into basal-like breast tumors. *Annals of oncology : official journal of the European Society for Medical Oncology / ESMO* **2013**, *24 Suppl 8*, viii13-viii21.
- (3) Rakha, E. A.; Reis-Filho, J. S.; Ellis, I. O.: Basal-like breast cancer: a critical review. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **2008**, *26*, 2568-81.
- (4) Vuong, D.; Simpson, P. T.; Green, B.; Cummings, M. C.; Lakhani, S. R.: Molecular classification of breast cancer. *Virchows Archiv : an international journal of pathology* **2014**, *465*, 1-14.
- (5) Prat, A.; Parker, J. S.; Karginova, O.; Fan, C.; Livasy, C.; Herschkowitz, J. I.; He, X.; Perou, C. M.: Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. *Breast cancer research : BCR* **2010**, *12*, R68.
- (6) Rakha, E. A.; El-Sayed, M. E.; Reis-Filho, J.; Ellis, I. O.: Patho-biological aspects of basal-like breast cancer. *Breast cancer research and treatment* **2009**, *113*, 411-22.
- (7) Foulkes, W. D.; Stefansson, I. M.; Chappuis, P. O.; Begin, L. R.; Goffin, J. R.; Wong, N.; Trudel, M.; Akslen, L. A.: Germline BRCA1 mutations and a basal epithelial phenotype in breast cancer. *Journal of the National Cancer Institute* **2003**, *95*, 1482-5.
- (8) Waddell, N.; Arnold, J.; Cocciardi, S.; da Silva, L.; Marsh, A.; Riley, J.; Johnstone, C. N.; Orloff, M.; Assie, G.; Eng, C.; Reid, L.; Keith, P.; Yan, M.; Fox, S.; Devilee, P.; Godwin, A. K.; Hogervorst, F. B.; Couch, F.; Grimmond, S.; Flanagan, J. M.; Khanna, K.; Simpson, P. T.; Lakhani, S. R.; Chenevix-Trench, G.: Subtypes of familial breast tumours revealed by expression and copy number profiling. *Breast cancer research and treatment* **2010**, *123*, 661-77.
- (9) Shastry, M.; Yardley, D. A.: Updates in the treatment of basal/triple-negative breast cancer. *Current opinion in obstetrics & gynecology* **2013**, *25*, 40-8.
- (10) Bayraktar, S.; Gluck, S.: Systemic therapy options in BRCA mutation-associated breast cancer. *Breast cancer research and treatment* **2012**, *135*, 355-66.
- (11) Colleoni, M.; Munzone, E.: Extended adjuvant chemotherapy in endocrine non-responsive disease. *Breast (Edinburgh, Scotland)* **2013**, *22S2*, S161-S164.
- (12) Brini, M.; Carafoli, E.: Calcium pumps in health and disease. *Physiological reviews* **2009**, *89*, 1341-78.
- (13) Carafoli, E.: Calcium signaling: a tale for all seasons. *Proceedings of the National Academy of Sciences of the United States of America* **2002**, *99*, 1115-22.
- (14) Venkatachalam, K.; van Rossum, D. B.; Patterson, R. L.; Ma, H. T.; Gill, D. L.: The cellular and molecular basis of store-operated calcium entry. *Nature cell biology* **2002**, *4*, E263-72.
- (15) Berridge, M. J.; Bootman, M. D.; Roderick, H. L.: Calcium signalling: dynamics, homeostasis and remodelling. *Nature reviews. Molecular cell biology* **2003**, *4*, 517-29.
- (16) Targos, B.; Baranska, J.; Pomorski, P.: Store-operated calcium entry in physiology and pathology of mammalian cells. *Acta biochimica Polonica* **2005**, *52*, 397-409.
- (17) Thastrup, O.; Dawson, A. P.; Scharff, O.; Foder, B.; Cullen, P. J.; Drobak, B. K.; Bjerrum, P. J.; Christensen, S. B.; Hanley, M. R.: Thapsigargin, a novel molecular probe for studying intracellular calcium release and storage. *Agents and actions* **1989**, *27*, 17-23.

- (18) Roberts-Thomson, S. J.; Peters, A. A.; Grice, D. M.; Monteith, G. R.: ORAI-mediated calcium entry: mechanism and roles, diseases and pharmacology. *Pharmacology & therapeutics* **2010**, *127*, 121-30.
- (19) Ebashi, S.: Calcium binding activity of vesicular relaxing factor. *Journal de chirurgie* **1961**, *50*, 236-44.
- (20) Ebashi, S.; Lipmann, F.: Adenosine triphosphate-linked concentration of calcium ions in a particulate fraction of rabbit muscle. *The Journal of cell biology* **1962**, *14*, 389-400.
- (21) Hasselbach, W.; Makinose, M.: [The calcium pump of the "relaxing granules" of muscle and its dependence on ATP-splitting]. *Biochemische Zeitschrift* **1961**, *333*, 518-28.
- (22) Fagan, M. J.; Saier, M. H., Jr.: P-type ATPases of eukaryotes and bacteria: sequence analyses and construction of phylogenetic trees. *Journal of molecular evolution* **1994**, *38*, 57-99.
- (23) Yu, X.; Carroll, S.; Rigaud, J. L.; Inesi, G.: H⁺ countertransport and electrogenicity of the sarcoplasmic reticulum Ca²⁺ pump in reconstituted proteoliposomes. *Biophysical journal* **1993**, *64*, 1232-42.
- (24) Moller, J. V.; Juul, B.; le Maire, M.: Structural organization, ion transport, and energy transduction of P-type ATPases. *Biochimica et biophysica acta* **1996**, *1286*, 1-51.
- (25) Toyoshima, C.: Structural aspects of ion pumping by Ca²⁺-ATPase of sarcoplasmic reticulum. *Archives of biochemistry and biophysics* **2008**, *476*, 3-11.
- (26) Toyoshima, C.: How Ca²⁺-ATPase pumps ions across the sarcoplasmic reticulum membrane. *Biochimica et biophysica acta* **2009**, *1793*, 941-6.
- (27) Toyoshima, C.; Inesi, G.: Structural basis of ion pumping by Ca²⁺-ATPase of the sarcoplasmic reticulum. *Annual review of biochemistry* **2004**, *73*, 269-92.
- (28) Toyoshima, C.; Mizutani, T.: Crystal structure of the calcium pump with a bound ATP analogue. *Nature* **2004**, *430*, 529-35.
- (29) Toyoshima, C.; Nakasako, M.; Nomura, H.; Ogawa, H.: Crystal structure of the calcium pump of sarcoplasmic reticulum at 2.6 Å resolution. *Nature* **2000**, *405*, 647-55.
- (30) Toyoshima, C.; Nomura, H.: Structural changes in the calcium pump accompanying the dissociation of calcium. *Nature* **2002**, *418*, 605-11.
- (31) Toyoshima, C.; Nomura, H.; Sugita, Y.: Crystal structures of Ca²⁺-ATPase in various physiological states. *Annals of the New York Academy of Sciences* **2003**, *986*, 1-8.
- (32) Toyoshima, C.; Nomura, H.; Tsuda, T.: Lumenal gating mechanism revealed in calcium pump crystal structures with phosphate analogues. *Nature* **2004**, *432*, 361-8.
- (33) Toyoshima, C.; Norimatsu, Y.; Iwasawa, S.; Tsuda, T.; Ogawa, H.: How processing of aspartylphosphate is coupled to lumenal gating of the ion pathway in the calcium pump. *Proceedings of the National Academy of Sciences of the United States of America* **2007**, *104*, 19831-6.
- (34) Toyoshima, C.; Yonekura, S.; Tsueda, J.; Iwasawa, S.: Trinitrophenyl derivatives bind differently from parent adenine nucleotides to Ca²⁺-ATPase in the absence of Ca²⁺. *Proceedings of the National Academy of Sciences of the United States of America* **2011**, *108*, 1833-8.
- (35) Albers, R. W.: Biochemical aspects of active transport. *Annual review of biochemistry* **1967**, *36*, 727-56.
- (36) Post, R. L.; Hegyvary, C.; Kume, S.: Activation by adenosine triphosphate in the phosphorylation kinetics of sodium and potassium ion transport adenosine triphosphatase. *The Journal of biological chemistry* **1972**, *247*, 6530-40.
- (37) Wuytack, F.; Raeymaekers, L.; Missiaen, L.: Molecular physiology of the SERCA and SPCA pumps. *Cell calcium* **2002**, *32*, 279-305.
- (38) Seidler, N. W.; Jona, I.; Vegh, M.; Martonosi, A.: Cyclopiazonic acid is a specific inhibitor of the Ca²⁺-ATPase of sarcoplasmic reticulum. *The Journal of biological chemistry* **1989**, *264*, 17816-23.

- (39) Goeger, D. E.; Riley, R. T.; Dorner, J. W.; Cole, R. J.: Cyclopiazonic acid inhibition of the Ca²⁺-transport ATPase in rat skeletal muscle sarcoplasmic reticulum vesicles. *Biochem Pharmacol* **1988**, *37*, 978-81.
- (40) Moncoq, K.; Trieber, C. A.; Young, H. S.: The molecular basis for cyclopiazonic acid inhibition of the sarcoplasmic reticulum calcium pump. *The Journal of biological chemistry* **2007**, *282*, 9748-57.
- (41) Riley, R. T.; Goeger, D. E.; Yoo, H.; Showker, J. L.: Comparison of three tetramic acids and their ability to alter membrane function in cultured skeletal muscle cells and sarcoplasmic reticulum vesicles. *Toxicology and applied pharmacology* **1992**, *114*, 261-7.
- (42) Thastrup, O.; Cullen, P. J.; Drobak, B. K.; Hanley, M. R.; Dawson, A. P.: Thapsigargin, a tumor promoter, discharges intracellular Ca²⁺ stores by specific inhibition of the endoplasmic reticulum Ca²⁺(+)-ATPase. *Proceedings of the National Academy of Sciences of the United States of America* **1990**, *87*, 2466-70.
- (43) Xu, C.; Ma, H.; Inesi, G.; Al-Shawi, M. K.; Toyoshima, C.: Specific structural requirements for the inhibitory effect of thapsigargin on the Ca²⁺ ATPase SERCA. *The Journal of biological chemistry* **2004**, *279*, 17973-9.
- (44) Sagara, Y.; Inesi, G.: Inhibition of the sarcoplasmic reticulum Ca²⁺ transport ATPase by thapsigargin at subnanomolar concentrations. *The Journal of biological chemistry* **1991**, *266*, 13503-6.
- (45) Yu, M.; Zhong, L.; Rishi, A. K.; Khadeer, M.; Inesi, G.; Hussain, A.: Specific substitutions at amino acid 256 of the sarcoplasmic/endoplasmic reticulum Ca²⁺ transport ATPase mediate resistance to thapsigargin in thapsigargin-resistant hamster cells. *The Journal of biological chemistry* **1998**, *273*, 3542-6.
- (46) Wootton, L. L.; Michelangeli, F.: The effects of the phenylalanine 256 to valine mutation on the sensitivity of sarcoplasmic/endoplasmic reticulum Ca²⁺ ATPase (SERCA) Ca²⁺ pump isoforms 1, 2, and 3 to thapsigargin and other inhibitors. *The Journal of biological chemistry* **2006**, *281*, 6970-6.
- (47) Brody, I. A.: Muscle contracture induced by exercise. A syndrome attributable to decreased relaxing factor. *The New England journal of medicine* **1969**, *281*, 187-92.
- (48) Benders, A. A.; Veerkamp, J. H.; Oosterhof, A.; Jongen, P. J.; Bindels, R. J.; Smit, L. M.; Busch, H. F.; Wevers, R. A.: Ca²⁺ homeostasis in Brody's disease. A study in skeletal muscle and cultured muscle cells and the effects of dantrolene and verapamil. *The Journal of clinical investigation* **1994**, *94*, 741-8.
- (49) Karpati, G.; Charuk, J.; Carpenter, S.; Jablecki, C.; Holland, P.: Myopathy caused by a deficiency of Ca²⁺-adenosine triphosphatase in sarcoplasmic reticulum (Brody's disease). *Annals of neurology* **1986**, *20*, 38-49.
- (50) Burge, S. M.; Wilkinson, J. D.: Darier-White disease: a review of the clinical features in 163 patients. *Journal of the American Academy of Dermatology* **1992**, *27*, 40-50.
- (51) Brini, M.; Cali, T.; Ottolini, D.; Carafoli, E.: The plasma membrane calcium pump in health and disease. *The FEBS journal* **2013**, *280*, 5385-97.
- (52) Schatzmann, H. J.: ATP-dependent Ca⁺⁺-extrusion from human red cells. *Experientia* **1966**, *22*, 364-5.
- (53) Niggli, V.; Adunyah, E. S.; Carafoli, E.: Acidic phospholipids, unsaturated fatty acids, and limited proteolysis mimic the effect of calmodulin on the purified erythrocyte Ca²⁺ - ATPase. *The Journal of biological chemistry* **1981**, *256*, 8588-92.
- (54) Szewczyk, M. M.; Pande, J.; Grover, A. K.: Caloxins: a novel class of selective plasma membrane Ca²⁺ pump inhibitors obtained using biotechnology. *Pflugers Archiv : European journal of physiology* **2008**, *456*, 255-66.
- (55) Cho, Y. S.; Go, M. J.; Kim, Y. J.; Heo, J. Y.; Oh, J. H.; Ban, H. J.; Yoon, D.; Lee, M. H.; Kim, D. J.; Park, M.; Cha, S. H.; Kim, J. W.; Han, B. G.; Min, H.; Ahn, Y.; Park, M. S.; Han, H. R.; Jang, H. Y.; Cho, E.

Y.; Lee, J. E.; Cho, N. H.; Shin, C.; Park, T.; Park, J. W.; Lee, J. K.; Cardon, L.; Clarke, G.; McCarthy, M. I.; Lee, J. Y.; Oh, B.; Kim, H. L.: A large-scale genome-wide association study of Asian populations uncovers genetic factors influencing eight quantitative traits. *Nature genetics* **2009**, *41*, 527-34.

(56) Lu, X.; Wang, L.; Chen, S.; He, L.; Yang, X.; Shi, Y.; Cheng, J.; Zhang, L.; Gu, C. C.; Huang, J.; Wu, T.; Ma, Y.; Li, J.; Cao, J.; Chen, J.; Ge, D.; Fan, Z.; Li, Y.; Zhao, L.; Li, H.; Zhou, X.; Chen, L.; Liu, D.; Duan, X.; Hao, Y.; Lu, F.; Liu, Z.; Yao, C.; Shen, C.; Pu, X.; Yu, L.; Fang, X.; Xu, L.; Mu, J.; Wu, X.; Zheng, R.; Wu, N.; Zhao, Q.; Liu, X.; Wang, M.; Yu, D.; Hu, D.; Ji, X.; Guo, D.; Sun, D.; Wang, Q.; Yang, Y.; Liu, F.; Mao, Q.; Liang, X.; Ji, J.; Chen, P.; Mo, X.; Li, D.; Chai, G.; Tang, Y.; Li, X.; Du, Z.; Dou, C.; Yang, Z.; Meng, Q.; Wang, D.; Wang, R.; Yang, J.; Schunkert, H.; Samani, N. J.; Kathiresan, S.; Reilly, M. P.; Erdmann, J.; Peng, X.; Chen, R.; Qiang, B.; Gu, D.: Genome-wide association study in Han Chinese identifies four new susceptibility loci for coronary artery disease. *Nature genetics* **2012**, *44*, 890-4.

(57) Schuh, K.; Cartwright, E. J.; Jankevics, E.; Bundschu, K.; Liebermann, J.; Williams, J. C.; Armesilla, A. L.; Emerson, M.; Oceandy, D.; Knobloch, K. P.; Neyses, L.: Plasma membrane Ca²⁺ ATPase 4 is required for sperm motility and male fertility. *The Journal of biological chemistry* **2004**, *279*, 28220-6.

(58) Mohamed, T. M.; Oceandy, D.; Zi, M.; Prehar, S.; Alatwi, N.; Wang, Y.; Shaheen, M. A.; Abou-Leisa, R.; Schelcher, C.; Hegab, Z.; Baudoin, F.; Emerson, M.; Mamas, M.; Di Benedetto, G.; Zaccolo, M.; Lei, M.; Cartwright, E. J.; Neyses, L.: Plasma membrane calcium pump (PMCA4)-neuronal nitric-oxide synthase complex regulates cardiac contractility through modulation of a compartmentalized cyclic nucleotide microdomain. *The Journal of biological chemistry* **2011**, *286*, 41520-9.

(59) Oceandy, D.; Cartwright, E. J.; Emerson, M.; Prehar, S.; Baudoin, F. M.; Zi, M.; Alatwi, N.; Venetucci, L.; Schuh, K.; Williams, J. C.; Armesilla, A. L.; Neyses, L.: Neuronal nitric oxide synthase signaling in the heart is regulated by the sarcolemmal calcium pump 4b. *Circulation* **2007**, *115*, 483-92.

(60) Ficarella, R.; Di Leva, F.; Bortolozzi, M.; Ortolano, S.; Donaudy, F.; Petrillo, M.; Melchionda, S.; Lelli, A.; Domi, T.; Fedrizzi, L.; Lim, D.; Shull, G. E.; Gasparini, P.; Brini, M.; Mammano, F.; Carafoli, E.: A functional study of plasma-membrane calcium-pump isoform 2 mutants causing digenic deafness. *Proceedings of the National Academy of Sciences of the United States of America* **2007**, *104*, 1516-21.

(61) Schutz, M.; Scimemi, P.; Majumder, P.; De Siati, R. D.; Crispino, G.; Rodriguez, L.; Bortolozzi, M.; Santarelli, R.; Seydel, A.; Sonntag, S.; Ingham, N.; Steel, K. P.; Willecke, K.; Mammano, F.: The human deafness-associated connexin 30 T5M mutation causes mild hearing loss and reduces biochemical coupling among cochlear non-sensory cells in knock-in mice. *Human molecular genetics* **2010**, *19*, 4759-73.

(62) Zanni, G.; Cali, T.; Kalscheuer, V. M.; Ottolini, D.; Barresi, S.; Lebrun, N.; Montecchi-Palazzi, L.; Hu, H.; Chelly, J.; Bertini, E.; Brini, M.; Carafoli, E.: Mutation of plasma membrane Ca²⁺ ATPase isoform 3 in a family with X-linked congenital cerebellar ataxia impairs Ca²⁺ homeostasis. *Proceedings of the National Academy of Sciences of the United States of America* **2012**, *109*, 14514-9.

(63) Rudolph, H. K.; Antebi, A.; Fink, G. R.; Buckley, C. M.; Dorman, T. E.; LeVitre, J.; Davidow, L. S.; Mao, J. I.; Moir, D. T.: The yeast secretory pathway is perturbed by mutations in PMR1, a member of a Ca²⁺ ATPase family. *Cell* **1989**, *58*, 133-45.

(64) Okorokov, L. A.; Tanner, W.; Lehle, L.: A novel primary Ca(2+)-transport system from *Saccharomyces cerevisiae*. *European journal of biochemistry / FEBS* **1993**, *216*, 573-7.

(65) Sorin, A.; Rosas, G.; Rao, R.: PMR1, a Ca²⁺-ATPase in yeast Golgi, has properties distinct from sarco/endoplasmic reticulum and plasma membrane calcium pumps. *The Journal of biological chemistry* **1997**, *272*, 9895-901.

(66) Van Baelen, K.; Vanoevelen, J.; Missiaen, L.; Raeymaekers, L.; Wuytack, F.: The Golgi PMR1 P-type ATPase of *Caenorhabditis elegans*. Identification of the gene and demonstration of calcium and manganese transport. *The Journal of biological chemistry* **2001**, *276*, 10683-91.

- (67) Missiaen, L.; Vanoevelen, J.; Parys, J. B.; Raeymaekers, L.; De Smedt, H.; Callewaert, G.; Erneux, C.; Wuytack, F.: Ca²⁺ uptake and release properties of a thapsigargin-insensitive nonmitochondrial Ca²⁺ store in A7r5 and 16HBE14o- cells. *The Journal of biological chemistry* **2002**, *277*, 6898-902.
- (68) Ton, V. K.; Mandal, D.; Vahadji, C.; Rao, R.: Functional expression in yeast of the human secretory pathway Ca(2+), Mn(2+)-ATPase defective in Hailey-Hailey disease. *The Journal of biological chemistry* **2002**, *277*, 6422-7.
- (69) Varki, A.: Factors controlling the glycosylation potential of the Golgi apparatus. *Trends in cell biology* **1998**, *8*, 34-40.
- (70) Kaufman, R. J.; Swaroop, M.; Murtha-Riel, P.: Depletion of manganese within the secretory pathway inhibits O-linked glycosylation in mammalian cells. *Biochemistry* **1994**, *33*, 9813-9.
- (71) Van Baelen, K.; Vanoevelen, J.; Callewaert, G.; Parys, J. B.; De Smedt, H.; Raeymaekers, L.; Rizzuto, R.; Missiaen, L.; Wuytack, F.: The contribution of the SPCA1 Ca²⁺ pump to the Ca²⁺ accumulation in the Golgi apparatus of HeLa cells assessed via RNA-mediated interference. *Biochemical and biophysical research communications* **2003**, *306*, 430-6.
- (72) Van Baelen, K.; Dode, L.; Vanoevelen, J.; Callewaert, G.; De Smedt, H.; Missiaen, L.; Parys, J. B.; Raeymaekers, L.; Wuytack, F.: The Ca²⁺/Mn²⁺ pumps in the Golgi apparatus. *Biochimica et biophysica acta* **2004**, *1742*, 103-12.
- (73) Zhang, H. Z.; Tian, H. Q.; Du, D. H.; Wang, G. J.; Yan, X. X.; Liu, H.; Zhou, G. Z.; Fu, X. A.; Yu, Y. X.; Yu, G. Q.; Liu, H. X.; Zhang, F. R.: Analysis of ATP2C1 gene mutations in Chinese patients with Hailey-Hailey disease. *Clinical and experimental dermatology* **2012**, *37*, 190-3.
- (74) Vanoevelen, J.; Dode, L.; Van Baelen, K.; Fairclough, R. J.; Missiaen, L.; Raeymaekers, L.; Wuytack, F.: The secretory pathway Ca²⁺/Mn²⁺-ATPase 2 is a Golgi-localized pump with high affinity for Ca²⁺ ions. *The Journal of biological chemistry* **2005**, *280*, 22800-8.
- (75) Hu, Z.; Bonifas, J. M.; Beech, J.; Bench, G.; Shigihara, T.; Ogawa, H.; Ikeda, S.; Mauro, T.; Epstein, E. H., Jr.: Mutations in ATP2C1, encoding a calcium pump, cause Hailey-Hailey disease. *Nature genetics* **2000**, *24*, 61-5.
- (76) Behne, M. J.; Tu, C. L.; Aronchik, I.; Epstein, E.; Bench, G.; Bikle, D. D.; Pozzan, T.; Mauro, T. M.: Human keratinocyte ATP2C1 localizes to the Golgi and controls Golgi Ca²⁺ stores. *The Journal of investigative dermatology* **2003**, *121*, 688-94.
- (77) Lee, J. M.; Davis, F. M.; Roberts-Thomson, S. J.; Monteith, G. R.: Ion channels and transporters in cancer. 4. Remodeling of Ca(2+) signaling in tumorigenesis: role of Ca(2+) transport. *American journal of physiology. Cell physiology* **2011**, *301*, C969-76.
- (78) Roberts-Thomson, S. J.; Curry, M. C.; Monteith, G. R.: Plasma membrane calcium pumps and their emerging roles in cancer. *World journal of biological chemistry* **2010**, *1*, 248-53.
- (79) Liu, L. H.; Boivin, G. P.; Prasad, V.; Periasamy, M.; Shull, G. E.: Squamous cell tumors in mice heterozygous for a null allele of Atp2a2, encoding the sarco(endo)plasmic reticulum Ca²⁺-ATPase isoform 2 Ca²⁺ pump. *The Journal of biological chemistry* **2001**, *276*, 26737-40.
- (80) Prasad, V.; Boivin, G. P.; Miller, M. L.; Liu, L. H.; Erwin, C. R.; Warner, B. W.; Shull, G. E.: Haploinsufficiency of Atp2a2, encoding the sarco(endo)plasmic reticulum Ca²⁺-ATPase isoform 2 Ca²⁺ pump, predisposes mice to squamous cell tumors via a novel mode of cancer susceptibility. *Cancer research* **2005**, *65*, 8655-61.
- (81) Okunade, G. W.; Miller, M. L.; Azhar, M.; Andringa, A.; Sanford, L. P.; Doetschman, T.; Prasad, V.; Shull, G. E.: Loss of the Atp2c1 secretory pathway Ca(2+)-ATPase (SPCA1) in mice causes Golgi stress, apoptosis, and midgestational death in homozygous embryos and squamous cell tumors in adult heterozygotes. *The Journal of biological chemistry* **2007**, *282*, 26517-27.
- (82) Curry, M. C.; Luk, N. A.; Kenny, P. A.; Roberts-Thomson, S. J.; Monteith, G. R.: Distinct regulation of cytoplasmic calcium signals and cell death pathways by different plasma membrane

calcium ATPase isoforms in MDA-MB-231 breast cancer cells. *The Journal of biological chemistry* **2012**, *287*, 28598-608.

(83) Curry, M. C.; Roberts-Thomson, S. J.; Monteith, G. R.: Plasma membrane calcium ATPases and cancer. *BioFactors (Oxford, England)* **2011**, *37*, 132-8.

(84) Lee, W. J.; Roberts-Thomson, S. J.; Monteith, G. R.: Plasma membrane calcium-ATPase 2 and 4 in human breast cancer cell lines. *Biochemical and biophysical research communications* **2005**, *337*, 779-83.

(85) Lee, W. J.; Robinson, J. A.; Holman, N. A.; McCall, M. N.; Roberts-Thomson, S. J.; Monteith, G. R.: Antisense-mediated inhibition of the plasma membrane calcium-ATPase suppresses proliferation of MCF-7 cells. *The Journal of biological chemistry* **2005**, *280*, 27076-84.

(86) Aung, C. S.; Kruger, W. A.; Poronnik, P.; Roberts-Thomson, S. J.; Monteith, G. R.: Plasma membrane Ca²⁺-ATPase expression during colon cancer cell line differentiation. *Biochemical and biophysical research communications* **2007**, *355*, 932-6.

(87) Tse, C.; Shoemaker, A. R.; Adickes, J.; Anderson, M. G.; Chen, J.; Jin, S.; Johnson, E. F.; Marsh, K. C.; Mitten, M. J.; Nimmer, P.; Roberts, L.; Tahir, S. K.; Xiao, Y.; Yang, X.; Zhang, H.; Fesik, S.; Rosenberg, S. H.; Elmore, S. W.: ABT-263: a potent and orally bioavailable Bcl-2 family inhibitor. *Cancer research* **2008**, *68*, 3421-8.

(88) Grice, D. M.; Vetter, I.; Faddy, H. M.; Kenny, P. A.; Roberts-Thomson, S. J.; Monteith, G. R.: Golgi calcium pump secretory pathway calcium ATPase 1 (SPCA1) is a key regulator of insulin-like growth factor receptor (IGF1R) processing in the basal-like breast cancer cell line MDA-MB-231. *The Journal of biological chemistry* **2010**, *285*, 37458-66.

(89) Glide: v5.8, Schrödinger. LLC, New York, NY, , 2012.

(90) Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; Shenkin, P. S.: Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *Journal of medicinal chemistry* **2004**, *47*, 1739-49.

(91) Halgren, T. A.; Murphy, R. B.; Friesner, R. A.; Beard, H. S.; Frye, L. L.; Pollard, W. T.; Banks, J. L.: Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. *Journal of medicinal chemistry* **2004**, *47*, 1750-9.

(92) Friesner, R. A.; Murphy, R. B.; Repasky, M. P.; Frye, L. L.; Greenwood, J. R.; Halgren, T. A.; Sanschagrin, P. C.; Mainz, D. T.: Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. *Journal of medicinal chemistry* **2006**, *49*, 6177-96.

(93) Eldridge, M. D.; Murray, C. W.; Auton, T. R.; Paolini, G. V.; Mee, R. P.: Empirical scoring functions: I. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. *Journal of computer-aided molecular design* **1997**, *11*, 425-45.

(94) Levitt, D. G.; Banaszak, L. J.: POCKET: a computer graphics method for identifying and displaying protein cavities and their surrounding amino acids. *Journal of molecular graphics* **1992**, *10*, 229-34.

(95) Laskowski, R. A.: SURFNET: a program for visualizing molecular surfaces, cavities, and intermolecular interactions. *Journal of molecular graphics* **1995**, *13*, 323-30, 307-8.

(96) Peters, K. P.; Fauck, J.; Frommel, C.: The automatic search for ligand binding sites in proteins of known three-dimensional structure using only geometric criteria. *Journal of molecular biology* **1996**, *256*, 201-13.

(97) Hendlich, M.; Rippmann, F.; Barnickel, G.: LIGSITE: automatic and efficient detection of potential small molecule-binding sites in proteins. *Journal of molecular graphics & modelling* **1997**, *15*, 359-63, 389.

(98) Liang, J.; Edelsbrunner, H.; Woodward, C.: Anatomy of protein pockets and cavities: measurement of binding site geometry and implications for ligand design. *Protein science : a publication of the Protein Society* **1998**, *7*, 1884-97.

(99) Brady, G. P., Jr.; Stouten, P. F.: Fast prediction and visualization of protein binding pockets with PASS. *Journal of computer-aided molecular design* **2000**, *14*, 383-401.

(100) Binkowski, T. A.; Naghibzadeh, S.; Liang, J.: CASTp: Computed Atlas of Surface Topography of proteins. *Nucleic acids research* **2003**, *31*, 3352-5.

(101) Dennis, S.; Kortvelyesi, T.; Vajda, S.: Computational mapping identifies the binding sites of organic solvents on proteins. *Proceedings of the National Academy of Sciences of the United States of America* **2002**, *99*, 4290-5.

(102) Kortvelyesi, T.; Silberstein, M.; Dennis, S.; Vajda, S.: Improved mapping of protein binding sites. *Journal of computer-aided molecular design* **2003**, *17*, 173-86.

(103) Ruppert, J.; Welch, W.; Jain, A. N.: Automatic identification and representation of protein binding sites for molecular docking. *Protein science : a publication of the Protein Society* **1997**, *6*, 524-33.

(104) Verdonk, M. L.; Cole, J. C.; Watson, P.; Gillet, V.; Willett, P.: SuperStar: improved knowledge-based interaction fields for protein binding sites. *Journal of molecular biology* **2001**, *307*, 841-59.

(105) Cheng, A. C.; Coleman, R. G.; Smyth, K. T.; Cao, Q.; Soulard, P.; Caffrey, D. R.; Salzberg, A. C.; Huang, E. S.: Structure-based maximal affinity model predicts small-molecule druggability. *Nature biotechnology* **2007**, *25*, 71-5.

(106) Coleman, R. G.; Salzberg, A. C.; Cheng, A. C.: Structure-based identification of small molecule binding sites using a free energy model. *Journal of chemical information and modeling* **2006**, *46*, 2631-7.

(107) Joughin, B. A.; Tidor, B.; Yaffe, M. B.: A computational method for the analysis and prediction of protein:phosphopeptide-binding sites. *Protein science : a publication of the Protein Society* **2005**, *14*, 131-9.

(108) Nayal, M.; Honig, B.: On the nature of cavities on protein surfaces: application to the identification of drug-binding sites. *Proteins* **2006**, *63*, 892-906.

(109) SiteMap: v2.6, Schrödinger. LLC, New York, NY,, 2012.

(110) Halgren, T.: New method for fast and accurate binding-site identification and analysis. *Chemical biology & drug design* **2007**, *69*, 146-8.

(111) Halgren, T. A.: Identifying and characterizing binding sites and assessing druggability. *Journal of chemical information and modeling* **2009**, *49*, 377-89.

(112) Tubert-Brohman, I.; Sherman, W.; Repasky, M.; Beuming, T.: Improved docking of polypeptides with glide. *Journal of chemical information and modeling* **2013**, *53*, 1689-99.

(113) Sastry, G. M.; Inakollu, V. S.; Sherman, W.: Boosting virtual screening enrichments with data fusion: coalescing hits from two-dimensional fingerprints, shape, and docking. *Journal of chemical information and modeling* **2013**, *53*, 1531-42.

(114) Palakurti, R.; Sriram, D.; Yogeewari, P.; Vadrevu, R.: Multiple e-Pharmacophore Modeling Combined with High-Throughput Virtual Screening and Docking to Identify Potential Inhibitors of β -Secretase(BACE1). *Molecular Informatics* **2013**, *32*, 385-398.

(115) Schrödinger: Schrödinger Suite 2012 Protein Preparation Wizard; Epik version 2.3, Schrödinger, LLC, New

York, NY, 2012; Impact version 5.8, Schrödinger, LLC, New York, NY, 2012; Prime version

3.1, Schrödinger, LLC, New York, NY, 2012. 2012.

- (116) Sastry, G. M.; Adzhigirey, M.; Day, T.; Annabhimoju, R.; Sherman, W.: Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *Journal of computer-aided molecular design* **2013**, *27*, 221-34.
- (117) Olsson, M. H. M.; Søndergaard, C. R.; Rostkowski, M.; Jensen, J. H.: PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical pKa Predictions. *Journal of Chemical Theory and Computation* **2011**, *7*, 525-537.
- (118) Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; Donini, O.; Cieplak, P.; Srinivasan, J.; Case, D. A.; Cheatham, T. E., 3rd: Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Accounts of chemical research* **2000**, *33*, 889-97.
- (119) Yu, Z.; Jacobson, M. P.; Friesner, R. A.: What role do surfaces play in GB models? A new-generation of surface-generalized born model based on a novel gaussian surface for biomolecules. *Journal of computational chemistry* **2006**, *27*, 72-89.
- (120) Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T.: Semianalytical treatment of solvation for molecular mechanics and dynamics. *Journal of the American Chemical Society* **1990**, *112*, 6127-6129.
- (121) Rastelli, G.; Del Rio, A.; Degliesposti, G.; Sgobba, M.: Fast and accurate predictions of binding free energies using MM-PBSA and MM-GBSA. *Journal of computational chemistry* **2010**, *31*, 797-810.
- (122) Hou, T.; Wang, J.; Li, Y.; Wang, W.: Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations. *Journal of chemical information and modeling* **2011**, *51*, 69-82.
- (123) Prime: v3.1, Schrödinger. LLC, New York, NY,, 2012.
- (124) Ghosh, A.; Rapp, C. S.; Friesner, R. A.: Generalized Born Model Based on a Surface Integral Formulation. *The Journal of Physical Chemistry B* **1998**, *102*, 10983-10990.
- (125) Zhu, K.; Pincus, D. L.; Zhao, S.; Friesner, R. A.: Long loop prediction using the protein local optimization program. *Proteins* **2006**, *65*, 438-52.
- (126) Zhu, K.; Shirts, M. R.; Friesner, R. A.: Improved Methods for Side Chain and Loop Predictions via the Protein Local Optimization Program: Variable Dielectric Model for Implicitly Improving the Treatment of Polarization Effects. *Journal of Chemical Theory and Computation* **2007**, *3*, 2108-2119.
- (127) Li, J.; Abel, R.; Zhu, K.; Cao, Y.; Zhao, S.; Friesner, R. A.: The VSGB 2.0 model: a next generation energy model for high resolution protein structure modeling. *Proteins* **2011**, *79*, 2794-812.
- (128) Tsui, V.; Case, D. A.: Molecular Dynamics Simulations of Nucleic Acids with a Generalized Born Solvation Model. *Journal of the American Chemical Society* **2000**, *122*, 2489-2498.
- (129) Ylilauri, M.; Pentikainen, O. T.: MMGBSA As a Tool To Understand the Binding Affinities of Filamin-Peptide Interactions. *Journal of chemical information and modeling* **2013**.
- (130) Lyne, P. D.; Lamb, M. L.; Saeh, J. C.: Accurate prediction of the relative potencies of members of a series of kinase inhibitors using molecular docking and MM-GBSA scoring. *Journal of medicinal chemistry* **2006**, *49*, 4805-8.
- (131) Du, J.; Sun, H.; Xi, L.; Li, J.; Yang, Y.; Liu, H.; Yao, X.: Molecular modeling study of checkpoint kinase 1 inhibitors by multiple docking strategies and prime/MM-GBSA calculation. *Journal of computational chemistry* **2011**, *32*, 2800-9.
- (132) Maestro: v9.3, Schrödinger. LLC, New York, NY, , 2012.
- (133) Sheldrick, G. M.: A short history of SHELX. *Acta crystallographica. Section A, Foundations of crystallography* **2008**, *64*, 112-22.

- (134) Jacobson, M. P.; Friesner, R. A.; Xiang, Z.; Honig, B.: On the role of the crystal environment in determining protein side-chain conformations. *Journal of molecular biology* **2002**, *320*, 597-608.
- (135) Jacobson, M. P.; Pincus, D. L.; Rapp, C. S.; Day, T. J.; Honig, B.; Shaw, D. E.; Friesner, R. A.: A hierarchical approach to all-atom protein loop prediction. *Proteins* **2004**, *55*, 351-67.
- (136) Rost, B.: Twilight zone of protein sequence alignments. *Protein engineering* **1999**, *12*, 85-94.
- (137) Irwin, J. J.; Shoichet, B. K.: ZINC--a free database of commercially available compounds for virtual screening. *Journal of chemical information and modeling* **2005**, *45*, 177-82.
- (138) LigPrep: v2.5, Schrödinger. LLC, New York, NY, 2012.
- (139) Brooks, W. H.; Daniel, K. G.; Sung, S. S.; Guida, W. C.: Computational validation of the importance of absolute stereochemistry in virtual screening. *Journal of chemical information and modeling* **2008**, *48*, 639-45.
- (140) Albers, H. M.; Hendrickx, L. J.; van Tol, R. J.; Hausmann, J.; Perrakis, A.; Ovaa, H.: Structure-based design of novel boronic acid-based inhibitors of autotaxin. *Journal of medicinal chemistry* **2011**, *54*, 4619-26.
- (141) Kenyon, V.; Chorny, I.; Carvajal, W. J.; Holman, T. R.; Jacobson, M. P.: Novel human lipoyxygenase inhibitors discovered using virtual screening with homology models. *Journal of medicinal chemistry* **2006**, *49*, 1356-63.
- (142) William, A. D.; Lee, A. C.; Blanchard, S.; Poulsen, A.; Teo, E. L.; Nagaraj, H.; Tan, E.; Chen, D.; Williams, M.; Sun, E. T.; Goh, K. C.; Ong, W. C.; Goh, S. K.; Hart, S.; Jayaraman, R.; Pasha, M. K.; Ethirajulu, K.; Wood, J. M.; Dymock, B. W.: Discovery of the macrocycle 11-(2-pyrrolidin-1-yl-ethoxy)-14,19-dioxo-5,7,26-triaza-tetracyclo[19.3.1.1(2,6). 1(8,12)]heptacos-1(25),2(26),3,5,8,10,12(27),16,21,23-decaene (SB1518), a potent Janus kinase 2/fms-like tyrosine kinase-3 (JAK2/FLT3) inhibitor for the treatment of myelofibrosis and lymphoma. *Journal of medicinal chemistry* **2011**, *54*, 4638-58.
- (143) Siddiquee, K.; Zhang, S.; Guida, W. C.; Blaskovich, M. A.; Greedy, B.; Lawrence, H. R.; Yip, M. L.; Jove, R.; McLaughlin, M. M.; Lawrence, N. J.; Sebt, S. M.; Turkson, J.: Selective chemical probe inhibitor of Stat3, identified through structure-based virtual screening, induces antitumor activity. *Proceedings of the National Academy of Sciences of the United States of America* **2007**, *104*, 7391-6.
- (144) Gadakar, P. K.; Phukan, S.; Dattatreya, P.; Balaji, V. N.: Pose prediction accuracy in docking studies and enrichment of actives in the active site of GSK-3beta. *Journal of chemical information and modeling* **2007**, *47*, 1446-59.
- (145) Wu, D.; Wang, Q.; Assary, R. S.; Broadbelt, L. J.; Krilov, G.: A computational approach to design and evaluate enzymatic reaction pathways: application to 1-butanol production from pyruvate. *Journal of chemical information and modeling* **2011**, *51*, 1634-47.
- (146) Bortolato, A.; Moro, S.: In silico binding free energy predictability by using the linear interaction energy (LIE) method: bromobenzimidazole CK2 inhibitors as a case study. *Journal of chemical information and modeling* **2007**, *47*, 572-82.
- (147) Durdagi, S.; Duff, H. J.; Noskov, S. Y.: Combined receptor and ligand-based approach to the universal pharmacophore model development for studies of drug blockade to the hERG1 pore domain. *Journal of chemical information and modeling* **2011**, *51*, 463-74.
- (148) CombiGlide: v2.8, Schrödinger. LLC, New York, NY, , 2012.
- (149) Schroeder, K. S.; Neagle, B. D.: FLIPR: A New Instrument for Accurate, High Throughput Optical Screening. *Journal of Biomolecular Screening* **1996**, *1*, 75-80.
- (150) Whiteaker, K. L.; Sullivan, J. P.; Gopalakrishnan, M.: Cell-based assays using the fluorometric imaging plate reader (FLIPR). *Current protocols in pharmacology / editorial board, S.J. Enna (editor-in-chief) ... [et al.]* **2001**, Chapter 9, Unit 9 2.

- (151) Matsuo, K.; Kimura, M.; Kinuta, T.; Takai, N.; Tanaka, K.: Syntheses and antimicrobial activities of 3-acyltetramic acid derivatives. *Chemical & pharmaceutical bulletin* **1984**, *32*, 4197-204.
- (152) Balasubramanian, S.; Nair, M. G.: An Efficient "One Pot" Synthesis of Isoflavones. *Synthetic Communications* **2000**, *30*, 469-484.
- (153) Woeste, M.; Steller, J.; Hofmann, E.; Kidd, T.; Patel, R.; Connolly, K.; Jayasinghe, M.; Paula, S.: Structural requirements for inhibitory effects of bisphenols on the activity of the sarco/endoplasmic reticulum calcium ATPase. *Bioorganic & medicinal chemistry* **2013**, *21*, 3927-33.
- (154) Mollah†, M. S. I.; Seo†, D. W.; Islam, M. M.; Lim, Y. D.; Cho, S. H.; Shin, K. M. O. O.; Kim, J. H.; Kim, W. G.: Synthesis and Characterization of Grafted Silicone Polycarbonates. *Journal of Macromolecular Science, Part A* **2011**, *48*, 400-408.
- (155) Bureeva, S.; Andia-Pravdivy, J.; Petrov, G.; Igumnov, M.; Romanov, S.; Kolesnikova, E.; Kaplun, A.; Kozlov, L.: Inhibition of classical pathway of complement activation with negative charged derivatives of bisphenol A and bisphenol disulphates. *Bioorganic & medicinal chemistry* **2005**, *13*, 1045-52.
- (156) Bar-Sela, G.; Epelbaum, R.; Schaffer, M.: Curcumin as an anti-cancer agent: review of the gap between basic and clinical applications. *Current medicinal chemistry* **2010**, *17*, 190-7.
- (157) Mayadevi, M.; Sherin, D. R.; Keerthi, V. S.; Rajasekharan, K. N.; Omkumar, R. V.: Curcumin is an inhibitor of calcium/calmodulin dependent protein kinase II. *Bioorganic & medicinal chemistry* **2012**, *20*, 6040-7.
- (158) Elavarasan, S.; Bhakiaraj, D.; Chellakili, B.; Elavarasan, T.; Gopalakrishnan, M.: One pot synthesis, structural and spectral analysis of some symmetrical curcumin analogues catalyzed by calcium oxide under microwave irradiation. *Spectrochimica acta. Part A, Molecular and biomolecular spectroscopy* **2012**, *97*, 717-21.
- (159) Salituro, F., G.; (US); Saunders, J., O.; (US) Therapeutically active compounds for use in the treatment of cancer characterized as having an IDH mutation. A61K 31/495 (2006.01), A61K 31/496 (2006.01), A61K 31/506 (2006.01), A61K 31/551 (2006.01), A61P 35/00 (2006.01), C07D 243/08 (2006.01), C07D 405/12 (2006.01) ed.: USA, 2011; Vol. WO/2011/072174.
- (160) Burnett, M. N.; Johnson, C. K.: ORTEP-III: Oak Ridge Thermal Ellipsoid Plot Program for Crystal Structure Illustrations. Oak Ridge National Laboratory Report ORNL-6895, 1996.
- (161) SMART, S. a. X.: Area detector control, data integration and reduction software. Bruker Analytical X-ray Instruments Inc.: Madison, WI, 1995.
- (162) Hamon; Christian; (Kriftel, D. S. J. F. a. M., DE) ; Becker; Wolfgang; (Eppstein, DE) ; Kienle; Stefan; (Frankfurt am Main, DE) ; Kuhn; Karsten; (Hofheim am Taunus, DE) ; Schaefer; Juergen; (Lauterbach, DE): Mass Label. Office, U. P. a. T., Ed.; Electrophoretics Limited, Cobham Surrey, GB, 2010.
- (163) Pierce, A. C.; Rao, G.; Bemis, G. W.: BREED: Generating novel inhibitors through hybridization of known ligands. Application to CDK2, p38, and HIV protease. *Journal of medicinal chemistry* **2004**, *47*, 2768-75.
- (164) Eckstein-Ludwig, U.; Webb, R. J.; Van Goethem, I. D.; East, J. M.; Lee, A. G.; Kimura, M.; O'Neill, P. M.; Bray, P. G.; Ward, S. A.; Krishna, S.: Artemisinins target the SERCA of Plasmodium falciparum. *Nature* **2003**, *424*, 957-61.
- (165) Lai, P.; Michelangeli, F.: Bis(2-hydroxy-3-tert-butyl-5-methyl-phenyl)-methane (bisphenol) is a potent and selective inhibitor of the secretory pathway Ca(2)(+) ATPase (SPCA1). *Biochemical and biophysical research communications* **2012**, *424*, 616-9.
- (166) Patrick, G. L.: An Introduction to Medicinal Chemistry. Fourth ed.; Oxford University Press Inc.: New York, 2009; pp 362-363, 407-418, 686.
- (167) Kuttan, R.; Bhanumathy, P.; Nirmala, K.; George, M. C.: Potential anticancer activity of turmeric (*Curcuma longa*). *Cancer letters* **1985**, *29*, 197-202.

- (168) Brouet, I.; Ohshima, H.: Curcumin, an anti-tumour promoter and anti-inflammatory agent, inhibits induction of nitric oxide synthase in activated macrophages. *Biochemical and biophysical research communications* **1995**, *206*, 533-40.
- (169) Zhu, J.; Bultynck, G.; Luyten, T.; Parys, J. B.; Creemers, J. W.; Van de Ven, W. J.; Vermorken, A. J.: Curcumin affects proprotein convertase activity: elucidation of the molecular and subcellular mechanism. *Biochimica et biophysica acta* **2013**, *1833*, 1924-35.
- (170) Sokolove, P. M.; Albuquerque, E. X.; Kauffman, F. C.; Spande, T. F.; Daly, J. W.: Phenolic antioxidants: potent inhibitors of the (Ca²⁺ + Mg²⁺)-ATPase of sarcoplasmic reticulum. *FEBS letters* **1986**, *203*, 121-6.
- (171) Scior, T.; Bender, A.; Tresadern, G.; Medina-Franco, J. L.; Martinez-Mayorga, K.; Langer, T.; Cuanalo-Contreras, K.; Agrafiotis, D. K.: Recognizing pitfalls in virtual screening: a critical review. *Journal of chemical information and modeling* **2012**, *52*, 867-81.
- (172) Kitamura, S.; Suzuki, T.; Sanoh, S.; Kohta, R.; Jinno, N.; Sugihara, K.; Yoshihara, S.; Fujimoto, N.; Watanabe, H.; Ohta, S.: Comparative study of the endocrine-disrupting activity of bisphenol A and 19 related compounds. *Toxicological sciences : an official journal of the Society of Toxicology* **2005**, *84*, 249-59.
- (173) Hengstler, J. G.; Foth, H.; Gebel, T.; Kramer, P. J.; Lilienblum, W.; Schweinfurth, H.; Volkel, W.; Wollin, K. M.; Gundert-Remy, U.: Critical evaluation of key evidence on the human health hazards of exposure to bisphenol A. *Critical reviews in toxicology* **2011**, *41*, 263-91.
- (174) Huang, M. T.; Newmark, H. L.; Frenkel, K.: Inhibitory effects of curcumin on tumorigenesis in mice. *Journal of cellular biochemistry. Supplement* **1997**, *27*, 26-34.

APPENDIX A

Literature review for molecular modelling methods section

The first step was to understand the background of Ca^{2+} ATPase, especially the role of Ca^{2+} ATPase in general and then specifically in cancer. National Centre for Biotechnology Information (NCBI) Pubmed was the main literature database used for literature review. Keywords used were cancers and SPCA1, calcium ATPase AND sarcoplasmic reticulum AND endoplasmic reticulum AND SPCA1. Other literature searches included work done prior from our collaborator from School of Pharmacy at the University of Queensland on calcium signalling in cancer.

Further literature searches were conducted to find any human-sourced structures of Ca^{2+} ATPase. Keywords used in the search were X-ray crystallography, X-ray diffraction and calcium ATPase. A total of 81 papers were found with 14 of them from human source. Research Collaboratory for Structural Bioinformatics (RCSB) Protein Data Bank (PDB) website has provided protein structures obtained from X-ray crystallography, nuclear magnetic resonance spectroscopy and/or homology modelling. There were many SERCA structures found in this website.

Ca^{2+} ATPase protein structures were chosen based on the protein source to see if it was from human and if the resolution was less than 3 Å. Actual search on PDB website has found 44 structures for SERCA, two structures for PMCA and none for SPCA. However, none of the Ca^{2+} ATPase structures found were from human, except one based on calmodulin- Ca^{2+} ATPase

complex for PMCA. Majority of the SERCA protein structures were contributed from the work of C. Toyoshima *et. al.* where majority of the protein source was from rabbit skeletal muscles. The keywords used to search for Toyoshima's work on SERCA protein structures on Pubmed were calcium ATPase or Toyoshima C. calcium ATPase and for PDB website were calcium ATPase. On the other hand, further searches were carried out on Pubmed in order to find any SPCA structures by using keywords such as secretory pathway calcium ATPase and X-ray crystallography. However there were no results that matched these keywords.

Since there were no SPCA protein structures available, the next step was to make a homology protein model of SPCA based on the structures available for SERCA. To do this, the actual protein sequence from a human-sourced SPCA was found on NCBI Protein. This would provide the actual protein backbone for the *in silico* SPCA protein model while using the actual SERCA crystal structure to provide the protein shape.

APPENDIX B

Cell feeding methods:

Dulbecco's Modified Eagle Medium (DMEM) was warmed in 37°C water bath prior to use. Working surface in the hood was sprayed and wiped with 70% ethanol. Cell confluency and contamination were checked before cell feeding. Old media was discarded and replaced with 10mL of fresh media. The cell flask was then placed into the incubator for further cell culture.

Cell splitting methods:

DMEM was warmed in 37°C water bath prior to use. Trypsin was left to be warmed at room temperature. Working surface, equipments and containers were sprayed and wiped with 70% ethanol prior to use. Cell culture flask was checked under the microscope for cell confluency and possible contamination. Old media was aspirated and replaced with PBS/EDTA (5mL) which was then discarded and replaced with trypsin (2mL). The cell flask was placed into the incubator for about half to two minutes. The flask was then checked under the microscope to ensure all cells were removed. Fresh media (5mL) containing PBS was added to the flask and aspirated to be centrifuged for two minutes at 1500rpm. Small cell pellet formed at the bottom of tube with supernatant removed and discarded. The cell pellet was then mixed to resuspend the cells. Fresh media (5mL) was added and mixed with the cells. The resuspended cell solution (0.5mL for 1:10 dilution or 1mL for 1:5 dilution) were aspirated and dispensed into a new flask. Fresh media (9.5mL for 1:10 dilution or 9mL for 1:5 dilution) was then added to the resuspended cell solution. The cell mixture was then added into the new flask and stored in the incubator at 37°C with 5% CO₂.

Cell plating methods:

After cell splitting took place, cell count was done prior to cell plating. A 1:2 dilution (50 μ L of cell solution was added to 50 μ L PBS) was used for the cell count. FLIPR-specific black cell 96-well plates were used. ViaFlo pipettor was used to aspirate and dispense 100 μ L of the cell solution into each well of the plate. The cell plate was stored in the incubator after completion before running the FLIPR.

APPENDIX C

The FLIPR protocol:

HEPES buffer solution (1L) contained potassium chloride powder (4.398g), magnesium chloride solution (14mL of 1M), HEPES powder (23.830g), sodium dihydrogen phosphate solution (12mL of 1M) and sodium bicarbonate powder (4.201g).

PSS buffer solution (1L) contained HEPES stock buffer (100mL), sodium chloride solution 9100mL or 8.182g of 1.4M), glucose powder (2.07238g) and calcium chloride solution (0.9mL of 2M). The actual buffer used in FLIPR was the PSS buffer without the Ca^{2+} component which was named PSS nominal Ca^{2+} buffer.

Day 1: Cells were plated in black 96-well plate for use in FLIPR in complete growth media (DMEM + L-glutamine + foetal bovine serum + penicillin-streptomycin). Each well was seeded with 15,000 cells for 48 hour at 37°C, 5% CO_2 (100 μL per well).

	1	2	3	4	5	6	7	8	9	10	11	12
A												
B												
C												
D												
E												
F												
G												
H												

Day 2: FLIPR

Layout:

Reagent plate 1	Reagent plate 2	Reagent plate 3
<ul style="list-style-type: none"> BAPTA 3× 	<ul style="list-style-type: none"> BAPTA 1× CPA 4× Compound 8 (ST4045533) 4× Compound 3 (ST4081102) 4× DMSO (i.e. matched to CPA) 	<ul style="list-style-type: none"> Ca²⁺ 5× BAPTA 1× DMSO 1×/ CPA 1× Compound 8 (ST4045533) 1× Compound 3 (ST4081102) 1×

REAGENT PLATE #1: BAPTA

Reagent	Initial/stock conc	Final plate conc	3X conc	Vol BAPTA	Vol PSS nominal
BAPTA STOCK	100mM stock BAPTA was prepared and added 56.4mg (0.0564g) to 1mL nominal Ca ²⁺ PSS				
BAPTA 3X	100mM	500μM	1.5mM	600μL	39.4mL

Plate 1: 100 μL × 72 wells = 7.2mL

Plate 2: 0.33mL × 24 wells (1× of BAPTA for 3 treatment groups × 8 rows of wells) ≈ 8mL

Plate 3: 0.33mL × 26 wells (making some extra) ≈ 9mL

Total volume of BAPTA = 25mL therefore made 40mL in total

	1	2	3	4	5	6	7	8	9	10	11	12
A	BAPTA	BAPTA	BAPTA	BAPTA	BAPTA	BAPTA	BAPTA	BAPTA	BAPTA	BAPTA	BAPTA	BAPTA
B	BAPTA	BAPTA	BAPTA	BAPTA	BAPTA	BAPTA	BAPTA	BAPTA	BAPTA	BAPTA	BAPTA	BAPTA
C	BAPTA	BAPTA	BAPTA	BAPTA	BAPTA	BAPTA	BAPTA	BAPTA	BAPTA	BAPTA	BAPTA	BAPTA
D	BAPTA	BAPTA	BAPTA	BAPTA	BAPTA	BAPTA	BAPTA	BAPTA	BAPTA	BAPTA	BAPTA	BAPTA
E	BAPTA	BAPTA	BAPTA	BAPTA	BAPTA	BAPTA	BAPTA	BAPTA	BAPTA	BAPTA	BAPTA	BAPTA
F	BAPTA	BAPTA	BAPTA	BAPTA	BAPTA	BAPTA	BAPTA	BAPTA	BAPTA	BAPTA	BAPTA	BAPTA
G	BAPTA	BAPTA	BAPTA	BAPTA	BAPTA	BAPTA	BAPTA	BAPTA	BAPTA	BAPTA	BAPTA	BAPTA
H	BAPTA	BAPTA	BAPTA	BAPTA	BAPTA	BAPTA	BAPTA	BAPTA	BAPTA	BAPTA	BAPTA	BAPTA

REAGENT PLATE #2 – COMPOUND 8 & 2 + CPA

Stock solutions of the compounds were prepared in 100% DMSO

1. Compound 8 (ZINC04189857) 0.0039g in 100 μ L
Actual measurement: 3.596mg in 93.4 μ L of DMSO
2. Compound 3 (ZINC08846492) 0.0048g in 100 μ L
Actual measurement: 3.657mg in 75.6 μ L of DMSO
3. Stock CPA solution was made up at 100mM.

Both compound 8 and 3 were in white powder form placed in amber glass vials. Due to static charges inside the vial, the powders were attached to the sides of the vial during the weighing process. Therefore, the actual amounts of the compounds weighed were less than expected.

Compound molar concentrations were reduced to increase solubility by diluting stock solution of 100mM with DMSO to 10mM (1:10 dilution) for C1, to 3mM (1:33) for C2 and to 10mM (1:10) for CPA.

Agonists	Initial / stock conc	Final conc	4 \times conc	Volume agonist	Volume PSS/ BAPTA 3 \times	Volume PSS nominal Ca ²⁺	Volume DMSO	TOTAL volume
Compound 8 ZINC04189857	10mM	10 μ M	40 μ M	4 μ L	0.333mL	0.663mL	0 μ L	1mL
	40 μ M	3 μ M	12 μ M	0.3mL	0.333mL	0.364mL	2.8 μ L	1mL
	12 μ M	1 μ M	4 μ M	0.333mL	0.333mL	0.331mL	2.68 μ L	1mL
	4 μ M	0.3 μ M	1.2 μ M	0.3mL	0.333mL	0.364mL	2.8 μ L	1mL
	1.2 μ M	0.1 μ M	0.4 μ M	0.333mL	0.333mL	0.331mL	2.68 μ L	1mL
	0.4 μ M	0.03 μ M	0.12 μ M	0.3mL	0.333mL	0.364mL	2.8 μ L	1mL
	0.12 μ M	0.01 μ M	0.04 μ M	0.333mL	0.333mL	0.331mL	2.68 μ L	1mL
	0 μ M	0 μ M	0 μ M	0mL	0.333mL	0.663mL	4 μ L	1mL

Agonists	Initial / stock conc	Final conc	4× conc	Volume agonist	Volume PSS/ BAPTA 3×	Volume PSS nominal Ca ²⁺	Volume DMSO	TOTAL volume
Compound 3 ZINC08846492	3mM	3μM	12μM	4μL	0.333mL	0.663mL	0μL	1mL
	12μM	1μM	4μM	0.333mL	0.333mL	0.331mL	2.68μL	1mL
	4μM	0.3μM	1.2μM	0.3mL	0.333mL	0.364mL	2.8μL	1mL
	1.2μM	0.1μM	0.4μM	0.333mL	0.333mL	0.331mL	2.68μL	1mL
	0.4μM	0.03μM	0.12μM	0.3mL	0.333mL	0.364mL	2.8μL	1mL
	0.12μM	0.01μM	0.04μM	0.333mL	0.333mL	0.331mL	2.68μL	1mL
	0μM	0μM	0μM	0mL	0.333mL	0.663mL	4μL	1mL

Agonists	Initial/ stock conc	Final conc	4× conc	Volume agonist	Volume PSS/ BAPTA 3×	Volume PSS nominal Ca ²⁺	Volume DMSO	TOTAL volume
CPA	10mM	10μM	40μM	4μL	0.333mL	0.663mL	0μL	1mL
	40μM	3μM	12μM	0.3mL	0.333mL	0.364mL	2.8μL	1mL
	12μM	1μM	4μM	0.333mL	0.333mL	0.331mL	2.68μL	1mL
	4μM	0.3μM	1.2μM	0.3mL	0.333mL	0.364mL	2.8μL	1mL
	1.2μM	0.1μM	0.4μM	0.333mL	0.333mL	0.331mL	2.68μL	1mL
	0.4μM	0.03μM	0.12μM	0.3mL	0.333mL	0.364mL	2.8μL	1mL
	0.12μM	0.01μM	0.04μM	0.333mL	0.333mL	0.331mL	2.68μL	1mL
	0μM	0μM	0μM	0mL	0.333mL	0.663mL	4μL	1mL

		CPA			Compound 8			Compound 3			10	11	12
		1	2	3	4	5	6	7	8	9			
10μM	A							PSS	PSS	PSS	PSS	PSS	PSS
3μM	B										PSS	PSS	PSS
1μM	C										PSS	PSS	PSS
0.3μM	D										PSS	PSS	PSS
0.1μM	E										PSS	PSS	PSS
0.03μM	F										PSS	PSS	PSS
0.01μM	G										PSS	PSS	PSS
0μM	H										PSS	PSS	PSS

REAGENT PLATE #3:

2mM Ca²⁺ was used for each well

Final well conc	Initial conc	Final Ca ²⁺ plate conc	5× conc	Vol Ca ²⁺	Vol PSS 3× BAPTA	Vol PSS 4× DMSO/ 4× CPA/ 4× Cpd 1/ 4× Cpd 2	Vol PSS nominal	TOTAL volume
10μM	2M	2mM	10mM	5μL	0.333mL	0.25mL CPA /Cpd 1 /Cpd 2	0.412mL	1mL
3μM	2M	2mM	10mM					
1μM	2M	2mM	10mM					
0.3μM	2M	2mM	10mM					
0.1μM	2M	2mM	10mM					
0.03μM	2M	2mM	10mM					
0.01μM	2M	2mM	10mM					
0μM	0M	0mM	0mM					

Twenty six batches of 130 μ L Ca²⁺ plus 8.58mL PSS 3 \times BAPTA and 10.79mL PSS nominal Ca²⁺ (19.5mL in total) were prepared. This was followed by adding 750 μ L of Ca²⁺ + PSS 3 \times BAPTA + PSS nominal Ca²⁺ into each tube (for plate 2) then added 250 μ L of PSS DMSO/CPA/Cpd1/Cpd2 to each tube (for plate 3) – made from plate 2 for matching concentrations.

		CPA			Compound 8			Compound 3			10	11	12
		1	2	3	4	5	6	7	8	9			
10μM	A	Ca ²⁺	Ca ²⁺	Ca ²⁺	Ca ²⁺	Ca ²⁺	Ca ²⁺	PSS	PSS	PSS	PSS	PSS	PSS
3μM	B	Ca ²⁺	Ca ²⁺	Ca ²⁺	Ca ²⁺	Ca ²⁺	Ca ²⁺	Ca ²⁺	Ca ²⁺	Ca ²⁺	PSS	PSS	PSS
1μM	C	Ca ²⁺	Ca ²⁺	Ca ²⁺	Ca ²⁺	Ca ²⁺	Ca ²⁺	Ca ²⁺	Ca ²⁺	Ca ²⁺	PSS	PSS	PSS
0.3μM	D	Ca ²⁺	Ca ²⁺	Ca ²⁺	Ca ²⁺	Ca ²⁺	Ca ²⁺	Ca ²⁺	Ca ²⁺	Ca ²⁺	PSS	PSS	PSS
0.1μM	E	Ca ²⁺	Ca ²⁺	Ca ²⁺	Ca ²⁺	Ca ²⁺	Ca ²⁺	Ca ²⁺	Ca ²⁺	Ca ²⁺	PSS	PSS	PSS
0.03μM	F	Ca ²⁺	Ca ²⁺	Ca ²⁺	Ca ²⁺	Ca ²⁺	Ca ²⁺	Ca ²⁺	Ca ²⁺	Ca ²⁺	PSS	PSS	PSS
0.01μM	G	Ca ²⁺	Ca ²⁺	Ca ²⁺	Ca ²⁺	Ca ²⁺	Ca ²⁺	Ca ²⁺	Ca ²⁺	Ca ²⁺	PSS	PSS	PSS
0μM	H	Ca ²⁺	Ca ²⁺	Ca ²⁺	Ca ²⁺	Ca ²⁺	Ca ²⁺	Ca ²⁺	Ca ²⁺	Ca ²⁺	PSS	PSS	PSS

FLIPR solutions were made fresh on the day of running FLIPR and PSS buffer was warmed to room temperature. The components of each FLIPR solutions were shown as below.

PSS/Ca²⁺ + Bovine serum albumin (BSA)

BSA, fatty acid free (4°C)..... 0.06g
PSS/Ca²⁺ 20mL

Fluo-4-AM (4mM)

Fluo-4 (Invitrogen) 1 tube
DMSO 11.4 μ L
wrap in foil; store at -22°C

Fluo-4-AM (4μM) (dye loading solution)

Fluo-4 (4mM)..... 8.5μL
 Growth media..... 8.5mL

Cell loading

The media was removed and Fluo-4 was loaded to each well using Viaflo pipettor and the plate was incubated for 30 minutes. The media was again removed and the plate was further incubated for 15 minutes at room temperature in PSS/Ca²⁺ in foil.

Cell washing

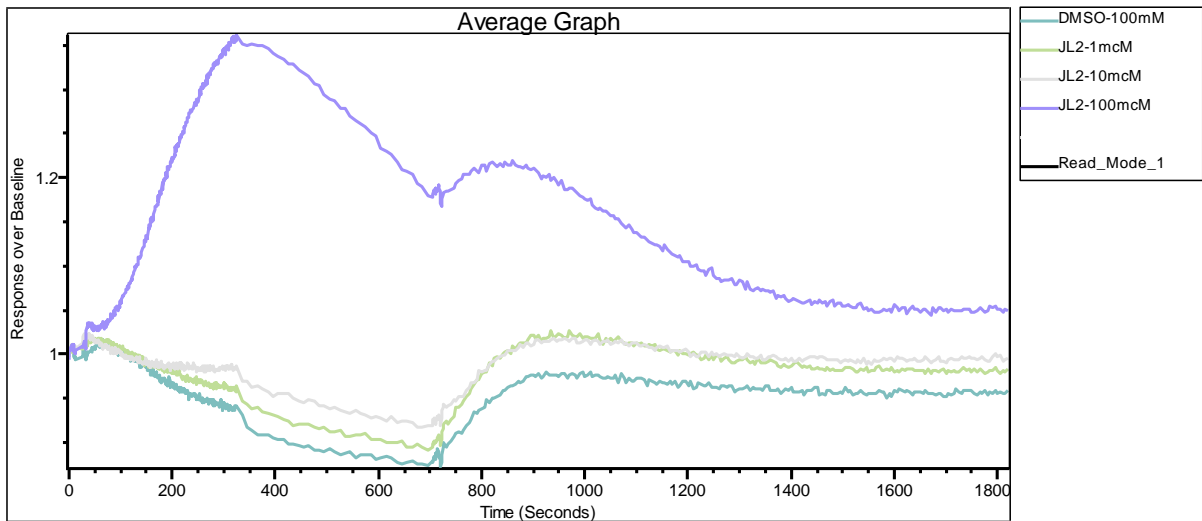
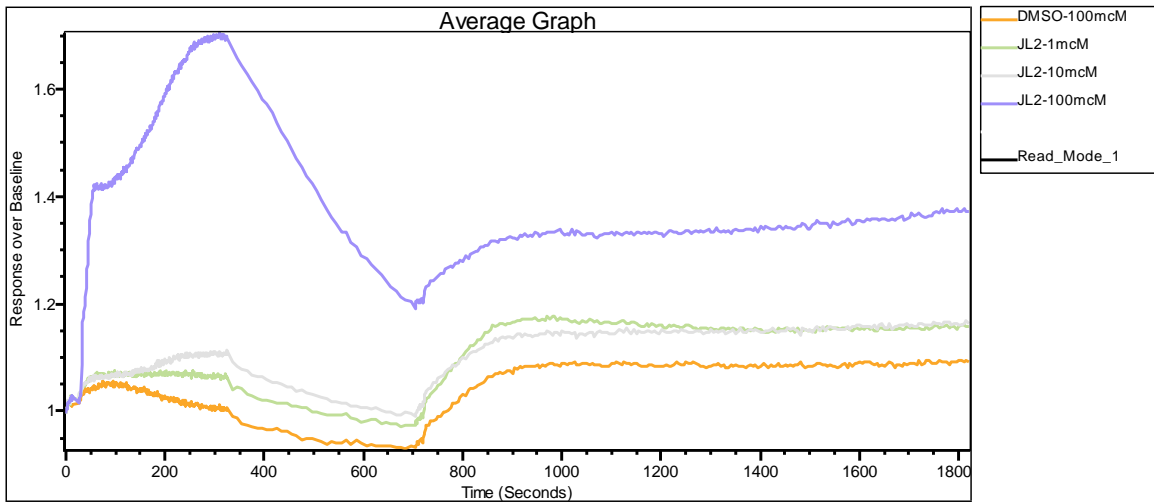
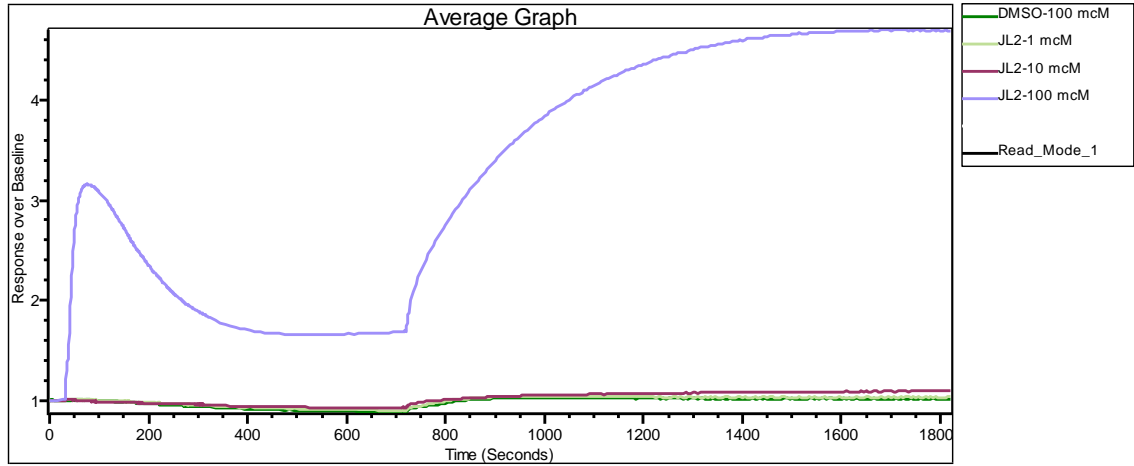
Cells were washed twice with PSS/Ca²⁺+BSA (100μL/well) and once with PSS/nominal Ca²⁺ (100μL/well). PSS/nominal Ca²⁺ (100μL/well) were added into each well and the plate was submitted to FLIPR to begin the run. The reagents were added in the order of plate 1, plate 2 and plate 3 to the cell plate.

Final reagent plate layout

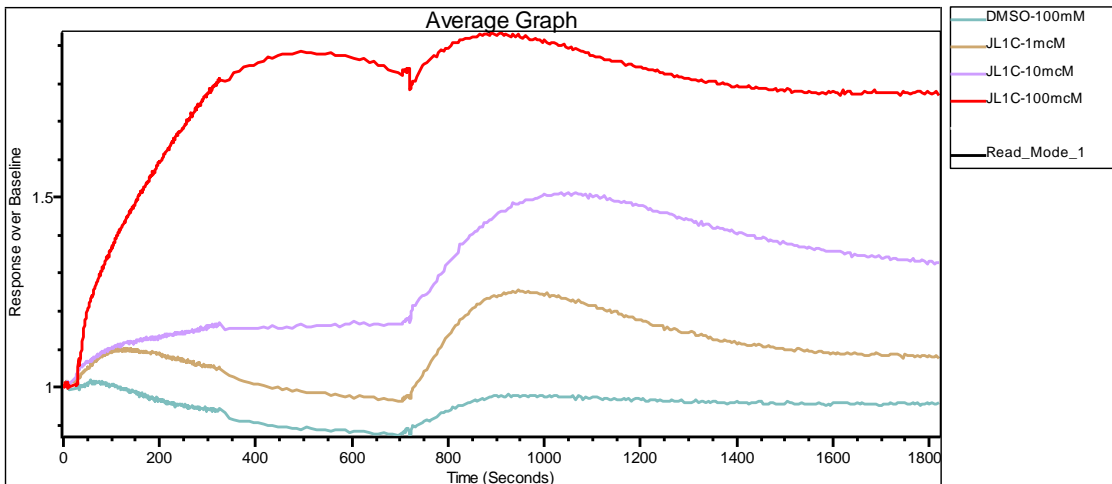
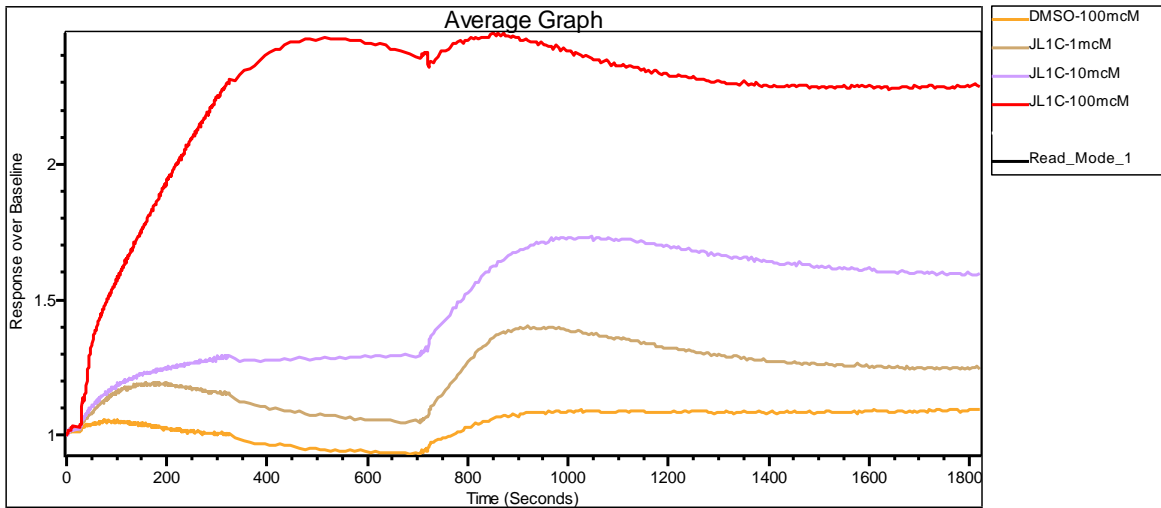
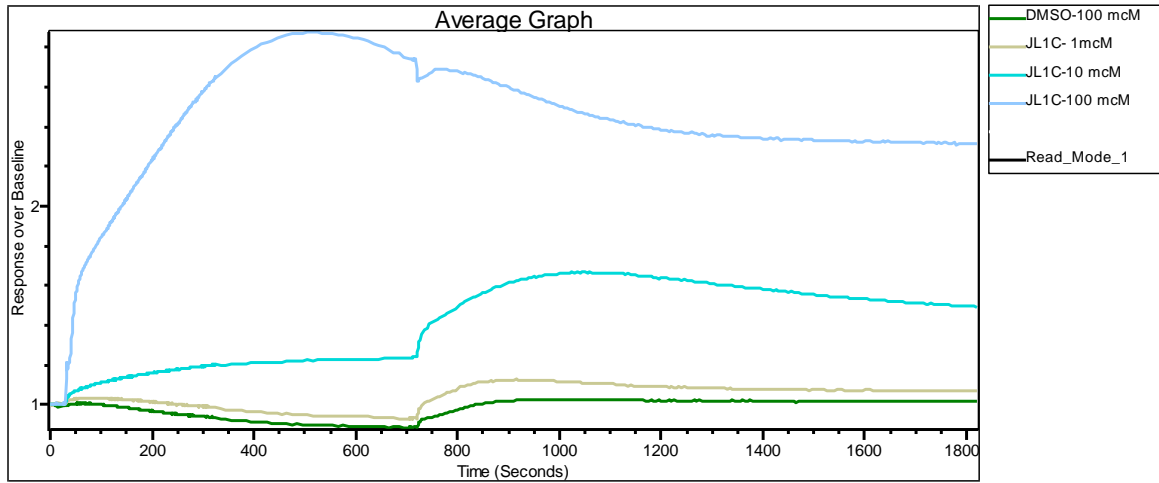
		CPA			Compound 8			Compound 3					
		1	2	3	4	5	6	7	8	9	10	11	12
10μM	A	BAPTA Ca ²⁺	BAPTA Ca ²⁺	BAPTA Ca ²⁺	BAPTA Ca ²⁺	BAPTA Ca ²⁺	BAPTA Ca ²⁺	PSS	PSS	PSS	PSS.	PSS	PSS
3μM	B	BAPTA Ca ²⁺	BAPTA Ca ²⁺	BAPTA Ca ²⁺	BAPTA Ca ²⁺	BAPTA Ca ²⁺	BAPTA Ca ²⁺	BAPTA Ca ²⁺	BAPTA Ca ²⁺	BAPTA Ca ²⁺	PSS	PSS	PSS
1μM	C	BAPTA Ca ²⁺	BAPTA Ca ²⁺	BAPTA Ca ²⁺	BAPTA Ca ²⁺	BAPTA Ca ²⁺	BAPTA Ca ²⁺	BAPTA Ca ²⁺	BAPTA Ca ²⁺	BAPTA Ca ²⁺	PSS	PSS	PSS
0.3μM	D	BAPTA Ca ²⁺	BAPTA Ca ²⁺	BAPTA Ca ²⁺	BAPTA Ca ²⁺	BAPTA Ca ²⁺	BAPTA Ca ²⁺	BAPTA Ca ²⁺	BAPTA Ca ²⁺	BAPTA Ca ²⁺	PSS	PSS	PSS
0.1μM	E	BAPTA Ca ²⁺	BAPTA Ca ²⁺	BAPTA Ca ²⁺	BAPTA Ca ²⁺	BAPTA Ca ²⁺	BAPTA Ca ²⁺	BAPTA Ca ²⁺	BAPTA Ca ²⁺	BAPTA Ca ²⁺	PSS	PSS	PSS
0.03μM	F	BAPTA Ca ²⁺	BAPTA Ca ²⁺	BAPTA Ca ²⁺	BAPTA Ca ²⁺	BAPTA Ca ²⁺	BAPTA Ca ²⁺	BAPTA Ca ²⁺	BAPTA Ca ²⁺	BAPTA Ca ²⁺	PSS	PSS	PSS
0.01μM	G	BAPTA Ca ²⁺	BAPTA Ca ²⁺	BAPTA Ca ²⁺	BAPTA Ca ²⁺	BAPTA Ca ²⁺	BAPTA Ca ²⁺	BAPTA Ca ²⁺	BAPTA Ca ²⁺	BAPTA Ca ²⁺	PSS	PSS	PSS
0μM	H	BAPTA Ca ²⁺	BAPTA Ca ²⁺	BAPTA Ca ²⁺	BAPTA Ca ²⁺	BAPTA Ca ²⁺	BAPTA Ca ²⁺	BAPTA Ca ²⁺	BAPTA Ca ²⁺	BAPTA Ca ²⁺	PSS	PSS	PSS

APPENDIX D - FLIPR repeats for compounds 10, 13, 16 and 17

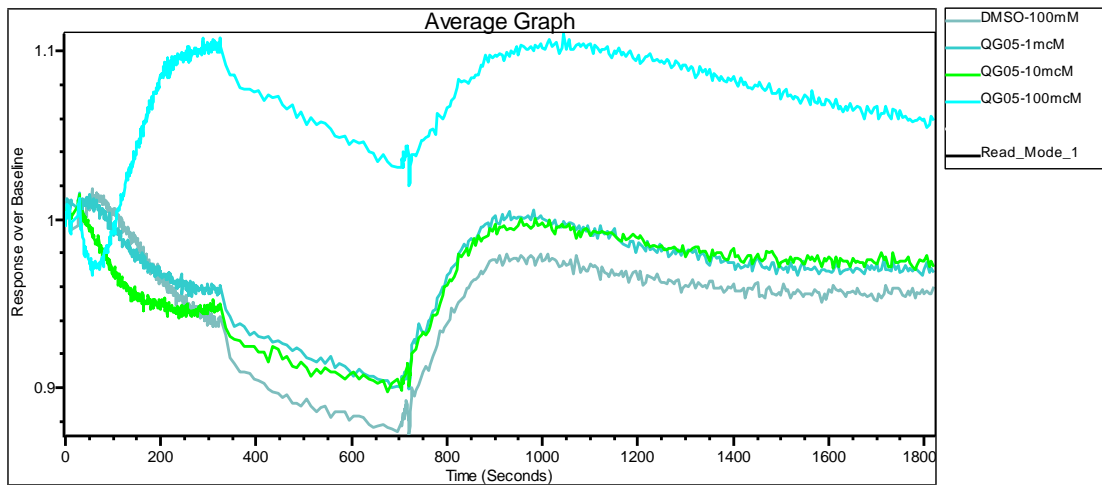
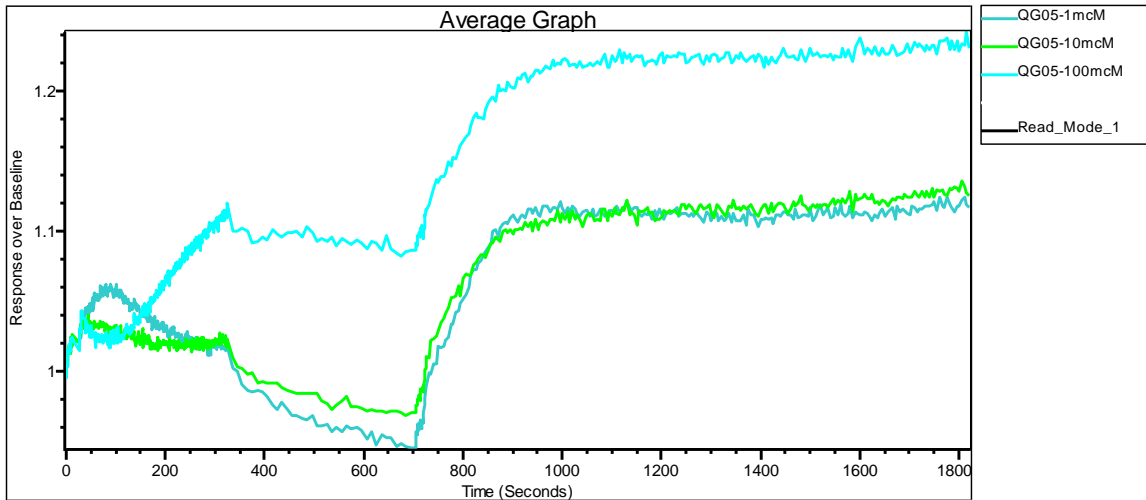
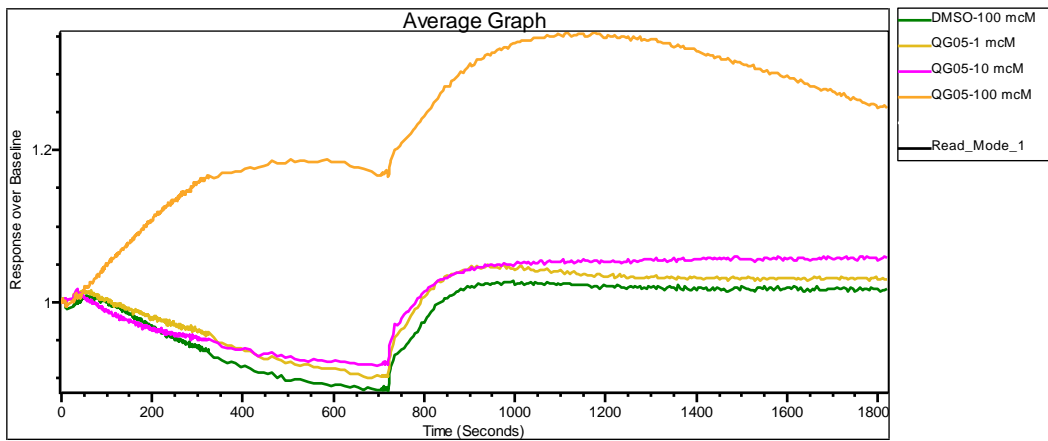
1. Compound 13 – three repeats in order of appearance



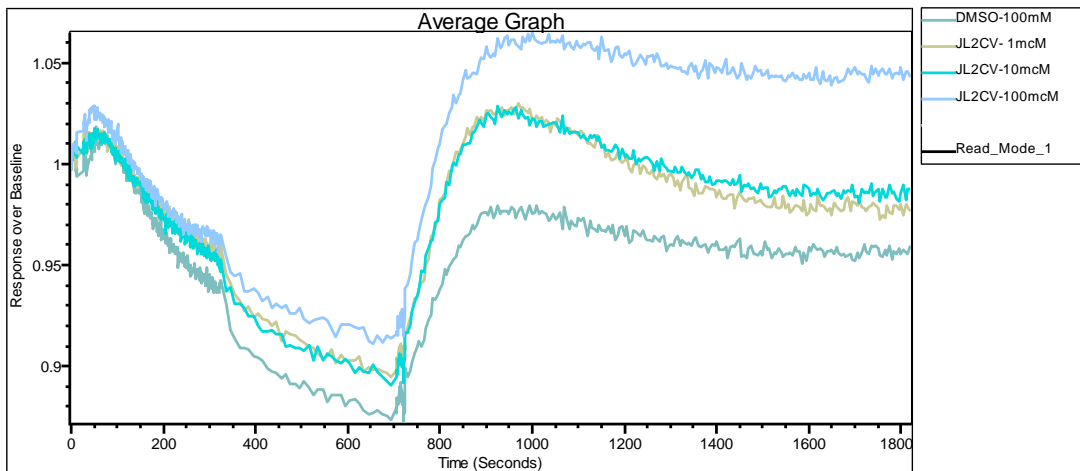
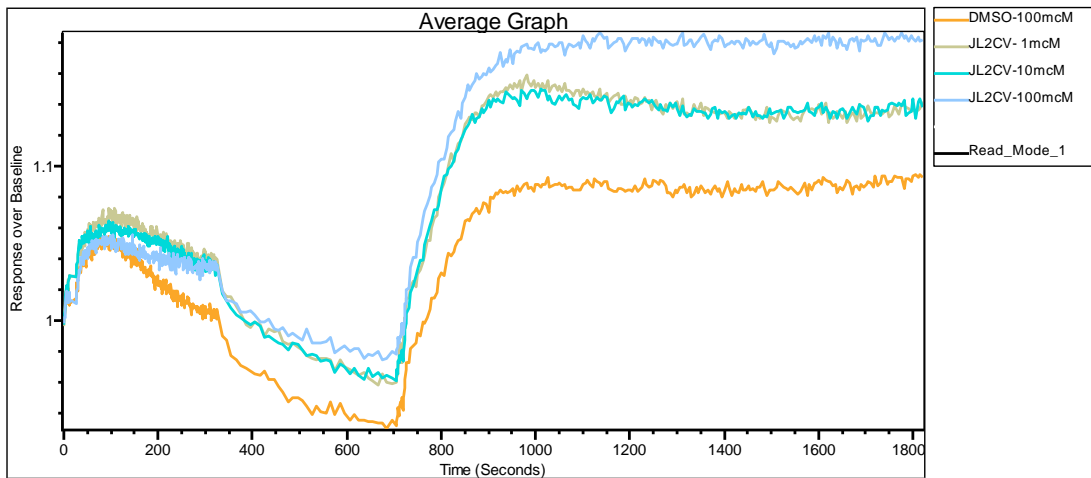
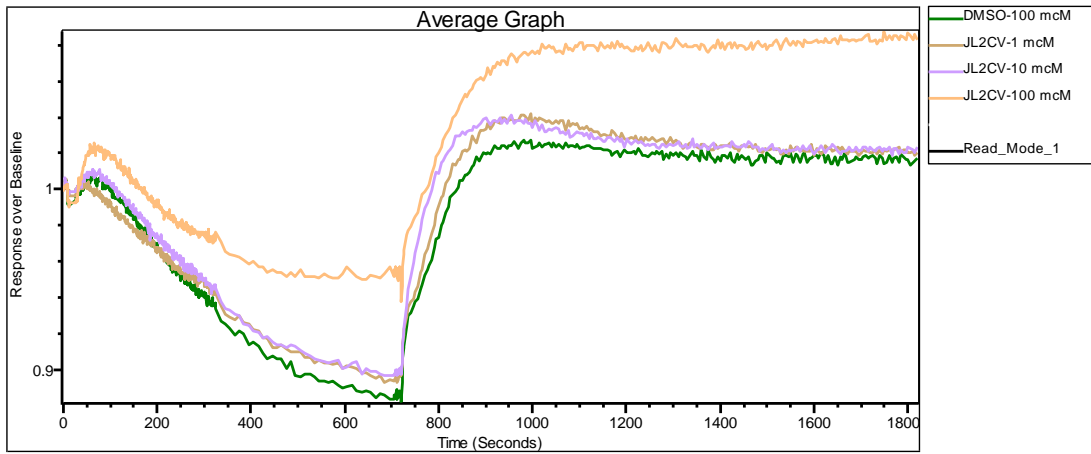
2. Compound 16 – three repeats in order of appearance



3. Compound 10 – three repeats in order of appearance

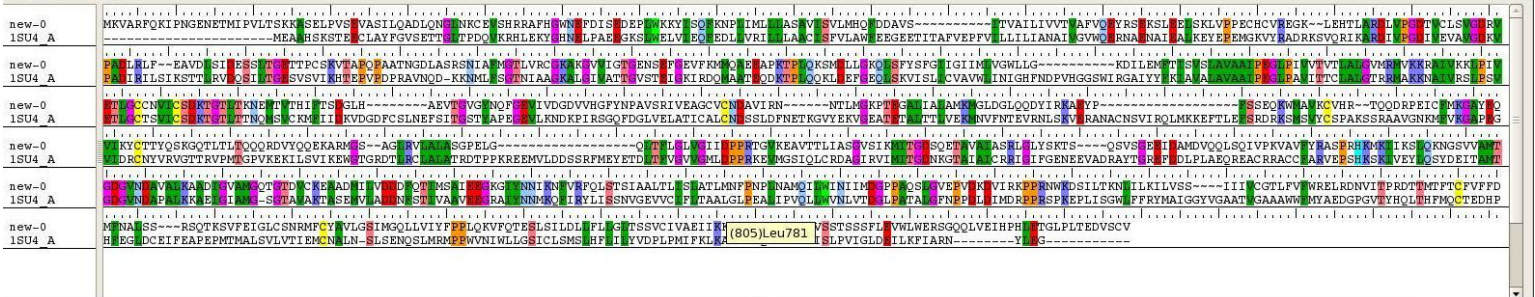


4. Compound 17 – three repeats in order of appearance



APPENDIX E – Homology model sequence alignments

SPCA1d model – Homolog: 1SU4



Find Homologs - Search for and/or import sequence homologs, then select those you wish to use as templates for structure building. Job Options...

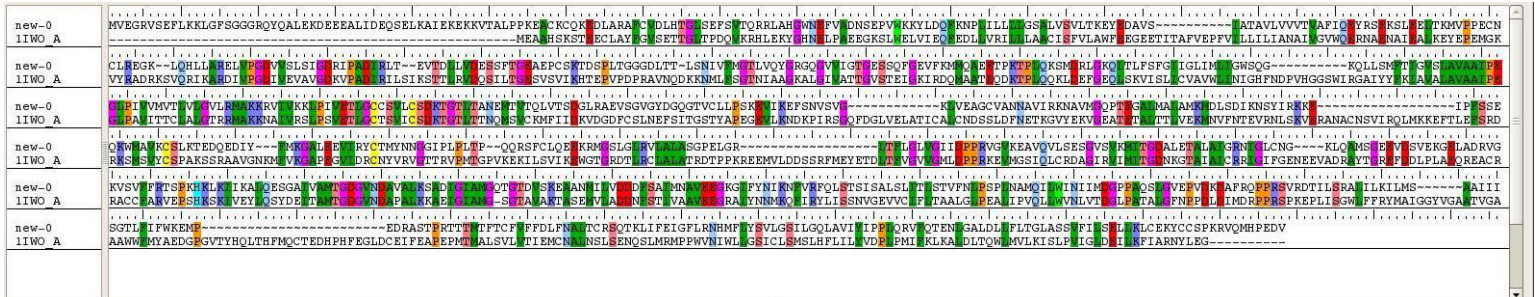
BLAST Homology Search Import:

Identify Globally Conserved Residues Family name:

Homologs:

ID	Name	Score	Expect	Identities	Positives	Gaps	Pfam	Title	Compound	Source	Experiment	Resolution	Ligands/Cofactors	Warning
1SU4_A	Imported Homolog			28%	45%	12%	E1-E2_ATPas...	CRYSTAL STRUCTUR...	HYDROLASE EC: 3.6...	ORYCTOLA	X-RAY DIFFRACTI...	2.4		

SPCA2 model – Homolog: 1IW0



Find Homologs - Search for and/or import sequence homologs, then select those you wish to use as templates for structure building. Job Options...

BLAST Homology Search Import:

Identify Globally Conserved Residues Family name:

Homologs:

ID	Name	Score	Expect	Identities	Positives	Gaps	Pfam	Title	Compound	Source	Experiment	Resolution	Ligands/Cofactors	Warning
1IW0_B	Imported Homolog			28%	45%	16%	E1-E2_ATPas...	CRYSTAL STRUCTUR...	HYDROLASE...	ORYCTOLA	X-RAY DIFFRACTI...	3.1	TG1 (OCTANOIC ACID [3S-[3ALPHA...	
1IW0_A	Imported Homolog			28%	45%	16%	E1-E2_ATPas...	CRYSTAL STRUCTUR...	HYDROLASE...	ORYCTOLA	X-RAY DIFFRACTI...	3.1	TG1 (OCTANOIC ACID [3S-[3ALPHA...	

SERCA1 model – Homolog: 3AR5

new-0
3AR5_A

new-0
3AR5_A

new-0
3AR5_A

new-0
3AR5_A

new-0
3AR5_A

(832)Gly831

Find Homologs - Search for and/or import sequence homologs, then select those you wish to use as templates for structure building.

Job Options...

BLAST Homology Search Options... Import: From File... From Project Table

Identify Globally Conserved Residues Family name:

Homologs:

ID	Name	Score	Expect	Identities	Positives	Gaps	Pfam	Title	Compound	Source	Experiment	Resolution	Ligands/Cofactors	Warning
3AR5_A	Imported Homolog			95%	97%	0%	E1-E2_ATPas...	CALCIUM PUMP CRY...	HYDROLAS...	ORYCTOLA...	X-RAY DIFFRACTI...	2.2	T61 (OCTANOIC ACID [3S-[3ALPHA...	

SERCA2 model – Homolog: 3AR4

new-0
3AR4_A

new-0
3AR4_A

new-0
3AR4_A

new-0
3AR4_A

new-0
3AR4_A

new-0
3AR4_A

Find Homologs - Search for and/or import sequence homologs, then select those you wish to use as templates for structure building.

Job Options...

BLAST Homology Search Options... Import: From File... From Project Table

Identify Globally Conserved Residues Family name:

Homologs:

ID	Name	Score	Expect	Identities	Positives	Gaps	Pfam	Title	Compound	Source	Experiment	Resolution	Ligands/Cofactors	Warning
3AR4_A	Imported Homolog			80%	88%	4%	E1-E2_ATPas...	CALCIUM PUMP CRY...	HYDROLAS...	ORYCTOLA...	X-RAY DIFFRACTI...	2.15	T61 (OCTANOIC ACID [3S-[3ALPHA...	

SERCA3 model – Homolog: 3AR7

```

new-0
3AR7_A  ...
new-0
3AR7_A  ...
new-0
3AR7_A  ...
new-0
3AR7_A  ...
new-0
3AR7_A  ...
new-0
3AR7_A  ...

```

Find Homologs - Search for and/or import sequence homologs, then select those you wish to use as templates for structure building. Job Options...

Import:

Family name:

Homologs:

ID	Name	Score	Expect	Identities	Positives	Gaps	Pfam	Title	Compound	Source	Experiment	Resolution	Ligands/Cofactors	Warning
3AR7_A	Imported Homolog			71%	83%	4%	E1-E2_ATPas...	CALCIUM PUMP CRY...	HYDROLAS...	ORYCTOLA...	X-RAY DIFFRACTI...	2.15	T61 (OCTANOIC ACID [3S-[3ALPHA...	