UCL SCHOOL OF PHARMACY
BRUNSWICK SQUARE

Synthesis of Novobiocin Based Anti-cancerous Hsp90 C- terminal Inhibitors

Guoxuan Sun

Supervisor: Dr Stephen Hilton

Date 2020.11.30

This research project is submitted in part fulfilment of the requirements for the PhD degree, UCL School of Pharmacy

Department of Pharmaceutical and Biological Chemistry

Abstract

The research described in this thesis focuses on heat shock protein 90 kDa (Hsp90), which is a molecular chaperone that contributes to the correct folding of around 200 proteins. It is an important protein that is associated with cancer survival and has been a focus of anticancer drug development for over two decades. Previous studies have shown that novobiocin (an FDA approved drug) acts as a weak inhibitor of Hsp90 via binding to the C-terminal domain (CTD) of Hsp90 and that modification of novobiocin has been shown to increase its anticancer activity. In this thesis, we designed and synthesized a series of analogues based on 4' substituted novobiocin that focused on four modifications to improve its anticancer efficacy: In **chapter 2** we aimed to: 1) replace the sugar moieties attached on the 4'-hydroxyl group; 2) substitute the 3'-amino group; 3) replace the novoise group with other sugars; in **chapter 3** we aim to: 4) Develop of a covalent targeting strategy to identify any active cysteine residues of Hsp90 CTD, which could react with Michael acceptors added to novobiocin 4'.

We designed and synthesized a number of glycosidic type novobiocin analogues via an optimized phase transfer catalysis. The resulted compounds showed increased antiproliferative activity. Protected sugar moieties are far less active than deprotected ones, however, difficulties of deprotection limited further development of such types of inhibitors. Nevertheless, we report compound **68**, an indole-amide analogue of novobiocin, to be a new lead structure with IC_{50} = 23.4 μM. Also, 4'-Acrylate Michael acceptor type novobiocin Hsp90 inhibitors mildly improve anti-proliferative activity and 4' substitution of cinnamic or sulfone fluoride group significantly increase the drug potency. Proteomic analysis of drug-Hsp90 complex confirmed the covalent modification of C597/C598. Overall, our project added new synthetic and biological knowledge into the design of non-covalent Hsp90 CTD inhibitors, more importantly, opened the gate for covalent Hsp90 CTD inhibitors drug design for the first time. Non-covalent candidate **68** and Michael acceptor type covalent candidate **108**, **138** are identified as lead structures.

Impact statement

Cancer is still the 2nd deadliest disease in 2018 and presents a significant threat to human health and life spent. Although a large number of targeted anti-cancer chemotherapy agents have been used and provided patient benefit, there is still a massive demand for new anticancer drugs due to the need for better drug efficacy and drug resistance of existed approved drugs.

Hsp90 is a promising anti-cancer drug target which is significantly overexpressed in cancer cells; for example, it occupies 55.6% of total protein expressed in colorectal cancer cells while only 3.7% in normal cells (Zhang, Guo et al. 2019). As a molecular chaperone that supports many biopathways, inhibition of Hsp90 will subsequently inhibit relevant oncoproteins, making it possible to inhibit multiple targets at the same time. However, currently, there is no approved drug of this type due to the high toxicity and harmful heat shock response found in Hsp90 NTD inhibitors. In recent decades, the Hsp90 CTD inhibitor has been shown to be an excellent alternative towards Hsp90 NTD inhibitors, triggering no heat shock response. Nevertheless, the lack of a drug-protein co-crystal structure limited its development and created a gap between detailed structural activity data and clinical trials.

In our first part of the project, we synthesised a series of glycosidic type Hsp90 CTD inhibitors based on the lead structure novobiocin. Our synthetic products significantly improved the anti-proliferative activity of lead structure, thus fulfilled the SAR in terms of glycosylation. We proved that such modification is beneficial and can be applied in future development or design. Also, we reported a phase-transfer glycosylation method for solving the challenging glycosylation step which can be referenced by any drug synthesis that requires non-traditional glycosylation.

More importantly, in our second part of the project, we applied the currently hot targeted covalent drug design strategy towards Hsp90 CTD, which has no synthetic covalent drug reported before. Using novobiocin as a lead scaffold, we designed and synthesised several types of covalent warhead substituted novobiocin analogues and evaluated them for their anti-proliferative activity, protein binding affinity, binding mechanism, and covalently modified residues.

We found that our covalent-warhead modification results in a significant increase of anti-proliferative activity, and Cys597/Cys598 of Hsp90 CTD are their covalent-modification targets. Through investigation of the binding mechanism, we also find that there are different binding affinity Kd values between quaternary dimeric Hsp90 inhibition and tertiary monomeric Hsp90 inhibition. Due to traditional mechanisms, we do not distinguish between dimeric and monomeric Hsp90 inhibition, but our research adds new knowledge to the understanding of Hsp90 inhibition, which will benefit the improvement of drug potency.

Since the first anti-cancer covalent drug afatinib was approved by FDA for the treatment of lung cancer in 2013, the development of covalent anticancer chemotherapy agents is soaring, bringing billions to the pharmaceutical industry and bringing hope to patients who have cancer. Our research no doubt opens the door for covalent drug development against Hsp90, of which no inhibitor (neither covalent nor non-covalent) has yet been approved. Taking advantage of previous successful experience in covalent drug design targeting EGFR and other targets, lots of new Hsp90 CTD, covalent drug candidates with different types of validated covalent warheads can be easily developed, as we have already proved the feasibility of this strategy in this project. Previous work from other groups worldwide has developed lots of promising non-covalent Hsp90 CTD scaffold and improved the drug potency into a 1- 10 micromolar level. Combining our covalent inhibition strategy, Hsp90 inhibitors may have a good chance to obtain nanomolar efficacy which is potent enough for a clinical trial. We identified some compounds as new inhibitors of Hsp90 CTD, which has not yet been reported (e.g. sulfone fluoride, disulfiram). These secondary outcomes expand the range of anti-Hsp90 drug discovery, and further investigation of these lead structures will result in new Hsp90 CTD inhibitors.

Acknowledgements

First of everything I would like to sincerely thank to my supervisor Dr. Stephen Hilton. I am so lucky to be his student during my last year of Ph.D. I thank him for his professional knowledge, patience, warmness and all kinds of help. Moreover, he saved my broken soul and heart, lightening my path of science. Without him, everything would be much worse upon life and academic study. I feel grateful and fortunate since he took me on as his student.

Secondly, I would like to thank all the experimental collaborators:

Prof. Andreas Schatzlein, Dr. Clemens Michael Smola and Dr. Francisco Humberto Xavier-Junior from Pharma & Biochemistry UCL School of Pharmacy for providing the facilities and expertise to carry out anti-proliferative tests. Also, I thank them for the important help and suggestions upon my project.

My friend Banghao Yuan from Pharmacology UCL School of Pharmacy for method development of native page gel electrophoresis as well as Hsp90 protein supply and Sizhu Lu from the University of Oxford for carrying out the add-ups anti-proliferative tests.

Dr. Kate Heesom from the Proteomics Facility University of Bristol for professional support of proteomic analysis. I am grateful for the fast response and patience.

Staff of ESPRC UK National Mass Spectrometry Facility for technical support and method development of MALDI-TOF mass spec analysis.

There are also people who helped me during my 4 years' life, and I would like to acknowledge:

Colleagues from the Hilton Group: Dr. Ben Bowles, Marpin, Dr Zaid Muwaffak, Dr. Ahtsham Ishaq and Dr. Matthew Penny. Especially Ash and Matt for the useful suggestions and discussions for my projects. Although I have been alone for most of the time due to autism, all the pieces of memory with you

guys are treasured deep inside in my mind. I will remember all your help and smiles.

Dr. Frank Kozielski from Pharma & Biochemistry UCL School of Pharmacy and his student Jiaqi Lou for the help and guidance of Microscale thermophoresis (MST). Also, I would like to sincerely thank Frank for his care and effort in solving my grant money crisis together with my upgrade.

Dr. Jasmina Jovanovic from Pharmacology UCL School of Pharmacy for the help of solving my grant money crisis and other non-academic difficulties.

Ms. Sabine Heinrich for the support of office and communications. Her kindness and caring will be always remembered.

Dr. Geoff Wells from Pharma & Biochemistry UCL School of Pharmacy for the help and guidance during my upgrade.

My Chinese fellows: Heyu Li, Liu Dou, Tiantian Ma, Yang Mai, Yixi Zhang for the golden friendship. You guys helped me out through my hardest time of the second and third year.

Previous colleague Dr. Xuexia Huang from UCL School of Pharnacy, Jialin Zhang from University Art London for the help about proteomic analysis and native page gel method development.

Y L from University of , my love. Special thanks for her support and love. You are one of the reasons that make me stand strong, stand still, and continue doing research.

Dr. Min Yang, my previous supervisor for introducing me to the field of Hsp90 C-terminal inhibitors.

Last, I would like to thank my sponsor, my mum and dad, Y Zhang, and Y Sun. I shame that I am the black sheep of the family who still asks for money from parents even in his 28's. Only God knows how hard you have worked through these years, trying to pay for everything of my life here with the income of a normal high school lecturer. This thesis is for science and for you.

Plagiarism Statement

I, Guoxuan Sun, hereby confirm that the work submitted in the thesis is my own. Any ideas submitted in this thesis are my own. Any ideas, quotations, and paraphrasing from other people's work and publications have been appropriately referenced. I have not violated the School of Pharmacy's policy on plagiarism.

Signature ....

Date...30/11/2020.....

Abbreviations

AHA1:	Activator of heat shock protein 90 kDa ATPase 1
Akt:	Protein kinase B
Apaf-1:	Apoptotic protease activating factor 1
ATP:	Adenosine triphosphate
A2780:	Human ovarian carcinoma
A549:	Adenocarcinomic human alveolar basal epithelial lung cancer cells
BCR-ABL:	Philadelphia chromosome
Bcl2:	B-cell lymphoma 2
BTAB:	Benzyltributylammonium bromide
BTK:	Bruton's tyrosine kinase
Cdk4:	Cyclin-dependent kinase 4
Cdk6:	Cyclin-dependent kinase 6
CHCA:	Alpha-cyano-4-hydroxycinnamic acid
CLL:	Chronic lymphocytic leukaemia
C-raf:	Proto-oncogene serine/threonine-protein kinase
COX:	Cyclooxygenase
CTD:	Carboxyl terminal domain
CV:	Column volume
DCM:	Dichloromethane
DHB:	2,5-Dihydroxybenzoic acid
DNA:	Deoxyribonucleic acid

EAE:	Experimental autoimmune encephalomyelitis
EDCI:	1-Ethyl-3-(3-dimethylaminopropyl)-carbodiimide
EGCG:	Epigallocatechin-3-Gallate
EGFR:	Epidermal growth factor receptor
EmL-ALK:	Echinoderm microtubule-associated protein-like - Anaplastic lymphoma kinase gene
eNOS:	Endothelial nitric oxide synthase
ErbB-2:	Human epidermal growth factor receptor 2
ESI:	Electro Spray Ionization
FDA:	United States Food and Drug Administration
GA:	Geldanamycin
HCl:	Hydrochloric acid
HER-2:	Receptor tyrosine-protein kinase erbB-2
HIF-1:	Hypoxia-inducible factors
HPLC:	High performance liquid chromatography
HSF1:	Heat shock factors 1
Hsp:	Heat shock protein
Hsp90:	Heat Shock Protein 90 kDa
H1299:	Human non-small cell lung carcinoma cell line derived the lymph node
IGF-1R:	Insulin-like growth factor 1
IgG:	Immunoglobulin G
IL-6:	Interleukin 6

INOS:	Inducible nitric oxide synthase
KA:	Kongensin A
LC:	Liquid Chromatography
Lck:	Lymphocyte-specific protein tyrosine kinase
MALDI:	Matrix-assisted laser desorption/ionisation
MCF-7:	Michigan Cancer Foundation-7
MCL:	Mantle cell lymphoma
MIA PaCa-2:	Human pancreatic cancer cell line
Myt-1:	Myelin transcription factor 1
min:	Minutes
MS:	Mass Spectrometer
MST:	Microscale thermophoresis
MTT:	3-(4,5-Dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide
NADPH:	Triphosphopyridine nucleotide
NMR:	Nuclear Magnetic Resonance
NSCLC:	Non-small cell lung cancer.
NTD:	Amino terminal domain
OTs:	Tosylate
PEN-A:	Penicisulfuranol A
Plk-1:	Serine/threonine-protein kinase
PPI:	Proton pump inhibitor
PTC:	Phase Transfer Catalysis

$P2Y_{12}$:	Chemoreceptor for adenosine diphosphate
P450:	Cytochrome 450
p53:	Tumor protein p53
Raf:	Proto-oncogene serine/threonine-protein kinase
RT:	Retention Time
SAR:	Structure-activity Relationship
SDS:	Sodium dodecyl sulfate polyacrylamide
sec:	Seconds
Src:	Steroid receptor coactivator-1
TCI:	Targeted covalent inhibition or Targeted covalent inhibitor
THF:	Tetrahydrofuran
TLC:	Thin Layer Chromatography
TOF:	Time of flight
UV:	Ultraviolet
U87-MG:	Human primary glioblastoma cell line
VEGFR:	Vascular endothelial growth factor
ViB:	Vibsanin B

Contents

List of Figures

List of Schemes

List of tables

Chapter 1: General Introduction

1.1 Cancer

Cancer is one of the predominant causes of morbidity and mortality in the world. According to the World Health Organization, about 8.8 million people worldwide died from cancer in 2018 (Miller, Fidler‐Benaoudia et al. 2020). The number of new cases is expected to rise by 70% over the next two decades. In 2000, a study proposed that the complex array of phenotypes displayed by cancer cells may be organized into six cancer-associated traits. These include the ability of cells: to proliferate indefinitely (immortalization), to become independent of extra-cellular growth or anti-growth signals, to avoid apoptosis, to induce a self-sustained supply of nutrients and oxygen (angiogenesis), and ultimately, to invade and metastasize to distant sites (Figure. 1, Hanahan and Weinberg, 2000).

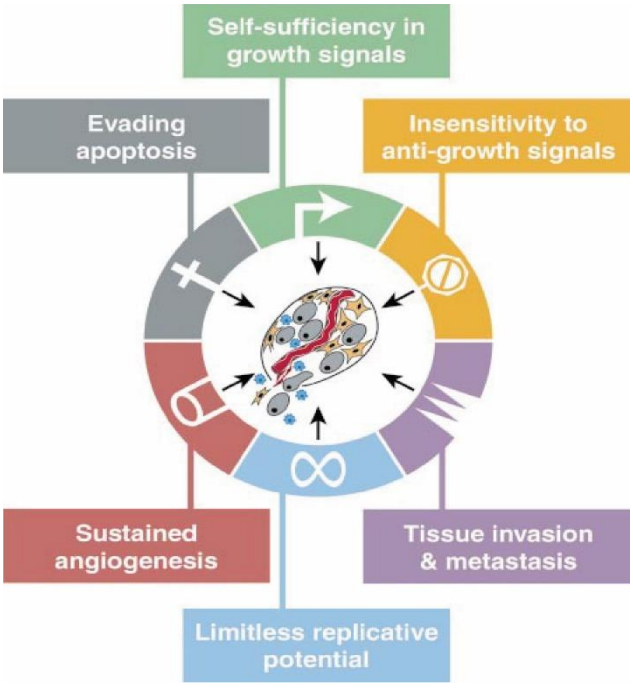


Figure 1: The hallmarks of cancer (Hanahan and Weinberg).

These cancer-associated hallmarks refer to genetic alterations in multiple ‘safe-guard’ genes responsible for the regulation and the tight co-ordination of diverse processes such as cell survival, proliferation, growth, differentiation, and motility(Blume-Jensen and Hunter 2001). As the discovery of cancer

hallmarks greatly enhanced the understanding of the molecular and biochemical alterations that occur during malignant transformation of normal cells, a sophisticated method of anticancer drug discovery was proposed and led to the design and development of selective and personal anticancer chemotherapies, acting on specific single molecular targets (Figure. 2, Collins and Workman, 2006).

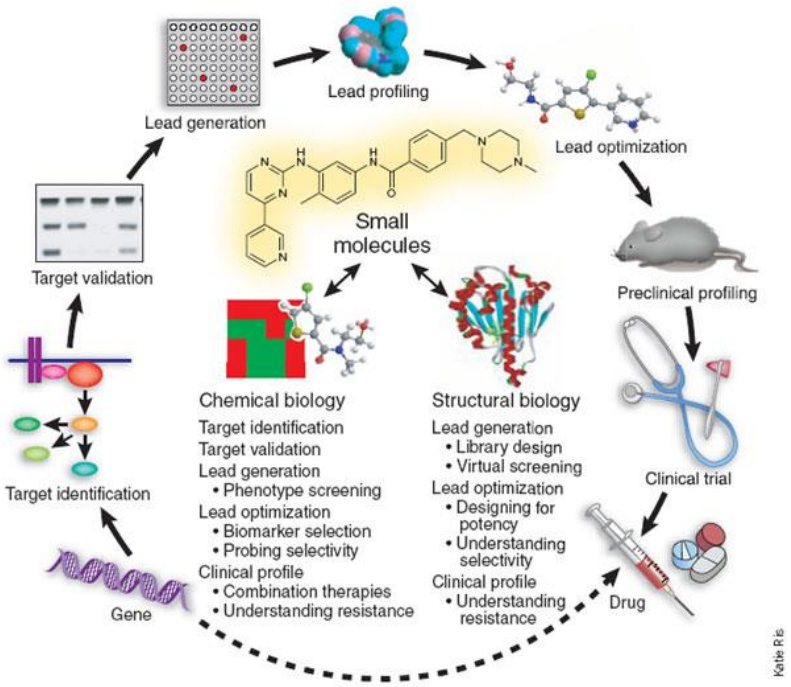


Figure 2**:** Targeted cancer therapy development cycle (Collins and Workman 2006).

This method bridges the knowledge between the biochemical effects of small molecules and the structural biology of characterized genes and proteins that support the survival of cancer cells. Through this method, many new anticancer drugs (Table 1) targeting different genetic targets related to the different hallmarks were produced.

Table 1**:** Examples of anticancer personalized drugs.

Drugs	**Type**	**Targets**	**Cancer type**
Dasatinib	Small molecule	BCR-ABL, Src	Lukemia
Trastuzumab	Monoclonal antibody	ErbB-2	Breast, stomach
Gefitinib	Small molecule	EGFR	Breast, lung
Sorafenib	Small molecule	VEGFR, C-RAF, PDGFR	Kidney
Bevacizumab	Monoclonal antibody	VEGF-A	Colorectal
BAY43-9006	Small molecule	RAF	Melanoma

Cancer cells are genetically unstable, living under specific stresses such as the presence of mutated and post-translational, dysregulated signalling proteins, chromosome and microsatellite instability, hypoxia, and low nutrient concentrations. The survival of cancer cells in such stressed environments depends on rapid selection of adaptive mutations or activating alternative signalling pathways (Xiao, Lu et al. 2006). As a result, the effectiveness of anticancer drugs that specifically target a single oncoprotein or signalling pathway may gradually be decreased, or even totally lost, due to the genetic and epigenetic variation of cancer cells, even within a tumour (Hong, Banerji et al. 2013). This indicates that a molecular target that affects more than one oncogenic protein and pathway will be advantageous. Therefore, an ideal candidate is the molecular chaperone, due to its role as a central node in signalling networks and coordinating with 60% of cell survival proteins (Figure. 3). Therefore inhibition of a molecular chaperone can lead to the combinatorial depletion of an array of client oncoproteins and simultaneously suppressing multiple oncogenic pathways (Sidera and Patsavoudi 2014).

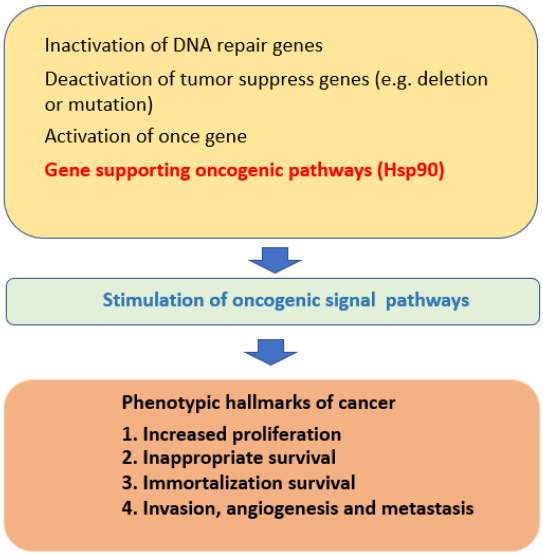


Figure 3: Hsp90 is a supportive element to the growth of cancer (Collins and Workman 2006)

1.2. Molecular chaperones and Hsp90

1.2.1. The function of molecular chaperones in Cancer

The molecular chaperone is a type of protein that interacts and aids in the conformational folding or unfolding of the assembly or disassembly of another protein without being part of its final structure (Hartl 1996). Chaperones are present when macromolecules perform their normal biological functions and have correctly completed the processes of folding and assembly. The first molecular chaperone discovered, was found to assist in the assembly of nucleosomes from folded histones and DNA (Richardson, Alekseev et al. 2006). Such chaperones, especially in the cell nucleus, are concerned with the assembly of folded subunits into oligomeric structures (Richardson, Alekseev et al. 2006). Among molecular chaperones, the most ubiquitous species are stress proteins and heat shock proteins (Hsp). Heat shock proteins are a family of molecular chaperone proteins that are produced by cells in response to exposure to stressful conditions (Ritossa 1962). They were first described in relation to heat shock, but are now known to be expressed during other stresses including exposure to cold, UV light, and during wound healing or tissue remodelling (Matz, Blake et al. 1995). When exposed to stressful

environments (e.g. heat, shock, acidosis, hypoxia, or heavy metals), cells will rapidly increase the expression of Hsp's to resist the stress and prevent cell death (Bagatell and Whitesell 2004). Hsp's are found in all living organisms ranging from bacteria to humans. They are named according to their molecular weight, for example, Hsp60 (60 kDa), Hsp70 (70 kDa, 71 kDa, and 72 kDa), Hsp90 (90 kDa), and Hsp100 (104 kDa and 110 kDa). The functions of Hsp's are important for protein-protein interactions such as folding and assistance in the establishment of proper protein conformation and prevention of unexpected protein aggregation. In eukaryotes, Hsp70 and Hsp90 are among the most prominent members of this family. Although Hsp70 has a broad substrate range mainly stabilizing the protein primary and secondary structure, Hsp90 is responsible for the stabilization and folding of protein tertiary steric confirmation after their primary folding and has a narrower range of clients (Karras, Yi et al. 2017). Hsp90 acts rather late in the maturation process of a client protein, which, as a result, is believed to add an additional regulation step (Jhaveri, Taldone et al. 2012).

Malignancy demands which cancer cells acquire new biological capabilities, collectively described as the hallmarks of cancer (Hanahan and Weinberg). Stresses induced by can arise from a host of intracellular and extracellular processes, including damage to macromolecules such as proteins, lipids, and nucleic acids; restricted access to nutrients; increased demand for biosynthetic processes; exposure to toxins; and attack from immune cells (Jaeger and Whitesell 2019). As a result, the stress response manifested by Hsp's is critical to the development of cancer hallmarks and the malignancy state (Hanahan and Weinberg 2011). The heat shock response of Hsp's, which in normal cells serves as a protective mechanism, is hijacked by cancer cells to assist in their growth and survival (Davidson, Jonas et al. 2017). Overexpression of molecular chaperones increases the ability of cancer cells to combat the most challenging issue, of protein misfolding, which is due to their nature of rapid reproduction (Hartl 1996). Decades of fundamental and systematic observations have led to the conclusion that the ability of cancer cells to tolerate the stresses associated with their malignant phenotype is intricately connected to their ability to combat protein misfolding

(Whitesell and Lindquist 2009). The authors also concluded that for cancer cells to combat protein misfolding, molecular chaperones, especially Hsp90, serve as a buffer from the very first stage of cancer development (Jaeger and Whitesell 2019).

1.2.2 Hsp90

Hsp90 is a 90 kDa Hsp (Pearl, Prodromou et al. 2008) which delivers essential ATP dependent chaperone functions to more than two hundred client proteins involved in cell growth and survival (Li, Soroka et al. 2012). More than 20% of them, including overexpressed or mutant oncogenic proteins ErbB2/HER2 (Breuer et al., 1994), Braf (Grbovic, Basso et al. 2006), Akt/PKB (Sato, Fujita et al. 2000), mutated p53 (Blagosklonny, Toretsky et al. 1996), transcription factors: hormone steroid receptors GR (Grad and Picard 2007), ER and AR angiogenic factors HIF-1α (Picard 2006, Kuduk 2000, Johnson and Toft 1995), telomerase (Forsythe, Jarvis et al. 2001), are associated with the six hallmarks of cancer. Up to 20 co-chaperones of Hsp90 have been discovered (Zhao, Yan et al. 2012). Furthermore, it is also reported that over 60% of the total cellular proteins are shown to be either co-chaperone or client proteins of Hsp90 (Figure. 4) playing a multitude of different roles in various kinds of bioactivities (Sofis, Jarmolowicz et al. 2017). Almost all proteins need assistance from the chaperone machinery to adopt their functional three-dimensional conformations within the crowded, highly concentrated environment of the cell. Many proteins that associate with cell proliferative activity but lack mature stability specifically require the Hsp90 pathway to achieve maturation stability (Verba and Agard 2017). HSP90 physically associates with these HSP90 clients and, through iterative steps of interaction, holds them in a metastable inactive state, subsequently transferring them into an active conformation (Jaeger and Whitesell 2019).

The scope of HSP90-client proteins is notably diverse, and various kinds of functional (Figure.4) categories are represented among them.

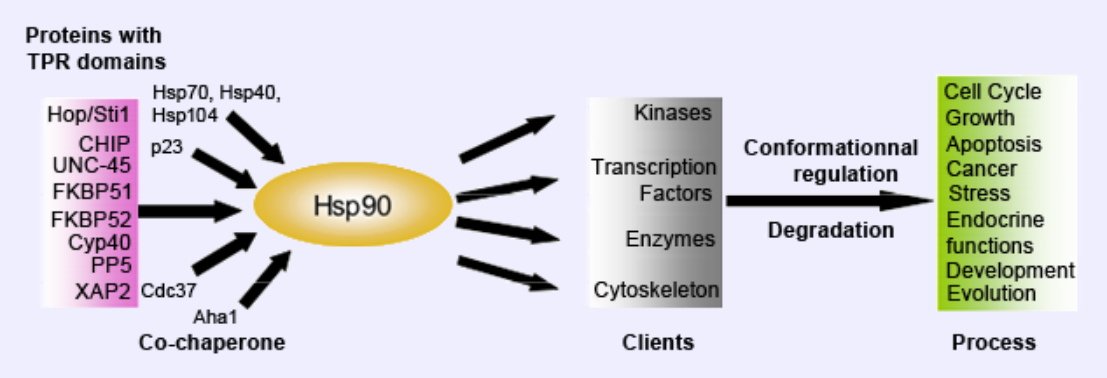


Figure 4**:** Hsp90 Protein partners and clients destabilized by Hsp90 inhibition (Jackson, Queitsch et al. 2004).

Several isoforms of Hsp90 have been isolated from the human proteome including Hsp90α (inducible form, cancer-related) and Hsp90 β (constitutive form, embryonic), which are located in the cytoplasm (Subbarao Sreedhar, Kalmár et al. 2004) and TRAP, which is located in mitochondria (Felts, Owen et al. 2000). Under non-stressful conditions, the quaternary structure of Hsp90 is established to be a dimeric complex. Moreover, recent studies had indicated that Hsp90 is not confined to the intracellular environment, and differences are found between intracellular and extracellular Hsp90. Pertaining to cancer cells, whilst intracellular Hsp90s functions to cell survival, the level of Hsp90 on the cell surface is higher in cancer cells than in normal cells and correlates with metastatic activity (Trepel, Mollapour et al. 2010). Selective inhibition of extracellular Hsp90, whilst not manifesting an anti-proliferative effect does achieve anti-metastasis activity (Stivarou, Stellas et al. 2016). This phenomenon presents potential options for late stage cancer treatment, as currently, most clinical cytotoxic drugs have a limited effect on preventing metastasis (Sidera and Patsavoudi 2014). Most recently, in 2018, Hsp90 was reported to directly mediate membrane deformation and control exosome release, which is totally independent of its chaperone function (Lauwers, Wang et al. 2018). This novel activity occurs through an evolutionarily conserved sequence located in the Hsp90 middle domain. Mutation, middle-domain inhibition, and C-terminus domain inhibition all result in silencing of this function (Lauwers, Wang et al. 2018). The discovery of this non-chaperone

machinery suggested that the current understanding of the role of Hsp90 is incomplete.

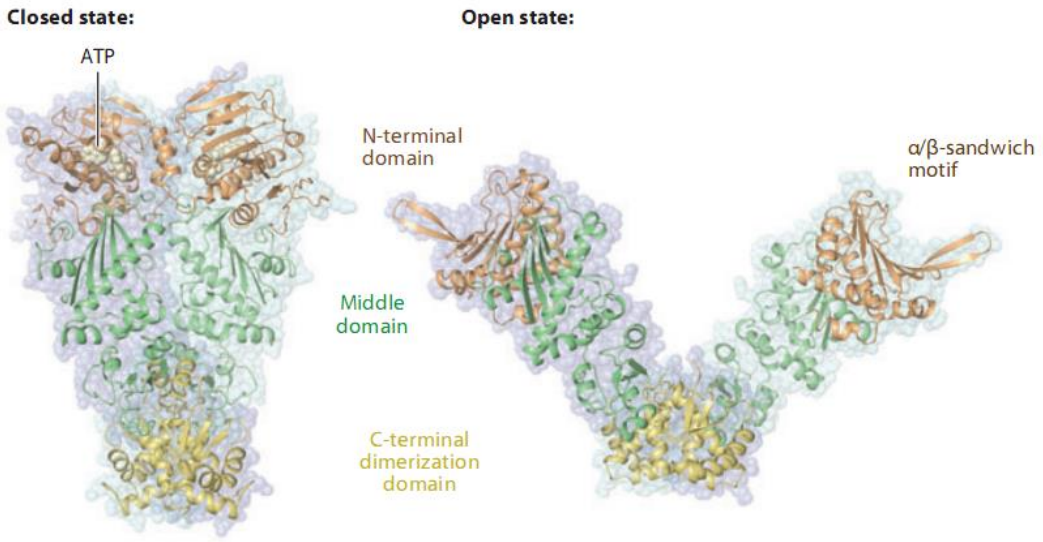


Figure 5: Structure of Hsp90 (2CG9) with 3 different domains and close/open states (Young, Schneider et al. 1997).

Hsp90 consists of three domains, the N-terminal domain (NTD), a middle domain implicates the client binding, and a C-terminal dimerization domain (CTD) (Harris, Shiau et al. 2004) (Shiau, Harris et al. 2006). The Hsp90 CTD is claimed to contain an alternative ATP-binding site (Sőti, Pál et al. 2005); however, there is no ATP bound Hsp90 CTD crystal structure recorded in the Protein Data Bank (Simon, Huang et al. 2017). The available structural information (Figure. 5) clearly shows that the CTD is involved in dimerization which is independent of nucleotide-binding or protein interactions (Mayer, Prodromou et al. 2009). On the other hand, the N-terminal domain will undergo a conformational change on binding of ATP (Mayer, Prodromou et al. 2009). Association of Hsp90 with client proteins is regulated by the activity of the N-terminal ATPase binding domain, which hydrolyses ATP to mediate a series of association-dissociation cycles between Hsp90 and its clients.

Recent structural evidence supports a unified model for the conformational changes of Hsp90, containing numerous distinct steps (Figure. 6) (Frydman, 2009; Neckers, 2009). The rate of ATP hydrolysis and conformational changes

are variable, while some forms of Hsp90 adopt the 'closed' conformation even in the absence of ATP (Krukenberg, Förster et al. 2008), thus, contrasting earlier results that suggested that conformational changes were strictly dependent on the nucleotide. A more recent study suggests that nucleotide shifts the equilibrium between distinct conformations (Sidera and Patsavoudi 2014).

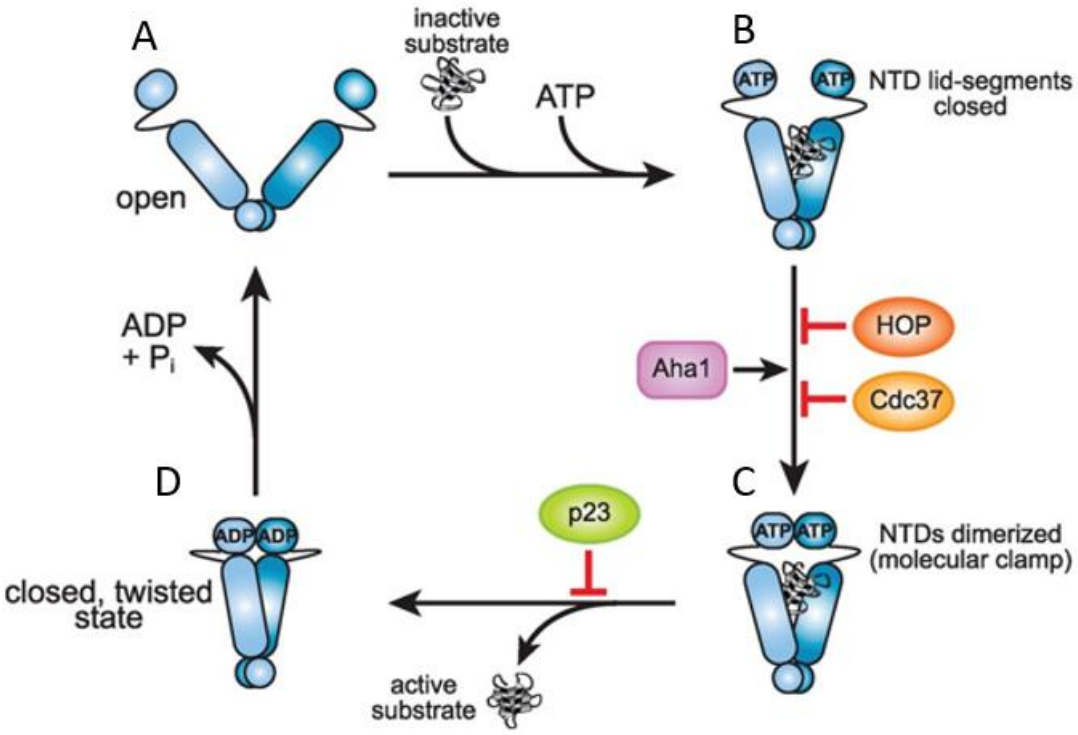


Figure 6**:** Hsp90 Conformational cycle: (A) In the absence of nucleotide, Hsp90 is dimerized at the carboxy-terminus resulting in an open conformation. (B) Nucleotide-binding induces the closing of a lid over bound nucleotide. This is followed by association of NTDs. The process was assisted by co-chaperon HOP, Cdc37 and Aha1 (C) In the closed conformation, portions of the amino-terminal domain 'cross-over' to associate with the other promoter. (D) The Hsp90 dimer then forms a compact twisted structure that results in the association of a flexible loop from the middle domain with the nucleotide-binding pocket. This structure is capable of ATP hydrolysis. After hydrolysis, Hsp90 returns to the open conformation (Mayer, Prodromou et al. 2009).

1.2.3 Hsp90 in Cancer

It is well established that Hsp90 clients are associated with all six hallmarks of cancer (Table 2, Sidera and Patsavoudi, 2014). This is due to Hsp90 serving as a predominant contributor to the development of all cancer hallmarks by supporting mutated oncoproteins, which should be removed in normal cells (Jaeger and Whitesell 2019).

Table 2**:** Multiple hallmarks of cancer and Hsp90 client protein (Sidera and Patsavoudi, 2014).

Hall Mark	**Client Protein**
Evasion of Apoptosis	Akt, Rip, p53, Survivin, Apaf-1, Bcl2, IGF-IR
Sustained Angiogenesis	VEGFR, HIF1, Akt, FAK, Src
Limitless replicative potential	Telomerase, n-TERT
Tissue invasion and metastasis	MMP-2, c-MET
Self-sufficiency in growth signals	EGFR, Raf, Bcr-Abl, ErbB-2, Src, Akt, MEK
Insensitivity to anti-growth signals	Plk-1, Cdk4, Cdk6, Myt-1, cyclin D

Hsp90 plays a vital role in maintaining transformation and supporting survival and growth potential of cancer cells, attributed to its essential chaperoning functions that facilitate the acquisition and maintenance of the malignant phenotype. Moreover, cancer cells experience various types of stress in their microenvironments such as acidosis, hypoxia, and nutrient deprivation. They therefore exhibit higher requirements for Hsp90 function in order to tolerate alterations and survive everchanging hostile, stressful conditions. Hsp90 is frequently regenerated in tumour cells, in transformed cells (Lebeau, Le Chalony et al. 1991) malignancies of the hematopoietic system (Chant, Rose et al. 1995) and in solid tumours of various entities including ovarian and endometrial carcinomas (Nanbu, Konishi et al. 1996), as well as in breast (Scaltriti, 2012; De Mattos-Arruda, 2012), lung (Ruiz, Floor et al. 2008), prostate (Solit, Scher et al. 2003) and gastrointestinal cancers (Li, Huang et

al. 2008). One clinical study reported Hsp90 occupied 55.6% of total protein expressed in colorectal cancer cells, in contrast to a significantly lower 3.7% of total protein expression in normal cells (Zhang, Guo et al. 2019). At the protein level, Hsp90 functions as a biochemical buffer for the numerous genetic lesions that are present within tumours, allowing mutant proteins to retain or gain function while permitting cancer cells to tolerate the imbalanced signalling of these oncogenic proteins and thus escape apoptotic death (Whitesell and Lindquist 2005). Inhibition of Hsp90 simultaneously inhibits multiple pathways that are crucial for cancer survival (Workman, Burrows et al. 2007). Furthermore, Hsp90 has not been found frequently to acquire gain-of-function mutations or amplification, indicating cancer cells do not utilize Hsp90 as a driver of abnormal growth but rather as an enabler of the hallmarks of the malignant phenotype (Whitesell and Lindquist 2005). Cancer cells are not only addicted to oncogenes but are also addicted to Hsp90 expression (Biebl and Buchner 2019). The level of intracellular Hsp90 has been used as a biomarker for the diagnosis of liver cancer (Sun, Zang et al. 2010), and has also been shown to be negatively correlated with poor prognosis of several types of cancer (Dimas, Perlepe et al. 2018). For both genetic evolution and cancer development, Hsp90 fulfils the same function of an accumulator of genetic verification by rescuing the function of proteins encoded through mutations or polymorphisms that would otherwise have a deleterious effect on stability or function (Rutherford and Lindquist 1998). Beyond affecting cancer development, this function also manifests cancer metastasis, drug resistance, and tumour heterogeneity (Jaeger and Whitesell 2019). As shown (Figure 7), Hsp90 is critical for cancer cell survival, and hence presents an important anticancer drug target.

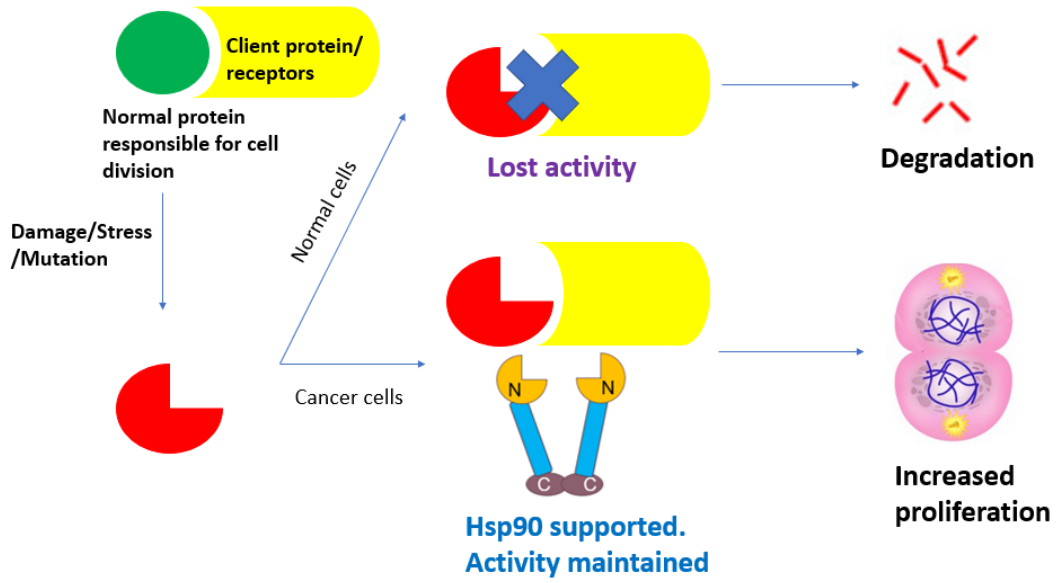


Figure 7: Mutated or damaged protein can maintain its activity via Hsp90 buffer; the mechanism is adopted by cancer for the demand of rapid proliferation.

Furthermore, despite Hsp90 being an abundant protein, Hsp90 inhibitors typically accumulate in tumours rather than in healthy tissues and thus are able to selectively destroy tumour cells over normal cells (Hong, Banerji et al. 2013). This therapeutic selectivity of Hsp90 inhibitors results from the presence of a predominantly high-affinity, activated form of Hsp90 in tumours, where in contrast Hsp90 in normal tissues is in a low-affinity, inactive form (Moulick, Ahn et al. 2011).

1.3 Hsp90 Inhibitors

In general, Hsp90 inhibitors work by binding to the NTD with an adenine binding pocket, which is responsible for the protein's ATPase activity (Figure 8, top side). Beside the NTD, a few small molecules including novobiocin (Marcu, Chadli et al. 2000), epigallocatechin gallate (Zhao and Blagg 2013), triazole-based chemicals (Terracciano, Chini et al. 2013), and other natural compounds are found to have a good binding affinity towards the C-terminal of Hsp90 (Figure 8, bottom side). In contrast to the widely established N-terminal binding drugs, there are only two human Hsp90 C-terminal crystal

structures with low resolution available in the Protein Data Bank (3Q6N, 3Q6M, Cheng-Chung Lee 2011) but no ligand bound Hsp90 C-terminal crystal structure has been reported. As a result, the binding mechanism towards CTD is poorly understood which has limited the discovery of Hsp90 CTD inhibitors. Other pathways of inhibition mostly target the assisted agent of Hsp90, for example targeting co-chaperone (Aha1, HOP) /Hsp90 interaction, Cdc37/Hsp90interaction, TPR co-chaperones/Hsp90 interaction, and client/Hsp90 association (Pearl 2005).

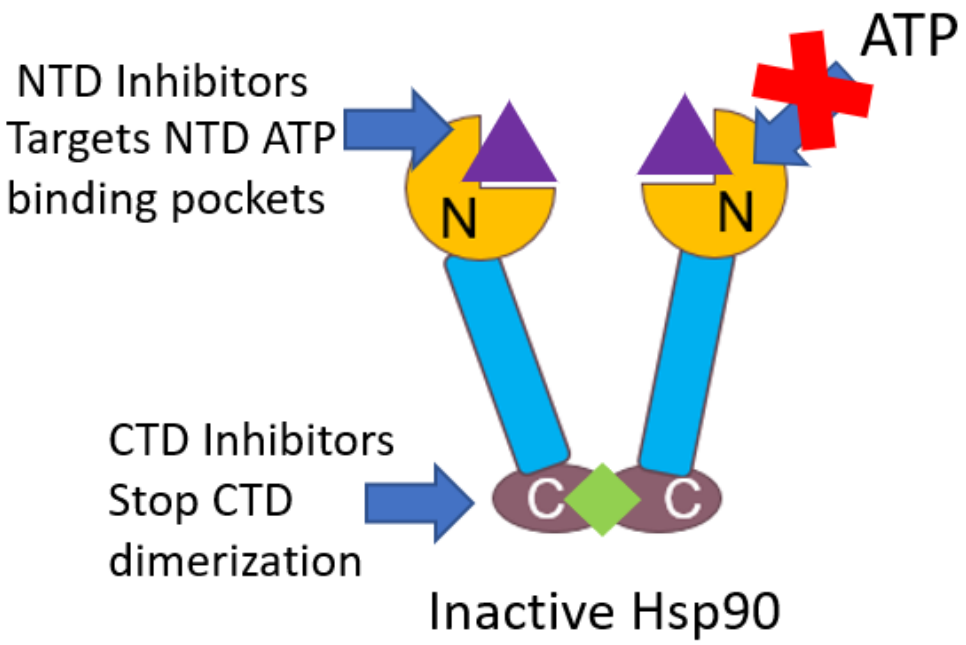


Figure 8: Two types of Hsp90 inhibitors' mode of action. Topside: Hsp90 NTD inhibitors (purple triangle) inhibit Hsp90 ATPase activity via binding to NTD ATP binding pocket. Bottom side: Hsp90 CTD inhibitors (green square) bind to Hsp90 CTD dimerization site thus stop the Hsp90 conformational change machinery.

1.3.1 Hsp90 NTD Inhibitors

In the early 1990s, natural products (Figure. 9, **1**) geldanamycin (GA) and radicicol (Figure. 9, **2**, RD) were found to have Hsp90 binding activity by simulating the structure adopted by ATP in the N-terminal nucleotide-binding pocket of Hsp90, thus leading to selective inhibition of ATP binding and hydrolysis, and in turn, the depletion of oncogenic Hsp90 clients (Mimnaugh, Chavany et al. 1996). Although GA and RD are proved to be highly toxic and

unstable for clinical use, some of their analogues have entered clinical trials (Garg, Khandelwal et al. 2016).

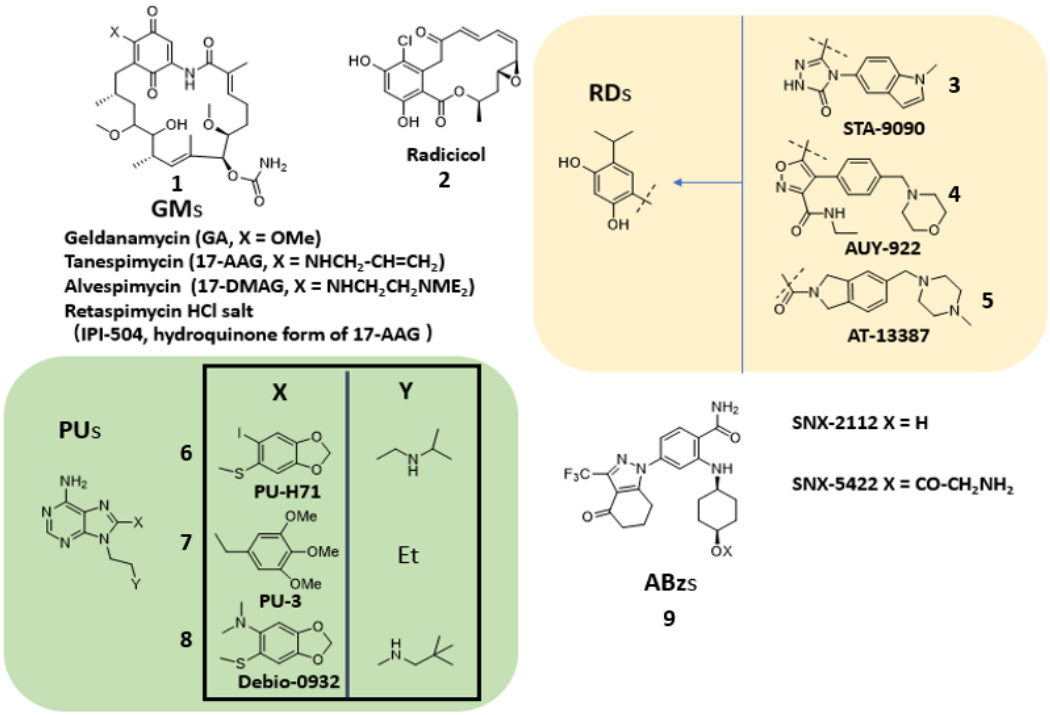


Figure 9**:** Hsp90 N-terminal inhibitors (Wang, 2016).

Tanespimycin (17-AAG, Figure.9 analogue of **1**) was the first Hsp90 inhibitor to have entered clinical trials but failed to reach its primary endpoint at phase III (clinicaltrial.gov) (Hong, Banerji et al. 2013). Despite the lack of drug efficacy, it is reported that benzoquinone sub-group of 17-AAG that undergoes reductive metabolism and detoxification by NADPH: quinone oxidoreductase (NQO1, also called DT-diaphorase) before it acts against Hsp90, resulting in significant hepatotoxicity hepatic toxicity (Hong et al., 2013). Other Hsp90 N-terminal inhibitors have since entered clinical trials. These molecules (**3**, **4**, **5** RDs; **6**, **7**, **8** PUs and **9**, ABzs shown in Figure. 9), plus **1** geldanamycin analogues, all failed to pass phase III clinical trials. While these synthetic molecules did exhibit less toxicity, another significant drawback that emerged, was an induced pro-survival heat shock response (Hall, Seedarala et al. 2016), which corroborated with a previous report that Hsp90 N-terminal inhibition caused overexpression of heat shock factor (HSF1)， a protein shown to

reduce apoptosis and sensitivity of HeLa cells to Tanespimycin (Walton-Diaz, Khan et al. 2013). Furthermore, Hsp90 inhibition leads through a negative feedback loop to activation of the heat shock transcription factor, HSF1, which causes transcriptional induction of Hsp70, Hsp27 and Hsp42, which further protect cancer cells from apoptosis. HSF1 is also believed to limit the activity of Hsp90 inhibitors (Neckers and Workman 2012). In addition, some N-terminal inhibitors precipitated unique side effects, such as, syncope (BIIB021) (Elfiky, Saif et al. 2008), retinal dysfunction (Alvespimycinand AUY922) (Kozloff, Martin et al. 2012), and supraventricular arrhythmias (MPC-3100 and AUY922) (Yu, Samlowski et al. 2010), which increased the difficulties in confirming its association of death rates in larger sample groups, especially when most of which have a high death rate in phase III (clinicaltrial.gov.com). Although it has been 18 years since the first N-terminal inhibitor entered clinical trials, it is still not clear why Hsp90 inhibition has shown unclear clinical activity. Currently, instead of screening new N-terminal lead compounds, companies have focused on developing suitable antibody carriers for N-terminal inhibitors, although none have yet been approved for clinical trial so far.

1.3.2 Hsp90 CTD Inhibitors

1.3.2.1 Novobiocin and its analogues as Hsp90 CTD inhibitor

An alternative approach involves the development of a promising new type of inhibitors that targets the Hsp90 CTD (Terracciano, Foglia et al. 2016). The previous study of Marcu revealed an alternative ATP-binding site on Hsp90 CTD (Marcu et al., 2000). Novobiocin (**0** in Figure. 10), an amino coumarin antibiotic that act as a DNA gyrase inhibitor, but also inhibits Hsp90 by targeting its CTD (Monica, 2000). The aminocoumarin-containing antibiotic novobiocin was found to bind at a second ATP-binding site in the CTD of Hsp90, thus disrupting the interaction of both p23 and Hsp70 co-chaperones with the Hsp90 complex (Marcu et al., 2000). Novobiocin is a weak inhibitor of Hsp90 (IC_{50} = 700 μM in SKBr3 cells) and depletes several Hsp90 client proteins, such as HER-2, v-Src, Raf-1, and mutated p53 (Burlison et al., 2006). Novobiocin binding to the Hsp90 CTD results in destabilization of the chaperone complex, the release of co-chaperones and substrates, and the

subsequent degradation of Hsp90 client proteins (Marcu et al., 2000). Analogues of novobiocin (Figure. 10, Figure 11) have been synthesized (Burlison, Avila et al. 2008) in an attempt to improve its Hsp90 inhibitory activity and establish a sophisticated Structure-Activity Relationship (SAR) of Hsp90 C-terminal inhibition.

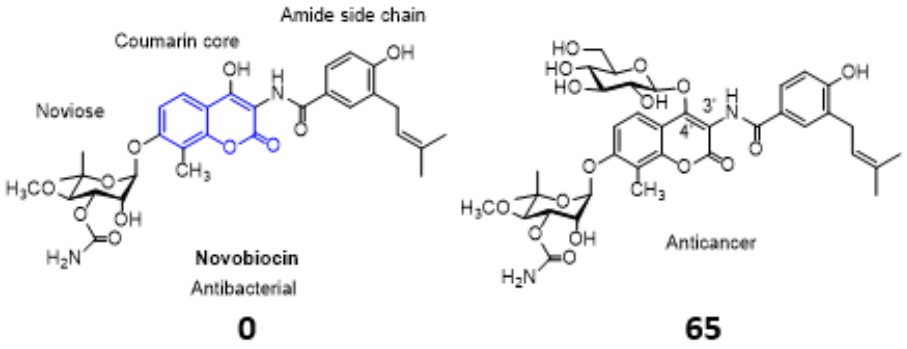


Figure 10**:** Glycosylation of novobiocin leads to increased anticancer activity (Patel, Fuente et al. 2011).

Novobiocin has three distinct parts: (Figure. 10) the amino side chain, the coumarin core and the noviose sugar. Modification on these three parts will result in different analogues which may exhibit improved anti-proliferative activities (Figure. 11 Simon et al., 2017).

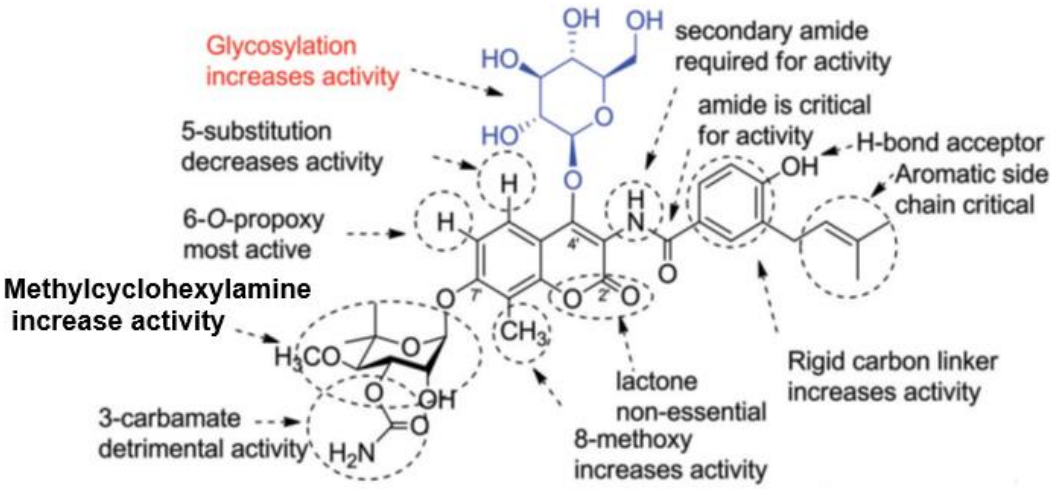


Figure 11**:** SAR of novobiocin analogues based on the previous report (Simon et al., 2017).

SAR Investigation on the amino side chain and noviose sugar

The first novobiocin analogue which was shown to have high inhibitory activity was compound **10** (Figure. 12) which consists of a shortened N-acyl side chain, no 4-hydroxy substituent, and no carbamoyl group on the noviose (Gaiser et al., 2009). It is a selective Hsp90 inhibitor ($SkBr_3$, IC_{50} = 10 µM) with only weak gyrase activity (Gaiser et al., 2009). **10** induced degradation of Hsp90-dependent client proteins at ~70-fold lower concentration than novobiocin (Yu et al., 2005). Interestingly, a later study on this compound and its 8'-methyl substituted form KU-32 (compound **11** in Figure. 12) showed that they mediate neuroprotective activity by triggering survival-supportive heat shock response in normal cells without cytotoxicity at a sub nanomolar concentration. **11** is also reported for an unwanted survival-supportive effects on certain cancer cell lines (e.g. HeLa) through the same mechanism (Farmer, Williams et al. 2012). **11** (also called RTA901 in clinical trials) is currently in phase I clinical trials for the treatment of diabetic neuropathies by Reata Pharmaceuticals, and is therefore the first Hsp90 CTD inhibitor to have entered clinical trials albeit not for its anticancer activity. Most Hsp90 CTD inhibitors do not trigger heat shock response, and hence it is unclear why A4 triggers a heat shock response. This suggests that Hsp90 CTD inhibition may have a more complex mechanism behind it and that it can be applied for the treatment of diseases beyond just cancer.

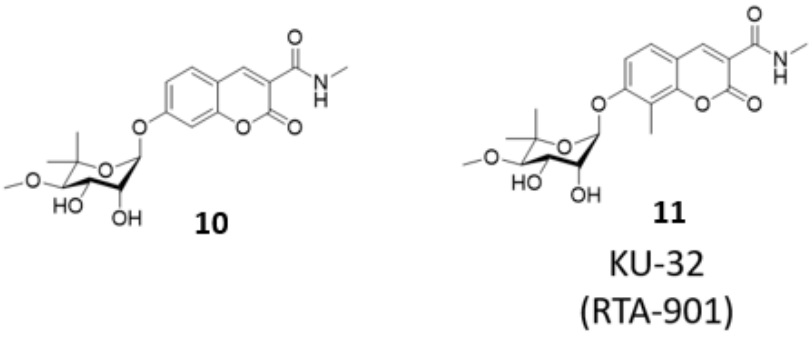


Figure 12: Novobiocin analogue A4 is an Hsp90 CTD inhibitor.

In 2006, Blagg's research group synthesized another series of novobiocin analogues in order to further establish SAR for Hsp90 CTD inhibition (Burlison, Neckers et al. 2006). The results showed that the 4-hydroxy moiety of the coumarin ring and the 3'-carbamate of the noviose appendage were

detrimental to Hsp90 inhibitory activity. In response, 3'-descarbamoyl-4-deshydroxynovobiocin (**12** in Figure 13 part 1, DHN2) was prepared and evaluated against Hsp90. **12** was significantly more potent than the natural product novobiocin (Burlison et al., 2006). Moreover, in order to know whether these modifications are important for DNA gyrase inhibition, these compounds were tested for their ability to inhibit DNA gyrase and were found to exhibit a significant reduction in gyrase activity (Burlison et al., 2006).

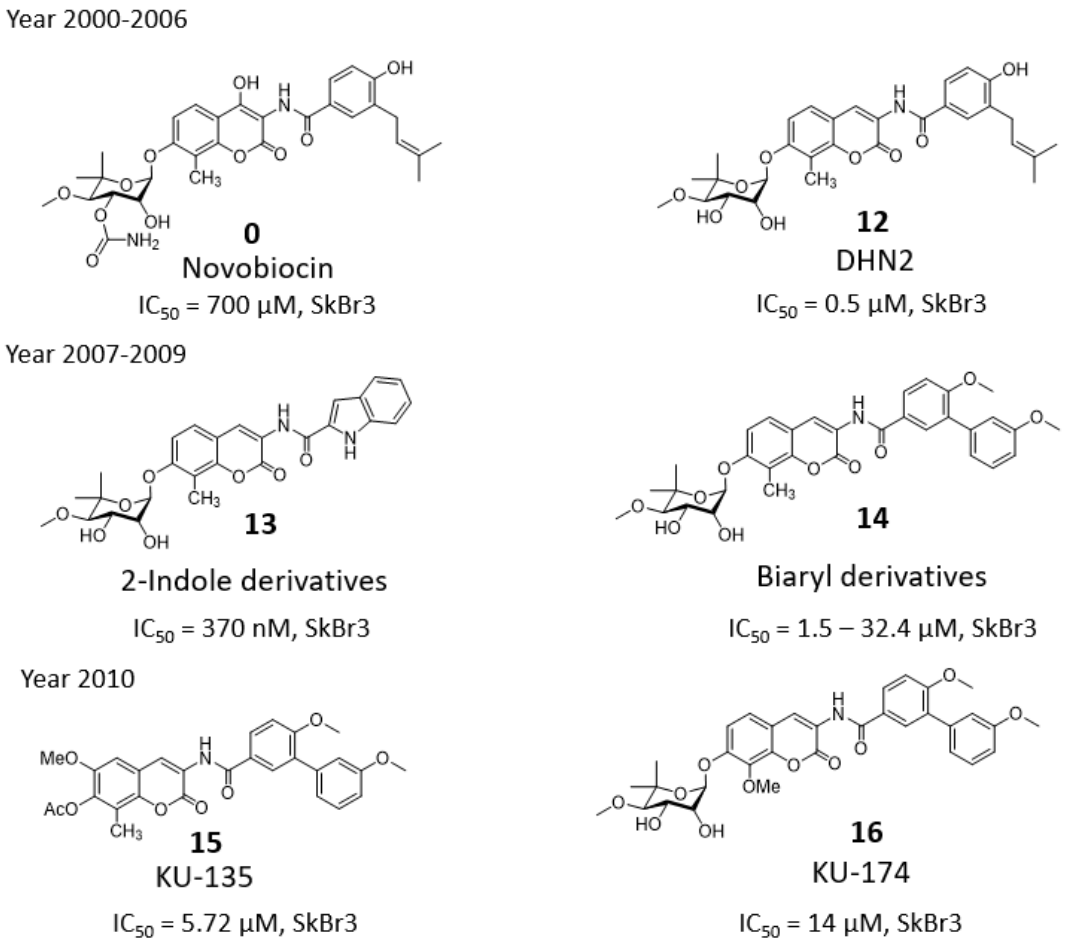


Figure 13: Novobiocin anti-cancer analogues targeting the Hsp90 CTD. (Part 1)

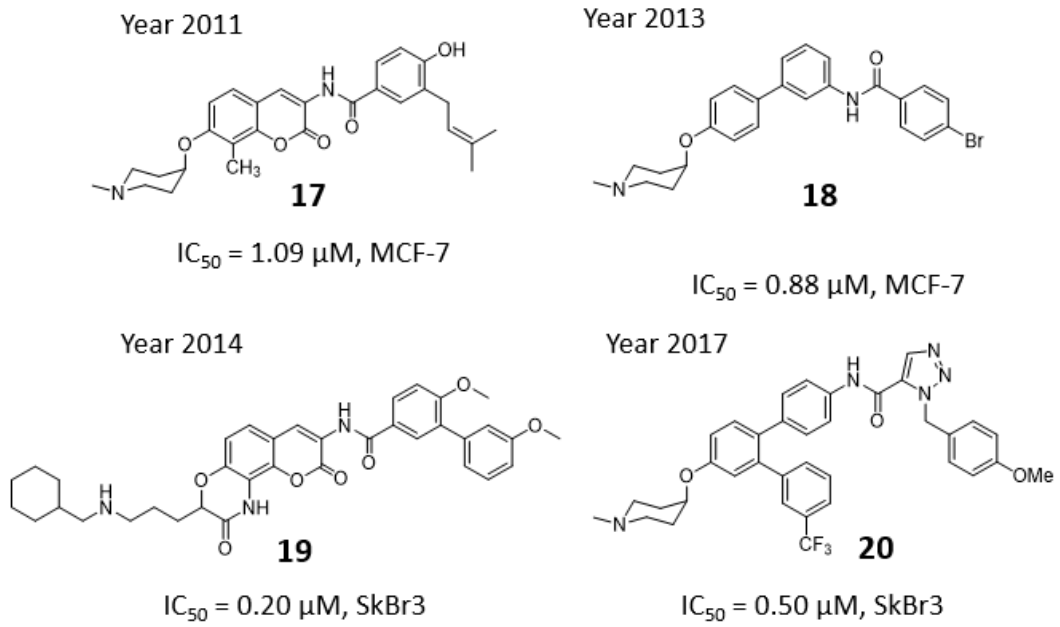


Figure 13: Novobiocin anti-cancer analogues targeting the Hsp90 CTD. (Part 2)

Thus, **12** was confirmed as a selective Hsp90 inhibitor. Blagg's group also synthesized another series of **12** based novobiocin analogues. These analogues have a substituted amide chain side with different functional groups. It was reported that when the amide side chain was changed into 2-indole (**13** in Figure 13 part 1) or biaryl (**14** in Figure 13 part 1), the anticancer activity dramatically increased against several cancer lines (Burlison et al., 2008). Therefore, experimental results obtained from these studies have produced the first series of novobiocin analogues (Figure. 13 part 1) that manifest anti-proliferative activity against several cancer cell lines (Burlison et al., 2008).

Investigation on the noviose sugar moiety was reported later in 2010, and showed that substitution of the sugar with an acetyl group results in an analogue that triggers the pro-survival heat shock response (Zhao, Garg et al. 2015). The Hsp90 CTD modulators (e.g. **15** KU-135 in Figure 13 part 1) that induce Hsp70 levels were shown to manifest a useful neuroprotective activity similar to **10** (Zhao et al., 2015). After another modelling-based design and systematic synthesis, the optimal substituent for 7-noviose sugar is 1-methylpiperidine (**17** in Figure 13 part 2) (Zhao, Donnelly et al. 2011) and

cyclohexyl amino – lactam (**19** in Figure 13 part 2) (Kusuma, Khandelwal et al. 2014). According to a computer-based ligand docking, the coumarin core of tested compounds made fewer contributions to domain binding, which only serves as hydrophobic backbone. In response, the coumarin core was replaced with a dibenzyl core to gain better molecular flexibility (**18** in Figure.13 part 2) (Kusuma et al., 2014). Most recently, this has been further substituted with a tribenzyl core substituent (**20** in Figure. 13 part 2) and its binding pocket was proved to be identical to novobiocin analogues, making it a new CTD inhibitor scaffold (Davis, Zhang et al. 2017).

Although some of those novobiocin analogues have shown improved antiproliferative activity, none of them have entered clinical trials for the treatment of cancer due to lack of drug-protein co-crystal structures. No research groups so far have reported any reliable drug-protein binding mechanisms.

SAR Investigation on 4'-hydroxyl:

While Blagg's group concluded that removal of the 4'-hydroxy moiety results in optimal antiproliferative activity, Jack-Michel Renoir reported 4TCNA (**21** in Figure. 14) that exhibits superior antiproliferative activity compared with novobiocin in 2007, his work showed that removal of the noviose moiety in novobiocin together with the introduction of a tosyl substituent at C'-4- on coumarin cores afford **21** as lead structures (Le Bras, Radanyi et al. 2007). His follow up work (Le Bras et al., 2007) proposed a 4-OTS, 7-acetylprotected-glucose novobiocin analogue (**22** in Figure 14) as the optimal structure, indicating the importance of the dual 4’ and 7’ substitutions on the novobiocin coumarin core. Since then, our research group in 2011 reported the 4'-glucosyl/galactosyl-novobiocin (**66** in Figure 14), which exhibited 27000+ fold increased anticancer selectivity with respect to novobiocin (Patel et al., 2011). Moreover, this work also proved that glycosylation on 4'-hydroxy of novobiocin would not only increase its anticancer activity by 100-fold but also greatly enhanced the anticancer selectivity by removing its unwanted antibacterial activity (Patel, 2011). Full understanding of the binding mechanism of

novobiocin C-terminal inhibitors remains elusive and research continues to this day.

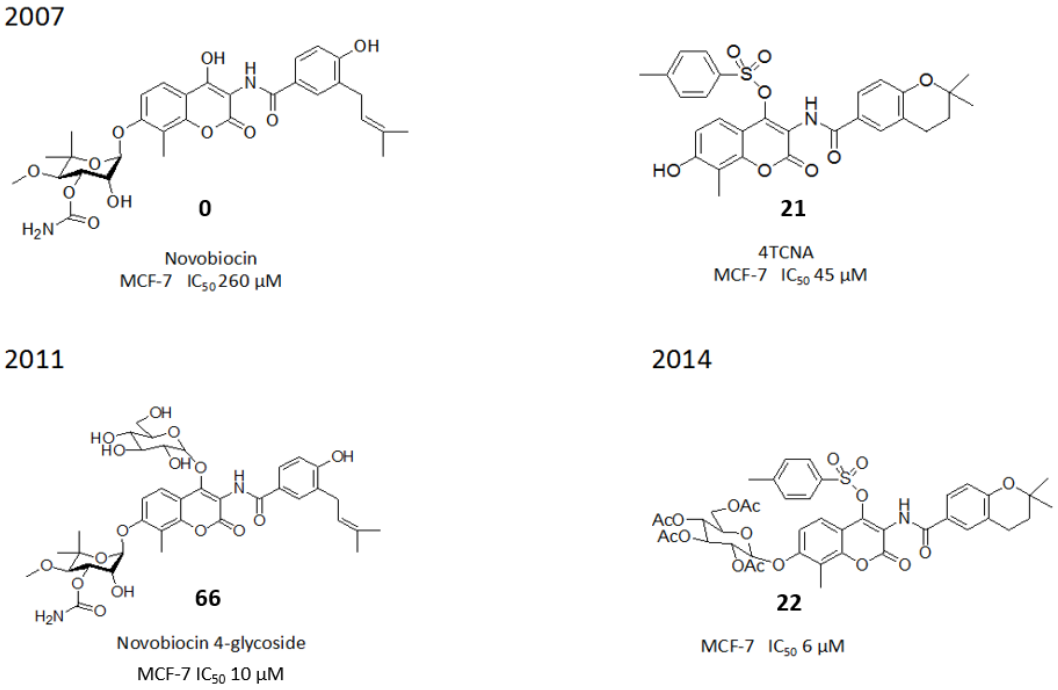


Figure 14**:** Evaluation of 4'-substituted novobiocin as Hsp90 CTD inhibitor. The IC_{50} value was measured after 72 hours.

Summary of SAR of Novobiocin:

Noviose part: Replacement of 3-carbamoyl into hydroxyl group would increase its activity. Substitution with 1-methylpiperzine manifests better activity (Zhao, Garg et al. 2015).

Coumarin part: The presence of 8-methyl increases activity, the presence of 6-methoxy increase solubility in water and activity. Although the presence of 4-hydroxy decreases the activity, glycosylation of 4-hydroxy resulted in an enhancement of activity. Moreover, lactones are not vital to Hsp90 inhibition (Patel et al., 2011).

Amide side chain part: The amide bond of secondary amine is required for anticancer activity. The aromatic side chain is important for activity. In addition,

the replacement of benzylamide into biaryl groups increases the anticancer activity as it serves as an important backbone that provides critical interactions with hydrophobic binding pocket (Donnelly and Blagg 2008).

1.3.2.2 Epigallocatechin-3-Gallate (EGCG)

EGCG (**23** in Figure 15) is a polyphenolic compound separated from green tea, and it is the most abundant catechin found in green tea. EGCG is well studied and reported for its anti-cancer, antioxidant, and antimicrobial activities (Matsumoto, Horie et al. 2004). Early studies showed that EGCG inhibits the activity of various kinds of proteins at a fixed concentration of 70 μM, including telomerase, the aryl hydrocarbon receptor (AhR), several kinases, and other transcription factors. All these proteins are well known Hsp90 client proteins (Yin, Henry et al. 2009). In 2005, Palermo and co-workers discovered that EGCG exhibits inhibitory activity against AhR partly via Hsp90 inhibition, using affinity purification experiments (Palermo, Westlake et al. 2005). Follow-up studies by Yin and co-workers reported that EGCG binds near the C-terminal dimerization site (residues 538–738) of Hsp90. However, unlike novobiocin and its analogues which will block dimerization of Hsp90, EGCG stabilizes the association of cochaperones Hsp70, Cyp40, and XAP-2 to Hsp90 (Yin, Henry et al. 2009). Moreover, it was found that EGCG induces concentration-dependent degradation of the Hsp90-dependent oncoproteins ErbB2, Raf-1, and pAkt with a slight increase of Hsp70 levels.

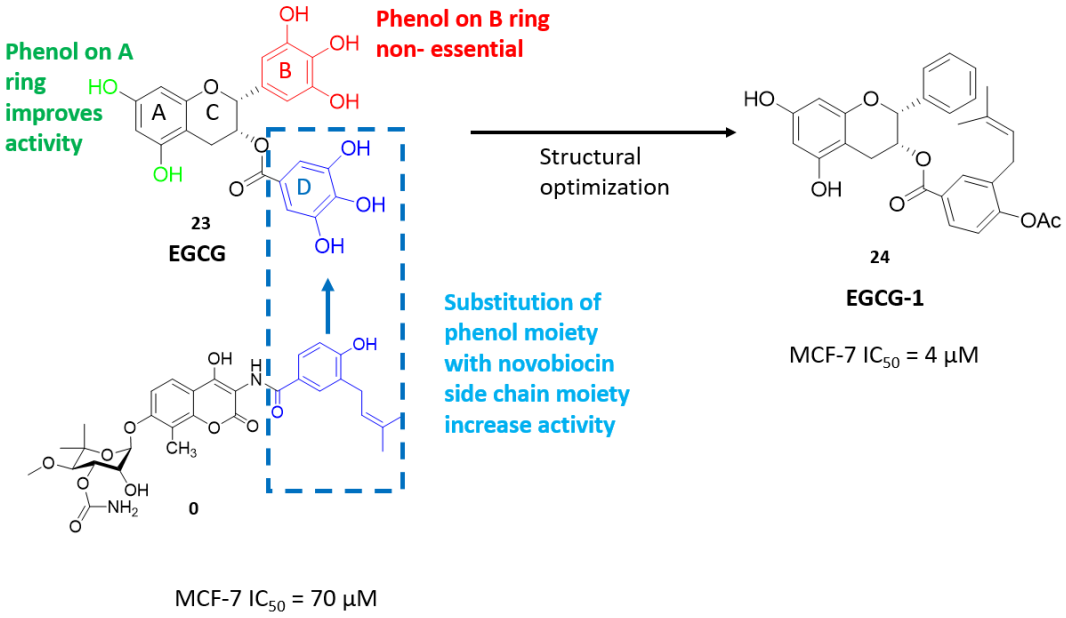


Figure 15: SAR of EGCG.

Since EGCG is confirmed to be a Hsp90 CTD inhibitor, Khandelwal and co-workers published the first structure-activity relationships studies on EGCG (Figure. 15), evaluated their cytotoxicity, and reduction of several Hsp90 client proteins' expression as drug potency validation (Khandelwal, Hall et al. 2013). In these studies, they observed that the phenols on the B- and the D-rings are detrimental to inhibitory activity, while syn-stereochemistry of the linker that connects the B- and D-rings with the benzopyran core is beneficial (Figure 15). The addition of benzamide present in the novobiocin side chain was shown to be a promising replacement for the gallic acid moiety of the EGCG D ring. This resulted in a 15-fold improvement in antiproliferative activity. This result led to the development of lead compound **24** (in Figure 15, EGCG-1 MCF-7, IC_{50} = 4 μM). Further studies by Bhat and co-workers confirmed that B- and D-ring phenols are non-essential to antiproliferative activity and established the ester linker connecting the C- and D-rings could be replaced with an amide (**25** in Figure 16) or sulphonamide (**26** in Figure 16) without compromising anticancer activity (Bhat, Adam et al. 2014). Nevertheless, the mechanism of how these modifications affect its interactions with Hsp90 remained to be researched.

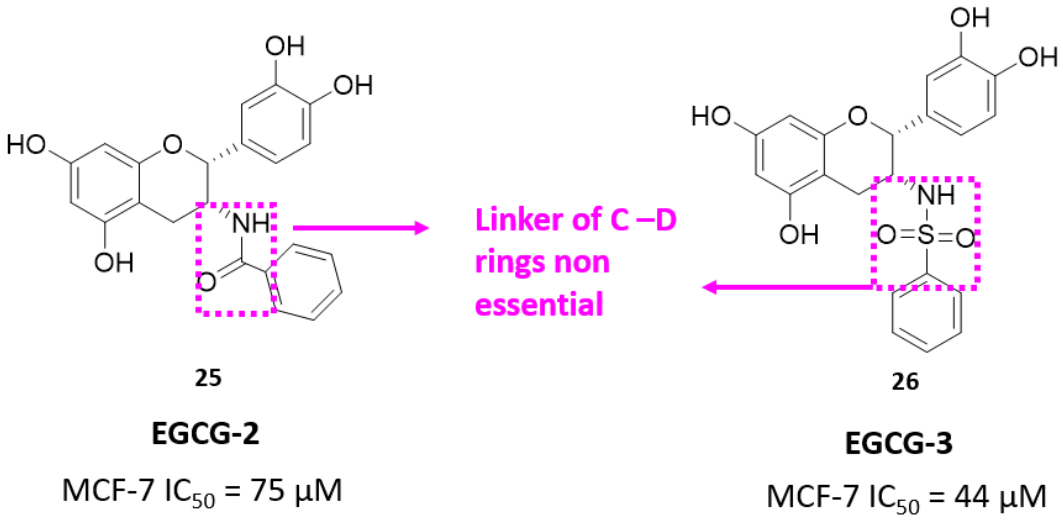


Figure 16: Structure of **25** EGCG-2 and **26** EGCG-3. Linker of EGCG C to D ring does not affect much about the antiproliferative activity.

1.3.2.3 Vibsanin B (ViB) and its analogues

Vibsanin B (**27** in Figure 17, ViB) is a natural terpene biosynthesis which isolated from *Viburnum odoratissimum* Ker-Gawl (Kawazu 1980). In 2015, Chen and co-workers observed that ViB has high selectivity for binding HSP90β rather than HSP90α, inhibiting leukocyte chemotactic migration in human monocytic cell lines as well as in a zebrafish model, and improved experimental autoimmune encephalomyelitis (EAE) in mice (Ye, Deng et al. 2015). Follow-up research by the same group discovered the Hsp90 CTD binding of ViB using a biotin pull-down assay together with an *in vitro* Hsp90 CTD binding assay, affirming ViB as a novel Hsp90 CTD inhibitor (Shao, Su et al. 2017) or it can also be called an Hsp90β CTD selective inhibitor. Furthermore, through an extensive SAR analysis of ViB (Figure. 17), they highlighted C-18-OH to be the most available position for substitution on ViB. Substitution of C-18-OH with tertiary amines, fluorine, and chlorine shows the most promise, with such modifications resulting in the antiproliferative activity increasing approximately 5 to 10-fold. While ViB C-18-OH substitution with azide, triazole, and esters generally maintained the compounds' anticancer activity, substitution with aldehyde and carboxylic acid significantly reduced activity. Besides C-18, C7 methylation reduces activity; however, removal of C7-OH or acetylation has no effect on activity. C4-C6 Michael moiety is critical for antiproliferative activity, with either substitution with C4-carbonyl or hydrogenated C4-C6 resulting in loss of antiproliferative activity. Interestingly, the location of the Michael acceptor moiety does not have to be on C4-C6. An analogue that possessed a C4-OH and C18-CHO, which creates a new Michael acceptor moiety on C2-C18, showed almost the same activity compared with original ViB (Shao, Su et al. 2017). Although the author did not mention, it is highly possible that ViB C4-C6 acts as a covalent modifier towards Hsp90 CTD. Moreover, C8 substitution of small ester groups is favoured; nevertheless, the introduction of either group showed a detrimental effect upon the antiproliferative activity (Shao, Su et al. 2017).

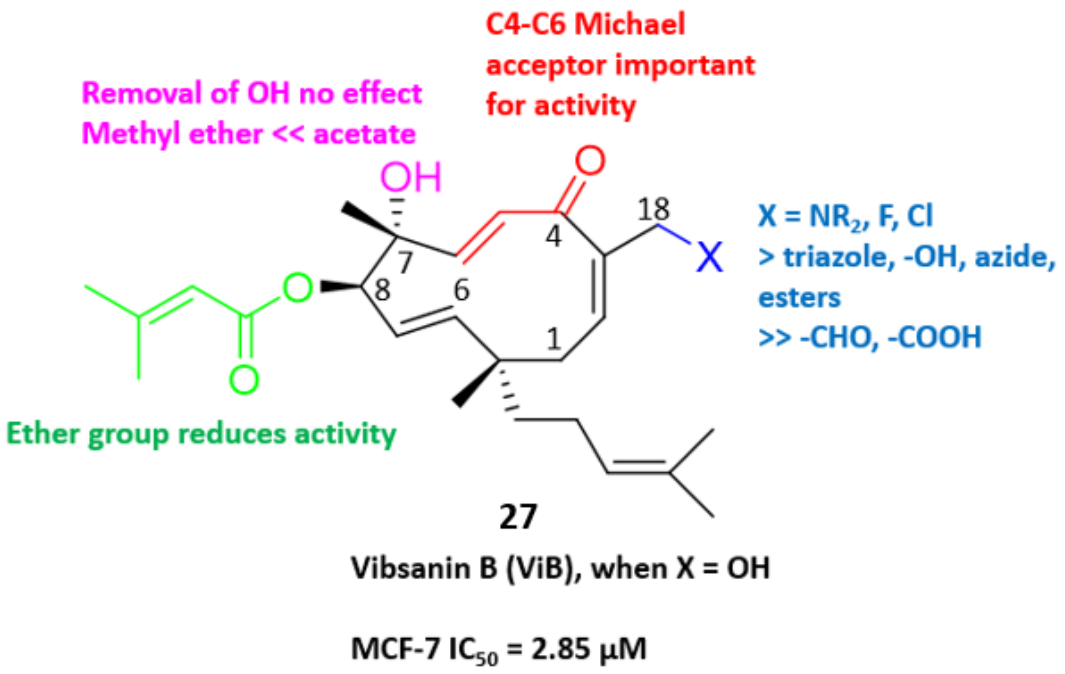


Figure 17: SAR of Hsp90β CTD inhibitor ViB **27**.

1.3.2.4 Cis-platin

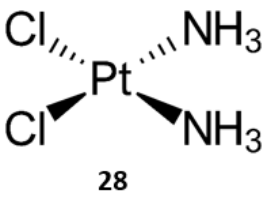


Figure 18: Structure of Cis-platin **28**

Cisplatin (28 in Figure. 18) is a platinum-containing coordination complex that is widely used for the chemotherapy of cervical, bladder, testicular, ovarian, and other solid tumours (Galanski 2006). The mechanism of anticancer activity of cisplatin is reported to be its ability to form intra-strand DNA adducts which are particularly lethal in several cancer cell types (Jordan and Carmo-Fonseca 2000). It has also been shown that due to its chemical reactivity, cisplatin can interact with various proteins, phospholipids, and RNA (Sreedhar, Söti et al. 2004). In 1999, Itoh and co-workers reported that cisplatin could inhibit Hsp90

chaperone activity (ITOH, OGURA et al. 1999). Affinity purification and protein fingerprinting experiments were used to confirm that cisplatin binds to the Hsp90 C-terminal domain. Subsequently, Söti and co-workers reported that cisplatin is an Hsp90 C-terminal inhibitor that binds near the previously identified C-terminal nucleotide-binding site (Söti, Rácz et al. 2002).

Further studies by Rosenhagen and collaborators revealed that the administration of cisplatin to neuroblastoma cells resulted in the degradation of steroid hormone receptors (androgen and glucocorticoid receptors), but no other Hsp90-dependent clients, such as Raf-1, lck, and c-rac (Rosenhagen, Sõti et al. 2003). Moreover, a heat-shock factor (HSF)-dependent luciferase reporter assay was used to showed that cisplatin does not induce the heat-shock response. These results suggest that cisplatin selectively inhibits some Hsp90 functions, does not trigger heat-shock response, and that its mode of action may be similar to novobiocin.

1.3.2.5 Silybin

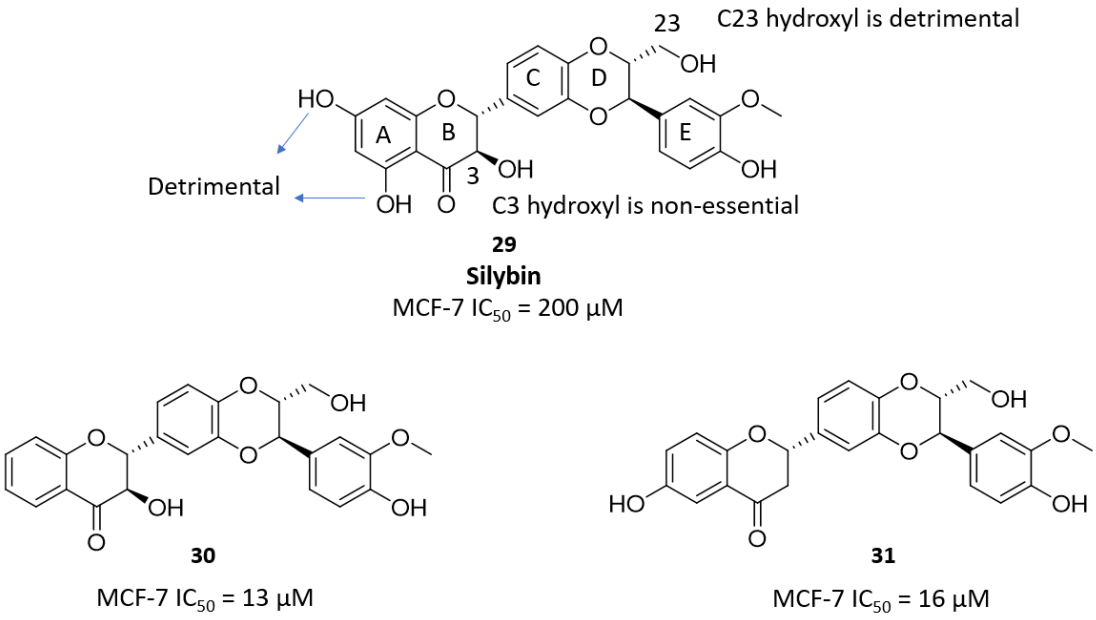


Figure 19: SAR of Silybin **29** as Hsp90 CTD inhibitor.

Silybin (**29** in Figure.19) is the major component of the flavonolignan extract isolated from the seed of milk thistle plants (*Silybum marianum*) and has displayed hepato-protective effects and antiproliferative activity against

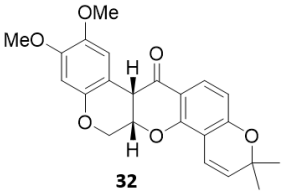
several cancer cell lines (Gazak, Walterova et al. 2007). It was reported by early studies that silybin induced cell cycle arrest and caused the depletion of CDK2, CDK4, cyclin E, and cyclin D1 proteins in colon cancer cells (Agarwal, Singh et al. 2003). Due to CDK2 and CDK4 being well-characterized Hsp90-dependent clients, Hsp90 could prove to be the primary target of silybin. In an attempt to establish whether silybin binds to Hsp90, Zhao and co-workers established a luciferase-refolding assay with silybin. They demonstrated that silybin inhibited the renaturation of heat-denatured luciferase, suggesting that Hsp90 could be a biochemical target for silybin (Zhao, Brandt et al. 2011). Subsequent experiments demonstrated that silybin induced a concentration-dependent degradation of the Hsp90-dependent client proteins Her2, Raf-1, and Akt, without affecting Hsp70 or Hsp90 levels (Zhao, Brandt et al. 2011). SAR study showed that the C-3 and C-23 hydroxyl groups were not required for activity; however, at least one substitution (preferably 4-hydroxyl) on the E-ring was important for activity (Zhao, Brandt et al. 2011). Furthermore, SAR studies suggested that the A-ring phenol was not required, and its removal led to the development of **30** and **31** (Figure.19), which manifest IC_{50} values of 13 and 16 μM against MCF-7 cell line, respectively (approx. 14-fold more potent than parent structure silybin IC_{50} = 200 μM). Another study showed that silybin binds to the C-terminal domain of Hsp90 and releases mature glucocorticoid receptors from the Hsp90 complex as demonstrated by NMR analysis (Riebold, Kozany et al. 2015).

1.3.2.6 Deguelin

Deguelin (**32** in Figure. 20) is a rotenoid isolated from Derris trifoliata or Mundulea sericea (Pellati and Rastelli 2016). It has been reported for its promising antiproliferative, antimetastatic, and apoptotic activity against several cancers both in vitro and in vivo (Garg, Khandelwal et al. 2016). This compound has been found to bind to the alternative ATP-binding pocket of Hsp90 CTD and disrupts Hsp90 function, causing ubiquitin-mediated degradation of hypoxia-inducible factor (HIF-1α) (Oh, Woo et al. 2007). Deguelin proved successful in treating mice with solid tumours, in which

tumour growth was significantly reduced by inducing apoptosis and reducing the expression of Hsp90 dependent client proteins, without any detectable side effects or unwanted toxicity (Oh, Woo et al. 2007). However, it was reported that deguelin in high doses caused Parkinson's disease-like syndrome (Caboni, Sherer et al. 2004). Other problems have limited its clinical application, for example, while deguelin and other rotenoids share some of the pharmacological hallmarks of Hsp90 inhibition, treatment of SkBr3 cells with different rotenoids did not significantly affect the level of Hsp90-dependent clients, such as ErbB2, known to be highly sensitive to Hsp90 inhibition (Garcia, Barluenga et al. 2010). Moreover, deguelin failed to induce the degradation of IP6K2, an Hsp90 client sensitive to Hsp90 CTD inhibition (Garcia, Barluenga et al. 2010). All these defects lead to the development of better analogues. Chang and co-workers reported SAR studies for deguelin validated by HIF-1α reduction and antiproliferative experiments(Chang, An et al. 2012). Their studies demonstrated that the 2,2-dimethyl-2H-chromene moiety and both methoxy groups at the C9 and C10 positions of deguelin are critical for biological activities. This SAR data led to the development of compounds **33** and **34** (Figure.20), which have cytotoxicity IC_{50} values of 0.14 and 0.49 μM against H1299 cell line, respectively. Notably, these two deguelin analogues exhibited anti-angiogenic activities in zebrafish embryos in a dose-dependent manner (Chang, An et al. 2012). **33** was shown to possess broad antitumor and anti-angiogenic activities in various non-small cell lung cancer (NSCLC) cell lines. Additionally, **33** displayed significantly reduced toxic effects on various normal cells and showed no apparent Parkinson's disease-like toxicity in the rat brain compared with deguelin (Chang, An et al. 2012). Furthermore, both two compounds inhibit Hsp90 CTD without inducing expression of Hsp70. Another promising deguelin analogue called L80 (**35** in Figure. 20) was developed by Lee and co-workers. **35** has significant antitumor and anti-angiogenic activities both in vitro and in vivo against multiple lines of NSCLC cells. **35** also has minimal toxicities to normal lung epithelial cells, hippocampal cells, retinal pigment epithelial cells, and vascular endothelial cells, whereas deguelin induced significant cytotoxicity in these cells. **35** was found to disrupt Hsp90 function by directly interacting with the C-terminal domain of Hsp90 and destabilizing multiple Hsp90 client proteins, such as HIF-1α (Lee, Min et al.

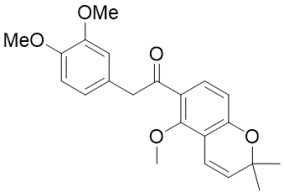
2015). **35** has no IC_{50} reported due to the author testing the compound at fixed concentrations rather than gradient concentration.



32

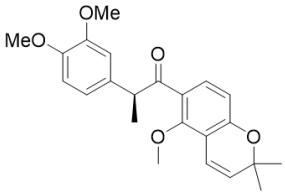
Deguelin

H1229 IC_{50} = 0.11 µM



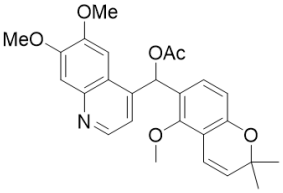
33

H1229 IC_{50} = 0.14 µM



34

H1229 IC_{50} = 0.49 µM



35

L80

Figure 20: Structure of deguelin **32** and its analogues **33**, **34** and **35**.

.

Chapter 2: Synthesis of novobiocin based glycosides as anticancer agent targeting Hsp90 C-terminal

2.1 Project Objectives

As a result of the structural diversity associated with the core structure of novobiocin and its anticancer activity against Hsp90, the research in this thesis focuses on an investigation as to whether its potency and binding can be improved via modification at three points around the coumarin core. Based on previous studies and reports in the literature, a series of novobiocin-based analogues were proposed in this project which focused on three key modifications to the core as shown below (Figure. 21).

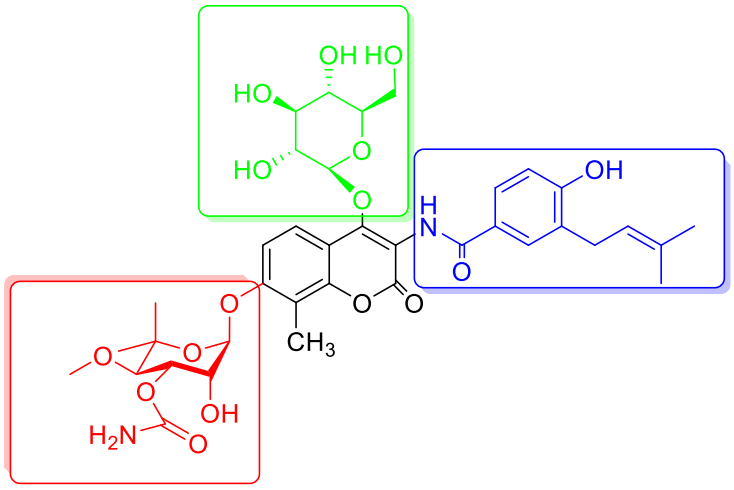


Figure 21: Building blocks of novobiocin. Blue = amide side chain, Red = novoise sugar, Green = 4'-substitutents.

These proposed changes to the core structure were driven by reports on the structure/activity relationship on the modification of novobiocin. Donelly and Blagg had shown that replacement of the amide side chain with biaryl derivatives increased the anticancer activity by up to 467-fold (Donnelly and Blagg, 2008). In addition, Burlison showed that removal of the 3''-carbamate increased activity (15-fold) together with the unmasking of the 4'-OH (93-fold) in the coumarin, which combined resulted in up to a 1400-fold increase in activity (Burlison et al., 2008). Moreover, a single glycosylation on the 4'-OH position, increased its anticancer activity by up to 100-fold, and this substitution also decreased the unwanted anti-bacterial activity by 267-fold (Patel et al., 2011). As a result, the three proposed structural variations consist

of modifications on the amino substitutions (with biaryl and 2-indoleacetic acid). Another planned R1 substitution involves the use of triazole analogues following reports from Burlison (Burlison et al., 2008).Different R1 and R2 substituted compounds will be synthesized using the same synthetic method developed. and 4'-OH glycosylation combined with the 3'-OH modification with various sugars as shown below (Figure. 21).

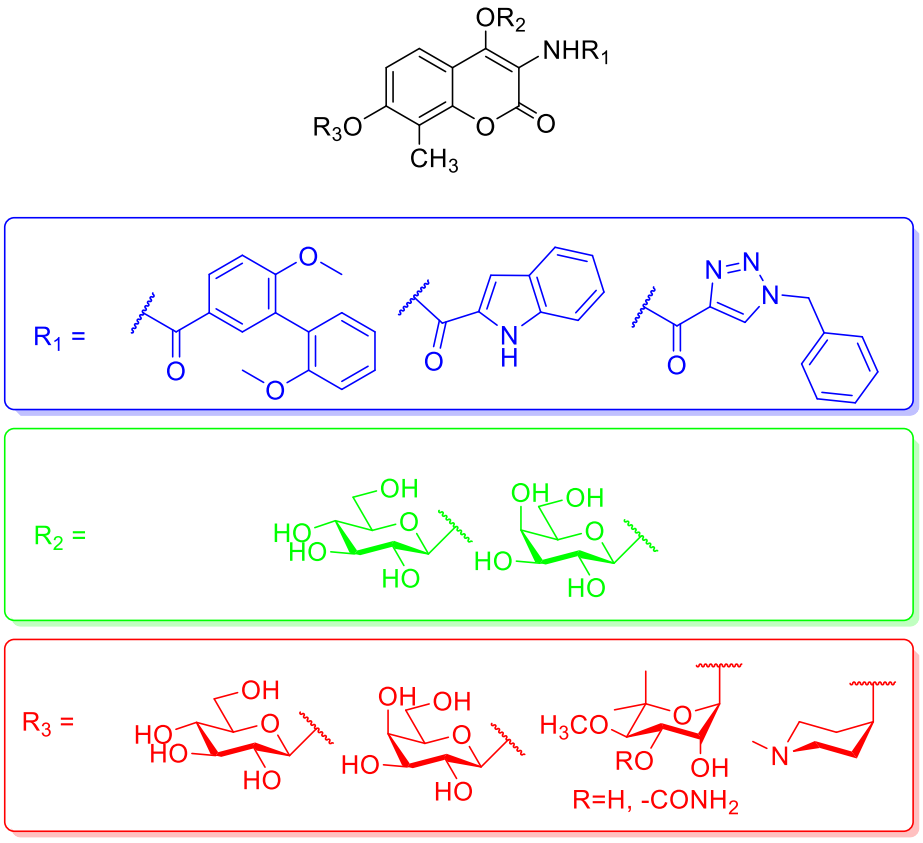
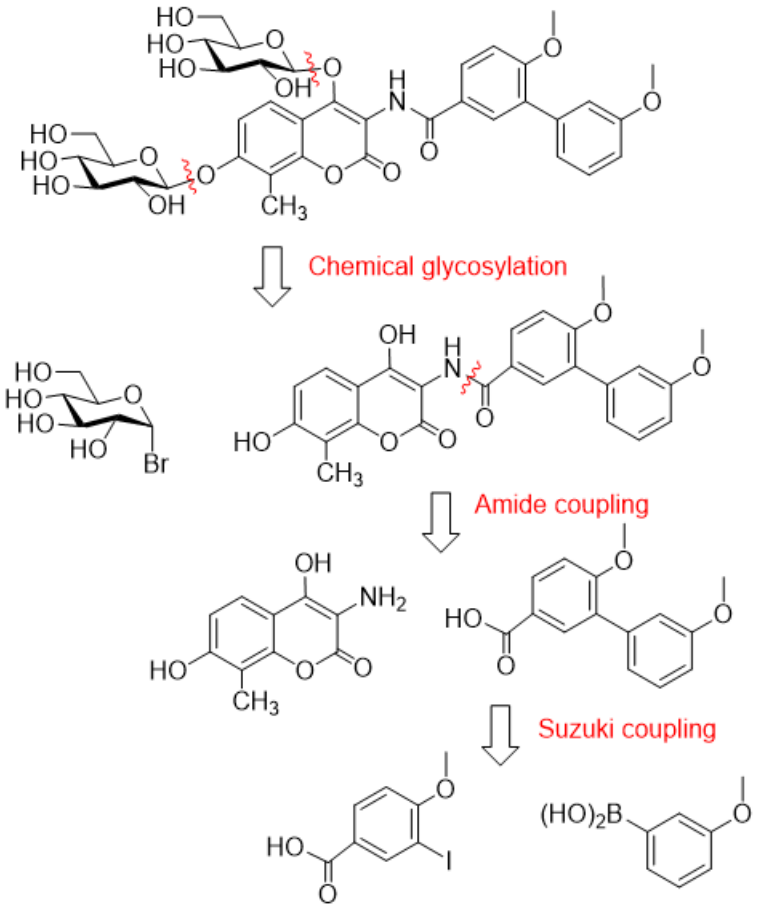


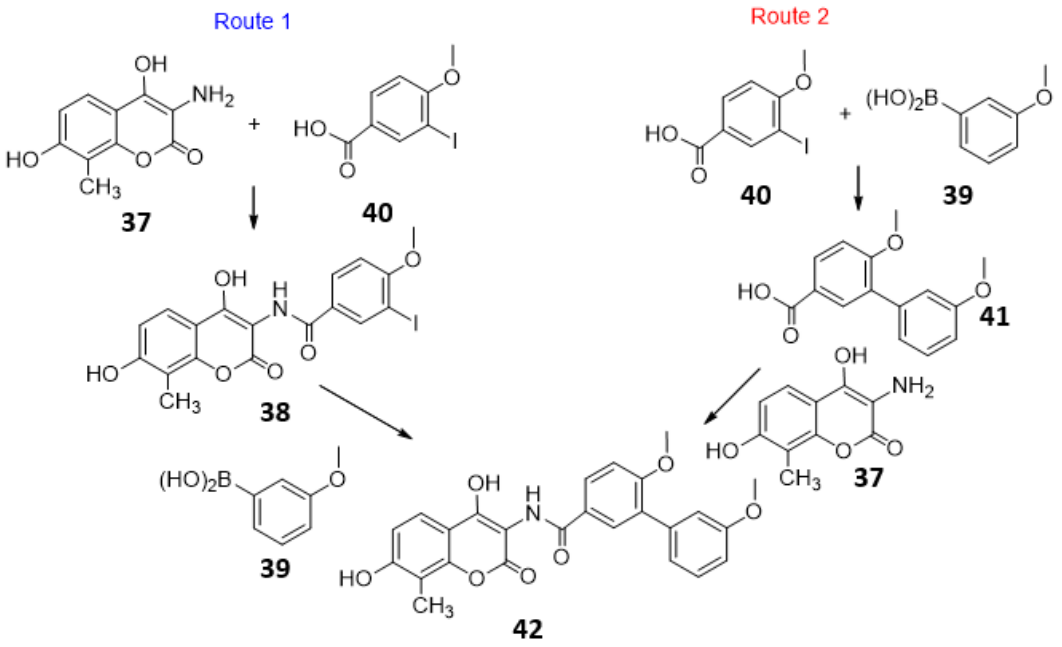
Figure 22: Proposed compounds to be synthesized in this project.

In order to develop a range of analogues, a straightforward retrosynthesis of the target compound was proposed in Scheme 1, as shown below. Initial disconnection of the two sugar moieties leads back to the dihydroxy coumarin core. Further unmasking of the amino group led back to compound **10** as the primary core structure. It was envisaged that the synthesis of the biaryl side chain could be achieved by Suzuki coupling and a challenging diglycosylation step on both the 4' and 7' -OH of the coumarin core would provide ready access to the desired target analogues (Scheme 1).



Scheme 1: Retrosynthesis of the target compound

While the last step involves glycosylation, there are two potential routes involved in the assembly of the biaryl unit (Scheme 2). As a result, there were two routes designed as potential means to access this structure. Route 1 was designed to allow for formation of the amide bond followed by subsequent Suzuki coupling to afford the biaryl amide; Route 2 involves initial Suzuki coupling followed by amide coupling with the coumarin core to afford the same final product. Both synthetic routes were designed to be explored in this thesis and evaluated for their ease of synthesis.



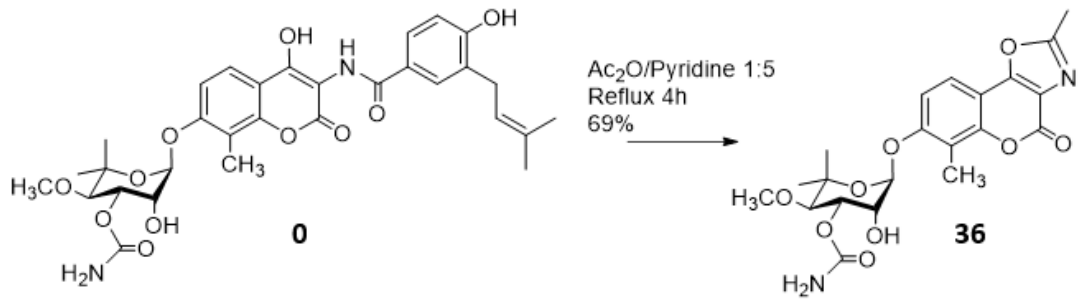
Scheme 2**:** Proposed two synthetic routes.

The most challenging part is the glycosylation step. According to previous work, it was indicated that 4'-OH on the coumarin core is not very chemically active, which makes the traditional glycosylation method using silver salts as catalyst unavailable. Therefore, extra method development of glycosylation using novobiocin as starting material was carried out to gain more knowledge in order to support the glycosylation of the biaryl amide.

2.2 Results and Discussion

2.2.1 Synthesis of 3-amino-4,7-dihydroxy-8-methylcoumarin (coumarin core 37)

In order to develop analogues of novobiocin, it was first necessary to gain access to the coumarin core. Novobiocin is a commercial compound and can be obtained in sufficient quantities and as such, provides ready access via removal of the extraneous parts to the highly substituted coumarin core **37** (Scheme 5). The synthetic route for the preparation of the coumarin core was modified from the original method as previously described by Hinman (Hinman, Caron et al. 1957). Firstly, novobiocin was reacted with acetic anhydride to remove the amide group giving compound **36** in 69% yield (Scheme 3).



Scheme 3**:** Synthetic route of **36** (Hinman, Caron et al. 1957).

Although Hinman did not proposed any mechanisms, another work (Berkov-Zrihen, Rutenberg et al. 2012) systemically studied the acetylation of novobiocin and concluded that the cleavage of the amide bond is somewhat unusual and involved in 3 distinct steps:

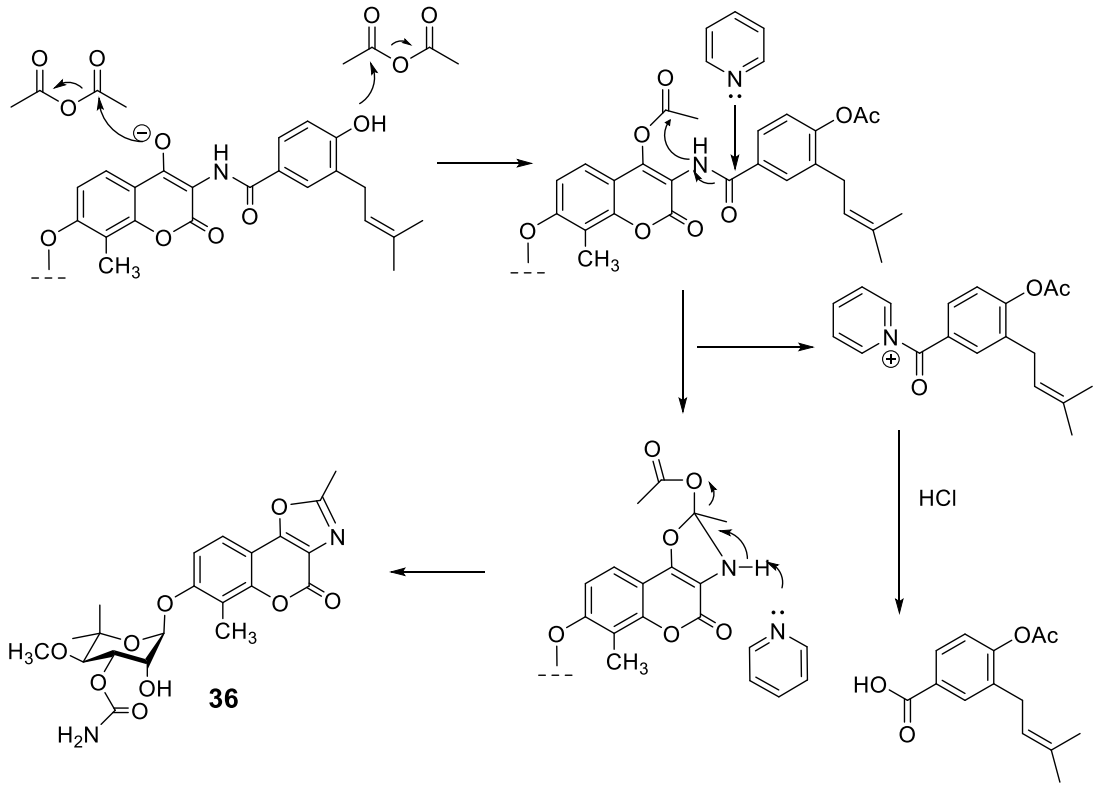
1) acetylation on 4’ -OH and phenol -OH resulted in di-acetylated product

2) pyridine assisted cleavage of amide bond

3) cyclization of 2-methyl- oxazole.

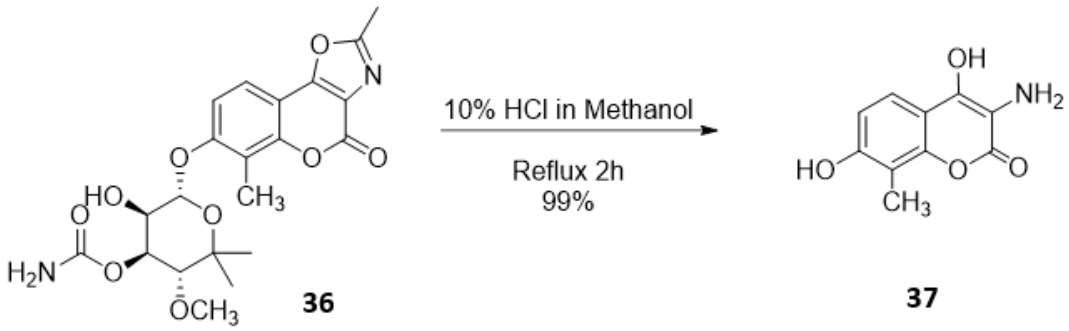
According to the literature (Berkov-Zrihen, Rutenberg et al. 2012), ESI-MS in-process data showed that phenol -OH must be acetylated for amide cleavage to happen, due to electro-withdrawing effect of acetyl group enhanced the

electrophilicity of amide carbonyl. The cleavage of the amide bond is believed to proceed via a nucleophilic catalysis of pyridine. Using non-nucleophilic base such as di-isopropyl ethylamine would result in no amide cleavage but only acetylation (Berkov-Zrihen, Rutenberg et al. 2012). 2'-amine will then react with 4'-acetyl to complete the cyclization and resulted in **36** (Scheme 4).



Scheme 4: Mechanism of novobiocin reaction with acetic anhydride in pyridine.

Compound **37** (the coumarin core) was finally formed by reacting compound **36** with a solution of HCl in methanol giving the free OH groups in quantitative yield (Scheme 5).

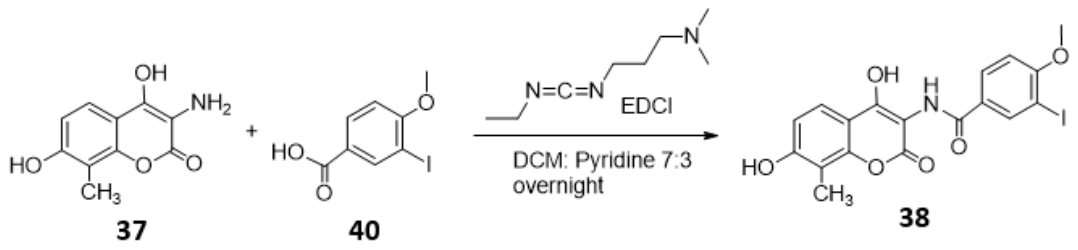


Scheme 5: Synthesis of novobiocin coumarin core **37**

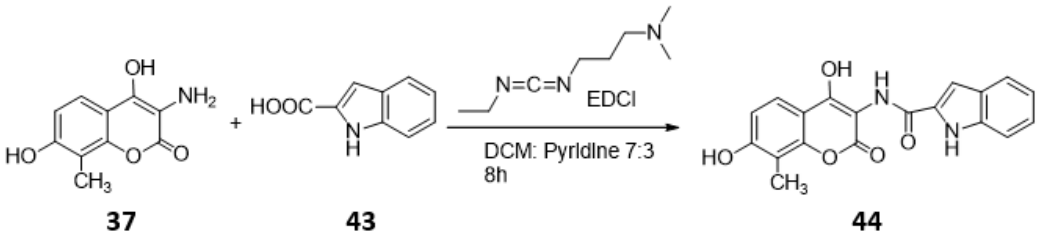
Once compound **37** was successfully obtained, it could now be used in subsequent coupling reactions to synthesise the desired analogues of novobiocin. One of the key challenges relating to its use was its poor solubility as it was poorly soluble in common organic solvents such as DCM, ethyl acetate or chloroform. However, it was soluble in pyridine and it proved to be very unstable on exposure to air, necessitating careful storage in a sealed and dry brown bottle in the freezer.

2.2.2 Route 1: Synthesis of N-(4,7-dihydroxy-8-methyl-2-oxo-2H-chromen -3-yl)-3-iodo -4-methoxy-benzamide 38

With the requisite obtention of the core compound **37**, attention turned towards its functionalization. The planned approach was for the formation of the amide bond to introduce the biaryl unit which involved a carbodiimide-mediated amidation that was catalyzed by *N*-(3-dimethylaminopropyl)-*N*'-ethylcarbodiimide hydrochloride (EDCI). The original conditions (in Scheme 7) used to complete the amide coupling of compound **37** to generate a range of analogues followed that of previous groups as shown below (Blagg 2008). The two key molecules that we envisaged synthesizing in this manner were the benzoyl unit and the indole moiety (Scheme 6 and 7) respectively.

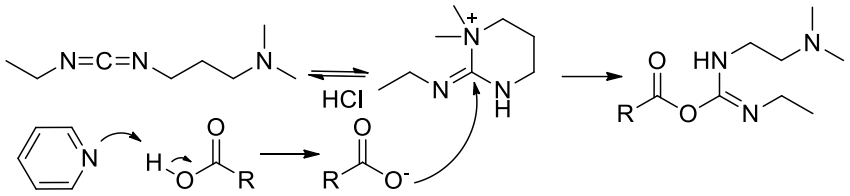


Scheme 6: Synthetic route of compound **38**.



Scheme 7**:** EDCI coupling of 2-indole carboxylic acid **43** with the coumarin core **37** results in formation of indole amide **44**.

Mechanistically, the reaction proceeds via self-rearrangement of EDCI to give a carbocation (Mojarradi 2010), which is then hydrolyzed into a urea derivate in the absence of a dissociated carboxylic acid. Addition of the carboxylate leads to attack on the carbocation giving an *O*-acylisourea (Scheme 8).



Scheme 8**:** Activation of carboxyl group by EDCI. R = organic groups.

However, on carrying out the planned coupling reaction of **37** and **40** catalyzed by EDCI it led to a complex mixture of products that were difficult to purify. This was surprising at first as we had previously have used this in a related coupling reaction on an analogous compound (Simon, Huang et al. 2017). Following literature precedent, we used 2 equivalents of EDCI in the reaction, but this led to a range of side products due to a number of active *O*-acylisourea groups that were formed, leading to multiple side products, which could easily be observed from the corresponding LCMS analysis of the crude reaction mixture (Figure.23 -26).

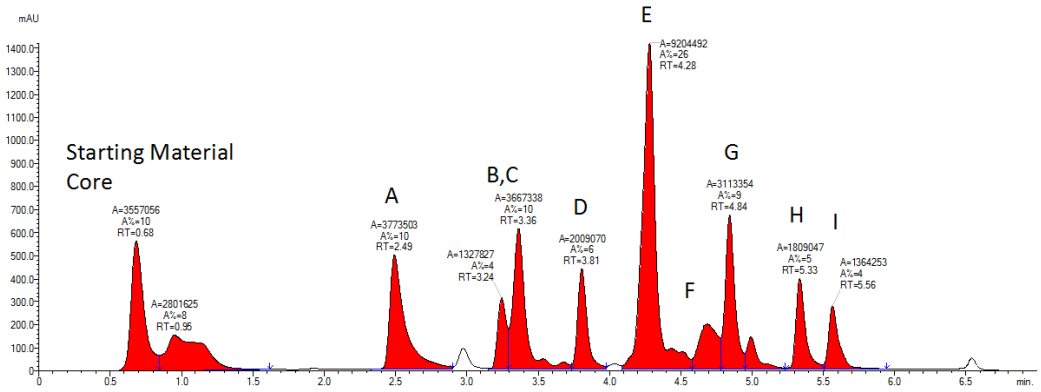
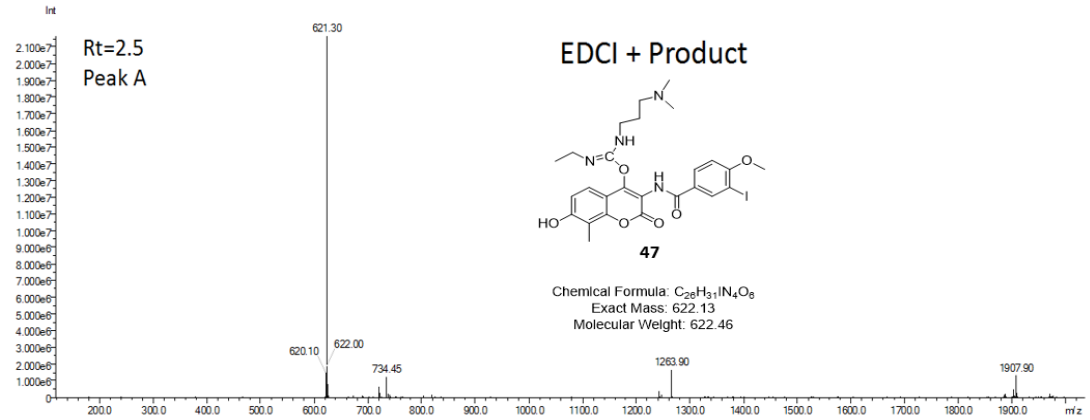
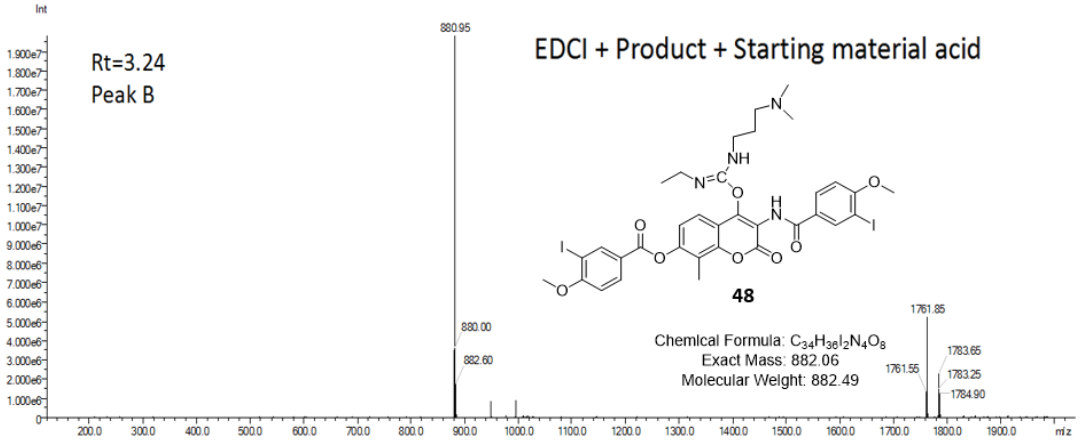
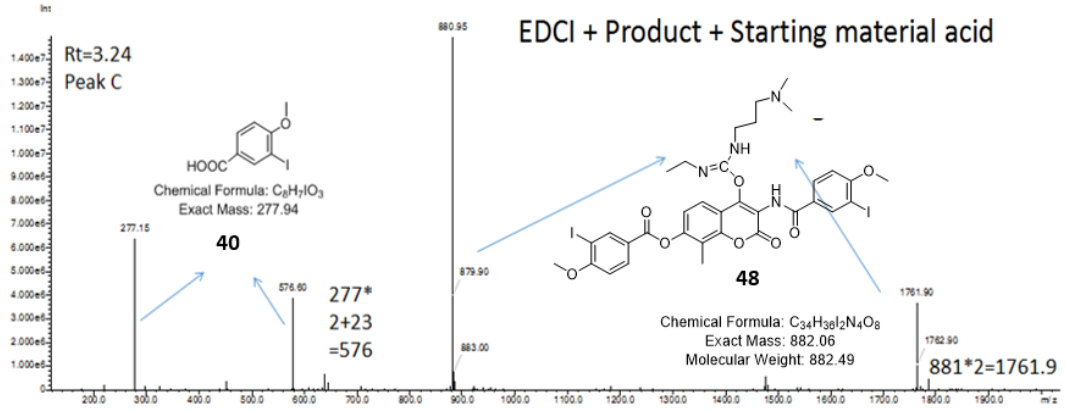


Figure 23**:** LC result of EDCI coupling reaction between **37** and **40**. Sampled from the crude product from the reaction shown in Scheme 6.


Rt=2.5
Peak A
EDCI + Product
47
Chemical Formula: $C_{26}H_{31}IN_4O_6$
Exact Mass: 622.13
Molecular Weight: 622.46
621.30
622.00
620.10
734.45
1263.90
1907.90


Rt=3.24
Peak B
EDCI + Product + Starting material acid
48
Chemical Formula: $C_{34}H_{36}I_2N_4O_8$
Exact Mass: 882.06
Molecular Weight: 882.49
880.95
880.00
882.60
1761.85
1761.55
1783.65
1783.25
1784.90


Rt=3.24
Peak C
EDCI + Product + Starting material acid
40
Chemical Formula: $C_8H_7IO_3$
Exact Mass: 277.94
277.15
576.60
277*
2+23
=576
48
Chemical Formula: $C_{34}H_{36}I_2N_4O_8$
Exact Mass: 882.06
Molecular Weight: 882.49
880.95
879.90
883.00
1761.90
1762.90
881*2=1761.9

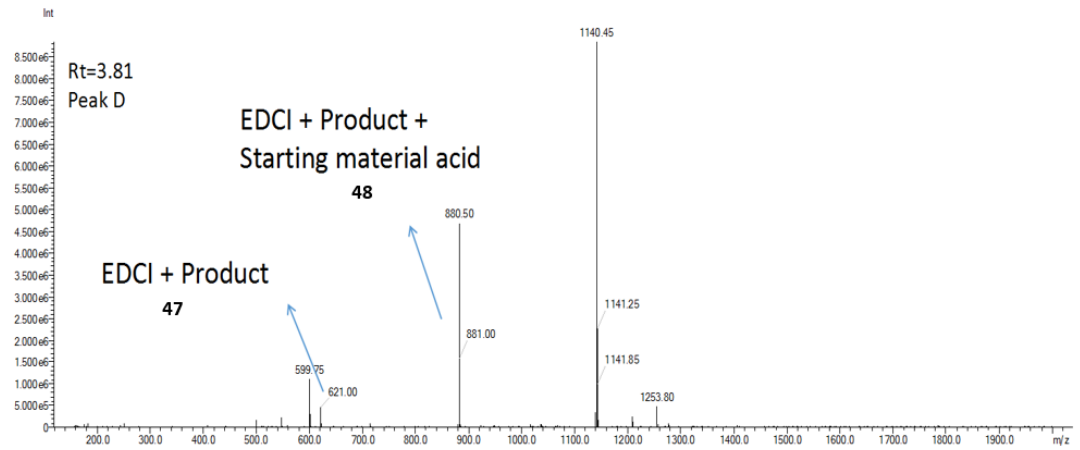
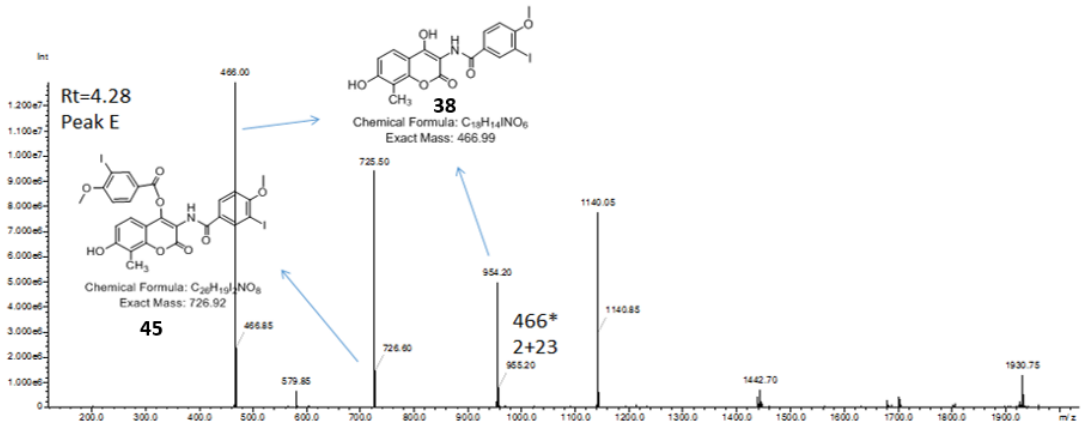
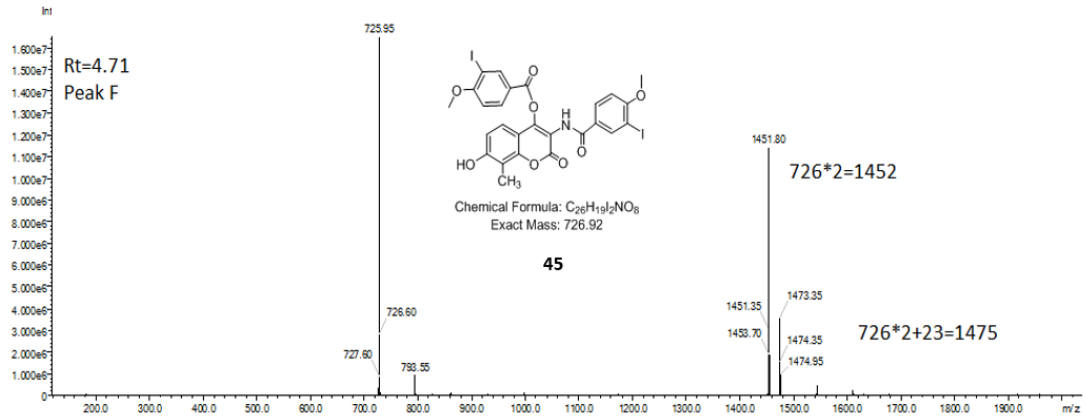


Figure 24**:** MS analysis of LC in Figure 23, which is the crude product of **37** reacted with **40** through EDCI coupling. This figure shows MS data from peak A to peak D.





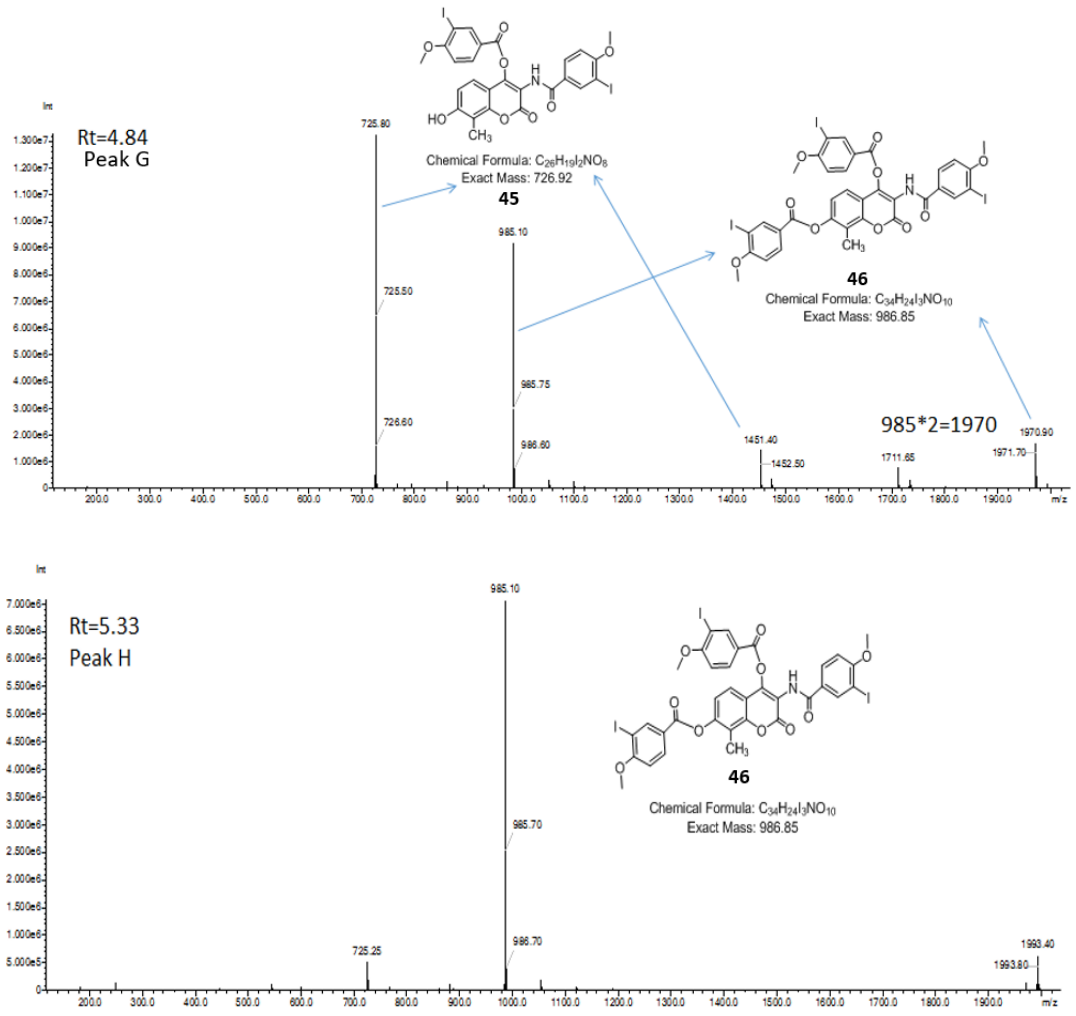


Figure 25**:** MS analysis of LC in Figure 23, which is the crude product of **37** reacted with **40** through EDCI coupling. This figure shows MS data from peak E to peak H.

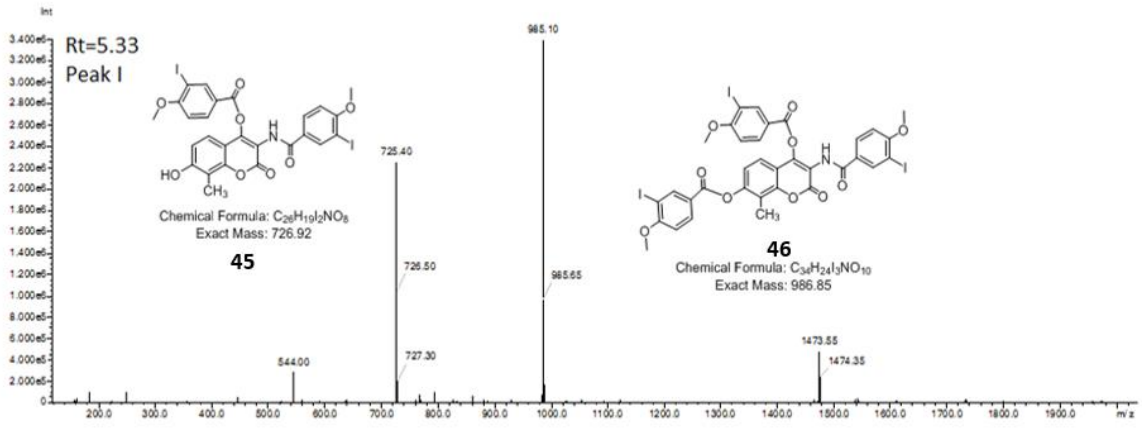


Figure 26**:** MS analysis of LC in Figure 23, which is the crude product of **37** reacted with **40** through EDCI coupling. This figure shows MS data of peak I

Whilst the reaction is, theoretically relatively trivial, one of the key challenges relates to the core coumarin structure of novobiocin **37**. In the core molecule, there are three possible sites for reaction with EDCI leading to the by-products shown and determined from the LCMS analysis. Coupling had occurred at the 3-, 4- and 7-positions (See side product structures 45, 46, 47 and 48, in the LCMS spectrum Figure 24-26). Whilst protection group strategies were considered as one way to mitigate these by-products, they were lowered down the list of preferred approaches as they would add further synthetic delays to the synthesis. We therefore elected to investigate modifications in the reactant ratios to try to optimize the reaction yield. We decided to explore the modification of the amount of base in the reaction, temperature and the number of equivalents of EDCI as shown below (Table 3 on next page).

Table 3**:** EDCI coupling experimental condition set. All entries start with 20 mg core, EDCI-1 as reference.

Entry	**Acid Equiv.**	**EDCI Equiv.**	**Solvent**	**Temp.**	**Note**
EDCI-1	2	2.5	30% pyridine/DCM	rt	
EDCI-2	1	1	30% pyridine/DCM	rt	
EDCI-3	2	2	30% pyridine/DCM	0˚C	Red colour
EDCI-4	1	1	DCM	rt	Not soluble
EDCI-5	1	1	10% pyridine/DCM	rt	Not soluble
EDCI-6	1	1	DMF	rt	Black solution
EDCI-7	1	1	50% pyridine	rt	
EDCI-8	1	1	30% pyridine/DCM	-78˚C	Red colour
EDCI-11	1	1	30% pyridine/DCM	-78˚C	1 drop/min

Each entry was sampled and sent for LCMS analysis when the reaction was complete to determine the best condition and to study the formation of side products. The raw LC traces are presented on the next page as a reference for further discussion. Although each LC traces varies in some details, there are 4 LC peaks (marked as A, B, C and D in Figure 27) appeared repeatedly in all entries. A detailed characterization of peak A, B, C and D in Figure 27 is also given on two pages after.

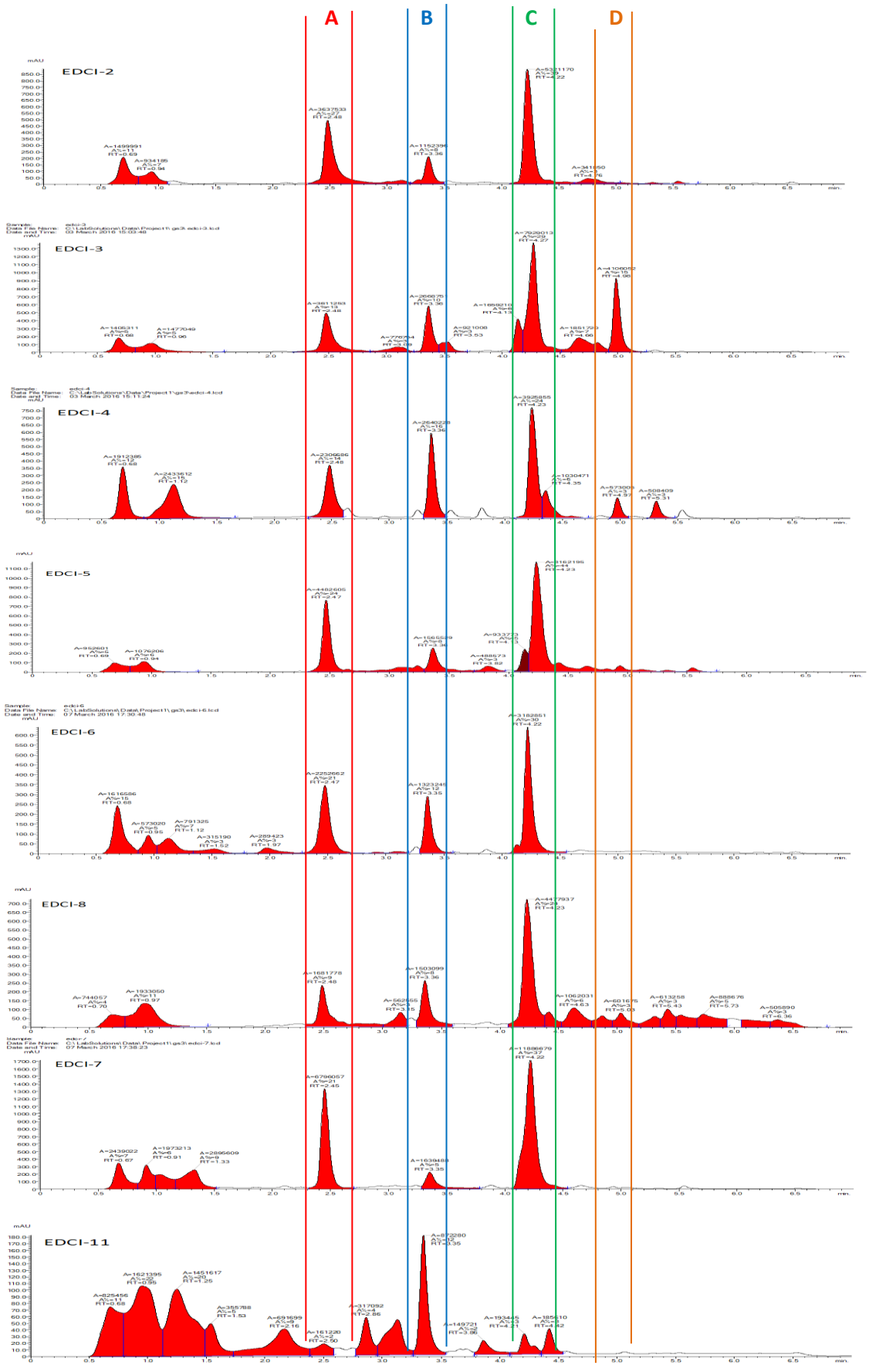


Figure 27**:** LCMS results of parallel experiments of **37** reacted with **40**.

According to the LCMS spectrums in Figure 27, the two peaks located from 0.5min-1min correspond to the starting material **37** (coumarin core).

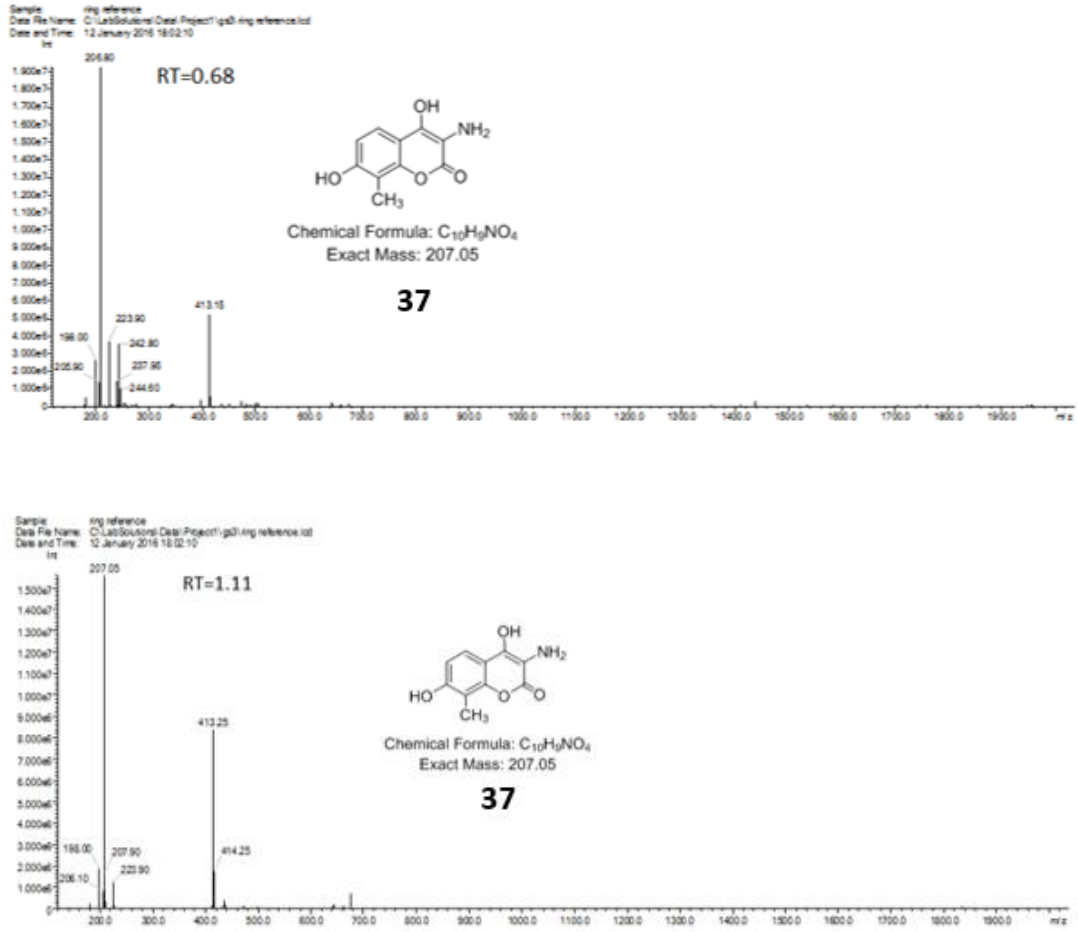


Figure 28**:** MS trace of starting material coumarin core **37**. Starting material 37 was traced by LC in all traces listed in Table 3 and Figure 27, from RT = 0.5 to 1min. RT is short for retention time.

Although the NMR of the starting material indicated it was a pure compound, under reverse-phase column two peaks with the exact same MS spectra can be observed. This is probably due to the different forms of ionized products at the pH of the HPLC solution.

MS for peaks (A, B, C, D) in Figure 27 are

Peak A:

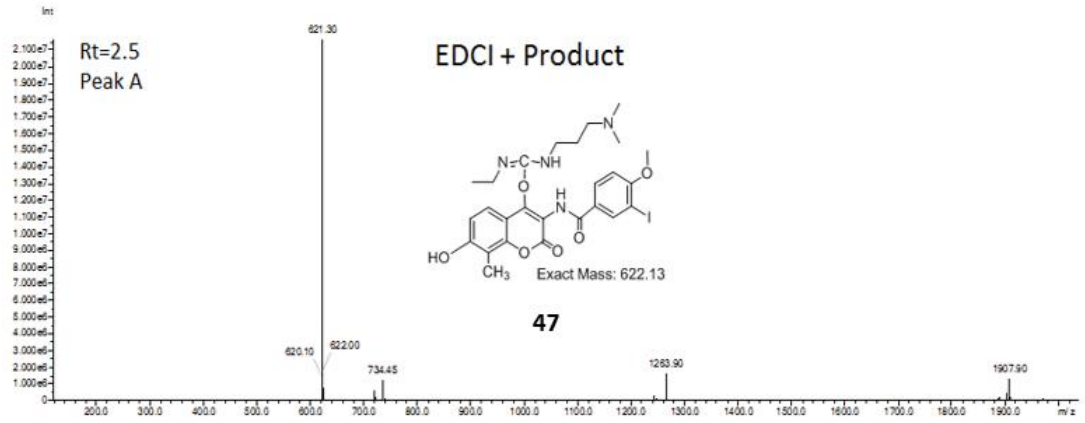


Figure 29**:** MS trace of product peak A in EDCI parallel experiments (Table 3 and Figure27). Peak A is confirmed to be a side product with the same mass as structure **47**. (Proposed structure for MS m/z value, not confirmed)

Peak A is when the product reacts with 1 molecule of EDCI on a certain hydroxyl group, either 4' or 7'. It is not certain if the EDCI assembled on either the 4 or 7 hydroxyl group, as a result, we presented a proposed structure **47** in Figure 29 as an example.

Peak B:

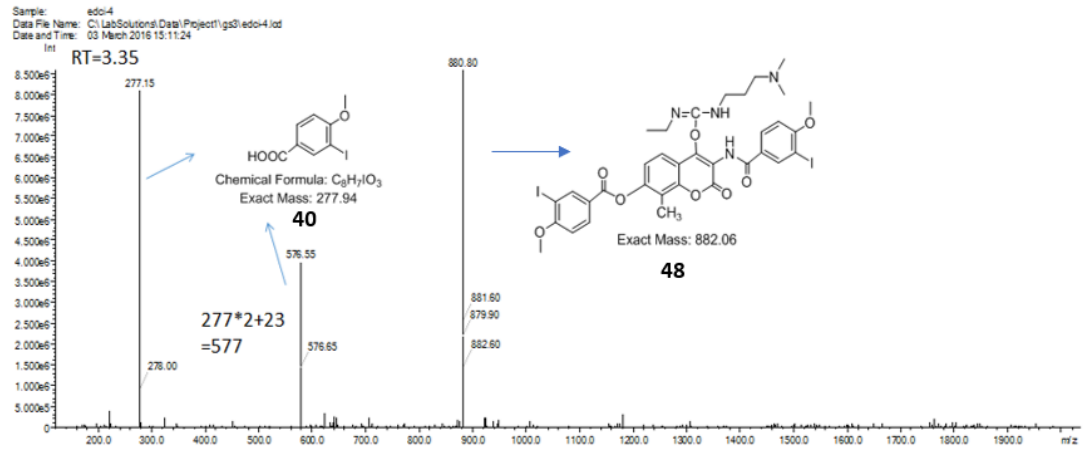


Figure 30**:** MS trace of product peak B in EDCI parallel experiments (Table 3 and Figure 27). Peak B contains starting material **40** and side product **48** (**48** is a proposed structure)

Peak B is a mixture of the acid starting material **40** and a side product **48**. m/z calculation indicated that **48** has the mass of compound **47** plus an additional addition of moiety **40.** These two compounds, **40** and **48**, have the same retention time Rt = 3.35 min on the spectrum shown in Figure 27. However, sometimes the MS trace at 3.35 minutes will change due to the basicity of the solvent, which will be discussed later.

Peak C:

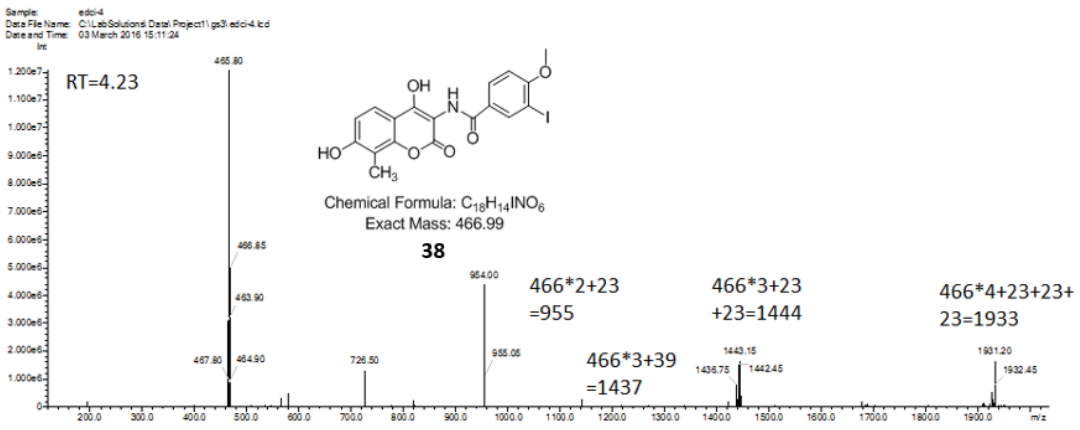


Figure 31: MS trace of product peak C in EDCI parallel experiments (Table 3 and Figure 27). Peak C represents the target product **38**

Peak C at 4.23 min is the desired product with either Na^+ (+23) or K^+ (+39).

Stoichiometry plays a critical role in terms of side reactions. Once it is reduced to 1:1:1 eqv. on each starting material (Table 3, EDCI-2 condition), only two forms of side product could be seen (**47** in Peak A and **48** in Peak B). Low-temperature caused an incomplete reaction (Table 3, EDCI-3 condition) which resulted in the formation of Peak D that represented the activated O-acylisourea form of carboxylic acid (Figure 3, line 2, entry EDCI-3).

Changing the solvent system with different percentages of pyridine did not affect the yield of **38** significantly (range from 24%-40%). However, either removal of pyridine (entry EDCI-4 in Table 3) or reduced concentration of

pyridine (entry EDCI-5 in Table 3) did reduce the solubility of the coumarin core significantly. In order to increase the solubility, a better solvent DMF (entry EDCI-6 in Table 3) was also tested. Although the LCMS results obtained did not show much difference when the solvent was changed to DMF, it generated an extra black colour and a few new side products with a retention time range from 1.0-2.5min (Figure 27 and Table 3, entry EDCI-6). In addition, the concentration of base affected the final amount of starting material left inside the crude product. For example, in entry EDCI-6 (condition see Figure 27 and Table 3), with no base added, peak B showed the existence of starting material carboxylic acid (Figure 32 top trace). On the contrary, in entry EDCI-7 (condition see Figure 27 and Table 3), where 50% of pyridine was used, there was almost no trace of acid starting material left in the system (Figure. 32, bottom trace) and MS trace of peak B in entry EDCI-7 showed only the trace of side product 48 (Figure 32, bottom trace).

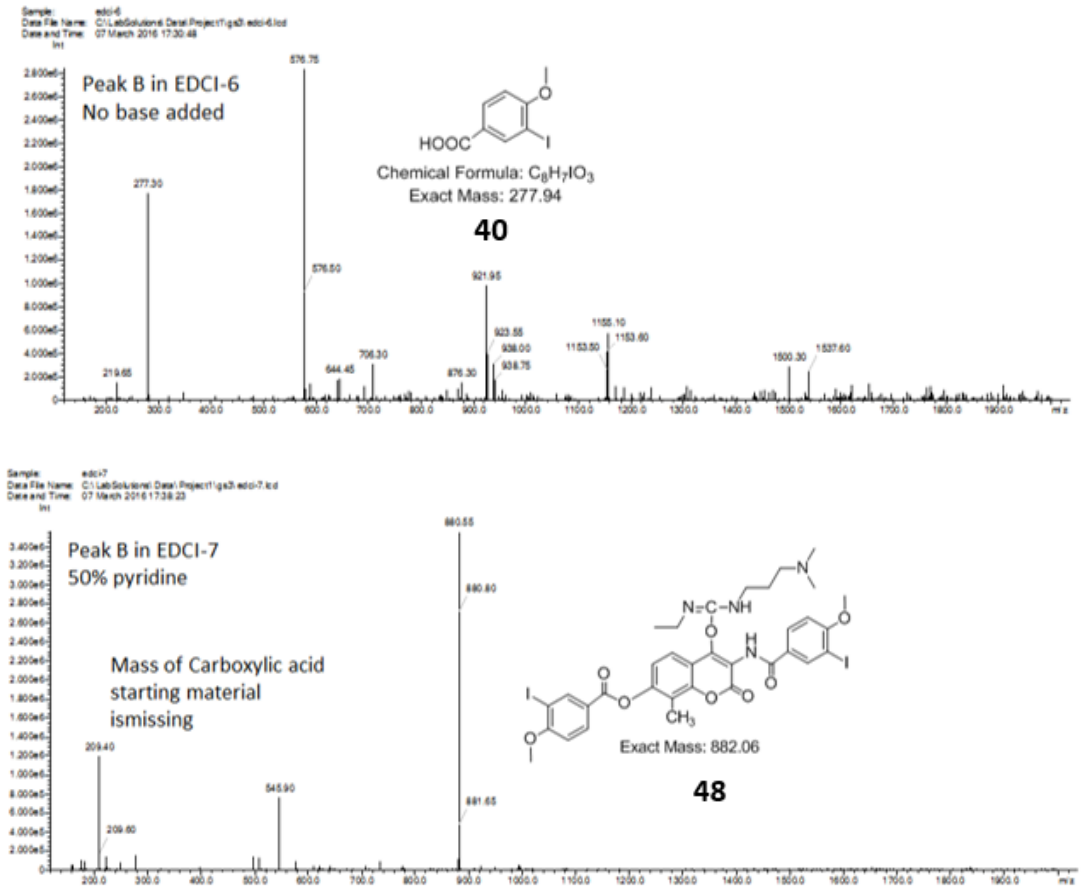


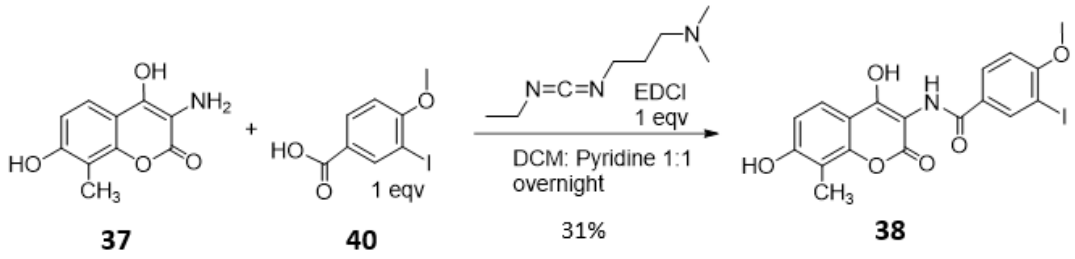
Figure 32**:** Concentration of base (pyridine) affects the amount of unreacted carboxylic acid left in the crude product. (side product 48Proposed structure for MS analysis)

Basicity is one of the most critical conditions that controlled the yield in this case. High basicity promoted carboxyl activation and increase solubility but contributed to the generation of side products. On the other hand, low basicity resulted in poor solubility and inadequate carboxyl activation, however, it depressed the generation of any side products.

Another problem that was encountered was the poor solubility of the side product **47** (see Figure 29 for peak A) in common solvents (including pyridine), which made it very difficult for column purification. Therefore, in order to reduce the generation of side product, a trial (see EDCI-8 in Table 3 and Figure 27) under -78°C was carried out in order to enhance the selectivity of a more active 3'-NH_2. However, although peak A which is side product **47** in entry

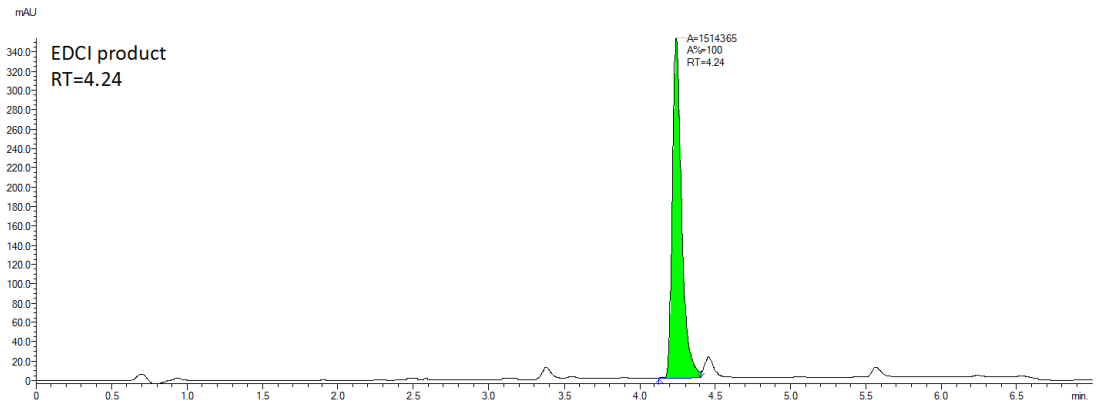
EDCI-8 was significantly reduced, it generated many new uncharacterized side products (see Figure 27, entry EDCI-8, from retention time = 4.5 to 6.5 min). A repeated reaction was carried out with a slower addition of acid. Surprisingly it did not give any product. It is possible that the low temperature significantly slowed down the amide coupling process and resulted in no product.

In conclusion, the condition of EDCI-7 in Table 3 was selected as the optimal conditions despite its limitations (Scheme. 9):



Scheme 9: Selected optimal condition (1 eqv. EDCI, 50% DCM in pyridine as solent and stirred overnight) for EDCI coupling of **37** and **40** with 31% yield.

After the reaction was complete, the mixture of crude products was purified by flash column chromatography using dry loading. The first spot separated with R_f = 0.33 which was the desired product (Figure. 33).



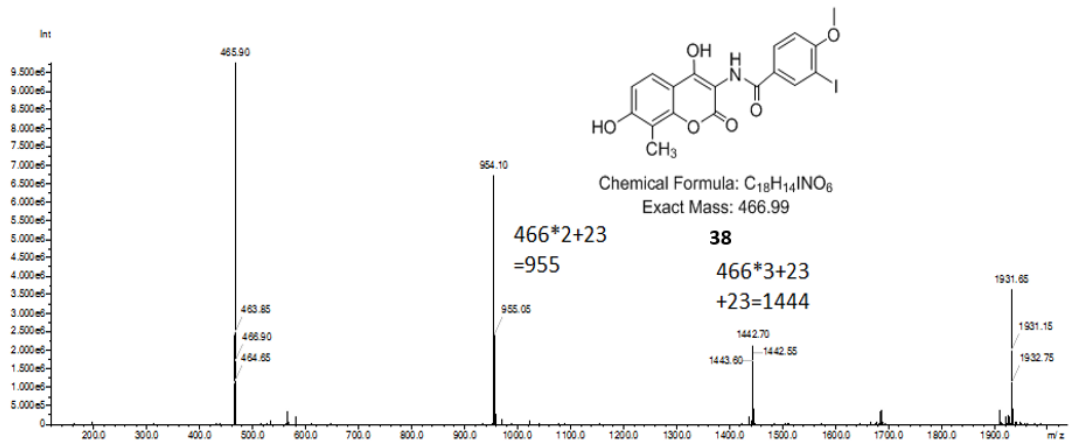
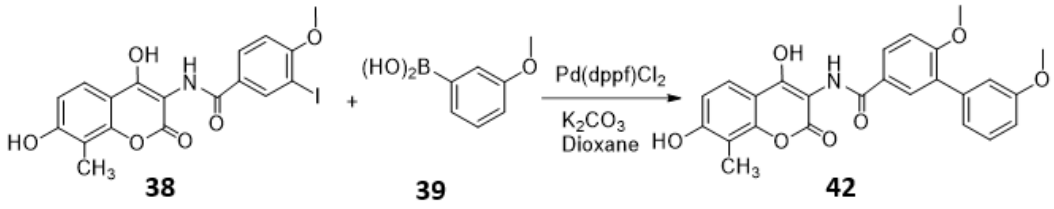


Figure 33**:** LCMS of pure product of N-(4,7-Dihydroxy-8-methyl-2-oxo-2H-chromen-3-yl)-3-iodo -4-methoxy-benzamide **38**.

2.2.3 Synthesis of 6,3'-Dimethoxy-biphenyl-3-carboxylic acid (4,7-dihydroxy-8-methyl-2-oxo-2H-chromen-3-yl)-amide 42 (Route1)

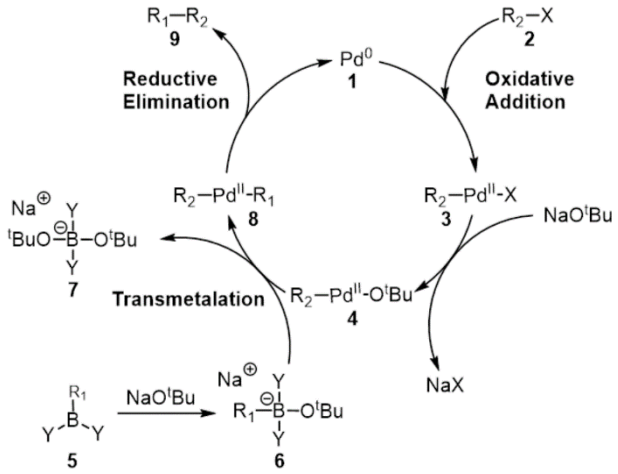
Suzuki coupling is one of the most powerful C-C bond formation reactions. Using easily prepared Pd salts as the catalyst, this reaction converts organohalide and boronic acid into C-C bond (Suzuki, Akira 1979). Since 1979, the Suzuki reaction has widely been used in different areas of synthetic chemistry as a preferred method of C-C bond formation and was rewarded with a Nobel Prize in 2010 (Lipshutz, Taft et al. 2012). The Suzuki reaction was used in our synthesis of compound **42** (Scheme 10), and the mechanism is shown in Scheme 11.



Scheme 10: Synthesis of biaryl-amide using Suzuki coupling. No yield is given due to unsuccessful purification.

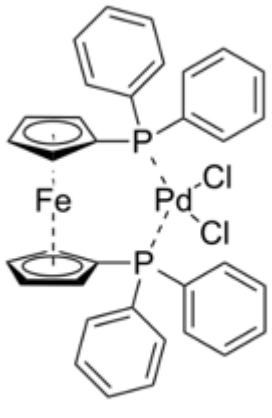
Pd metal (**1**) undergoes oxidative addition with the starting material organohalide (**2**), resulting in the formation of **3**. Then the base will remove the

halide attached to the metal to give an intermediate **4**. Before further reaction, boronic acid (**5**) must be deprotonated by the base to form a salt (**6**). Then the boronic salt will take a ligand from Pd but, in return giving back an organic group R1 to Pd through transmetalation. The newly formed Pd coordinates (**8**), which contain R1 and R2, then undergoes reductive elimination to give the C-C coupled product (Scheme 11).



Scheme 11: Mechanism of Suzuki coupling (Kim and Yu 2003).

Taking advantage of the development of Pd catalyst, Pd (dppf) Cl_2 was selected as the metal catalyst due to the electronic donating effect of dppf group (Burlison, Neckers et al. 2006).



49

Figure 34: Chemical Structure of 49 Pd(dppf)Cl_2.

To carry out the experiment, firstly, boronic acid **39** (2 eqv) and compound **38** was dissolved in dioxane with K_2CO_3 (2 eqv) and stirred for 30mins. This allows enough time for boronic acid **39** to be transformed into boronic salt. Then metal catalyst **49** $Pd(dppf)Cl_2$ (3% eqv) was added to start the reaction. The colour of the system started as a red colour (caused by Pd salt) and ended up as a dark colour after 2 days. The crude product was sampled for LCMS analysis and the result obtained (Figure. 35) showed a very reasonable yield (60%) of product **42** and the existence of the product's ion was observed by MS trace at LC Rt = 4.43 (see Figure 35 bottom).

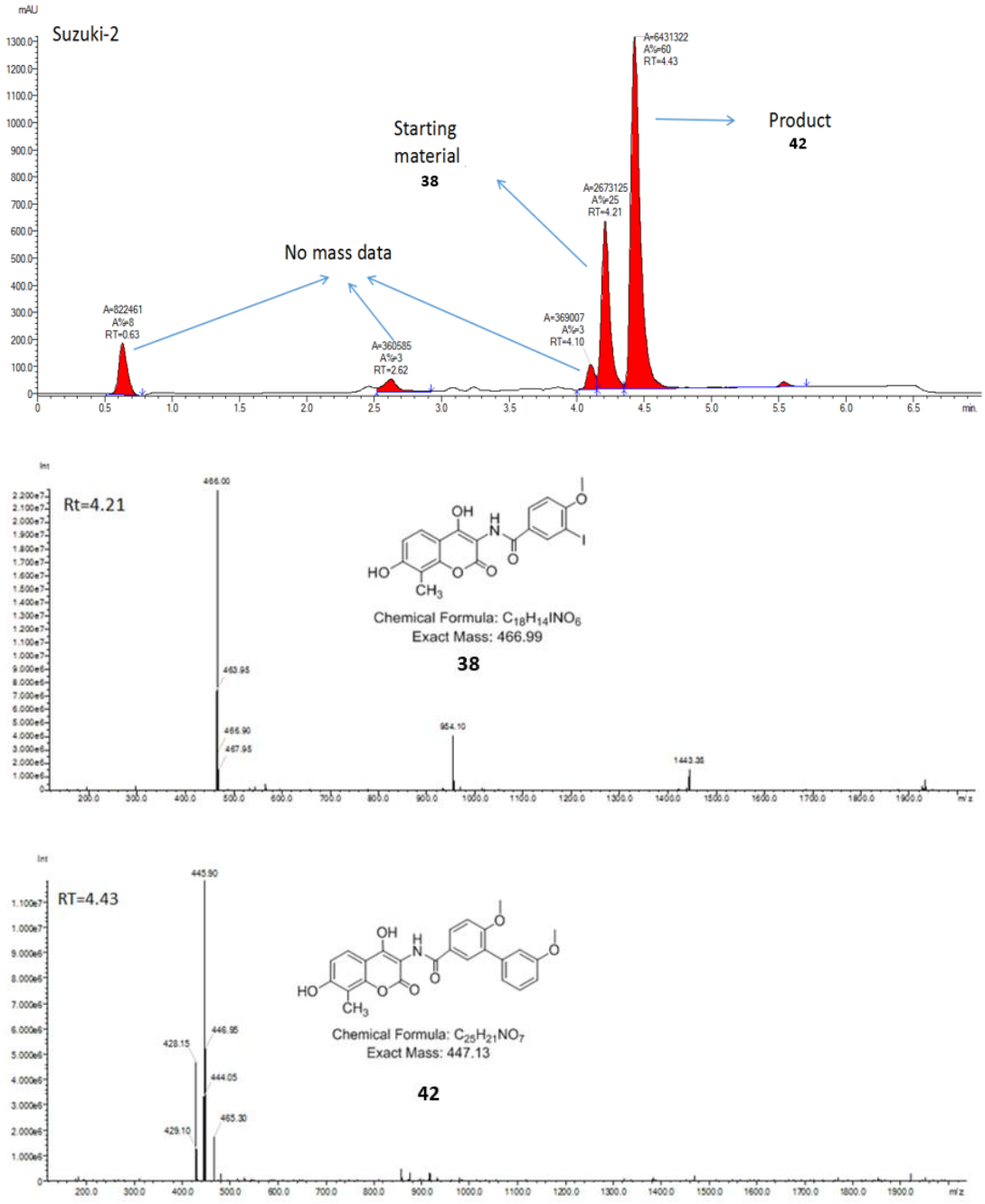


Figure 35: LCMS result which sampled from the crude product of the reaction showed in Scheme 10, in which **38** was reacted with **39** catalyzed by **49** through Suzuki reaction. Product **42** found at Rt = 4.43 min of LC trace.

The crude product **42** was purified by flash column chromatography. However, at this point, several problems occurred. First of all, the Pd catalyst **49**, boronic acid **39**, and product **42** all had a very similar R_f value on TLC (R_f 0.88-0.9). As a result, a slow flow rate was needed (1 mL / 6 s) to separate a fraction of the product from the column. Flow rate higher than 1 mL / 6 s caused incomplete separation.

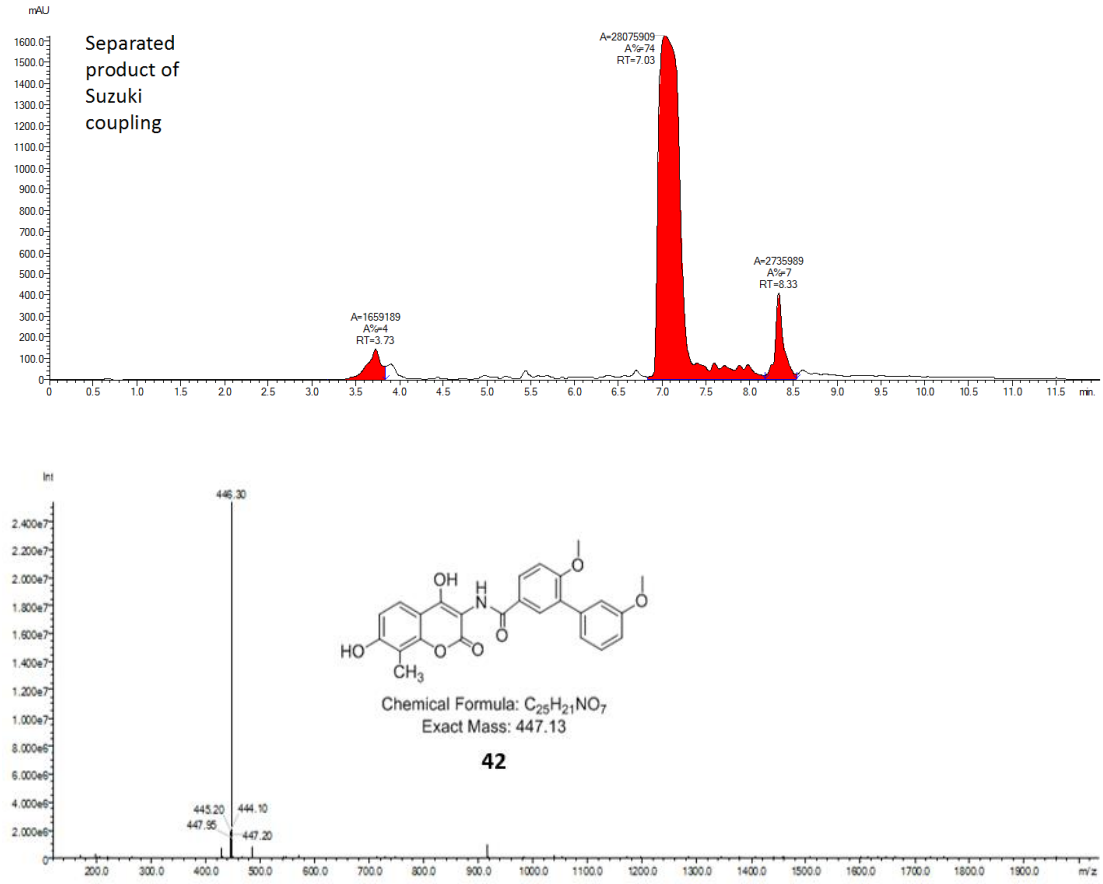


Figure 36: Biaryl amide product **42** separated by a flash column method. The separated **42** was sampled for LCMS and shown above.

Secondly, although the LCMS result showed that the biaryl amide product **42** did exist in the sample (Figure. 36), the NMR test result showed a different conclusion, as the NMR spectra indicated that the sample was not the product but could actually be catalyst **49**. Distinct and huge benzyl group peaks were found from 6 ppm to 8 ppm, whereas all H peaks of **42** were missing

These results indicated that the separated product **42** from its first synthetic trial is not pure and may mixed with catalyst **49** after the separation. In order to investigate whether product **42** and catalyst **49** share a similar elution time through the column. A sample prepared by mixing all starting materials (**38**, **39**, **49** and **42**) was sent for LCMS (LC trace in Figure 37 top side) and compared with the crude product **42** (LC trace in Figure 37 bottom side) It was noticed that product **42** and catalyst **49** indeed similar retention time in LCMS (Figure. 37).

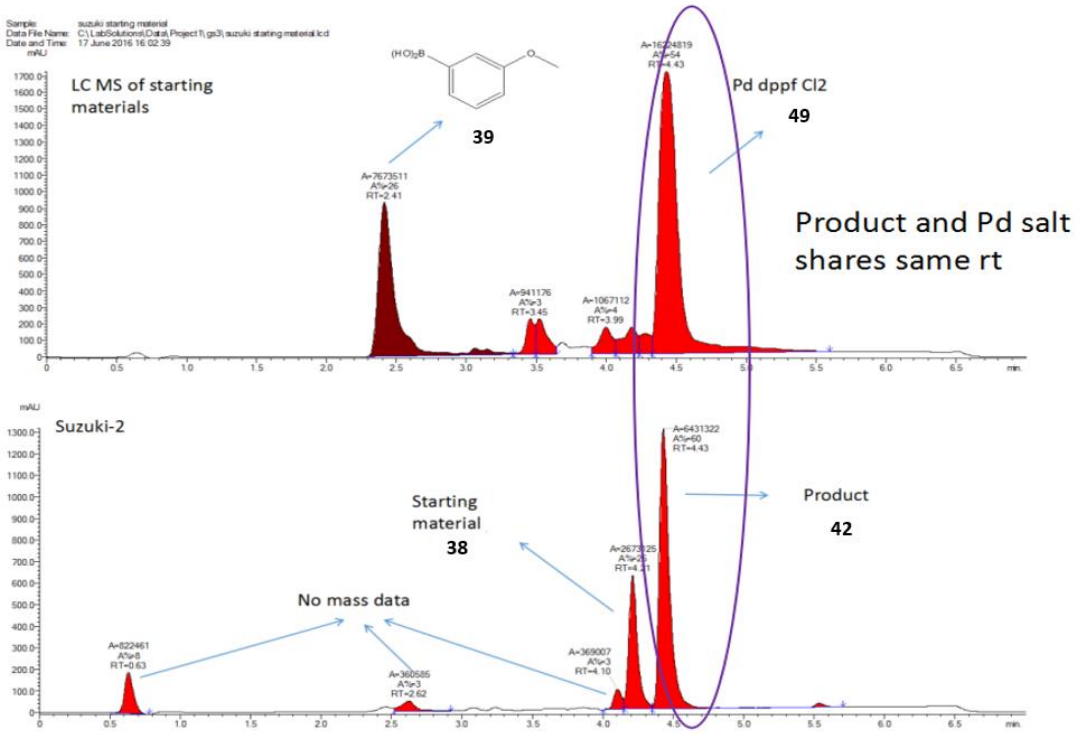
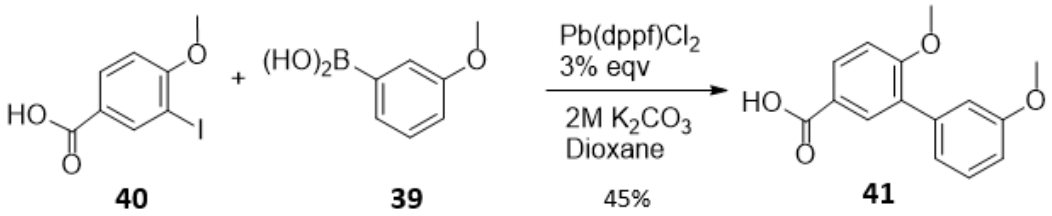


Figure 37: Biaryl amide **42** and Pd(dppf)Cl_2 **49** has the same retention time on LC spectra.

While flash column chromatography was not very applicable for this separation, reverse phase HPLC was probably not an option either because LCMS already showed that the product shares the same retention time (Figure. 37). Much effort was made to remove the catalyst using silica, celite, or activated carbon. However, after several trials, this method was totally banned due to the separation issue.

2.2.4 Synthesis of 6,3'-Dimethoxy-biphenyl-3-carboxylic acid 41 (Route2)

Synthetic route 2 was next carried out where the biaryl carboxylic acid **41** was synthesized first (Scheme 12).



Scheme 12**:** Synthesis of biaryl acid **41** via Suzuki coupling catalyzed by **49**.

The reaction used the same Suzuki coupling conditions as the previous one (conditions applied to Scheme 10); however, the basicity was increased from 2 eqv. to 2M K_2CO_3 for better activation. From the LC/MS spectrum (Fig. 37), starting materials and products were observed and Peak B was the biaryl-acid product of.

However, the separation of the product was still a challenging process, mainly due to the difficult removal of Pd metal and the solubility of the product. When this reaction was quenched, if the resulting residue directly went to the flash column, only the catalyst residue would elute, even though the product was entirely separated on TLC. The rest of the compound left on the flash column could not be washed off except using methanol (Figure. 38). The solvent system used in this case was 3:1 ethyl acetate: hexane. After testing the pH of the mixed crude product, it was discovered that the biaryl acid product **41** was existed in the salt form (due to pH =12, no acid form **41** should exist).

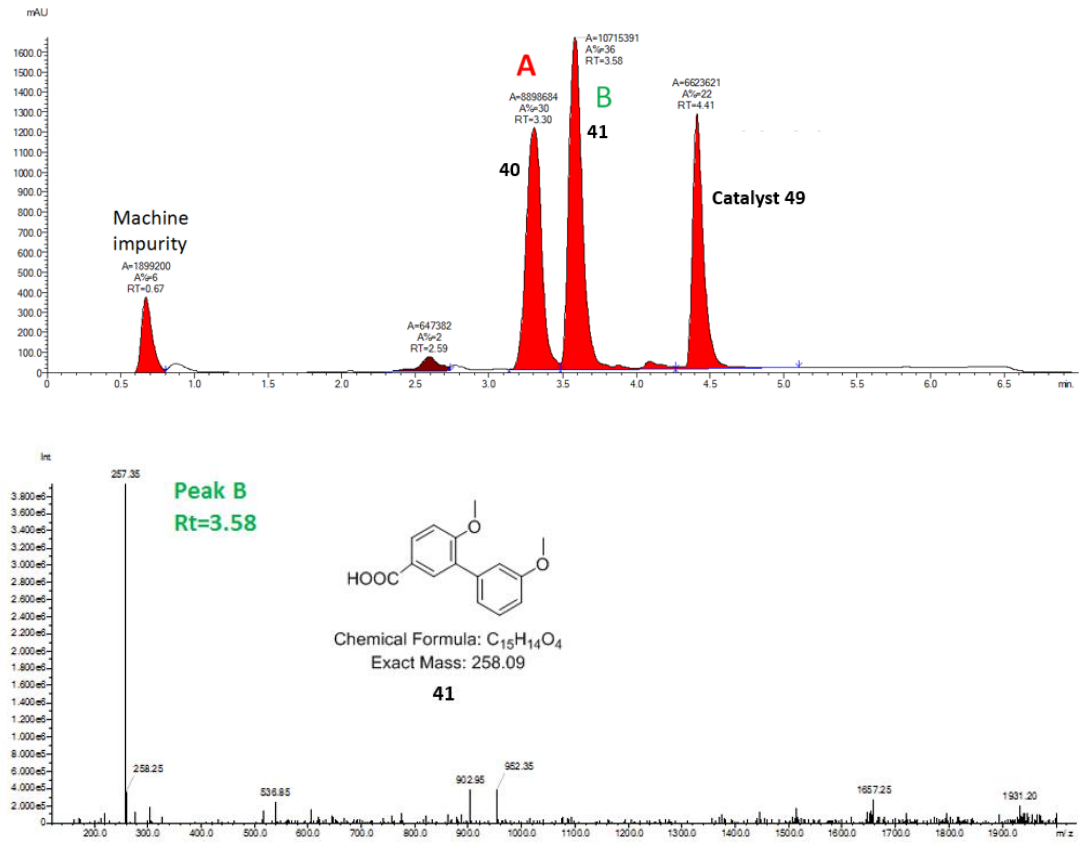


Figure 38: LCMS result of the crude product which sampled from the reaction showed in Scheme 12. m/z in MS trace proved that peak B, Rt = 3.58 is product **41**。

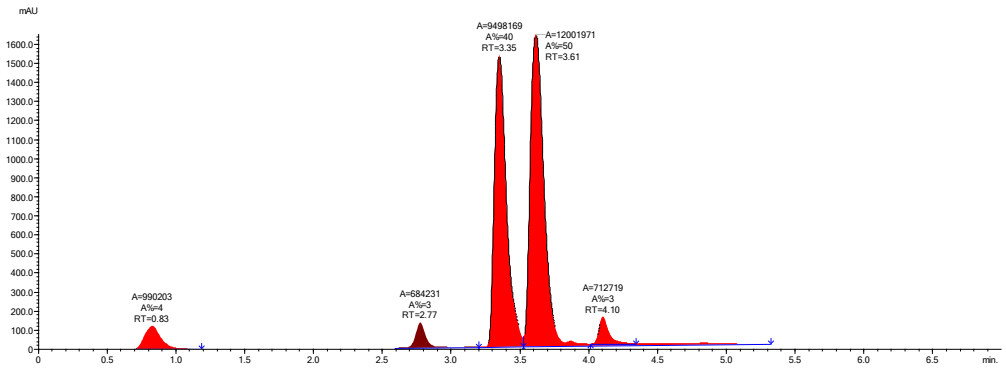


Figure 39: Flash column is not able to separate the product **41** from its crude product. After the crude product purified by flash column using 1: 1 ethyl acetate: hexane solvent system, there are still starting material **40** left inside the residue.

The crude product **41** was separated using 1: 1 ethyl acetate: hexane mobile phase, however, the separated product spot fraction contains product **41** and

starting material **40** （Figure 39）. This is due to both of **41** and **40** are organic salt at pH = 12, which does not dissolve in organic solvent. Therefore, a 7:3 acetone: DCM solvent system (acetone can dissolve salts) was developed for the flash column purification of crude **41**. Pure **41** was separated using acetone based solvent system, however with a very low recovery yield of **41** (14%). This was due to the salt-form of the product **41** dissolves poorly in the organic phase. A later choice of using a reverse phase column on the Biotage automatic purification, which fully dissolved the salt form crude product, achieved a reasonable recovery yield (>56%), and the LCMS of product **41** purified by automated reverse-phase LC (Biotage) is shown below (Figure 40):

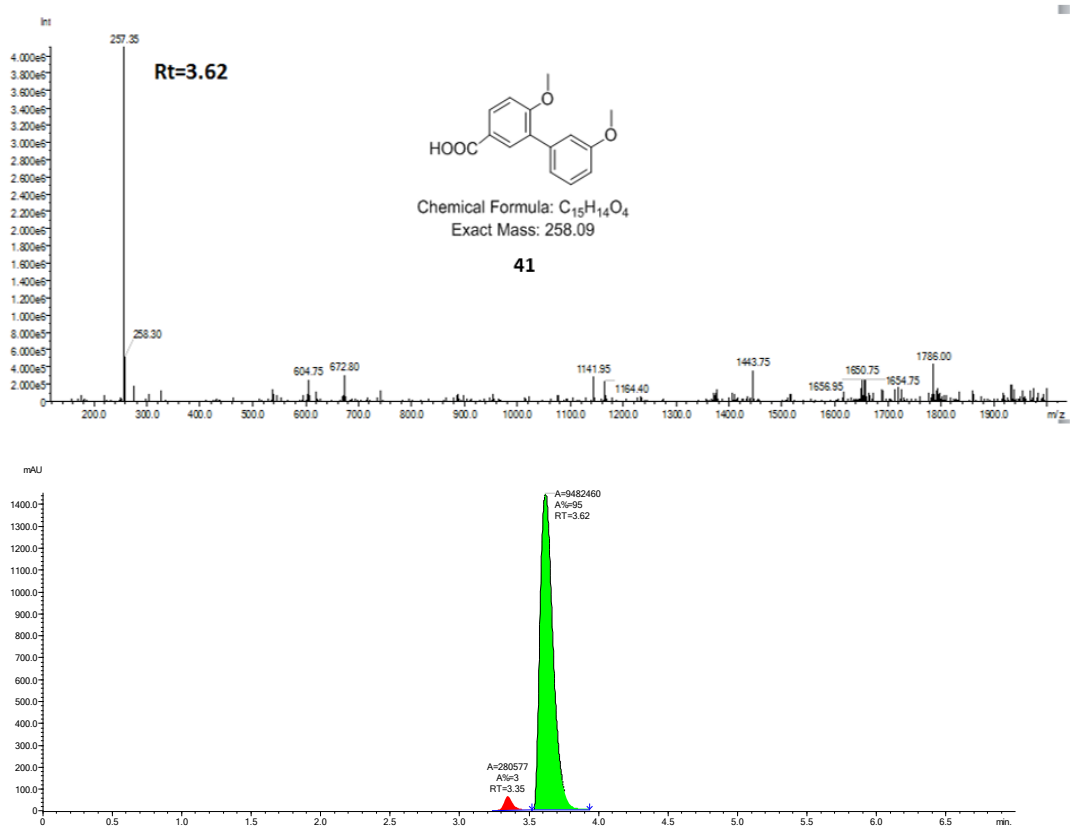
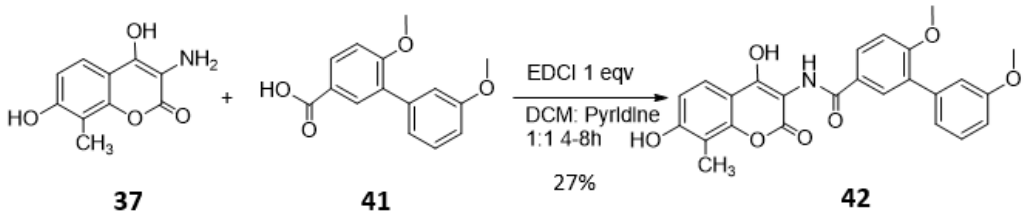


Figure 40**:** LCMS of pure biaryl carboxylic acid **41**, purified using automatic reverse phase column system Biotage. Pure **41** was purified from crude product **41** (the sample shown in Figure 38)

2.2.5 Synthesis of 6,3'-Dimethoxy-biphenyl-3-carboxylic acid (4,7-dihydroxy-8-methyl-2-oxo-2H-chromen-3-yl)-amide 42 (Route 2)

Using the same method developed in the previous study (Section **2.2.2**), the first trial of route 2 amide coupling, entry named icde-1, was carried out with 1:1 and 50% pyridine in DCM (Scheme 13). After compound **41** was added dropwise into the solution of compound **37**, the system was stirred overnight. However, the LCMS result showed that the reaction not only generated the desired target product but also generated several side products (Figure. 41).



37 41 42

Scheme 13**:** Synthesis of biaryl-amide **42** through synthetic route 2 designed in project aim section.

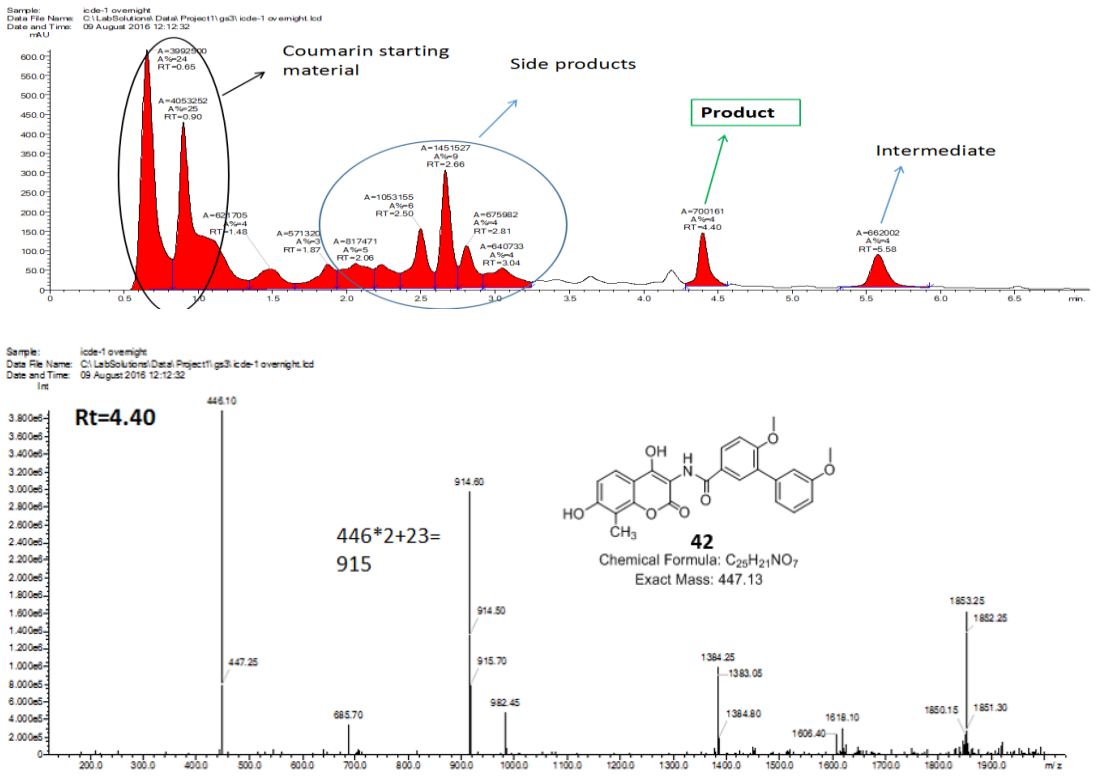


Figure 41**:** LCMS result of crude product sampled from the reaction shown in Scheme 13, entry icde-1. Product **42** was found at LC Rt = 4.40.

The result from the first trial (entry icde-1, Figure 41) showed that the correct product was formed but in low yield. In order to obtain adequate product, same reaction was repeated again (entry icde-2) and the crude product was sent for LCMS (Figure 42). Nevertheless, Figure 42 showed that although the product can be identified from the spectrum, it was very difficult for it to be purified.

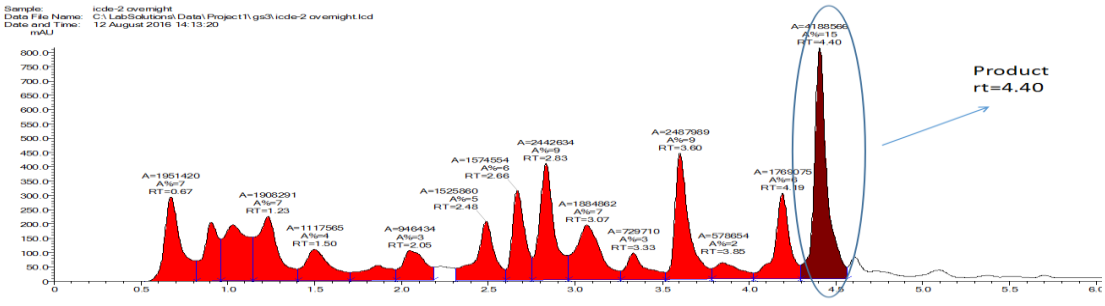


Figure 42**:** LCMS result of crude product sampled from the reaction shown in Scheme 13, entry icde-2. Product **42** was found at LC Rt = 4.40.

To find out the reason why this EDCI coupling generated so many side products, all experimental factors (including reaction time, temperature, carboxyl activation, and mixing) were considered. Most likely, time and mixing were the two key factors in this case. On the one hand, the generation of side products can be a sign of overreaction, meaning that a proper and precise timing for this reaction is important. On the other hand, experimental data in the previous discussion (Table 3 and Figure 27) suggested that inappropriate drop speed of activated carboxylic acid also resulted in the generation of side products.

In order to optimize the yield of product **42** and reduce the generation of side products, a process optimization in-process control study was established using the same condition demonstrated in Scheme 13 (entry icde-4), except the EDCI activated carboxylic acid was added directly instead of dropwise. Figure 43 is the LC in-process controlled results monitored in the first 5 hours, together with a 48-h sample. We monitored the variation of 4 peaks (marked as peak A, B, C and D in Figure 43) to study the in-process transformation of the product and side products. Detailed peak characterization is given from Figure. 44 to Figure 46. The most important conclusion came from this trial was that EDCI activated carboxylic acid must be added directly into the coumarin core solution to eliminate most of the side product. This was confirmed by repeating experiments with the same conditions.

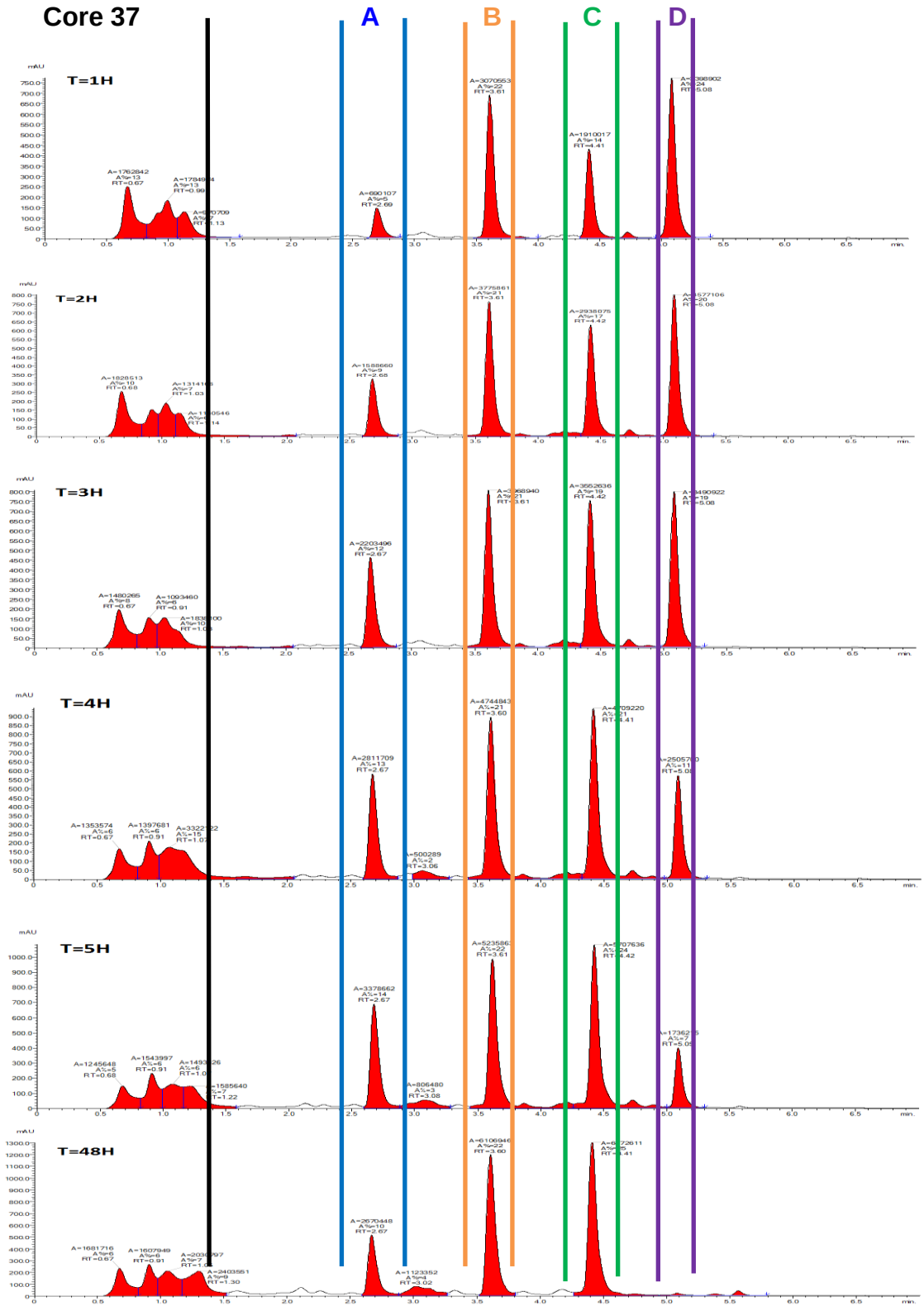


Figure 43: Reaction shown in Scheme 13 was monitored by LCMS hourly in first 5 hours, together with a 48-h check point. For peak A, B, C and D characterization, pleases see Figures 44 – 46. Entry icde-4.

Peak A in Figure 43: Side product **50**

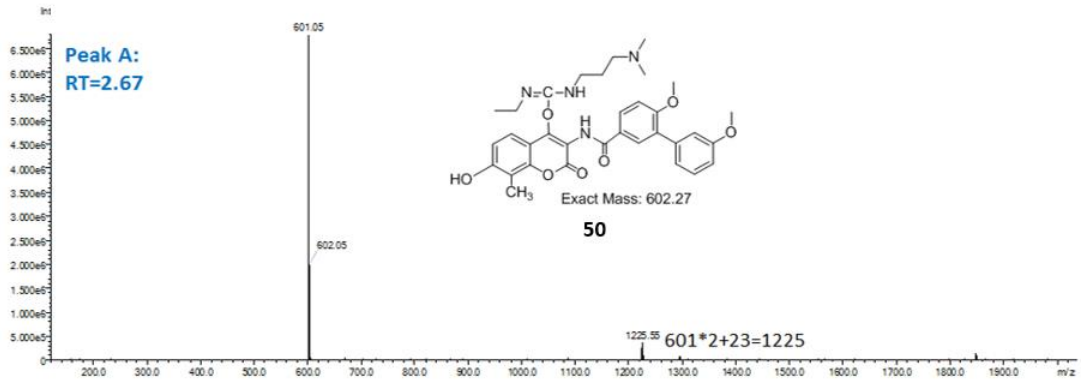


Figure 44: Peak A is the side product **50** where **42** further reacted with 1 molecule of EDCI, either on 4' or 7' position. We proposed structure 50 as an illustration.

Peak B in Figure 43 Starting material of biaryl carboxylic acid and side product

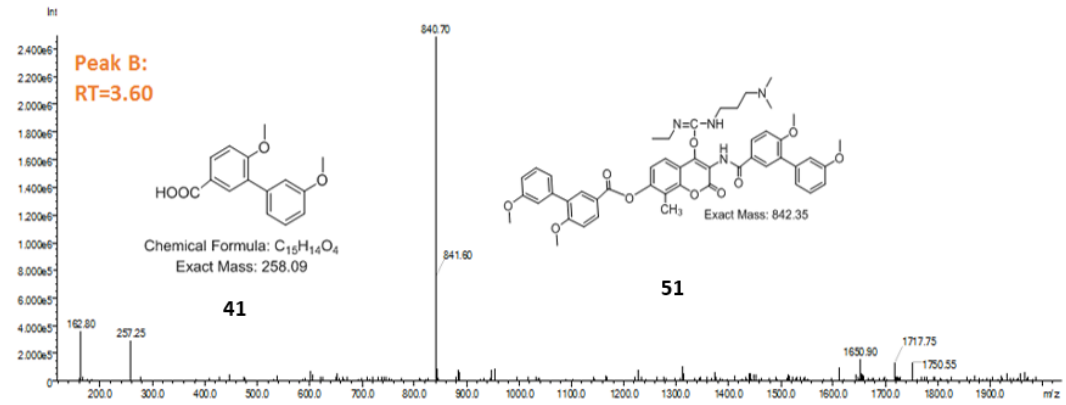


Figure 45: Peak B is the starting material and side product **51** where side product 50 further reacted with one molecule of **41**. **51** is a proposed structure for demonstration.

Peak C in Figure 43: Product **42** + Side product **52**

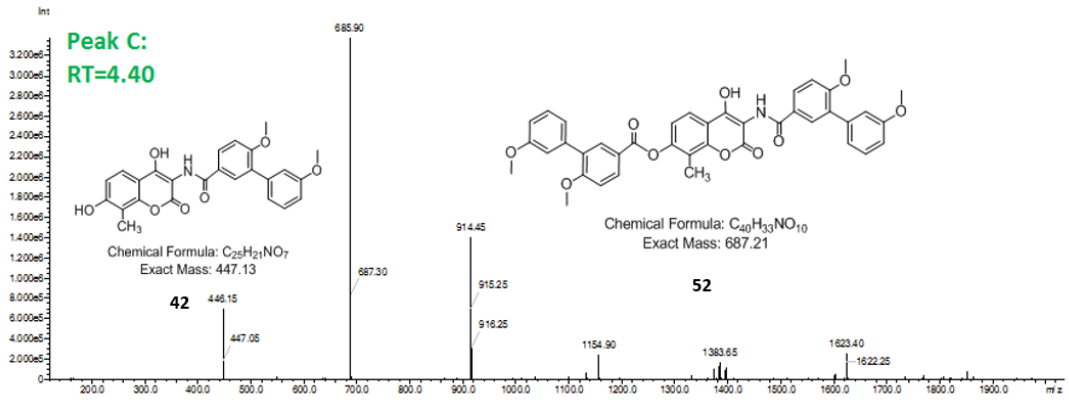


Figure 46**:** Peak C is product **42** + another side product **52**. **52** is formed by **42** further reacted with one molecule of **41**, either on 4' or 7' position. **52** is a proposed structure for demonstration

Peak D: Activated carboxylic acid intermediate.

From MS data obtained, Peak B is the biaryl amide product added **42** one molecule of EDCI attached to it. Product **42** was found in Peak C, but it probably shared a similar retention time with another side product **52** with a mass of 667.21 (Figure 46, right side).

Peak D is the intermediate form of carboxylic acid activated by EDCI (active form of **41**). The decrease of peak D marks the process of the reaction. As a result, product formation curves are plotted based on data in Figure 43 to calculate the speed of reaction (Figure 47).

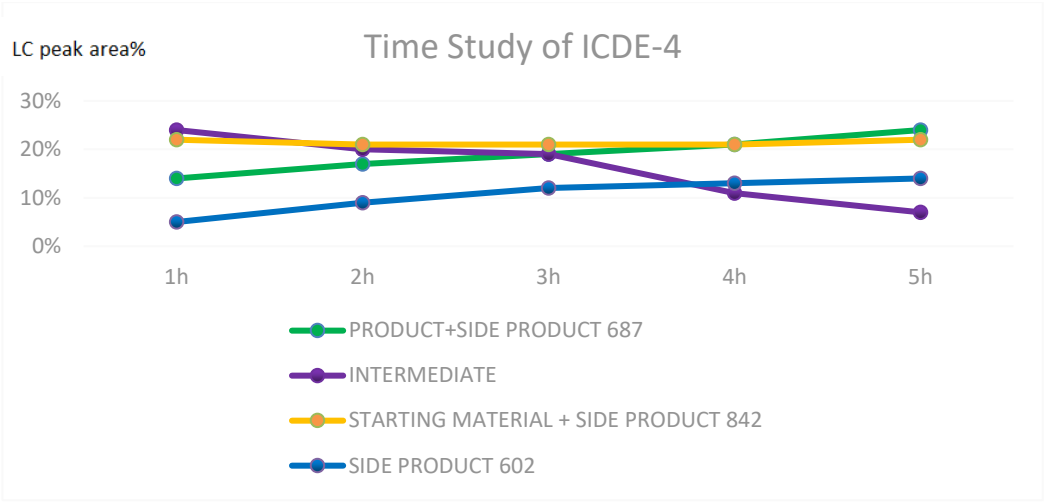


Figure 47: Line chart plotted using date in Figure 43. Blue= Peak A side product **50**, Orange = Peak B side product **51** and starting material **41**, Purple = Peak D activated **41**, Green = Peak C product **42**.

As shown in Figure 47, the amount of activated intermediate **41** (peak D) gradually decreased over time and only 7% was left at the end of 5 hours. As a result, it could be estimated that the reaction would finish in approximately 6-7 hours. The decrease of intermediate means it was transformed into either product or side products. Thus, it was evident that peak A (side product **50**) and C (product **42**) kept increasing at a similar rate (about 2.35% / hour) and the intermediate peak D (activated starting material **41**) transformed into peak A (side product **50**) and C (product **42**). Although peak B, the starting material **41**, should decrease with time, it did not decrease due to it becoming mixed with a new generated side product **51**. In conclusion, the reaction should be quenched within 8 hours (estimated from the curve, under ambient temperature) and both side products and product are formed at the same time.

Due to product and side product **52** shared a same retention time on LCMS, it remains unknown which form was the majority inside the same peak. Fortunately, these two compounds could be separated by a flash column using 3: 2 ethyl acetate: hexane as mobile phase with R_f = 0.71 and R_f = 0.66 respectively. Following purification, on a scale of 200 mg of the coumarin core, 127mg of side product **52** vs 36 mg of biaryl amide product **42** was recovered. In order to achieve a higher yield, 3 eqv of K_2CO_3 was reacted with side product **52** for 2 hours to break the unwanted 7' – substituted ester bond while

unreactive to 3'- substituted amide bonds, through which the side product was 'deprotected' into the desired amide product **42** (top side scheme in Figure. 48). The crude product from this 'deprotection' was sent for LCMS and the result indicated that 3 eqv K_2CO_3 can effective cleave the unwanted biaryl group (Figure 48. Middle LC trace, bottom MS trace) to give the target product **42.**

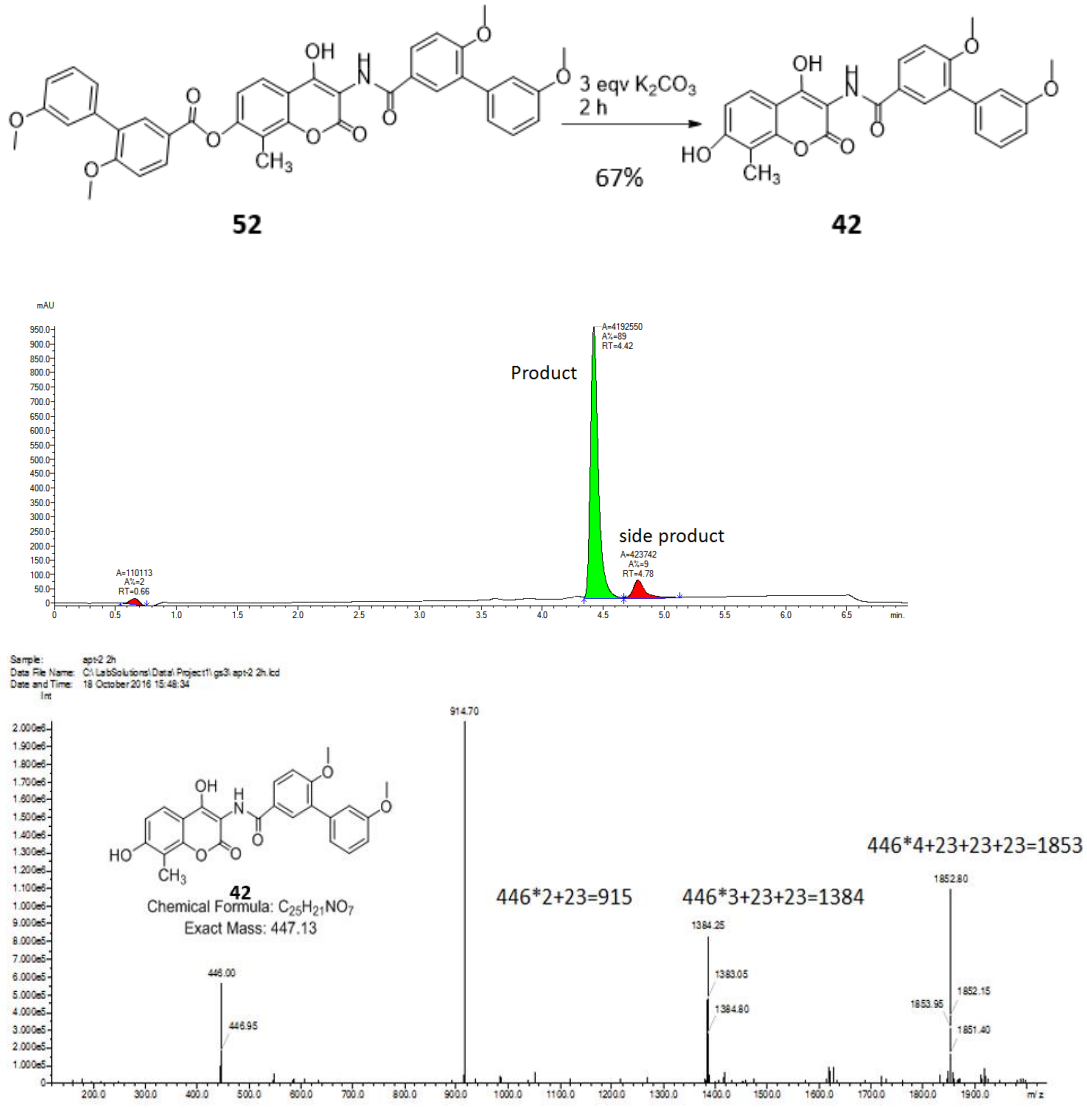
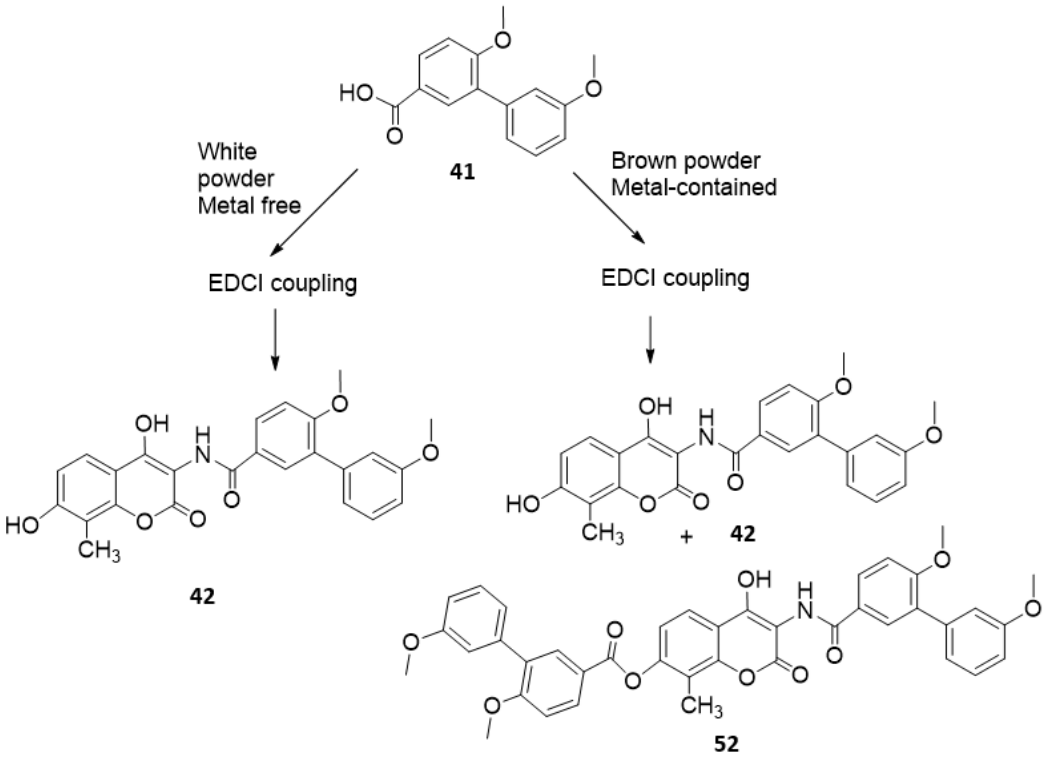


Figure 48: Top scheme: Side product **52** 'deprotected' by K_2CO_3 solution and cleaved the unwanted biaryl group linked by an ester bond. Middle: LC result of the crude product showed the effectiveness of ester hydrolysis. **52** almost transformed into **42**. Bottom: MS trace confirmed that the dominate peak is product **42**.

Usually, the organic acid starting material **41** often contains traceable amount of metal catalyst brought from the last Suzuki-coupling step. When the

traceable metal catalyst in **41** was totally removed by several-times repeated reverse phase column, the brown powder **41** turns into white powder form. Brown and white **41** behaves differently during the amide coupling reaction. Because we found that unlike brown powder **41** contributes to the formation of side product **52** (Scheme 14 left route)**,** applying the metal-free white powder **41** to amide coupling step did not cause the formation of side product **52** (Scheme 14, right route), which meant the catalyst metal probably promoted the side reaction (Scheme 14). The crude product obtained from white powder 41 reacted with 37 was sent for LCMS (Figure 49). MS trace of product **41** peak at Rt = 4.43 presented a significantly reduced signal of side product **52.**



Scheme 14: Metal catalyst impurities in starting material **41** affect the generation of side product **52** in amide coupling step.

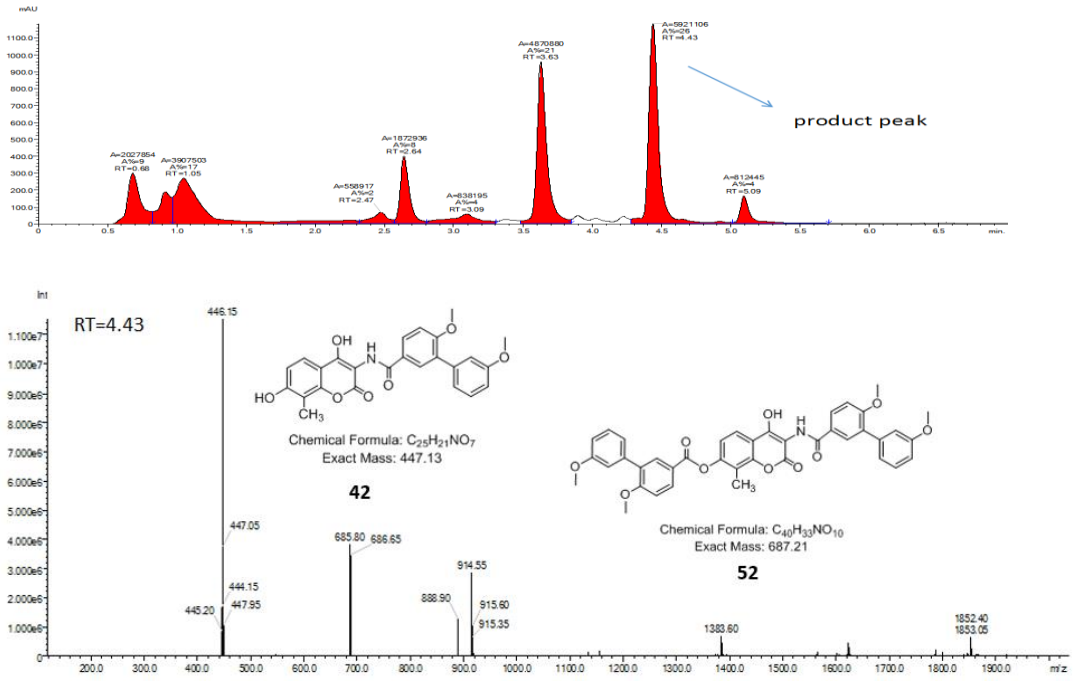
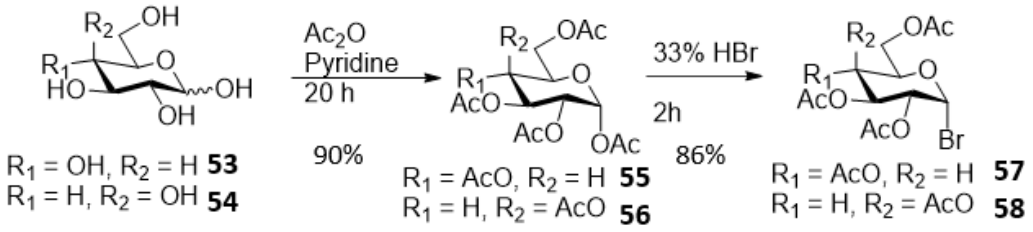


Figure 49**:** LCMS result of the crude product **42**, originated from metal-free white powder form starting material **41**. A reduced *m/z* signal of side product **52** was detected.

Later, during flash column separation, unlike the previous trial, there was no side product **52** eluted from the column, with only the biaryl amide product **42** obtained in a reasonable yield (35%) which suggested that under the same experimental condition, metal remains from the previous Suzuki coupling can have an effect on EDCI coupling. However, purifying the product of Suzuki coupling using a reverse-phase column, even though it contributes to better purity is not recommended, because it can cause significant damage to the expensive reverse-phase columns.

2.2.6 Synthesis of glycosyl donor 57 and 58

In this project, glucose **53** and galactose **54** were selected for glycosylation as we have previously demonstrated that glycosylation increased anticancer activity (Patel, 2011). The glucosyl **57** and galactosyl **58** bromide donors were prepared as described in Scheme 15.

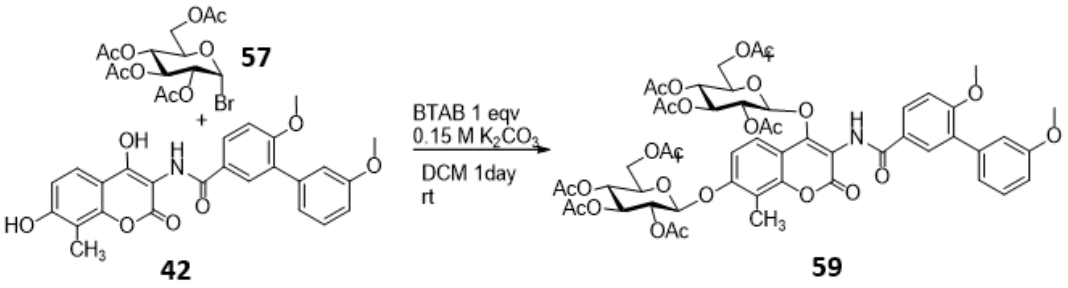


Scheme 15**:** Preparation of Glucosyl donors

The glycosylation method that we had planned to use was that reported by Koenigs-Knorr, a traditional glycosylation method using Lewis acid like $Hg(CN)_2$, Ag_2CO_3 etc. as promoters and sugar-bromide (**57** or **58**) as the donor.

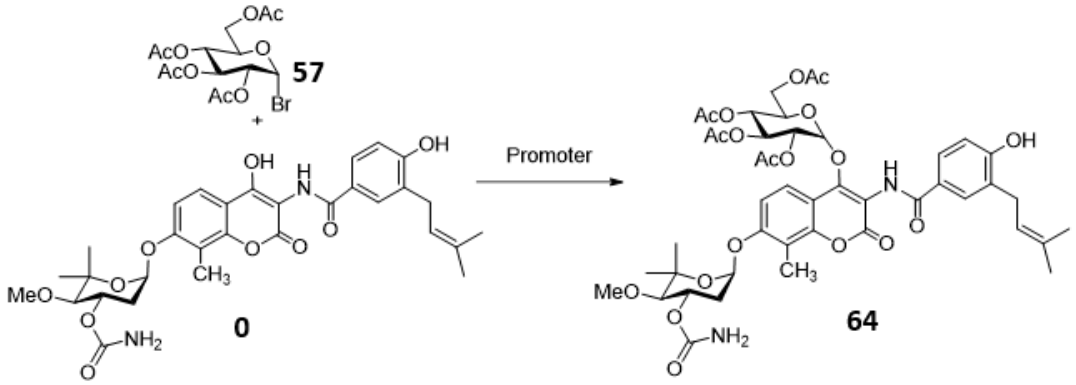
2.2.7 Glycosylation of Biaryl amide led to di-glycosylated product 59

Synthesis of Biaryl-diglucoside (**59**) was completed using a phase transfer reaction, in which the starting material **42** was dissolved in water and sugar bromide donor **57** was dissolved in the organic phase.



Scheme 16**:** Phase transfer glycosylation of biaryl amide **42** to afford the di-glycosylated product **59**.

We did not try any traditional glycosylation promoter because we had carried out a systematic method development in order to find an optimal catalyst for glycosylation of novobiocin scaffold (Scheme 17) and published relevant data (Sun *et al*., 2017), in which we reported that traditional Koenigs-Knorr reaction did not work with our specific compound (novobiocin **0**) with various Lewis acid as the catalyst such as Ag_2CO_3, AgOTf, $Hg(CN)_2$, $HgCl_2$. Other catalysts and conditions were also tested (Table 4).



Scheme 17: Glycosylation of novobiocin with promoter. General scheme. Process optimization data please see Table 4.

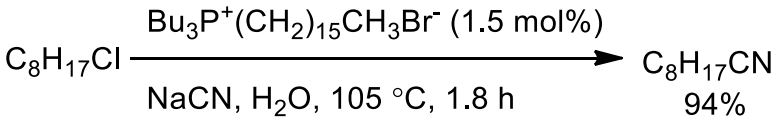
Table 4**:** Promoter screening results of glycosylation of novobiocin (refer to the reaction in Scheme 17).

Entry	Promoter	Solvent	Donor (/eqv)	Temp/° C	Compound 64 Yield
GN-0	$Hg(CN)_2$	THF	1.2 eqv	rt*	52%
GN-1	K_2CO_3	THF	1.2	68	Side product
GN-2	K_2CO_3	THF	1.2 + 4Å molecular sieves	68	Side product
GN-3	K_2CO_3	THF	3.0	68	Side product
GN-4	no	DMF	1.1 + 2.0	rt	11%
GN-5	$AgBF_4$	THF	1.1	rt	No reaction
GN-6	AgOTf	THF	1.1	rt	Traceable product
GN-7	CuOTf	THF	1.1	rt	Side product
GN-9	$AgCO_3$	DMF	1.1	rt	14%
GN-10	$BiONO_3$	DMF	1.2	rt	5%
GN-11	AgOTf	DMF	1.1	rt	No reaction
GN-12	K_2CO_3	DMF	1.1	rt	9%
GN-13	LiBr	DMF	1.1	rt	No reaction
GN-14	No	Propylene carbonate	1.1	rt	15%
GN-15	No	DMF	1.1	60	15%

* rt: room temperature (only in this table). (Sun, de Resende et al. 2017)

It is important to note that DMF as a solvent can act as a weak promoter in the reaction in entry GN-4 (Table 4), resulting in a small amount of glycoside product. Additives can have either promoting (entry GN-9, Table 4) or detrimental (entries GN10-13, Table 4)) effect on the reaction using DMF as the solvent. In general, the Koenigs-Knorr reaction is not ideal for glycosylation of novobiocin, as 4'-OH is not a typical alcoholic hydroxyl, which is not very reactive and sensitive to S_N2 attack (Sun, de Resende et al. 2017).

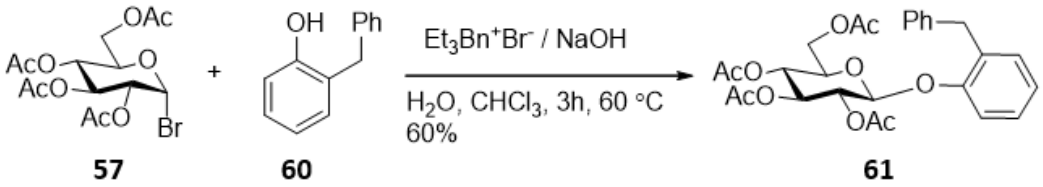
An alternative method toward classic Koenigs-Knorr reaction is phase transfer glycosylation. In 1971, Starks introduced the term phase-transfer catalysis (Scheme 18) to explain the critical role of tetra-alkyl ammonium or phosphonium salts in the reactions between two substances located in different immiscible phases (Starks, 1971).



Scheme 18**:** Phase-transfer reaction discovered by Starks (1971).

The reaction showed in scheme 18 will not happen if there is no presence of quaternary phosphonium bromide. The key element which contributes to this high reactivity enhancement is the generation of quaternary phosphonium cyanide, which renders the cyanide anion organically soluble and sufficiently nucleophilic (Campos-Delgado, Romo-Herrera et al. 2008).

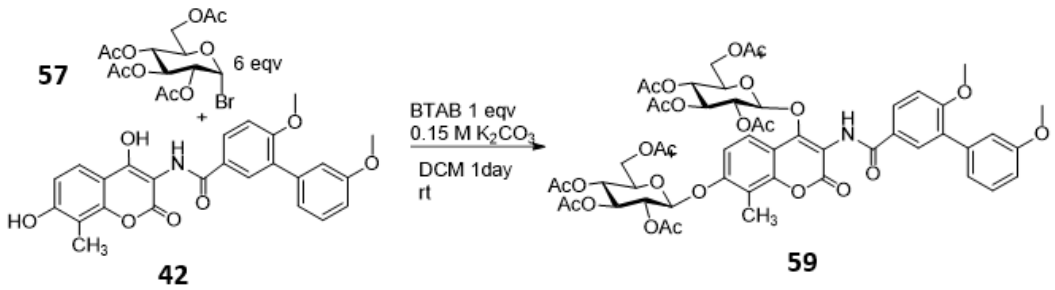
Based on the principle of phase transfer condition (PTC), glycosylation reaction could be applied. PTC glycosylation can be date back to 1984 when the synthesis of aryl glycosides **61** was completed with protected sugar bromide **57** as a donor (Scheme 19) (Harris, Henry et al. 1984).



Scheme 19**:** PTC glycosylation of aryl compounds (Harris et al., 1984).

As an alternative for the Koenigs-Knorr method, PTC glycosylation clearly has several advantages. Firstly, PTC catalysis is metal-free, which is good for the environment and secondly, the method uses easily accessible chemicals. Moreover, the reaction conditions are not strict. For example, it does not require anhydrous conditions and also, it is observed that the glycosylation position can be controlled by adjusting the pH value(Sun, de Resende et al. 2017).

The first trial with the biaryl compound **42** reacted with sugar donor **57** was carried out at room temperature (Scheme 20, entry apt-1), catalyzed by 1 eqv. of phase transfer catalyst benzyl tributyl ammonium bromide (compound **63**, BTAB). After 1-day, organic phase of the reaction was sampled for LCMS analysis (Figure 50)



Scheme 20: Glycosylation of biaryl amide **42** with sugar donor **57** via phase transfer catalysis to form diglycoside **59**. The condition was mentioned as entry apt-1.

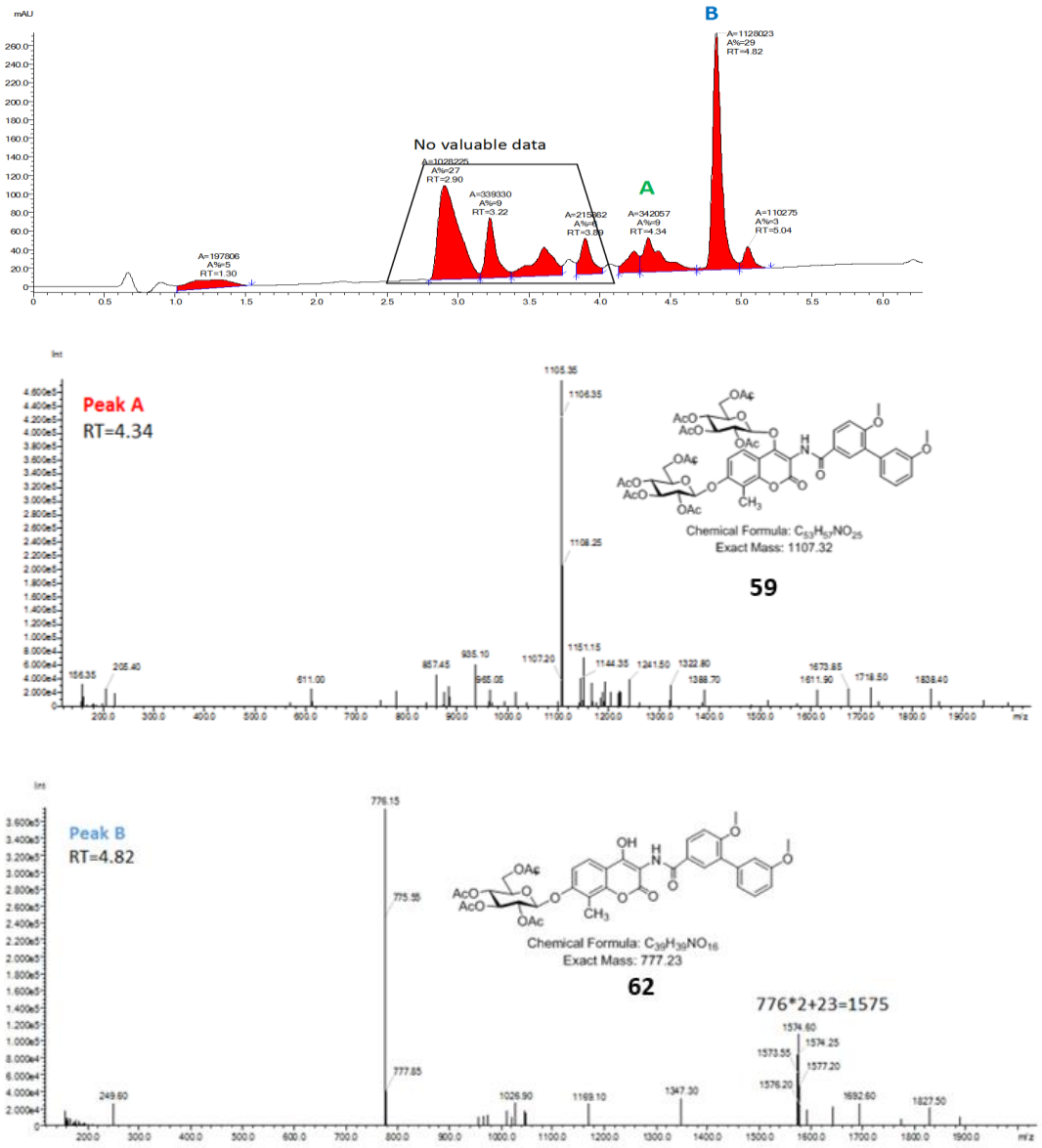


Figure 50: LCMS result of crude product sampled from the organic phase in entry apt-1. MS spectrum that refer to peak A (Rt = 4.34) was identified as diglycoside product **59** due to the matched *m/z* value 1106 found at 1106.

From the data, it can be observed that there is no starting material (**42**) left inside the system, meaning that biaryl amide **42** has an excellent reacting activity. Secondly, according to the MS spectrum, peak A (Rt = 4.34 min) has the mass of diglycoside product **52** (ESI negative mode m/z = 1107-1, found at 1106, Figure 50) although with very low yield. Lastly, it is observed that at room temperature, the major product is mono glycoside for which rt = 4.82 (peak B). However, the glycosylation position remains unknown. A time course study (results see Figure. 50, entry apt-2) was carried out over several days

to investigate how much di-product **52** could form under the reaction condition being tested in entry apt-1, scheme 20.

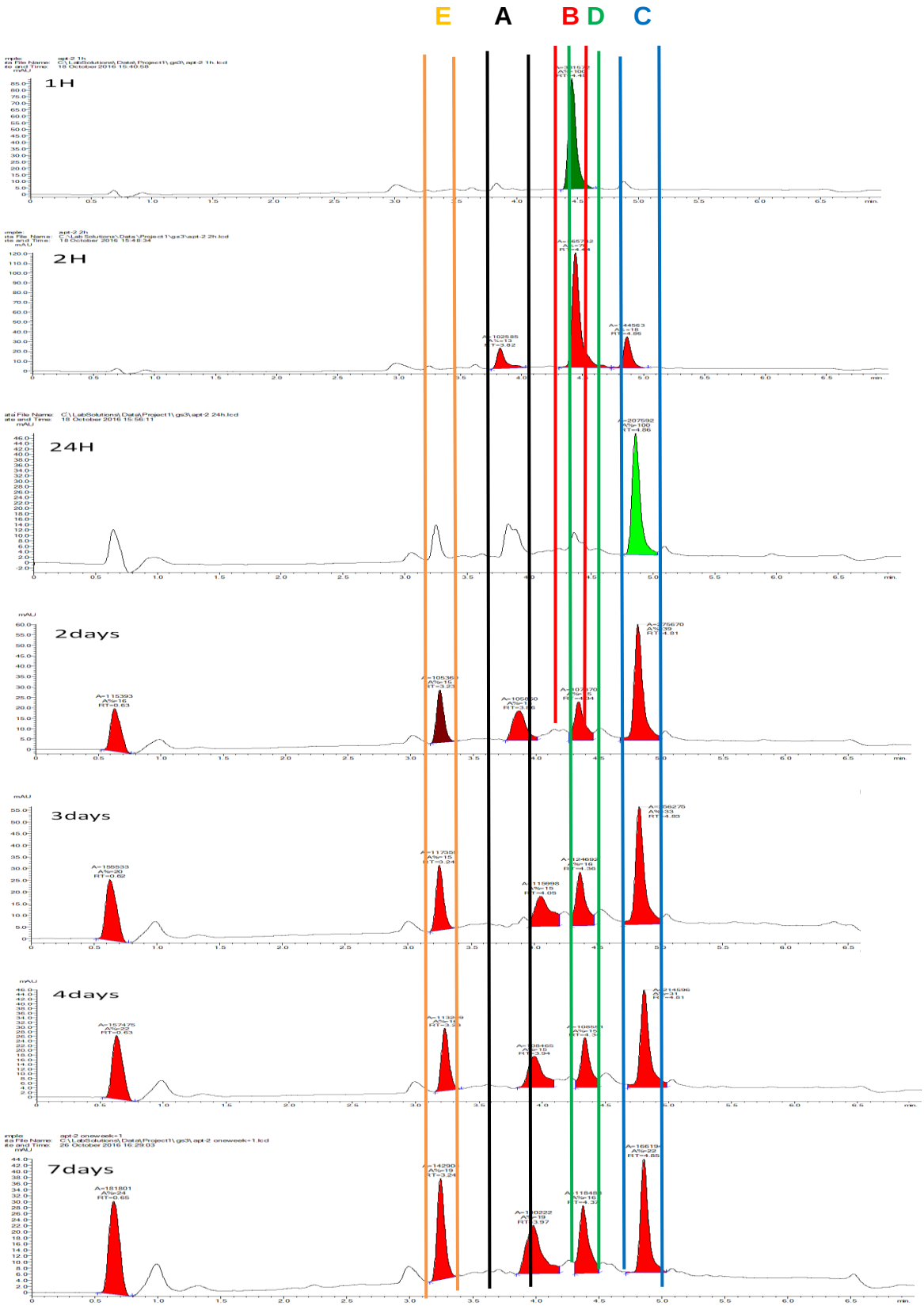


Figure 51**:** Time study of **apt-2**, LC/MS results.

Peak A in figure 51: This was assigned as one of the mono-glycoside products. It is not certain whether the product is a 4' or 7' glycoside. However, according to previous studies (Sun, de Resende et al. 2017) (Simon, Huang et al. 2017), glycosylation on 4 positions tend to reduce retention times on reverse phase HPLC, meaning that in this case, peak A could be the 4-mono-glycoside. However, in general, we demonstrate all monomeric glycoside product as structure **62** in figure 52

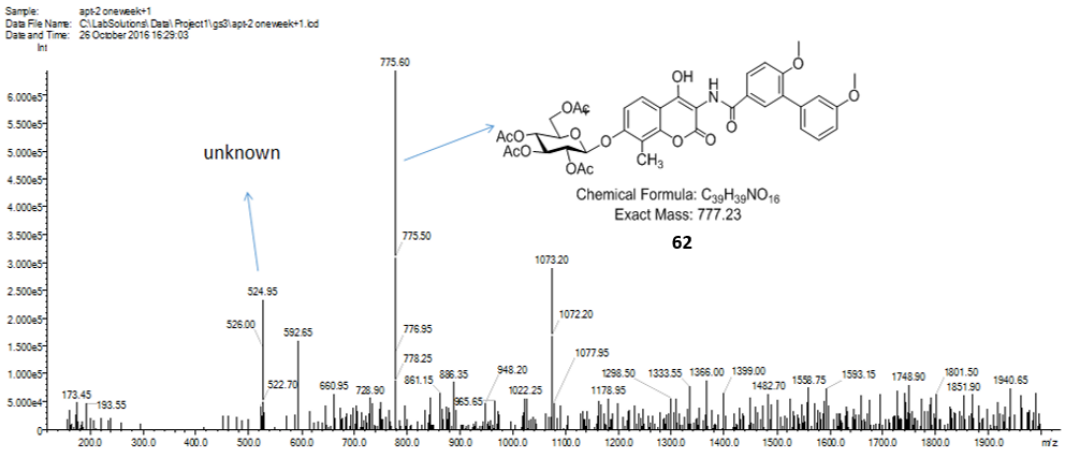


Figure 52: MS result of peak A in figure 51. It was identified as starting material **42** reacted with one molecule of sugar donor **57**. **62** is a proposed structure for this side product.

Peak B in figure 51: Starting material, biaryl amide **42**. Furthermore, starting material **42** peak B and diglycoside product **59** peak D shares a very similar retention time (4.45 vs. 4.37 min**,** respectively). m/z value of **42** disappeared since day 2.

Peak C in figure 51: Another mono-glycoside product. Although it is not possible to confirm whether it is a 4 or 7 glycoside, based on the experience, glycosylation on the 7 position tends to result in increased retention time on reverse phase HPLC. As a result, it is highly possible that peak C is the 7-glycoside product (Figure.53).

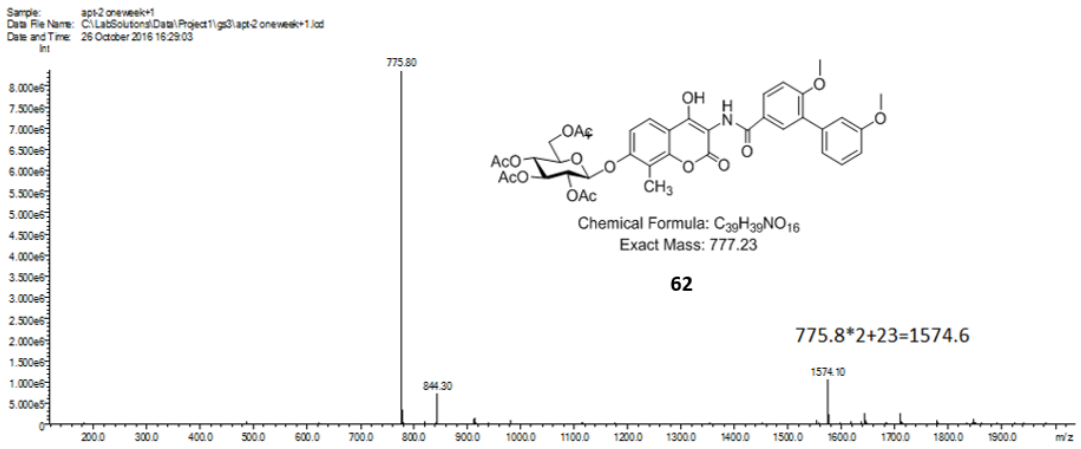


Figure 53**:** MS result of peak C in figure 51. It was identified as starting material **42** reacted with one molecule of sugar donor **57**. **62** is a proposed structure for this side product.

Peak D in figure 51: Diglucosyl product **59** which was already identified in the apt-1 trial. Mass spec ESI^- calculated for 1107 found at 1106 (Figure 54).

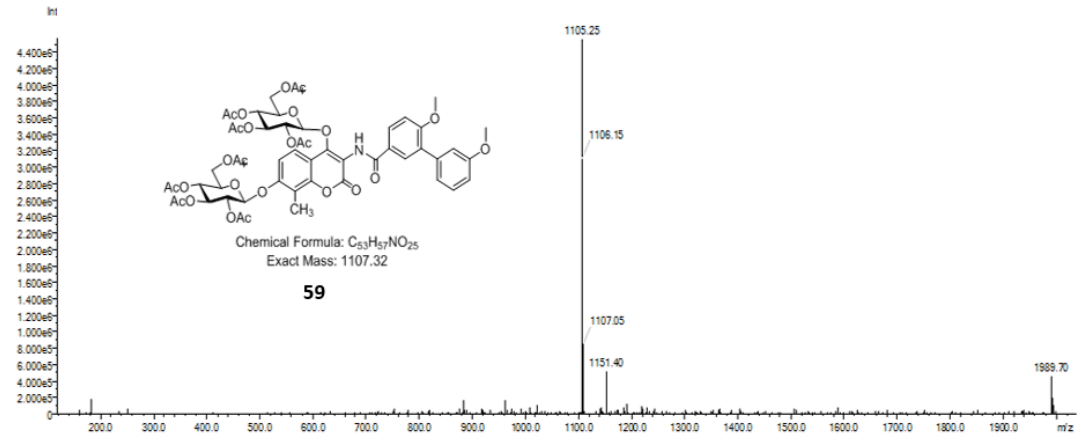


Figure 54: MS result of peak D.

Peak E in figure 51: Phase transfer catalyst fragment **63** (Figure 55).

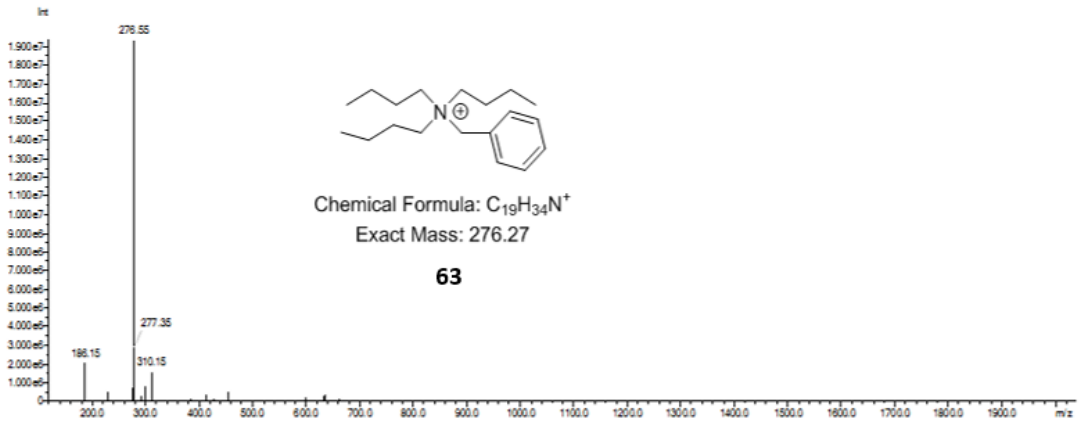
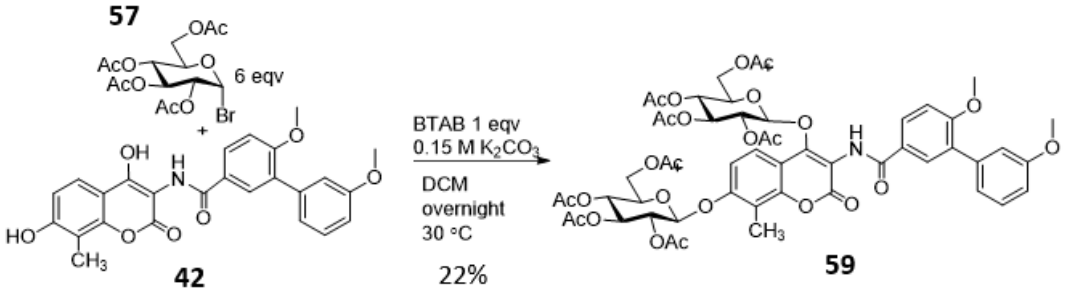


Figure 55: MS result of peak E in figure 51 was characterized as **63** BTAB.

The time course study of entry apt-2 (Figure. 51) showed that indeed starting material biaryl amide **42** (Peak B) has a high activity under phase transfer conditions, as two glycosyl products started to appear within just 2 hours and it takes less than 24 hours for the bi-aryl-amide to be fully consumed. It is also obvious that the mono glycoside product **62** (Peak C) is the favoured product form (31%-39%) under room temperature while another mono glycoside product (Peak A) stays the same after 2 hours. Diglycosyl product **59** (Peak D) started to appear after 24 hours (15%); however, the chromatogram results did not change much since day 1, indicating that the reaction ends after 24 hours with no further mono glycoside **62** being transformed into diglycoside **59**. Therefore, referring to previous experience (Table 4, entry GN-15) to generate more diglycoside product **59**, entry apt-3 was established and applied with a higher temperature from room temperature to 30 °C (Scheme 21).



Scheme 21**:**Synthesis of **59** via phase transfer method with higher temperature (30 °C) applied. This condition is mention as entry apt-3.

After the reaction was stirred overnight at 30 °C, Analysis of the data showed that, it mainly generated diglycoside product **59** and the resulted LCMS of the crude product is shown below (Figure 56).

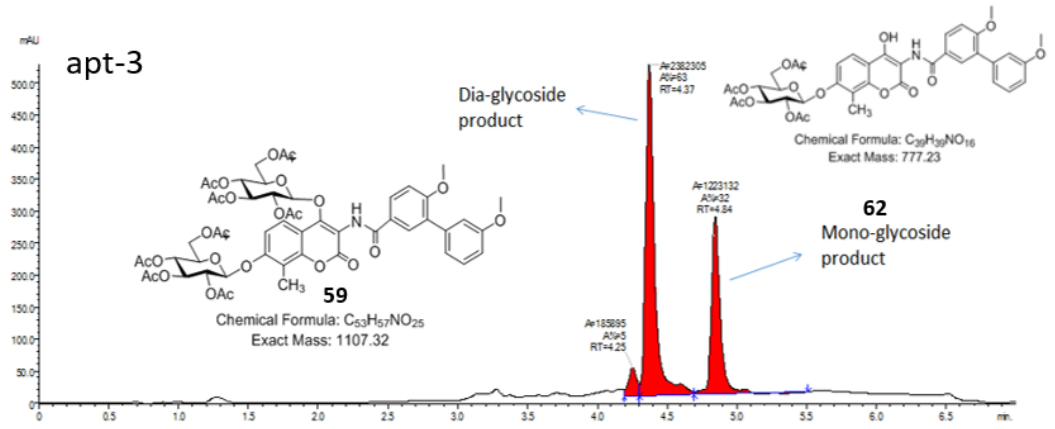


Figure 56: LCMS result of crude product sampled from entry apt-3. Diglycoside product **59** is the dominance product.

Figure 56 showed that when the temperature reached 30°C, the LC yield of diglycoside product **59** dramatically increased from 15% (at 20 °C) to 63%, and **59** became the main dominant form of the product rather than monoglycoside **62**. There was a small amount of mono-glycoside left **62** inside the system.

The organic phase of the reaction was concentrated and purified by reverse-phase column chromatography using automatic Biotage purification using MeOH/Water solvent system, and 165 mg (25%) of the product **59** was recovered. The LCMS data for the purified product **59** is given below (Figure. 57).

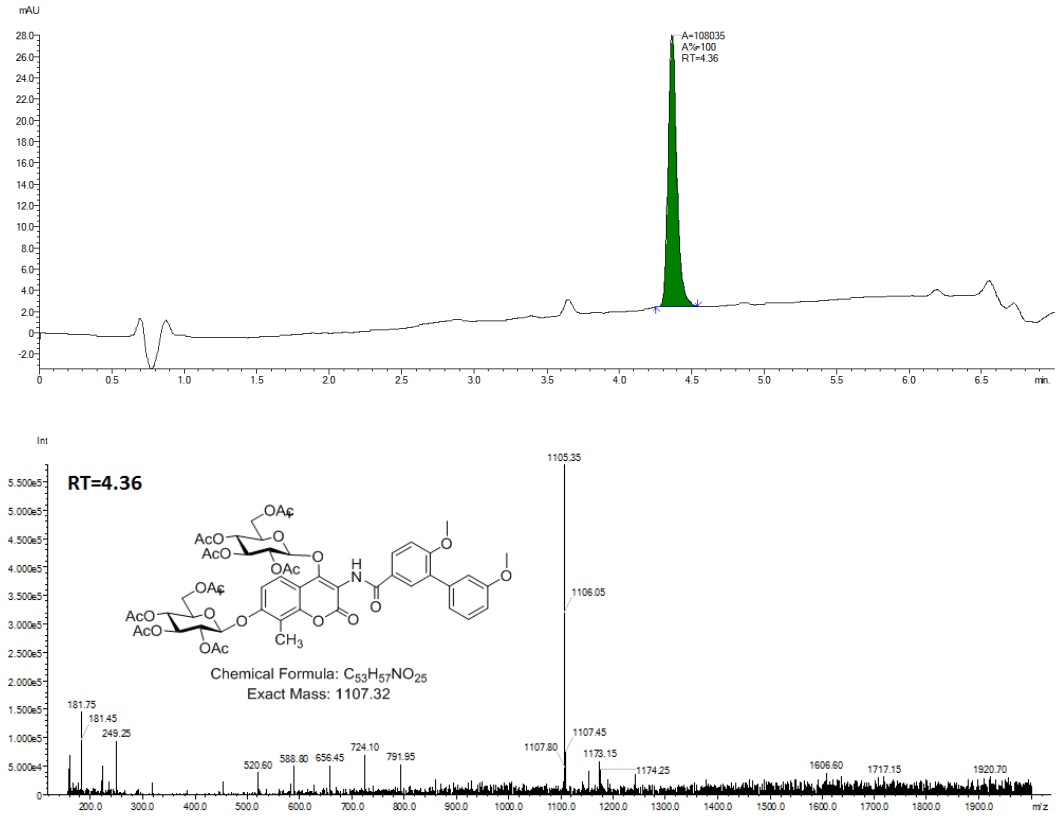
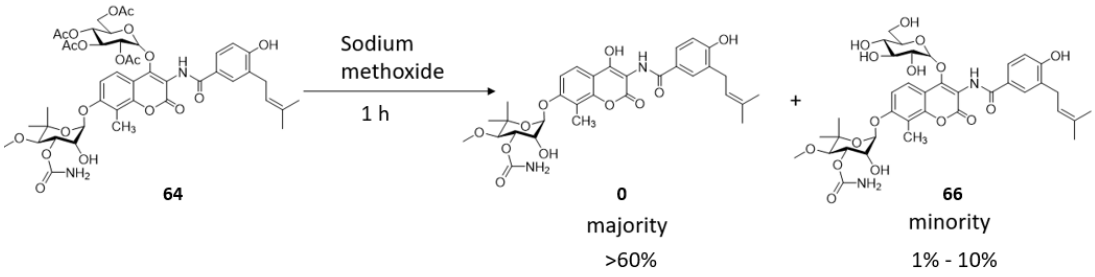


Figure 57**:** LCMS result of purified product **59**.

In conclusion, the Hsp90 C-terminal domain targeted anticancer drug candidates were designed based on the previous experimental data. The main approach was to adopt the modification on the 3-NH_2 that and glycosylation on the 4, or 7'- position, which proved to increase its activity. A synthetic route has been applied, and unexpected challenges have been solved using phase transfer glycosylation using methodology studies. The analytical method has been shown to be an accurate way of following the reaction as it provides detailed information on how modifications can be used to improve the reaction.

2.2.8 Deprotection of glycoside products.

Acetyl protective groups are usually removed by sodium methoxide in dry methanol. For the experimental work, round bottomed flasks were dried at 200 °C in the oven overnight. Relevant glycoside products and 0.1 eqv. of sodium methoxide were mixed and stirred under nitrogen atmosphere in dry methanol for 60 mins. However, through this method, only 4'-mono glycosylated deprotected product **62** could be recovered through reverse-phase column purification in 20% yield. 4',7'-Diglycosylated product **59** were far less sensitive to deprotection and the LC-MS results of such deprotection showed no trace of fully deprotected product. A more challenging problem was that the glycosidic bond at the 4' position was very unstable. The disassociation of 4' glycosidic bond happened at the same time with acetyl deprotection, and the undesired degradation product was found to be the dominant form of the product as shown below (Scheme 22).



Scheme 22: Deprotection of 4'-acetylglucosyl-novobiocin **64** leads to disassociation of sugar moiety, with an unideal yield of **66** range from 1%-10% percentage.

To solve the instability problem, sodium methoxide was replaced with less basic potassium carbonate as the catalyst for deprotection. Nevertheless, the issues discussed above remained with no improvement. Later, we reviewed the literature and found that this is due to the substitution at the 3' position of coumarin. It is reported that for coumarin scaffold, deprotection of 4-glycosides will only go smoothly when the coumarin 3 position remained unsubstituted

(Spero, Ballou et al. 1949), (Ikawa, Stahmann et al. 1944). As in our case, we cannot remove 3'-amide substitution of novobiocin; it is extremely difficult to get the desired deprotection of 4',7'-diglycosylated product. Finally, we decided to stop deprotection at this point, saving protected compounds for bioassays, as these diglycosidic products are considerably hard to synthesize and large quantities of them were used in previous deprotections attempts.

2.2.9 Antiproliferative activity of synthetic glycosidic novobiocin analogues.

Synthetic glycosidic Hsp90 CTD inhibitors were evaluated for their antiproliferative activity. Experimental details are reported in both the experimental sessions (p303) and in the Chapter 3 section **3.3.10.1** (p178). This part of work was carried out using the same experimental as in section **3.3.10.1** (p178). The experimental was finished by collaborator Dr. Clemens Michael Smola from Prof. Andreas Schatzlein's group. The raw MTT data was analyzed and calculated for IC_{50} by Guoxuan Sun from Hilton's group, then reviewed by Prof. Andreas Schatzlein which confirmed that all calculations were convincing. In this section, instead of starting with the aim and design, the data is presented in a conclusion-discussion style for ease of interpretation.

2.2.9.1 Substitution of 4'-OH with 4'-deprotected sugar moiety increase antiproliferative activity

Novobiocin **0**, 4'-acetylglucosyl-novobiocin **64**, and 4'-glucosyl-novobiocin **66** were tested against the MCF-7 breast cancer cell line. Raw data were plotted in Figure 58. Novobiocin **0** exhibited a non-toxic pattern in the 24-h MTT test. It was already reported in our previous work (Patel, Fuente et al. 2011) that 4'-glucosyl-novobiocin **66** has promising anti-proliferative activity; however, the effect of acetyl protection of the sugar moiety remained unknown. From Figure 58, it is obvious that, the protection of glucosyl moiety decreases antiproliferative activities dramatically.

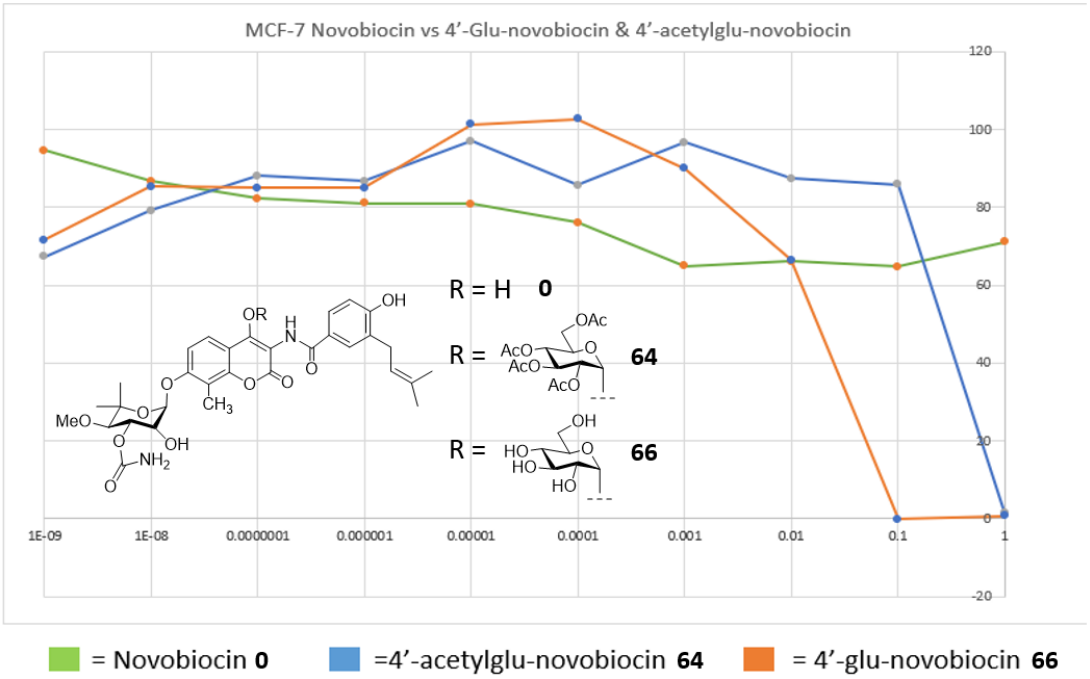


Figure 58: Novobiocin **0**, 4'-substituted glucosyl **66** and 4'-acetylglucosyl novobiocin analogue **64** in MTT assay against MCF-7 cell line, original data plotted. It is a 24h assay. Concentrations are in mM.

We then calculated the IC_{50} of both compounds. The IC_{50} value for deprotected **66** and protected analogues **64** were 11.5 μM (Figure 59) and 702.8 μM (Figure 60), respectively. As a result, we concluded that deprotection of 4'-acetylglucosyl novobiocin **64** increase antiproliferative activity around 61.1-fold. Since novobiocin **0** has no antiproliferative activity in 24h assay against MCF-7 cells, glycosylation of protected and deprotected glucose moiety would increase the activity; however, the deprotected glucose moiety is a more promising lead structure.

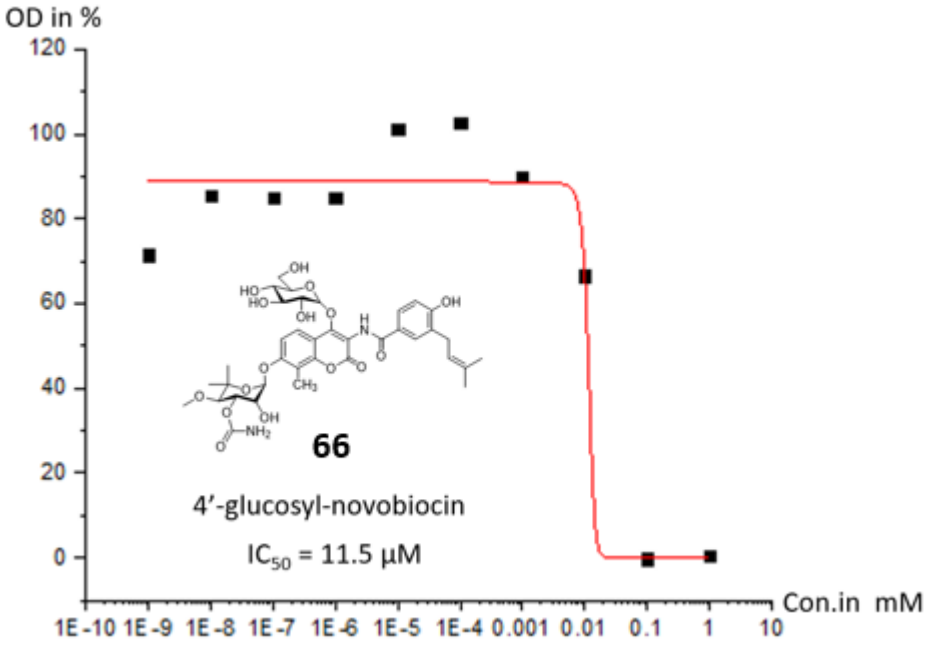


Figure 59: Calculation of IC_{50} of 4'-glu-novobiocin **66**. IC_{50} = 11.5 µM

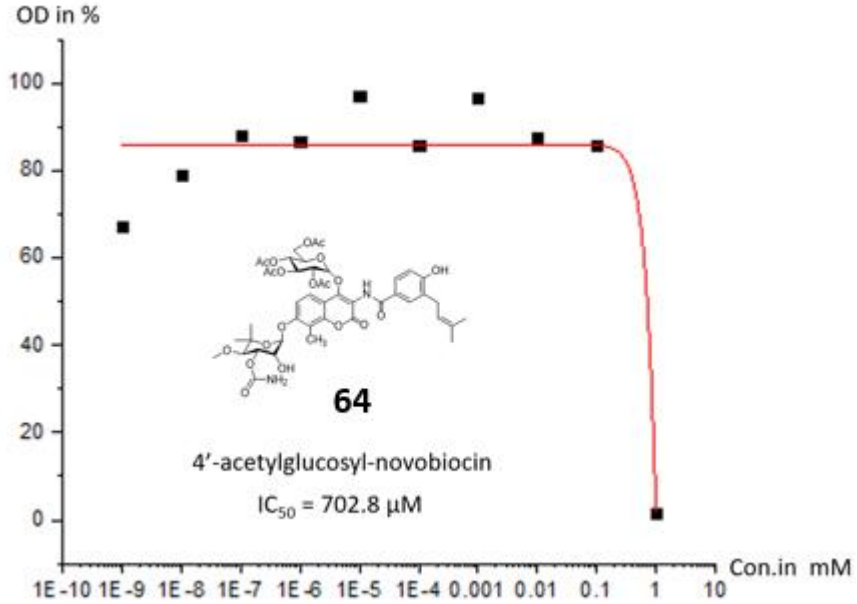


Figure 60: Calculation of IC_{50} of 4'-acetylglucosyl-novobiocin **64**. IC_{50} = 702.8 µM

2.2.9.2 4’-Galactosal novobiocin is less potent than 4’-glucosal novobiocin

Galactose is the isomeric form of glucose with the varied direction of 4-OH on the hexane scaffold. Nevertheless, this tiny modification significantly varies the antiproliferative activity. While 4’-acetylglucosyl-novobiocin **64** showed a mild antiproliferative activity (IC_{50} = 702.8 µM), 4’-acetylgalactosyl-novobiocin **65** showed no anti-cancer activity (Figure. 61). Deprotection of 4’-acetylgalactosyl novobiocin **65** (led to compound **67**) increase the antiproliferative activity, again, demonstrating the importance of deprotection. However, 4’-galactosyl-novobiocin **67** was 8.22-fold (IC_{50} = 97.9 µM, Figure. 62) less potent than 4’-glucosyl-novobiocin **66** (IC_{50} = 11.5 µM, Figure 59), indicated the importance of glucose moiety.

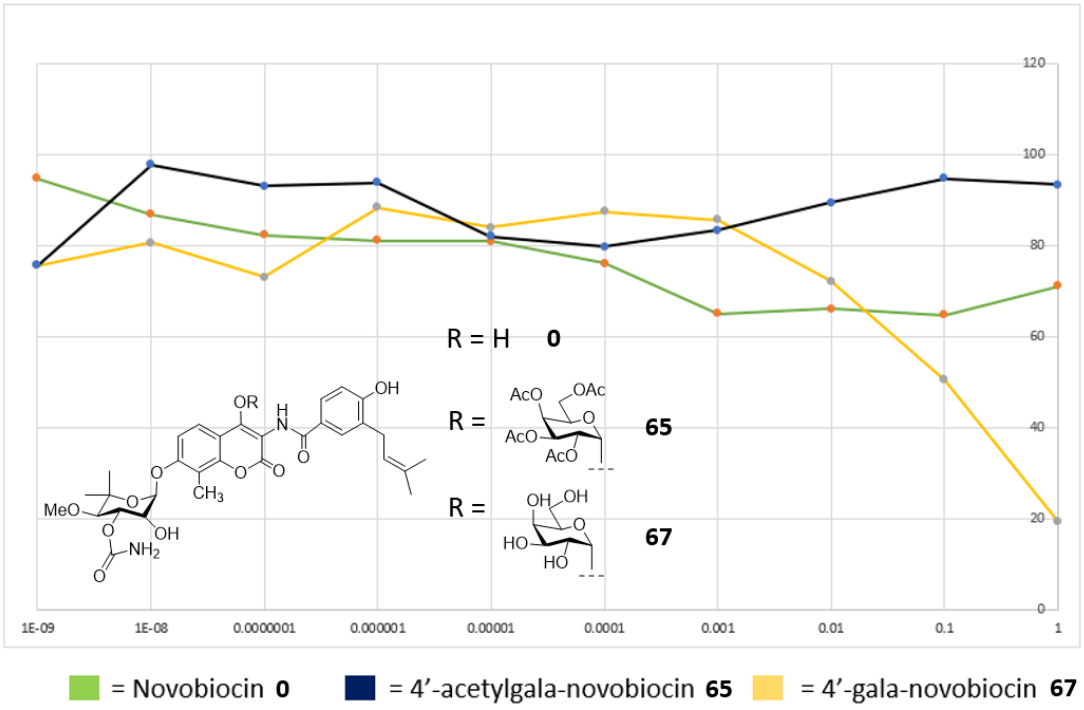


Figure 61: 4’-substituted galactosyl **67** and acetylgalactosyl **65** novobiocin analogue in MTT assay against MCF-7, original data plotted. It is a 24h assay. Concentrations are in mM.

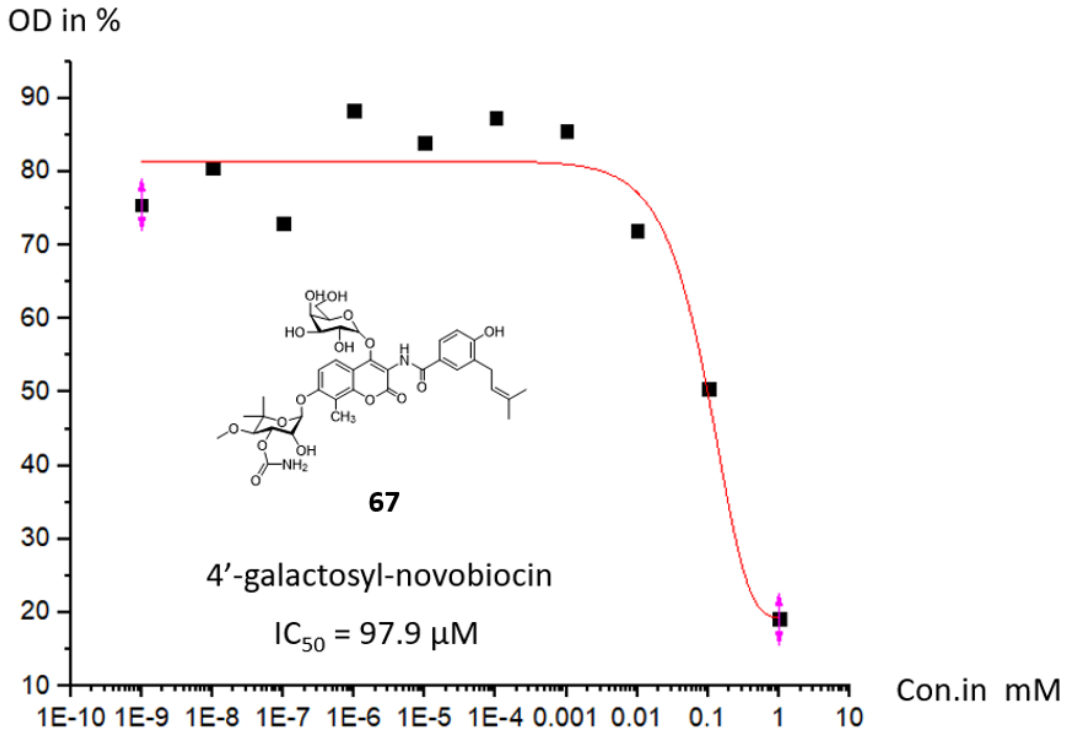
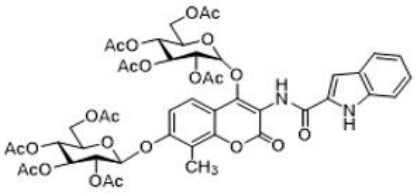


Figure 62: Calculation of IC_{50} of 4'-galactosyl novobiocin **67**. IC_{50} = 97.9 µM

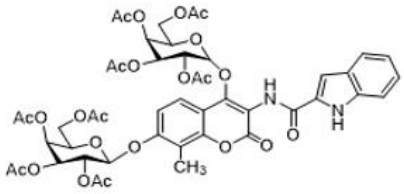
2.2.9.3 4'7'-Acetyldiglucosal substitution is more potent than 4',7'-acetyldigalactosal substitution

4'7'-Diacetylglucosyl-indole amide (**68** in Figure 63) and 4'7'-Diacetylgalactosyl-indole (**69** in Figure 63) amide-type novobiocin analogues were synthesized during a previous MRes project within the Dr. Min Yang's group (Sun, de Resende et al. 2017) using the same phase-transfer catalysis reaction reported in the previous session (p82 Scheme 21). These two compounds were sent for MTT assay; in this body of research and as a result, their antiproliferative activity is reported in this thesis. MCF-7 breast cancer cell line was used for the determination of antiproliferative activity. As discussed before, due to the difficulties of deprotection of 4'7'-diglycosylated product, there were no 4'7'-deprotected novobiocin analogues be tested.



68

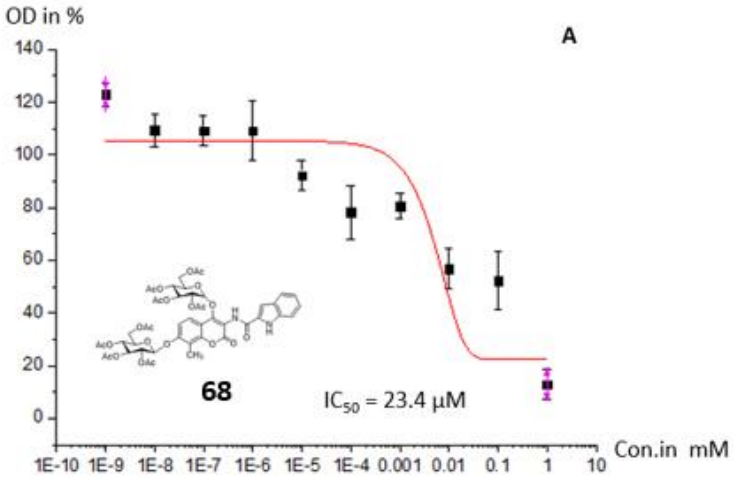
4'7'-diacetylglucosyl-indole amide



69

4'7'-diacetylgalactosyl-indole amide

Figure 63: Structure of Indole amide-type glycosylated novobiocin analogues, which are glucosyl **68** and galactosyl **69.**



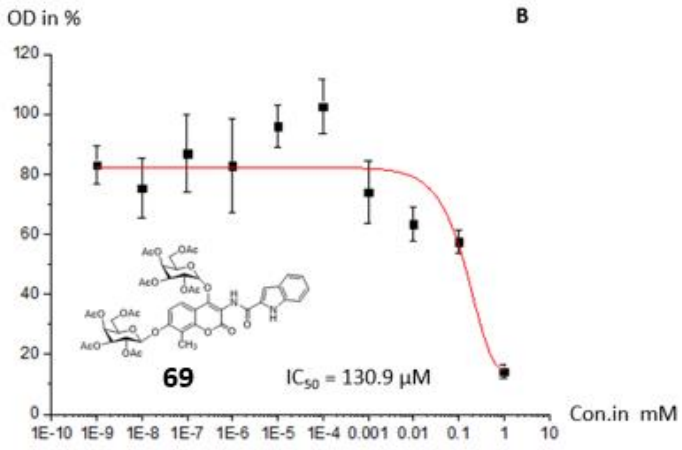
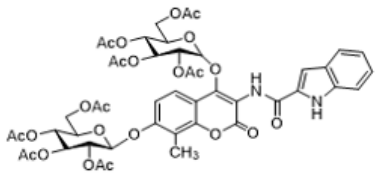


Figure 64: MTT test results of compound **68** and **69** with calculated IC_{50} = 23.4 µM and 130.9 µM, respectively. Against MCF-7 cell line incubated for 24 h.

MTT data of **68** and **69** is shown in Figure 64. Both analogues exhibited promising antiproliferative activity (< 200 µM) compared with the original novobiocin scaffold (non-toxic in 24h against MCF-7). Constituted of the same indole amide moiety, 4'7'-acetylglucosyl substitution analogue **68** exhibited 5.59-fold better antiproliferative activity than 4'7'-acetylgalactosyl analogue 69 (23.4 µM vs 130.9 µM). From this result, we concluded that glucose type substitution is more potent than galactose substitution.

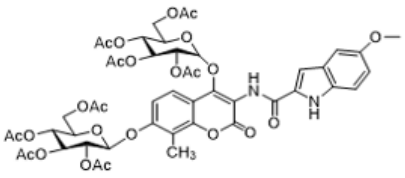
2.2.9.4 3'- Indole amide side chain modification exhibited exceptional promising anticancer activity than other amide side chain modifications.

In the previous section **2.2.7**, we reported the synthesis of a number of amides with side chain 4'7'-diacetylglucosyl modifications. These 3'-amide chain modifications included 2-indole **68**, bi-aryl **59** and 5-methoxy-2-indole **70**. All amide precursor compounds were glycosylated into the corresponding 4'7'-diglycosyl products for MTT testing.



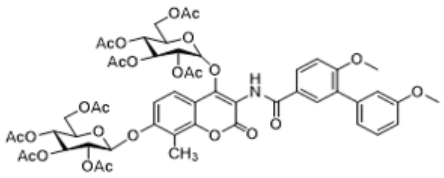
68

4'7'-diacetylglucosyl-indole amide



70

4'7'-diacetylglucosyl-5-methoxy-indole amide



59

4'7'-diacetylglucosyl-bi-aryl amide

Figure 65: Structure of 3'-substituted novobiocin glycosylated analogues.

Surprisingly, MTT assay results showed that compound **59** and **70** exhibited no antiproliferative activity in the 24 h assay against MCF-7 cells, **68** had a potent IC_{50} of 23.4 μM. This indicated that the addition of the methoxy group on the amide side chain would bring negative effects on drug potency. Although such modification was reported to be beneficial in the in literature for 3',7' substituted novobiocin coumarin core (Burlison, Neckers et al. 2006), it was not applicable to the 3',4',7' tri-substituted coumarin core scaffold. The reason for this may be due to the addition of the methoxy group altering the mode of binding in the binding pocket.

2.2.9.5 Conclusions of SAR

In summary, in the Chapter 2 of the project, we designed and synthesized 4'-mono glycosylated analogues (**64, 65**, **66** and **67**) as well as 4'7'-diglycosylated analogues (**59**, **68**, **69** and **70**) and tested these against the MCF-7 cancer cell line to evaluate their antiproliferative activity. The overall SAR results are shown in Figure 66 B and the IC_{50} values are summarized in Figure 66 A.

Glycosylation modification increases antiproliferative activity. For the 4' position, glucose substitution (**64**, **66** and **68**) is more potent than galactose substitution (**65**, **67** and **69**). Acetyl protection of sugar moiety decreases activity, but 4'-acetylglycosyl-novobiocin **65** still maintained a better activity than the original novobiocin **0** moiety. 4'7'-Diglycosylated novobiocin analogue **68** increase activity. For 3'- amide side-chain modification, 2-indole **68** is the optimized substitution for anticancer activity. Methoxy substitution (**59** and **70**) decreases activity.

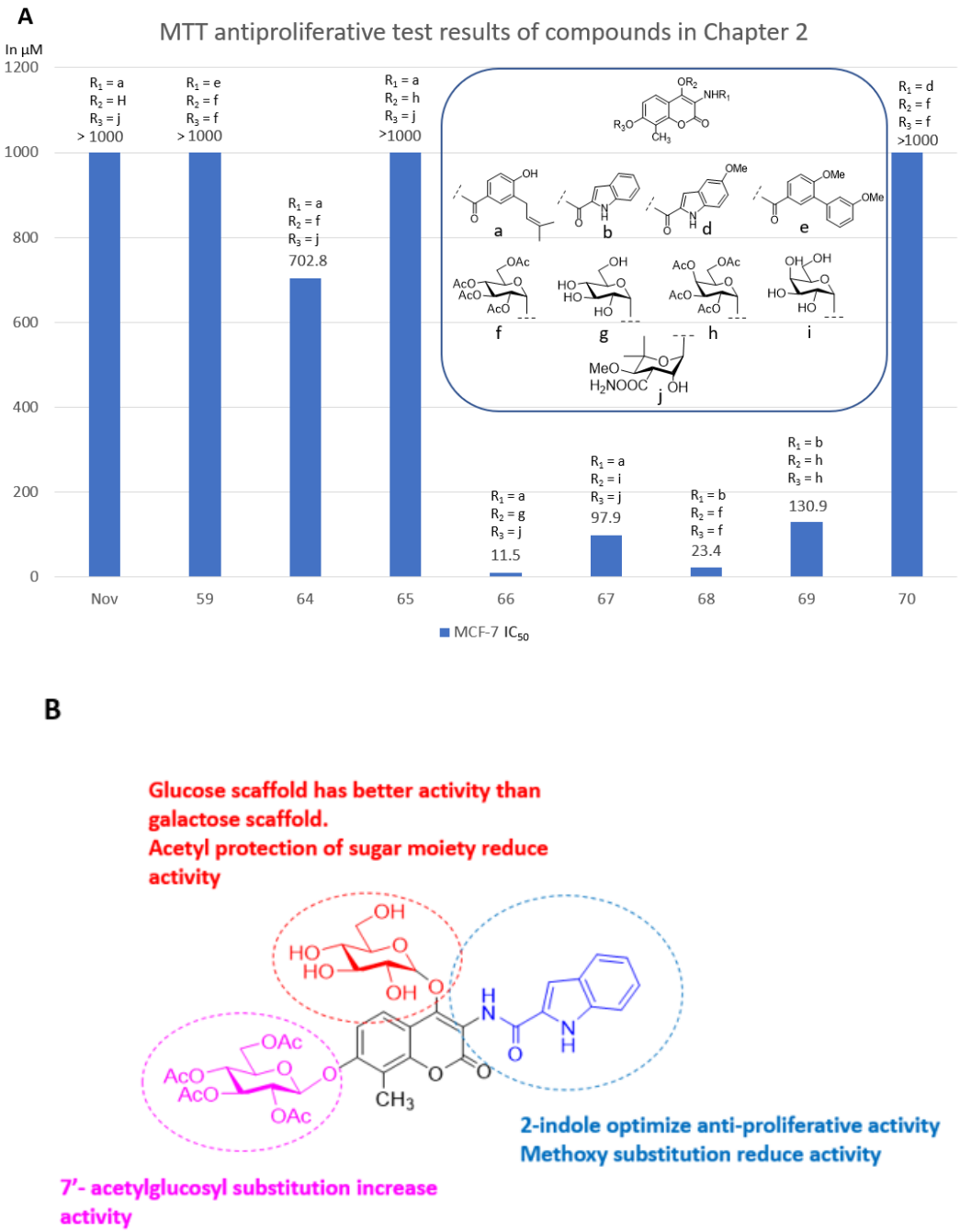


Figure 66: A: Summary of IC_{50} values of compounds tested in this chapter. B: Summary of SAR in this chapter.

From the conclusion, we can propose that if acetyl protected diglycosidic analogue **68** is deprotected, it may exhibit even better activity. However, this is not applicable due to difficulties of deprotection. Therefore, the glycosylation design of novobiocin based anticancer analogues have reached its maximum potential, and an alternative drug design strategy should be applied for better antiproliferative activity.

Chapter 3: Application of targeted covalent drug design strategy to the synthesis of Hsp90 novobiocin based Michael acceptor style anticancer covalent agents.

3.1 Introduction

3.1.1 General Introduction

In the previous chapter, we designed and synthesized several novobiocin glycoside analogues through phase transfer catalysis and tested them for their anti-cancer activity. While the coumarin core 4' substituted and deprotected glycosides showed potent activity, 4'-acetyl protected and 4',7'–acetyl protected glycosylated analogues showed reduced activity. The deprotection methodology was proven to be extremely difficult. Therefore, glycosylation is not an ideal strategy in terms of improving the antiproliferative activity of novobiocin.

Currently, despite several promising lead candidates, there are still no Hsp90 C-terminal inhibitors that have entered clinical trials. Although Hsp90 CTD inhibitors do not have the serious drawback of triggering the cancer-supportive heat shock response caused by administration of Hsp90 NTD inhibitors, the main reason is that CTD inhibitors are far less close to the stage of entering clinical trial due to the fact that there is no convincing drug-protein interaction mechanism that has been proposed. Hsp90 NTD inhibitors has been develop quickly since the Hsp90 NTD was disclosed after being solved via obtention of a high-resolution crystal structure (up to 1.0 – 1.5 Å). As such, any new compound with a good potency can be easily co-crystallized with the protein, and the mode of action can therefore be readily solved. However, for Hsp90 CTD inhibitors, despite the increases in drug potency, there is still no high-resolution crystal structure for human Hsp90 protein families. As a result, there are fewer published references demonstrating any non-in-silico validated binding pocket or any key residue that has been reported to interact with Hsp90 CTD inhibitors. Traditionally, papers that discuss Hsp90 CTD inhibitors only validate drug-Hsp90 CTD binding through the quantitative Western-Blot method, measuring the reduction of Hsp90 client protein such as Akt. This type of target validation only adds "good or bad" knowledge to SAR, but never contributes to protein structure site-directed based drug design. Learning from the story of Hsp90 NTD inhibitors and other successful well-solved anticancer drug target protein (e.g. EGFR), development of Hsp90 CTD inhibitors urgently

needs detailed information about active key residues located at Hsp90 CTD. The solving of Hsp90 CTD active residue will not only contribute to the discovery of potential binding pockets but will also serve as a direct, precise guidance for future drug design.

Cysteine is an amino acid that performs important functions in protein folding, acting as an active reducing residue as well as forming disulfide bridges between peptides. Cysteine residues can be readily inhibited by oxidative agents such as Michael – acceptors. The human body takes advantage of its reducing activity using glutathione to detoxify oxidative stress. Inhibition of cysteine has been an outstanding strategy adopted by anti-cancer drug design. Nowadays, it is possible for a chemical scaffold with a cysteine modifier to be designed to covalently bond to a specific cysteine residue. The resulting covalent bond would be long-lived enough to inhibit the protein irreversibly. This so-called “targeted covalent drug’ has been proved to be a validated approach towards drug design, especially after FDA announced the approval of covalent EGFR inhibitors afatanib, ibrutinib, and osimertinib (Schwartz, Kuzmic et al. 2014).

To develop a targeted covalent drug design approach towards Hsp90 CTD inhibitors, a review of Hsp90 CTD structures was first needed. This aimed to find any reference that provides information on cysteines in human Hsp90. Human Hsp90 alpha protein contains 7 cysteine residues: C374, C420, C481, C529, C572, C597, and C598. The earliest reports of reactive cysteine of Hsp90 date to the year 2000, in which C597/C598 were reported to be able to reduce cytochrome complex but do not reduce either disulfide bonds or produce an NADPH-quinone oxidoreductase activity (Nardai, Sass et al. 2000). Then in 2005 (Martínez-Ruiz, Villanueva et al. 2005), Hsp90 was confirmed to undergo a covalent S-nitrosylation regulated by nitric oxide synthase (eNOS). S-Nitrosylation of Hsp90 resulted in reduced Hsp90 ATPase activity and reduced eNOS activity, which suggested that inhibition of Hsp90 cysteine residue will reduce its activity and may affect subsequent biological pathways. The authors also digested the S-nitrosylated Hsp90 for proteomic mass spec analysis (Martínez-Ruiz, Villanueva et al. 2005). The result showed that C597/598 residues were modified. Then, the follow-up work

from a different research group came out in 2009 (Retzlaff, Stahl et al. 2009). In this research, they concluded that C597/598 regulates Hsp90 as an important switch. Mutation of C597A and C598A simply decreases the ATPase activity, which has the same effect of C597/598 S-nitrosylation. Moreover, in 2012 (Zhao, Yan et al. 2012), another research group reported that nitration of Hsp90 tyrosine induces cell death. 2 of 5 modified tyrosines were located at Hsp90 CTD, and modification of Y603 (this residue located on the same peptide as C597/598) induced cell death. From all these reports, it is obvious to conclude that the chemical flexibility of Cys597/Cys598 is critical to the control of Hsp90 ATPase activity. Inhibition of these two cysteines with chemical measurements will reduce Hsp90 ATPase activity. There could be a potential binding pocket located around that region due to the modification of other residues beside those cysteines that also induces negative biological effects. These kinds of literature overall suggest that Cys597/Cys598 is targetable through drug design.

Since the literature had confirmed two potential targetable residues of Hsp90 CTD, it is important to develop a chemical structure that non-covalently and specifically binds to the Cys/597/Cys598 region. This is usually achieved by in-silico screening using protein modelling and docking techniques where thousands of structures are simulated and calculated for the best results. However, it is not applicable to human Hsp90 due to the lack of high-resolution Hsp90 CTD crystal structures. Any modelling and docking under this situation may not be precise enough to give validated results. Nevertheless, our research group (Dr. Min Yang's group) previously developed a novobiocin based photoaffinity diazirine labelling reagent probe (compound **119**, Scheme 31 on p134) to investigate the binding pocket of 4'-substituted analogues.(Simon, Huang et al. 2017). The probe act as a non-covalent Hsp90 inhibitor under normal conditions. When the probe was incubated with Hsp90 for an adequate period, UV light was applied to the sample, and an extremely active carbene reacts with any C-H, N-H and O-H bonds that are then quickly quenched by water to give a low photoaffinity yield. By far, this project successfully discovered that the probe binds specifically into yeast Hsp90 CTD on residue W585. Due to the fact that the yeast Hsp90 has better resolution

crystal structure 2CG9, docking of novobiocin scaffold towards yeast Hsp90 CTD is potentially applicable. The built-up drug-protein docking model proposed a hydrogen bond between novobiocin 5-OH and W585. Yeast Hsp90 residue W585 refers to human Hsp90 W606 according to sequence alignment. Since both proteomics data and *in-silico* docking data imply a binding pocket around W585 (W606 in human), which is close to A576/A577 (C597/C598 in human), the novobiocin scaffold may potentially non-covalently bind to the C597/C598 region to deliver cysteine modifier groups. As a result, a series of 4'- covalent modifiers substituted novobiocin analogues were designed and synthesized. In this project, they were tested to evaluate the possibility of applying targeted drug design theory toward Hsp90 CTD inhibitor drug design.

3.1.2 Targeted covalent inhibitors (TCI)

The traditional drug design approach mentioned in Chapter 1 has led to the development of numerous successful small molecule inhibitors. Most of them are designed and act based on their ability to interact with their protein target under non-covalent equilibrium binding modes (Figure 67. A). Such non-covalent inhibition pattern is rapid and reverse. Its inhibition power relies on the intrinsic binding affinity between drug and protein (Bauer 2015), causing target aggregation and lost biological activity. As a result, the traditional way to screen a ligand or to optimize drug potency is by changing the steric shape chemical structure or non-covalent interactions (e.g. hydrogen bonds, van der Waals interactions, salt bridge) thus altering its biochemical and bio-physical property for better fitting of the binding pocket (Lonsdale and Ward 2018). However, the reversible mode of action of non-covalent binding was limited by off-target inhibition and drug resistance, caused by the non-specific shape of target binding pocket and steric shape-altering mutations from binding site residues (Jackson, Widen et al. 2017).

With the development of drug design theory, the limitation of non-covalent inhibition was overcome by targeted covalent inhibition. In targeted covalent inhibition, a small molecule is designed not only to bind to a protein through traditional reversible interactions but also to undergo a bond-forming process

with a targetable residue which produces a durable drug-protein linkage (Bauer 2015). The covalent bond that originate from chemical addition can be sufficiently long lasting and persistent that it can be considered irreversible within the half-life of the target protein, resulting in a drug-protein complex that is not subject to classical equilibrium kinetics. Covalent inhibitors that work in this pattern must have a binding pocket with targetable steric and covalent-modifiable amino acid residues. Among all reactive modifiable amino acids, cysteine is the most available and most studied covalent drug target. Cysteine is an underrepresented residue in the protein sequence (3.3% frequency) (King and Jukes 1969) but is disproportionately responsible for different protein functions, with >50% of cysteine residues being solvent-exposed and implicated in an extremely large number of biochemical processes (Requejo, Hurd et al. 2010). According to structural biology, cysteine serves as the reactive nucleophile in many hydrolases (such as cysteine proteases) and can catalyze redox reactions (e.g., protein disulfide isomerase) (K Hallenbeck, M Turner et al. 2017). Oxidized forms of cysteine with sulfenic acid or nitrosothiol functionality are recognized as playing a role in cellular signalling (Leonard, Reddie et al. 2009). Another very important function of cysteine is the formation of a disulfide bond between two cysteine residues, which has been long appreciated as contributing to tertiary protein structure. Disruption of these functions will lead to a decrease of activity or total muting of protein activity. Therefore, these features make cysteine an attractive target for covalent modification by electrophile-composed small molecules. Using cysteine covalent inhibition as an example, the mechanism of which is demonstrated in figure below (Figure. 67 B):

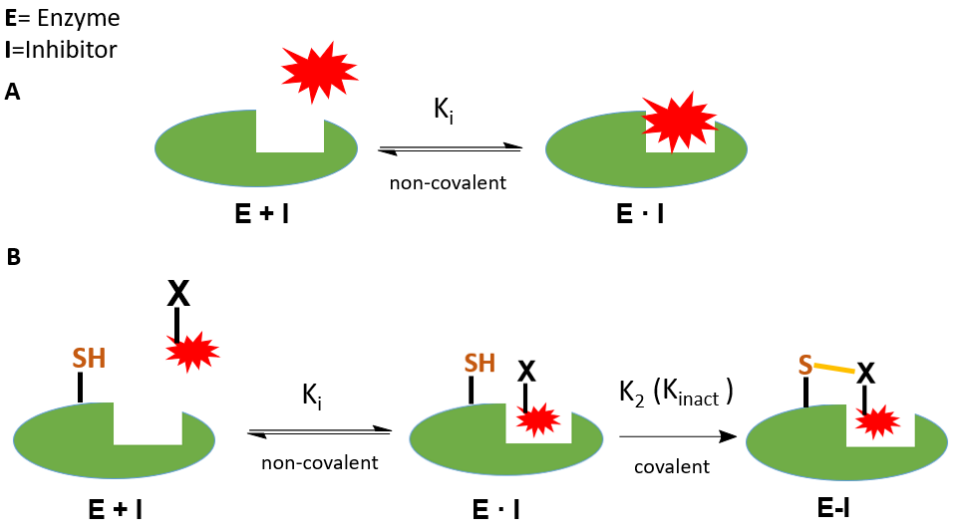


Figure 67: **A**: non-covalent inhibition mode of action. K_i represents its binding affinity. **B**: Covalent binding mode of action. X represents cysteine reactive electrophile chemical groups such as thiol and Michael acceptors.

The first step of covalent inhibition is identical to non-covalent inhibition, where the backbone structure interacts with the protein binding pocket and finally reaches an equilibrium (K_i) between ligand + protein and ligand-protein complex (Bauer 2015). After this process, the electrophile moiety of ligand will react with available nucleophilic residues, e.g. cysteine, and form an irreversible drug-protein complex product (K_{inact}). Therefore, covalent drugs rely on two binding interactions, covalent and non-covalent, that can be independently optimized to obtain the necessary selectivity and potency to be useful chemical probes or drug leads (K Hallenbeck, M Turner et al. 2017). Different from conventional drugs, irreversible inhibitors can theoretically achieve complete neutralization of biomolecular targets if given enough time. Therefore, it is suggested that covalent inhibitors should not easily be ranked using traditional IC_{50} measurements due to its time-dependent nature, and instead require a consideration of the rate of inactivation of a target (K_{inact} in Fig. 66 B). Quantification of covalent drug binding affinity using K_{inact}: Ki ratio is generally preferred over IC_{50} values to evacuate the potency of covalent inhibitors against its target (Bauer 2015).

3.1.3 History and overview of covalent drugs

The very first usage of a small molecule covalent inhibitor to inhibit specific enzymic targets for the treatment of human disease dates back to the starting point of drug discovery 1897 when aspirin **71** was discovered by Bayer as a treatment of inflammation and pain (Figure. 68). However, this covalent drug was not identified by design due to lack of relevant knowledge. It was not known until in the 1970s that aspirin acts by the covalent and irreversible inhibition of cyclooxygenase (COX)-1 and -2, enzymes responsible for the biosynthesis of prostaglandins (Roth, Stanford et al. 1975), through acetylation of a serine located at the protein active site. Another historical famous covalent drug is penicillin **72** (Figure. 68), an antibiotic which its β-lactam moiety covalently binds to the active site serine of bacterial DD-transpeptidase (also known as penicillin-binding protein) to disable the enzymic activities that are used to catalyze a key step in bacterial cell-wall synthesis (Waxman and Strominger 1983).

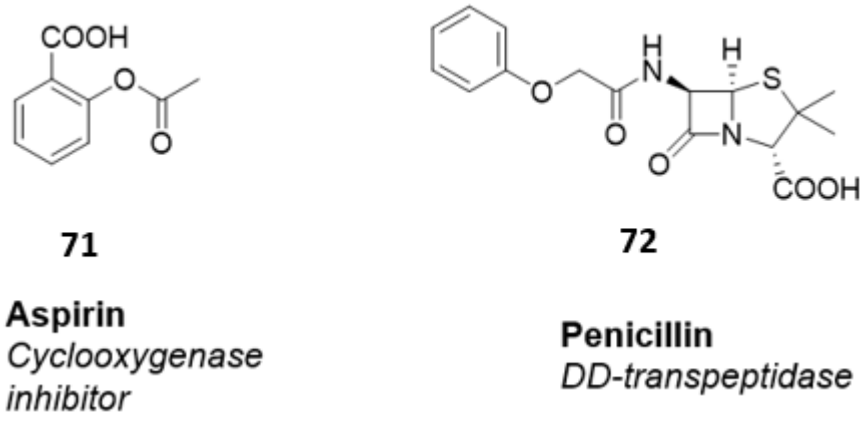


Figure 68: Early covalent inhibitors as drugs.

Later, in the 1980s, proton pump inhibitors (PPI) which are another type of covalent drug (e.g. omeprazole **73**) (Figure. 69), were developed by pharmacological screening while its covalent mode of action was not realised at the outset (Strand, Kim et al. 2017). Omeprazole scaffold does not contain any reactive electrophile group; however, it acts as a prodrug which is locally converted into the proton-pump covalent drug **74** form at the gastric compartment catalyzed by stomach acid. This unique mode of action minimized the exposure of the active form towards other unwanted protentional protein targets, leading to evolutionary success in the treatment

of gastrointestinal reflux disease (Olbe, Carlsson et al. 2003). Similar to PPIs, an antiplatelet prodrug clopidogrel **75** (Figure. 69) was invented without knowing its covalent mode of action originally. Clopidogrel metabolites **76**, which possessed an active thiol is responsible for its irreversible modification towards $P2Y_{12}$ subtype of ADP receptor, causing antiplatelet clinical effect (Savi, Pereillo et al. 2000)

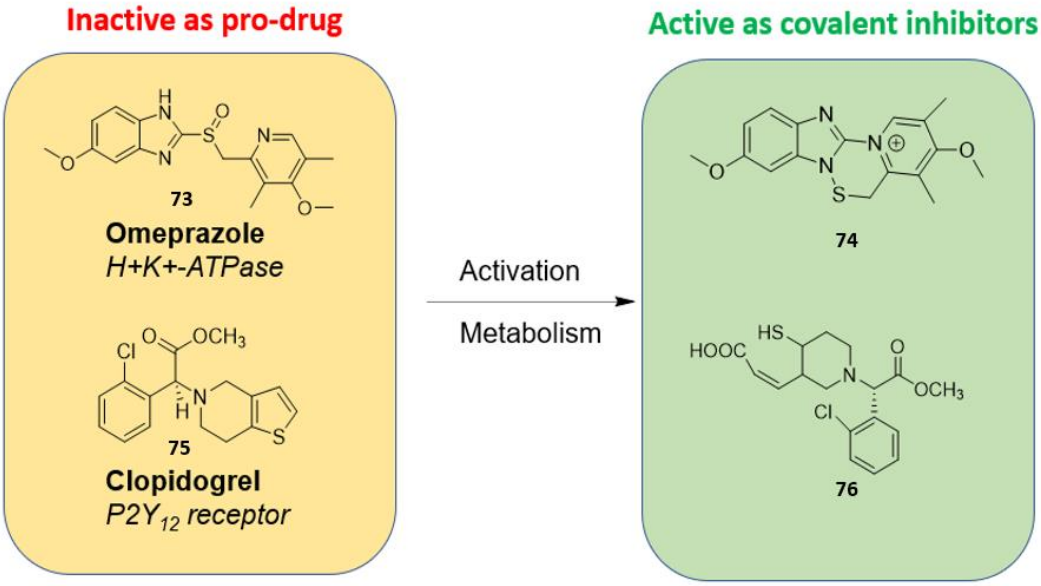


Figure 69: Covalent inhibitors developed in the 1980s and 1990s. PPI inhibitor omeprazole **73** and $P2Y_{12}$ inhibitor clopidogrel **75** require metabolism to become active covalent drug form with reactive electrophiles.

Although covalent type irreversible drugs were a success in the early ages of modern drug discovery, the pace of covalent drug development slowed in recent decades because of the rise of structural-based and target-based drug design (Singh, Petter et al. 2011). After the 1990s, the drug discovery method shifted from biological effect screening – target identification loop to target protein-based drug screening – biological effect validation. More and more structural based toxicity is being considered and covalent modifiers (or covalent metabolites), which may undergo universal non-specific covalent inhibition, were not favoured by drug design in order to reduce off-target toxicity. These electrophiles which act as irreversible covalent modification groups (usually called covalent warhead) are usually epoxide, aziridine, ester, ketone, α, β-unsaturated carbonyl, and nitrile (Bauer 2015). For example, a

classic drug acetaminophen (paracetamol **77**) can be metabolized by cytochrome P450 into highly reactive quinone intermediates *N*-acetyl-*p*-benzoquinone imine **78** (NAPBQI), a primary drug metabolite that non-specifically oxidizes nucleophiles such as cysteines and glutathione presented in proteins for covalent modifications (Jollow, Mitchell et al. 1973). This side effect occurs when a huge dose is administrated, causing unwanted toxicity (Figure. 70).

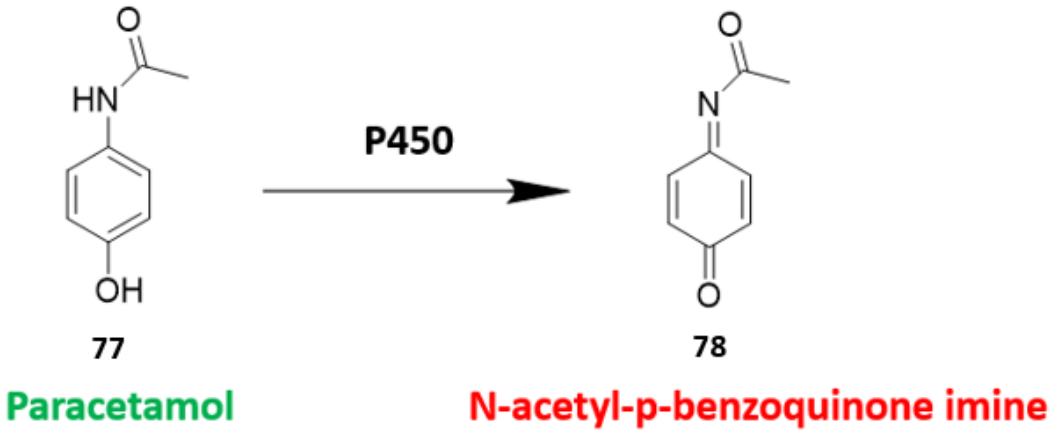


Figure 70: Paracetamol **77** is metabolized to a toxic structure **78** with covalent modification functional groups.

The toxicity of covalent modification attracted serious concerns among drug researchers; however, later on, it was realized that toxicity could be easily muted by proper and target-specific construction of a non-covalent chemical backbone. The viewed toxicity of covalent modifying electrophiles is not necessarily correct when the molecule is designed effectively.

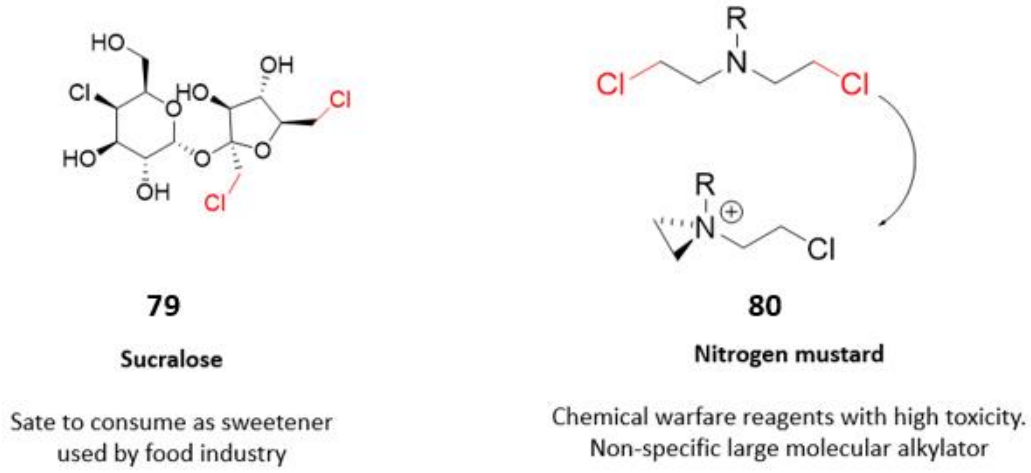


Figure 71: Sucralose **79** and nitrogen mustard **80** both possess of covalent modifying chloride, but with different non-covalent structure, they display opposite toxicity.

For example, sucralose **79** and nitrogen mustards **80** are two chlorinated structures with reactive electrophile chlorines (Figure. 71). However, the different non-covalent backbone fundamentally changes their biological effect. While sucralose has been widely used as a safe sweetener of food despite its high bioavailability (15%) (Baird, Shephard et al. 2000), nitrogen mustard was used as a chemical weapon and has been used for the treatment of late-stage cancers considering its super-reactive chlorides and associated aziridinium salt that results in non-specific alkylation of macromolecules, including DNA and plasma proteins (Thompson and DeCaprio 2013) (Bauer 2015). This example demonstrated that the non-specific binding of a covalent warhead could either be muted or controlled with the assemble of a designed non-covalent scaffold. The toxicity can also be minimized with less reactive or milder electrophiles such as acrylamide which has been widely applied to FDA approved EGFR covalent inhibitors (Ghosh, Samanta et al. 2019). Nowadays, covalent drug design can take advantage of easy-accessible protein-ligand co-crystal structure analysis as well as computer aid drug design approaches for reliable determination of initial structural design which maximizes drug-target binding, then selected scaffolds are synthesized for traditional SAR study. With the approval of the first anticancer covalent EGFR inhibitor afatinib in 2013, the development of covalent inhibitors has become an attractive

research area. As more and more studies about modern covalent inhibitors are published, many unique advantages of covalent inhibitors have emerged.

Covalent inhibition helps to break the limitations of non-covalent drug design. Lots of non-covalent drugs have failed during clinical phase II or III trials due to a lack of selectivity and present unwanted toxicity (Kola and Landis 2004). A solution to this problem at the clinical phase usually focuses on statistic parameters such as a more careful selection of patients (Knowles and Gromo 2003), rather than prospective drug design. This problem was less observed among covalent type inhibitors due to covalent drugs presenting superior biochemical efficiency in many reports, contributing to improved therapeutic margins (Swinney 2004).

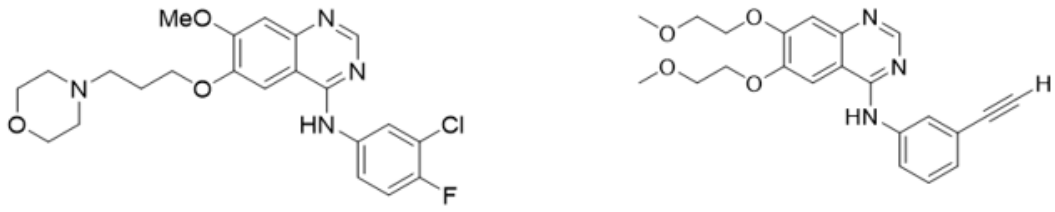
Non-covalent drugs kept shifting between bound and unbound states; when they are at unbound states, endogenous substrates which accumulated during therapy will act as competitors occupying the same binding site as a non-covalent drug (Swinney 2004). Therefore, it is calculated that 80% of approved drugs suffer from this detox mechanism and display reduced activity after long time administration of non-covalent inhibitors. (Swinney 2006). However, covalent type drugs do not suffer from this issue as they permanently block the binding pocket once they are anchored to the target residue, giving no chance for endogenous substrates, e.g. ATP to compete for the same binding site (Barf and Kaptein 2012).

In conclusion, the discovery and development of covalent drugs have a significant effect on human health. Targeted covalent drug design theory is currently being applied to more and more protein targets range from oncology, virology, anti-inflammatory, and other areas.

3.1.4 Oncology successful stories of covalent inhibitors

3.1.4.1 Covalent type EGFR inhibitors

As reviewed in Chapter 1, EGFR belongs to the receptor tyrosine kinase family, which catalyzes protein tyrosine phosphorylation to control signal transduction. EGFR is a cell-surface protein that binds to its original ligand, epidermal growth factor (EGF), to induce tyrosine autophosphorylation and signals cell proliferation (Stamos, Sliwkowski et al. 2002). Mutation of the EGFR gene may subsequently result in the development of non-small-cell lung cancer (Gazdar 2009). Through non-covalent site-directed drug design, the first EGFR ATP binding competitor gefitinib **81** was approved by the FDA in 2003 as the first generation of EGFR targeted inhibitor (Figure. 72). However, it is reported that nearly half of patients treated with gefitinib **81** or another non-covalent EGFR inhibitor erlotinib **82** was found to have a significant drug resistance caused by T790M single point mutation, which decreases the efficacy of drugs by 50% (Schwartz, Kuzmic et al. 2014). This is due to the specific T790M mutation that reduces the binding affinity of both ATP and inhibitors; as a result, an increased dose of inhibitors must be given to reach a reasonable drug response (Ghosh, Samanta et al. 2019).



81 Gefitinib **82** Erlotinib

Figure 72: 1st generation non-covalent EGFR inhibitors, gefitinib **81** and erlotinib **82**.

Development of anti T790M mutation EGFR inhibitors started in the 1990s when Pfizer applied the mild electrophile acrylamide structure to substitute 3-methoxy group, resulted in the first experimental covalent type drug candidate canertib **83** (Figure. 73). Canertib **83** and other experimental covalent

irreversible drug candidate developed during this age (Bauer 2015), for example, PD168393 **84** (Figure. 73) (Fry, Bridges et al. 1998), dacomitinib **85** (Engelman, Zejnullahu et al. 2007) and pelitinib **86** (Yoshimura, Kudoh et al. 2006) failed at clinical trial phase II due to lack of long-term efficacy against T790M mutated EGFR (Ghosh, Samanta et al. 2019). Different from the non-covalent inhibition mechanism adopted by first-generation EGFR inhibitors, these EGFR covalent inhibitors all equipped with acrylamide as a covalent warhead which undergoes Michael addition with the conserved C797 residue in the EGFR active site, forming a covalent bond and increases the drug action time to binding pocket (Engelman, Zejnullahu et al. 2007).

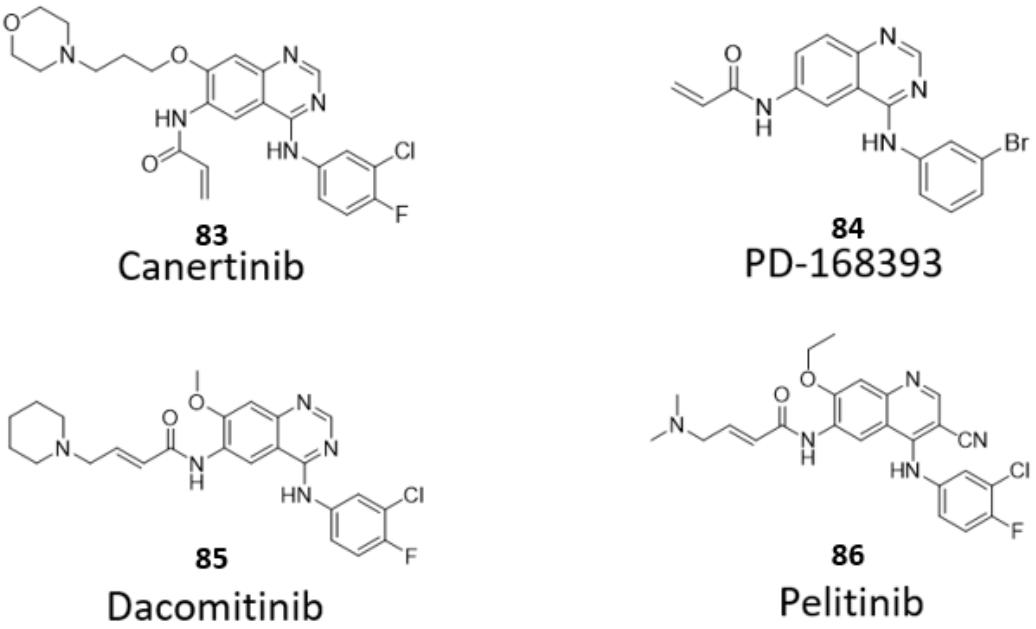


Figure 73: Experimental EGFR T790M covalent drug candidates.

Another well studied EGFR T790M covalent inhibitor is neratinib **87** of which its co-crystal structure with EGFR T790M clearly showed a covalent bond at Cys797 (Figure. 74).

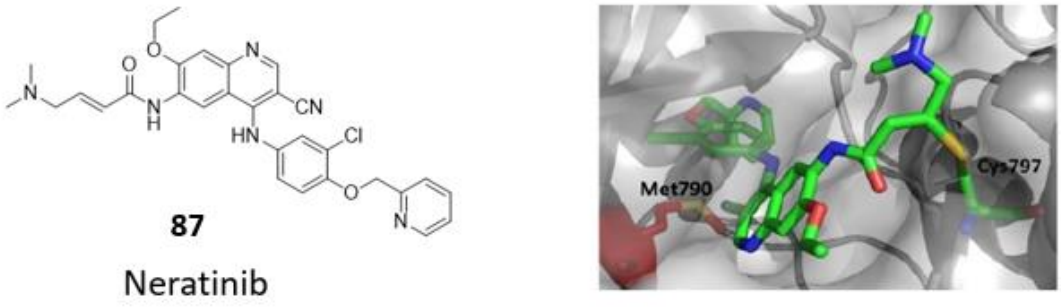


Figure 74: EGFR T790M covalent inhibitor neratinib **87** and its co-crystal structure indicate a covalent modification upon Cys797 (PDB: 2JIV).

Nevertheless, **87** failed the clinical test despite its excellent protein level activity, where both targeted-covalent drug treated patients and untreated patients have poor drug response caused by a diarrhoea-imposed dose limit (Sequist, Besse et al. 2010).

The EGFR T790M mutated drug resistance was finally overcome by modification of gefitinib **81** with the replacement of the 3-morpholino propyl group with a 4-dimethylaminobutenamide to produce covalent EGFR inhibitor afatinib **88**, proved by the FDA in 2013 for treatment of NSCLC. Afatinib **88** (Figure. 75), usually classified as a 2nd generation EGFR inhibitor, that binds covalently to the ATP binding pocket of EGFR and partially hinders its tyrosine kinase activity (Ghosh, Samanta et al. 2019).

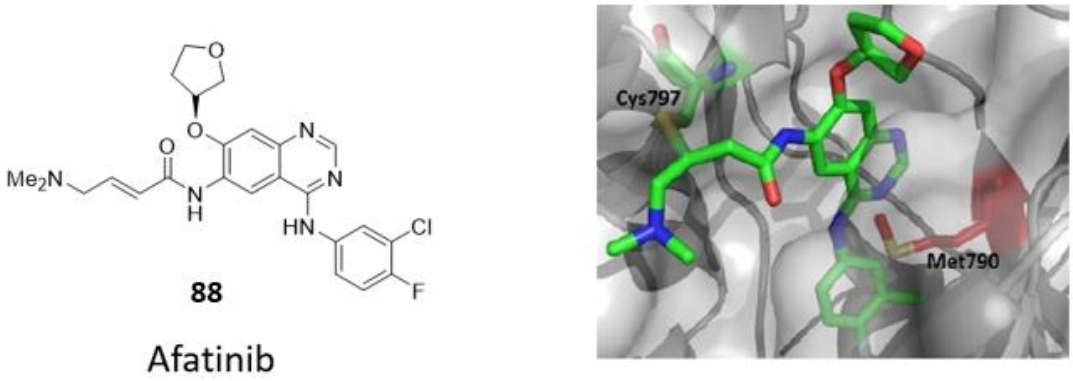
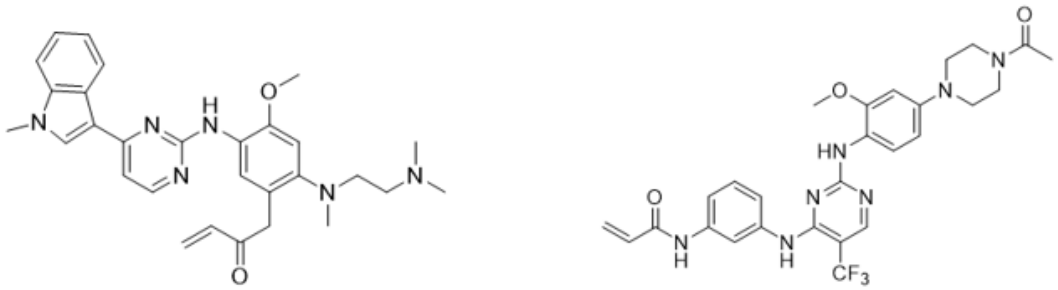


Figure 75: EGFR T790M covalent inhibitor afatinib **88** and its co-crystal structure indicate a covalent modification upon Cys797 (PDB: 2G5P).

X-ray crystallography and *in situ* labelling followed by LC-MS/MS analysis confirmed a covalent bond between afatinib **88** and Cys797 (Figure. 75, PDB:2G5P)

However, it has been reported that the T790M mutation lowers the affinity of the initial non-covalent binding event of afatinib **88** before covalent linkage to Cys797 based on X-ray co-crystal analyses, which may have led to toxicity and lack of efficacy in the clinic (Miller, Hirsh et al. 2012). This led to the development of 3rd generation EGFR covalent inhibitors (Figure 76, for example, osimertinib **89** and rociletinib **90** that were designed to increase the selectivity against T790M only, rather than wild type and the gefitinib **81** targeted type L585R. Osimertinib **89** (Figure. 76, also known as AZD9291, Tagrisso) is an irreversible EGFR inhibitor developed by AstraZeneca, which was approved in November 2015 for the treatment of non-small cell lung cancer in patients with the T790M mutation of EGFR. Osimertinib **89** has an IC_{50} value of 12 nM for inhibition of EGFR with the gefitinib-resisted L858R/T790M mutation and an IC_{50} value of 480 nM for wild-type EGFR (Cross, Ashton et al. 2014). Rociletinib **90** (Figure 76, also known as CO-1686, AVL-301) is another mutant selective covalent inhibitor of L858R/T790M EGFR, with a binding affinity of 21 nM against this mutation and a binding affinity of 303 nM against wild-type EGFR (Walter, Sjin et al. 2013). Although rociletinib **90** showed promising target selectivity and good bioavailability similar to approved drugs afatinib **88** and osimertinib **89**, its clinical trial was terminated due to inconsistency in the published data (Ghosh, Samanta et al. 2019).



89
Osimertinib

90
Rociletinib

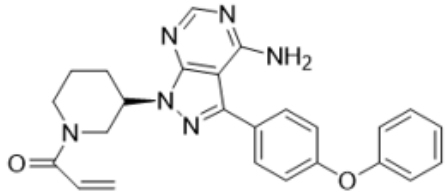
Figure 76: EGFR T790M 3rd generation covalent inhibitor osimertinib 89 and rociletinib **90** specifically target L858R/T790M EGFR.

Similar to previous-generation drugs such as **81**, secondary acquired resistance was reported for osimertinib **89**, usually observed after eight or nine months of treatment (Mok, Wu et al. 2017). It was reported that such resistance is due to mutation of C797S, where covalent linkage formation residue cysteine is mutated into less nucleophilic residue serine, which reduced the activity of covalent EGFR inhibitors of acrylamide type (Thress, Paweletz et al. 2015). The 4th generation EGFR L858R/T790M/C797S targeted covalent drugs are currently under development and early-stage clinical trials.

In conclusion, there is no doubt that EGFR covalent drugs are the most successful applications of targeted covalent inhibition strategy in oncology, resulting in a series of FDA approved drugs bringing considerable profit to the pharmaceutical industry. The story of covalent drug development will inspire the application of covalent drug theory towards more oncoproteins.

3.1.4.2 Covalent type Bruton's tyrosine kinase (BTK) inhibitors

BTK is a non-receptor protein tyrosine kinase that belongs to the Tec (tyrosine protein kinase) family, playing a key role in the maturation of B cells. Due to its specific mode of action, targeting BTK is an effective treatment for B cell lymphoma and leukaemia (Novero, Ravella et al. 2014).



91

Ibrutinib

Figure 77: Structure of BTK covalent inhibitor Ibrutinib **91**.

Ibrutinib **91** (Figure. 77, also called PCI32765), a covalent BTK inhibitor, was first developed by Pharmacyclis LLC and approved by the FDA as an irreversible covalent kinase inhibitor drug. It was approved for the treatment of chronic lymphocytic leukaemia (CLL), mantle cell lymphoma (MCL), and Waldenstrom's macroglobulinemia (Castillo, Treon et al. 2016). Ibrutinib **91** inhibits BTK by disrupting the phosphorylation at its Y223 residue via a relevant binding pocket. The acrylamide moiety of ibrutinib undergoing Michael addition with C481 to form a covalent linkage in the BTK kinase domain, which locks BTK in an allosteric inhibitory state (PDB: 3GEN). Despite the fact that ibrutinib **91** was successful in the treatment of B cell lymphoma and leukaemia, it led to side effects such as bleeding, diarrhoea, rash, and atrial fibrillation due to off-target inhibition of EGFR and other Tec family proteins which also possess a similar cysteine residue at the active site (Fabbro, Smith et al. 2015) (Li, Zuo et al. 2014). Resistance to ibrutinib **91** was also observed during chronic treatment (Furman, Cheng et al. 2014).

In order to overcome the side effects of ibrutinib **91**, several second-generation BTK covalent inhibitors have been developed (Wu, Liu et al. 2016). Among these inhibitors, acalabrutinib **92** (Figure. 78, also called ACP-196), exhibits promising high inhibition potency, rapid oral absorption, a short half-life, and

reduced binding to EGFR as well as other Tec family proteins (Byrd, Harrington et al. 2016). A combination of acalabrutinib **92** with other antibodies, such as obinutuzumab is a promising therapy for the treatment of B cell malignancies (Iragavarapu, Mustafa et al. 2015). Another example is CHMFL-BTK-11 **93**, a covalent BTK inhibitor that suppresses the activation of B cells effectively and also blocks the secretion of various cytokines like IgG1, IgG2, and IL-6 (Wu, Huang et al. 2017). Rest of the second-generation covalent BTK inhibitors such as CC-292 **94** (Evans, Aslanian et al. 2013), BGB-3111 **95** (Wu, Liu et al. 2016), PRN1008 **96** (Smith, Krishnarajah et al. 2017) ONO-4059 **97** (Yasuhiro, Sawada et al. 2017) have also been developed and undergoing early clinical trials.

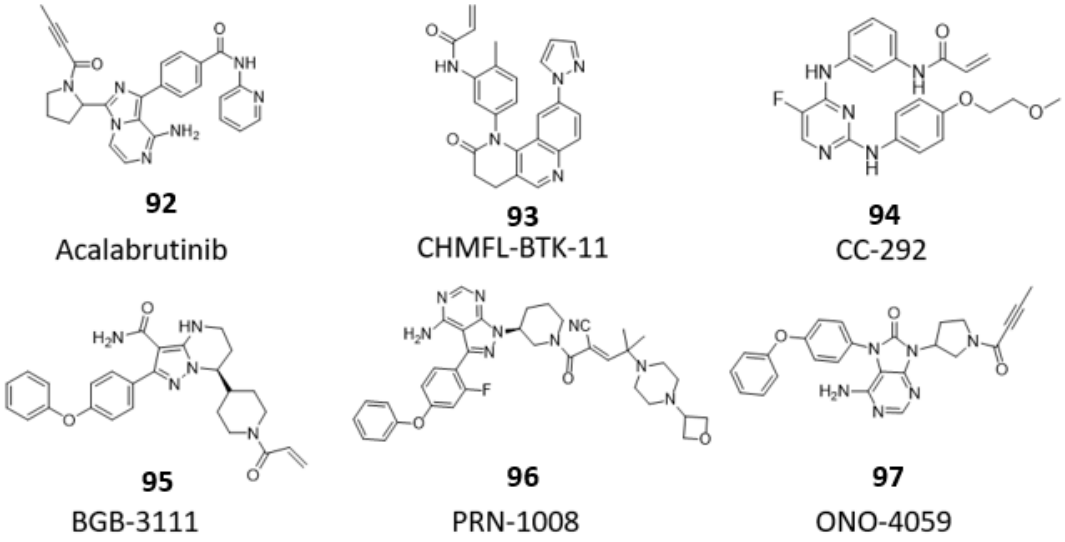


Figure 78: Structures of 2nd generation of covalent BTK inhibitors.

3.1.5 Hsp90 CTD reactive cysteine

As reviewed in chapter 1, Hsp90 CTD is responsible for the dimerization of inactive Hsp90 monomer into a biological active dimer, and Hsp90 is naturally observed in a predominately dimeric form. Removal of Hsp90 CTD will cause the loss of dimerization ability. (Krishna, Reddy et al. 1997) In 1995, Nemoto and co-workers mutated Hsp90 alpha human to discover any critical residues responsible for dimerization. They reported that residues 533-732 (Uniport code: P07900) in the CTD were sufficient to form dimers. They also reported that the same residues on each monomer accounted for its dimerization. The binding site of Hsp90 CTD is symmetrical. (Nemoto, Ohara‐Nemoto et al. 1995) Later, it was confirmed that all Hsp90 homologues from bacterial to human share a similar mechanism of dimerization. Since then, lots of efforts have been made to explore the specific function of CTD residues.

Cysteine is an amino residue that is usually undergoing bioactivity determined post-translational modifications, including oxidation, disulfide bond formation, sulfenic, palmitoylation, *N*-acetylation and *S*-nitrosylation. (Chung, Wang et al. 2013). Interestingly, cysteine residues are only conserved in higher animals, missing in bacteria and fungi (Retzlaff, Stahl et al. 2009). Cysteine oxidation of Hsp90 is a sign of a response to oxidative stress, thus causing the proteasome-mediated degradation of several known Hsp90 client proteins, including Akt, cyclin D1, Cdk4 Raf-1, and mutant p53 (Mollapour and Neckers 2012) & (Chen, Chang et al. 2008). The investigation of Hsp90 active cysteines was firstly carried out by Gabor and co-workers, in which they reported human Hsp90 and Hsp90 fragment peptides containing Cys597/598 are able to reduce cytochrome c. These two cysteines can neither reduce disulfide bonds of insulin nor possesses an NADPH: quinone oxidoreductase activity (Nardai, Sass et al. 2000). As a result, Cys597/Cys598 were identified as two reactive cysteines located at Hsp90 CTD. Later, in 2005, Ruiz and co-workers reported that the S-nitrosylation of Hsp90 leads to reduced Hsp90 ATPase and endothelial nitro oxide synthase regulatory activities. Endothelial nitro oxide synthase (eNOS) is a co-client protein of Hsp90. In this research, eNOS

activity was reported to be regulated by S-nitrosylation of a specific Hsp90 residue Cys597 (Figure.79) (Martínez-Ruiz, Villanueva et al. 2005). Hsp90 ATPase activity and its positive effect on eNOS activity are both inhibited by S-nitrosylation, indicating that modification of Cys597 will result in an inhibition effect towards Hsp90.

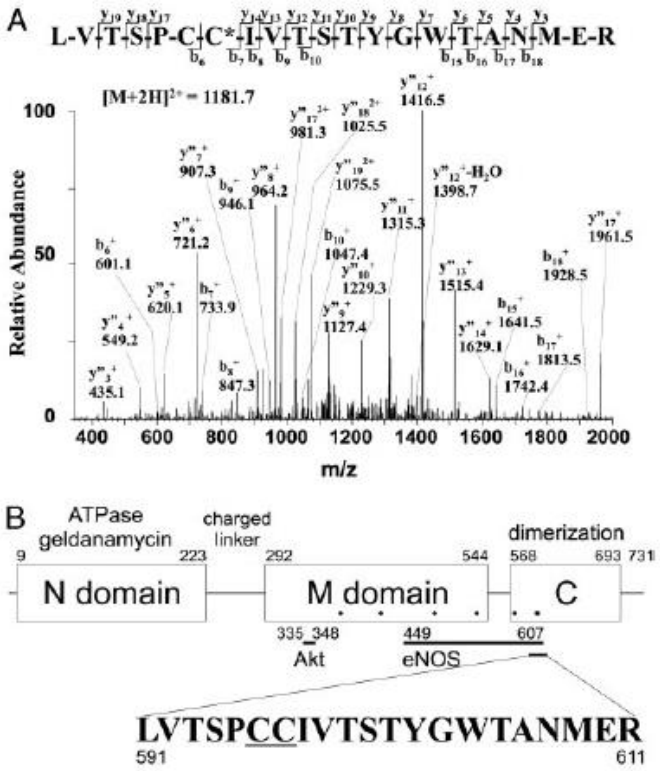


Figure 79: **A**: Proteomic analysis confirmed Cys 597 is S-nitrosylated.

B: Location of Hsp90 active cysteines, they are modified by NO. (Martínez-Ruiz, Villanueva et al. 2005)

The modification of Cys597 was further investigated by Retzlaff and co-workers in 2009. They confirmed that the negative effect of S-nitrosylation of Cys597 is not limited to the decreasing of N-terminal based ATPase activity but also Hsp90 CTD dimerization, which acts as a switch point.

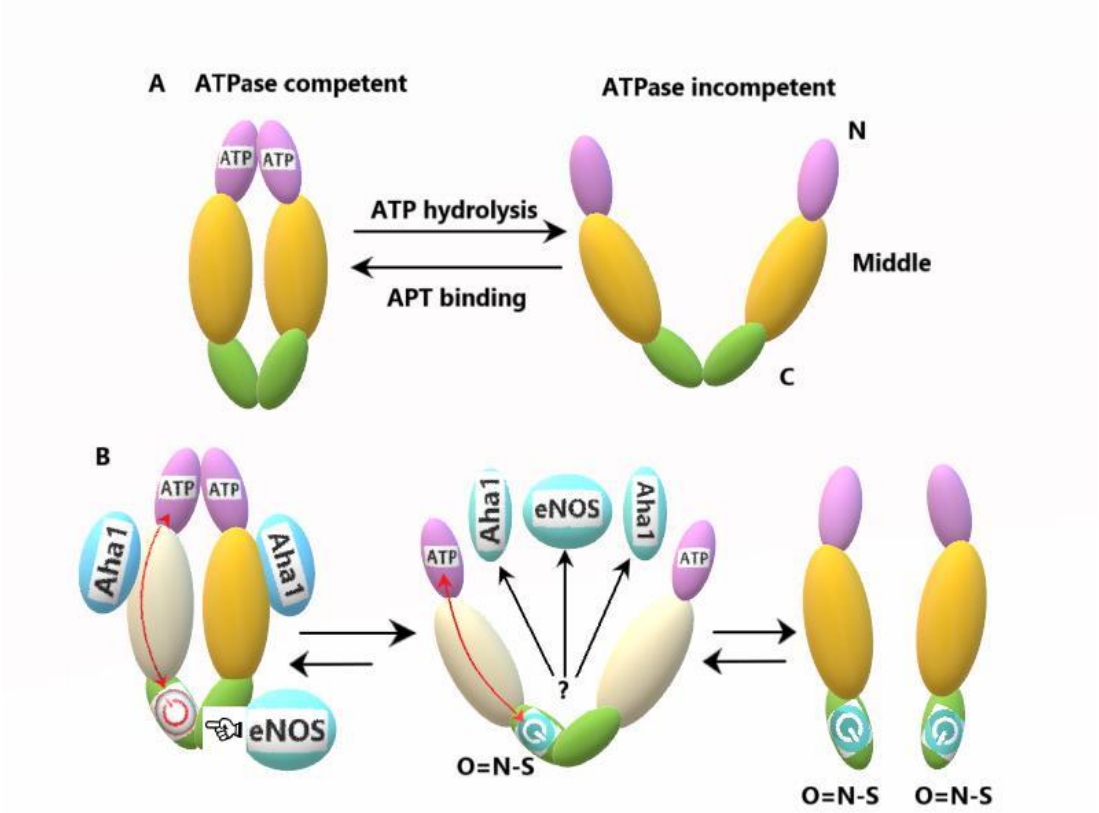


Figure 80: Hsp90 conformational dynamics. (**A)** Hsp90 cycles between ATPase-competent and ATPase-incompetent states. ATP binding to the amino-terminal (N) domains of Hsp90 stabilizes their transient association, which is essential for subsequent ATP hydrolysis. Hsp90 is constitutively dimerized through motifs in the carboxy-terminal (**C**) domain of each protomer. (**B)** ATP binding to the N-terminal domain propagates a conformational signal to the C-terminal domain concurrent with the acquisition of the ATPase-competent conformation. ATPase activity is enhanced, and Hsp90 cycling increased by the binding of the co-chaperone Aha1 to the middle domain of Hsp90. *S*-nitrosylation of Cys 597 in the C-terminal domain disrupts ATP-induced signal propagation from N-terminal to C-terminal domains (and vice versa), inhibits ATPase activity by preventing the necessary conformational change and Aha1-mediated stimulation, and weakens C-terminal association of Hsp90 protomers. (Retzlaff, Stahl et al. 2009)

S-nitrosylation of Cys597 is a reversible process controlled by covalent modification of NO, which is released by eNOS. The figure above (Figure.80) showed how NO regulates the activity of Hsp90. Partial *S*-nitrosylation of

Hsp90 would help Hsp90 to maintain an open conformation, generating steric space allowing the co-client protein to bind or to release. This modification works by increasing the energy barrier necessary for conformational changes. Mutation of human Hsp90 alpha C597A decrease ATPase activity, and mutation of yeast Hsp82 A577C (yeast A577 in sequence alignment is human Hsp90 C597) increase ATPase activity, which indicates the importance of NO-modifiable cysteine residue in Hsp90 in higher eukaryotes allows the fine-tuning of Hsp90 activity by reversible post-translational modification. (Retzlaff, Stahl et al. 2009) Total *S*-nitrosylation of Hsp90 monomer will also prevent its dimerization, therefore muting the molecular chaperone function of Hsp90. As a single modification of Cys597 allows the efficient and fast regulation of the whole Hsp90 catalytic circle, it works like a switch that controls its biological machinery.

Both NO and Hsp90 CTD inhibitors work in the same pattern, which inhibits Hsp90 CTD and results in the disassociation of Hsp90 dimer into monomer. As inhibition of Hsp90 can be achieved through covalent modification of Cys597 by NO, a similar inhibition effect should be achieved using covalent Hsp90 inhibitors, as long as the cysteine covalent modifier moiety is correctly delivered. Literature in this section provides the theoretical basis of the Hsp90 CTD covalent drug design. By turning the reversible cysteine modification into irreversible modification, the ATPase function of Hsp90 is expected to reduce and results in the degradation of multiple Hsp90 client proteins, thus, presenting anti-proliferative activities.

To find out a scaffold that is suitable for Hsp90 Cys597 inhibition. We took advantage of our previously published paper, in which photoaffinity label novobiocin was synthesized (Simon, Huang et al. 2017). Tandem MS/MS proteomic analysis confirmed that this molecule modified W585 of yeast Hsp82 (Huang 2018). *In silico* docking of 4'-glucosyl-novobiocin with yeast Hsp82 2CG9 revealed a binding pocket where 4'-glucosyl moiety has direct interactions with G584, G581, and F583 while the hydroxy group on the amide side interact with W585 (Figure. 81) (Huang 2018).

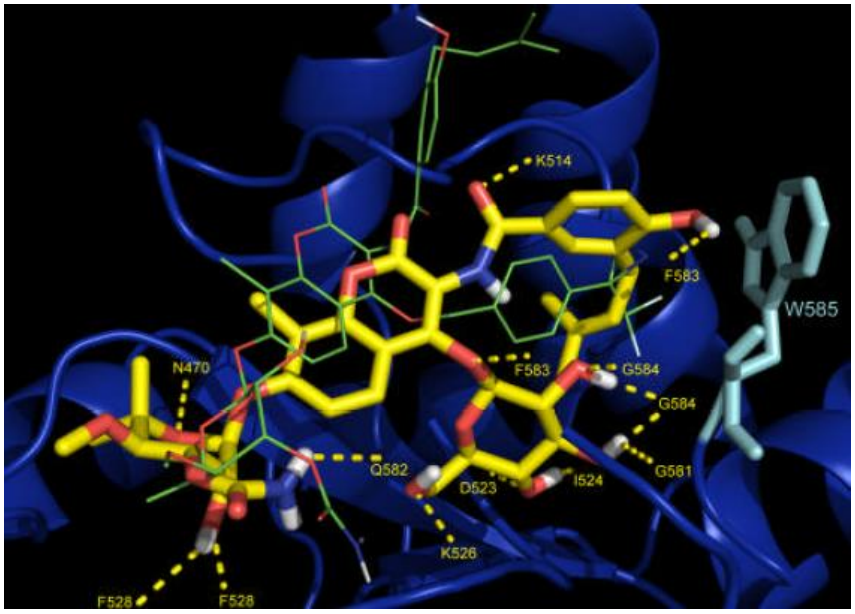


Figure 81: 4'-glucosyl-novobiocin revealed a binding pocket near W585 of yeast Hsp82.

Due to the fact that W585 residue is only 8 residues away from A577 and the structure is a continuous beta-sheet-loop, 4'-substituents of novobiocin may potentially interact with the human Hsp90 C597 region (equal to yeast A577 in sequence alignment), and the unmutated human Hsp90 W606 (equal to yeast W585) will still interact with novobiocin amide side chain.

3.1.6 Natural Hsp90 covalent modifiers

Although there is no report of synthetic Hsp90 covalent drugs, some natural compounds composed of active electrophiles were reported as Hsp90 covalent modifiers. These covalent ligands modify more than just cysteines, other residues of which mainly lysine, are also reported as target residue. Covalent modification of Hsp90 will result in a conformational shift and variation of ATPase activity and thus cause antiproliferative activity or anti-inflammatory activity (Pellati and Rastelli 2016).

3.1.6.1 Natural products that modify Hsp90 lysines

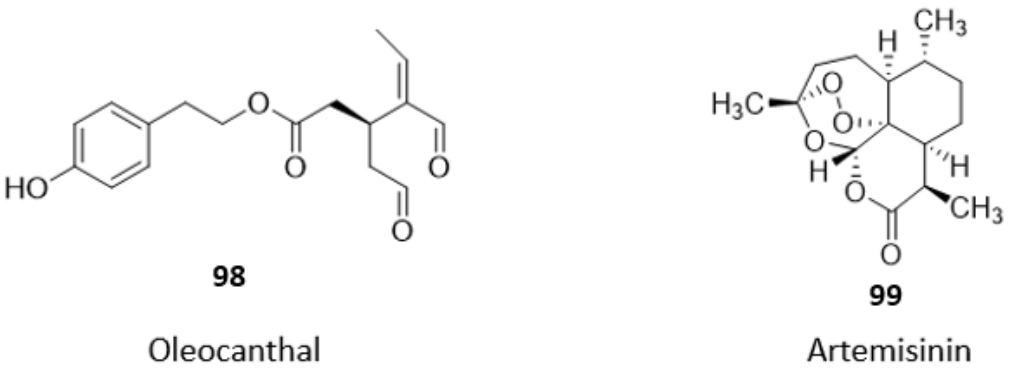


Figure 82: Oleocanthal **98** and Artemisinin **99** are two kinds of Hsp90 lysine covalent modifiers.

Oleocanthal **98** (Figure. 82) is a phenolic compound rich in extra virgin olive oil reported to interfere with different pathways of human disease such as cancer, Alzheimer's, and inflammation (Margarucci, Monti et al. 2013). Comprehensive identification of its interactome found this molecule significantly inhibits human Hsp90 ATPase activity with a dose-dependent pattern, similar to Hsp90 NTD inhibitors radicicol. In silico modelling successfully docked the molecule into Hsp90 NTD ATP binding pocket, which also suggested a potential covalent binding between its aldehyde group and Hsp90 NTD residues (Margarucci, Monti et al. 2013). MS/MS proteomic analysis confirmed that oleocanthal **98** modifies Hsp90 NTD Lys112 and Lys58 covalently. Unlike radicicol, which its administration would cause unwanted overexpression of Hsp70 and Hsp27 (Heat shock response, reviewed in chapter 1 intro), oleocanthal **98** will not induce such side effects

and was reported to reduce the amount of bio-active oligomeric form Hsp90 (Margarucci, Monti et al. 2013). This suggests that although both radicicol and oleocanthal binds to the same Hsp90 NTD ATP binding pocket, their effect towards biological pathways are varied. Oleocanthal showed promising cytotoxic effects against HeLa and U937 cancer cell lines.

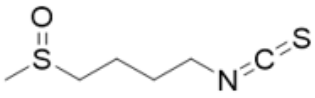
Another reported Hsp90 lysine modifier is artemisinin **99** (Figure. 82), an anti-malaria drug discovered by Chinese scientist Youyou Tu who won the 2015 Nobel prize in medicine. Despite its well-characterized anti-malaria mechanism through its unusual natural existed peroxide scaffold, its application as an anticancer, antiviral and anti-inflammation agent is currently under investigation (Konstat-Korzenny, Ascencio-Aragón et al. 2018). One study in an attempt to find the direct protein target of artemisinin **99** responsible for anti-inflammation using a quantitative chemical proteomics approach reported that this molecule directly targets Hsp90 (Wu, Cheng et al. 2019). Proteomic analysis and molecular docking confirmed that artemisinin **99** covalently binds to Hsp90 NTD Lys58. The anti-inflammation mechanism of artemisinin was reported to be its disruption of binding between Hsp90 and inducible nitric oxide synthase (iNOS), causing less NO, the inflammation signal molecule, is released (Wu, Cheng et al. 2019).

3.1.6.2 Natural products which modify Hsp90 cysteines.

Human Hsp90 alpha has 7 cysteines, and Hsp90 beta has 6 cysteines, located in the middle and C-terminal domain. Therefore, the natural product consists of cysteine reactive electrophile; for example, acrylate can potentially bind to one or multiple cysteine residues of Hsp90.

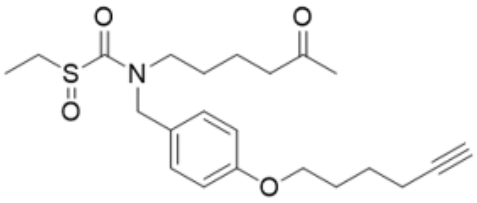
Sulforaphane **100** (Figure. 83) is a kind of isothiocyanate which widely found in cruciferous vegetables, e.g. broccoli. Plants synthesize the compound as a protective response of physical injury such as cutting and chewing (Fahey, Zalcmann et al. 2001). Sulforaphane **100** has been long investigated for its anticancer chemopreventive activity and reported to covalently tagged to Hsp90 NTD Ile72 – Arg81 without mentioning the specific residue being modified (Li, Karagöz et al. 2012). Such modification resulted in disruption of the Hsp90-p50-Cdc37 complex and resulted in anticancer activity. Later, to

enhance the covalent modification effect of **100**, a small library of sulphoxythiocarbamate analogues (which were synthesized from sulforaphane) were synthesized and tested to determine its target protein. It was reported that one of the synthetic analogues sulphoxythiocarbamate alkyne **101** (Figure 83, STCA) caused overexpression of Hsp70 by inhibition of the Hsp90-HIF-1 complex. Induction of STCA **101** also causes destabilization of Hsp90 co-client onco protein such as HER2 and Raf-1 without affecting Hsp90 ATPase activity (Zhang, Naidu et al. 2014), suggesting that STCA **101** is an Hsp90 CTD inhibitor. Due to the fact that the electrophilic sulphoxythiocarbamate group of STCA **101** may react with nucleophilic cysteine residues in its protein targets, further investigation of STCA **101** binding site was carried out using LC-MALDI TOF proteomic analysis. STCA **101** was incubated with human Hsp90 beta, and the sample was analyzed confirming that STCA modifies Cys564 & Cys589/Cys590 of Hsp90 beta (Zhang, Naidu et al. 2014). Cys589/Cys90 in Hsp90 beta equals Cys597/Cys598 in Hsp90 alpha according to sequence alignment, which again emphasizes the importance of the switch point double cysteine residues in terms of Hsp90 CTD inhibition.



100

Sulforaphane



101

Sulphoxythiocarbamate alkyne (STCA)

Figure 83: Natural Hsp90 covalent inhibitor sulforaphane **100** and its synthetic analogue sulphoxythiocarbamate **101** (STCA).

Kongensin A **102** (Figure. 84, KA) is a natural product separated from Croton kongensis plant and being reported as an Hsp90 middle domain inhibitor, which covalently modifies Cys420 of human Hsp90.

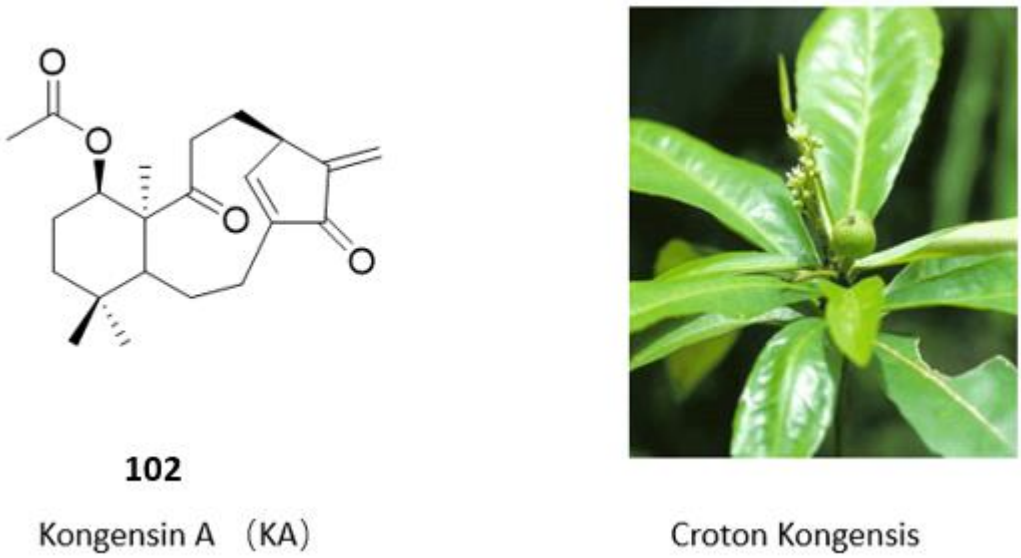


Figure 84: Kogensin A **102** was an extracts of Croton Kongensis.

KA **102** displays outstanding activity in its initial screening of receptor-interacting kinase 3 (RIP3) dependent necroptosis inhibitors, acting as a necroptosis inhibitor and apoptosis inducer. However, KA **102** does not bind to any protein directly involved in this process, and the hydrolyzed product of KA **102** lost activity. Therefore a chemical probe-based pull-down experiment using KA **102** and hydrolyzed KA identified Hsp90 as its target protein (Li, Li et al. 2016). To validate the importance of covalent binding, Cys to Ala mutated Hsp90s were incubated with KA **102**, and C420A mutants were reported to lose activity, which was further confirmed by MS/MS proteomic analysis demostrating that KA **102** covalently binds to Cys420. Cys420 of human Hsp90 belongs to the middle domain. The middle domain (MD) of Hsp90 is much less studied than NTD and CTD, let alone the report of Hsp90 MD inhibitors. Discovery of KA **102** as an Hsp90 MD covalent modifiers also identified a novel mechanism, which KA **102** inhibited Hsp90 MD will dissociate from its co-chaperone Cdc-37, therefore, subsequentially inhibiting RIP3 dependent necroptosis and promoting of apoptosis in multiple cancer cell lines (Li, Li et al. 2016).

Penicisulfuranol A **103** (PEN-A) (Figure. 85) is a newly identified compound consisting of a rare 3H-spiro [benzofuran-2, 2′-piperazine] ring isolated from mangrove endophytic fungus Penicillium janthinellum HDN13-309 (Dai, Chen et al. 2019). It exhibited strong antiproliferative activities. Removal of the disulfide bond (**104** in Figure 84, Penicisulfuranol D) resulted in dramatically reduced anticancer activity, suggesting that the disulfide bond of PEN-A may serve as a cysteine reactive moiety.

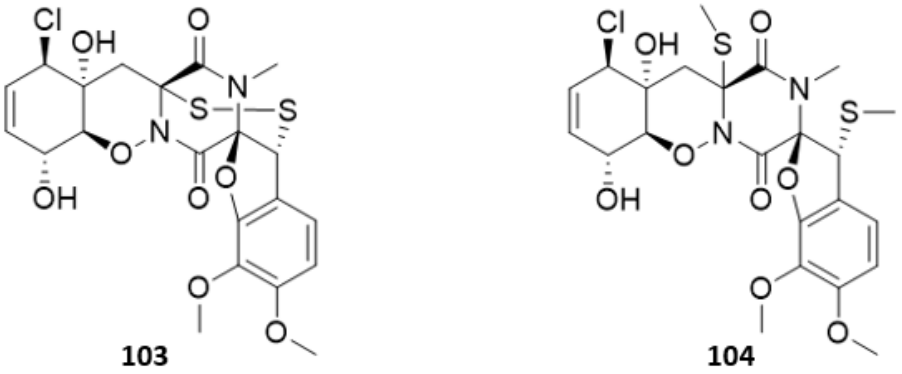
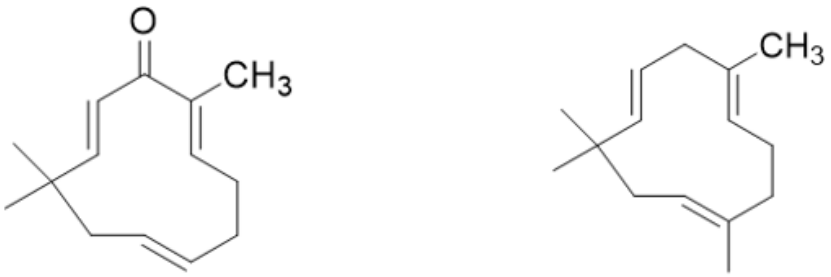


Figure 85: Penicisulfuranol A **103** consists of 3H-spiro [benzofuran-2, 2′-piperazine] ring and an α, β-disulfide bridge. Removal of the disulfide bridge leads to penicisulfranol D **104**.

PEN-A **103** was confirmed as an Hsp90 CTD through proteolytic fingerprinting assay, in which PEN-A **103** protected the Hsp90 alpha CTD from trypsin degradation other than Hsp90 NTD or Hsp90 MD. Also, PEN-A **103** exhibits similar effects on Hsp90 dimerization and interaction of co-chaperones with Hsp90 (Dai, Chen et al. 2019). Although the author did not report the specific residue modified by PEN-A **103**, it is concluded that it may covalently bind to C592, C597/C598.

Zerumbone **105** (Figure. 86) is a cyclic sesquiterpene that was first isolated from the rhizome oil of Zingiber zerumbet and was reported to present anti-inflammatory effects (Murakami, Takahashi et al. 1999). The most studied biological activity of zerumbone **105** is its inhibition of proliferation of a wide variety of cancerous cells including cells derived from colon and breast cancers (Kirana, McIntosh et al. 2003), leukaemia, and myeloid and liver cancers (Yodkeeree, Sung et al. 2009). Zerumbone **105** was reported to have antiproliferative activity through covalent binding via Hsp90 mid domain cysteines, thus, disrupting the binding of Cdc37 towards Hsp90 and leading to antiproliferative activity (Nakamoto, Amaya et al. 2018).



105
Zermbone

106
Humulene

Figure 86: Structure of zerumbone **105** and its reduced form humulene **106**.

Different from other Hsp90 inhibitors, which reduce ATPase activity, induction of zerumbone results in an increase of ATPase activity, especially for human Hsp90 alpha (increased 467%); however, it does not affect yeast Hsp82 due to no cysteine residue existing in this protein (Nakamoto, Amaya et al. 2018). Humulene **106** (Figure. 86) is the reduced structure of zerumbone **105**, which contains no cysteine reactive Michael receptor groups. Although humulene **106** slightly reduces Hsp90 ATPase activity, it showed no antiproliferative activity. It was also concluded that zerumbone would alter the conformation of Hsp90 N and Middle domain, but the report lacked a detailed mechanism of action. As a result, zerumbone **105** became an exceptional case of which its covalent modification causes positive-correlated effects between Hsp90 ATPase activity and antiproliferative activity. Other reports came to similar conclusions, of which Hsp90 ATPase positive modifications (ATPase activity

increasing mutations and induction of ATPase activating agents) results in growth defects under normal conditions (Nathan and Lindquist 1995) (Yokoyama, Ohtaki et al. 2015). Therefore, covalent inhibition of Hsp90 may either increase or decrease ATPase activity, and Hsp90 ATPase activity is not necessarily correlated with overall cell survival status. Considering cysteines in Hsp90 are reported as switch point, the correct, controlled and normally organized duration of different Hsp90 confirmation may be the key that decides the overall chaperone function, which can be interfered with by Hsp90 cysteine covalent modification (Nakamoto, Amaya et al. 2018).

3.2 Project Objectives

In previous section (chatper 3, 1.0 Introduction) we have reviewed the feasibility of applying the covalent drug strategy towards Hsp90 protein and concluded that Hsp90 CTD active cysteine residues may be modified by small molecules, making Hsp90 CTD cysteine a good covalent drug target.

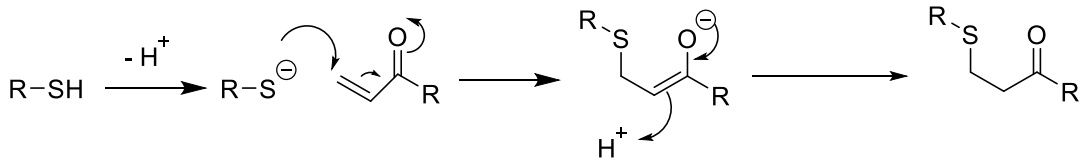
As a result, in this project, we aimed to design and synthesize several covalent type Hsp90 CTD inhibitors based on Hsp90 CTD non-covalent scaffold novobiocin **0,** and then determine their antiproliferative activities in order to generate a systematic SAR study of covalent modification. We also aimed to develop a protein-based assay to visualize the effect of Hsp90 CTD inhibition and another assay to quantify the binding affinity.

If results from these Hsp90 CTD inhibition assays show some evidence of successful covalent binding (e.g. super high binding affinity), then we aimed to explore the actual modified residues through mass spec based proteomic analysis. Data obtained from proteomic analysis would be a solid proof of successful Hsp90 CTD covalent binding thus fulfilled the blank knowledge of Hsp90 CTD covalent drug design.

3.3 Result and Discussion

3.3.1 Design of cysteine targeted electrophiles.

Cysteine residue can be covalently oxidized by electrophiles, with one example being it reacting with a Michael acceptor through a Michael addition reaction. In this reaction, cysteine act as a nucleophile and attacks the electrophilic double bond of α,β-unsaturated carbonyl compounds, resulting in a carbonyl sulphide (Scheme 23).



Scheme 23: Mechanism of Michael addition when thiol act as a nucleophile.

In this project, we adopted a “rocket – warhead” design strategy (Figure. 86).

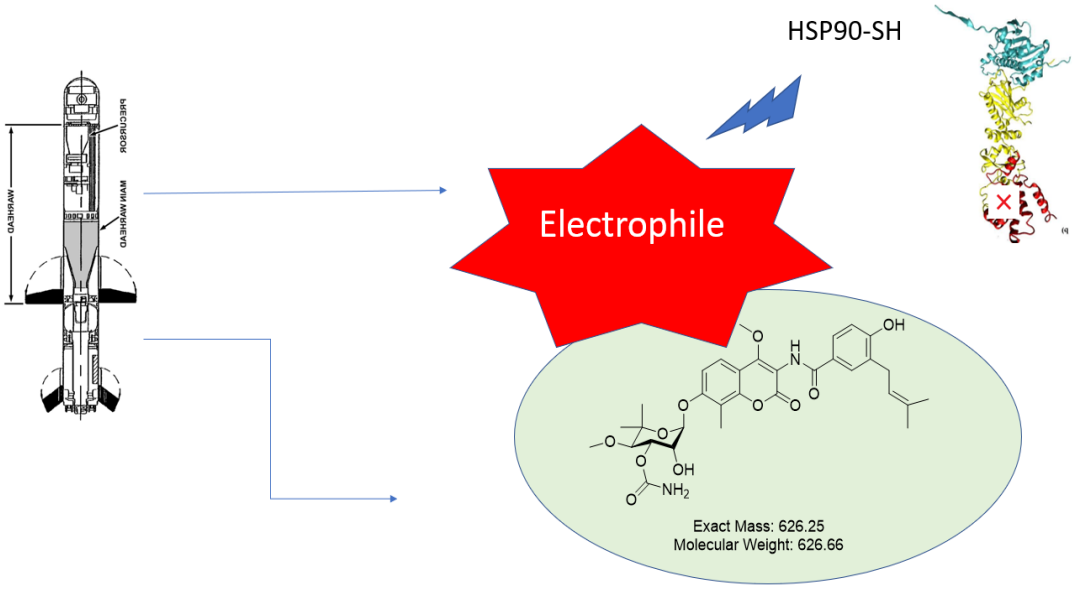
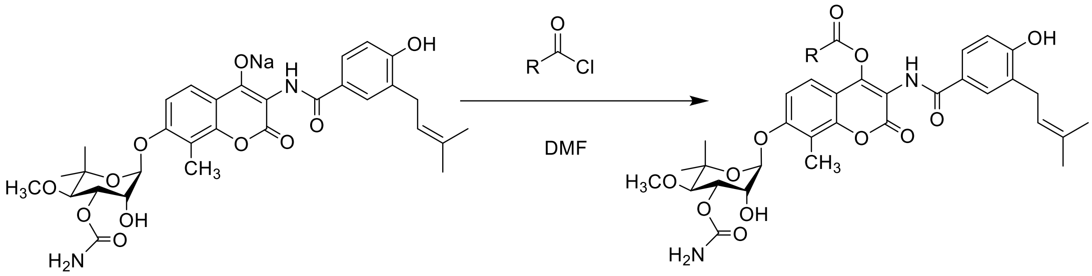


Figure 87: Illustration of “rocket – warhead” strategy. Novobiocin act as a site-directed “rocket” function that assures the electrophile group be delivered to the right pocket to react with active cysteine.

A site-directed missile weapon usually has two parts: the rocket part, which delivers the warhead to the location, and the non-directive warhead contains explosives to cause the damage. Covalent drug design is quite the same, in which the non-covalent backbone acts as the rocket, and the electrophile act as a warhead. There are numerous proteins inside a cell, and any of them would contain active cysteine residues. In other words, without a proper “rocket” scaffold, the “warhead” electrophile has a chance to be destroyed by cysteines from other proteins rather than attack by cysteines from Hsp90. The “rocket” must non-covalently binds to the binding pocket of the desired target and then carry out the reaction with cysteine.

As discussed in the introduction section, novobiocin is a classic well-studied Hsp90 CTD inhibitor. Although it can be deduced from its mild antiproliferative activity (IC_{50} = 700 µM) that the binding affinity of novobiocin towards Hsp90 CTD was mild, after so many years of study, there is no report of any off-target activity. Novobiocin scaffold will therefore ensure the specific binding toward human Hsp90 alpha. In this way, novobiocin acts as the “rocket” of the missile. In addition, novobiocin is commercially available as its 4’-sodium salt, making it easier to synthesize 4’- substituted products through a nucleophilic attack.

Method development in chapter one concluded that DMF as a solvent would facilitate good regioselectivity to give 4'- substituted addition products. The reaction of novobiocin with acid chloride to from the 4'- substituted product is straight forward and easy to handle (Scheme 24).



Scheme 24: Proposed synthesis of novobiocin with carbonyl chloride.

For the selection of warheads, the main class of electrophile that we intended to use, were Michael acceptors. Due to the 4'-hydroxy property of novobiocin, this type of analogues will be 4'- acrylate novobiocin type. Low – molecular weight acrylate compounds have been wildly used as cysteine biomarkers in structural biology. It generally only reacts with the protein cysteine residue, which ensures specificity. Vinyl sulfonamide is a sulfur-based Michael acceptor that shares a similar activity with acrylate. It was synthesized in this project to expand diversity. An inactive 4'-propionyl was synthesized as an inactive form of 4'- acrylate to compare data.

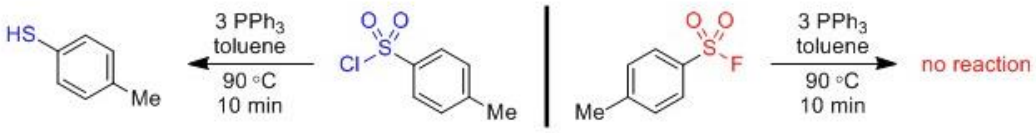
Also, the "warhead" pool was not designed to be only limited to Michael acceptor types. Tosylate groups have been used as leaving groups in chemical reactions. It is reported that 4'- tosyl substitution on the novobiocin-like scaffold will increase its antiproliferative activity. As a result, we included this modification.

Another essential design we adopted was the synthesis of 4'- benzenesulfonyl fluoride substituent. The benzenesulfonyl grouup is a type of newly reported protein probes. Sulfur (VI) fluoride exchange (SuFEx) has been identified as the a chemistry alternative reaction owing to the balanced properties of fluoride displacement (Scheme 24).

Scheme 25: Mechanism of sulfur fluoride exchange.

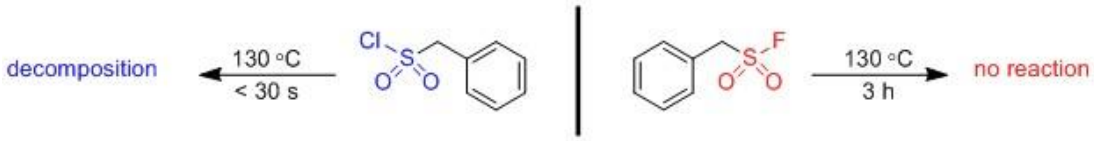
This covalent modifier is superior to traditional ones for the following reasons:

1. **Resistance to reduction**: Different from other halides, sulfonyl fluoride cleavage is heterolytic and thus resistant to reduction (Scheme 26) (Dong, Krasnova et al. 2014).



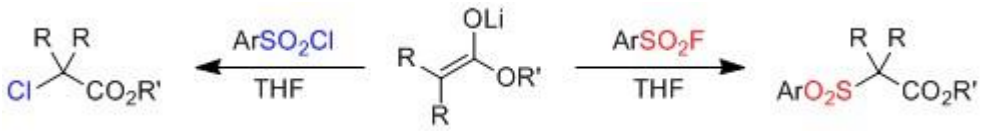
Scheme 26: SO_2F group resist to reduction. While SO_2Cl is reduced with specific conditions (left column) , SO_2F is stable enough to resist the reduction with the same condition (right column).

2. **Thermodynamic stability**: Sulfonyl fluorides are stable to thermolysis and nucleophilic substitution (Scheme 27) (Krutak, Burpitt et al. 1979).



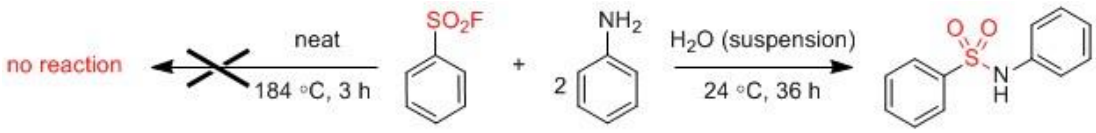
Scheme 27: SO_2F group has thermodynamic stability: While SO_2Cl group undergoes decomposition at 130 ℃ (left column), SO_2F group is stable at this temperature with no reaction happens.

3. **Exclusive reaction at sulfur**: Sulfonyl fluorides react fast. Compared to sulfonyl chlorides, chemoselectively produce only sulfonylation products (Scheme 28) (Dong, Krasnova et al. 2014)



Scheme 28: SO_2F group has an exclusive reaction at sulfur

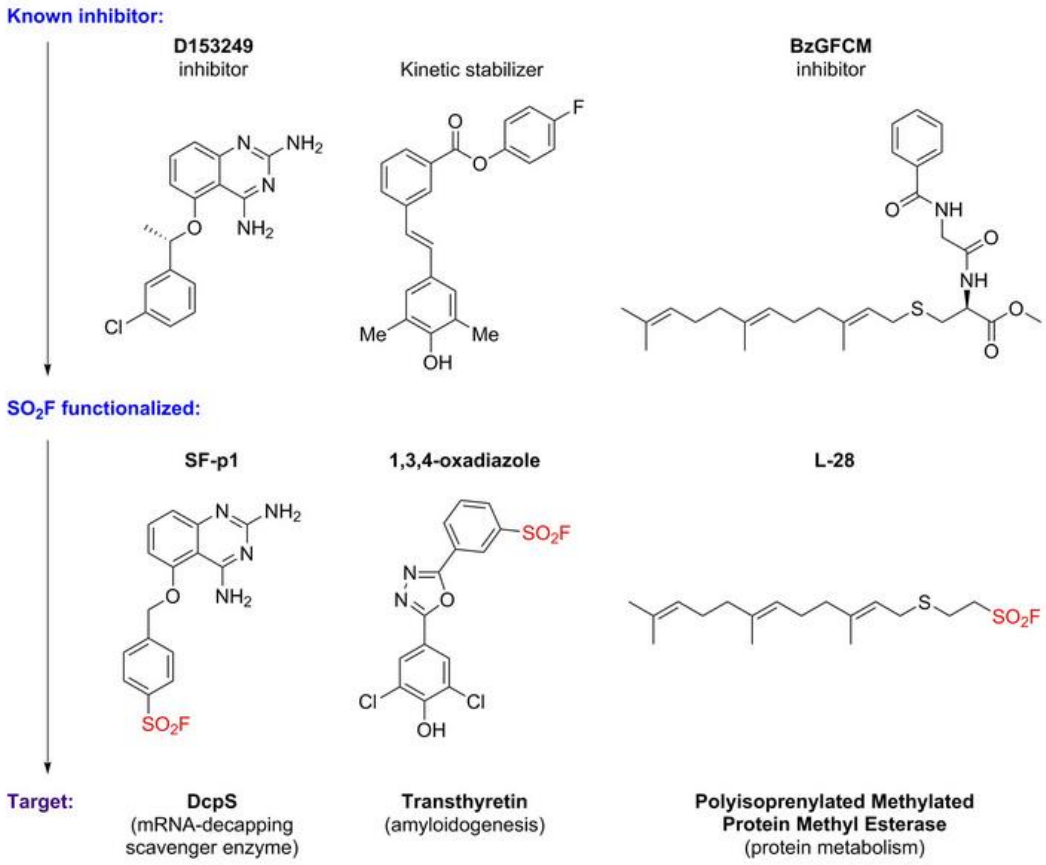
4. **Special nature of the fluoride-proton interaction**: Stabilization of the fluoride ion in water affords chemistry in aqueous environments, making it ideal for biological activation (Scheme 29) (Dong, Krasnova et al. 2014).



Scheme 29: SO_2F presents good reactivity only in aqueous condition.

The SO_2F group covalently binds to many protein residues but in a context-specific manner: serine, threonine, tyrosine, lysine, cysteine, and histidine (Dong, Krasnova et al. 2014). Therefore, sulfonyl fluorides can be uniquely applied as chemical probes or covalent inhibitors.

Some recent examples of the functionalization of previous inhibitors and enzyme stabilizers with SO_2F lead to the chemoselective modification of proteins of interest (Scheme 30): DcpS on Tyr, (Hett, Xu et al. 2015) transthyretin on Lys, (Grimster, Connelly et al. 2013) and polyisoprenylated methylated protein methyl esterase on Ser (Aguilar, Amissah et al. 2011).



Scheme 30: Some examples of SO_2F substituted compound leading to targeted protein modification.

However, unlike Michael acceptors, that exclusively modify cysteines, sulfonyl fluoride, as mentioned before, react with multiple residues type including lysine and serine. As a result, the test data of SO_2F equipped analogue will broaden the knowledge of Hsp90 CTD covalent modification.

Based on all designs, a summary of synthetic targets is given in Figure 88.

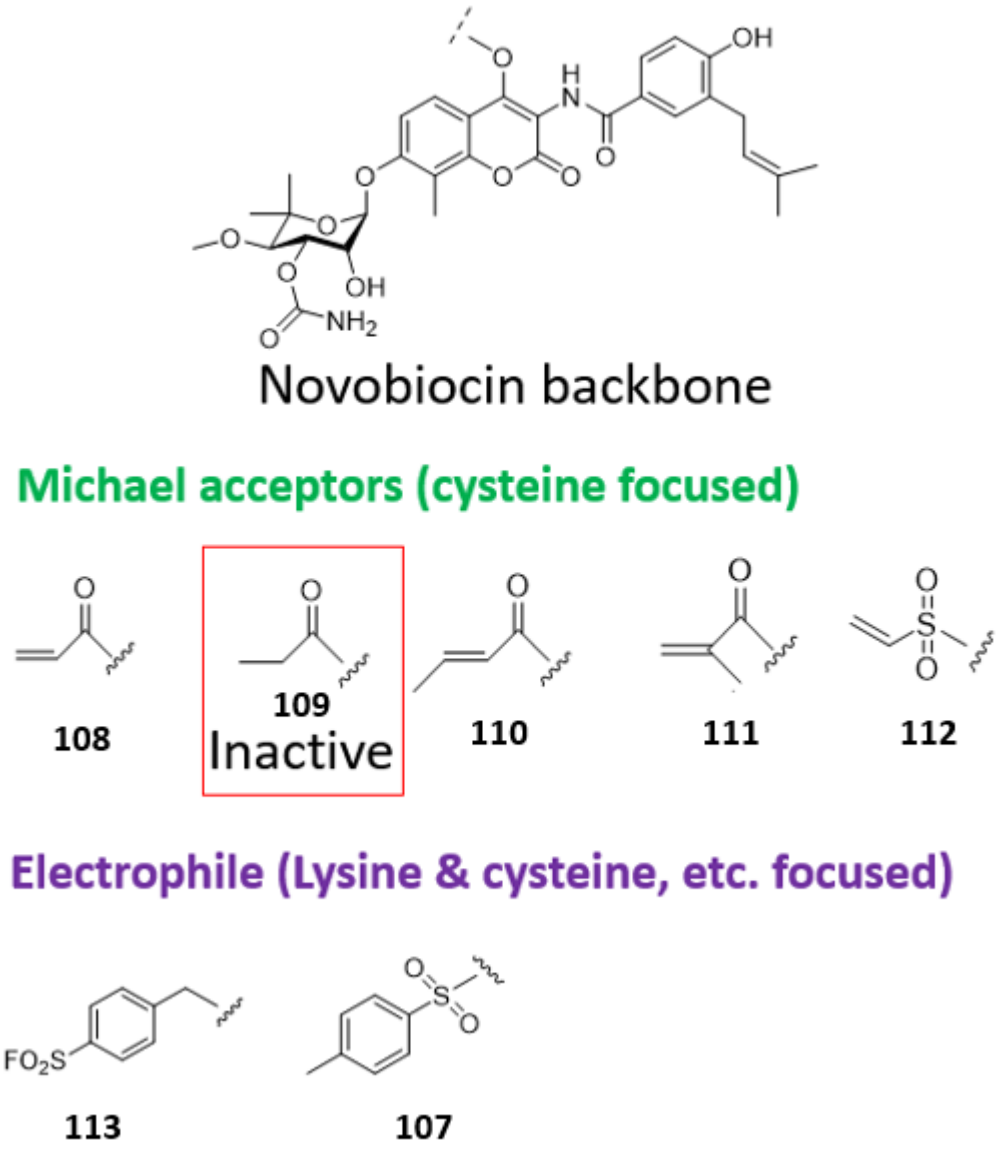


Figure 88: Summary of synthetic targets.

The reduced, cysteine unreactive form of Michael acceptor moiety was also synthesized **109** (marked as inactive in the figure above), as a comparison.

Moreover, several 4' - substituted coumarins were synthesized (Figure 89) to explore the importance of 4' substitution of the coumarin core structure. Nevertheless, because of limited time and resources, all the 4'-substituted novobiocin analogues are built with a default 4'- ester bond linker originated from 4'-hydroxy of novobiocin. Therefore, 4-amino coumarin **116** and 4–thiol

coumarin **117** were synthesized and tested to see whether a 4'–linker type affects the antiproliferative activity.

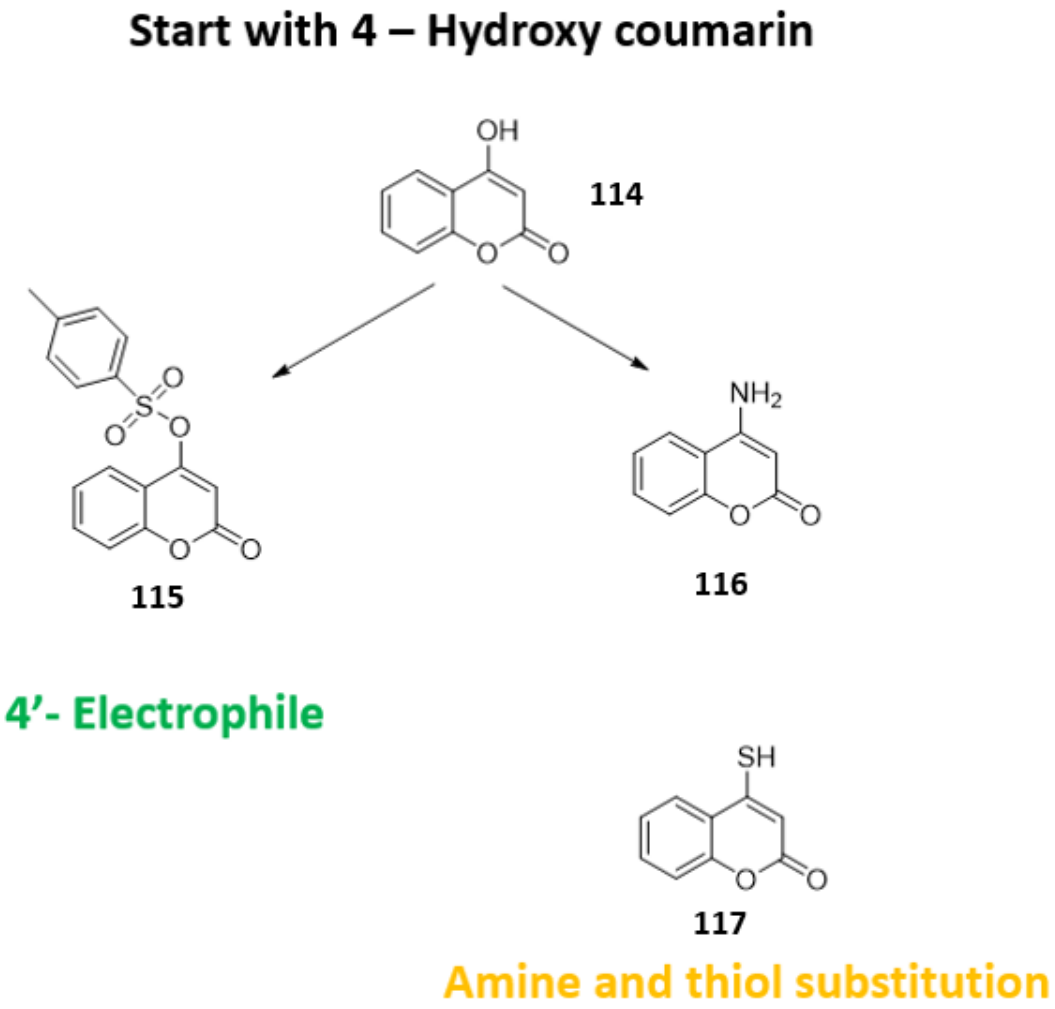


Figure 89: Synthetic targets of 4-substituted coumarins.

3.3.2 General synthetic method of designed 4' – substituted novobiocin analogues

As planned in section 3.1 Chapter 3, the main synthetic task of this chapter focuses on the synthesis of 4' – electrophile - novobiocin products. To achieve this goal, novobiocin sodium salt **0** (Figure 90) was reacted with chlorides of electrophiles in DMF to give the desired products.

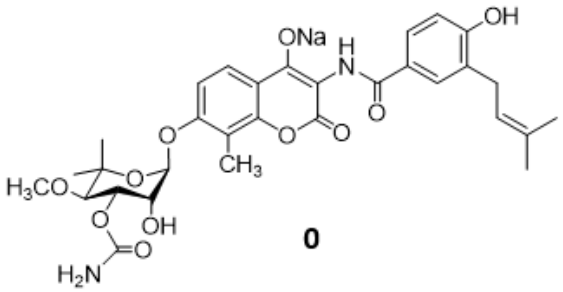
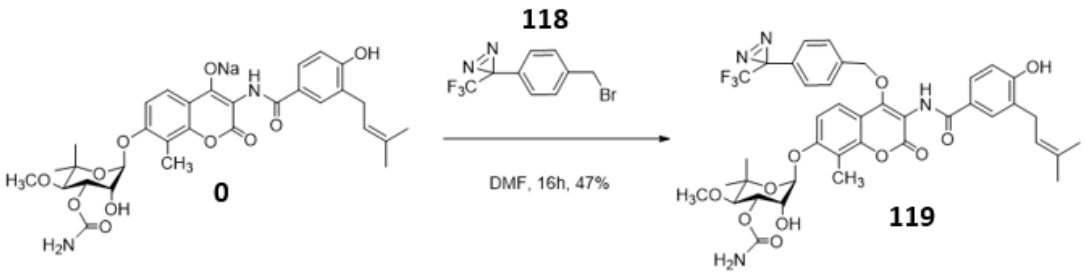


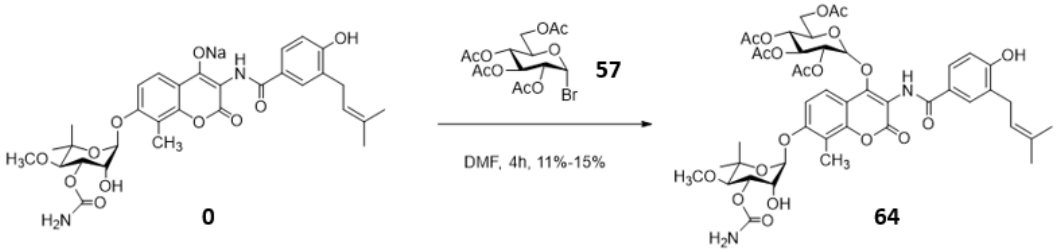
Figure 90: Novobiocin sodium salt **0** from Sigma-Aldrich (Merk) N 1628

It is worth mentioning that all the novobiocin analogues synthesized in this thesis came from this product. Because this product has a deprotonated 4'–hydroxy, the S_N2 reaction usually is usually favoured at the 4'–OH rather than 5-OH of the amide side chain and 2''-OH of noviose sugar. Previously in Min's group, we have published two papers that reported this type of reaction (Simon, Huang et al. 2017) (Sun, de Resende et al. 2017). The first one refers to the reaction of novobiocin **0** sodium with diazirine bromide **118** to give the 4' substituted analogue **119** （Scheme 31）. It is worth mentioning that the polar aprotic solvent DMF enhances regioselectivity towards the 4'-OH of novobiocin (Simon, Huang et al. 2017).



Scheme 31: Formation of novobiocin-photo affinity label analogue through regioselective nucleophilic substitution at 4'-OH. (Simon, Huang et al. 2017)

The second example mentioned in chapter one (p69), where novobiocin sodium salt **0** reacts with acetyl-glucose-bromide **57** to give the glycosylated product **64** albeit in low (10% - 15%) yield was developed to replace the traditional toxic catalyst $Hg(CN)_2$ used previously (Scheme 32) (Andreia 2011). Adding a mild base or Lewis acid in the reaction system has almost no effect on the yield (Sun, de Resende et al. 2017). Note that only a traceable amount of side product was found in this reaction; most of the novobiocin remained unreacted under room temperature.



Scheme 32: Novobiocin reacts with acetyl-glucose-bromide to give 4'-glycoside product of novobiocin.(Sun, de Resende et al. 2017)

Another experimental detail which was not discussed in chapter one but needs to be addressed here is the solubility of novobiocin. In most common solvents stored in the laboratory (e.g. chloroform, toluene, DCM, THF, acetone, and ethyl acetate), novobiocin sodium salt **0** simply does not dissolve well. It only dissolves in MeOH, water, and DMF due to its ionized salt form. For those two published examples discussed before (Sun, de Resende et al. 2017) (Simon, Huang et al. 2017), we reported a problem in purification using normal phase column. Due to the boiling point of DMF being too high to be evaporated, traditionally, water is added into DMF solution and then the crude product is extracted by hydrophobic organic solvents. However, in this case, extraction is not applicable as once the reaction was finished, novobiocin would lose its salt form and turns into the original protonated form and then will dissolve in the organic layer. As a result, extraction will not separate the product from the starting material. Purification using flash column also suffers from a poor solubility of the starting material and the product, as we were unsuccessful in developing an efficient normal-phase flash column purification method due to

too much volume of solvent needed for loading. This causes incomplete separation, and it was extremely time-consuming to wash the crude product to get pure products.

To solve this problem, the best way we found was to load the DMF solution directly onto the automated reverse phase purification system (Biotage) once the reaction was complete. The Biotage system is a fully automated flash purification system using a commercial packed disposable column. The system can achieve high-performance efficacy of separation like HPLC. As DMF solution is water-soluble, a gradient solvent system using MeOH / water was selected.

Figure 91: The automated flash purification system Biotage.

The purification process is monitored through UV, and fractions are collected by the auto fraction collector. In this way, it was easy to observe those fractions with the highest UV reading for better purity.

There are 3 process variables to be settled for each purification process:

1. Initial segment: The system runs with a fixed percentage of solvent mixture dominated by water. This process will slowly distance different components inside the column to prohibit instant wash off caused by loading. This is useful when the sample must be loaded with 100% organic phase. The initial segment should last no more than 3 CV.

2. Gradient: This is the same process of being used by HPLC. Herein, the percentage of organic phase will gradually increase by time. The more hydrophilic the compound is, the faster it comes out. The default time for this process is 10 CV.

3. Final segment: Usually, during this phase, 100% organic solvent will be applied to wash off anything left inside the column, which lasts 1 to 3 CV or until no reading on the UV monitor.

Herein, the purification of 4'-acetyl-glucose novobiocin is used as an illustrative example of the methods that we employed. The DMF reaction solution 10 (mL) was directly injected into the top of the column and the purification (Biotage method 1) settings are shown below:

Biotage method 1:

Solvent: Water / Methanol

Initial segment: 0% Methanol for 3 CV

Gradient: 0% Methanol to 100% Methanol in 10 CV

Final stage: 100% Methanol to 100% Methanol in 3 CV

However, we quickly spotted that 10 mL of DMF solution was too much for the wet loading capacity of the column. The system quickly washed everything off from the column without any separation (Figure 92).

Figure 92: When 10 ml DMF loaded into 136 ml per CV reverse phase column, the system will wash everything off instantly.

Commonly, the sample is loaded in organic phase solvent with a volume of <10% CV. In other words, for a 136 mL CV column here, loading a sample with < 13.6 mL polar organic solvent (e.g. MeOH, THF) will not induce any of the problems shown in Figure 89. The reason why >10% organic volume should be avoided is that the organic solvent of loading will also be added into the column solvent system, changing the percentage of the original solvent mix and pushing all components to travel together.

DMF is usually banned by reverse phase purification due to its overwhelming pushing power resulting from its polarity. Since 10 ml DMF was too much for the system; obviously, a limitation of loading volume should be defined. After a few experiments, 5 mL DMF was defined as the maximum volume to be used for loading. To make the purification in accordance with the reaction, 5 ml DMF was set as a standard for the synthesis of any novobiocin based analogues, and 300 mg novobiocin sodium / 5 mL DMF was found to be the maximum concentration that could be applied.

Then we tested with 5 mL DMF loaded sample, and it turned out that the product and the starting material novobiocin peaks merged at about 9 CV position, indicating that this method was not applicable towards this case. To optimize the process, we took advantage of the well-adapted LC separation method used by the LC-MS machine mentioned in the experimental section

(p288). We applied the LC condition to the Biotage method, changing the initial segment from 0% MeOH to 30% MeOH and increasing the CV of gradient from 10 CV to 14 or 16 CV. Details are given below

Biotage method 2:

Solvent: Water / Methanol

Initial segment: 30% Methanol for 3 CV

Gradient: 30% Methanol to 100% Methanol in 14 CV

Final stage: 100% Methanol to 100% Methanol in 3 CV

Biotage method 3:

Solvent: Water / Methanol

Initial segment: 30% Methanol for 3 CV

Gradient: 30% Methanol to 100% Methanol in 16 CV

Final stage: 100% Methanol to 100% Methanol in 3 CV

A successful example of purification using Biotage method 3 is shown below (Figure 93):

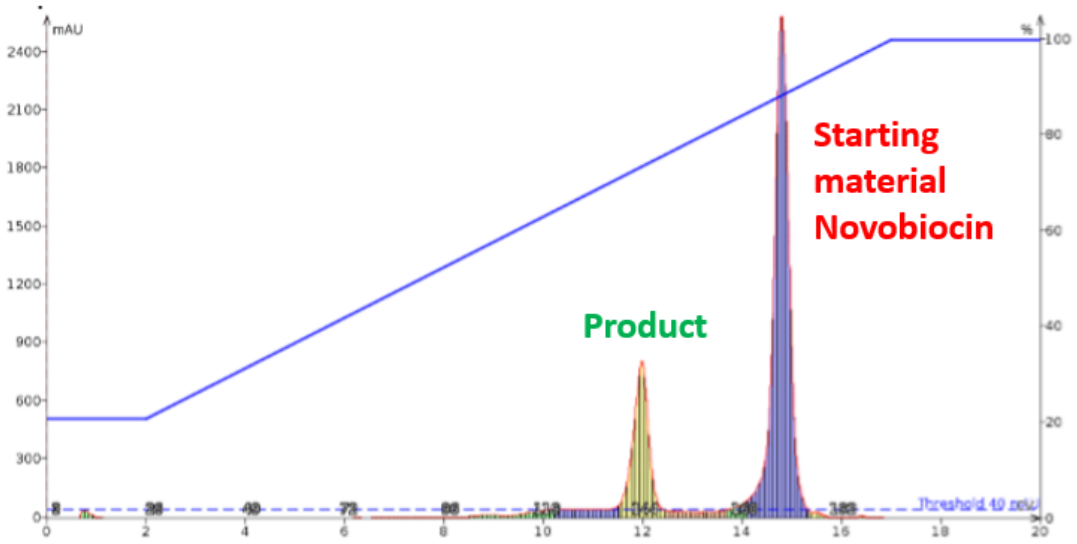


Figure 93: Purification of 4'-acetyl-glucosyl novobiocin **64** using Biotage method 3. Product came out at CV = 12, starting material came out at CV = 15.

Biotage method 2 and Biotage method 3 successfully purified most of the compounds mentioned in this thesis. Compared with the purification of dia-glycosylated novobiocin analogue from the last chapter, purification was much easier taking advantage of this method. All of the 4'-electrophile-novobiocins (**108** – **113**) were produced and purified following the same workflow in Figure **94**. Unless there is any particular case, purification of those 4'-substituted novobiocin analogues will not be discussed.

The standard workflow is shown in Figure 94. 300 mg novobiocin sodium **0** and relevant carbonyl chloride were mixed and dissolved in 5 ml DMF. When the reaction was complete, the DMF solution was directly transported into the Biotage reverse-phase column and purified by Biotage method 2 or 3 to afford the target products as colourless solids. This process will be mentioned as "standard process" in the rest of this chapter.

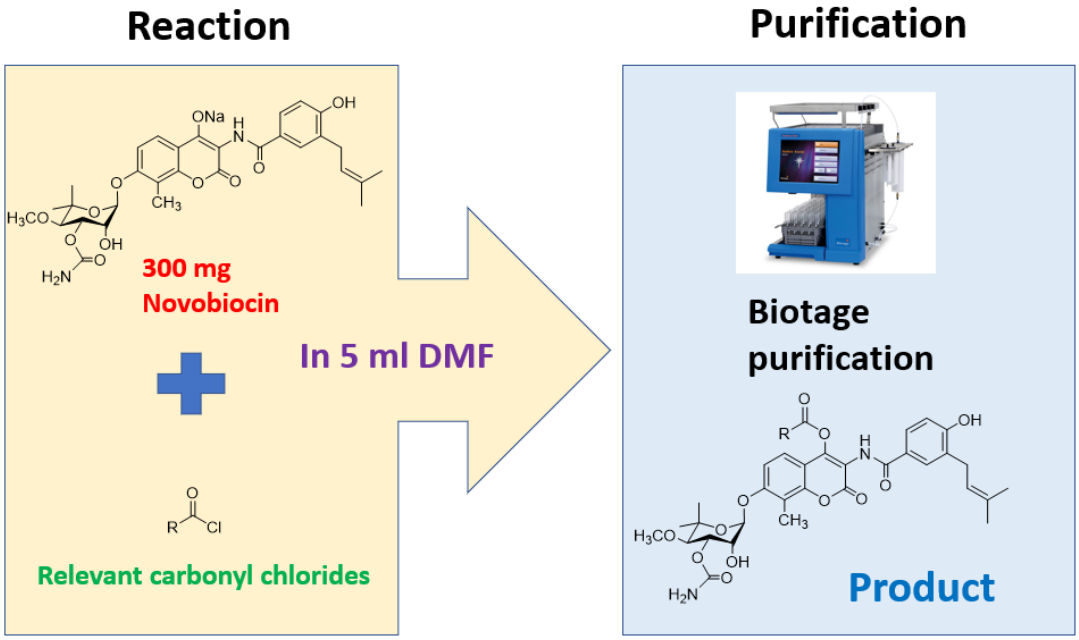
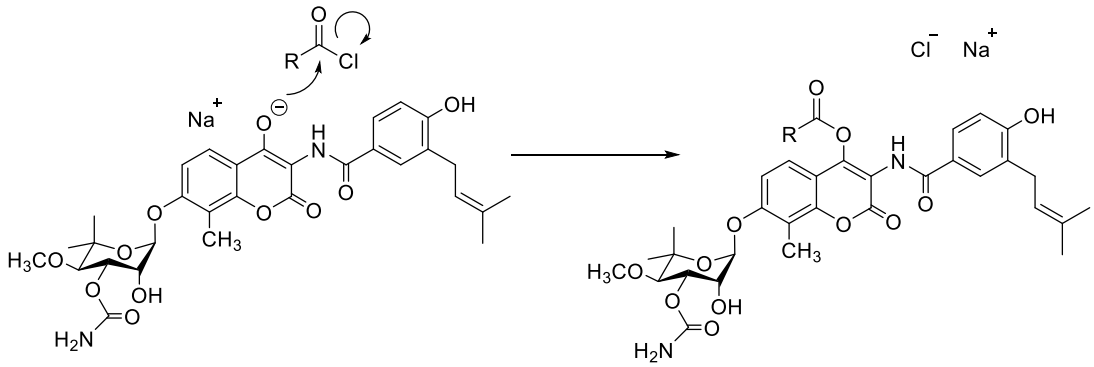


Figure 94: A standard experimental workflow to synthesize 4'-electrophile-novobiocin (**108** – **113**).

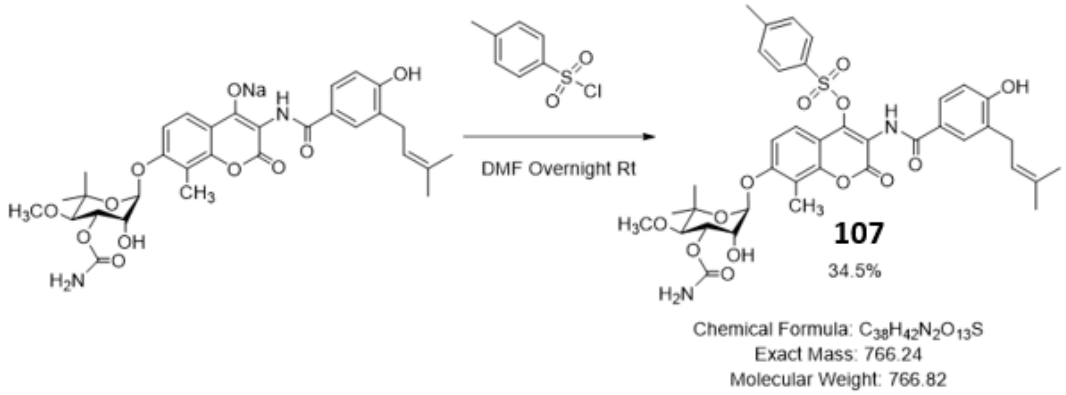
A general mechanism for this reaction is given below:



Scheme 33: Mechanism of synthesis of 4'-electrophile-novobiocins.

This reaction is a standard nucleophilic substitution following the S_N2 mechanism. Firstly, novobiocin sodium salt was dissolved in DMF to give a sodium cation and a 4'-deprotonated novobiocin inion. Then the negative charge act as a nucleophile attacking the electrophilic carbon atom of the carbonyl and displacing the chloride to form a new ester bond.

3.3.3 Synthesis of 4'–tosyl–novobiocin (107)



Scheme 34: Synthetic route of compound **107**

The synthesis of **107** followed the standard process showed in Figure 94 in section **3.3.2** chapter 3, p142. The reaction was left to stir overnight. After this period, the crude product was sampled and sent for LCMS. The LC trace (Figure 95) showed that the starting material novobiocin (**0**, RT = 4.26 in Figure 95) was almost completely consumed.

.

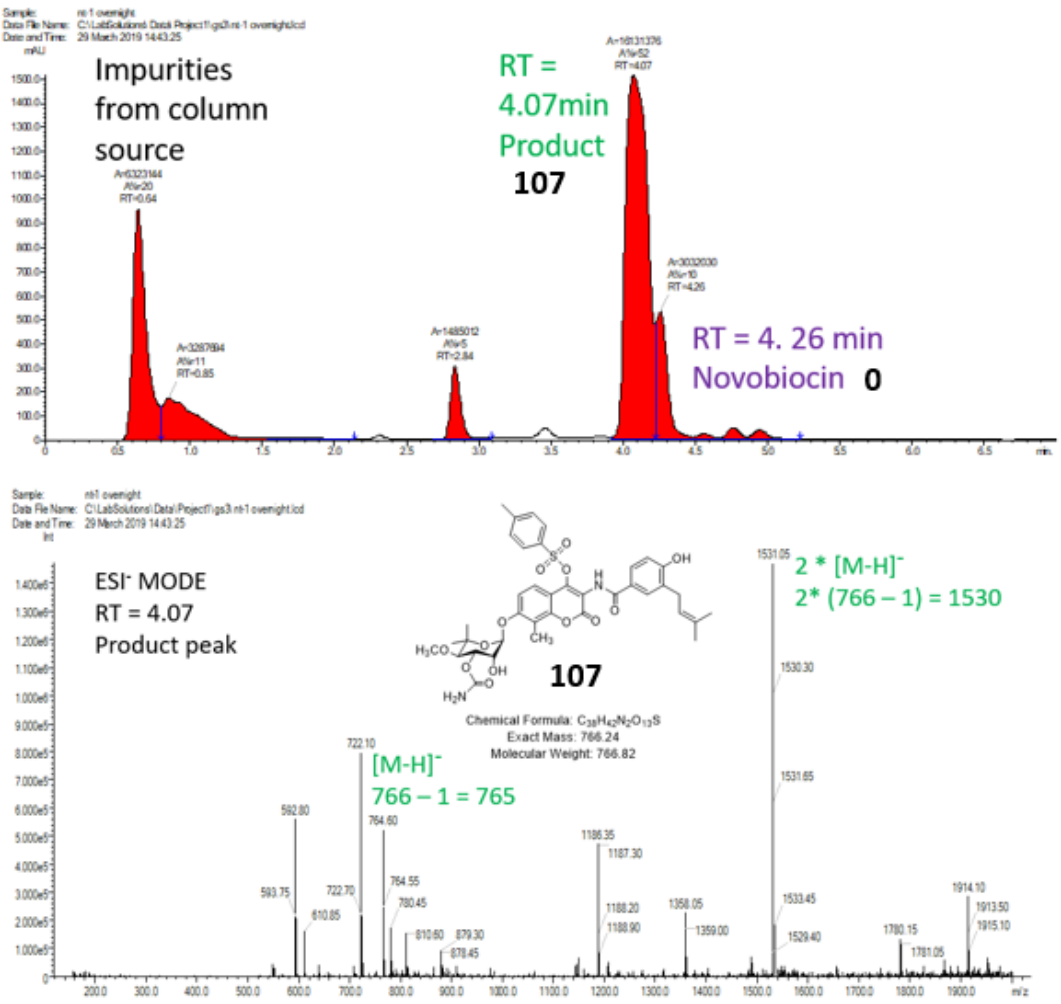


Figure 95:LCMS result of crude product sampled from reaction showed in scheme 34. Product **107** was found at RT = 4.07 min.

UV signal of the peak at RT = 4.07 min occupied 52% of total UV absolution (Figure. 95). MS result of the peak at RT = 4.07 min confirmed the formation of the product **107**. The starting material novobiocin **0**, as in the previous chapter 2, has an RT near 4.30 min (RT = 4.26 min in Figure. 95, the mass spectra is given below in Figure. 96). There is no valid mass spectra data for peak RT = 2.84 min due to poor ionization.

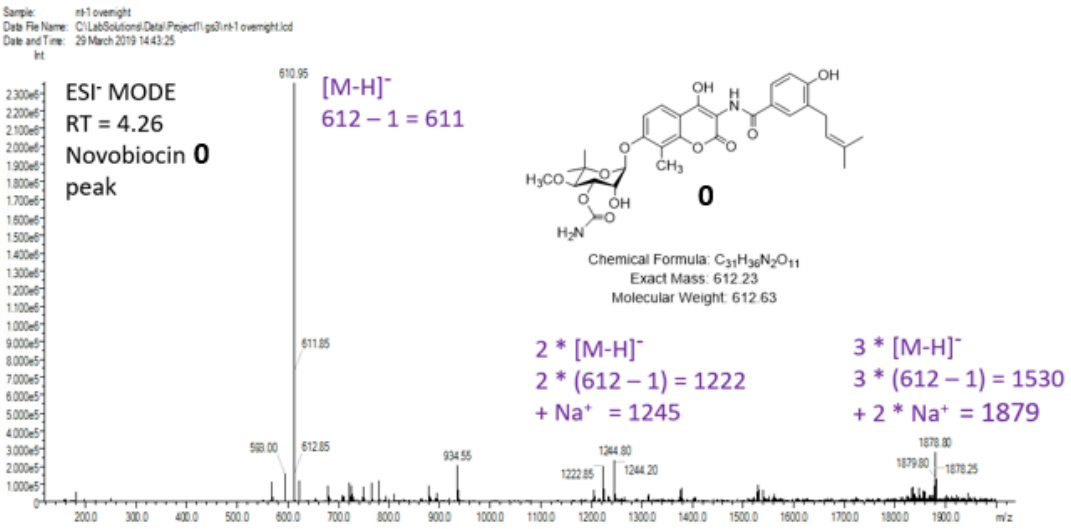


Figure 96: MS result of the peak at RT = 4.26 min in figure 95 which from the reaction in scheme 34. Peak refers to novobiocin **0**.

Although it has been discussed in **3.1** (p134) that extraction is not applicable for such a reaction due to poor organic solubility, herein, **107** was tested for extraction to be used as an example to explain why extraction is not applicable. Theoretically, substituting a hydrophilic hydroxy group with a hydrophobic tosyl group would significantly increase the organic solubility of the molecule. Since novobiocin sodium salt dissolves very well in water, hydrophobic **107** should be able to separate from novobiocin through extraction. To carry out the extraction, 50 ml of water was added to the DMF solution and then extracted with 2 x 50 ml DCM. Both two layers were sampled and were sent to LC-MS for analysis. The LC result of the water layer contained only small amount of product and novobiocin (Figure. 97). Nevertheless, the main component of the product and starting material were both transferred into the organic phase (Figure. 98). LC-MS data proved that the solubility of novobiocin was changed after the reaction, no longer in the form of a water-soluble salt.

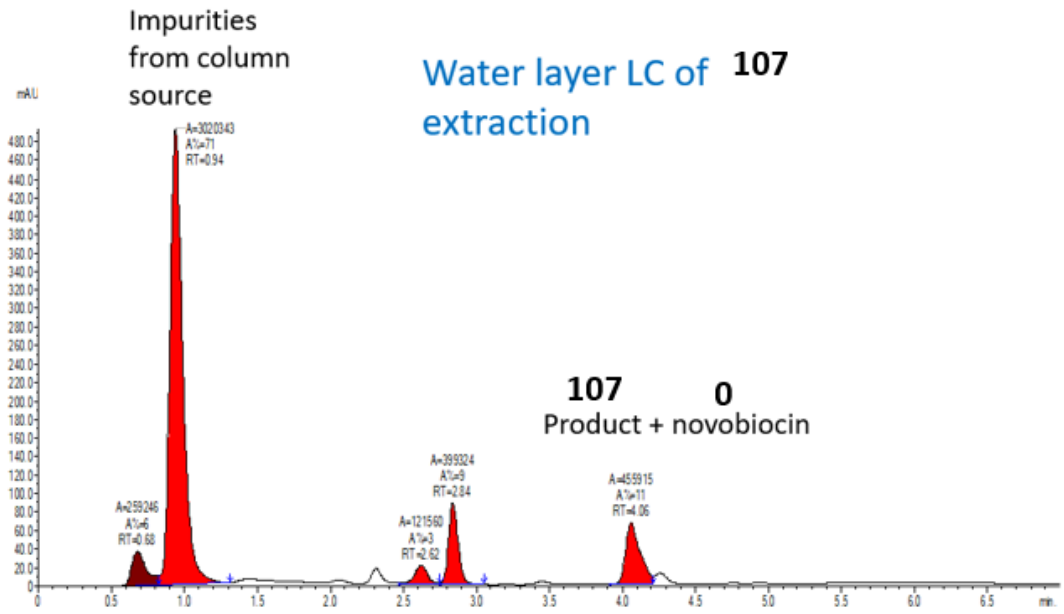


Figure 97: LC of water layer from **107** extraction, only little amount of **107** was found in water phase.

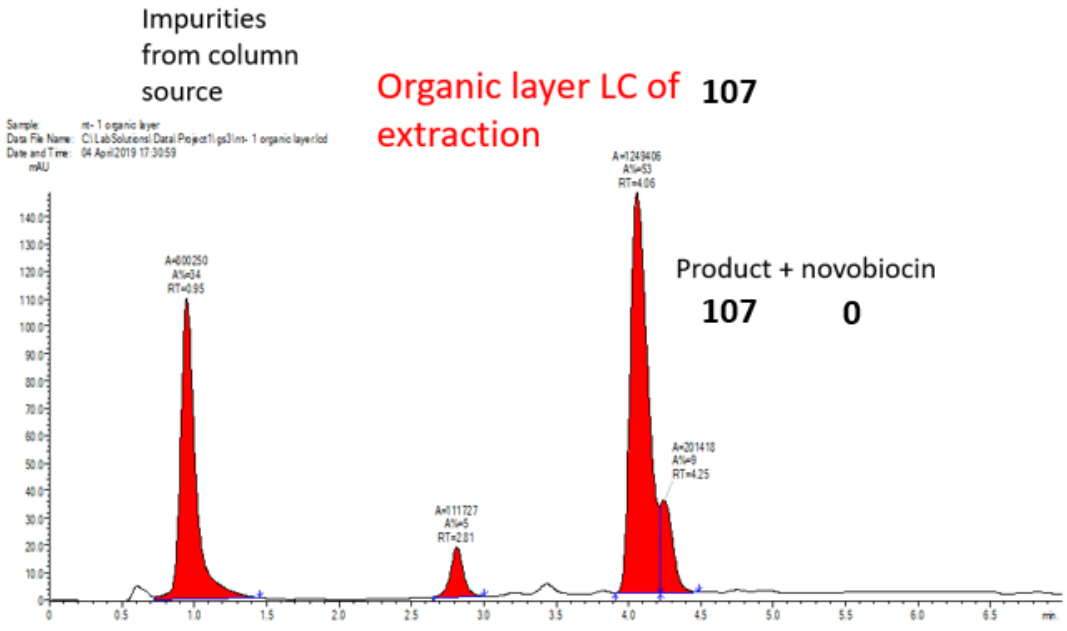


Figure 98: LC of the organic layer from **107** extraction.

Usually, during a reverse-phase column gradient process, hydrophilic compounds travel faster than hydrophobic compounds; thus, the smaller the

RT is, the more hydrophilic it is. However, even with a hydrophobic group added on the 4'-OH of novobiocin, the resulting theoretically more hydrophobic product **107** still eluted earlier then novobiocin, which is unusual. 4'-Substituted novobiocin analogues, with hydrophilic or hydrophobic group attached, all showed improved water solubility according to our observations.

As **107** failed to be purified by extraction, the organic layer was concentrated and purified using Biotage method 2.

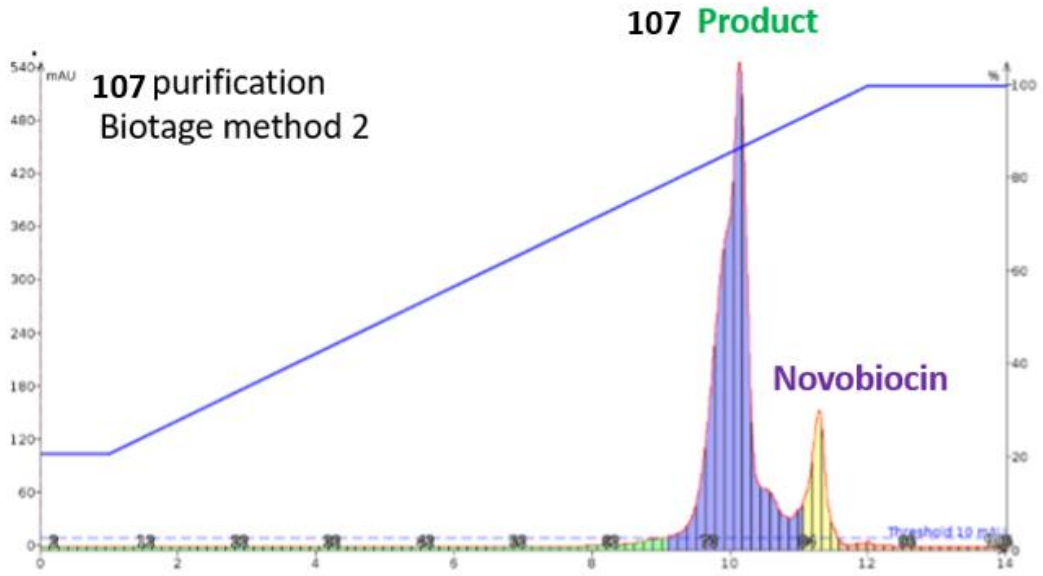


Figure 99: **107** Biotage method 2 purification results. Product came out at 10 CV (Column Volume)

The purification UV result was in accordance with LC-MS. Product **107** was collected at 10 CV (Figure. 99). The purified sample was sent for LC-MS to check the purity. LC-MS result of pure **107** is given in Figure 100. The purified sample has 93% purity, with 7% novobiocin content. For such a reaction with 300 mg novobiocin as a starting material, the final yield after purification was 34% (125 mg collected as a white solid). The low yield in this case was due to the reverse phase purification process, in which only the fraction with the highest UV absorption was collected for improved purity (Figure. 100).

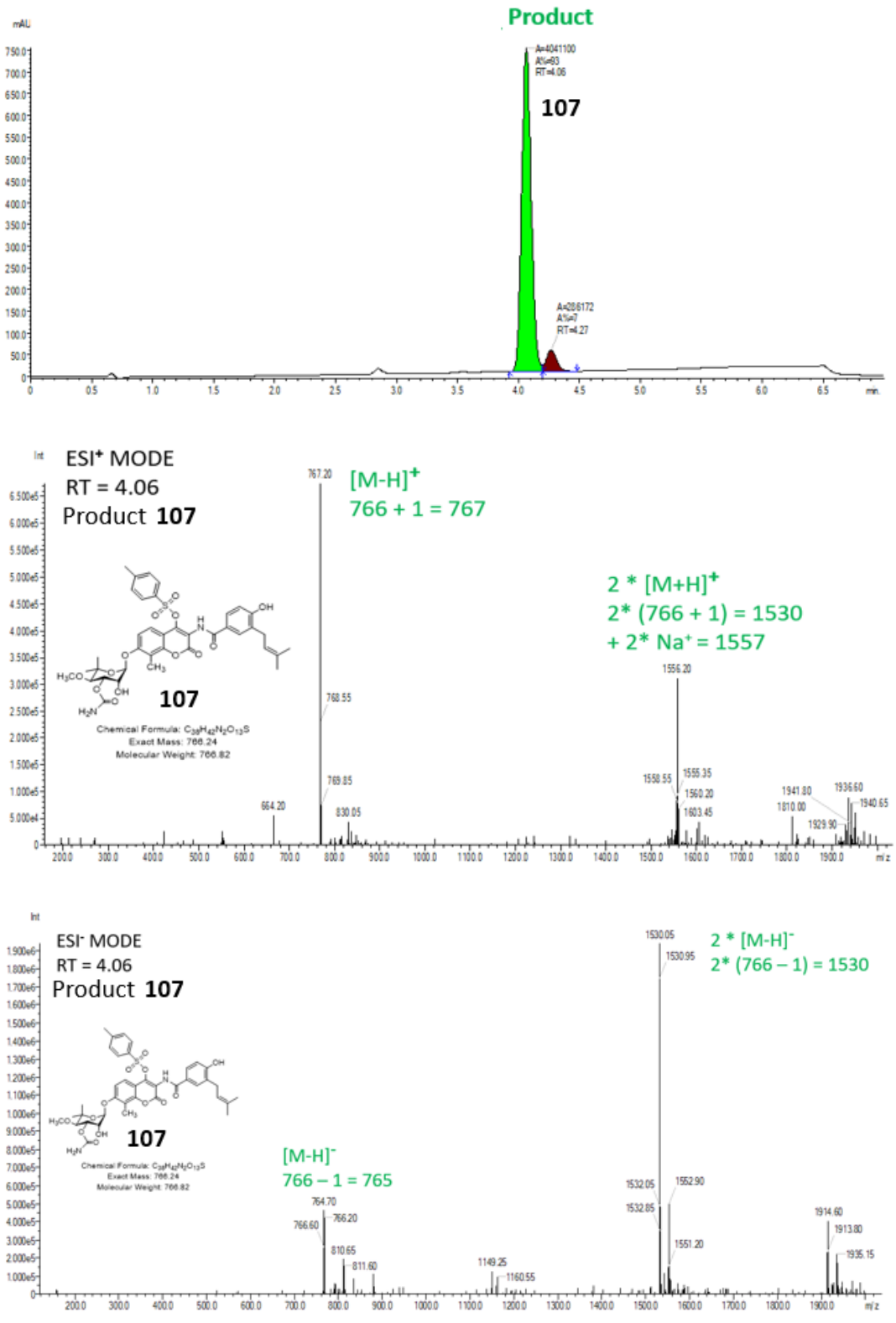
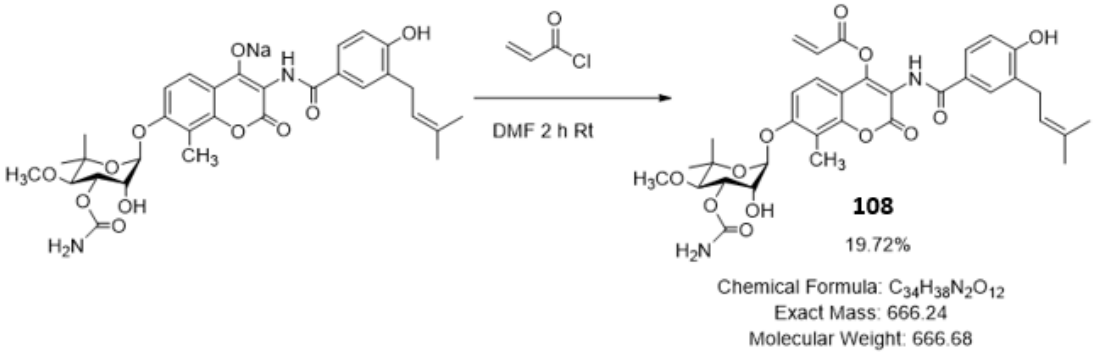


Figure 100: LC-MS of purified **107**. ESI^+ mode found at *m/z* = 767, ESI^- found at *m/z* = 764.70 (approx. *m/z* = 765).

3.3.4 Synthesis of 4'-acrylate-novobiocin (108)



Scheme 35: Synthesis of 4'-acrylate-novobiocin **108**.

The synthesis followed the 'standard process' (p142). As a result of the fact that acryloyl chloride is more reactive than tosyl chloride, the flask for the reaction was pre-dried overnight to ensure anhydrous condition. The first batch of synthesis showed a distinct LC-MS result, of which novobiocin was almost completely converted into the product **108** (Figure 101).

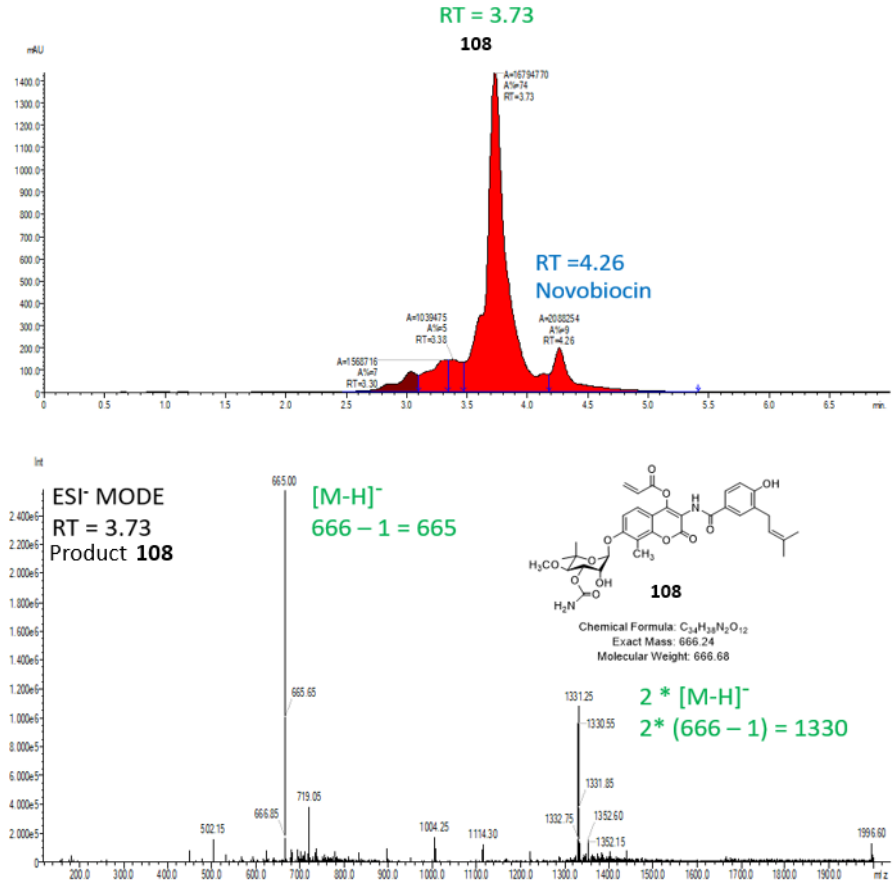


Figure 101: LC-MS result of the synthesis of 4'-acrylate-novobiocin **108**.

The 4'-acrylate-novobiocin product **108** peak was observed in the LC trace at RT = 3.73 min. ESI$^-$ mass spec analysis of the peak confirmed its identity of the desired product (Figure. 101). The crude product was purified by Biotage method 3 (Figure. 102) to achieve a colourless solid in a yield of 42%.

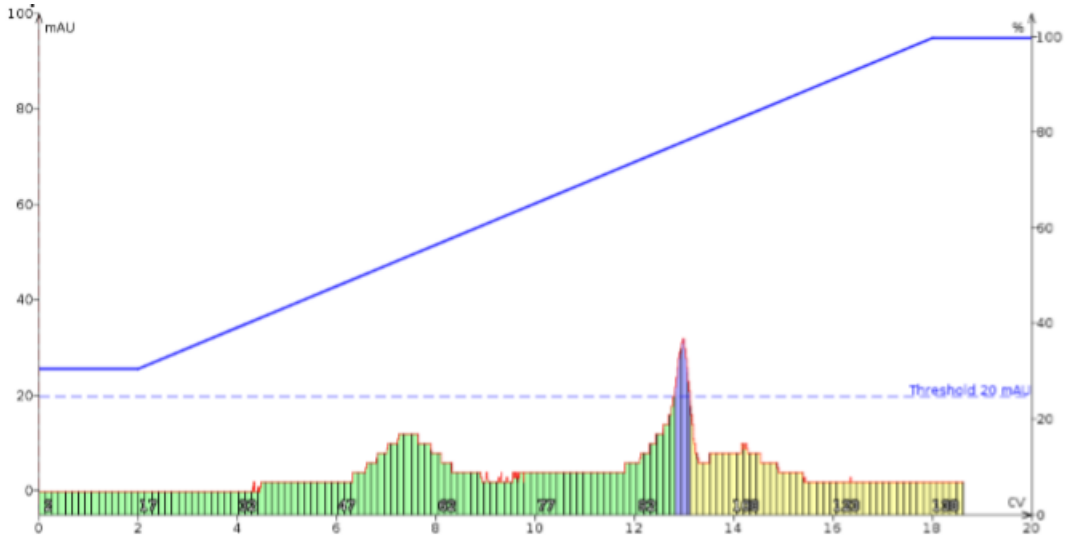


Figure 102: **108** crude products from the first batch, purified by Biotage method 3.

Prue product fraction marked in blue in Figure 102 was collected and tested for its purity by LC-MS (Figure. 103):

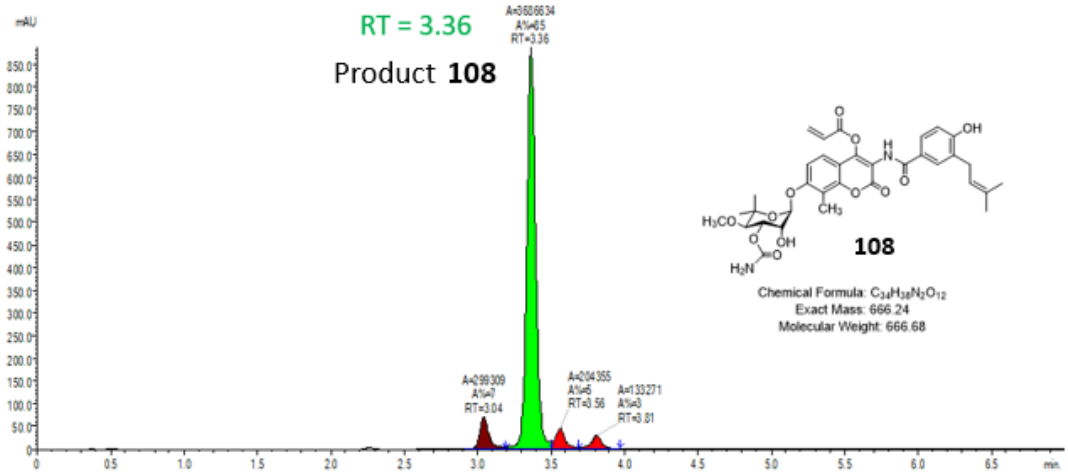


Figure 103: LC-results of purified compound **108**, which showed good purity.

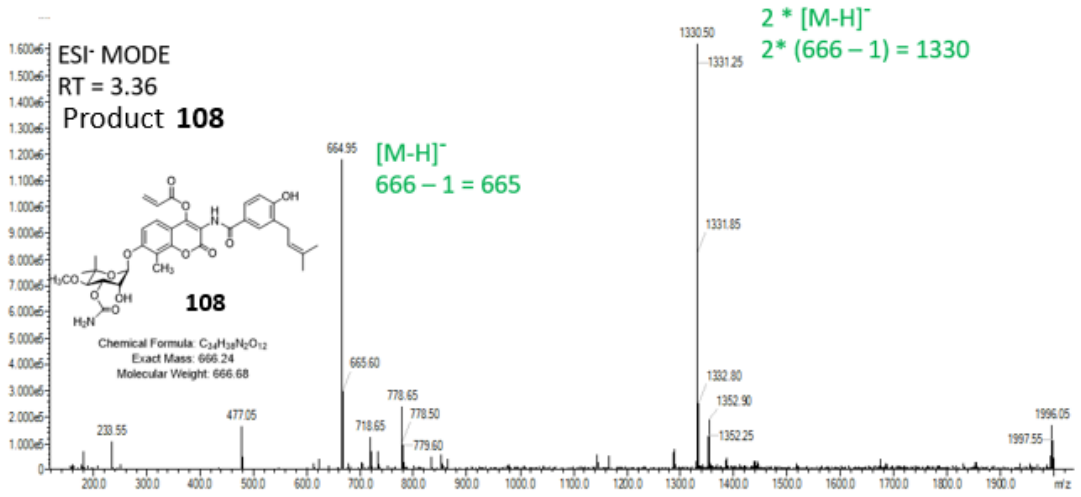


Figure 104: MS result of pure **108** peak in Figure 103, RT = 3.36 min.

There was a problem with the RT of the product which was changed from RT = 3.73 min (Figure 101) to RT = 3.36 min (Figure. 103). We analyzed the peak RT = 3.81 min in Figure 103 and found out it was the peak of novobiocin **0** (Figure 105, also changed from RT = 4.26 min in Figure 101 to RT = 3.81 min in Figure 103)

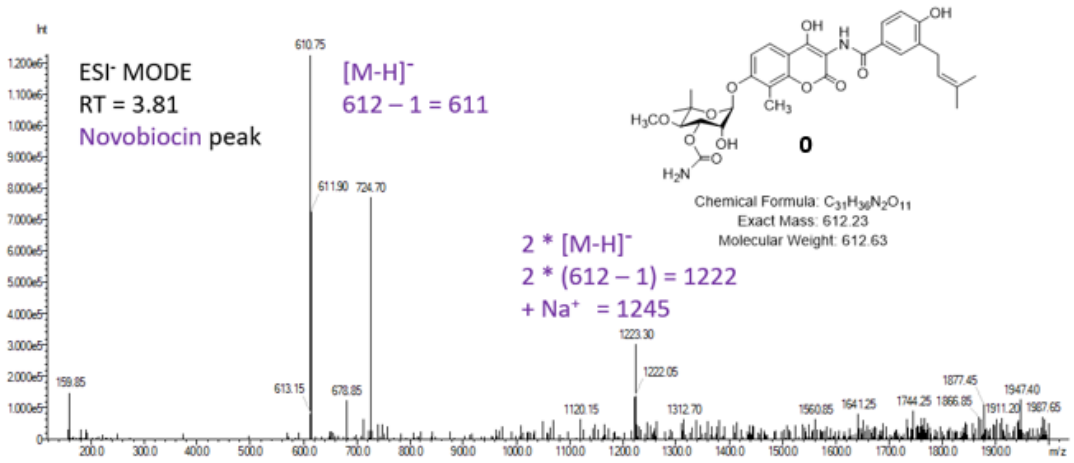


Figure 105: MS result of peak RT = 3.81 in Figure 103, refers to traceable novobiocin in purified **108** sample.

The synthesis of **108** proved challenging to repeat. Initially, the reaction shown in Figure. 101 was carried out in December. However, the figure below (Figure. 106) is the LC-MS result of the same **108** synthesis carried out in May/June. In comparison, it was about 15 °C in December, while it was > 25 °C in

summer. Ice cooling was therefore applied to optimize the yield at 15 °C, which was the optimal temperature for this reaction.

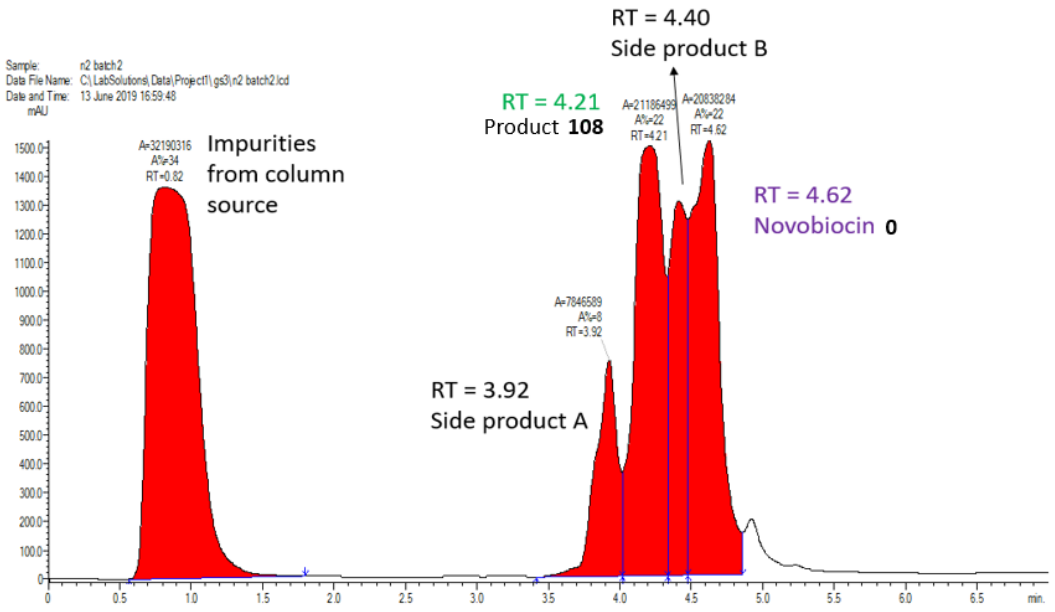


Figure 106: LC result of batch 2 **108** synthesis, temperature for batch 2 is 25 °C.

Herein, as mentioned before, all the RT of components gained an increased RT due to decreased column pressure, presenting an incomplete separation in the LC machine.

Product peak (in Figure 106) was found at RT = 4.21 min (MS trace see Figure. 107):

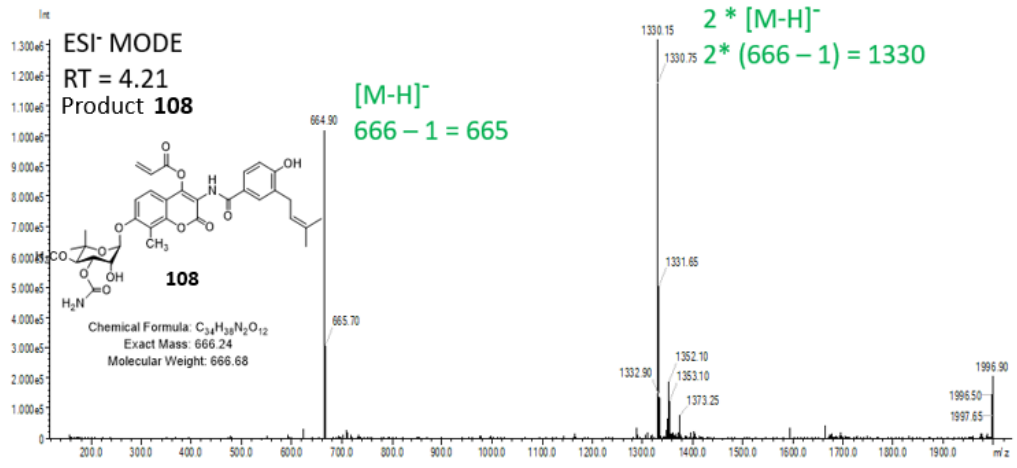


Figure 107: LC-MS of peak RT = 4.21 min in Figure 106 is product **108**

Starting material **0** peak (in Figure 106) was found at RT = 4.62 min (MS trace see Figure. 108):

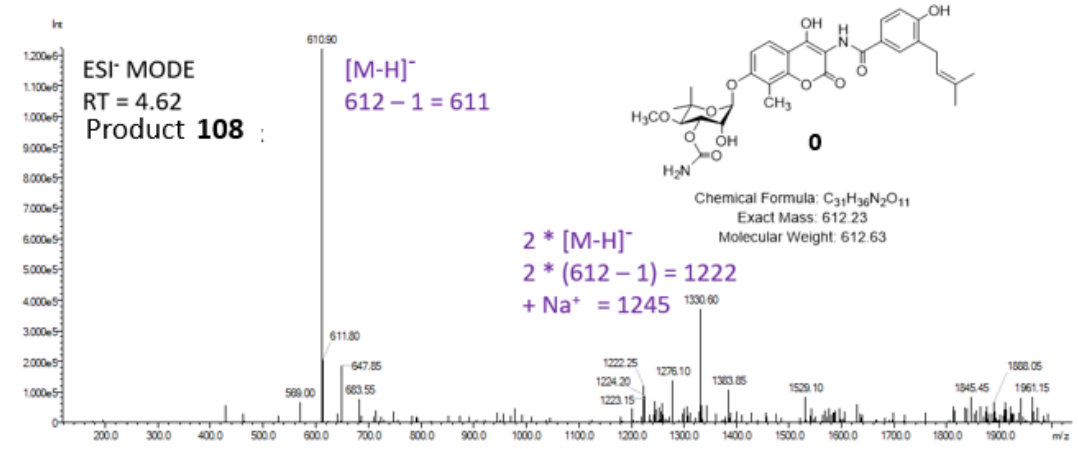


Figure 108: LC-MS of peak RT = 4.62 min in Figure 106 is characterized to be starting material **0**.

There were two types of side products found in this particular reaction. RT = 4.21 min in Figure 106 was found to be the isomer of **108**. The structure showed in Figure. 109 is believed to be the 5-OH substituted isomer as an example.

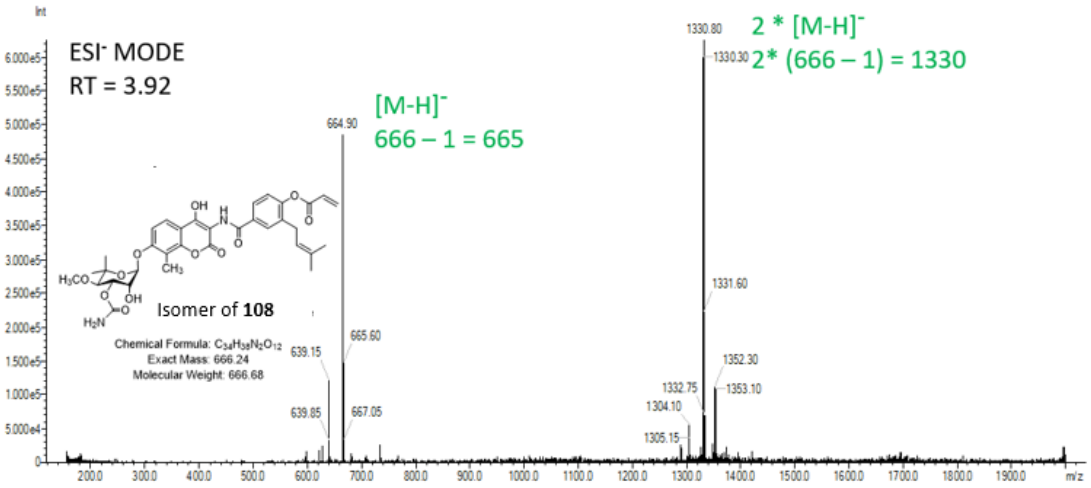


Figure 109: LC-MS of peak RT = 4.21 min in Figure 106 is believed to be a 5-OH substituted isomer of 108.

RT = 4.40 in Figure 106 referred to di-acrylate substituted novobiocin (Figure. 110):

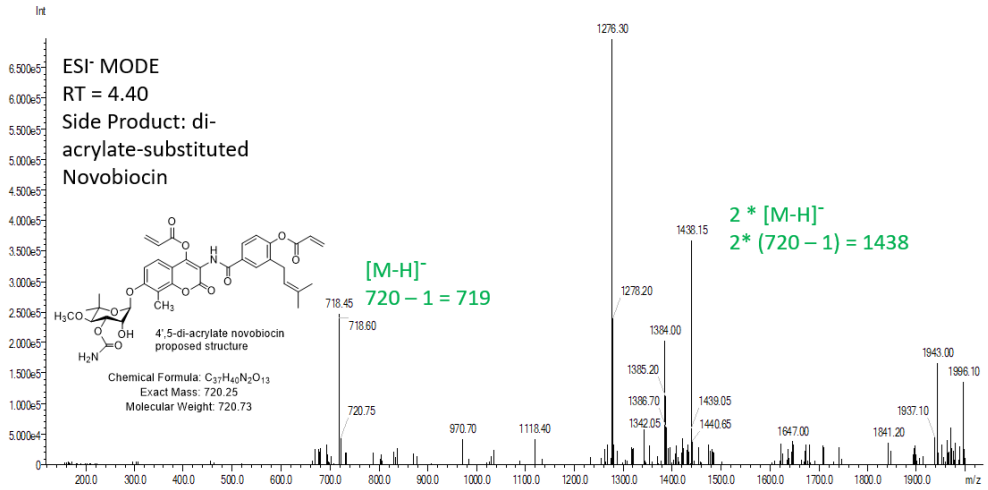


Figure 110: LC-MS of peak RT = 4.40 min in Figure 106 was characterised to be di-acrylate-substituted novobiocin.

As a result, the temperature significantly influenced the outcome of the reaction. The DMF controlled regio-selective reaction would be less effective under increased temperature, generating isomer, and di-substituted products. Therefore, to minimize the formation of side products, the reaction should not continue for more than 2 hours, which is a reduction from 6 hours compared to the previous experiments

The reaction mixture was directly purified by Biotage method 3 (Figure. 111):

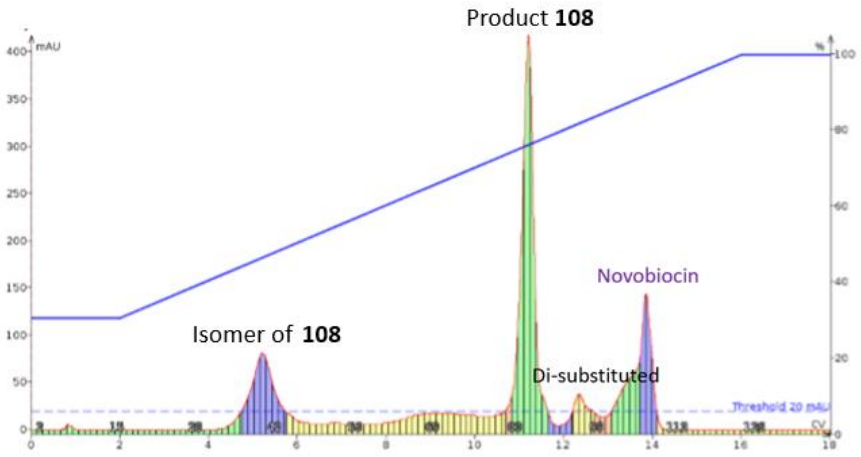


Figure 111: Biotage purification result of the crude product shown in Figure 106. Product 108 came out from the column at approx. 11 CV.

The UV absorption pattern (Figure 111) was the same as the LC-MS (Figure 106); both showing 4 peaks. The highest green fraction was collected and sent for LC-MS, which confirmed the identity of product **108**, with 14% isomer as impurities.

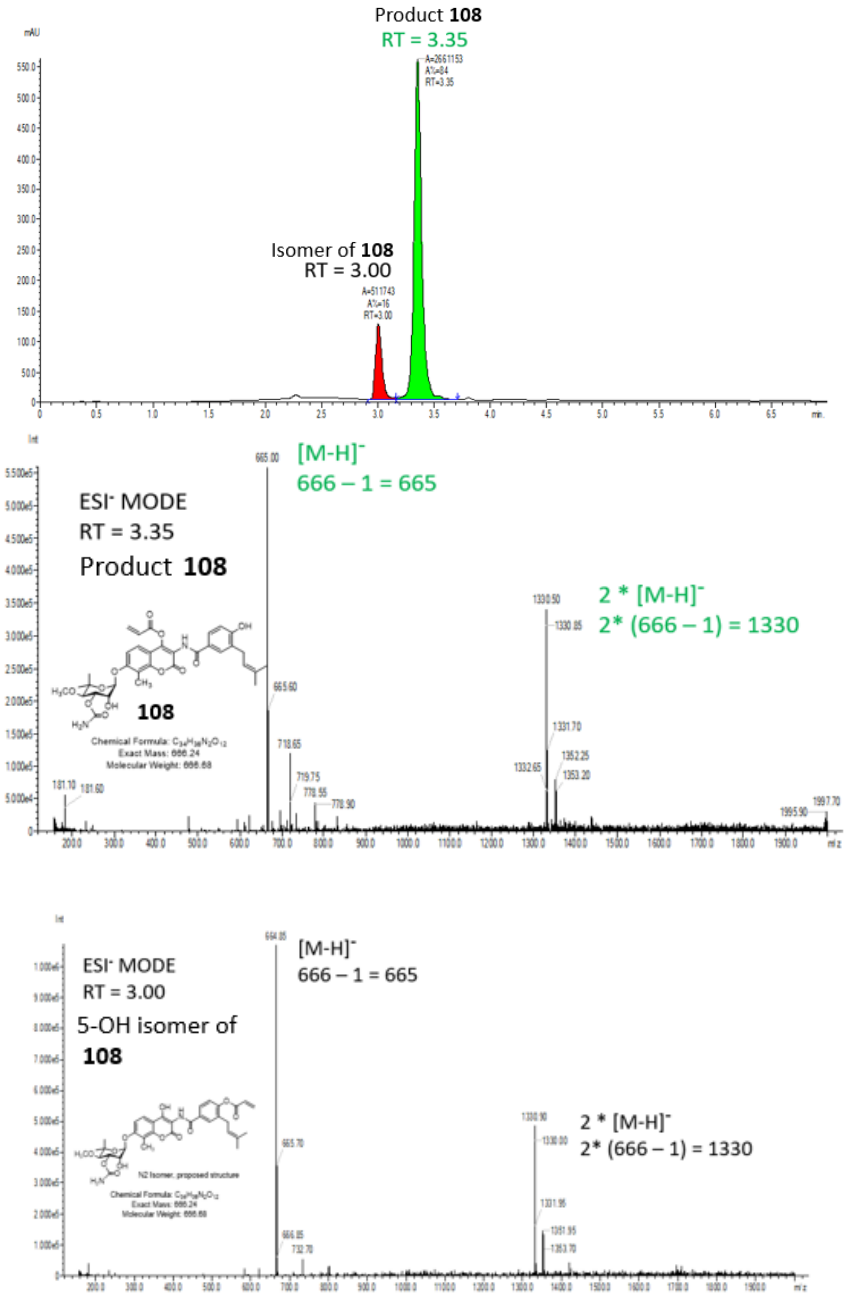
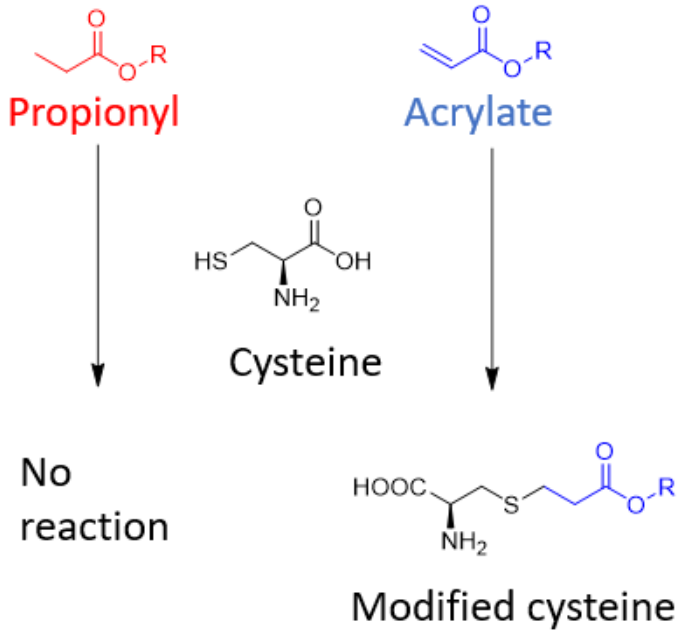


Figure 112: Biotage purified **108** from the crude product shown in Figure 106.

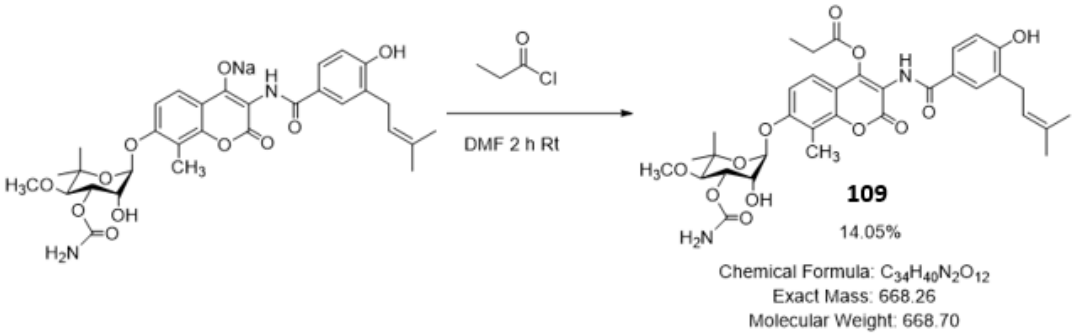
Again, the RT of product and side product was changed due to abnormal column pressure of LC-MS. Obviously, the reverse-phase gradient method adopted by Biotage performed much better than LC-MS.

3.3.5 Synthesis of 4'-propionyl-novobiocin (109)

The propionyl group is the reduced form of the acrylate group, which possesses no covalent cysteine binding ability (Scheme 36). Herein, the compound 4'-propionyl-novobiocin was synthesized (compound **109**) to compare the biological property with cysteine-binding active form **108** (Scheme 37). If **108** does covalently modify the cysteine residue, there should be a difference of results in terms of the antiproliferative activity, binding affinity amongst other data.

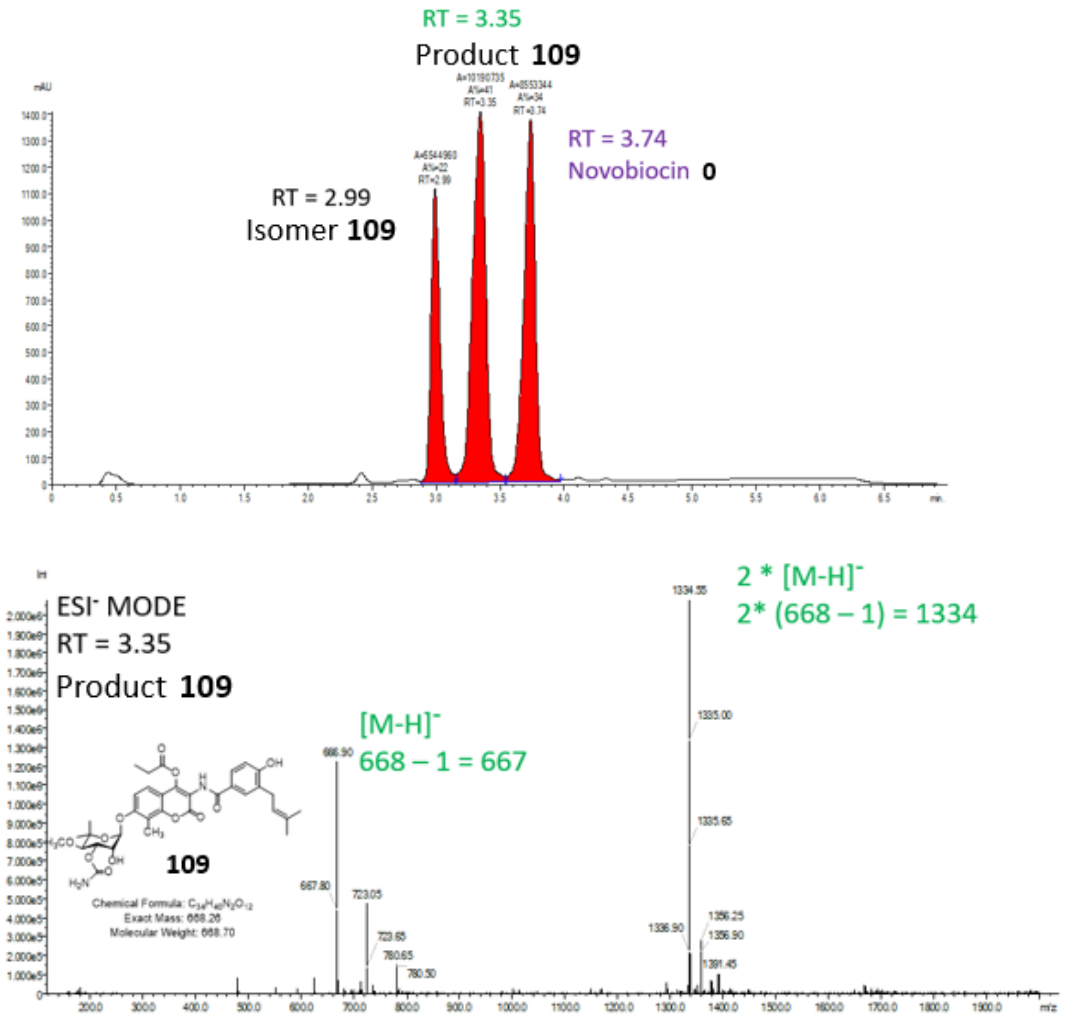


Scheme 36: Acrylate is a cysteine modifier (right blue) while its reduced form propionyl (red left) is not.



Scheme 37: Synthesis of 4'-propionyl-novobiocin **109**.

The reaction followed the "standard process" with a duration of 2 hours. An LC-MS of the 2 hours' reaction is given below (Figure. 113):



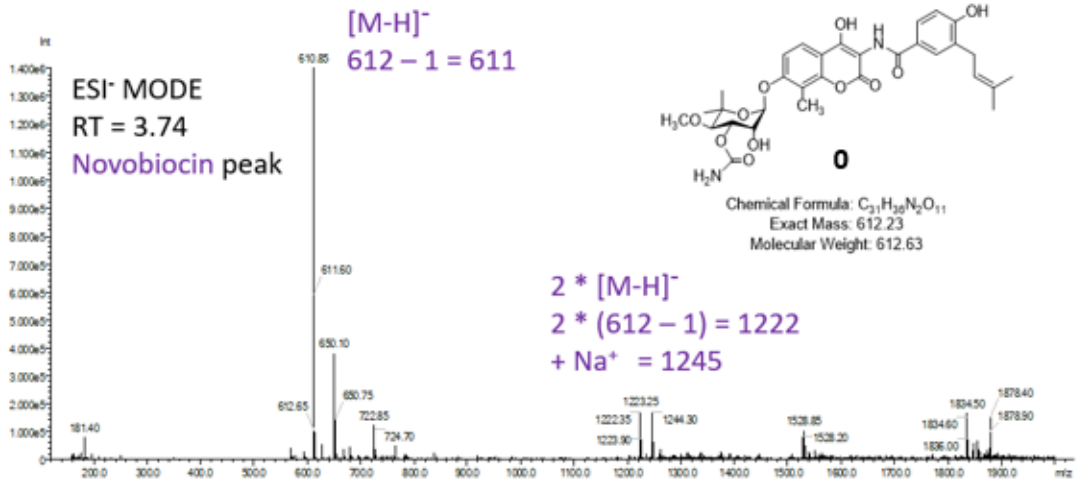


Figure 113: LC-MS result of the crude product sampled from the reaction shown in Scheme 37. Peak RT = 3.35 was identified as **109.**

Product **109** was found at RT = 3.35 min, the same as the RT of **108** (RT = 3.35 min in Fig. 112). Novobiocin showed a RT = 3.74 min under this column status (Figure 113), different from RT = 3.81 min (in Figure 105) and RT = 4.62 min (in Figure 108) observed previously.

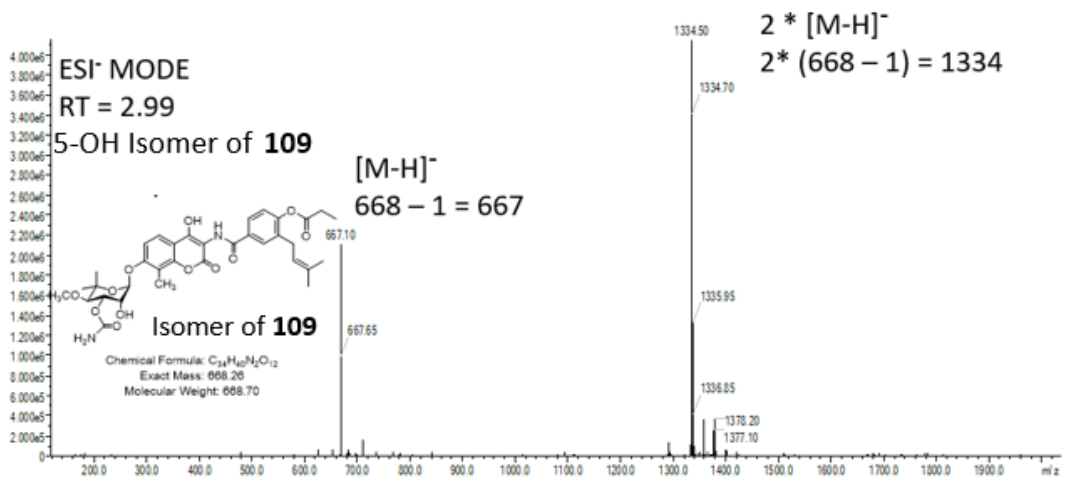


Figure 114: MS result of peak RT = 2.99 min in LC trace Figure 113. It is believed to be 5-OH isomer of **109**.

For the **109 isomer** side product, it also showed the same RT as **108 isomer**, at RT = 3.00. The synthesis of **109** generated more side products (22% > 8%) than what was observed in the synthesis of **108**. There was also more unreacted novobiocin left in the system (34% in **109** synthesis > 22% in **108** synthesis).

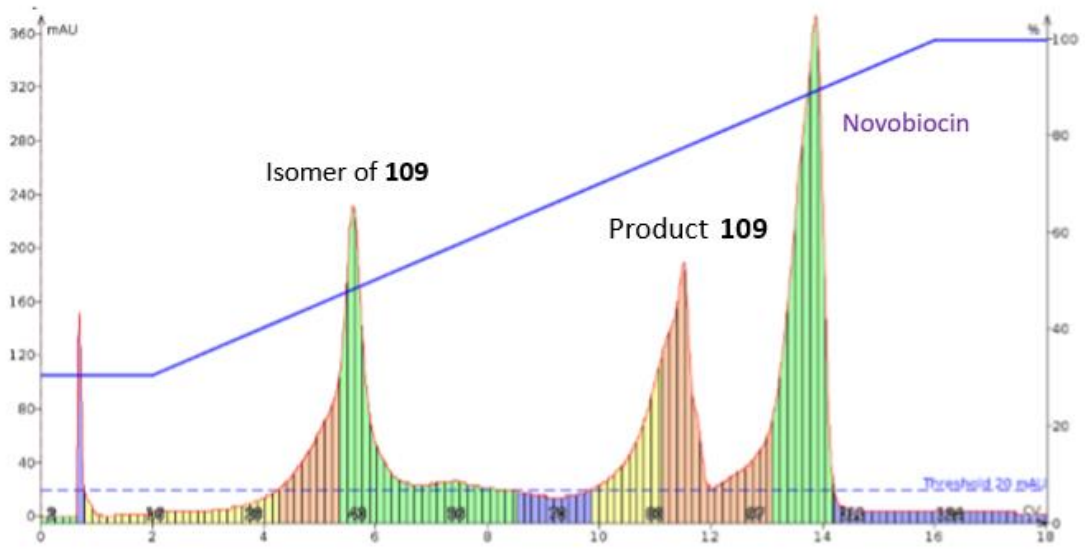


Figure 115: Biotage purification of crude product **109.** Product 109 was found at CV 11.

Crude product **109** was purified by Biotage method 2 (Figure 115) to produce a white powder with a yield of 14.05%. Although the yield was not ideal, it was enough for bio-assay tests. A LC-MS result of purified **109** is given below (Figure. 116):

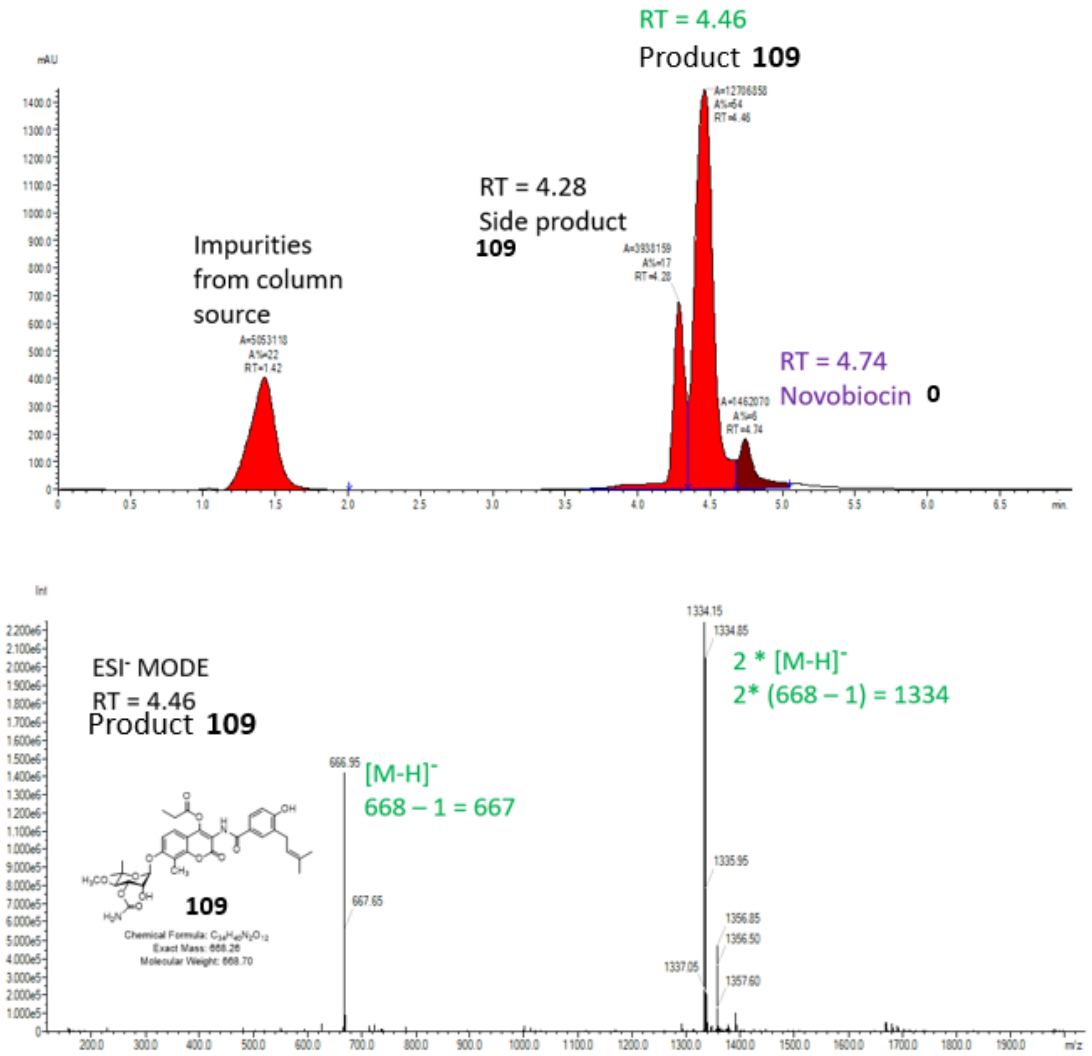


Figure 116: Biotage purified **109**.

Note that the column pressure of LC-MS dropped again, and the RT of **109** changed from 3.35 min in Figure 113 to 4.46 min in Figure 116.

3.3.6 Synthesis of 4'–crotonyl–novobiocin (110) and synthesis of 4'-methacryl-novobiocin (111).

Crotonyl **((2*E*)-but-2-enoic,** Figure 117**)** and methylacryl (**2-Methyl-2-propenoic,** Figure 117) are isomers with substitution from H to CH_3 on the α proton and β proton respectively (Figure. 117). These analogues aimed to explore the SAR of 4' substitution of novobiocin, exploring possible steric effects.

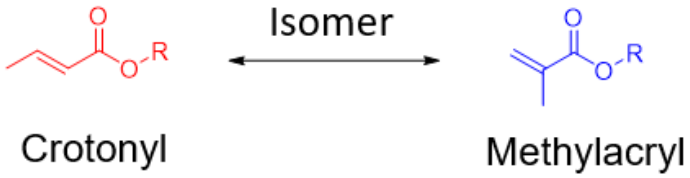
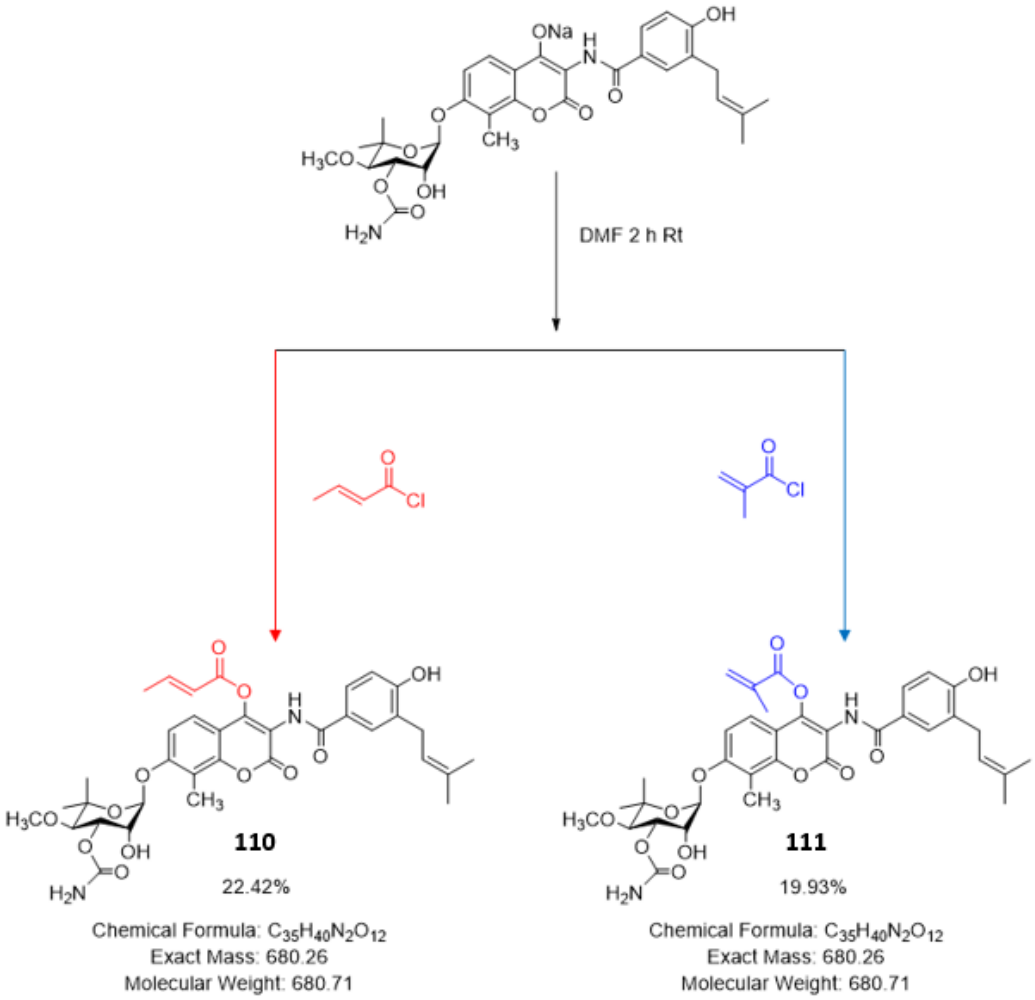


Figure 117: Crotonyl (red left) and methylacryl group (blue right) are isomers.



Scheme 38: Synthesis of **110** and **111.**

After two hours, the solution was directly sent to Biotage to be purified by Biotage method 3. From the UV result of Biotage (Figure 118), it can be concluded that the reaction has a pattern very similar to the synthesis of **108** and **109**. Products peak was suspected to be the one that came out at 13 CV (Figure. 118).

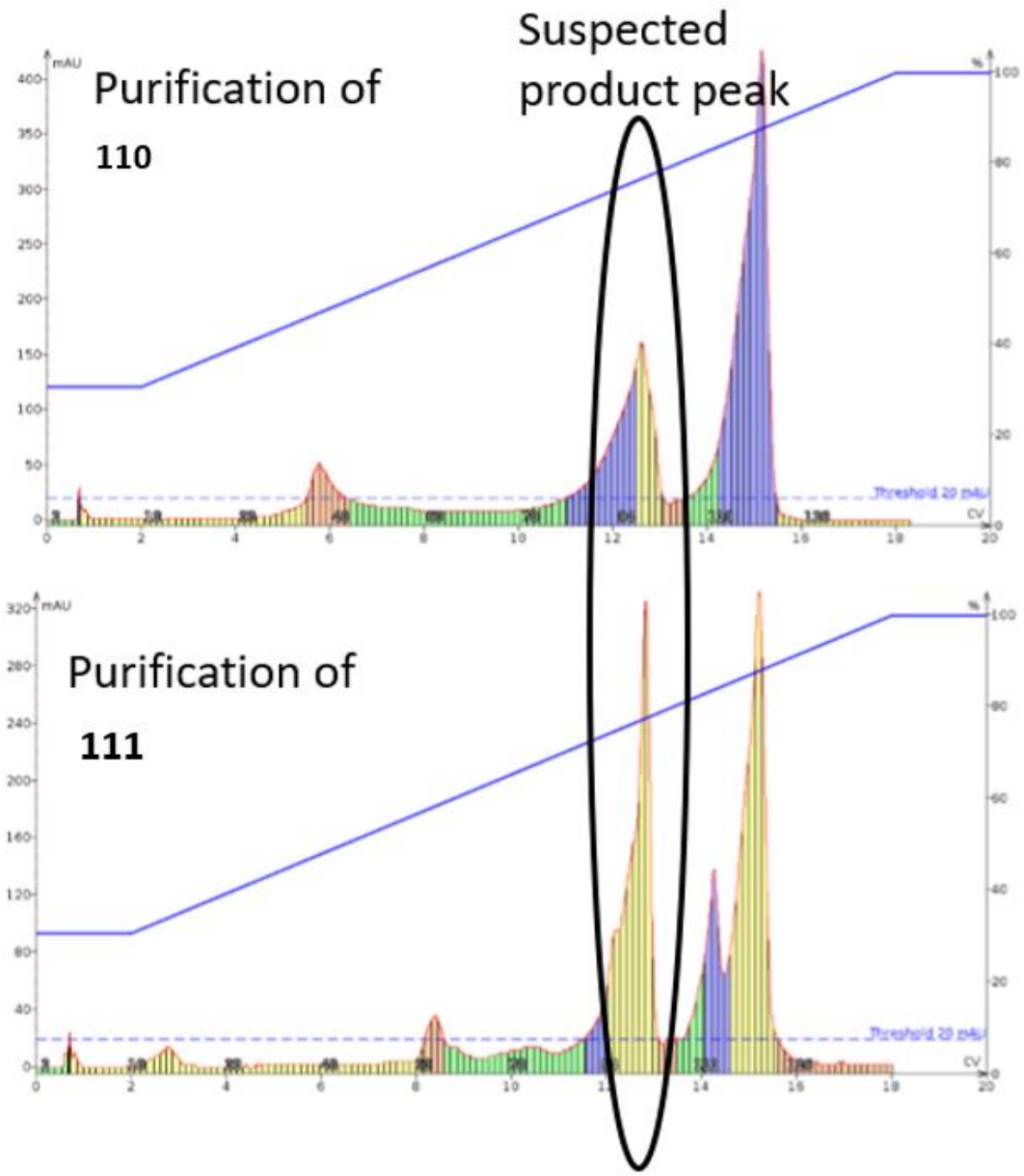


Figure 118: Biotage purification results of the crude product of **110** and **111**.

Relevant fractions of products were collected and sent for MS analysis, and the result showed that it is the product (Figure. 119). The yield was calculated for **110** and **111** as 22.42% and 19.93%, respectively.

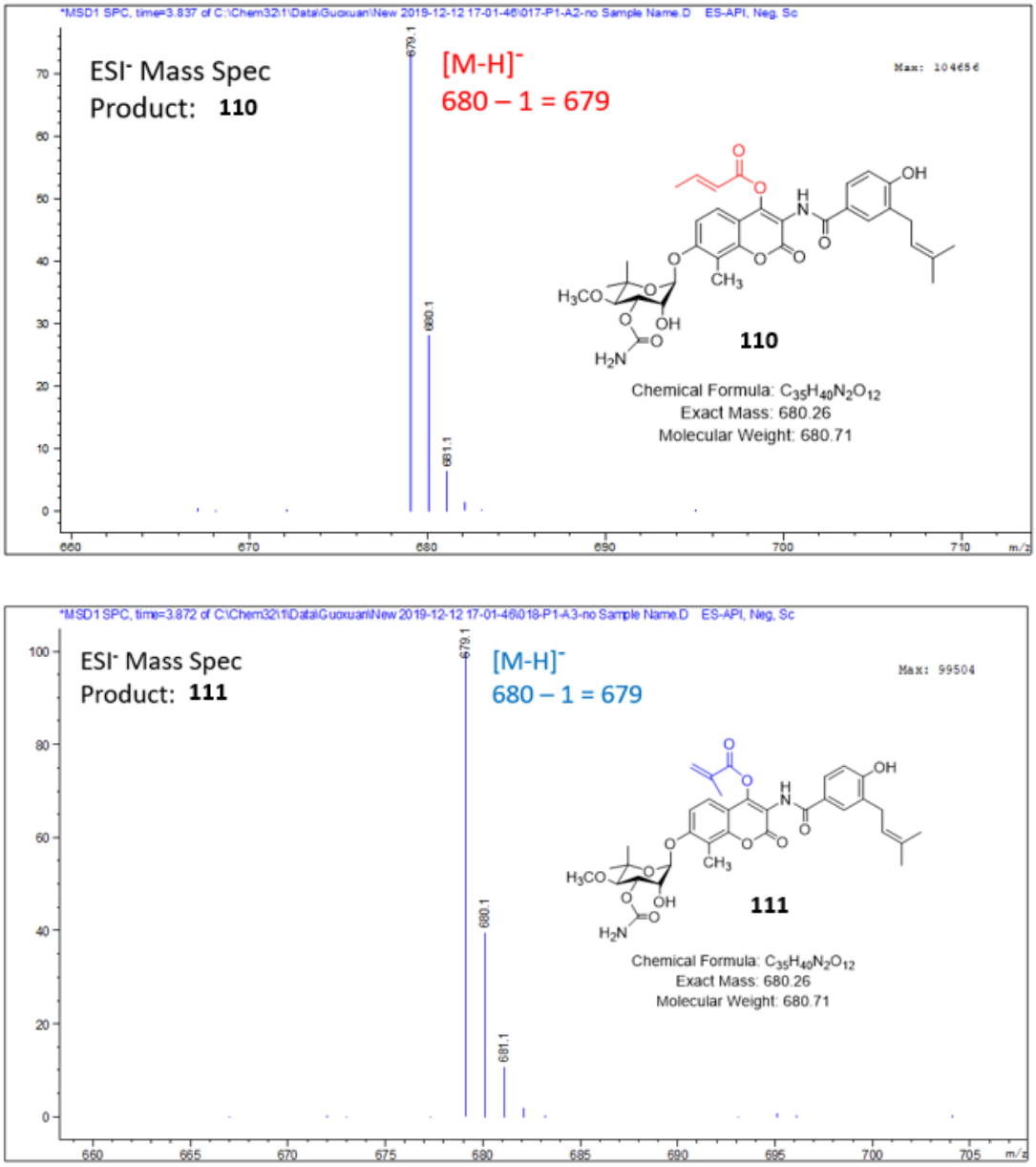
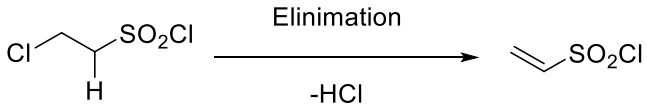


Figure 119: MS data of **110** and **111**.

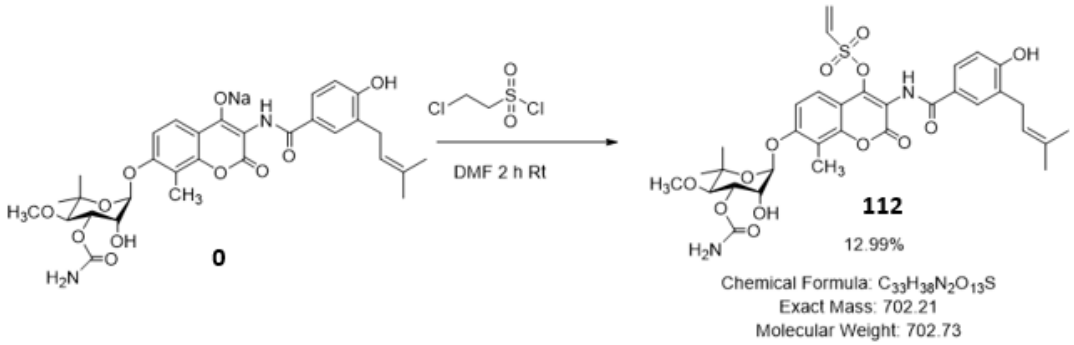
3.3.7 Synthesis of 4'–ethenesulfonyl–novobiocin (112)

Vinyl sulfone is a highly reactive group that can act as an electrophile. It is a well-known type of Michael acceptor and served as key structural units of many biologically active compounds as well as versatile building blocks for various organic transformations (Schneider, Rieger et al. 2015). In this synthesis, commercial ethenesulfonyl chloride was not available due to its high reactivity with water or moisture. A commercial protected form 2-chloro-1-ethanesulfonyl chloride was therefore used. The masked protection will be lost by elimination of HCl to transform it into its active form ethenesulfonyl chloride (Scheme 39).



Scheme 39: 2-Chloro-1-ethanesulfonyl chloride lost HCl in the solvent to form the active ethenesulfonyl chloride

This process was reported to happen naturally in solvent and can be catalyzed by base (Moriarty and Tyagi 2010). However, because the induction of base could deprotonate 5-OH on the novobiocin amide side chain, the base was not used in this reaction to ensure the DMF catalyzed regioselectivity (Scheme 40).

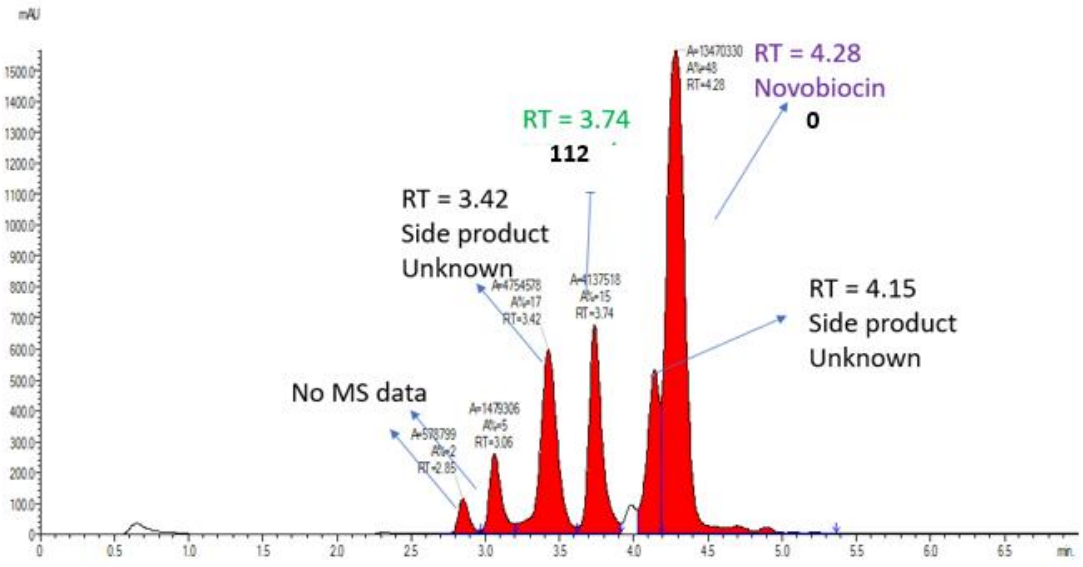


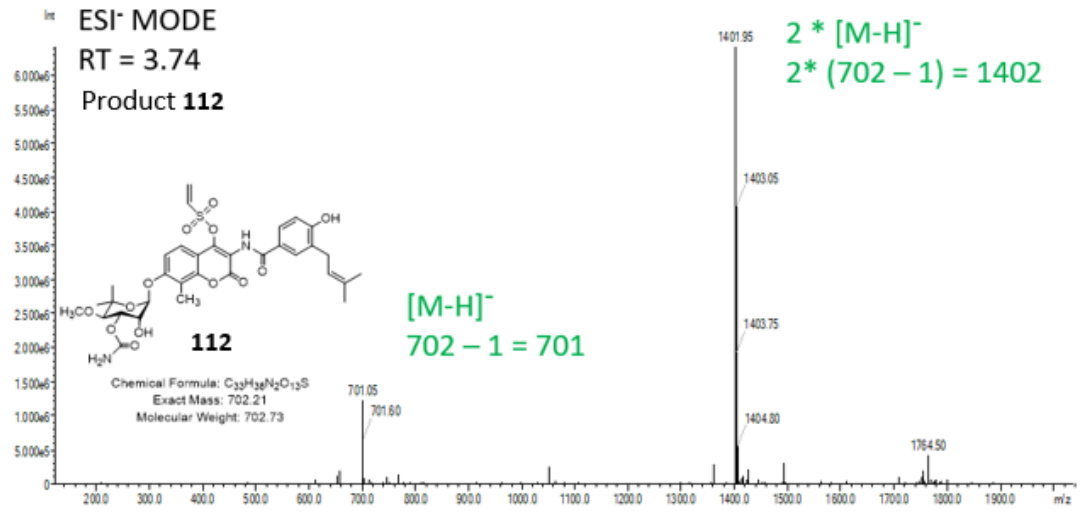
Scheme 40: Synthesis of 4'-ethenesulfonyl-novobiocin **112.**

The reaction was prepared following the standard procedure (Scheme 40), with a duration of 2 hours. In-process control LC-MS (in Figure 120) indicated that this reaction generated multiple side products. Product peak was found at

RT = 3.74 min, and novobiocin was found at RT = 4.28 min, validated by MS. Peak RT = 2.85 min and 3.06 min were unrecognizable as there were too little products in the sample for reasonable ionization (Figure. 120). Peak RT = 3.42 min represented a side product with a mass of *m/z* = 640 (Figure. 121), which was not the isomer side product of **112**. It remained unsolved.

Peak RT = 4.15 min in Figure 120 was solved to be novobiocin lost an amine on noviose sugar and rearranged into carbonate (Figure. 121).





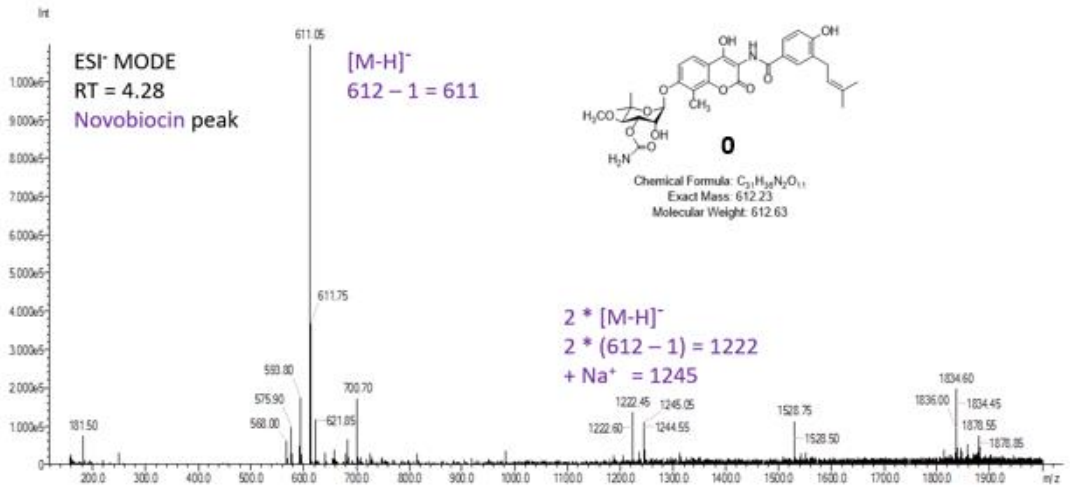


Figure 120: LC-MS result of the crude product sample from the reaction shown in Scheme 40.

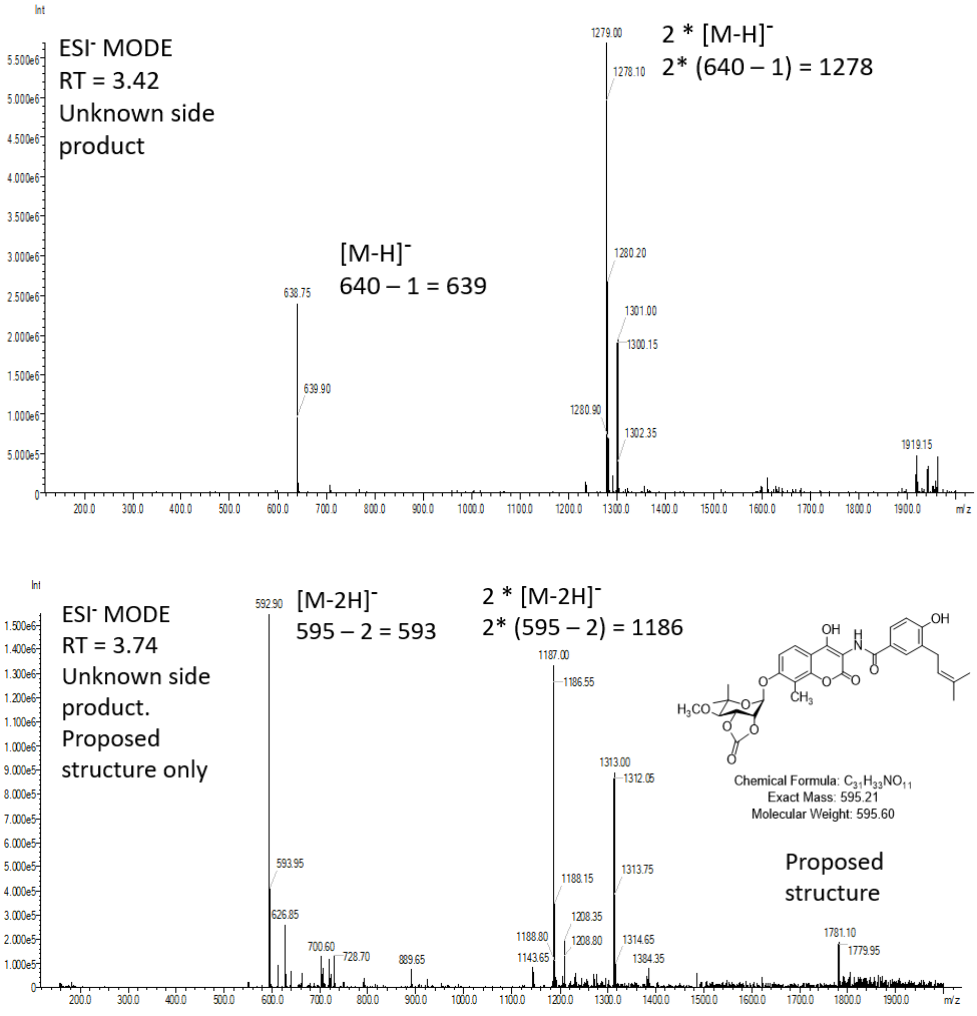


Figure 121: MS result of side products in the synthesis of **112,** refer to LC in Figure 120.

The reaction mixture shown by Figure 120 was purified by Biotage method 2 (Figure. 122) to afford a white solid with a 12.99% yield.

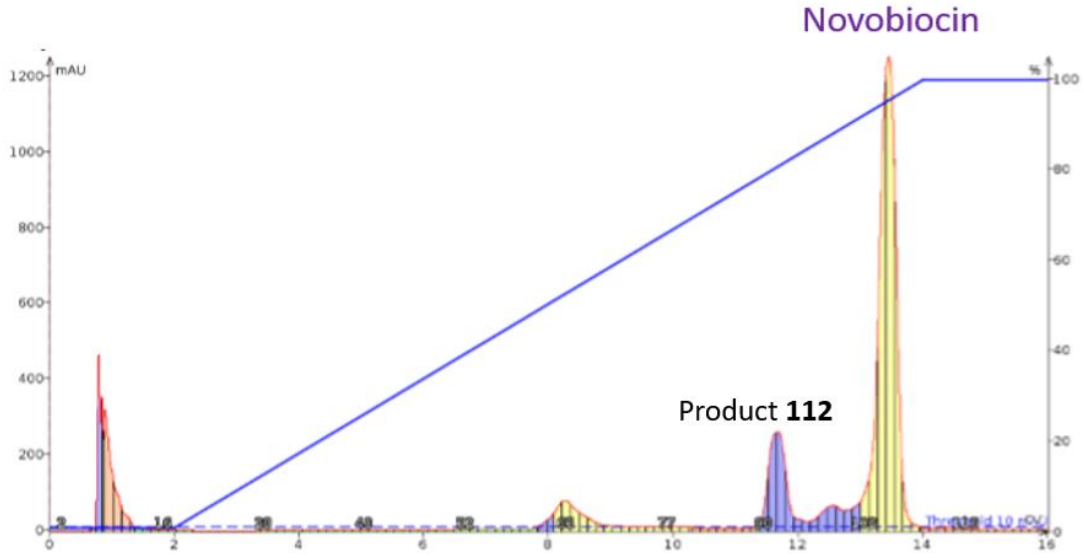


Figure 122: Biotage purification of **112**. Product 112 fraction came out at approx. CV 12.

LC-MS for the purified product **112** (which is the fraction came out at CV 12 in Figure 122) is given below (Figure. 123):

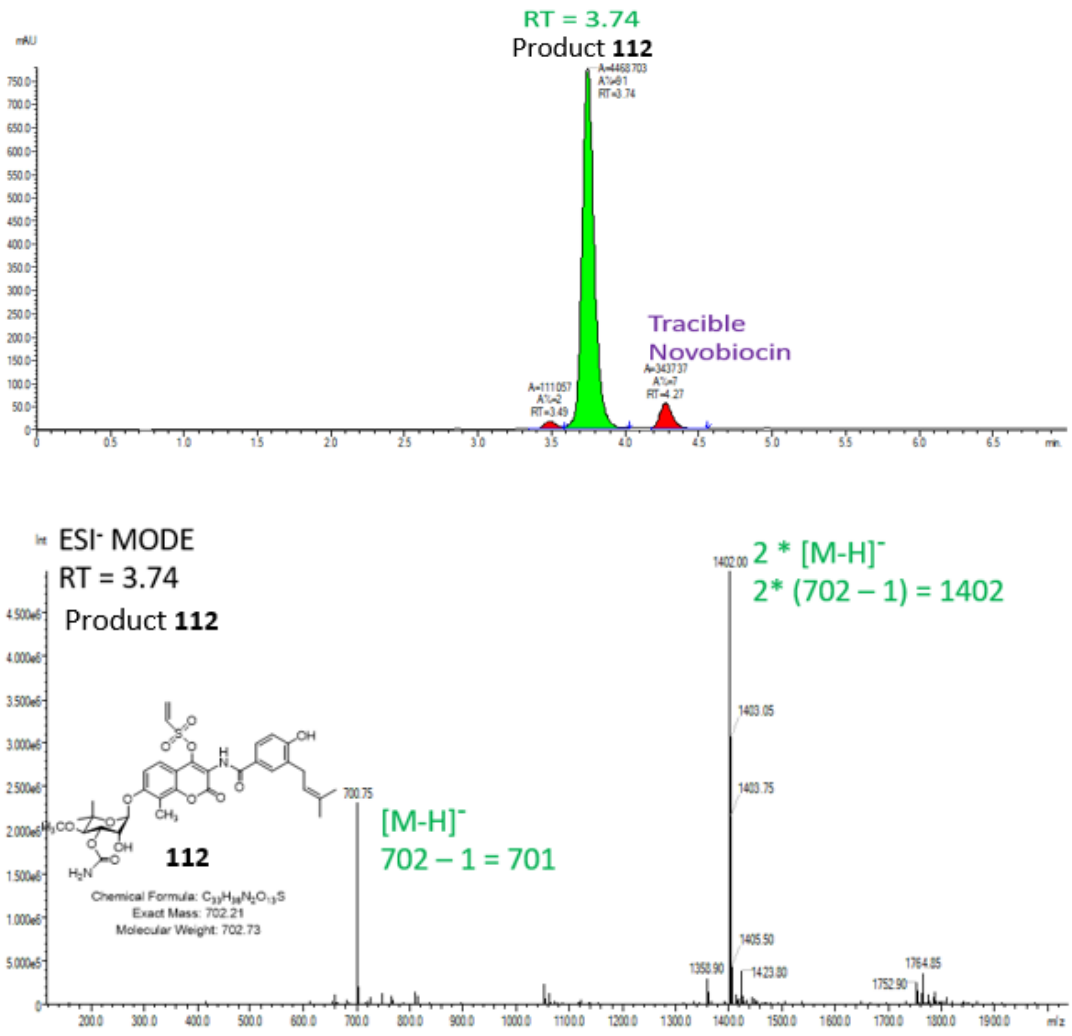
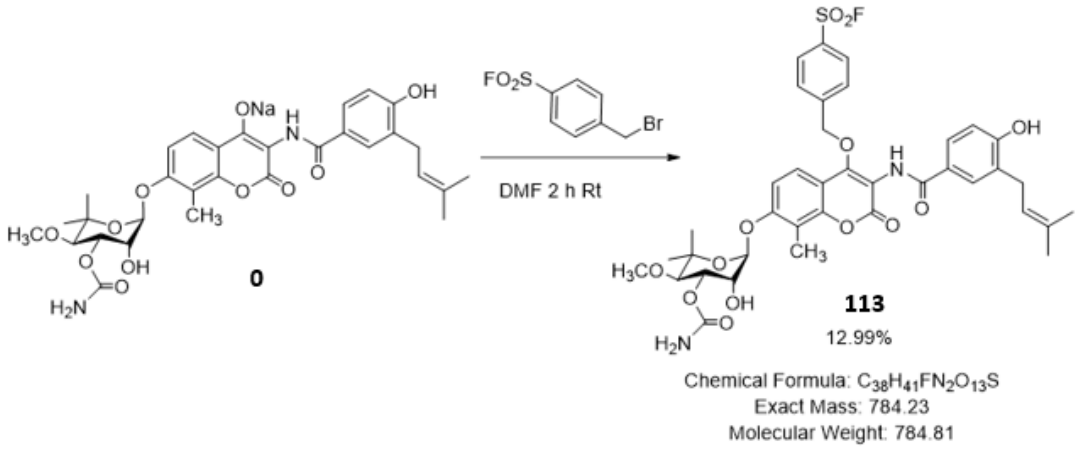


Figure 123: LC-MS of purified **112** from the crude product shown in Figure 120.

3.3.8 Synthesis of 4'–para-benzylsulfonyl fluoride–novobiocin (113)

The aim and the benefit of sulfonyl fluoride in biological research have been reviewed in **3.3.1** (p129). Herein, we selected 4-(bromomethyl) benzylsulfonyl fluoride as a reagent. Unlike previous product analogues which had 4'-ester bond linker, **113** was a 4'-ether bond linked analogue. Reaction still followed the standard process with two hours' duration (Scheme 41).



Scheme 41: Synthesis of **113**.

LC of the crude product **113** is given in Figure 124, after two hours' reaction:

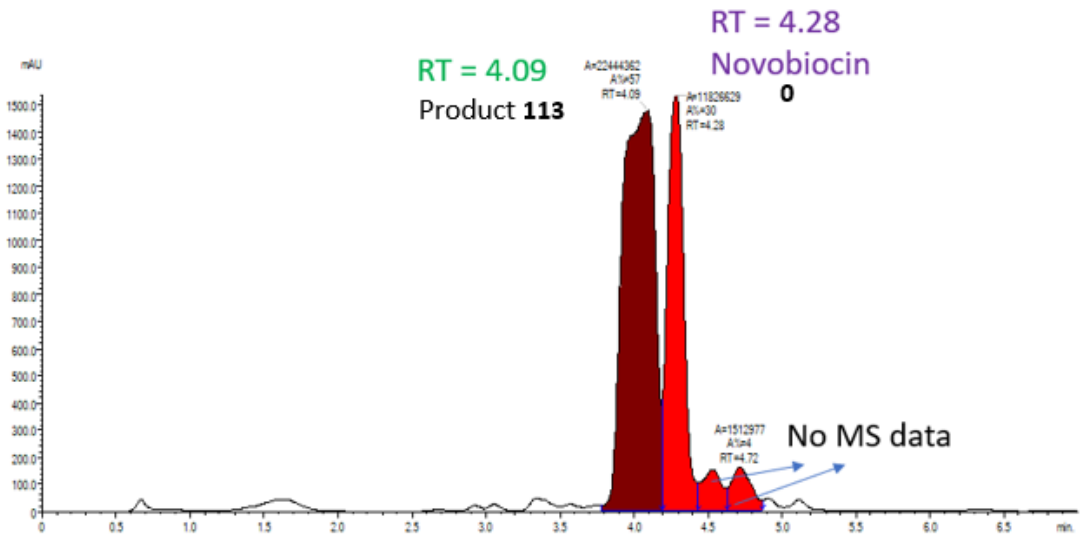


Figure 124: LC-MS results of the crude product **113** sampled from the reaction shown in Scheme 41.

As shown above (Figure 124), product **113** peak RT = 4.09 min (MS data see Figure 125) occupied 57% of UV absorption was the dominant substance inside the sample. Novobiocin, RT = 4.28 in Figure 124, left about 30% unreacted. The rest of the peaks which come after 4.5 min could not be characterized by mass spec due to poor ionization.

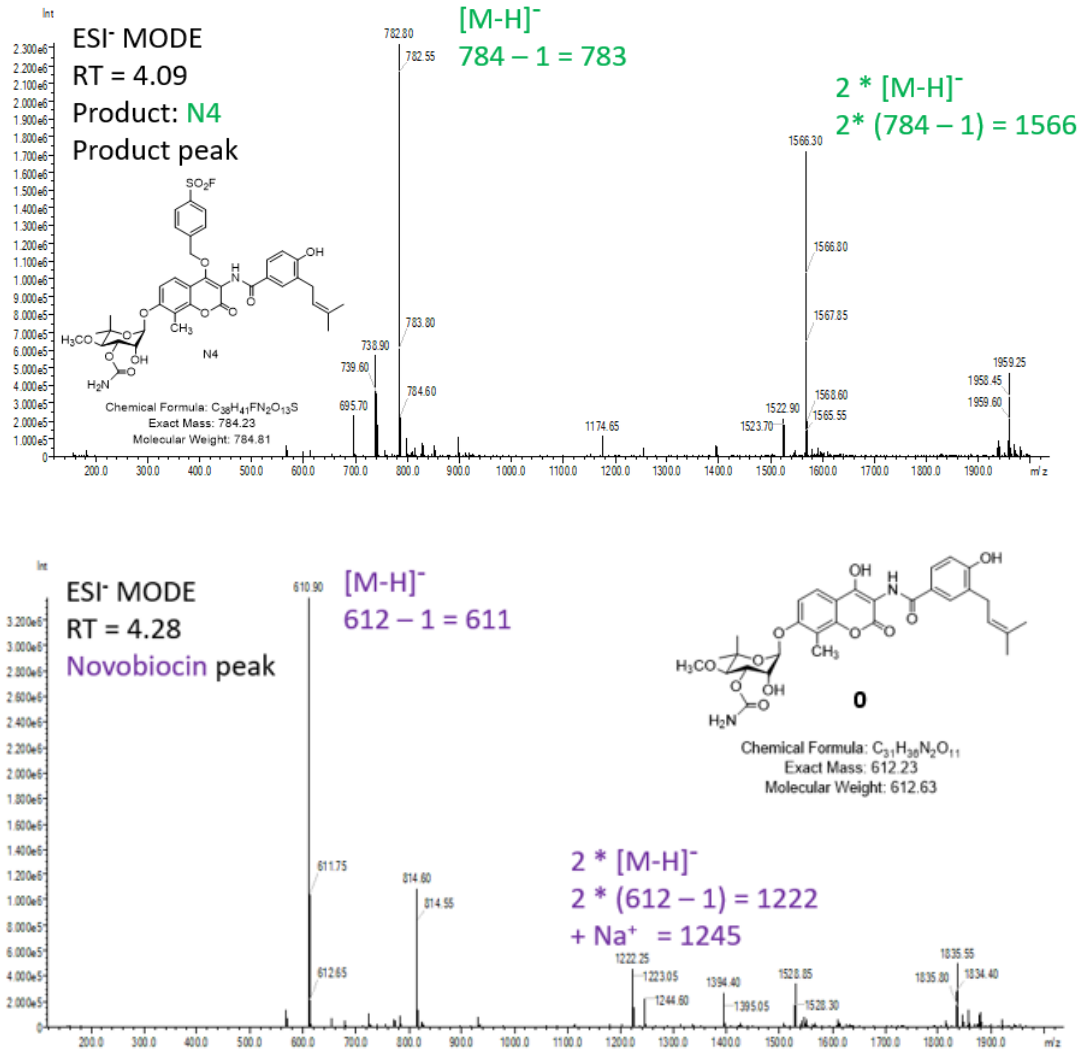


Figure 125: MS result of the LC trace shown in Figure 124. Peak at RT = 4.09 was found to be the product **113**.

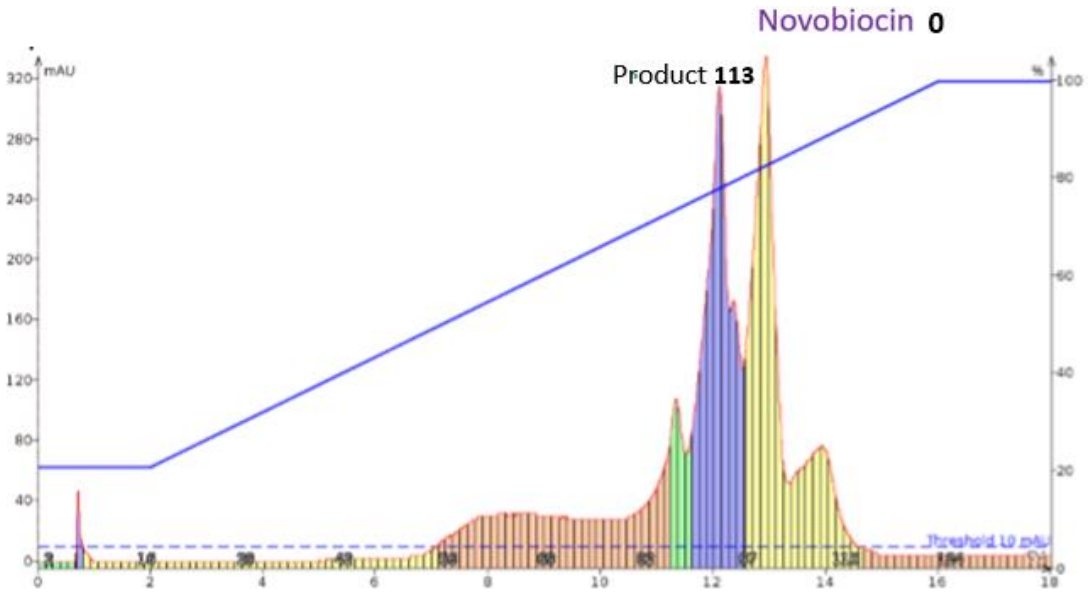
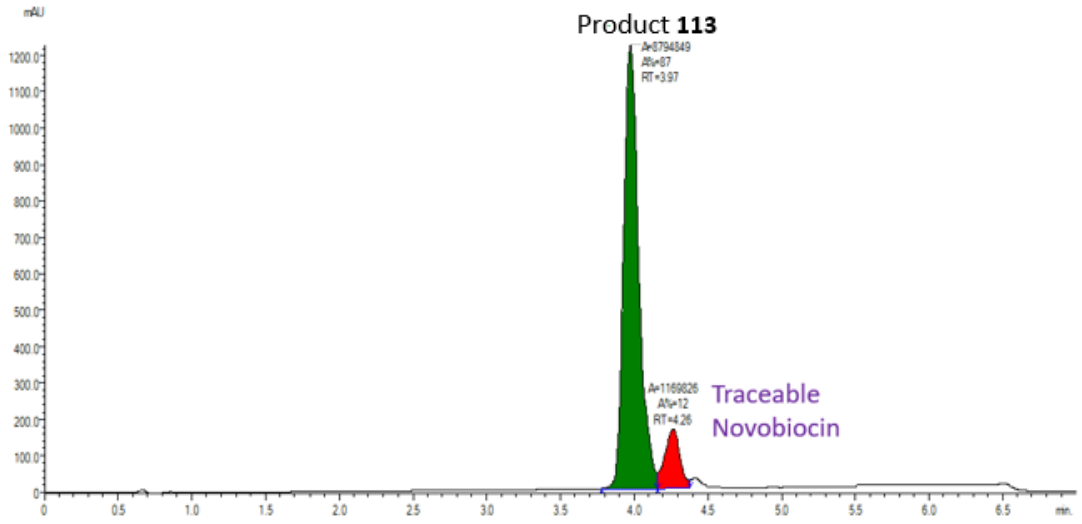


Figure 126: Biotage purification of **113.** Product **113** came out at 12 CV.

Crude product of **113** was purified by Biotage method 3 to afford a pale-yellow solid with a yield of 14% (Figure. 126). The low yield was because of only pure peak fractions were taken to ensure excellent purity of the product.

LC-MS of Purified **113** is given below (Figure. 127):



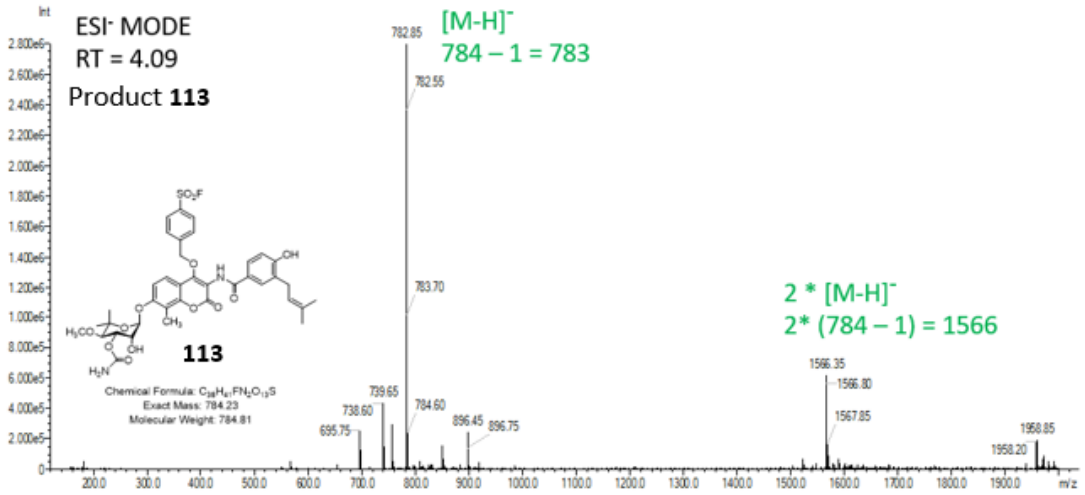
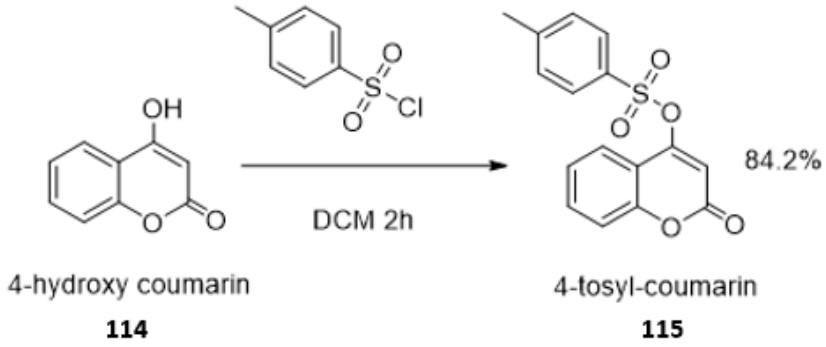


Figure 127: Biotage purified **113**, from the crude product shown in Figure 124.

A side product from this reaction was separated and named compound **120** which was found to be 2'-substituted isomer of **113**; detailed characterization and compound data of **120** can be found at the experimental part (p342).

3.3.9 Synthesis of 4-substituted-coumarin analogues.

3.3.9.1 Synthesis of 4 – tosyl – coumarin (115)



Scheme 42: Synthesis of 4-tosyl-coumarin **115**.

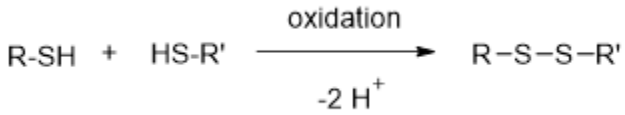
The synthesis of 4-tosyl-coumarin (**115**) used 4-hydroxy-coumarin (**114**) as a starting material, it reacts through S_N2 nucleophilic attack towards tosyl chloride. Because **114** has poor solubility in DCM while **115** was expected to

have good solubility in DCM, this phenomenon can be applied as a process indicator of reaction.

When **114** was added into DCM, it resulted in a suspension, so to ensure the completion of the reaction, the reaction was continued for a further 2 hours to afford **115** as a white solid, with a yield of 84%.

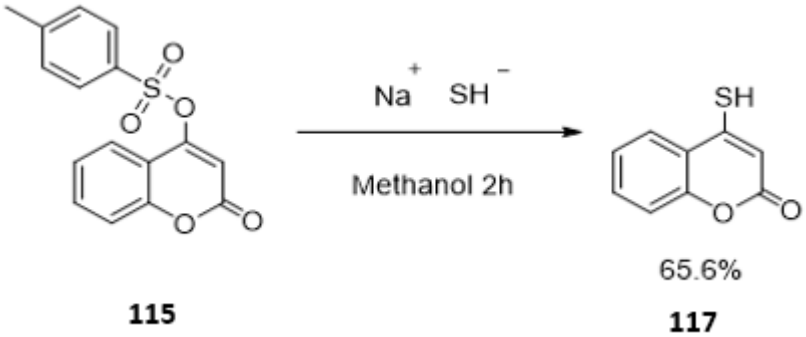
3.3.9.2 Synthesis of 4–thiol–coumarin (117)

The thiol group can act as a potential disfulfide bond donor (Scheme 43). In biology, disulfide bond bridges are formed between two cysteine molecules, which proved to be an important configuration of protein secondary and tertiary structures (Futami, Tada et al. 2000). Substitution of hydroxy to thiol may add specific covalent binding property towards the molecule.



Scheme 43: Two thiol groups form a disulfide bond through oxidation.

117 was synthesized from **115**, by reacting **115** with sodium hydrosulfide (Scheme 44) (Majumdar and Ghosh 2002).



Scheme 44: Synthesis of **117** (Majumdar and Ghosh 2002).

This method was originally developed by Majumdar and Ghosh. While the original solvent was ethanol, in this case, anhydrous methanol was used as an alternative for easier purification (We found that the product did not dissolve in methanol, which is ideal for easy filtration. The product did dissolve in

ethanol). As **115** has poor solubility in methanol, the adequate solvent should be added to ensure it is dissolved completely.

The two reactants (**115** and sodium hydrosulfide) were mixed and stirred for two hours. Although the original method suggests column purification, it can be purified through filtration directly. The solution in the flask was carefully evaporated until a large amount of precipitation was formed. The precipitate was then collected and analyzed by NMR, confirmed to be pure **115** product.

The rest of the filtrate was evaporated to a crude product, then either acetone or methanol was added. The precipitates were filtered off again to collect more products. This process was repeated 5-10 times until no more precipitate was formed. Finally, for 500 mg of **115**, 184 mg **117** was collected (yield 66%).

3.3.9.3: Synthesis of 4–amino–coumarin (116)

116 was synthesized to further investigate whether a stronger 4'- nucleophile group would increase or decrease its antiproliferative activity.
Previously, a novobiocin like analogue with 4'-amino group **121** showed reduced activity than 4'-OTs analogue **22** (Figure. 128) (Audisio, Methy-Gonnot et al. 2014).

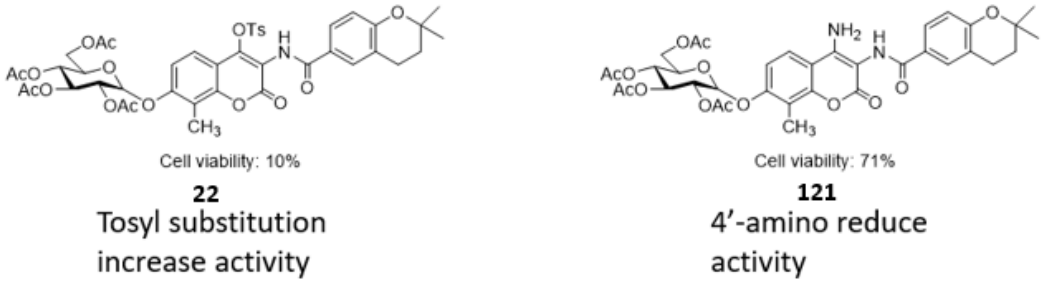
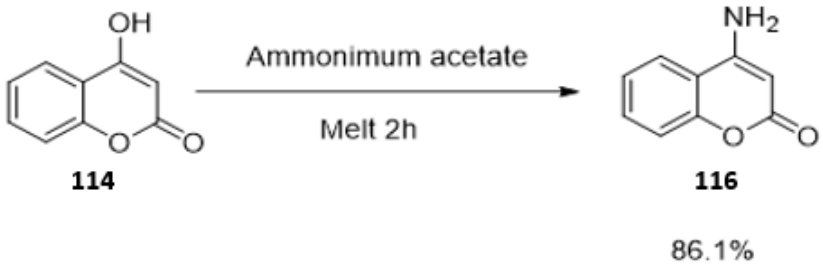


Figure 128: 4'-amino substitution reduces antiproliferative activity (Audisio, Methy-Gonnot et al. 2014).

Synthesis of **116** was used to a developed a method of which 4-hydroxy coumarin **114** was mixed with melted ammonium acetate. (Ghalehshahi, Balalaie et al. 2019) The system was heated to 110°C and stirred for 2 hours.



114 **116**

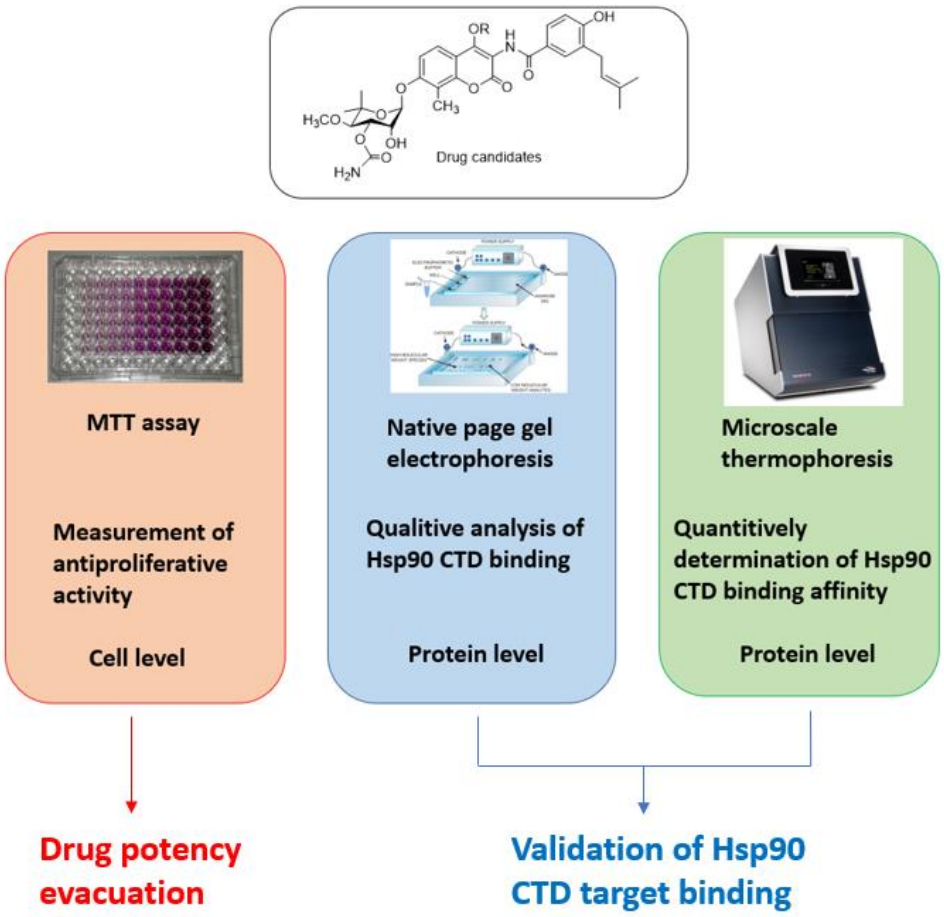
86.1%

Scheme 45: Synthesis of **116**. (Ghalehshahi, Balalaie et al. 2019)

The orange crude product **116** was recrystallized in 30% methanol to afford a white needle-like pure product, with a yield of 86%

3.3.10 Biological assays of synthetic compounds.

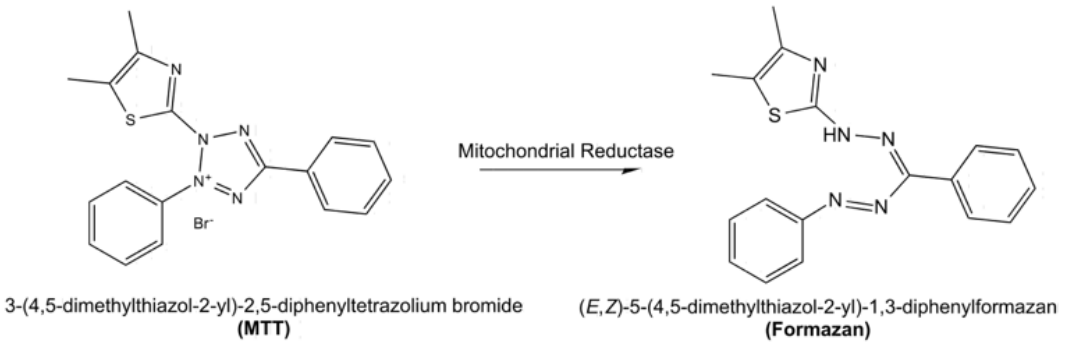
Following on from the synthesis of designed covalent (and non-covalent) drug candidates, we then elected to evaluate its anticancer biological effect through several biological assays. This part of experiments was collaborated with Prof. Andreas Schatzlein and Dr. Francisco Humberto Xavier-Junior from UCL school of pharmacy and Miss Sizhu Lu from University of Oxford. Several tests were designed to measure the drug potency as well as target validation. While all synthetic compounds were tested for their antiproliferative activity using MTT (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide) assay (carried out by Dr. Humberto and Miss Lu), only some key analogues were tested for their Hsp90 binding affinity through native polyacrylamide page gel electrophoresis as well as microscale thermophoresis (MST) (carried out by Guoxuan Sun from Hilton group). In this section, the macro effect of covalent binding will be discussed. It was expected that the success of covalent binding should have an increased biological activity e.g. antiproliferative activity, Hsp90 binding affinity (Scheme 46).



Scheme 46: Biological Tests to be discussed in this section

3.3.10.1 MTT antiproliferative activity assays (Collaborated with Dr. Humberto)

MTT assay is a colourimetric assay used to measure the cell metabolic activity of living cells. Under defined conditions, NAD(P)H-dependent cellular oxidoreductase enzymes may reflect the number of viable cells present. These enzymes are capable of reducing the tetrazolium dye MTT 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyl tetrazolium bromide to its insoluble formazan, which indicates a purple colour (Stockert, Horobin et al. 2018). This purple colour can be quantified by the spectrophotometric device. Under correctly developed conditions, the obtained absorbance value is directly proportional to the number of living cells (Stockert, Horobin et al. 2018). As a result, tetrazolium dye assays can be used to measure cytotoxicity (loss of viable cells) or cytostatic activity (shift from proliferation to quiescence) of potential medicinal agents and toxic materials. MTT assays are usually performed in the dark since the MTT reagent is sensitive to light (Figure. 129).



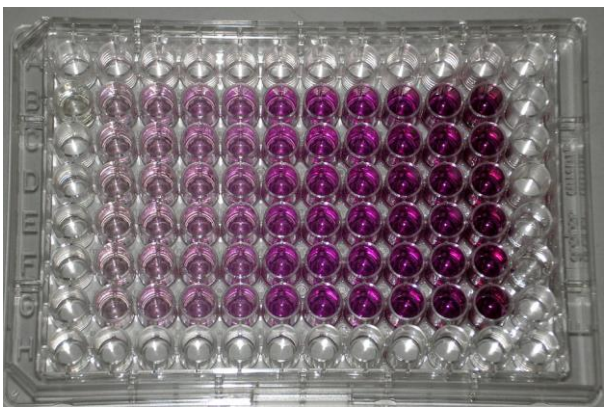


Figure 129: Example of MTT assay.

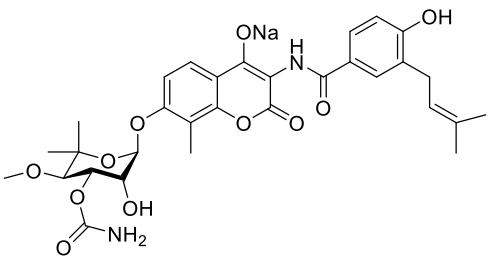
In this project, the MTT assay was used to quantify the antiproliferative activity of drug candidates. Selected cancer cell lines were incubated with each drug candidate, and a solubilization solution (usually DMSO solution) was added to dissolve the insoluble purple formazan product into a coloured solution. Then the well plate was sent for analysis. The absorbance of this coloured solution can be quantified by measuring at a particular wavelength (usually between 500 and 600 nm) by a spectrophotometer. The degree of light absorption is dependent on the degree of formazan concentration accumulated inside the cell as well as on the cell surface. The higher the formazan concentration, the darker the purple colour is, and thus, the higher the absorbance. Higher absorbance means there are more living cells in the untreated sample.

To screen for the most potent drug candidate, all compounds synthesized in **3.3.3 - 3.3.8** and **3.3.9** were tested against Michigan cancer foundation-7 (MCF-7) cell line. MCF-7 is a well-known breast cancer cell line used to measure anti-cancerous drug potency:

Table 5: Cancer cell lines used in the study.

Name of cell line	Type of cancer
A2780	Ovarian
U87MG	Brain
A549	lung
MIA PaCa-2	Pancreatic

Note that some of the compounds to be tested were from the previous chapter 2 (results discussed in p87). As the MTT test on my second year did not include some of the later synthesized novobiocin glycosides (compound **59, 68, 70**). A full list of MTT-tested compounds is given on next page.

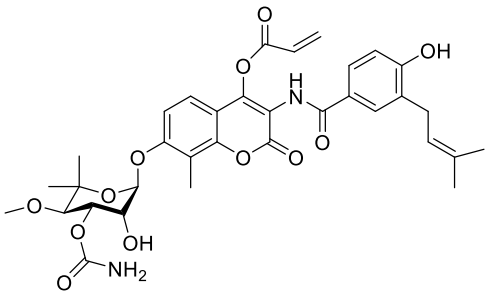


Exact Mass: 634.21
Molecular Weight: 634.61

0

Novobiocin Sodium salt

Soluble in water, DMSO, MeOH



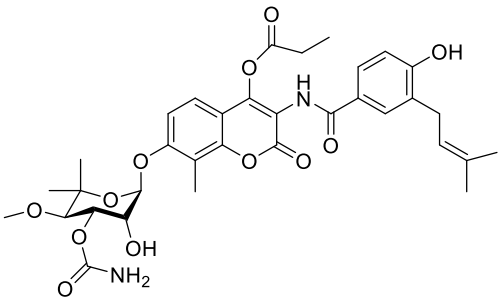
Exact Mass: 666.24
Molecular Weight: 666.68

108

4-Micheal acceptor- Novobiocin

Insoluble in water

Soluble in MeOH and DMSO



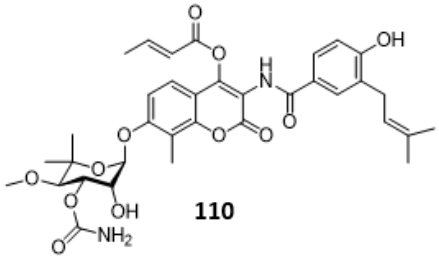
Exact Mass: 668.26
Molecular Weight: 668.70

109

Inactive form Michael acceptor novobiocin

Insoluble in water

Soluble in MeOH and DMSO



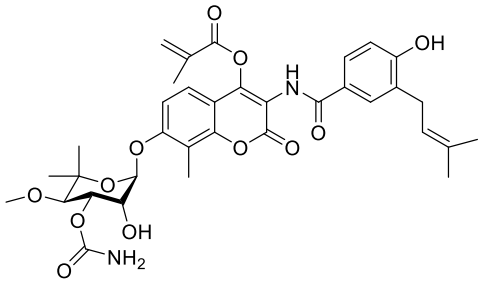
Exact Mass: 680.26
Molecular Weight: 680.71

110

Methyl Michael acceptor novobiocin

Insoluble in water

Soluble in DMSO



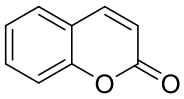
Exact Mass: 680.26
Molecular Weight: 680.71

111

Isoform of Methyl novobiocin

Insoluble in water

Soluble in and DMSO



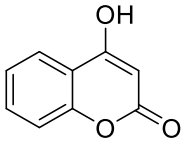
Exact Mass: 146.04
Molecular Weight: 146.14

122

Coumarin

Insoluble in cold water

Soluble in and DMSO



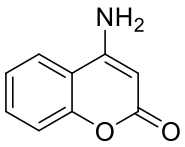
Exact Mass: 162.03
Molecular Weight: 162.14

4-hydroxy coumarin

Insoluble in water, DCM

Soluble in DMSO, THF, MeOH, Acetone

114



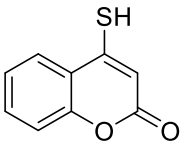
Exact Mass: 161.05
Molecular Weight: 161.16

4-amino coumarin

Insoluble in water

Soluble in DMSO, DMF

116



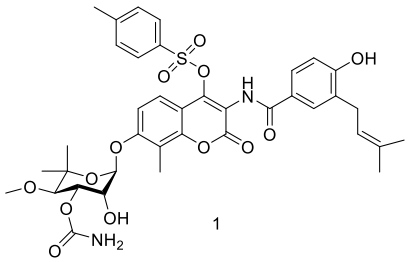
Exact Mass: 178.01
Molecular Weight: 178.20

4-thiol coumarin

Insoluble in water

Soluble in DMSO, DMF

117



1

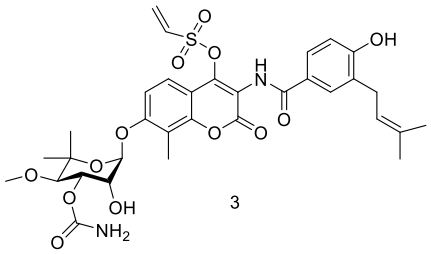
Exact Mass: 766.24
Molecular Weight: 766.81

4-tosyl-novobiocin

Insoluble in water

Soluble in DMSO

107



3

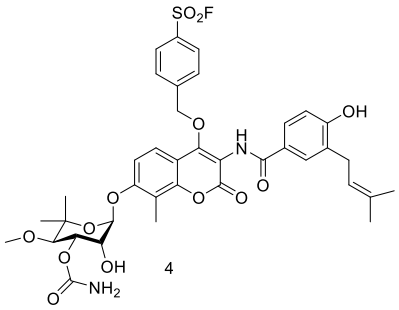
Exact Mass: 702.21
Molecular Weight: 702.73

4-vinyl sulfone-novobiocin

Insoluble in water

Soluble in DMSO

112



4

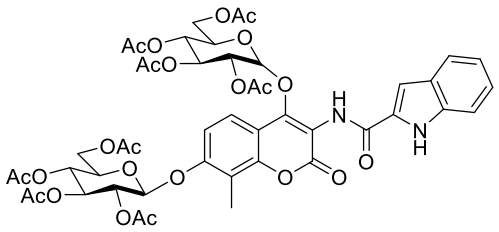
Exact Mass: 784.23
Molecular Weight: 784.81

4-SO2F-Novobiocin

Insoluble in water

Soluble in DMSO

113



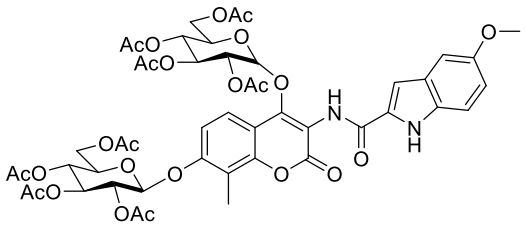
Exact Mass: 1010.28
Molecular Weight: 1010.91

5

Insoluble in water

Soluble in DMSO

68



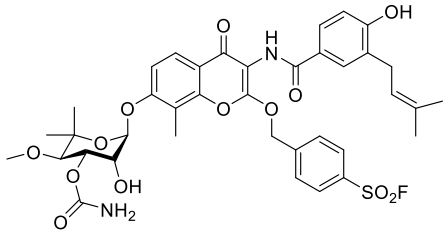
Exact Mass: 1040.29
Molecular Weight: 1040.93

6

Insoluble in water

Soluble in DMSO

70

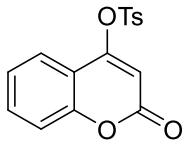


Exact Mass: 784.23
Molecular Weight: 784.81

Insoluble in water

Soluble in DMSO

120



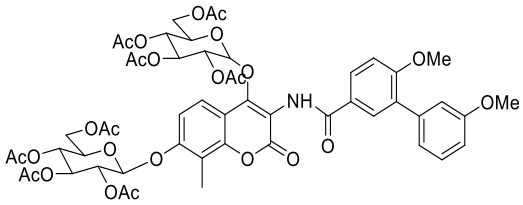
Exact Mass: 316.04
Molecular Weight: 316.33

4-Tosyl-coumarin

Insoluble in water

Soluble in DMSO

115



Exact Mass: 1107.32
Molecular Weight: 1108.02

Insoluble in water

Soluble in DMSO

59

Compounds tested against multiple cell lines:

0, 108, 109, 110, 111, 122, 114, 112

All compounds were tested against MCF-7; where compound **123** was identical to compound **108**. However, they were produced in different synthetical batches. This was to ensure the robustness of synthetic procedures.

3.3.10.2 MTT assay test results (General)

Because the amount of results in this section is overwhelmingly large, only the original results will be posted here. A discussion of data is on p193. Experimental and data analysis in this section was done by Dr. Humberto from UCL.

Test results of drug sensitivity

Table 6 shows the IC_{50} values of different compounds against A2780, MCF7, U-87 MG, MIA PaCa-2 and A549 human cancer cells to assess the drug sensitivity towards different cell lines

IC50 (μM)	A2780	SD	MCF7	SD	U-87 MG	SD	MIA PaCa-2	SD	A549	SD
0	567.3	6.6	1339.7	176.0	871.8	69.8	666.0	20.0	726.7	43.5
108	568.0	15.0	537.6	7.4	557.4	10.4	532.5	31.8	550.0	14.0
109	538.4	10.2	526.7	27.6	530.7	33.6	469.6	8.7	532.6	20.3
110	497.2	35.4	494.7	14.3	553.0	20.9	457.4	8.7	533.5	29.1
111	657.2	13.4	532.2	37.5	557.7	13.6	491.9	60.0	505.5	19.4
122	NT		NT		1327.2	114.8	NT		NT	
114	NT		NT		NT		NT		NT	
Deleted	NT		606.5	47.0	1346.4	175.9	639.4	6.5	766.8	62.7
112	573.5	1.3	513.5	93.1	539.2	9.4	582.5	11.9	615.5	35.1

IC_{50} (μM) values were determined for each compound in a panel of human cancer cell lines. The cytotoxicity of the compounds was evaluated by the colourimetric MTT assay after 24 h of continuous exposure. NT = Non-toxic

Table 7 *Correlation coefficient* obtained to determine the IC_{50} values for different compounds in a panel of human cancer cell lines (A2780, MCF7, U-87 MG, MIA PaCa-2 and A549)

IC50 (μM)	A2780	MCF7	U-87 MG	MIA PaCa-2	A549
0	0.997	0.983	0.976	0.997	0.954
108	0.997	0.995	0.996	0.998	0.998
109	0.999	1.000	0.980	0.995	0.998
110	0.995	0.983	0.984	0.959	0.999
111	0.999	1.000	0.992	0.997	0.992
122	0.056	-0.099	0.812	0.044	0.073
114	0.084	-0.036	0.110	0.139	0.022
Deleted	0.547	0.955	0.481	0.997	0.994
112	0.991	0.997	0.995	0.998	0.985

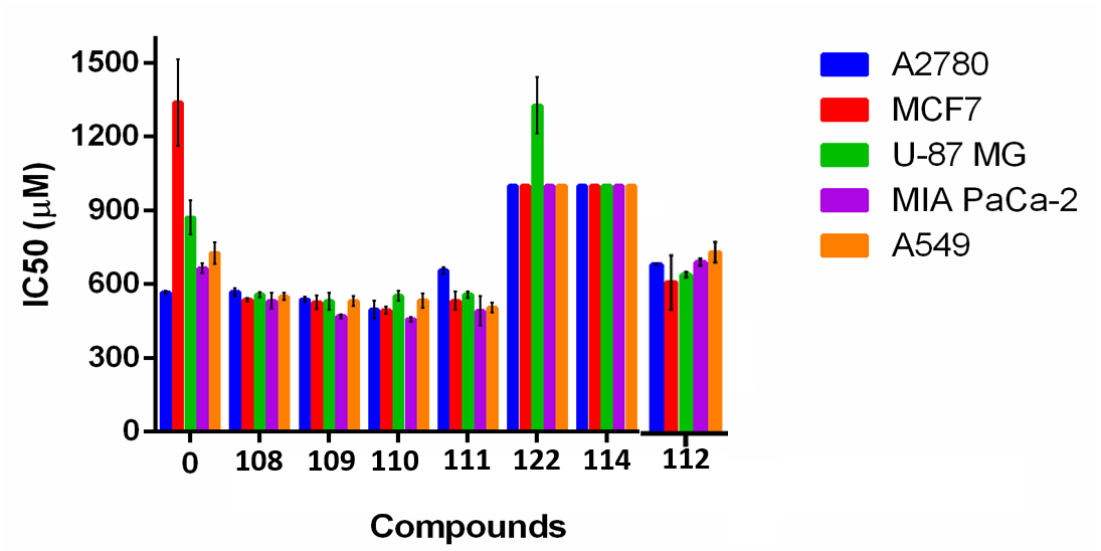
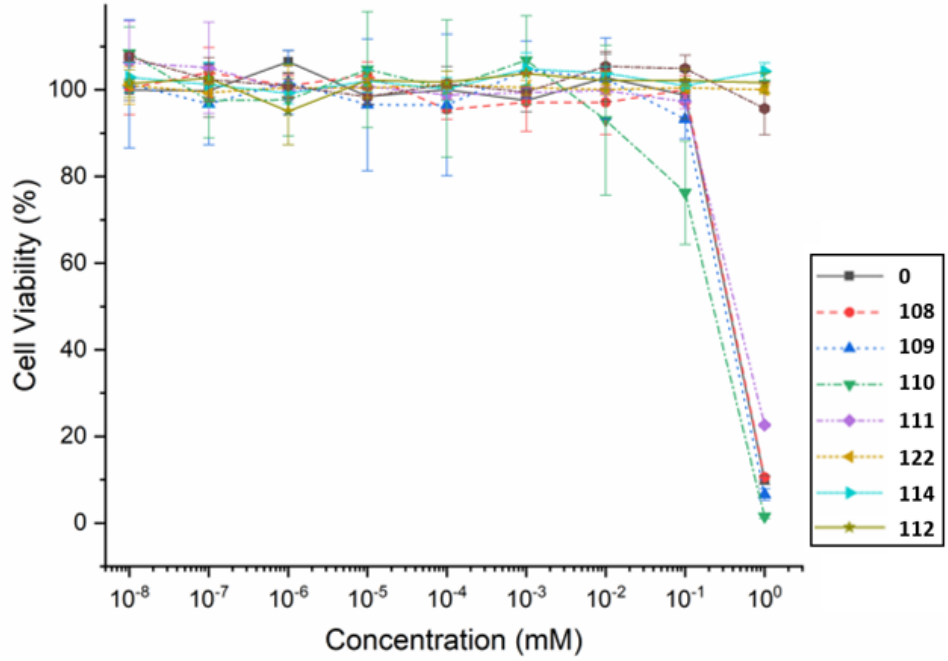


Figure 130**:** IC_{50} (μM) of different compounds against A2780, MCF7, U-87 MG, MIA PaCa-2, and A549 human cancer cells. Values reproduced from Table 7.

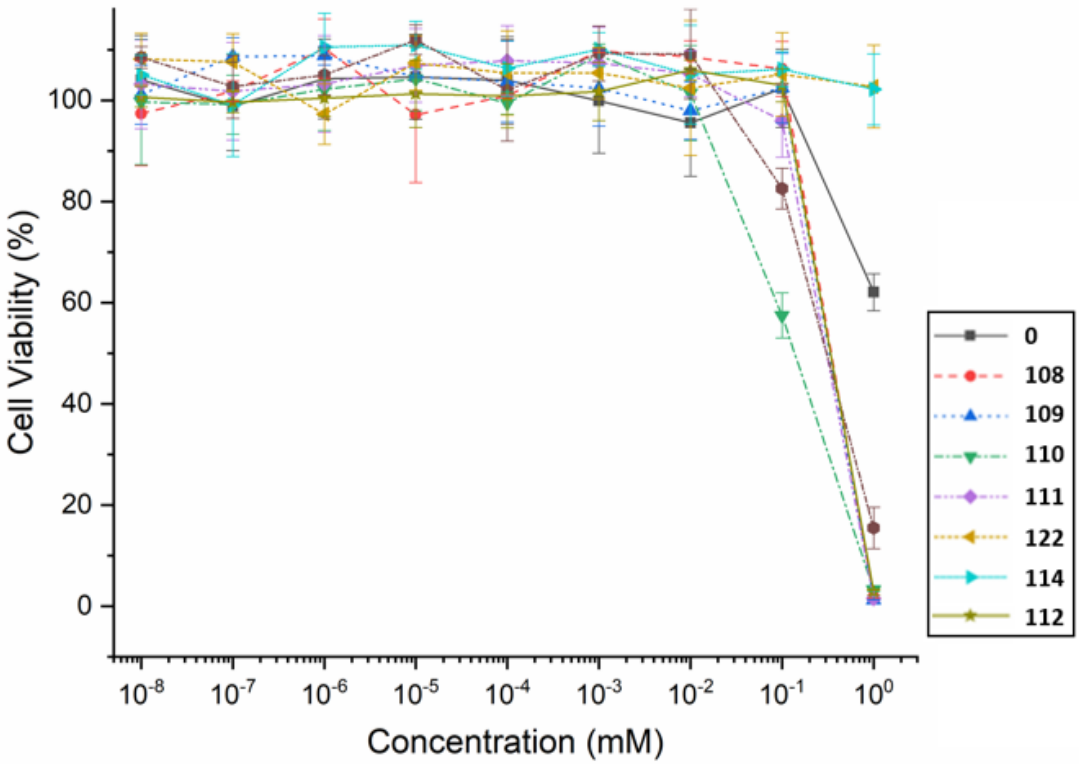
Dose-response curves of drug sensitivity tests

Dose-response curves obtained after 24h of continuous exposure to different compounds against A2780, MCF7, U-87 MG, MIA PaCa-2, and A549 human cancer cells. If the curve has no trend of sloping down, the relative compounds were marked as “Non-toxic” in Table 6

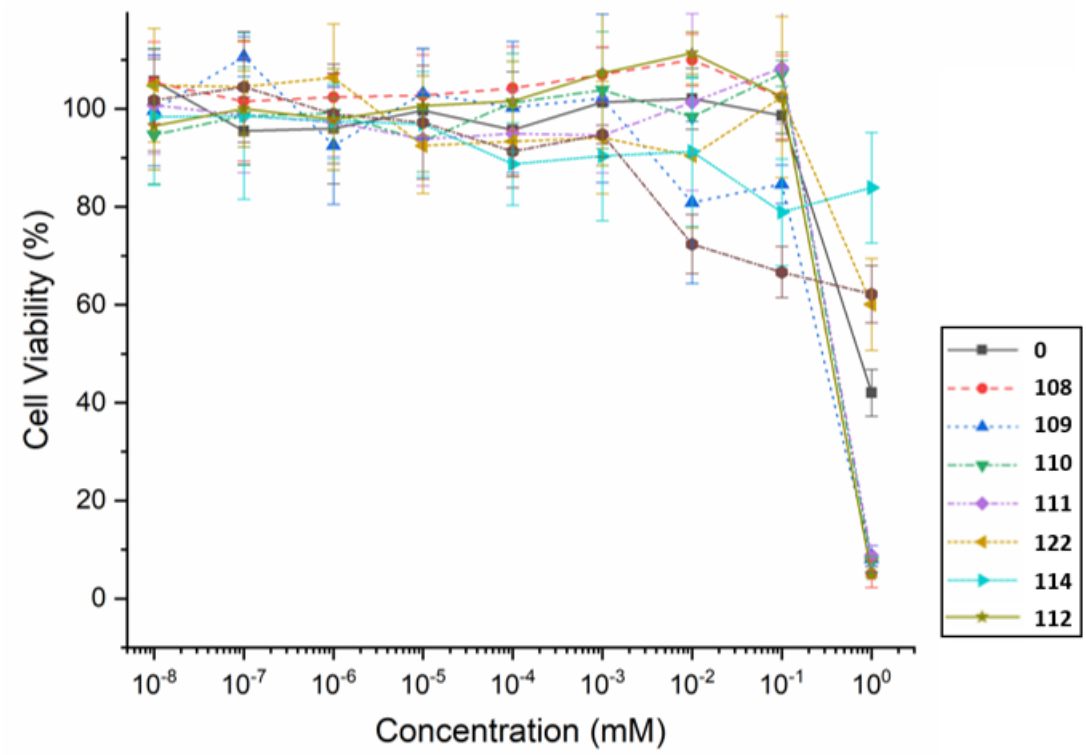
A2780


Cell Viability (%)
Concentration (mM)
0
108
109
110
111
122
114
112

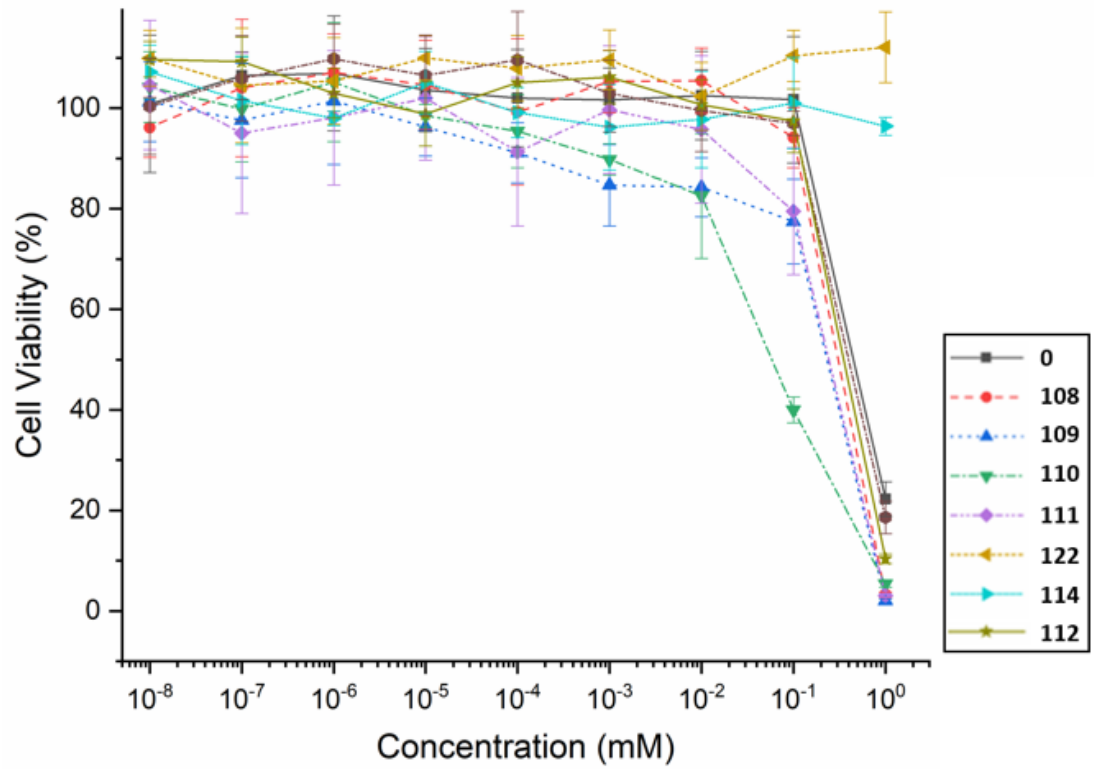
MCF7


Cell Viability (%)
Concentration (mM)
0
108
109
110
111
122
114
112

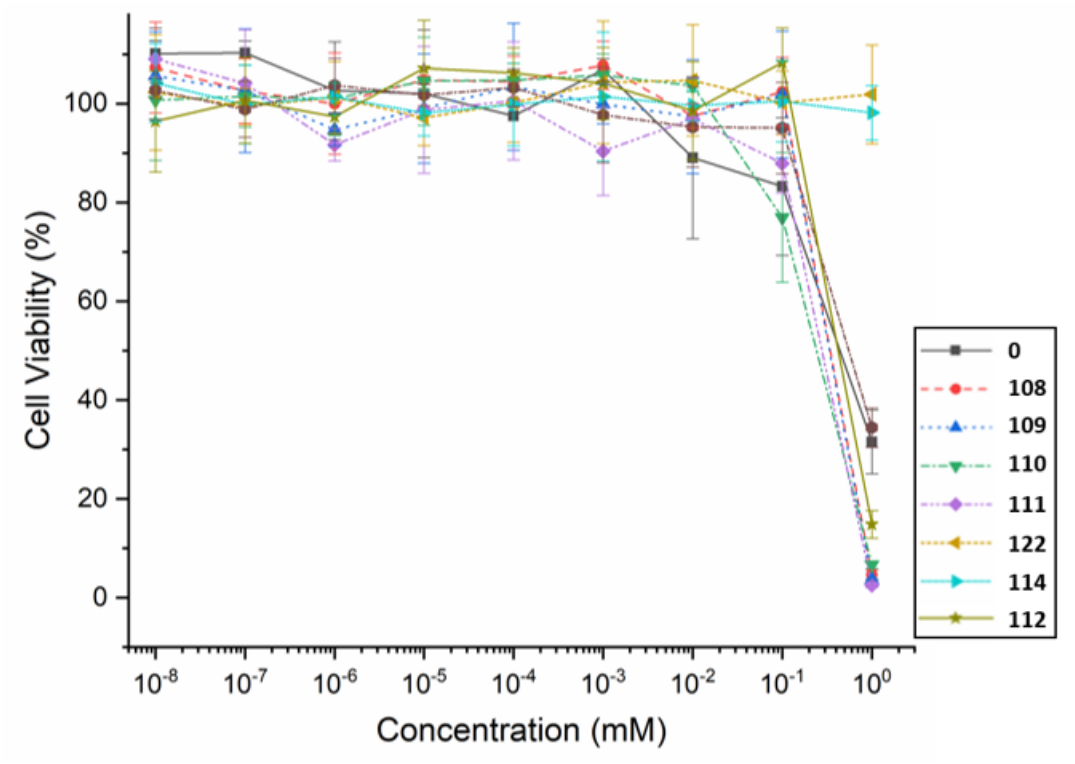
U-87MG



MIA PaCa-2



A549


Cell Viability (%)
100
80
60
40
20
0
10^{-8}
10^{-7}
10^{-6}
10^{-5}
10^{-4}
10^{-3}
10^{-2}
10^{-1}
10^{0}
Concentration (mM)
0
108
109
110
111
122
114
112

Drug potency screening against MCF-7

Table 8 shows IC_{50} values of different compounds against MCF7 cells. NT = non-toxic

IC50 (µM)	MCF7	SD	r2
116	933.6	148.6	0.962
117	525.2	28.2	0.998
107	171.3	15.0	0.981
113	112.0	31.8	0.952
68	1583.3	192.0	0.784
70	NT		0.186
120	689.7	47.1	0.963
115	515.1	38.5	0.999
59	NT		-0.107
123	538.5	5.6	0.999

The cytotoxicity of the compounds was evaluated by the colourimetric MTT assay after 24 h of continuous exposure.

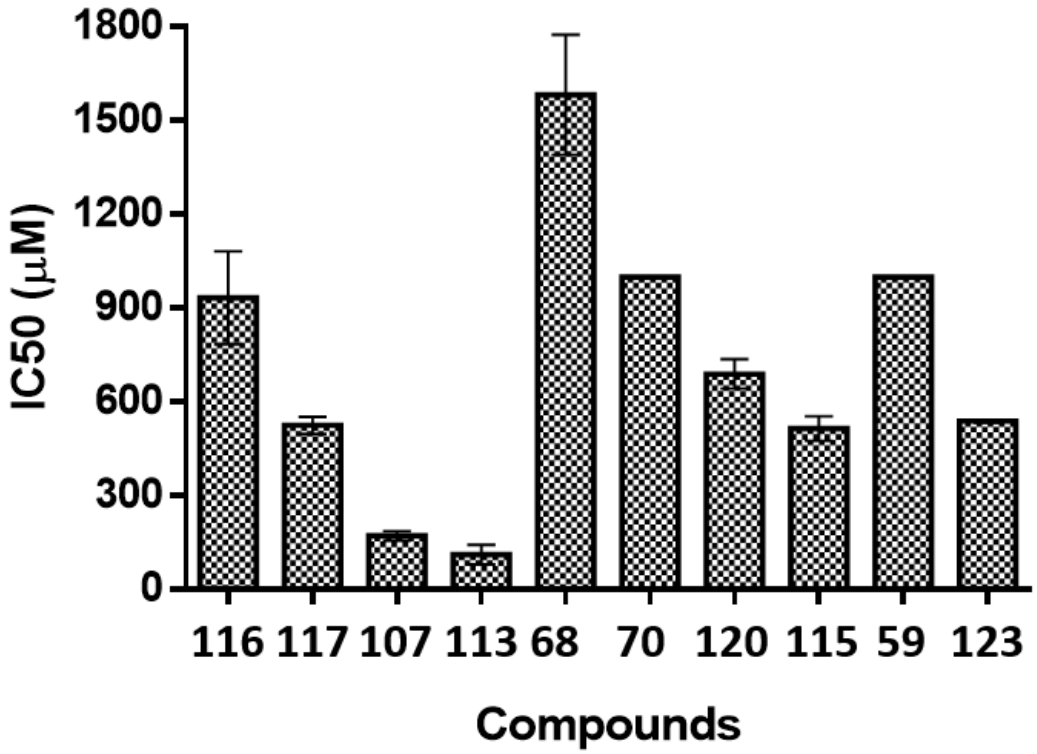
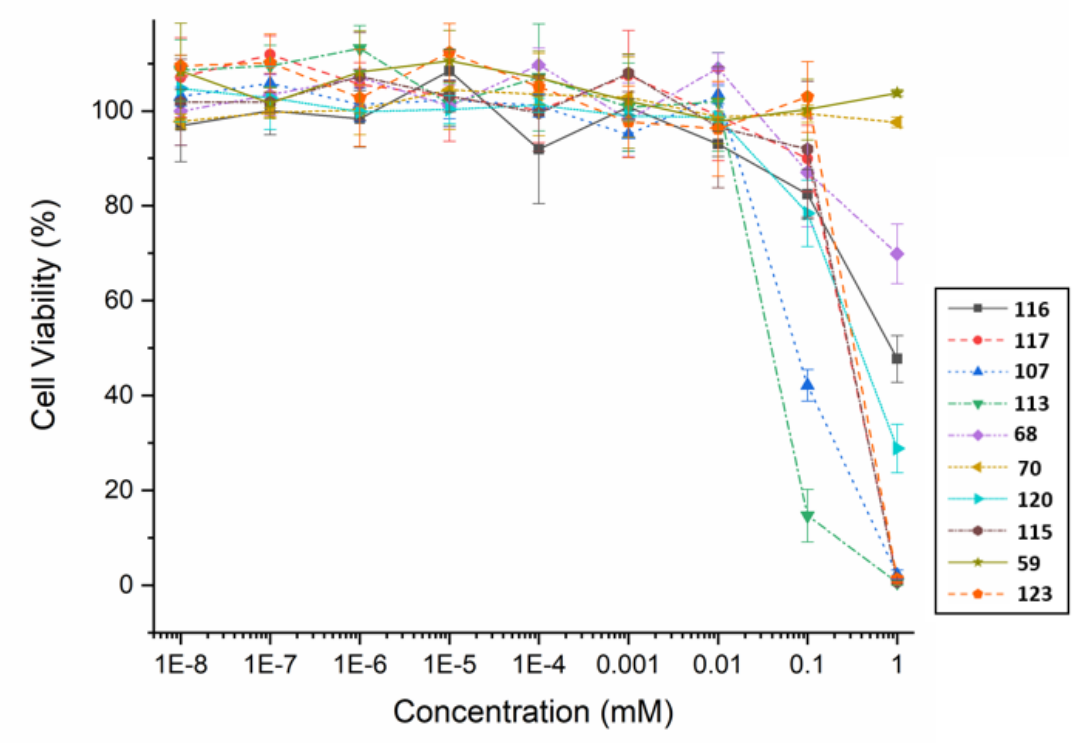


Figure 131**:** IC_{50} (µM) of different compounds against MCF7 cells. Values reproduced from Table 3

Dose-response curve of MCF-7 experiments

Dose-response curves obtained after 24h of continuous exposure for different compounds against MCF7 cells. If the curve has no trend of sloping down, relative compounds were marked as “Non-toxic” in table 8



3.3.10.3 Discussion of MTT test results

Analysis of MCF-7 drug potency screening results

A bar chart which summarized all results of novobiocin-based analogues is given below (Figure. 132):

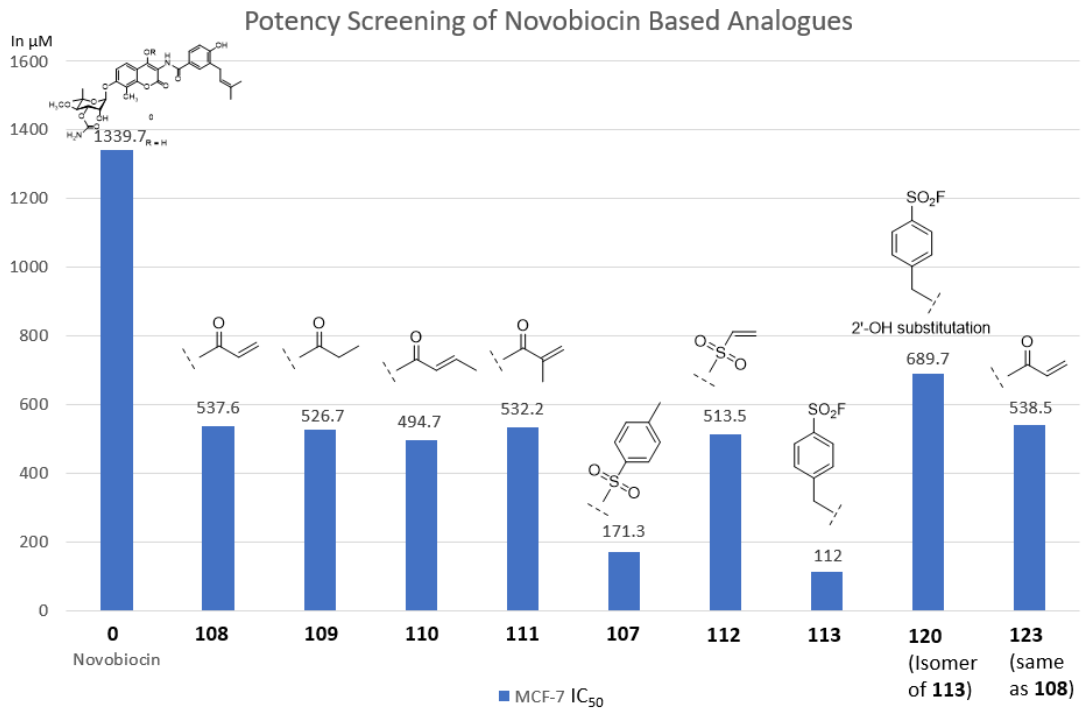


Figure 132: Antiproliferative activity screening results of novobiocin analogues. Compound **120** is the only 2'-OH substituted product, which is the isomer of compound **113**. Compound **108** and **123** are the same compound from different synthetic batches.

Although the value of these results may not show high potency, it is worth mentioning that all the data was observed after 24 hours' drug incubation, which was shorter than some published papers, for example, in a previous paper published, novobiocin **0** has an IC_{50} value of 237 μM against the MCF-7 cell line when incubated for 72 hours (Audisio, Methy-Gonnot et al. 2014). It depends on the researcher to decide the duration of the MTT test. Herein, in order to preserve the same standard with the previous work of our group in 2011 (Patel, Fuente et al. 2011), we decided to use the same duration of 24 hours as before. Longer exposure will increase the value of the final potency.

In the example of novobiocin, its antiproliferative value increased about 5.6 folds with a 48 hours incubation compared the 24-hours incubation. As a result, all drug candidates would be more potent if the incubation duration was increased.

Compound **113** and compound **107** were the two analogues with the best antiproliferative activity, with IC_{50} values of 112 μM and 171.3 μM, respectively. The potency compared with novobiocin **0** was increased by 11.9-fold and 7.8-fold. This was due to the hydrophobicity of 4'-substitution type, which indicated that Hsp90 CTD might have a hydrophobic binding pocket. However, it was not certain whether the electrophile substitution would covalently modify any residues of Hsp90 CTD. For compound **113**, further investigation is needed to confirm the importance of sulfonyl fluoride substitution, as the benefit might come from benzene ring substitution only. Moreover, our result of compound **107** summarized the same conclusion with the literature, in which 4'-tosyl substitution increased the anti-proliferative activity (Figure.133) (Audisio, Methy-Gonnot et al. 2014).

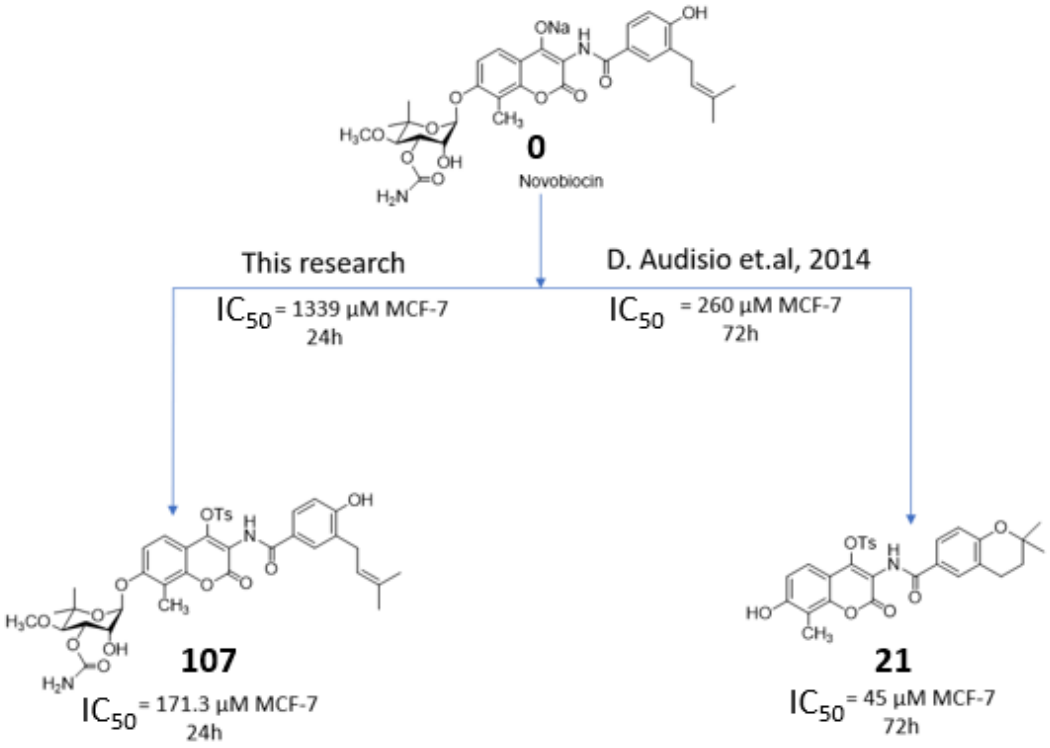


Figure 133: 4'-tosyl substitution of **0** resulted in **107** and **21,** increases activity. (Audisio, Methy-Gonnot et al. 2014).

Due to the 2'-substitutated enol tautomeric isomer of compound **113** (which is compound **120**) showed a worse activity, 4'-substitution was more potent than 2'- substitution (Figure.132).

4'-substituted saturated and unsaturated esters (compound **108,109,110,** and **111**), and 4'-vinyl sulfone (compound **112**) mildly increase activity range from 2.5 fold to 2.7 fold. These modifications increase lipophilicity like benzene ring substitution (compound **107** & **113**), but much weaker. We can conclude from the results that hydrophobic 4'-substitution increases the anti-proliferative activity.

Nevertheless, the test result did not show a covalent binding benefit when comparing compound **108** and compound **109**. According to the literature, the success of covalent binding would tell the difference between the active form of unsaturated acrylate and the inactive hydrolyzed form. For example, in the

research of developing covalent inhibitor **124** for enterovirus 71 3C protease, reduction of the double bond will lead to loss of anti-viral activity (**125** in Figure. 134) (Ma, Li et al. 2019).

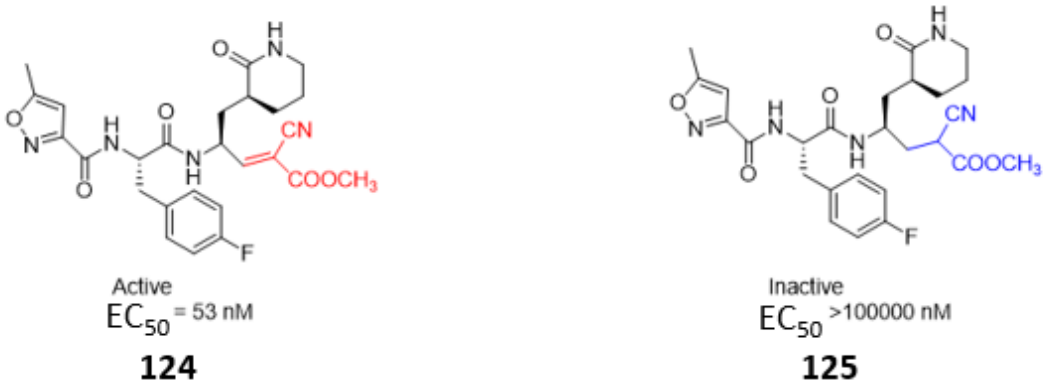


Figure 134: When the electrophilic double bond was reduced, the viral protease inhibitor **124** lost activity, indicating the success of covalent inhibition.

In this case, active Michael acceptor 4'-acrylate-novobiocin (**108**) showed an activity of 537.6 μM, without any improvement compared to the inactive reduced form of 4'-propionyl-novobiocin (**109**) for which the antiproliferative activity was 526.7 μM. Active analogues of 4'-acrylate-novobiocin (**110** and **111**) also failed to show increased activity; only compound **110** showed a minimal improvement (494.7 μM, 0.08 fold increase). Overall, these results showed that hydrophobic substitution on 4' increase activity. Active form Michael acceptors did not act as a covalent inhibitor in cell level MTT test.

There are some potential reasons for why these Michael acceptors did not act as covalent modifiers. Firstly, these compounds might bind to different proteins rather than Hsp90 CTD. To validate this issue, the drug-protein interaction assay will be discussed in the next section (p205). Secondly, active electrophiles consume electrons and act as oxidative stress to cells. Living cells have developed specialized pathways to detox active electrophiles, for example, by producing glutathione, which contains active cysteine to detox Michael acceptors. As a result, acrylates compound, in this case, may already be reduced into inactive form then lost covalent binding activity but only non-covalent inhibition. Finally, some review articles pointed out that the acrylate type of Michael acceptor at first did not draw any attention when researchers

failed to develop prototype covalent inhibitors due to its instability under bio condition. The ester linkage bond is too easy to be metabolized and lose its activity until a more stable acrylamide linkage can replace the acrylate, then, the first-generation acrylamide type covalent inhibitor afatinib was successful and put into the market (Jackson, Widen et al. 2017). Our result here proved the importance of acrylamide bond of those approved covalent inhibitors, such as the BTK inhibitor ibrutinib. Covalent warheads need adequate degrees of bio-resistance before it can be transported towards the drug target.

In this project, there were experiments designed to validate the questions mentioned above, and they will be discussed with the drug-protein assay together in the next section (p205).

For the coumarin study to expand the possibility of 4'- linker type, a summary bar chart is given below (Figure. 135):

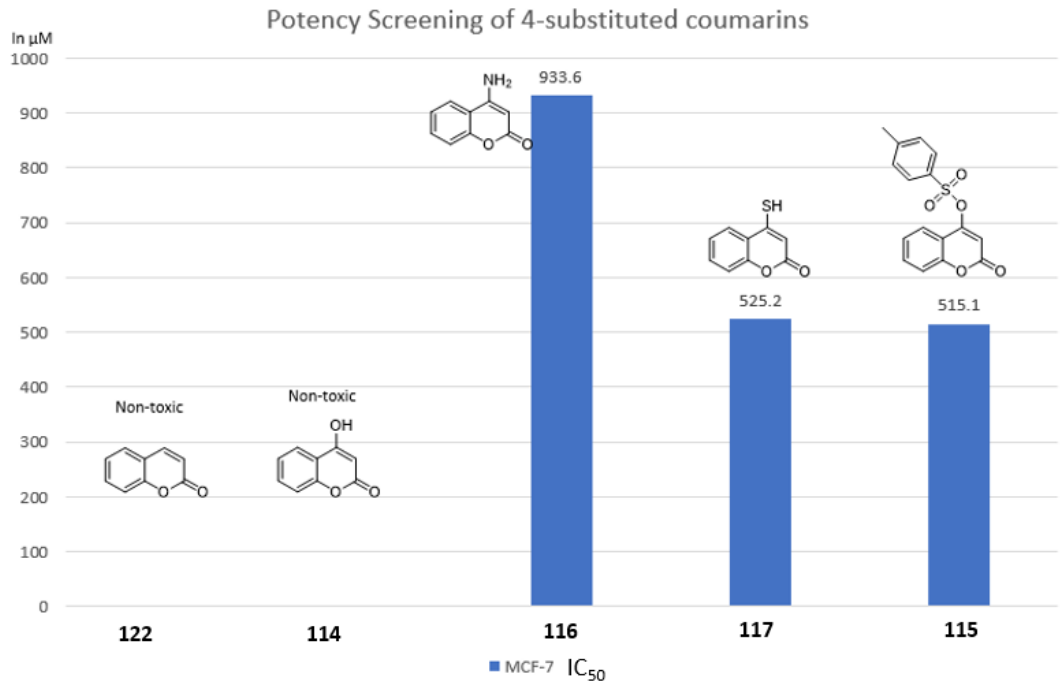


Figure 135: Antiproliferative activity screening result of coumarin **122** and coumarin 4-substituted analogues **114** - **117**.

Coumarins, also known as benzopyran-2-ones, form an elite class of naturally occurring compounds that possess promising therapeutic perspectives. This type of compounds belongs to the flavonoid class of plant secondary

metabolites and have a variety of biological activities, usually associated with low toxicity. However, due to diversity in their structural complexity, these range from simply substituted coumarins to polysubstituted coumarins (Klenkar and Molnar 2015) Coumarin analogues are also be found in micro-organisms, including our core scaffold novobiocin from *Streptomyces* and aflatoxins from *Aspergillus* species (Klenkar and Molnar 2015).

Coumarin derivatives process antiproliferative activities. Marshall M. E. et al. (1994) showed that coumarin and 7-hydroxycoumarin could inhibit growth in human cancer cell lines, including MCF-7 (Marshall, Mohler et al. 1994). Coumarin itself has mild toxicity; however, it can be metabolized into different inactive forms and become less toxic. The main metabolic pathway in humans is hydroxylation. Possible metabolites are 4, 5, 6, 7 or 8 hydroxy coumarin (Fig. 136), mainly at 7 position (Lake 1999).

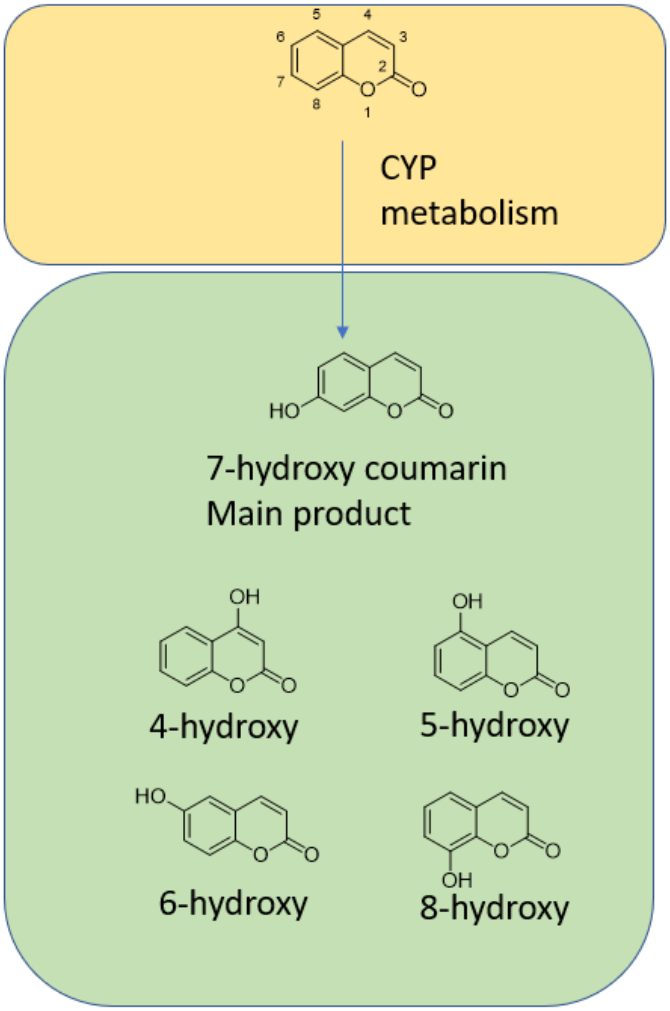
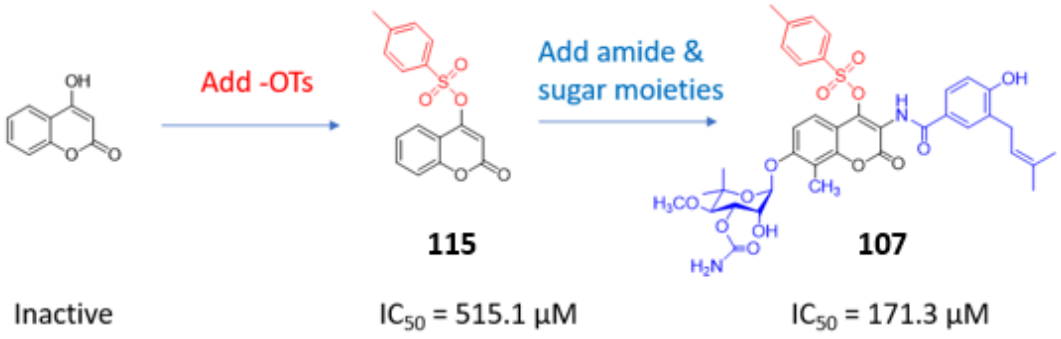


Figure 136: Coumarin non-toxic metabolites.

Therefore, it was expected that coumarin (compound **122**) and 4-hydroxy coumarin (compound **114**) induce no cytotoxicity against MCF-7. Substitution of hydroxy to either nucleophile like amino (compound **116**) or thiol (compound **117**) gained antiproliferative activity. While amino substitution **116** resulted in a mild activity 933.6 μM, thiol substitution compound **117** increased its activity with 1.7-fold (525.2 μM). 4'-thiol was very promising as the thiol group can behave either as a nucleophile or forming disulfide bond as an electrophile. Its reactive flexibility may add to its outstanding antiproliferative activity.

Hydrophobic substitution on coumarin 4 position also increases activity. 4-tosyl substitution again proved to have a positive effect on increasing cytotoxicity, for which compound **115** has a value of 515.1 μM. Compared to novobiocin analogue compound **107**, the anti-proliferative activity of **115** showed the importance of both Tosyl substitution and novobiocin amide side chain and noviose sugar moieties.



Scheme 47: Importance of 4-OTs and novobiocin moieties substitution

The coumarin study provided promising results that can be concerned with future drug design. While benefits of 4-tosyl had already been validated in the study, 4-thiol type of coumarin or 4'-thiol novobiocin has excellent potential to generate more potent lead structures.

Notably, these coumarins were not validated by their drug target. They were to confirm their target protein, not necessarily Hsp90 CTD.

Analysis of drug sensitivity data

A bar chart (Figure. 137) which summarized of all novobiocin based analogues is given below:

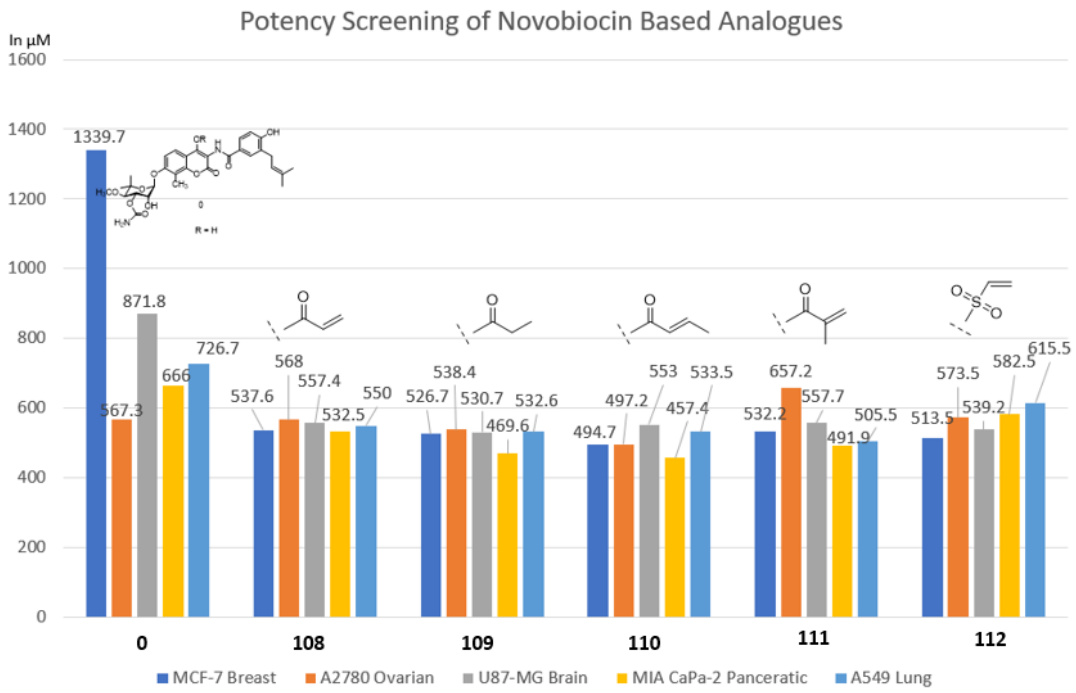


Figure 137: MTT test results of novobiocin 4'-substituted analogues against different types of cancer cell lines.

Pertaining to the 24 hours' MTT assay, while novobiocin **0** exhibited significant low activity against breast cancer (IC_{50} > 1000 µM MCF-7 cell line), it had better cytotoxicity against other cancer types shown above (Figure. 137). Ovarian cancer cell line A2780 (IC_{50} = 567.3 µM) and pancreatic cancer cell line MIA CaPa-2 (IC_{50} = 666.0 µM) were two types of cancer cells that were the most sensitive to novobiocin induced novobiocin-induced Hsp90 CTD inhibition, with increased potency of 2.36-fold and 2.01-fold respectively. Furthermore, novobiocin was slightly more potent against brain cancer cell line U87-MG (871.8 µM) and lung cancer cell line A549 (726.7 µM) then breast cancer cell line MCF-7 (> 1000 µM).

Surprisingly, all 4'-substituted analogues (**108** – **112**) showed very similar activity against different types of cancer cells, suggesting that such modification added robustness to Hsp90 CTD inhibition. There was no doubt that these analogues are most successful against breast cancer by increasing

the activity > 2-fold on average. Increased growth inhibitory effect could also be found in brain, lung, and pancreatic cell lines with average of increased potency of 1.59-fold, 1.33-fold and 1.31-fold, respectively. Only the ovarian cell line did not significantly be changed of drug sensitivity towards 4'-modifications comparing with original novobiocin moiety, while compound **111** even showed a decreased activity of -89.9 μM.

To discuss the reason for this phenomenon, the increased activity needs to be verified if it was due to the Hsp90 CTD covalent modification. The activity of cysteine reactive Michael acceptor compound **108** has mostly the same IC_{50} as cysteine inactive compound **109** among all cell lines. It was evident that the covalent inhibition did not happen and the reason of which has been discussed before (p196). It can be concluded here that detox of acrylate or vinyl sulfone would not only happen to occur in breast cancer cell lines but universal to all kinds of cancer cell lines tested.

If the increase of anti-proliferative activity in Figure. 137 did not caused by covalent modification, then it is highly possible that these data showed a trend of cell membrane permeation. In the MTT assay, only intracellular Hsp90 rather than extracellular was inhibited. As a result, the capability of cell membrane permeation for small molecules is vital for bioavailability or drug potency. There are usually 3 ways of cytosolic entry: passive diffusion, transported-mediated entry, and endocytosis (Figure.138) (Yang and Hinner 2015).

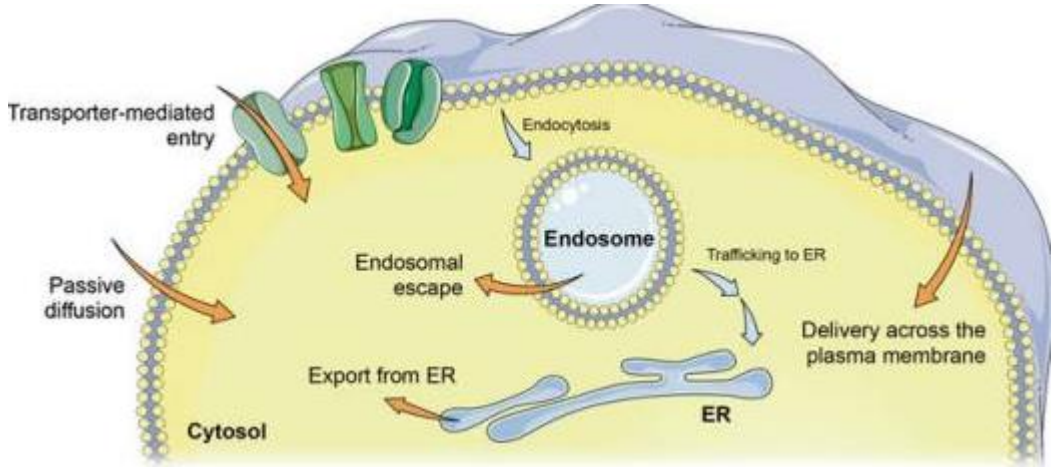


Figure 138: 3 ways of membrane permeation, which are passive diffusion, transporter-mediated entry and endocytosis. Molecules may passively diffuse across the cell membrane or be shuttled in via natural or artificial delivery mechanisms. Membrane transporters allow the passage of various ions and metabolites. (Yang and Hinner 2015).

While small molecules are usually not transported through the endocytosis pathway, they are most likely to complete membrane permeation through passive diffusion (Veber, Johnson et al. 2002). Small molecule drugs can also be transported through surface transporter proteins, however, that depends on the structure and cell type. Novobiocin was reported to penetrate the cell membrane through passive diffusion (Mandler, Baidin et al. 2018). As for small molecule drugs to passively penetrate the lipophilic cell membrane, a high degree of lipophilicity is required to allow interaction between lipid and small molecules (Zhang, Qin et al. 2019). In this case, all these 4'-modifications of novobiocin increase lipophilicity of the scaffold, making the molecule easier to pass through and present an inhibitory effect. This benefit was significant towards MCF-7 breast cancer cells and U87-MG brain cancer cells, indicating that those analogues being tested were less supported by surface transporter protein for permeation. On the contrary, such modification of novobiocin did not increase the drug potency against ovarian cancer. Although there is no reported specific transporter protein of ovarian cancer cells to transport small molecule like novobiocin, ovarian cancer cell A2780 apparently produce more of this kind of pathways naturally which presents a better drug delivery on account of better drug potency for unmodified novobiocin. Thereby, our results

suggested that, for novobiocin scaffold, it is possible to control its anticancer selectivity towards the modification of its lipophilicity, which would be useful in clinical trials to target specific cancer types.

For the coumarin study, a summary bar chart is given below (Figure. 139):

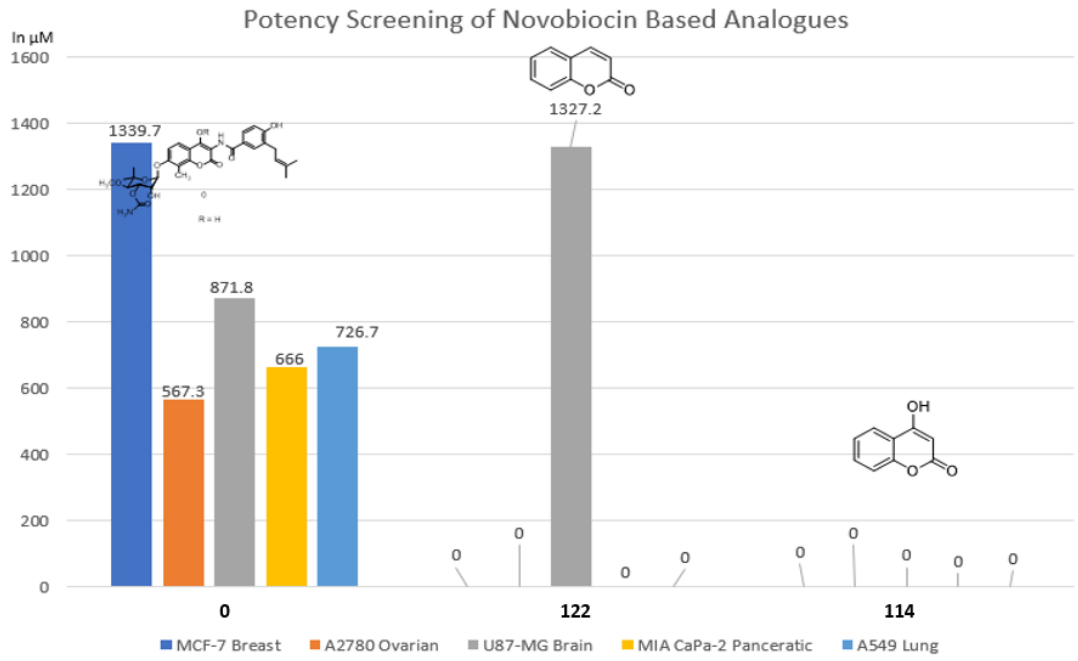


Figure 139: Coumarin antiproliferative sensitivity test.

As have been discussed before, coumarin presented no cytotoxicity towards human cells due to being metabolized into non-toxic 4-hydroxy coumarin. Therefore, compound **122** (coumarin) and compound **114** showed no growth inhibitory effect against breast, ovarian, pancreatic and lung cancer. The only unusual result was coumarin itself which showed mild cytotoxicity (IC_{50} = 1327.2 μM) towards brain cancer, whereas 4-hydroxy coumarin still had no antiproliferative effect. There was no literature reporting any cytotoxic data of coumarin. Nevertheless, we proposed that maybe brain cancer cells could not detox coumarin into its inactive form due to the lack of a certain oxidase. More experiments are needed to map its biological pathway in U87-MG brain cancer cells. This unique mechanism may contribute to the development of anti-brain-cancer drugs. Also, because coumarin is generally harmless to other cells, it

has the potential to be applied as a safe drug for the treatment of brain cancer, delivered by brain cancer-targeted antibodies.

The conclusion of this section can be summed up with the SAR figure below (Figure. 140):

SAR summary of this section:

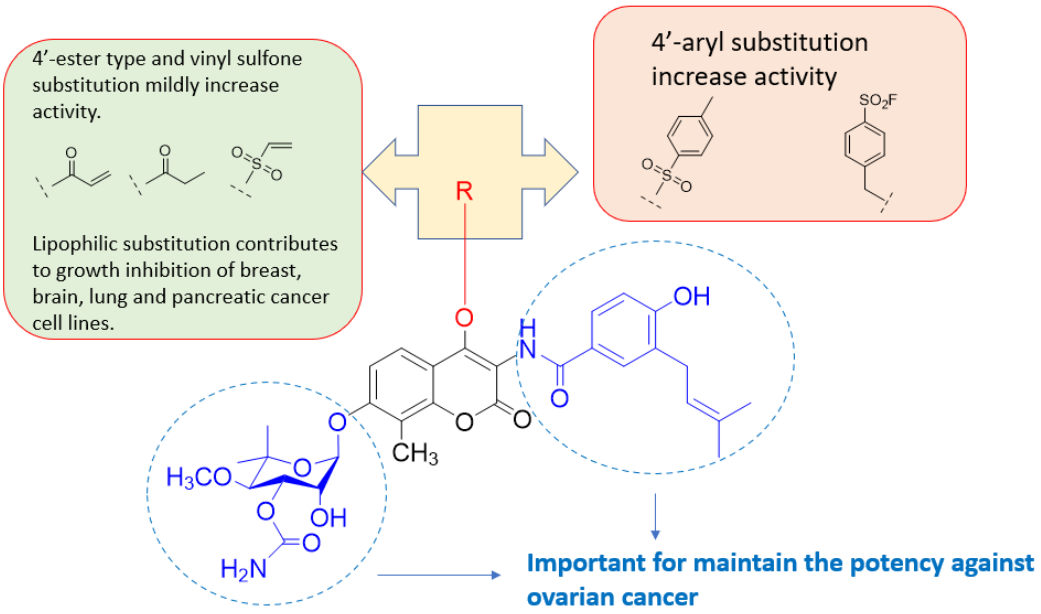


Figure 140: SAR summary from MTT assay test results.

3.3.10.4 Durg-Hsp90 native-page gel binding assay.

MTT tests in the previous section revealed the importance of 4'-substitution, which lead to increased antiproliferative activity. However, whether this benefit came from inhibition of Hsp90 or inhibition of other protein remained to be a question. Despite that bioactivity of Hsp90 CTD inhibitors can be easily tested through various assays, including MTT test, validation of Hsp90 CTD inhibitors to target specifically on Hsp90 CTD has always been a challenge. Most of the published papers using the only indirect method to carry out target validation by quantifying the reduction of Hsp90 client protein using western blot. For example, the figure below Figure. 141 showed Joseph A. Burlison (2008) validated the Hsp90 CTD target binding using western blot. A negative correlation can be found between Hsp90 CTD inhibitors and well-known Hsp90 client protein Her2, Raf and Akt in a MCF-7 cell model.

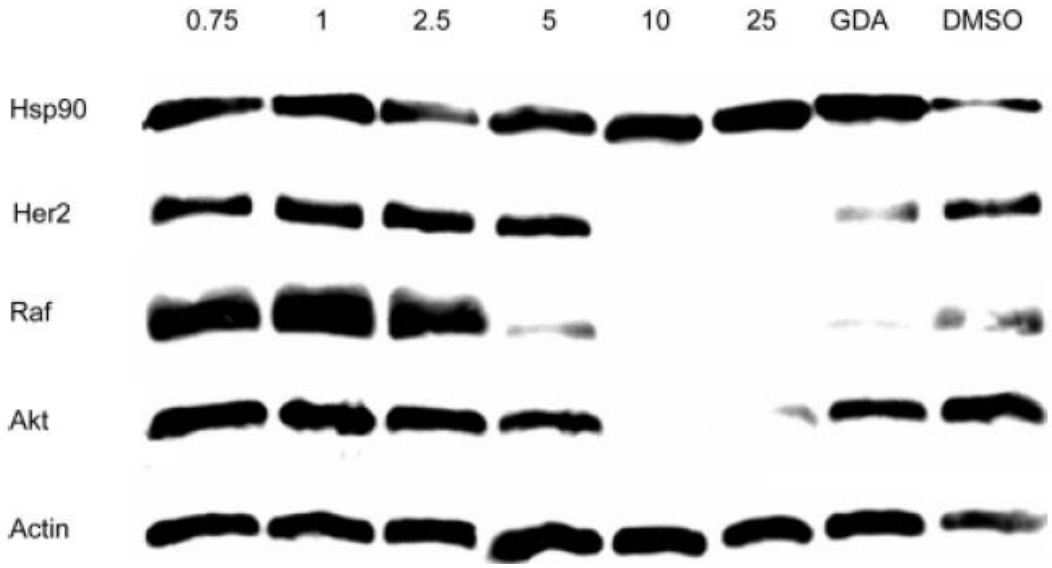


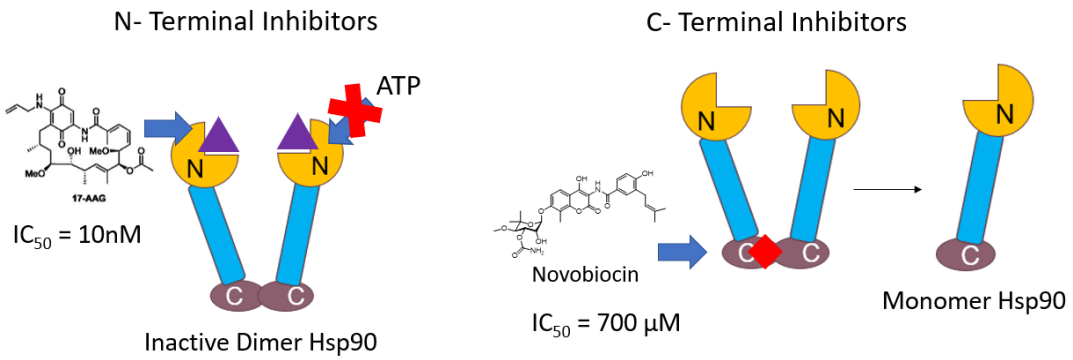
Figure 141: Western blot analyses of Hsp90 client protein degradation. With the increasing concentration of Hsp90 CTD inhibitor, a fade away of Hsp90 client protein (Her2, Raf, Akt) band was observed.

The concentration of novobiocin based analogues (in μM) is denoted above each lane. GDA (geldanamycin) and DMSO were used as positive and negative controls, respectively (Burlison, Avila et al. 2008).

This method of Western blot in Figure. 141 has many flaws in regard to the validation of Hsp90 CTD drug binding. For example, it never showed the direct reduction or disassociation of Hsp90 dimer protein. As a result, the molecule could also bind to other proteins and add up to its growth inhibitory effect. Also, the western blot method never tells the difference between Hsp90 NTD inhibition and Hsp90 CTD inhibition. As shown in Figure 141, both Hsp90 CTD inhibitor and N-terminal inhibitor GDA had very similar effects in terms of reducing Hsp90 client protein expression. In contrast, a novobiocin analogue could be a potential 'Hsp90 NTD inhibitor' after the chemical modification.

The reason why western blot method would never show the reduction of Hsp90 is due to Hsp90's dimeric activity.

In the introduction section, we discussed that Hsp90 only exhibits ATPase activity under the dimeric form, controlled by dimerization C-terminus. Hsp90 CTD inhibitors can bind to Hsp90 CTD, therefore, block the formation of its active dimeric form. Hsp90 NTD adopted a different well-studied mechanism of binding through the Hsp90 NTD ATP binding pocket (Scheme 48). This process does not affect the dimerization process (Cruz, Zhang et al. 2013). As a result, the potency of Hsp90 CTD inhibition should be visualized by manifesting the reduction of Hsp90 dimeric form into its monomeric form.



Scheme 48: Hsp90 NTD inhibitor and Hsp90 CTD inhibitor have different mechanisms of inhibition. While Hsp90 NTD inhibitors competitively binds to Hsp90 NTD ATP binding pocket, Hsp90 CTD inhibitors binds to CTD and disrupts the dimerisization.

Western blot method uses SDS-page (sodium dodecyl sulfate polyacrylamide gel electrophoresis) to separate proteins according to their molecular weight. SDS (sodium dodecyl sulfate) is used to break any steric conformation of the protein and drag them into a linear peptide. Consequently, the original uninhibited Hsp90 dimer or Hsp90 CTD inhibited monomer will have no difference but form the same band located at 84 kDa (Figure. 142, left side). To measure the CTD inhibition, the protein sample must be preserved in the native dimeric form to distinguish from its inactive monomeric form. Native polyacrylamide gel electrophoresis (native page gel) is the only option in this case. This method was developed for the analysis of proteins' native form and separating them without denaturing its oligomeric quaternary structure. In commercial Hsp90 human alpha protein sample, most of the protein was in bioactive dimeric form showed at 180 kDa position. In comparision, the minor inactive monomer is shown at 90 kDa position (Figure. 140, right side). With the incubation of Hsp90 CTD inhibitors, it is expected to see an increase of 84 kDa band and a fade-away of 180 kDa band. The more 84 kDa band is counted, the better the inhibition ability.

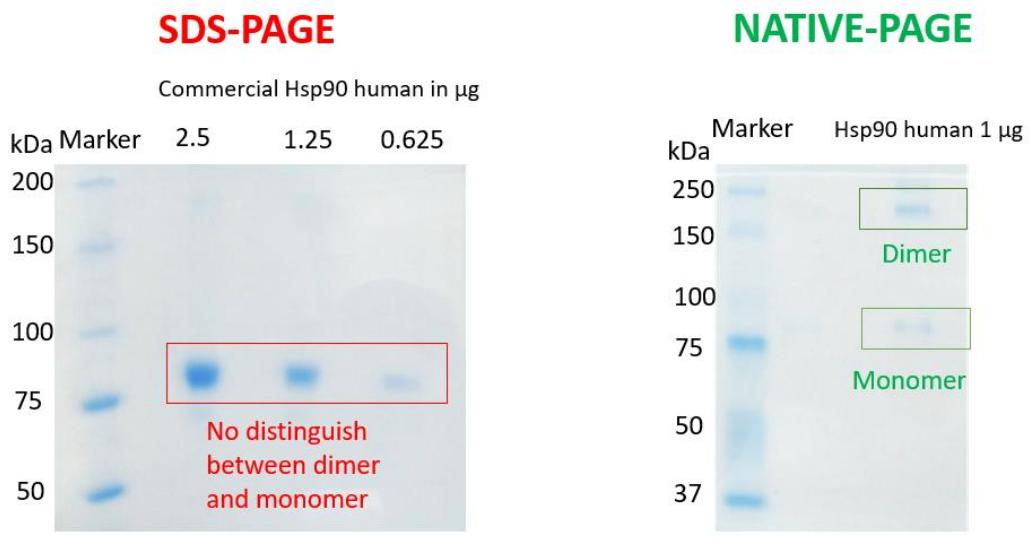


Figure 142: Pure commercial Hsp90 protein. SDS page will not distinguish between dimer and monomer.

Previously, our research group developed an Hsp90 binding assay measuring the binding of Hsp90 whole protein with Hsp90 CTD inhibitors (Cruz, Zhang et

al. 2013). In this native page gel method, drug candidates and Hsp90 were incubated under 37 ˚C overnight. Then the samples were separated using 10% native page gel and visualized by instant blue dye for qualitative analysis. It can also be used for quantitative analysis if the sample was prepared in gradient concentrations and visualized by silver stain dye. Sliver stain quantitatively dyes the protein band and can be quantified through optical density (Cruz, Zhang et al. 2013). However, by the time this project was ongoing, the optical density meter no longer existed. Thus, the native page gel method was only used for the qualitative validation of Hsp90 CTD binding.

This method was also validated against Hsp90 NTD inhibitors and confirmed that the method would not show NTD inhibition. According to our previous result, Hsp90 CTD inhibitors will visualize its inhibition through a heavily stained monomer band on the gel. Hsp90 NTD inhibitors do not cause Hsp90 dimer to separate, and the gel page results were just identical to the negative control group (Cruz, Zhang et al. 2013).

To carry out the native gel analysis, a supply of Hsp90 human alpha is needed. There were two products of Hsp90 being used in this project. The laboratory produced recombinant Hsp90 human alpha, supported by a collaborator Banghao Yuan from Dr Jasmina Jovanovic's research group. The protocol for protein synthesis can be found in the experimental part (p307). Another protein that was used was the Hsp90 alpha recombinant protein (ab48801) manufactured by Abcam. These two types of Hsp90 protein products shared the exact same sequence with His-tag modification on N-terminal (for protein sequence, see p306). There was another type of Hsp90 C-His-tagged product (ab80369) which was used once but then banned from the experiment, for which the reason will be discussed later (p221, Figure. 154). It is also worth noting that any protein-involved assay in this thesis was using full-length Hsp90 instead of Hsp90 CTD only. Results from experiments using full-length Hsp90 would be more convincing than using Hsp90 CTD protein (also added difficulties compared with using Hsp90 CTD protein only), as inside the cell, Hsp90s work in full length.

At first, we tried to apply lab-produced Hsp90 human alpha for the assay. The concentration of protein was adjusted to 2 mg/L solution, then 2.5 µL of protein was mixed with 7.5 µL solution of the ligand with relevant fixed concentrations and incubated for overnight under 37 ˚C. In order to save the stock of synthetic novobiocin analogues, we used only novobiocin and some other ligands of interest for early method development. This first trial was just to see if the literature method (Cruz, Zhang et al. 2013) works. The compounds being used for the first-time native page trial is given below (Figure. 143):

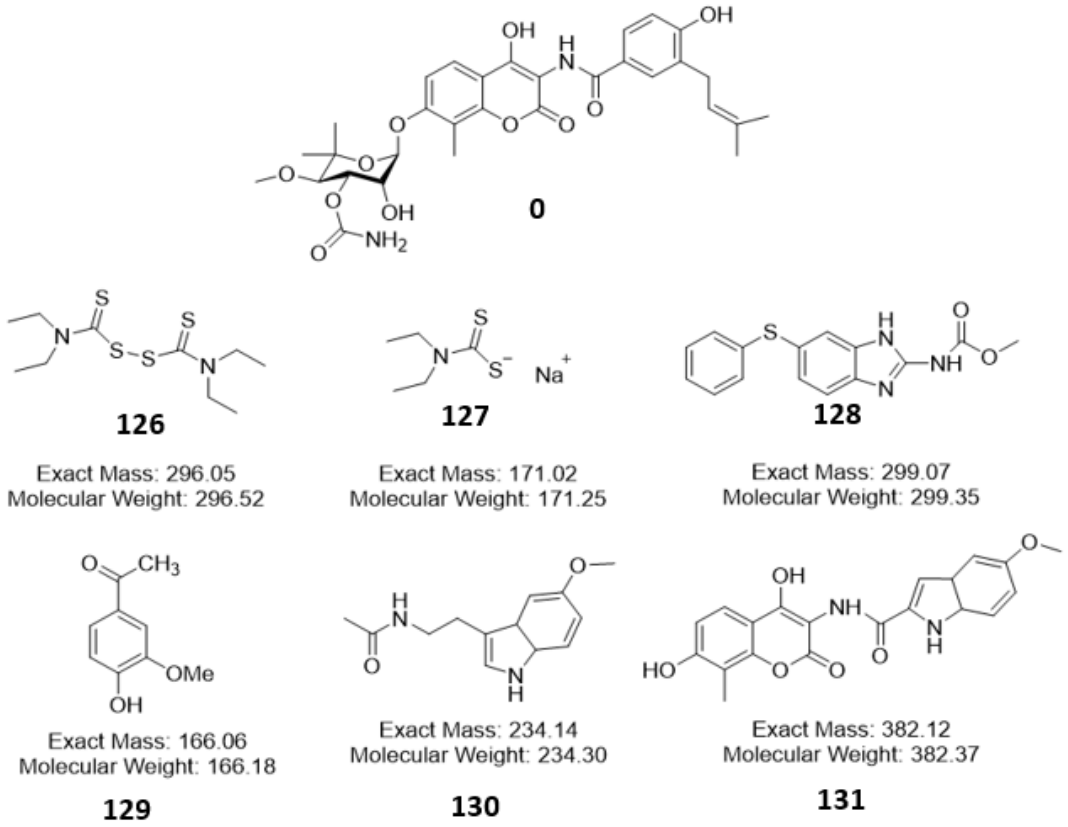


Figure 143: Tested ligand for the first time using native page gel Hsp90 CTD assay.

The formulation of polyacrylamide gel and the value of constant current during electrophoresis greatly affects the performance of the assay. We firstly adopted the recipe of gel same as the literature (Cruz, Zhang et al. 2013) by preparing 10% polyacrylamide gel and running the electrophoresis under 8 mA. It is important to note that native gel usually requires a much lower current and much longer duration of electrophoresis for good results. This is very

different from the SDS page where it could be as fast as TLC in chemistry, which is to ensure the native structure of the protein are not affected by either electric field or heat generated during electrophoresis. Finally, after approximately 6 hours, the result showed that the protein was not well separated (Figure. 144):

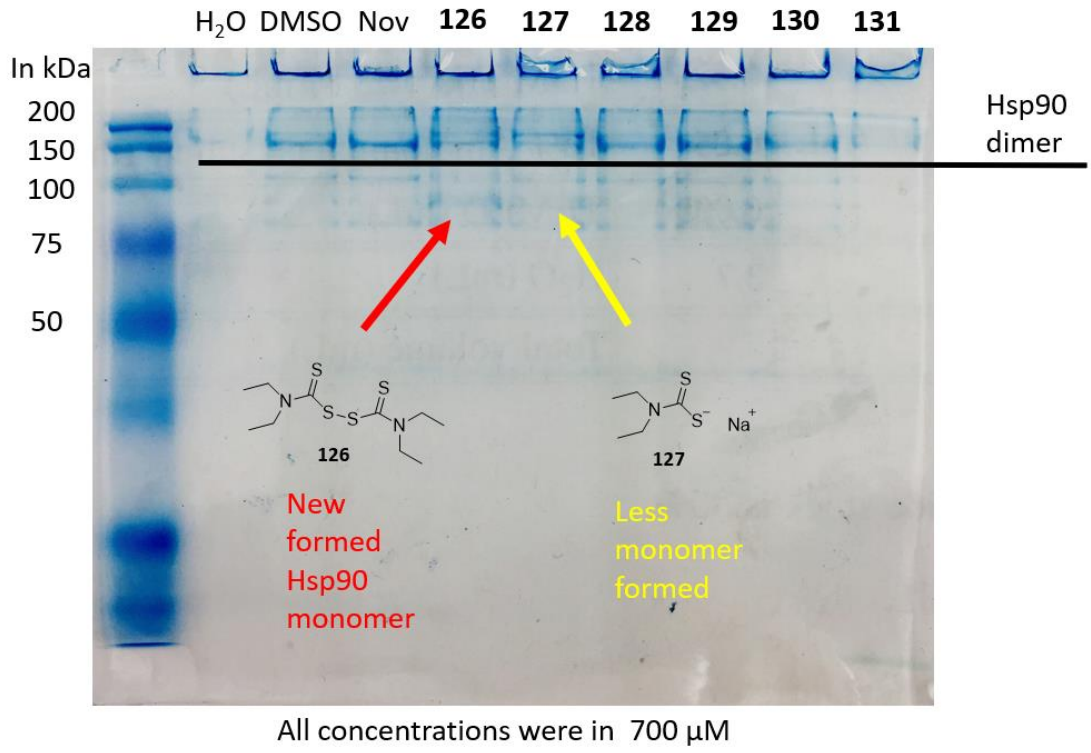
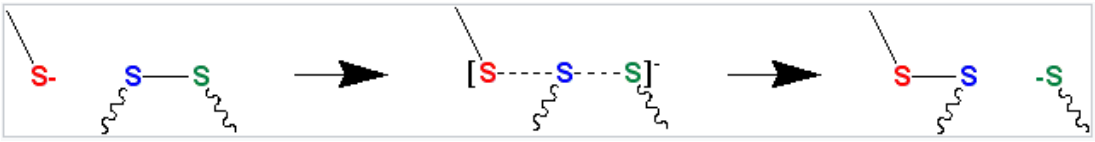


Figure 144: 1st trial of drug-Hsp90 binding native gel binding test.

From the distribution of marker, this gel condition (10% polyacrylamide) could separate low weight proteins under 100 kDa. Since we were looking at Hsp90 dimer at 180 kDa and Hsp90 monomer at 90 kDa, the percentage of polyacrylamide needs to be decreased for larger weighted protein to travel smoothly. In Figure. 144, the water control group and tested compound **130** and **131** were observed with regards to aggregation, and stacking at baseline, which did not flow throw electrophoresis; suggesting that percentage composition of acrylamide should be lower.

At this time of method development, the MTT test results were not yet returned from the collaborator. Therefore, we initially aimed to use novobiocin as a positive control under a reported 700 µM IC_{50} in the literature (Burlison, Avila et al. 2008). However, 700 µM novobiocin failed to show a significant Hsp90

CTD inhibition in this trial (as well as the rest of the native page gels in this section at this concentration). The low CTD binding activity of novobiocin shown here can be reasonable, as we reported the mild anti-proliferative test result of IC_{50} = 1339 µM (against MCF-7). Nevertheless, we surprisingly found that molecule **126** disulfiram showed a potent Hsp90 CTD inhibition in this case. The result even showed the importance of disulfide bond for C-terminal inhibition, as its precursor form diethyldithiocarbamate **127** showed much weaker inhibition. The disulfide bonds can exchange between cysteines (Scheme 43) (Futami, Tada et al. 2000). As a result, the inhibition power of disulfiram **126** may come from covalent inhibition of active cysteine residue on Hsp90. We added compound **126** and **127** into the first trial with an aim to investigate whether disulfide bond exchange could happen, but it was surprising to see this may have happened.



Scheme 43: Mechanism of thiol-disulfide exchange in cells.

Disulfiram, a very safe drug used for the treatment of alcoholism, was reported to have antiproliferative activity as well via p97 inhibition by its ability of Cu^{2+} coordination (Skrott, Mistrik et al. 2017). However, its connections to Hsp90 protein are yet to be reported (especially p97 and Hsp90 work in line in terms of deleting misfolded proteins). It would be promising to launch a further investigation to design and test more disulfiram analogues as a novel type of Hsp90 CTD inhibitor and explore its pharmacology effect of Hsp90 inhibition. As it was not the aim of this thesis, we did not continue this research idea. However, this was an indirect proof of the potential of Hsp90 covalent binding.

We then lowered the concentration of the 10% native page gel to 8% and prepared a second trial using 4'-substituted novobiocin analogues and novobiocin (Nov **0**) (for structures of tested compounds in second trial, see Fig. 145).

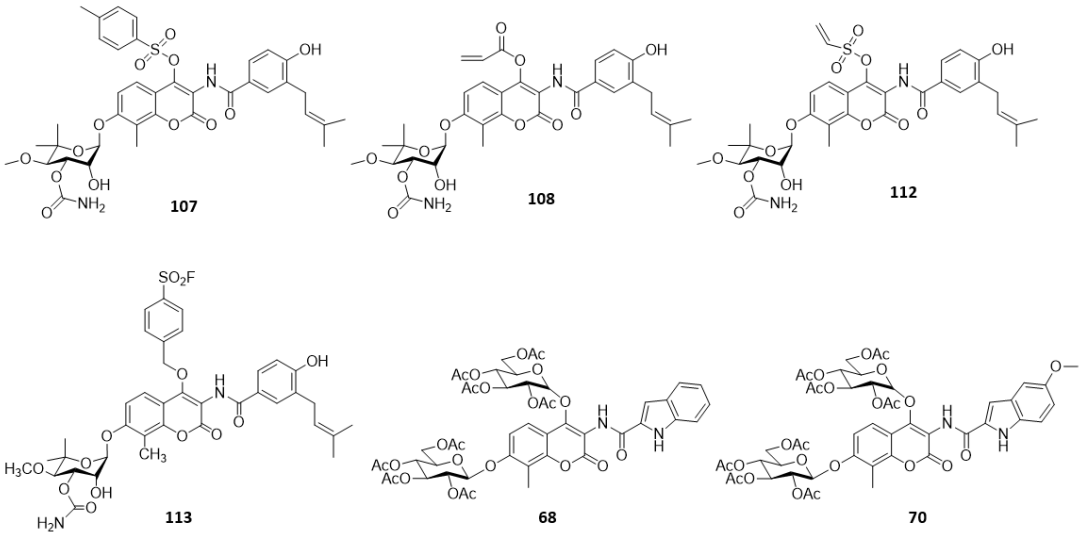


Figure 145: Tested ligands for the second trial of Hsp90 CTD native page gel binding assay.

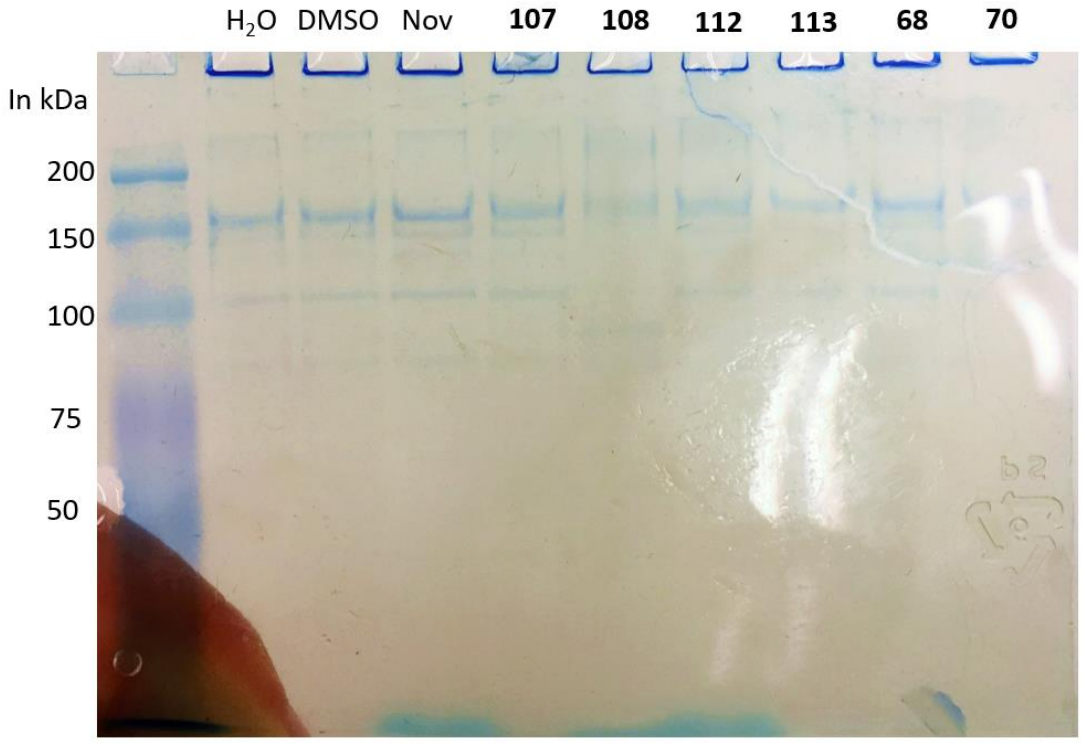


Figure 146: Hsp90 CTD native gel binding assay in the second trial.

The gel was prepared, sampled, and run for 6 hours. The resulted gel page (Figure. 146) was improved in terms of separation of proteins > 100 kDa, which

was adequate for observation of Hsp90 dimers. To view the results clearly, the gel was silver stained (Figure. 147).

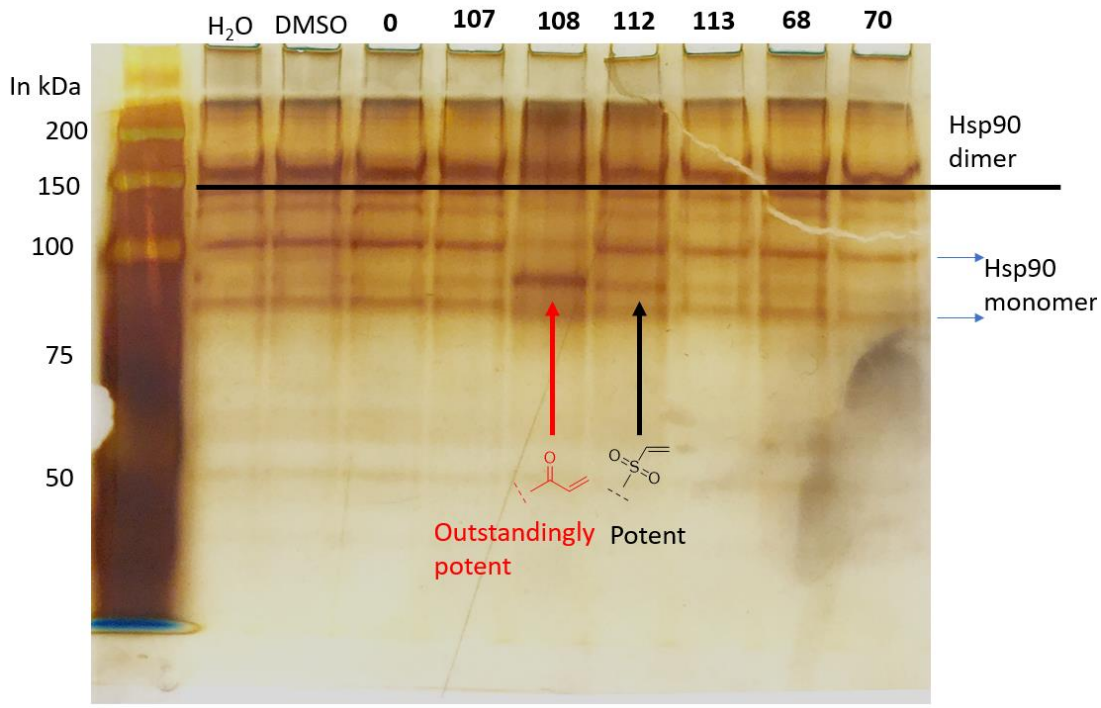


Figure 147: Hsp90 CTD native gel binding assay second trial silver stained. A fixed concentration of ligands = 700 μM

From this result, it can be clearly seen that most of the Hsp90 under native form were in dimeric form, as suggested in the literature (Cruz, Zhang et al. 2013), and 84 kDa monomeric form can be found between 75 kDa and 100 kDa marker. There was another protein that existed at 100 kDa. While literature (Cruz, Zhang et al. 2013) merely reported these two bands as “two types of Hsp90 monomer”, we cut the relevant gel band then sent this for proteomic analysis (this work was done by Dr. Kate Hessom from the University of Bristol). MS-MS spec database searching confirmed those two bands were all Hsp90 alpha human, not impurity protein from bacteria. For native page gel, bands are distributed according to the weight and steric conformation of the protein; therefore, 100 kDa band might not have an actual 100 kDa weight (for SDS-page, proteins are distributed by weight only, as they lost native steric confirmation). This suggested that there were two types of native form Hsp90 monomer.

Although **107** and **113** were the two most potent compounds in the MTT assay, they did not show Hsp90 CTD inhibition in this test, most likely because their binding pocket or drug target had changed with such chemical modification. Novobiocin, still showed no CTD inhibition under 700 μM concentration. The main breakthrough in this test was the confirmation of CTD inhibitor Michael acceptors **108** and **112**. There were two distinguished new-formed Hsp90 monomer bands, and **108** showed much more potent inhibition than **112**. It also can be seen that **108** not only reduce the amount of 180 kDa Hsp90 dimer but also faded the 100 kDa, Hsp90 monomer band. Overall, the second trial proved that Michael acceptor types analogues significantly improved the drug-Hsp90 CTD interaction, revealed by generating more Hsp90 monomers. This strong inhibition activity could highly link with its cysteine covalent binding potential.

68 and **70** were two novobiocin based glycosides analogues. The test result of these two ligands indicated that substitution with unprotected sugar moieties reduces not only antiproliferative activity but also CTD binding affinity.

3.10.5 Hsp90 CTD binding assay results of novobiocin analogues.

Next, to validate if the increased Hsp90 CTD inhibition came from active acrylate moiety, we introduced the inactive 4'-propionyl analogue **109** to the test to confirm this issue. Also, the concentration of novobiocin was increased to investigate the correct concentration of visible Hsp90 CTD inhibition on native page. The resulted gel page was given in Figure. 148.

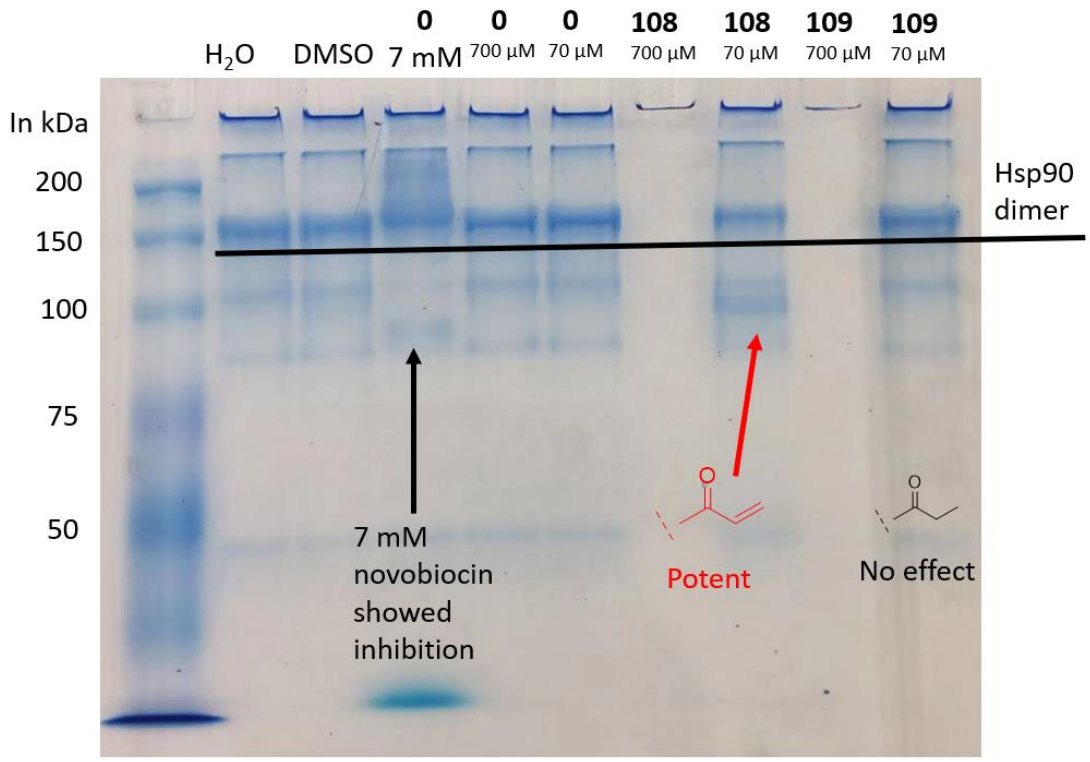


Figure 148: Hsp90 native gel binding assay for novobiocin **0**, **108** and **109.**

The concentration of novobiocin was increased to an extremely high 7 mM before a clear visualization of CTD inhibition could be found, which was abnormal to any reported MTT IC_{50} value (300 – 700 µM). Due to most of the literature' MTT assay was carried out with 72 hours' incubation, the same gel was repeated under the exact same procedure, except the incubation duration increased from overnight to 72 hours. The resulted gel after 72 hours (Fig. 149) was identical to the previous result, meaning that time was not the factor responsible for the low binding affinity of novobiocin **0**.

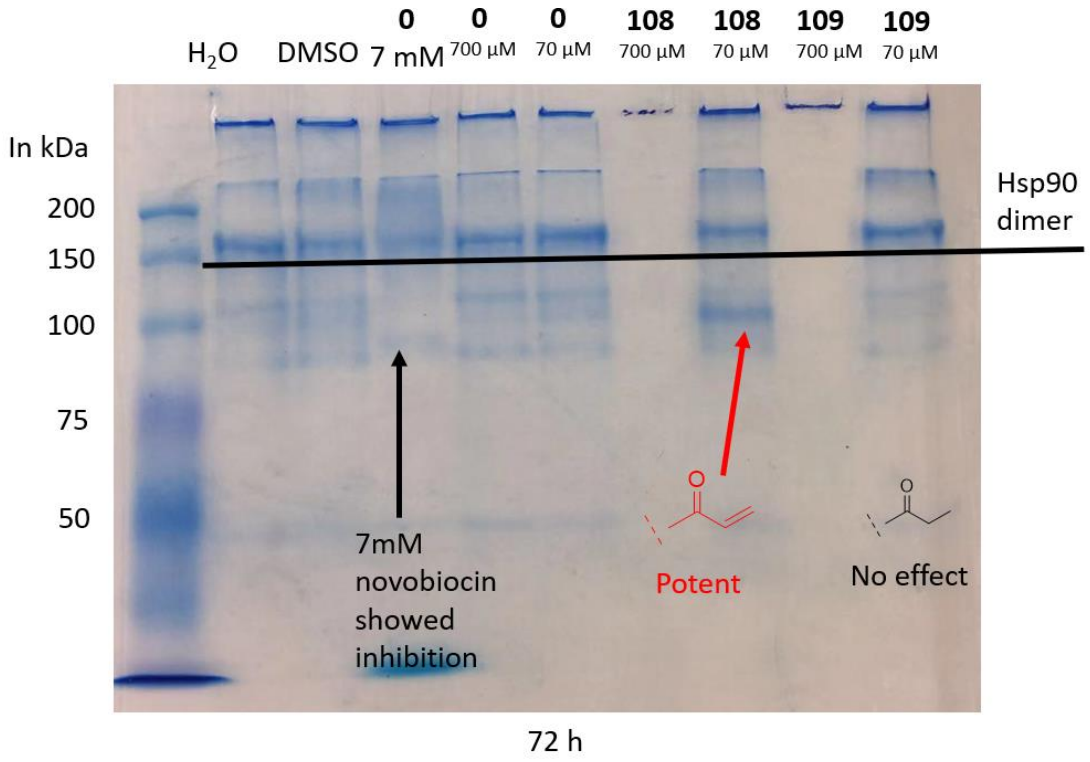


Figure 149: Hsp90 native gel binding assay for novobiocin **0**, **108** and **109**, incubation duration increased to 72 hours.

For active **108** and inactive **109** Michael acceptors, surprisingly, at a very low concentration of 70 µM (- 7.5-fold lower than MTT IC_{50} value), **108** still showed a distinct CTD inhibition power. In contrast, inactive form **109** already lost the sign of CTD binding, which strongly suggested that a covalent inhibition could had happened. However, for some reason, 700 µM **108** and **109** samples did not undergo electrophoresis (in both overnight and the 72-h incubation). It was essential to confirm whether they still have different inhibition results under higher concentrations as **109** could show CTD binding effect with higher concentration. To solve this issue, a few factors were being considered. First, manual mistakes (that forgot to add the drug samples) were investigated. Thus, the experiment was repeated and ensured no manual mistake was taken place. However, the same issue was encountered (Figure.150, 1). Secondly, the Hsp90 protein could be contaminated or denatured and being less sensitive to such ligands. We resynthesized a new batch of Hsp90 protein

and repeated the assay, and again, the problem was still present. (Figure 150, 2)

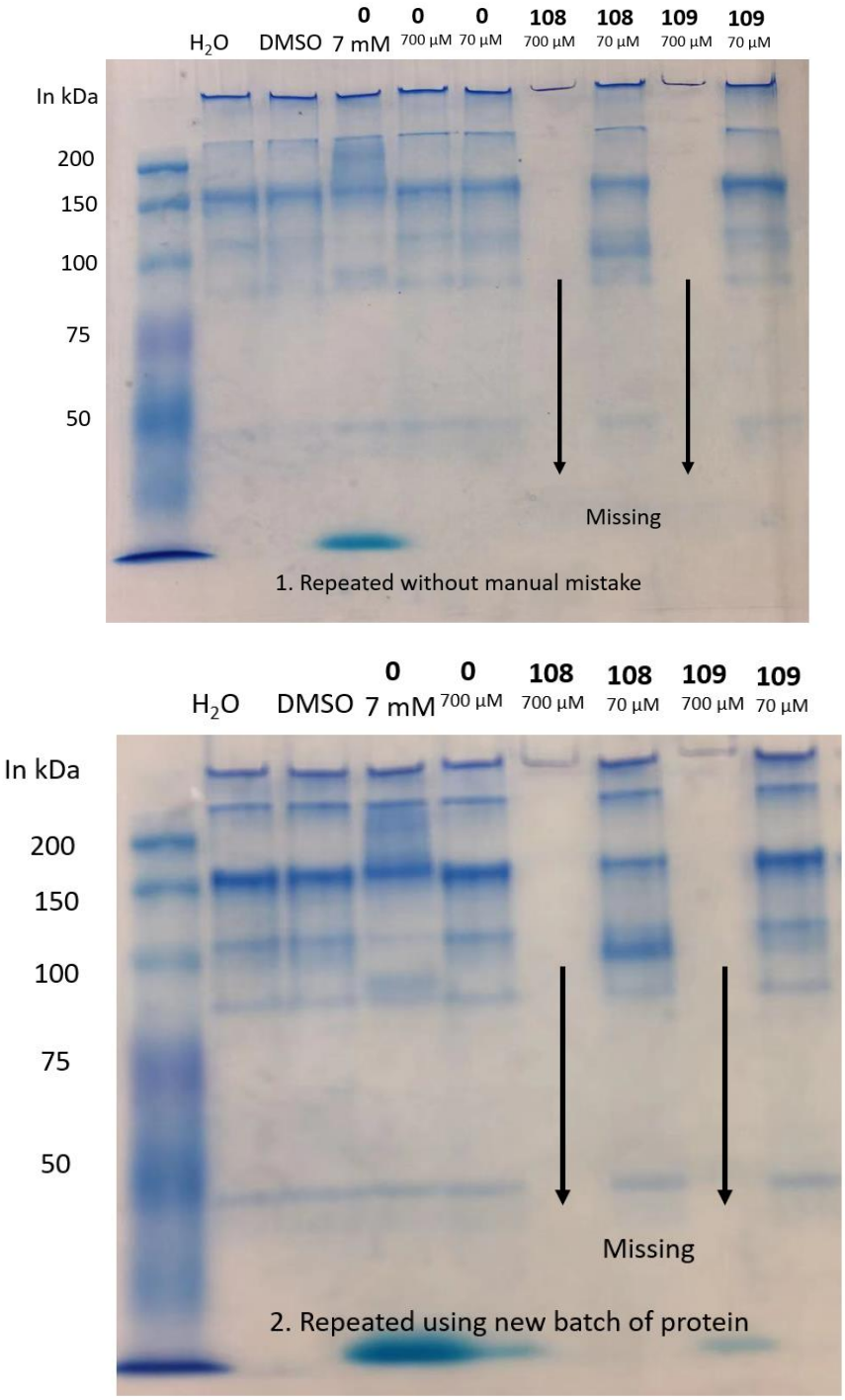


Figure 150: Repeated experiment of Hsp90 CTD native gel binding assay for novobiocin **0**, **108** and **109**.

To investigate if it the issue was caused by a high concentration, we carried out a gradient screening of **108** and **109** from 1 µM to 1mM (Figure.150 and Figure 152).

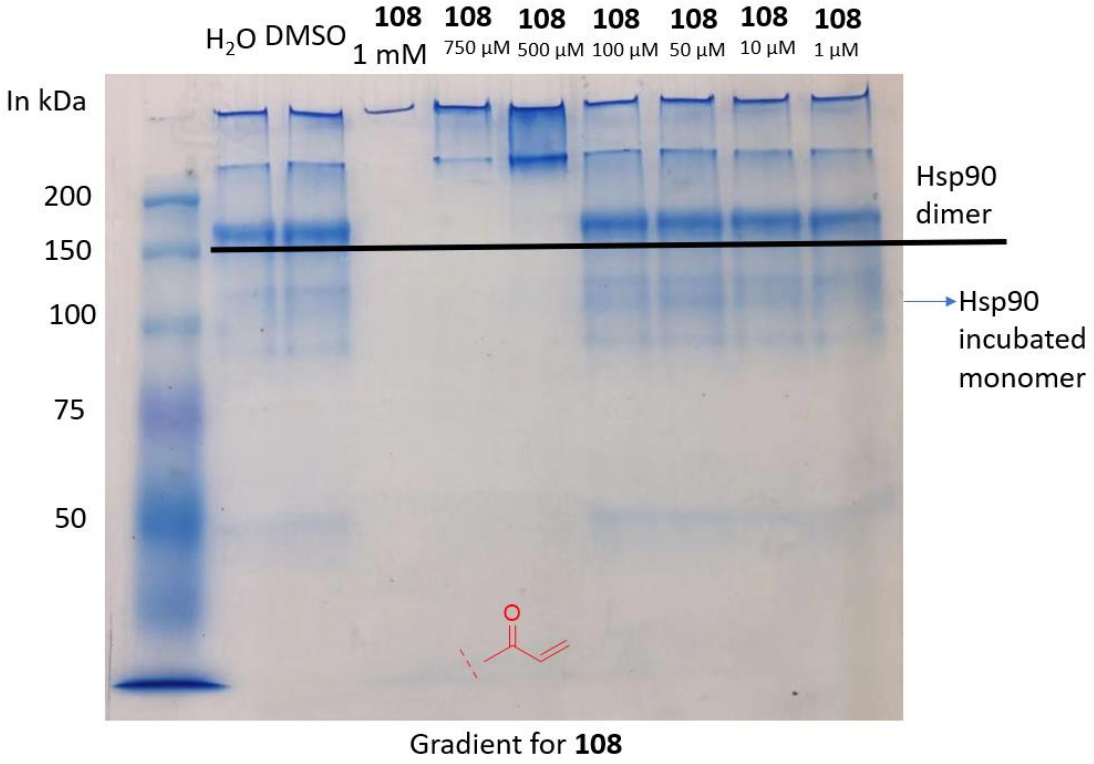


Figure 151: Hsp90 native page gel binding assay for **108** using gradient concentration.

It can be concluded from Figure. 151 that the aggregation happened when **108**'s concentration >500 µM. Moreover, due to **108** inhibited Hsp90 at 50 µM but no inhibitory effect when it went down to 10 µM. It can be deduced that the binding affinity of **108** (IC_{50} value) fails into 10 µM – 50 µM region.

The result of **109** gradient screening was shown in Figure. 152 on the next page.

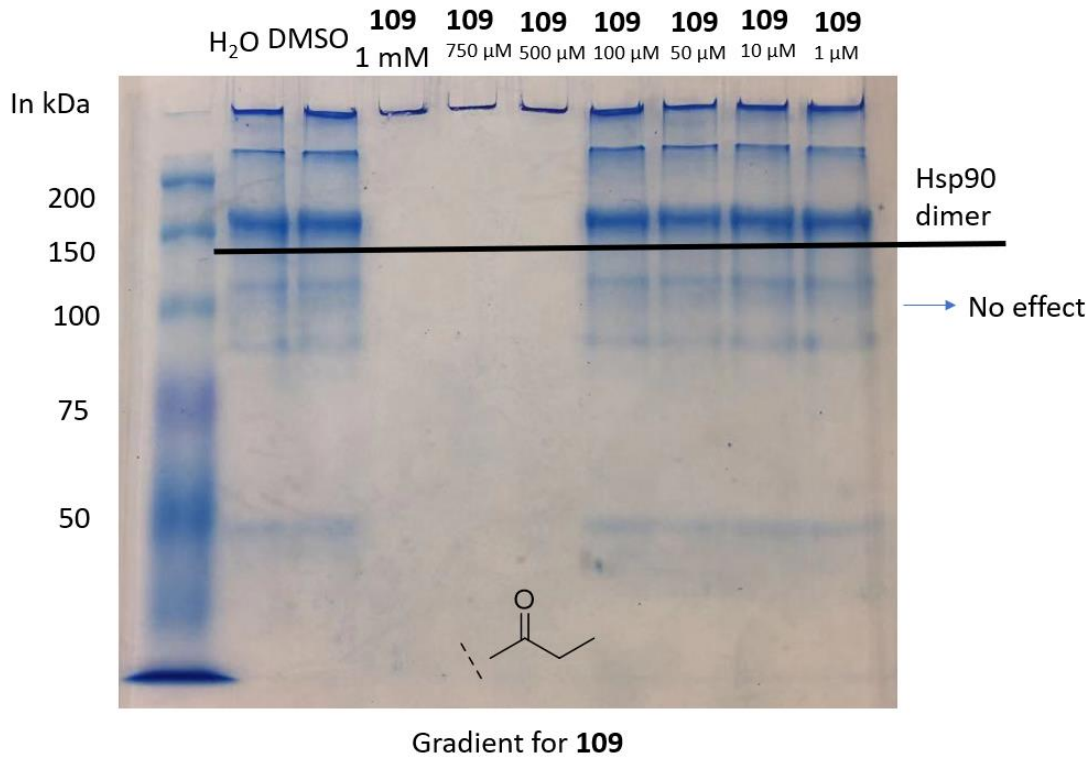


Figure 152: Hsp90 native page gel binding assay for **109** using gradient concentration.

The situation of **109** was the same as **108**. The sample started to aggregate when the concentration of **109** was above 500 µM. Although it seems like it was the nature of the compound that accounts for the Hsp90 protein aggregation, **108** was successfully electrophorized during the method development stage. After careful reviewing of the experimental process, buffer dilution could have been a contributing factor. The batch of protein sample consumed during method development was diluted by PBS buffer while the protein for recent native gels was diluted using HPLC water. Therefore, the test was repeated using same Hsp90 samples, however, added with either PBS buffer or water for comparison. It finally confirmed that the buffer additive was necessary for electrophoresis of the **108** or **109**-incubated protein sample with > 500 µM concentration (Figure 153), as PBS buffer prevents the aggregation.

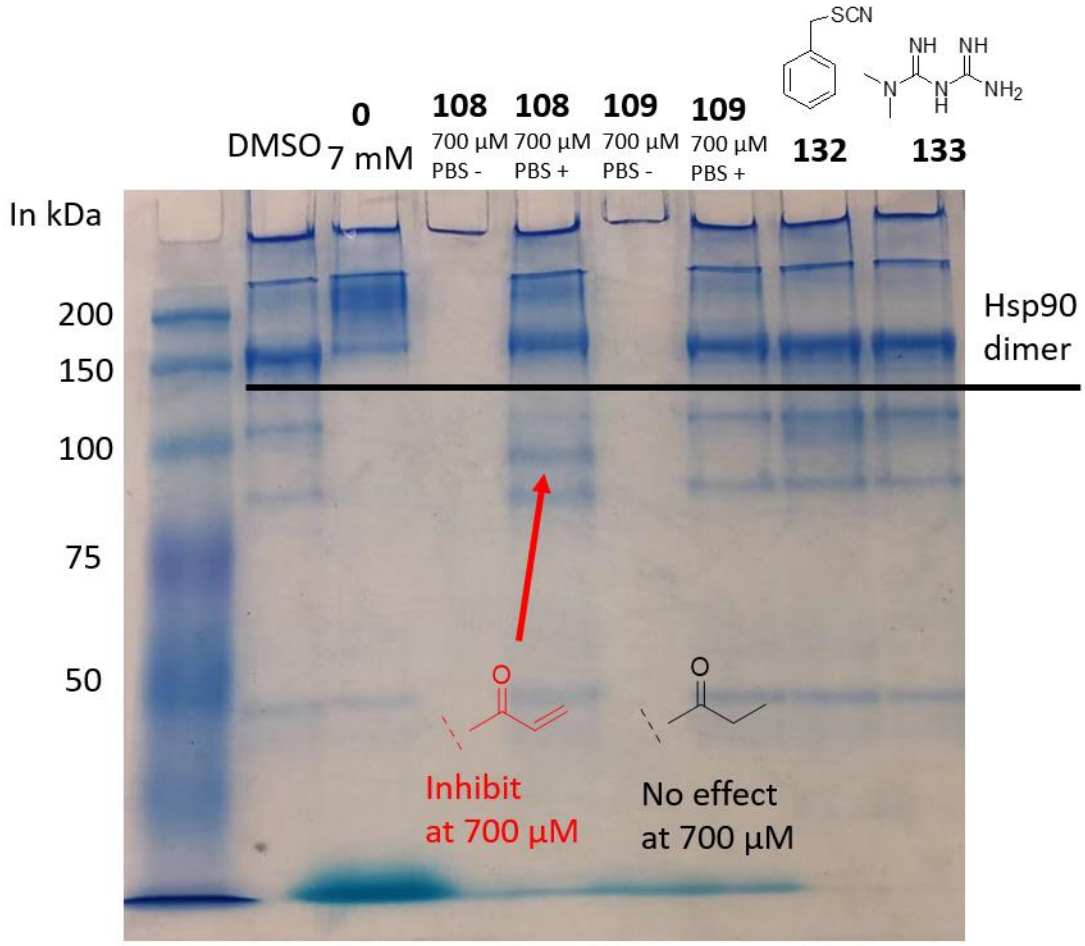


Figure 153: PBS dilution was essential for native page assay.

Entries in Figure 153 which taken by compound **132** and **133** in this trial were used for screening of the other two ligands. The result showed that benzene thiocyanate acts as an Hsp90 CTD inhibitor, while metformin does not affect Hsp90 dimerization.

In this trial, ultra-high (7 mM) concentration novobiocin incubated samples failed the electrophoresis. While **109** showed no binding activity at 700 µM (+1.32-fold of antiproliferative activity), **108** strongly inhibits the bio-active Hsp90 dimer and generates more inactive Hsp90 monomers. Therefore, the varied Hsp90 CTD binding affinity between cysteine reactive **108** and cysteine unreactive **109** could highly because of extra covalent binding potency of **108**. The incubated sample warrants further investigation through proteomic analysis to determine the binding site.

108 showed visible protein binding activity at 700 µM, 100 µM, 70 µM, and 50 µM. Although the optical absorption for each band could not be quantified, there was a huge gap of **108**'s IC_{50} between cell level MTT assay and protein level protein binding assay. **108** was no doubt to be more potent on protein level (50 µM > IC_{50} > 10 µM) than cellular level (IC_{50} = 537.6 µM MCF-7). Data from drug-Hsp90 binding assays suggested that 4'-acrylate moiety may detoxed by cells before it reached the targeted Hsp90 CTD.

To ensure our results were reliable and repeatable, we purchased commercial Hsp90 alpha human protein from Abcam and tested against compound **108** and **IA2** again.

The first product we purchased was ab80369 and found that **108** lost CTD binding activity (Figure. 154).

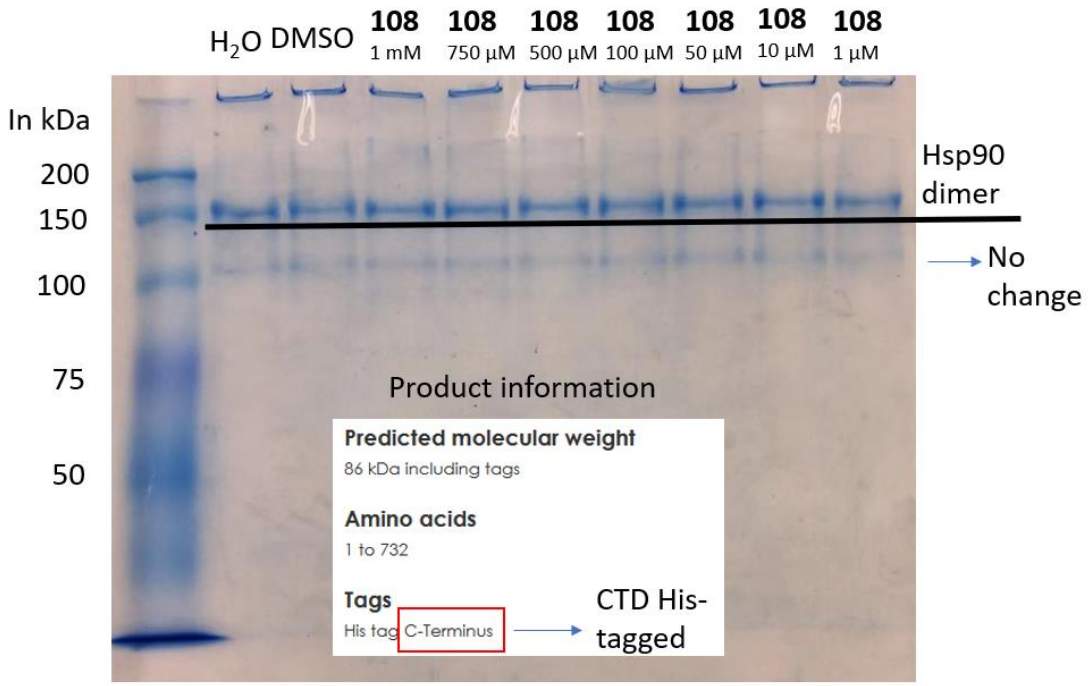


Figure 154: **108** lost Hsp90 inhibition towards commercial CTD His-tagged Hsp90 protein.

The addition of His tag towards Hsp90 CTD will not affect the ATPase activity (tested by Abcam). Because we were the first to test ab80369 using the native page gel method, there was no more information could be provided from the supplier. In order to find out whether the binding affinity was affected by CTD

His tag modification, NTD His-tagged Hsp90 protein product ab48801 was purchased and tested. The results are shown in Figure. 155.

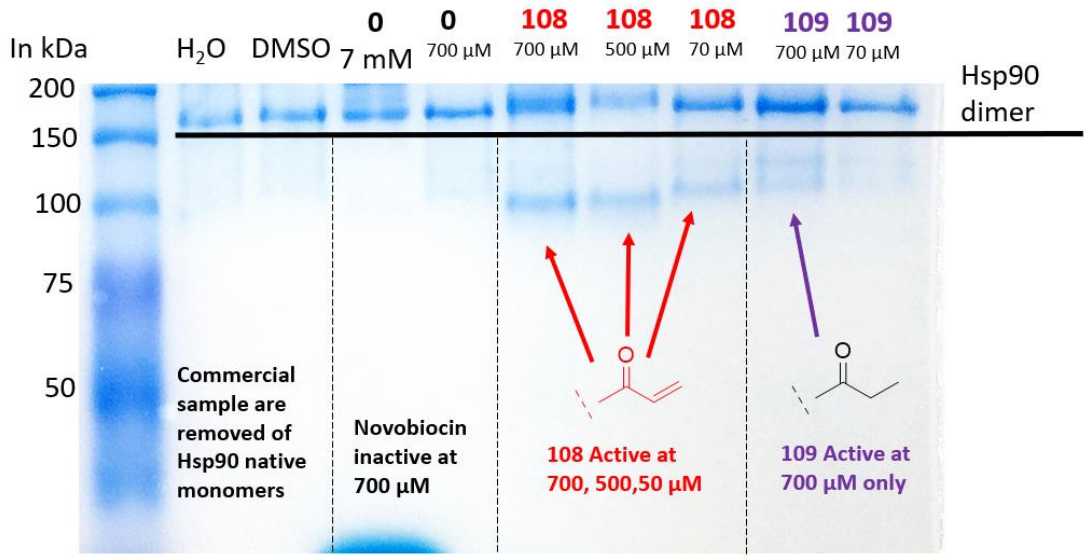


Figure 155: Hsp90 native page gel binding assay for **108** and **109** using commercial NTD His-tagged Hsp90 product ab48801.

Due to ligands only exhibiting protein inhibition effect towards Hsp90 NTD His-tagged products, it can be concluded that CTD His-tagged modification on Hsp90 would significantly interfere with its dimerization property, making it harder to disassociate from dimer to monomer. Any tests which involve in Hsp90 dimerization should avoid using CTD His-tagged Hsp90 protein.

As illustrated in Figure. 155, commercial Hsp90 protein was purified by FPLC to remove its inactive monomers. Inhibition activity against commercial samples only refer to the dissociation of Hsp90 dimer into monomers, which manifested through the formation of a new monomer band. Novobiocin **0** did not inhibit Hsp90 dimer at 7 mM and 700 µM. **108** performed similarly to the previous test, strongly inhibited Hsp90 dimer at 700, 500 and 70 µM. Despite that **109** still showed no activity at 70 µM, it displayed some degree of inhibition at 700 µM, which was weaker than **108**. Moreover, **108** tends to dissociate Hsp90 dimer into two monomer bands instead of one. The reason for this phenomenon may be due to the difference between Hsp90 monomer inhibition. Details will be discussed together with MST tests in the next section (p252).

The varied Hsp90 inhibition between **108** and **109** was a potential sign of successful covalent binding. Also, **108** was much more active at the protein level than the cellar level. Therefore, to discover possible factors which may detox or modify **108** and result in loss of activity, two mini experiments were designed to test the stability of acrylate electrophilic warhead.

4'-acrylate novobiocin **108** stability test 1:

Glutathione detox acrylate through reduction.

Glutathione (GSH) is a bio-synthesized antioxidant found in humans, fungi, plants, and some bacteria (Pompella, Visvikis et al. 2003). The active thiol group of glutathione can reduce many kinds of oxidative stress, including peroxides, heavy metals, free radicals, etc. Cancer cells can overexpress this molecule as protection against chemotherapies (Balendiran, Dabur et al. 2004). Therefore, 4'-acrylate novobiocin could have been reduced, resulting in a loss of its covalent modification activity.

Herein, we designed a reaction of 30 mg **108** mixed with 1 eqv. of glutathione in water under 30 °C for 24 hours. The resulted crude product was analyzed by LC-MS to find out if the reaction had happened.

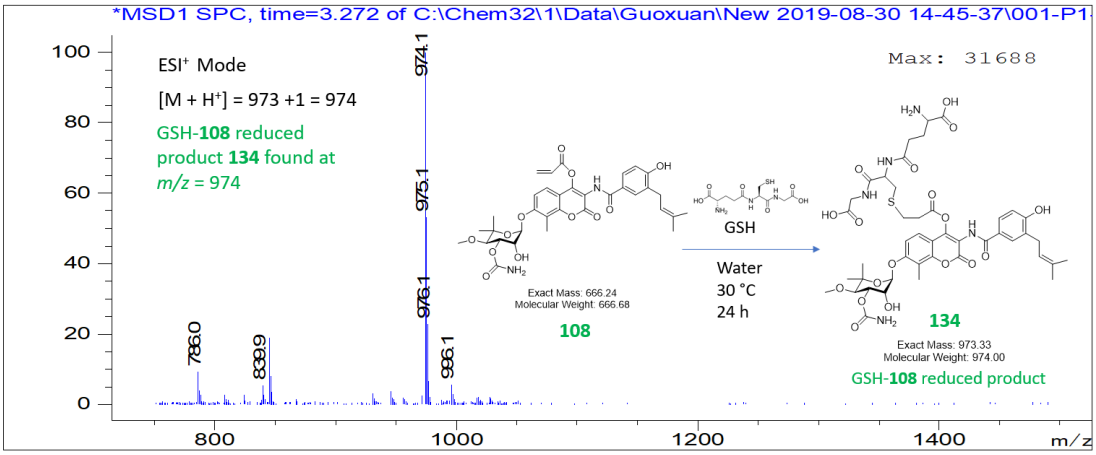


Figure 156: Glutathione reduced **108** into the inactive form.

GSH-**108** reduced product **134** was characterized by LC-MS; the product mass of **134** was calculated and found at *m/z* = 974 Da (Figure 156). Our test result illustrated the reduction of **108** by glutathione, provided a possible

mechanism for drug metabolism, and explained the nerfed drug potency under the intracellular environment.

4'-acrylate novobiocin **108** stability test 2:

Ester bond cleavage test

Along with the easy-react alkene moiety of acrylate, the ester bond of acrylate is also unstable to enzymes such as esterase, which are commonly found throughout the body (Fukami and Yokoi 2012). This factor was usually applied to the design of the prodrug (Stepan, Mascitti et al. 2013). To test whether esterase would cleave **108** into other products, one drop of commercial pork esterase (used in the pharmaceutical industry to test the bond stability (de Maria, Garcia-Burgos et al. 2007) was added to 20 mg **108** dissolved in water, and the mixture was shaken for 24 hours at 37 °C.

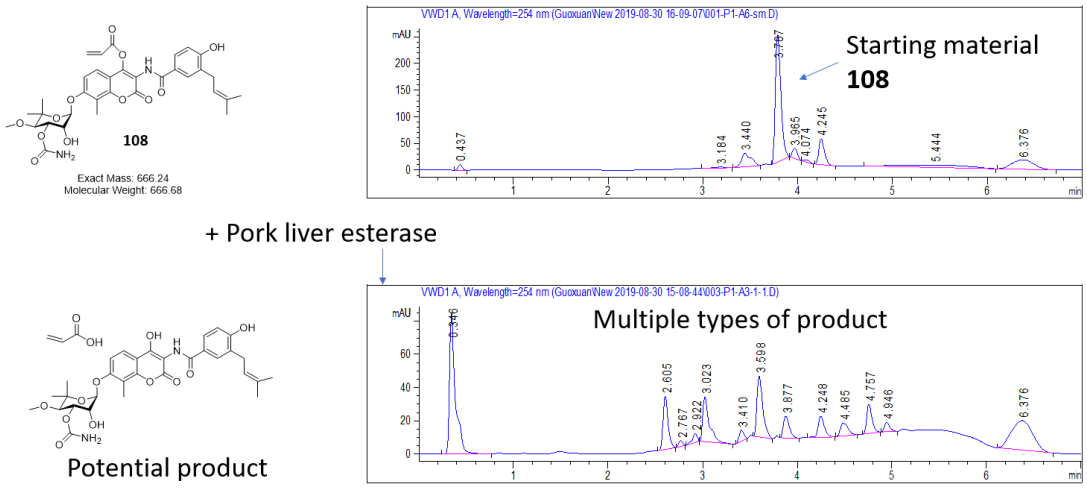


Figure 157: **108** was cleaved by pork liver esterase into multiple products.

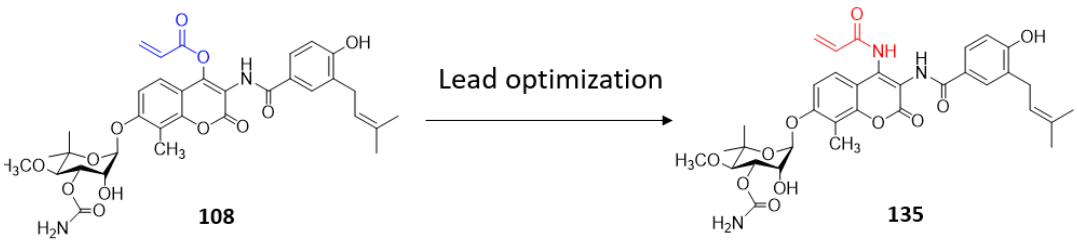
After 24 hours incubation, LC trace of the reaction (Figure. 157) demonstrated a decreased UV absorption of starting material and the formation of new products. We attempted to characterize the peaks of the product by MS spec, which proved to be too complicated to confirm any structure. Pork esterase can, not only break 4'-ester linkage but also other parts of the **108** molecule. As a result, the ester bond of **108** would be cleaved by cancer cells then lose its covalent modification activity.

In conclusion, we designed two mini-experiments to reveal the instability of acrylate moiety, which may be a good explanation of the contradiction between protein and cellar activity. **108** was more potent than its inactive form **109**, but they exhibited almost the same antiproliferative activity due to either bio-detoxing or cleavage of the ester bond. These two molecules would act as the same non-covalent inhibitor once the acrylate group was reduced or cleaved.

3.3.11 Synthesis of 4'-cinamic-novobiocin analogues to improve the instability of acrylate.

From Hsp90 native page binding tests, we concluded that Michael acceptor type novobiocin analogues might covalently bind to Hsp90 CTD. In contrast, the covalent binding pathway was inhibited by cellar common detoxing pathways and lost part of the antiproliferative activity.

A similar difficulty was also brought up in the history of covalent drug discovery targeting other oncoproteins. The first solution was to substitute the acrylate to acrylamide. As mentioned before, acrylamide presents a chemically stronger linkage of an amide bond instead of the easy-degradable ester bond. Besides, acrylamides are weakly electrophilic and even sometimes unreactive towards thiols, nevertheless, making it the most successful electrophile used in targeted covalent inhibitors (Jackson, Widen et al. 2017). We can see from previous results that if the covalent warhead was too reactive, then it will be detoxed by cellar enzyme or other factors before it could reach the targeted cysteine residue, explained the necessity of reactivity (Scheme 49).



Scheme 49: Designed next generation acrylamide novobiocin **135** analogues for better antiproliferative activity than **108**.

Synthesis of 4'-amino novobiocin **135** was not straightforward. Therefore, this plan was not applicable due to the limited resources of this project; this modification will be discussed in the future work section.

Another design which do not change the constitution of acrylate scaffold will be adding high-hydrophobicity substituents to the ethene group, for example, benzene. There are two reasons for such modification. First, it will slow the native hydrolysis of acrylate by pushing away the water molecule (Jackson, Widen et al. 2017). Second, it will decrease the reactivity with GSH. For example, Mondal et al. tested and compared the reactivity of acrolein with cinnamaldehyde towards GSH and cinnamaldehyde. While cinnamaldehyde showed good resistance towards GSH reduction, acrolein was a stronger electrophile (Figure. 158).

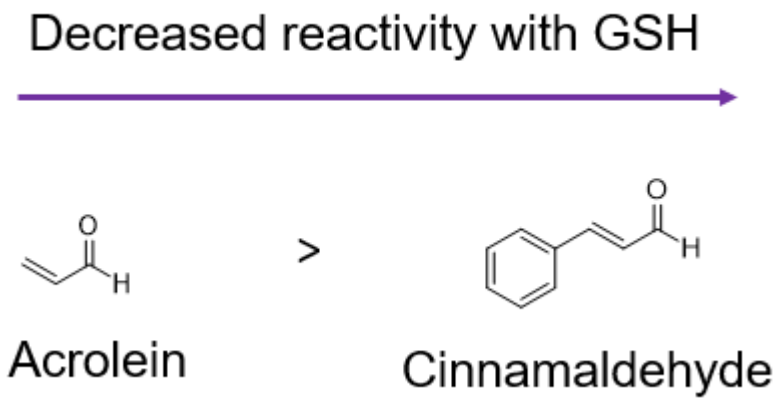


Figure 158: Relative experimental rates of GSH addition to α, β-unsaturated carbonyls.

Because cinnamic moiety is more stable then acrylate moiety, we designed several novobiocin based 4'-cinnamic-novobiocin analogues (Figure. 159), synthesized using the same method developed previously.

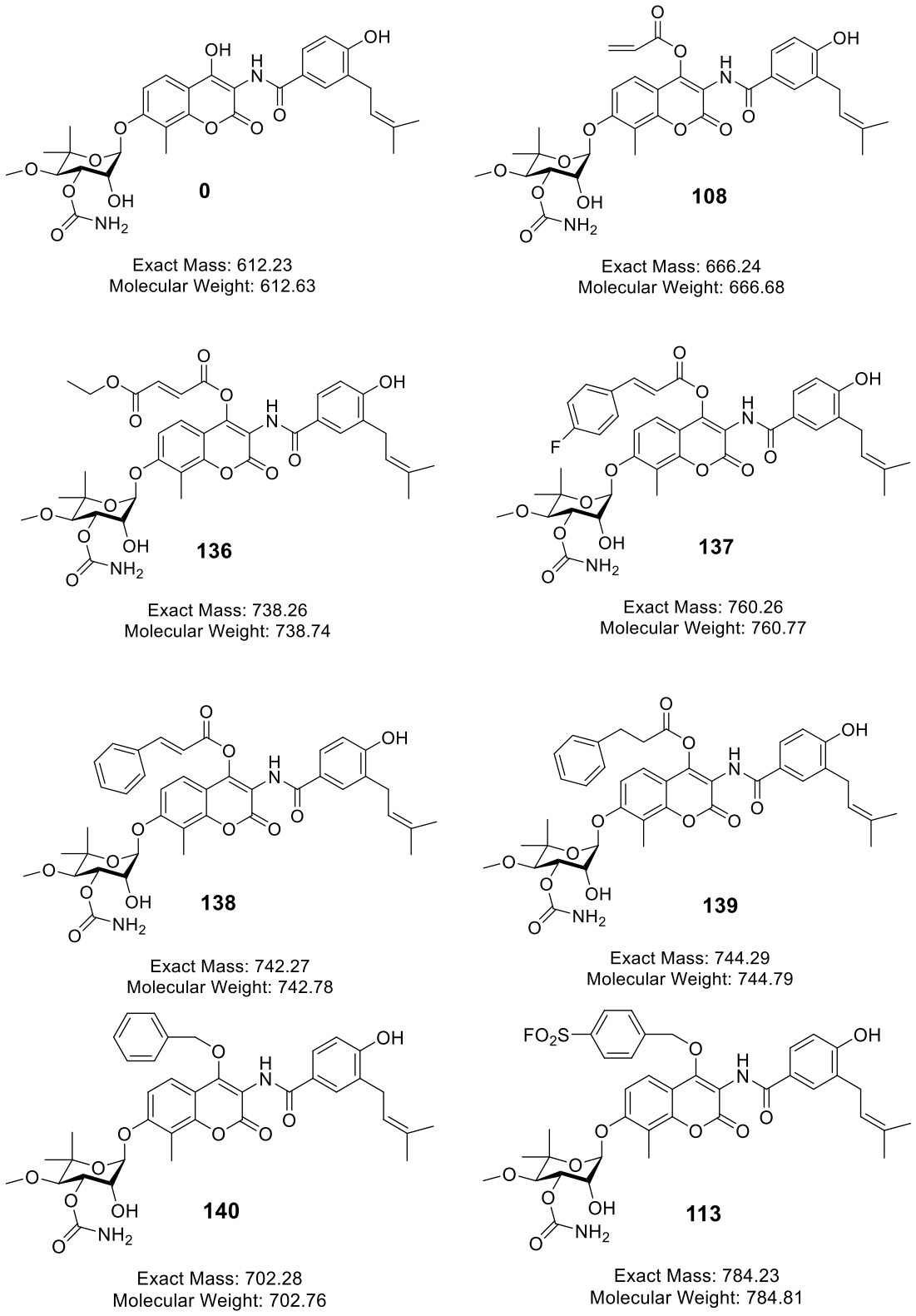


Figure 159: Second batch compounds which were sent for 48-h MTT tests.

MTT tests of these compounds (in Figure 159) were handed over to Miss Sizhu Lu from the University of Oxford. Apart from new synthesized cinnamic

novobiocin analogues **137** and **138**, hydro-cinnamic novobiocin **139** was added as the same purpose as the adding of **109** to **108** for covalent power validation. **136**, 4'-ethyl fumaroyl-novobiocin, was different from **108** with an ethyl ester substituent on β proton. This modification was expected to have more potent activity because the α proton of acrylate was also activated as an electrophile, just the same as the β proton. **140** came from the **113** scaffold missing the -SO_2F covalent warhead. As **113** was the most potent drug candidate in the previous test, it is worth finding out whether its antiproliferative activity was because of the SO_2F moiety or just the addition of benzene ring.

According to our collaborator (Miss Sizhu Lu), the fastest available cell line was MCF-7. We decided to increase the incubation time from 24 h to 48 h, as novobiocin showed poor antiproliferative activity under 24 h. Figure. 160 showed the dose-response curve of the 48 h MTT assay. Miss Sizhu Lu also carried out data analysis and calculated IC_{50} values for each compound.

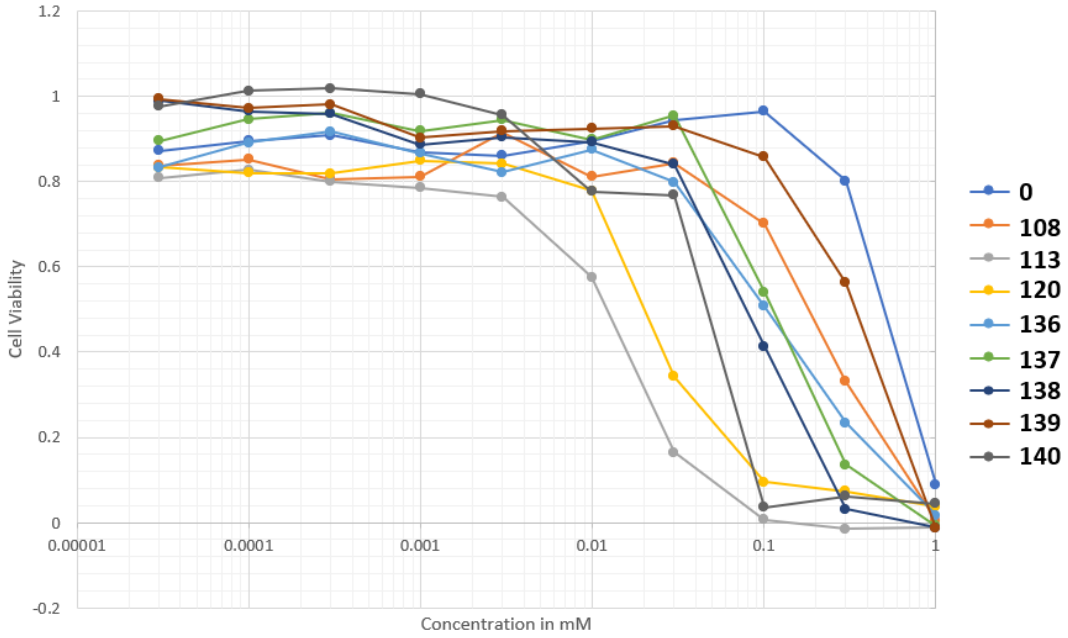


Figure 160: MTT test result of second batch compounds in Figure 159.

IC_{50} of each compound is given in table 9:

IC50 (μM)	MCF7
0	638.2
108	214.5
113	10.6
120	25.8
136	354.6
137	77.9
138	76.7
139	315.5
140	419.2

Table 9: MTT IC_{50} of second batch compounds in Figure 159.

The data in table 9 was plotted into a bar chart for convenience, see Figure. 161:

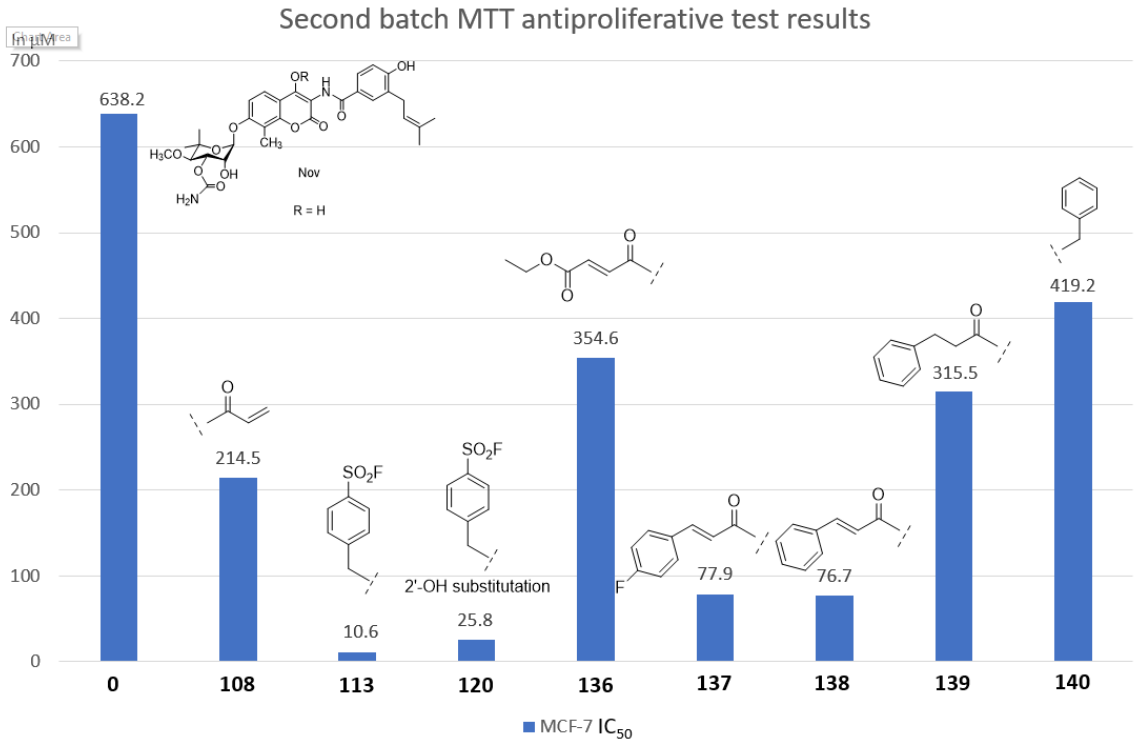


Figure 161: IC_{50} of second batch MTT assay.

In this batch of results, **0**, **108**, **113,** and **120** were previously tested for 24 hours MTT assay against MCF-7. Under 48 hours of drug-exposure, their drug

potency increased 2.09, 2.50, 10.56, and 26.7-fold, respectively. In general, all modifications have increased lipophilicity; as a result, they all presented an increased antiproliferative activity, which was in accordance with our previous conclusion.

108 was still the most potent drug candidate in 48 hours of MTT assay. By comparing **108** with its functional covalent warhead removed form **140,** obviously, adding a sulfone fluoride structure increased the antiproliferative by 39.5-fold. Benzene substituent alone increased the activity with 1.52-fold only (**0** 638.2 μM vs **120** 419.2 μM). We again verified if the 4'-substitution of sulfone fluoride **113** was better than its 2'-OH tautomeric substituted isomer **120**. In conclusion, our result confirmed the promising application of sulfone fluoride warhead. Nevertheless, such modification of novobiocin may change its original binding pocket, of which in native page gel test showed no inhibition of Hsp90 CTD. According to different references (Grimster, Connelly et al. 2013, Hett, Xu et al. 2015), it not likely the case that sulfone fluoride only acts as a traditional non-covalent inhibitor. More effort is needed to define its binding towards Hsp90 CTD or NTD or even other proteins.

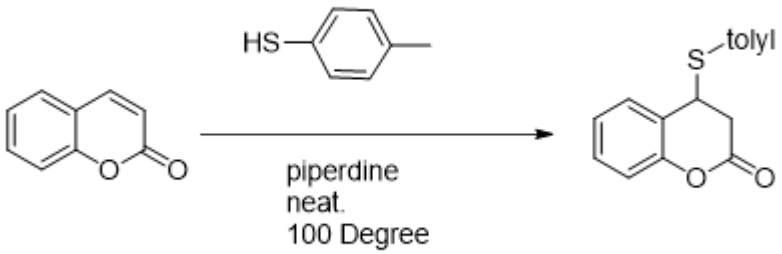
Another very promising outcome in this trial was the varied activity between cysteine reactive 4'-cinnamic-novobiocin **138** and cysteine unreactive 4'-hydrocinnamic-novobiocin **139**. With a single modification of changing one single bond to a double bond, this result demonstrated the importance of electrophilic activity which is identical to what has been reported in literatures (Jackson, Widen et al. 2017). This variation also showed the success of GSH reduction resistance, of which **108** and **109** can not overcome. Moreover, as shown in Figure. 161, the acrylate **108** scaffold was a better modification then the cinnamic type (**108** IC_{50} = 214.5 μM vs **139** IC_{50} = 315.5 μM). This non-covalent growth-inhibitory reducing effect of cinnamic ester moiety furtherly enhanced the necessity of cinnamic α, β unsaturated acrylate moiety, as the electrophilic characteristic of **138** even brought the worse scaffold **139** to have higher activity then **108**. 4-fluoride cinnamic ester **137** presented the same activity as **138**, suggesting that fluoride substitution was non-essential to increase the anti-proliferative activity of 4'-cinnamic novobiocin analogues. On the contrary of being designed for better anti-proliferative activity, α, β double

activated analogue **136** had 1.65-fold reduced activity when referring to **108**. As a result, such modification is not favoured by Hsp90 CTD inhibition.

In conclusion, the MTT assays and drug-Hsp90 native page gel binding assays altogether gave strong evidence that covalent modification of Hsp90 CTD had happened. Further investigation of the binding site will be measured by the mass spec method, which will be discussed in later section (p278).

3.3.12 Hsp90 CTD binding assay results of coumarins

Coumarin contains a cyclic acrylate moiety and is a theoretical Michael acceptor. However, it does not react with thiols under either room or body temperature, unless a higher temperature is applied (Scheme 50):



Scheme 50: Coumarin reacts as a Michael acceptor with thiol under 100°C (Mustafa, Kamel et al. 1956).

Since there was no report of any identified coumarin derivatives except novobiocin (Marcu, Chadli et al. 2000), which act as an Hsp90 CTD inhibitor, we herein screened coumarin and coumarin derivatives to determine any potential coumarin-type Hsp90 CTD inhibitors through native page gel assay.

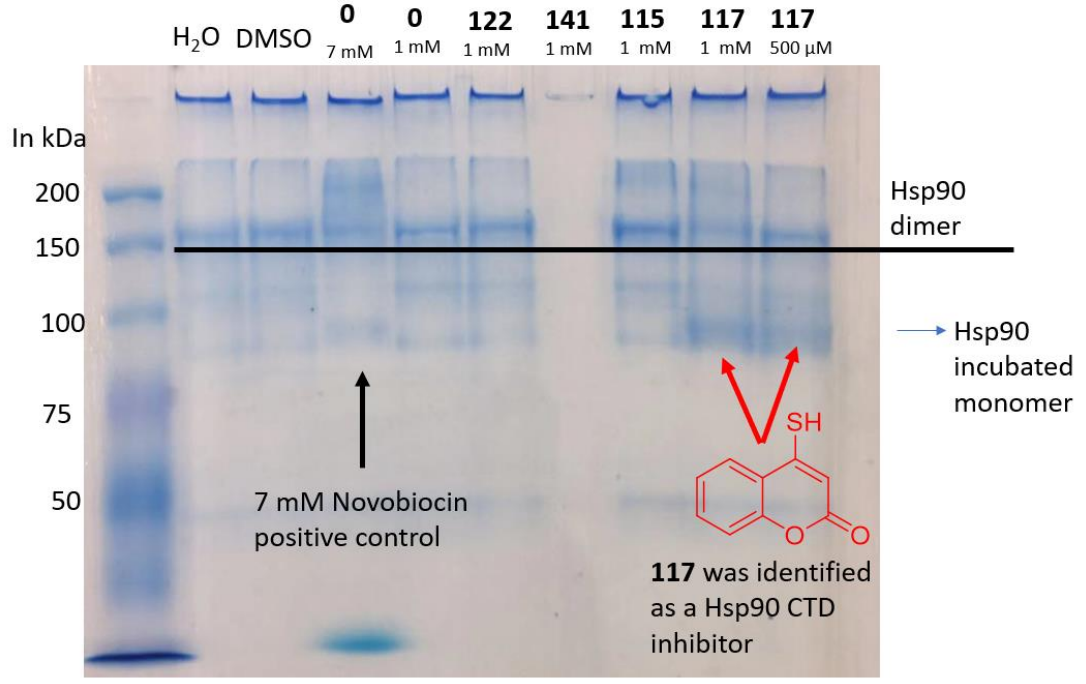


Figure 162: Hsp90 CTD native page gel binding assay of coumarins.

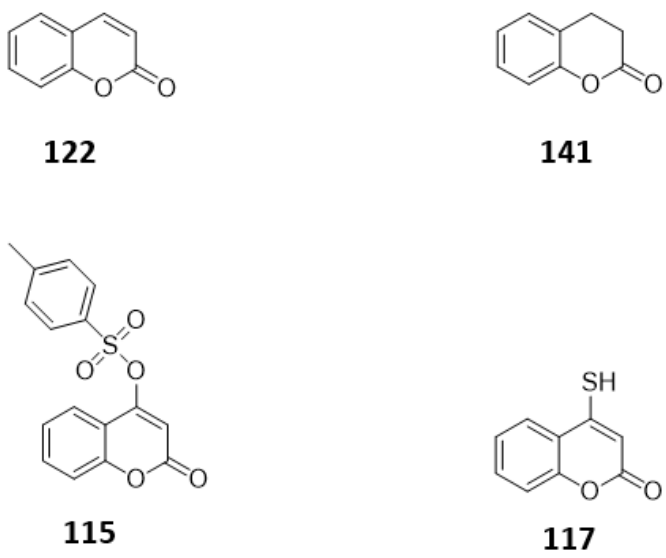


Figure 163: Structure of tested coumarin derivatives.

7 mM novobiocin was selected as a positive control. A new monomer band was formed after 24 hours incubation process (Figure 162).

We included 3,4-dihydrocoumarin **141** (Figure 163) to compare with active form coumarin **122** for potential covalent binding manifestation. The electrophilic double bond of 3,4-dihydrocoumarin was hydrolyzed and became inactive to cysteine thiol Michael addition. However, coumarin **122** failed to show any Hsp90 CTD inhibition, and its inactive form **141** was aggregated (Figure 162).

We found that coumarin analogue **115** and **117** share similar IC_{50} values of anti-proliferative activities (both compounds have an IC_{50} about 500 μM). Despite that **115** had no Hsp90 CTD inhibitory effect, thiol modified analogue **117** was confirmed to be a new Hsp90 CTD inhibitor, showing a CTD binding effect at 1 mM and 0.5 mM. Although whether the thiol group may potentially bind to any Hsp90 CTD residue remained unknown, there were references that reported thiol as a covalent drug (or prodrug, e.g., clopidogrel) for other proteins (Klingler, Wichelhaus et al. 2015).

3.3.13 Microscale thermophoresis (MST) test.

Our data obtained from drug-Hsp90 native page gel electrophoresis experiments suggested that a covalent binding of **108** against Hsp90 may exist due to **108** visually showed a much stronger inhibition than its cysteine-unreactive analogue **109**. In order to quantify the binding affinity of **108** and **109** against Hsp90 full length protein for further study of covalent binding, we applied the microscale thermophoresis (MST) assay to determine this Kd value, due to MST assay uses native condition same as native page gel electrophoresis.

Microscale thermophoresis (MST) is a technique used in biology and biochemistry to quantify biomolecular interactions (binding affinity Kd) (Wienken, Baaske et al. 2010). The measurement is based on thermophoresis, the directed movement of molecules in a temperature gradient, which strongly depends on a range of molecular parameters, e.g., size, charge, hydration shell, or conformation. Thus, this technique is highly sensitive to virtually any change in molecular properties, results in accurate quantification of molecular events independent of the size or nature of the investigated specimen. The measurement is carried out under native condition, which is suitable for Hsp90 in this case as its active form requires a native dimeric form. MST has a wide range of applications, including interactions of almost all types of biomolecules, including DNA, RNA, peptides, protein, small molecules, fragments and ions (Asmari, Ratih et al. 2018). Compared with other binding affinity determent assay, MST only requires a nanomolar concentration of the sample to perform the test with high-quality results. The solution buffer for biomolecules also ranges from organic solvents, water to serum, cell lysates, which allows excellent adaptability and flexibility (Seidel, Dijkman et al. 2013). This technique is a fluorescence-based technique in which the sample biomolecules are labelled with specifical fluorescence dye with proper fluorescent absorbance range. For drug-protein interaction, only the target protein should be labelled with fluorescence dye. Then the labelled protein is incubated with ligands with gradient concentration in vials before they are sampled into specialized capillaries for the measurement (Wienken, Baaske et al. 2010).

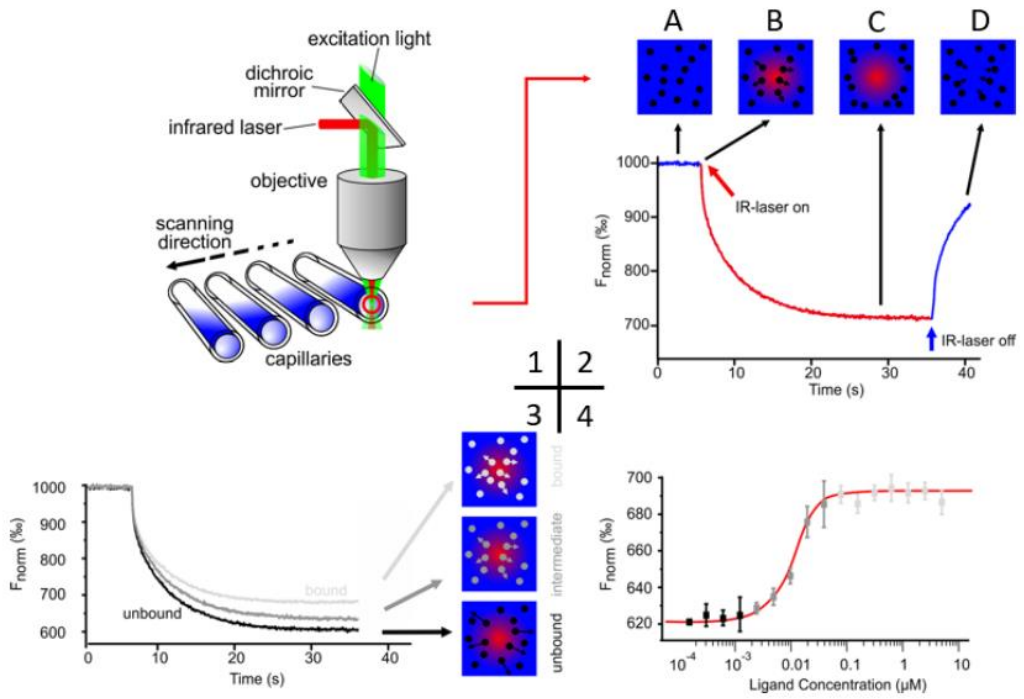


Figure 164: Demonstration of MST technique.

Figure 164 shows a breakdown of one MST test process. For a single run of the MST binding affinity test, the official inventor of MST, Nanotemper, set a standard of 16 gradient samples (50% dilution from highest to lowest concentration) for each run. The initial fluorescence count for the first capillary is measured for 2 seconds as an initial background (Figure 162, 2A).

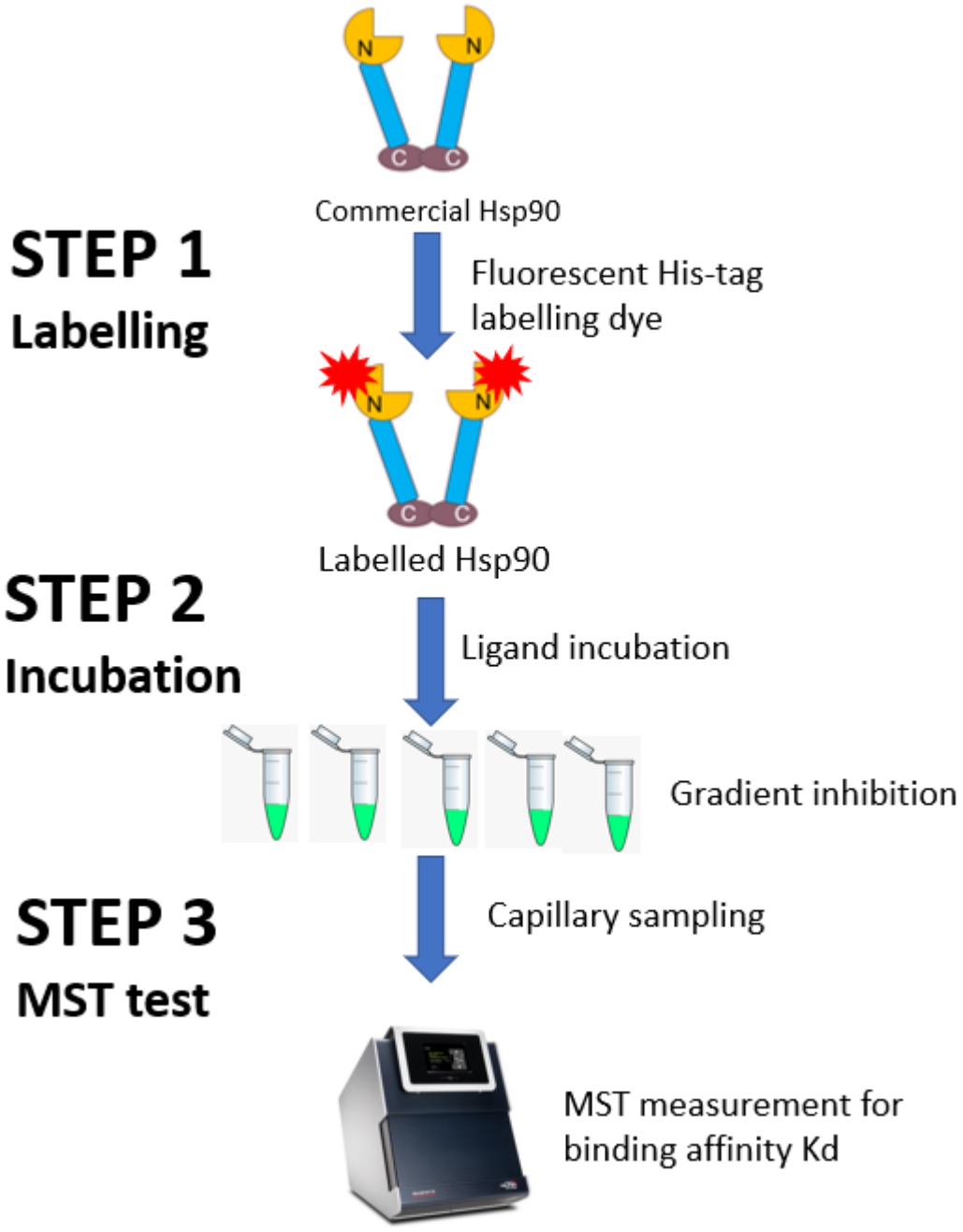
Then, the infrared laser will heat the liquid inside the first capillary (Figure 164, 1). The heat will cause the molecule to move away from the fluorescent detective centre, which will result in a reduction of fluorescent count (Figure 164, 2B – 2C). When this process reaches its maximum or equivalence, the fluorescent count stops to drop, and the absorbance curve moves horizontally, marking the completion of one MST trial (Figure 164, 2C). This equivalence is called a steady state. The IR-laser will be turned off, and molecules move back to the UV detective centre, increasing the UV reading again (Figure 164, 2D, this part is not counted into the final data analysis).

Such a process was repeated from capillary 1 (highest ligand concentration) to capillary 16 (lowest concentration), and 16 curves were generated on the screen. If the drug/ligand molecule binds to the labelled protein, it will change

its biophysical property, which results in more or less labelled biomolecules at the fluorescent detective centre during steady-state and cause the varied fluorescent count. This will be reflected in the fluorescent absorbance curves as the generation of intervals between each MST fluorescent absorbance curve. There will be either a move up or move down for bond sample and unbound sample (Figure. 164, 3).

To calculate the binding affinity, firstly, a time point where the UV curves start to divide must be selected, usually time between 10 s and 40 s. Then the inbuilt software will plot values of fluorescence as Y-axis versus ligand concentration as X-axis, fitted with an appropriate mathematical mode to produce a dose-response curve. Binding affinity Kd (or EC_{50}) can be calculated from the curve (Figure 164, 4).

In this project, we established a protocol for measuring drug-Hsp 90 full protein binding affinity tests using MST. Since novobiocin was discovered as an Hsp90 CTD inhibitor, its binding affinity towards Hsp90 has never been reported. An overall workflow of MST assay is given below (Scheme 51):



Scheme 51: Drug-Hsp90 binding MST test workflow.

3.3.13.1 Hsp90 full protein fluorescent labelling pretests.

Full length Hsp90 human alpha protein purchased from Abcam (ab48801) 1 mg/L was used in this test. The fluorescent labelling dye was purchased from Nanotemper with monolith His-tag labelling kit RED-tris-NTA 2ND generation, specifically labels the His-tag of recombinant Hsp90 located at NTD, which will not affect any CTD function. His-tags are common protein tags that are routinely used for affinity purification. As a result, this His-tag labelling strategy is highly specific, requires only nM concentration of His-tagged Hsp90 protein and no need for dye-removal.

A pre-test process was needed before a labelled protein could be incubated with any ligands. By sending pure labelled sample alone to the MST test, the pre-test resulted in a curve that will determine whether the protein was labelled correctly. Sometimes, due to different chemical and physical properties of a protein, the labelling dye will show adequate or inadequate binding affinity towards the protein's His-tag. We started with the standard process of labelling, in which 100 nM Hsp90 protein solution was labelled with 100 nM labelling dye (1:1 eqv.), incubated at dark for 30 mins under room temperature. After this period, the sample was centrifuged for 10 min at 4 °C at 15000 g. The sample was then sent for the pretest by Guoxuan Sun from Hilton group.

There was no ligand added in the pre-test. Two capillaries of labelled protein samples were run by MST. The pre-test result (Figure. 165) showed that, obviously, the actual application of this standard method on Hsp90 was not successful.

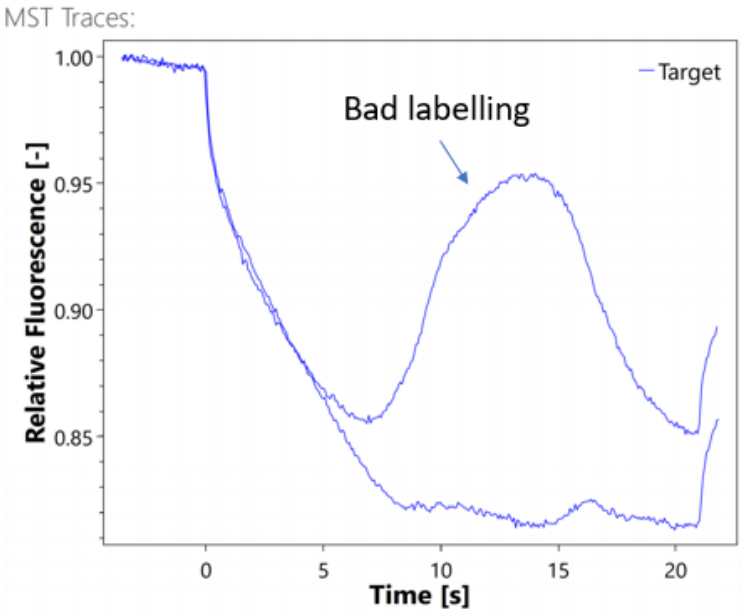


Figure 165: Pretest MST trace of labelled Hsp90 within 30 mins. The huge peak shown in figure is a sign of bad labelling.

In this case, the labelling was not straight forward, and a better labelling procedure should be developed.

Although the user manual of MST machine only described the cause of the hill shape of MST curve as 'protein has a bad binding affinity towards labelling dye or aggregation,' it took a considerable time to notice that 30 min was far less adequate for Hsp90 protein to be adequately labelled. This was due to the standard procedure does not take into account of the situation of such bulky proteins like Hsp90, especially its native dimeric form even reaching 180 kDa mass. The proposed reason why the protein did not reach a steady-state during MST after 5 s was that the protein did not entirely consume the labelling dye, resulted in the hill shape peak between 10-15s. The hill peak could be a sign of unbound dye remaining in the centre of the capillary. The proteins were far bigger than the small dye molecules, at first, the fluorescent count of free dye was masked by proteins (Figure 166, A), when proteins were swimming away from the fluorescent centre, .free dye was revealed and caused abnormal absorbance (Figure 166, B).

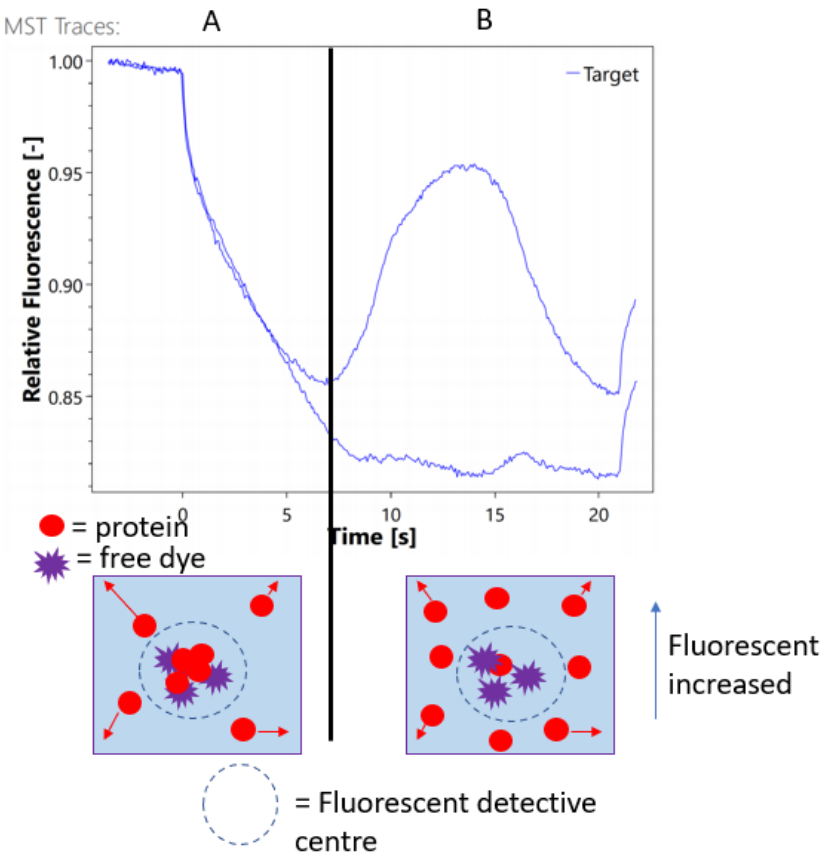


Figure 166: Troubleshoot of the pretest. Illustration of the formation of abnormal hill-shape fluorescent peak in MST pretest curve.

To test the above hypothesis, another batch of labelled Hsp90 protein was applied with 4 hours incubation and tested again (Figure 167). Although the curve was still not perfect as there was a little rebound from 15 – 20 secs, the status of labelled protein was significantly improved without any abnormal hill shape. Under this state, it is robust enough for ligand binding.

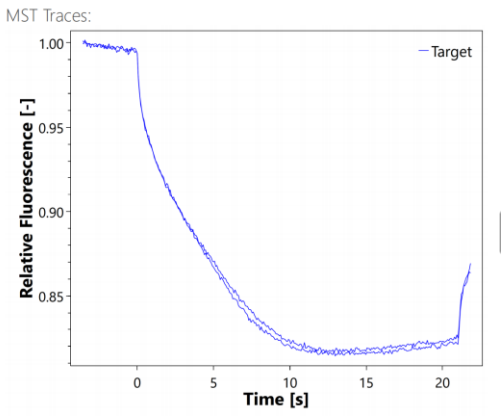


Figure 167: MST pretest result of labelled Hsp90 with 4 hours incubation.

After another series of experiments, it was found that for commercial Hsp90, the best protocol should be to incubate the sample in the dark overnight and then store it at 4 °C in a fridge for 3 – 4 days (Figure. 168). The reason for storing the sample in the fridge was that the enormous geometric protein would consume most of the dye very quickly; nevertheless, the minority unconsumed dye which left in the solution can only be absorbed slowly due to steric hindrance effect of the labelled protein molecule. Using this method described above, we developed, a labelling method that led to better pretest results (Figure 168).

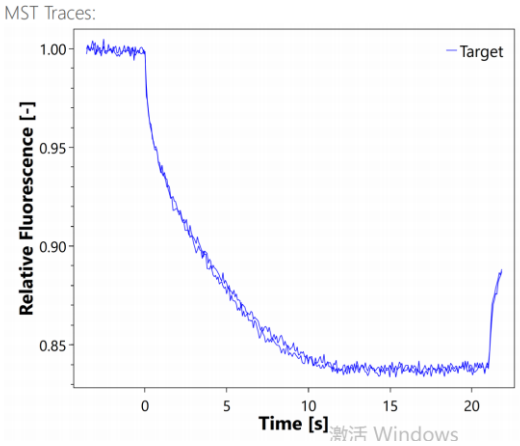


Figure 168: MST pretest result of labelled Hsp90 with 3 -4 days. The rebound issue found in Figure 167 no longer exist, suggesting that Hsp90 full length protein needs a 3 – 4 day incubation with the dye for the best labelling result.

The curve displayed in Figure 168 was a more reasonable labelling result than the sample which incubated for only 4 hours (Figure 167). Our test results showed that for protein with high molecular weight, the protocol from the user manuscript (provided by the MST machine manufacturer) should be modified for better results.

3.3.13.2 Ligand-Hsp90 binding check.

Since we have successfully labelled the Hsp90 protein for the MST test, the next step was to launch a binding check between drug ligand and labelled Hsp90 protein. There were about 40% of the biomolecular interaction process,

range from protein-protein, DNA-protein, ligand-protein interaction, which were not applicable for MST technique to manifest its effect (Jerabek-Willemsen, André et al. 2014). Binding check is the MST test of the ligand-protein incubated sample under a fixed concentration of ligand, to verify if there is a shift of MST trace curve compared with pure protein MST curve. If the incubated sample shows an identical curve as pure protein sample, then it indicates that the binding process cannot be monitored by MST technique.

The novobiocin analogue we tested here for validation of MST binding was compound **108.** The solution of the ligand **108** was made into 700 μM concentration.

40 nM labelled Hsp90 solution was mixed with 700 μM **108** water solution to make a standard 20 nM protein-ligand solution (incubated with 350 μM **108**). Both pure protein and protein-ligand mix was dipped for 4 capillaries and run 4 times. While the user manuscript of MST suggests testing immediately after all capillaries are prepared, it gave unreasonable results in this case (Figure 169).

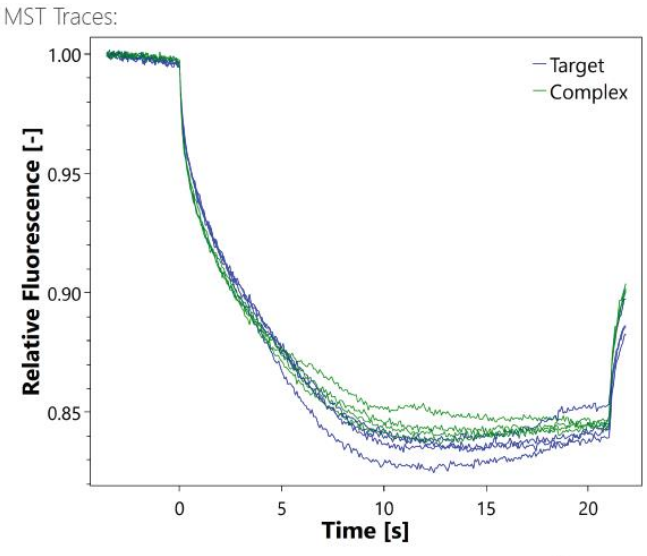


Figure 169: Immediate MST binding check of Hsp90 vs **108**.

The result in Figure 169 showed that the protein-drug complex curve (green) did not confirm binding, as it almost overlapped with the curve of pure protein

(blue). We then increased the inhibition duration to 4 hours, and it showed some degree of inhibition (Figure. 170):

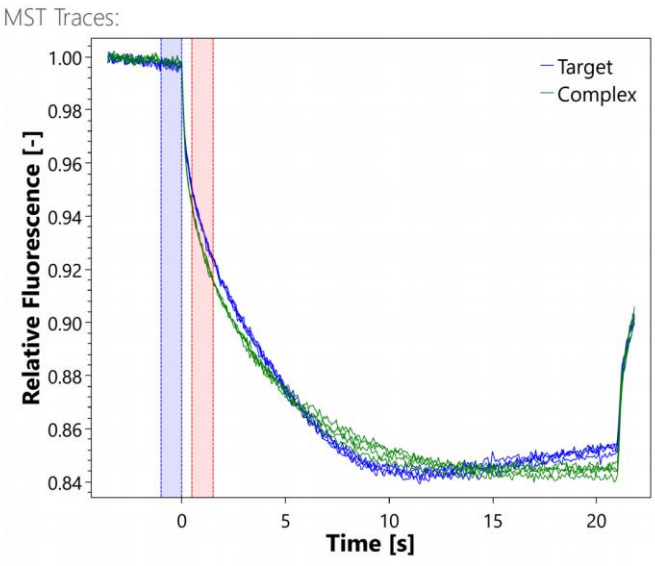


Figure 170: 4-hours MST binding check of Hsp90 vs **108.**

As mentioned before, the MST analytical software 'Nano' will automatically pick the time when two curves were most separated from each other, as shown by the marked in red region in Figure. 170. The interval is called response amplitude; it needs to be large enough to confirm binding. In this trial, the response amplitude was sampled at 1.5 s calculated to be 8.3 (this value has no units) by software, which confirmed the binding (Figure 170).

In accordance with our MTT and native page gel test, we performed another two binding check involving **108** and **109**, incubated with Hsp90 under 37 °C for 24 hours. The concentration of ligand was maintained at 350 μM as before.

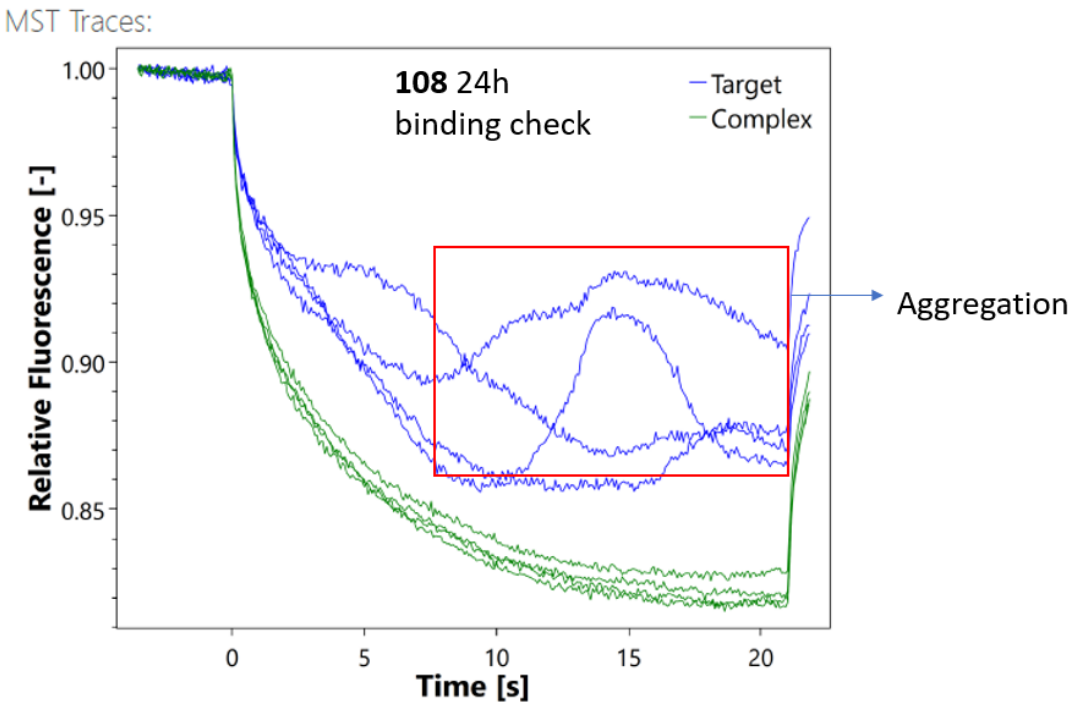


Figure 171: 24-hours MST binding check of Hsp90 vs **108**.

After 24 hours of incubation (Figure. 171), while the ligand-protein complex showed a reliable and robust pattern (green), the pure protein sample showed the problem of aggregation (blue). As a result, the addition of ligand stabilized the protein and prevented the aggregation for some reason. Despite the failure of providing a reliable MST curve of pure protein, it still can be concluded to that increasing the time of incubation would increase the response amplitude.

24 hours MST binding check result for **109**:

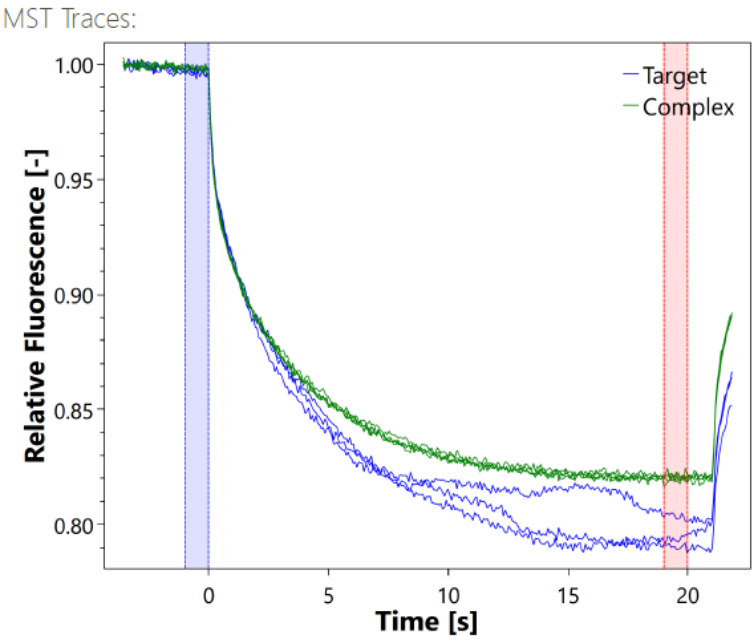


Figure 172: 24-hours MST binding check of Hsp90 vs **109**.

The result of **109**-Hsp90 complex presented very robust MST traces, as 4 curves (green) were almost identical (Figure 172). Pure Hsp90 MST traces (blue) seemed to be less robust, and one capillary showed aggregation. Two binding checks all indicated that adding ligands of novobiocin analogues would also add stability to the protein, or, the stability of Hsp90 monomer (generated by drug inhibition) was much higher than Hsp90 dimers (the native original purchased form). The calculated response amplitude was 23.4, which was about 3 times larger than 4-hours' at 8.3.

Our binding check concluded that the binding affinity of ligand-Hsp90 could be measured through MST technique. However, the MST data may have reduced quality when the ligand concentration was low, because pure Hsp90 samples were less stable according to Figure 171 and 172.

3.3.13.3 Determination of Novobiocin-Hsp90 binding affinity.

Novobiocin **0** was the first reported Hsp90 CTD inhibitor (Marcu, Chadli et al. 2000); however, its binding affinity towards Hsp90 full protein remained unknown. To determine the binding affinity Kd value of novobiocin, we prepared a gradient dilution of novobiocin and incubated with 20 nM Hsp90 full-length protein for an aimed duration and then measured by MST.

A total of 16 dilution samples were prepared. The dilution started with adding 20 μL highest concentration of ligand solution towards the test vial, then 10 μL of which was removed to the second vial. 10 μL of PBS buffer was added into the second vial to dilute the concentration into half. After adequate stirring, 10 μL of liquid from the second vial was transferred into the third vial. This dilution process was repeated until the 16^{th} vial was prepared.

The first difficulty of method development was to define the maximum concentration used in this test. Usually, a good set of dilution should fix the estimated IC_{50} concentration in the middle range of dilution. For example, if novobiocin's binding affinity was estimated at around 700 μM, to put the 700 μM curve in the middle, the first vial should have a concentration of 89.6 mM, which is high. Nevertheless, we decided to start with 25 mM concentration as the highest concentration sample to have a rough estimation of binding affinity. We firstly performed a 4 hours' test using samples prepared under room temperature to investigate whether the inhibition process was complete (Figure.173).

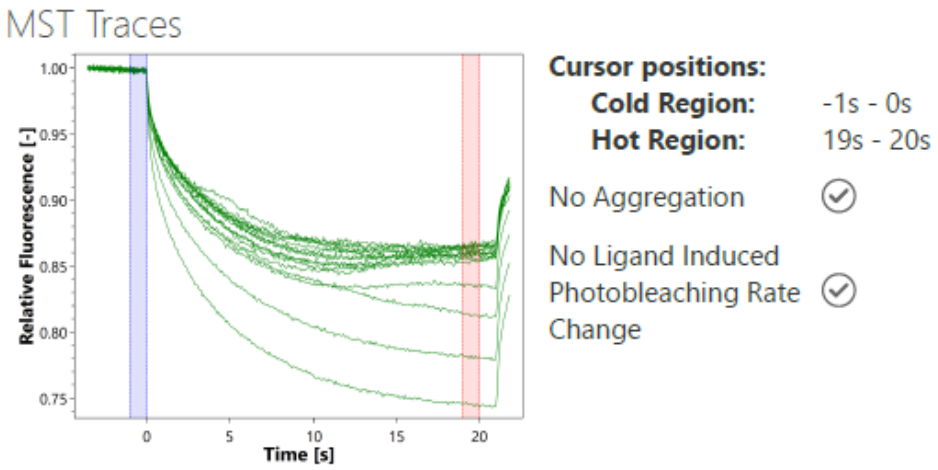


Figure 173: MST trace of Max. 25 mM novobiocin **0** binding affinity test, sample was incubated for 4 hours.

The MST trace curves showed good quality as the trend was apparent (Figure. 173), the higher the ligand concentration, the lower the fluorescence count.

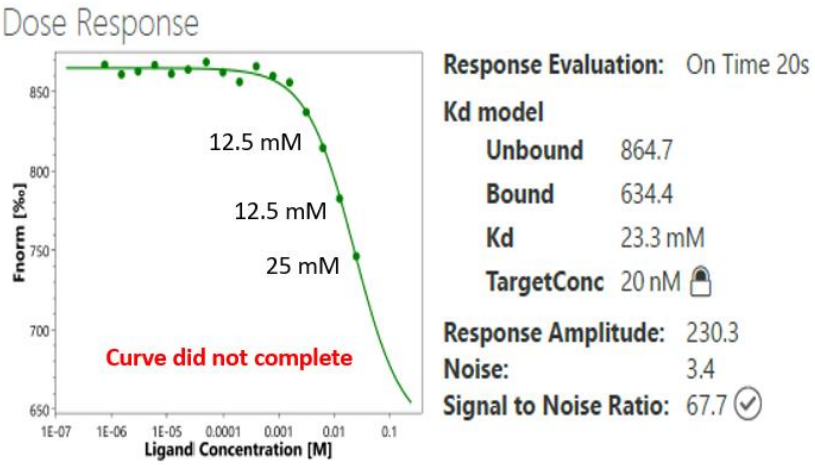


Figure 174: MST dose-response curve of novobiocin **0** Hsp90 binding affinity test, 4 hours' incubation. Binding affinity was calculated as 23.3 mM.

Apparently, 23.3 mM with 4 hours' inhibition did not reach the maximum inhibition 'twist point'. As a result, the software was unable to know whether higher concentration would significantly increase/decrease the fluorescence count or not, and the calculated Kd of 23.3 mM was not a valid value (Figure. 174).

To optimize the experiment, a higher concentration of 100 mM ligand was prepared and incubated with two different conditions (Figure. 175 left = 4 h rt. Right = 24 h 37 °C).

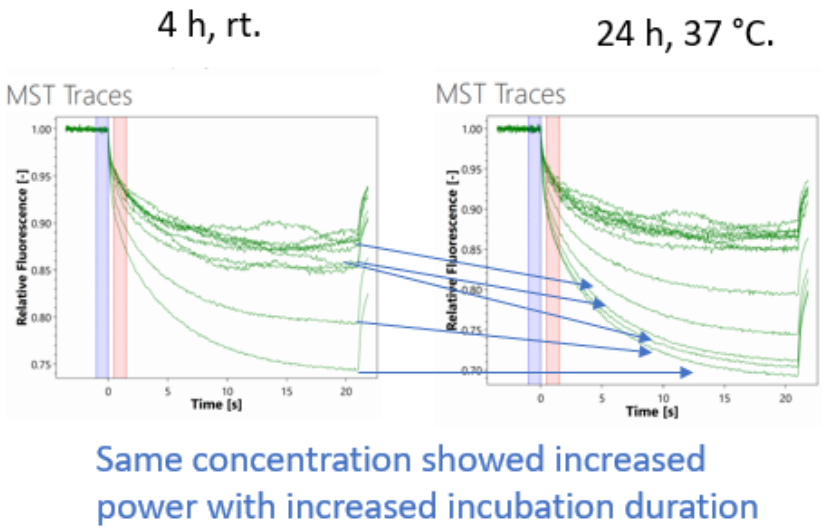


Figure 175: MST trace of Max. 100 mM novobiocin **0** binding affinity test, with incubation duration left = 4 hours room temperature, right = 24 hours 37 °C. Clearly, 24 hours incubation under 37 °C was a optimal condition.

It was clear that novobiocin 100 mM, 24 h, 37 °C MST trace showed increased endpoint activity (if sampled at 20 s) at high concentrations compared with the same curve under 4 h, rt. The fluorescence count of 100 mM (0.75 in 4 h to 0.70 in 24 h), 50 mM (0.80 in 4 h to 0.73 in 24 h), 25 mM, 12.5 mM and 6.25 mM (these 3 points were >0.85 in 4 h to >0.7 in 24 h) significantly dropped, meaning 24 h at, 37 °C condition was better for the inhibition to complete. This suggested that the non-covalent inhibition of novobiocin was slow under room temperature, and 24 h at 37 °C was enough for novobiocin to reach maximum inhibition. Therefore, the Kd calculated from the sample incubated using the condition of 24 h at 37 °C would be convincing.

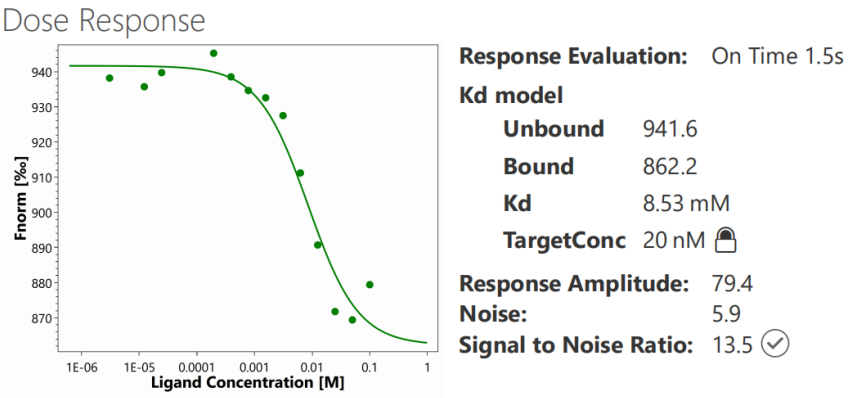


Figure 176: MST dose-response curve of novobiocin **0** Hsp90 binding affinity test, 24 hours, 37 °C. Binding affinity of **0** was calculated as 8.53 mM.

Surprisingly, the calculated Kd for novobiocin was 8.53 mM (Figure. 176), which was far less potent than either previously reported MTT IC_{50} 700 µM (SkBr3 cell line) (Burlison, Avila et al. 2008) in reference or 500 – 1339 µM (Patel, Fuente et al. 2011). The experiment was repeated for several times; however, the results were quite the same; the average value was 8.91 mM.

Considering that the protein sample we used in this trial was pure commercial dimer and 7 mM novobiocin did not show any effect on native page gel against the same commercial Hsp90, the MST defined Kd can be reasonable. There could be either some unknown mechanism under both protein and cellular level for novobiocin to present much higher antiproliferative activities.

3.3.13.4 Determination of 4'-acrylate novobiocin (108)-Hsp90 binding affinity.

Based on the discussion in section **3.13.3**, we attempted to make the highest concentration of compound **108** into 100 mM. Nevertheless, unlike water-soluble novobiocin sodium salt, the maximum concentration of **108** in 5% DMSO water as a solvent was discovered to be 23 mM through dilution tests. Thus, 23 mM of **108** was prepared into the gradient sample solution and incubated with Hsp90 for 24 hours under 37 °C. MST binding affinity results are given below (Figure. 177):

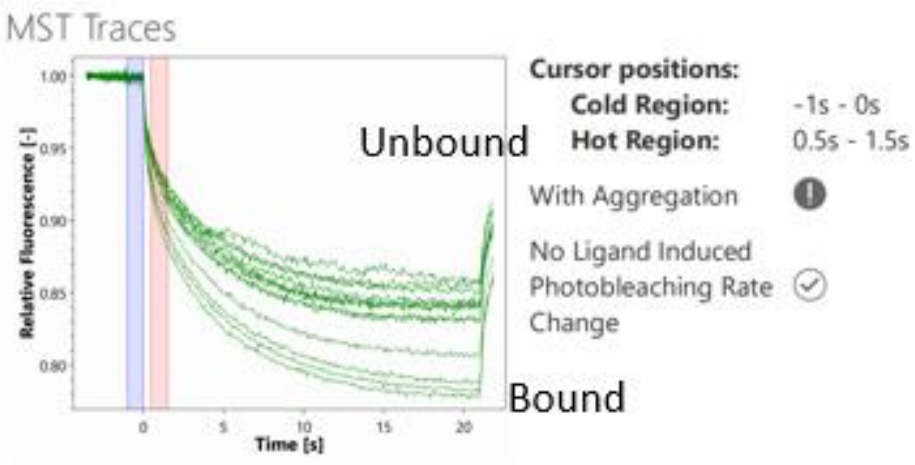


Figure 177: MST trace of Max. 23 mM **108** binding affinity test, samples were incubated for 24 hours at 37 °C.

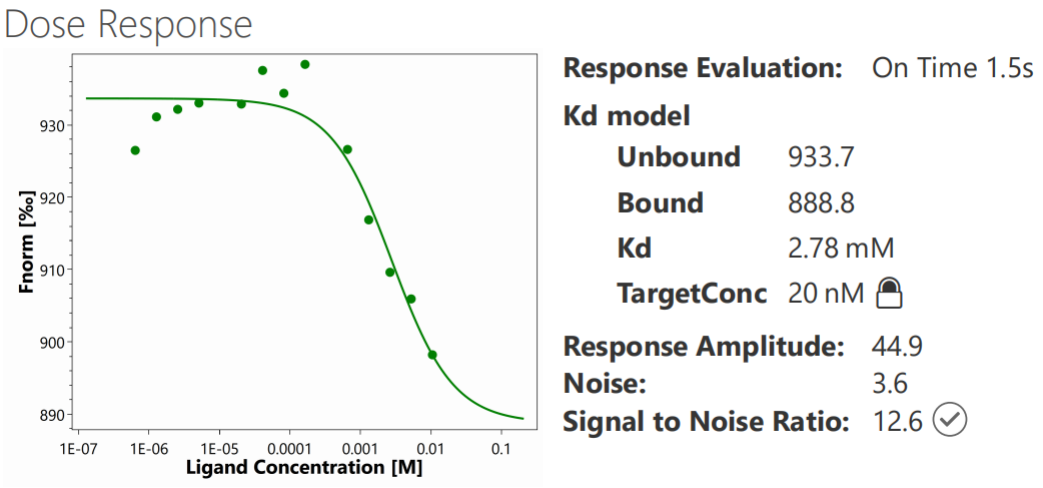


Figure 178: MST dose-response curve of **108** Hsp90 binding affinity test, samples were incubated for 24 hours at 37 °C. Binding affinity of **108** was calculated as 2.78 mM

The sample at the highest concentration showed a pattern of aggregation; thus, the MST trace of the concentration was removed. The result of the MST results was plotted (Figure. 178), and the Kd was calculated as 2.78 mM. The data will be discussed together with MST results of **109** in **3.13.6**

3.3.13.5 Determination of 4'-propionyl novobiocin (109)-Hsp90 binding affinity.

The preparation procedure was the same as the procedure applied for sample preparation of **108**, of which the maximum concentration of ligand was set to 23 mM. Gradient samples were incubated with Hsp90 for 24 hours under 37 °C, and the MST results are shown in Figure. 178:

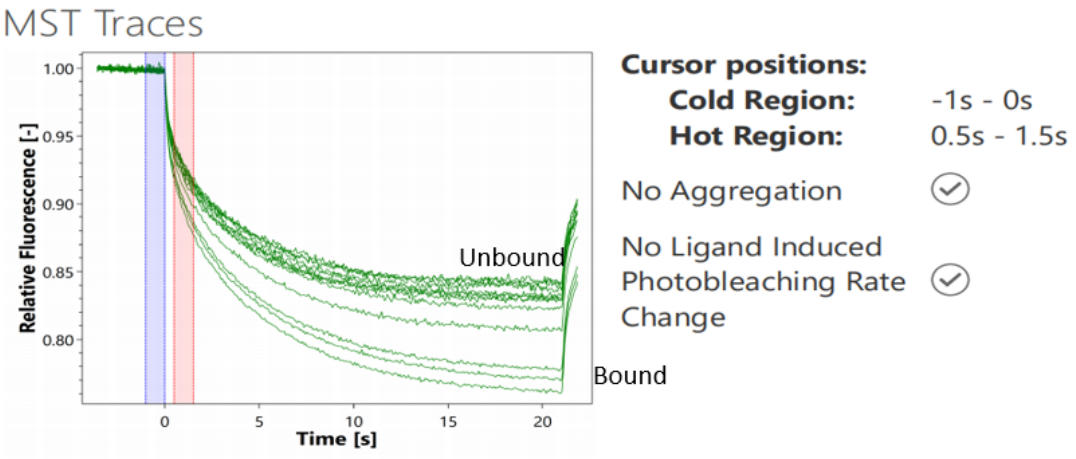


Figure 179: MST trace of Max. 23 mM **109** binding affinity test, samples were incubated for 24 hours at 37 °C.

MST trace of **109** showed curves with high quality and similar to MST traces of **108** in Figure 177, where ligand-bound traces were at the bottom and unbound traces were at the top. The Kd value measured in this test for **109** is 949 µM (Figure. 180).

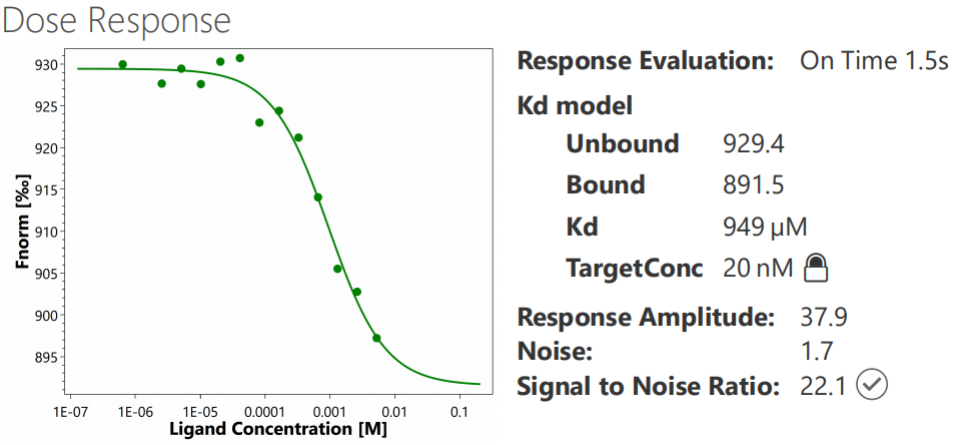


Figure 180: MST dose-response curve of **109** Hsp90 binding affinity test, 24 hours, 37 °C. Binding affinity was calculated as 0.949 mM

3.3.13.6 Discussion of MST test results.

MST assay quantifies the native interaction between drug and target protein; the test results summarized in Figure 181 showed that all 3 ligands were not very potent to inhibit Hsp90 protein.

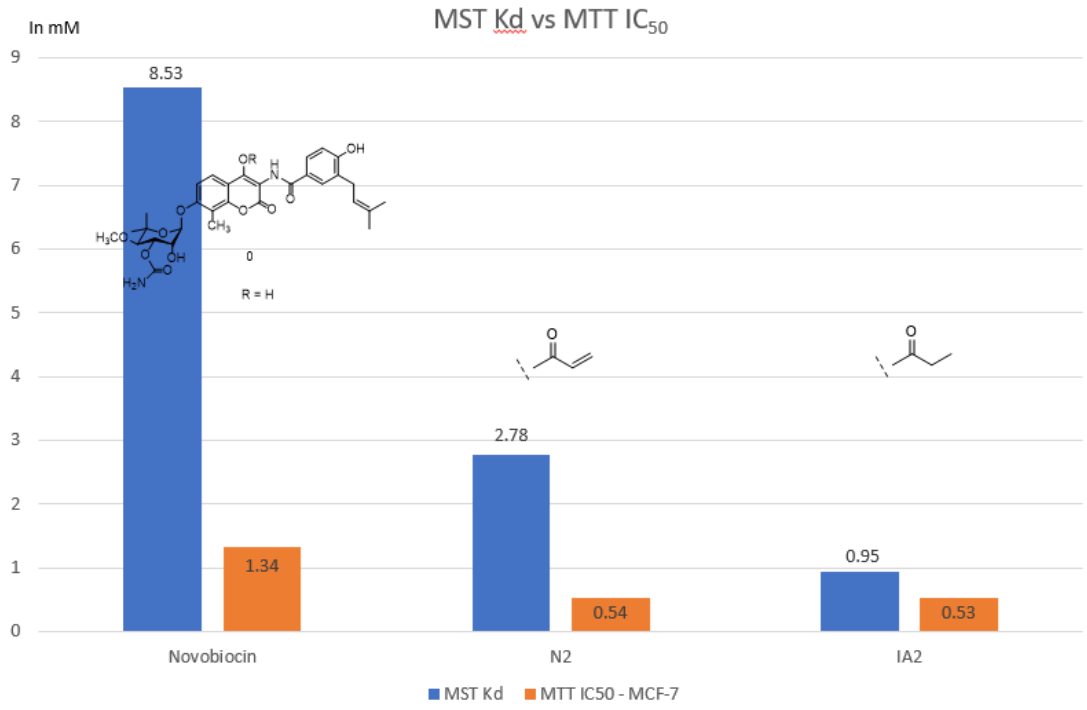


Figure 181: A comparison between MST Kd and MTT IC_{50} of novobiocin and its analogues.

Although all 3 compounds showed millimolar level inhibition towards Hsp90 (Figure. 181), it should be noted that all our research was carried out using Hsp90 full-length protein, which was closer to the condition inside the cells. Usage of full-length dimeric protein would add difficulties for the ligand to access the binding pocket; obviously, the well-preserved quaternary steric structure of Hsp90 would block the attack from ligand molecules thus reduce the chance of effective impact between ligands and residues. Therefore, if only Hsp90 C-terminus residues were applied, the Kd was expected to be higher due to the removal of the steric hindrance from mid & N terminus domains. As a result, our data revealed that the ligand – 'full length Hsp90' binding process

measured by MST was less active than expected (compared with MTT IC_{50} data).

MST assay and native page gel assay both manifested the native process of drug-ligand interaction. While the native page gel assay marked the superior activity of cysteine active form **108** (active under 0.07 mM) compared with its inactive from **109** (inactive under 0.07mM), quantified Kd of **108** by MST showed decreased affinity and lost its covalent superiority towards **109**.

As two native assays shared the same incubation condition, it was not likely the issue discussed for Figure 181 was caused by environmental factors. Also, unlike MTT test where cellular enzymes act as an uncontrollable factor, ligands were not modified by any factors in both native tests which means **108** were proposed to have a lower Kd than **109** according to previous native page conclusions. To discuss the reason why binding affinity significantly varies between MST and native page, we reviewed relevant literature (Lonsdale and Ward 2018) and reanalysed our native page gel graphs, and we found that these two native assays measured different stages of Hsp90 inhibition. This brought out a new possible mechanism of Hsp90 CTD inhibition that has not been reported. Despite that the definition of enzyme inhibitor only described its inhibition towards the target protein, it never mentioned whether the substrate inhibits the active site of its dimeric form or monomeric form. In this case, inhibition against Hsp90 dimer and Hsp90 monomer were different situations that must be distinguished.

According to the literature (Jackson, Widen et al. 2017), the reported classic inhibition mechanism of novobiocin type inhibitors was a reverse process of Hsp90 dimerization. It was believed that novobiocin binds to Hsp90 C-terminal and generates two monomers; thus, the target lost its bio-function (Figure 182).

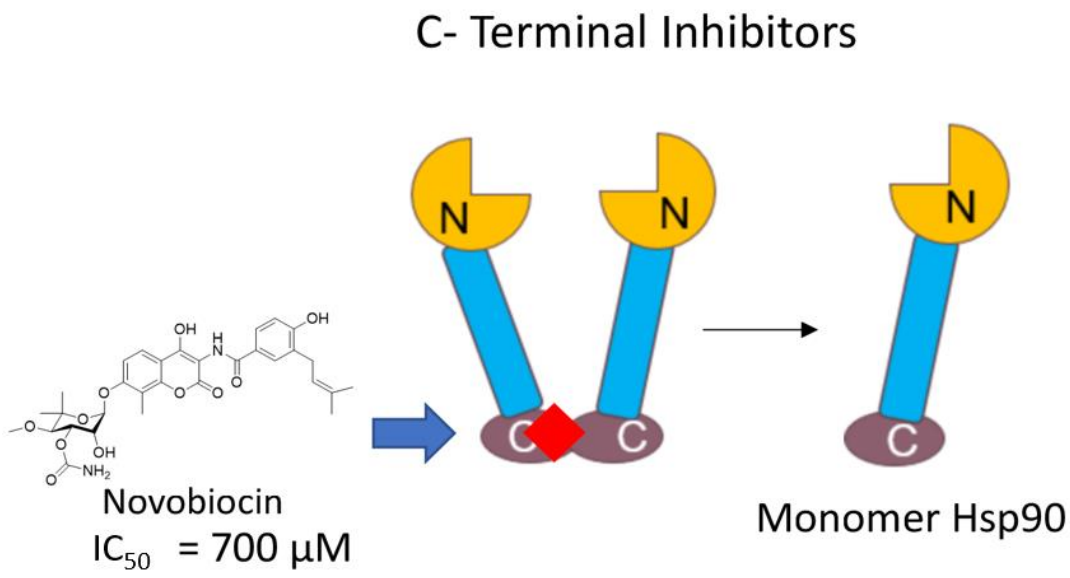


Figure 182: Inhibition of novobiocin like Hsp90 C-terminal inhibitors.

From this model, the binding affinity of Hsp90 CTD inhibitor came from its inhibition towards Hsp90 dimer, while its binding against dissociated Hsp90 monomer was ignored. To find out if there was possible inhibition among Hsp90 monomers rather than dimers, we reviewed our native gel results (Fig. 183).

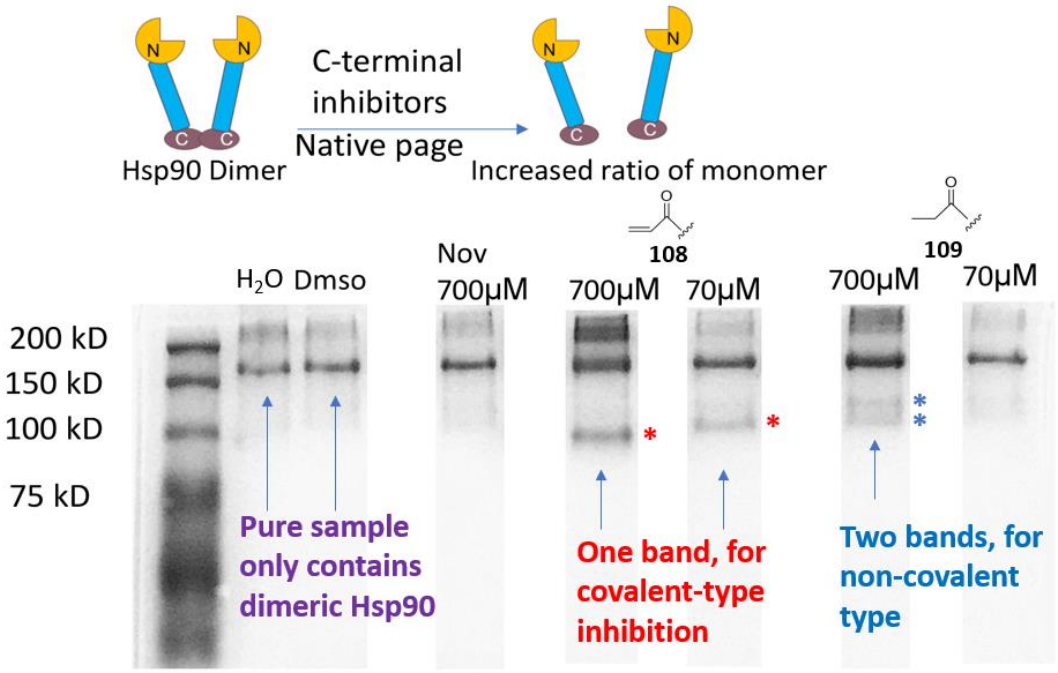


Figure 183: Repost of drug-Hsp90 native page gel binding assay. Commercial Hsp90 protein was in pure dimeric from, no monomeric Hsp90 was found.

Hsp90 sample we used for the gel above (Figure. 183) and for MST tests were commercial Hsp90 ab48801 in which Hsp90 only existed as bio-active dimeric form. Protein bands on native page gels were not necessarily distributed by their molecular weight. Therefore, proteins that shared a same molecular weight may locate at different bands on the native page gel due to their varied physical and bio-electrical properties. According to traditional Hsp90 CTD inhibition mechanism, Hsp90 CTD inhibitors would cause the dimeric Hsp90 to separate into two monomers (if they do not overlap on the native page). Herein, for noncovalent inhibitor **109**, there were two bands instead of one (Figure 183), while **108** separated Hsp90 dimer into new monomer band. This could be due to **108** and **109** would both inhibit dimeric Hsp90 into two monomers, and then **108** covalently binds to Hsp90 monomer, transferring two kinds of Hsp90 monomer to one kind. The best way to validate this theory was to synthesis the Hsp90 monomer only and then carry out the native page gel test. Nevertheless, the company Abcam has no such product due to Hsp90 proteins mostly were in dimeric form; also, there was no method to purify Hsp90 monomer from crude Hsp90 product (FPLC will not show any monomer

peak). An alternative plan was to use our lab-produced sample, which contains both monomer and dimer of Hsp90. Native page using pure dimeric Hsp90 has proved the traditional mechanism of reverse dimerization inhibition pattern. It can be deduced that, if there is a decreasing of monomer band on the native page using Hsp90 monomer-contained sample, Hsp90 CTD inhibitors may inhibit both dimeric and monomeric Hsp90.

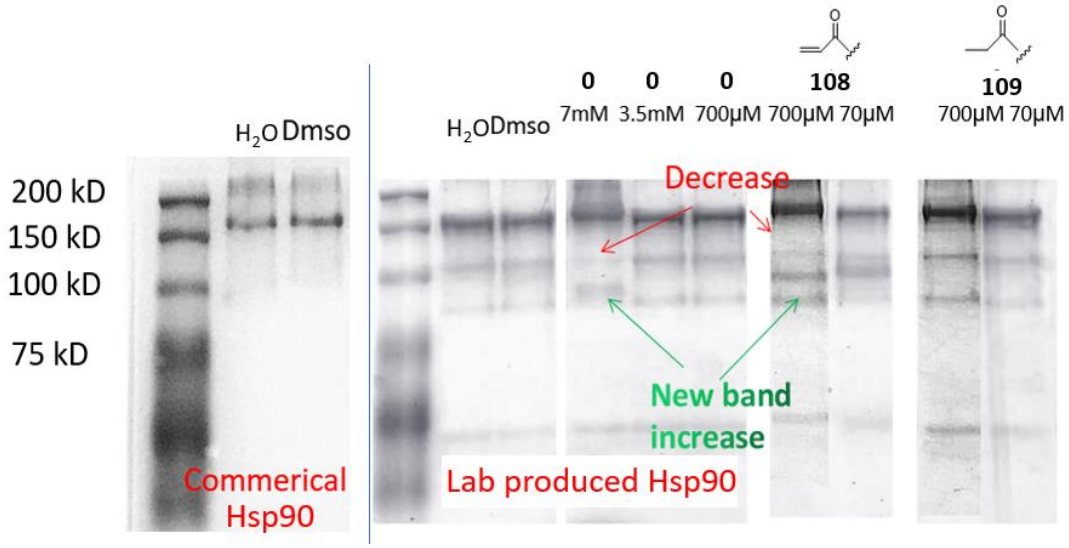


Figure 184: Repost of drug-Hsp90 native page gel binding assay. Commerical Hsp90 contains Hsp90 dimer only (left column). Lab-produced Hsp90 protein was a mixture of Hsp90 dimer and monomer (right column). Novobiocin **0** and **108** can inhibit Hsp90 monomer located at 100 kD and cause the formation of a new Hsp90 monomer band between 75 kD and 100 kD.

In Figure. 184, we can see the difference between the control groups of commercial and lab produced Hsp90. There were two more Hsp90 bands located at 100 kDa and 75 – 100 kDa respectively (no literature validated those two bands, but we asked our collaborator Dr. Kate Hessom from University of Bristol to do the proteomics analysis and confirmed they are Hsp90 protein). This was reasonable as Hsp90 may be assembled from these two monomers to active dimers. In the native page gel prepared from lab produced Hsp90, 7 mM novobiocin and 700 μM entry showed a decreased amount of 100 kDa Hsp90 monomer band, together with the appearance of a new band among 75 – 100 kDa. This a direct sign of Hsp90 monomer inhibition caused by

inhibition of novobiocin **0** and **108**, apart from Hsp90 dimeric inhibition. **109** did not show any visible effect on the same native page in Figure 184. However, due to **109** was confirmed to have Hsp90 dimeric inhibition via MST, Hsp90 monomers generated from inhibition of **109** may overlapped with original 100 kDa and 75 – 100 kDa bands. Generally, the missing of 100 kD monomer and the presence of new monomer band, caused by inhibition of novobiocin **0** and **108,** proved a distinct monomeric inhibition pattern which has never been reported.

More evidence was found in our group's previous research (Cruz, Zhang et al. 2013); nevertheless, the decreasing of monomer band was ignored.

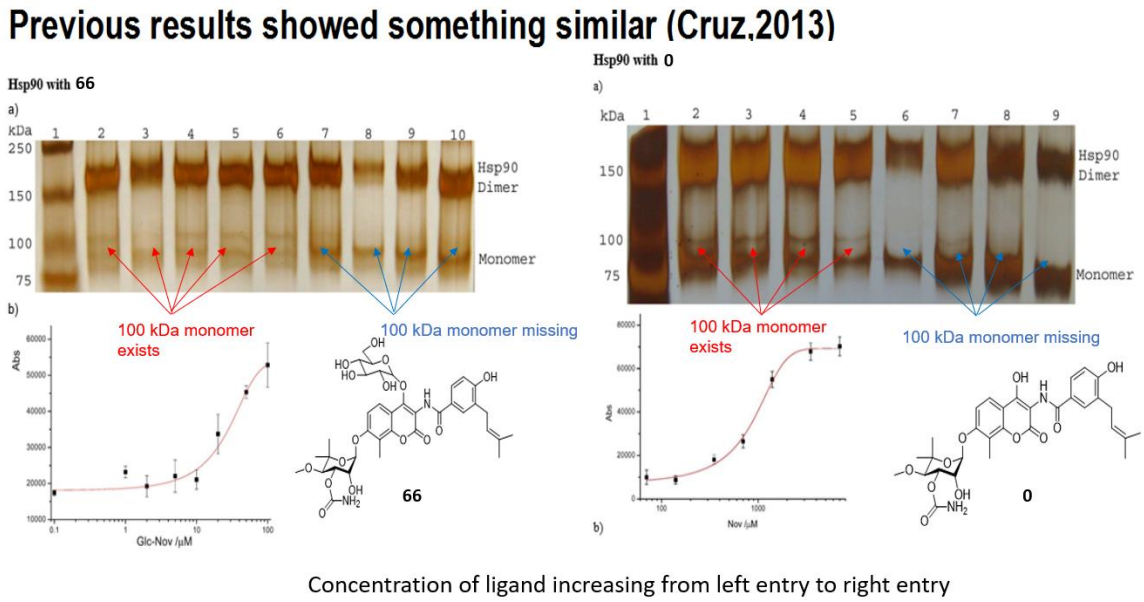
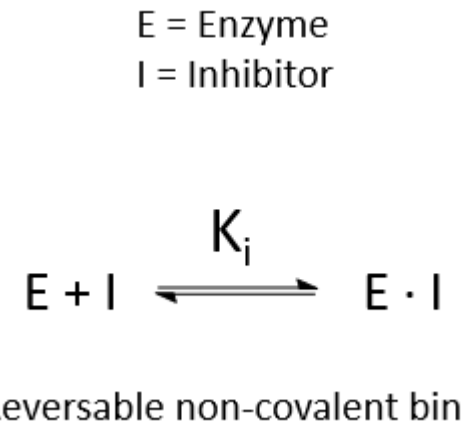


Figure 185: Previously published drug-Hsp90 native gel results (Cruz, Zhang et al. 2013). Glu-novobiocin **66** and novobiocin **0** inhibit Hsp90 monomer. With the increasing of ligand concentration, Hsp90 100 kDa monomer band faded.

With the increasing of ligand concentration, Glu-novobiocin **66** and novobiocin **0** was able to fade the 100 kDa Hsp90 monomer band (Figure. 185), identical to our observation (In Figure.184, novobiocin 7 mM caused the same fade away of 100 kDa Hsp90 band). Our current native page results (Figure. 184) together with previous results (Figure. 185) all confirmed that Hsp90 CTD

inhibitors inhibit not only Hsp90 protein through reverse dimerization (dimeric inhibition), but also monomeric inhibition.

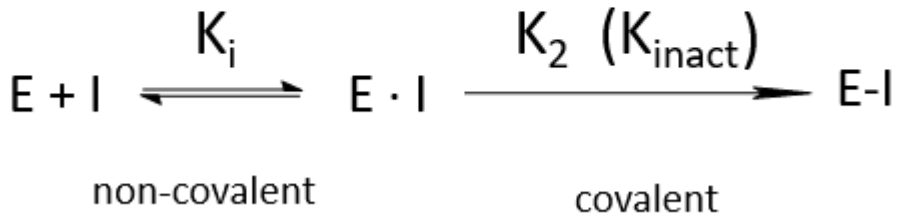
Herein, a new mechanism can be proposed. Traditionally, the binding affinity of reversible non-covalent inhibition was expressed in this function, see equation 1:



Equation 1: Equation of non-covalent binding affinity K_i mode.

K_i is initial binding specificity, defining the non-covalent binding affinity of traditional non-covalent inhibitors. The potency and selectivity of covalent drugs also affected by both their non-covalent affinity (Ki) and the second-order rate constant for covalent binding (K_{inact} or K_2) (Bauer 2015). In other words, to let the irreversible covalent inhibition happens, the non-covalent scaffold of the molecule must have adequate power to deliver the molecule to the binding-active site then covalent moiety will bind to active residue through an addition reaction (see equation 2).

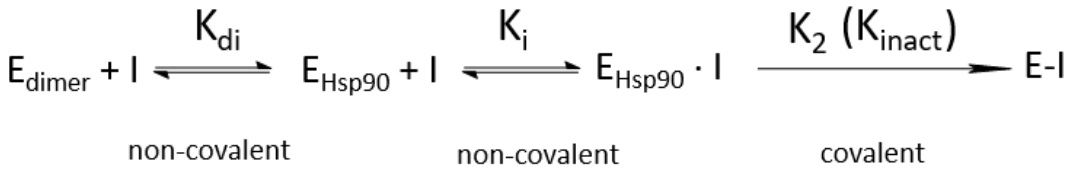
E = Enzyme
I = Inhibitor



Equation 2: Equation of covalent binding affinity K_2 mode

Instead of summarizing all native forms of target protein into E, inhibition of Hsp90 will require a new reverse dimerization stage, in which active dimeric Hsp90 (E_{dimer}) must be inhibited into monomers in a non-covalent pattern and then monomers (E_{Hsp90}) will undergo a traditional non-covalent inhibition (see equation 3).

E_{Hsp90} = Hsp90 monomer
E_{dimer} = Hsp90 180 kDa dimer
I = Inhibitor



Equation 3: Equation of Hsp90 dimeric covalent binding affinity mode. Kdi refers to the binding affinity between ligand and Hsp90 dimer.

Thus, based on the traditional non-covalent inhibition mechanism, the binding affinity between ligands and Hsp90 dimer 'Hsp180' can be defined as K_{di}. K_{di}

quantifies the power of reverse dimerization distinguished from monomeric Hsp90 binding affinity K_i.

In our previous discussion, we have reviewed the same incubation process adopted by both native page gel and MST. From native page gel, we confirmed that the process being quantified by MST was reverse dimerization (because there was no original Hsp90 monomer in the sample). Therefore, MST test results showed that K_{di} of novobiocin **0**, **108** and **109** were at low milligram range potency. Monomeric Hsp90 binding affinity (K_i) of these compounds remained unknown; however, it can be easily deduced that K_i would have a much smaller value than K_{di} because these compounds all have improved potency according to their much smaller MTT assay IC_{50}. The drug potency of Hsp90 CTD compounds relied more on Ki rather than K_{di}; otherwise, the MTT IC_{50} value of Hsp90 CTD inhibitors should also fall into millimole level range instead of micromole level. Inhibition of monomeric Hsp90 would be faster than 'open-up' of the bulky Hsp90 dimer molecule. MST assay would not quantify Ki and K_2 unless the pure monomeric form of Hsp90 was used during the incubation.

Currently, there is no reported literature distinguishing the two stages inhibition of Hsp90 c-terminal (K_{di} and K_i) This creates a large gap for the understanding of the binding mechanism of novobiocin and other Hsp90 CTD inhibitors. Usually, target validation of Hsp90 CTD inhibitors is through immune blots where the protein is denatured into liner form (no difference between monomer and dimer.). The native form of Hsp90 has been ignored for so long. By understanding the two non-covalent inhibition process, it will help not only the design of Hsp90 CTD inhibitor in the future but also benefits the pharmacology study.

To sum up, through a cross review of MST and native page gel results, we managed to discover a new non-covalent reverse dimerization process for Hsp90 full-length protein inhibition. MST quantified the binding affinity, K_{di}, of novobiocin, **108** and **109** with 8.78 mM, 2.78 mM and 0.95 mM respectively.

3.3.14 MALDI-TOF mass spec analysis of 108-Hsp90 full-length protein complex.

Since non-covalent inhibition will not change the overall mass of the target protein, one of the direct ways to validate the covalent binding between cysteine modifier **108** and Hsp90 was to search for protein mass and ligand mass value in mass spec results using suitable mass spectroscopy technique. The suitable mass spectroscopy technique for ionization of 90 kDa mass, suggested by National Mass Spectrometry Facility Swansea (NMSF), was matrix-assisted laser desorption/ionization (MALDI). MALDI is a soft ionization method which involves a laser striking a matrix of small molecules to lift the analyte molecules into the gas phase without fragmenting or decomposing them. This characteristic was beneficial in this case as we would like to keep the initial mass of protein or protein-ligand complex. Some biomolecules are too large and can decompose when heated, and traditional techniques will fragment or destroy macromolecules (Karas and Hillenkamp 1988). MALDI is appropriate to analyze large biomolecules like peptides, lipids, saccharides, or other organic macromolecules, up to 300 kDa range.

MALDI uses a laser energy absorbing matrix to mix with sample and assist the ionization. Standard matrix adopted are 3,5-dimethoxy-4-hydroxycinnamic acid, α-cyano-4-hydroxycinnamic acid (α-CHCA, alpha-cyano or alpha-matrix) and 2,5-dihydroxybenzoic acid (DHB). These matrices are mixed with analytes in an organic/water solvent system, of which the organic solvent is usually acetonitrile or ethanol. Then the matrix-analyte solution is spotted on a tin plate. After solvents vaporize, it will leave only the recrystallized matrix, but now with analyte molecules embedded into MALDI crystals. Prepared sample as such is ready for ionization.

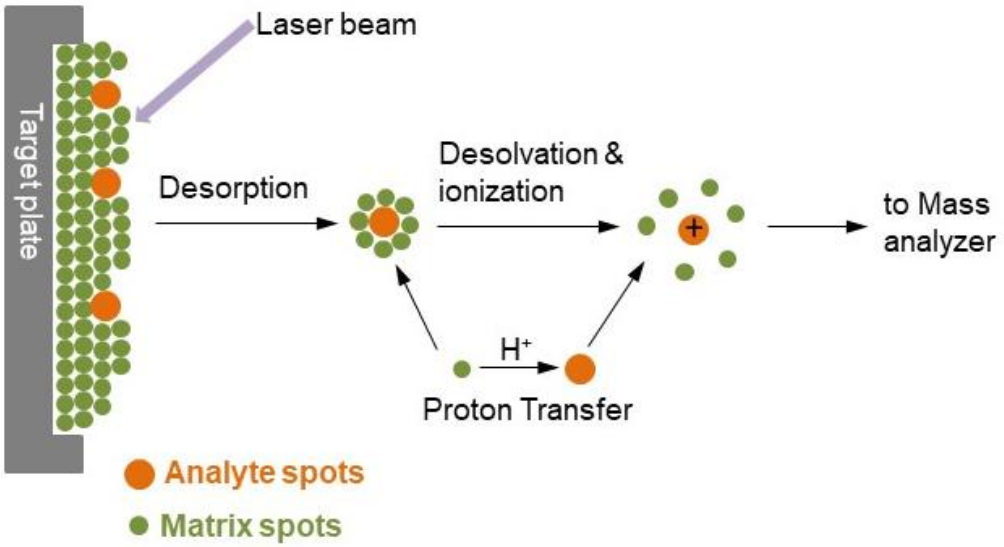


Figure 186: Ionization of analytes by MALDI.

Illustrated in Fig. 184, the analyte is embedded in a considerable excess of a matrix compound deposited on a solid surface called a target, usually made of conducting metal and having spots for several different samples to be applied. After a very brief laser pulse, the irradiated spot is rapidly heated and becomes vibrationally excited. The matrix molecules energetically ablated from the surface of the sample, absorb the laser energy, and carry the analyte molecules into the gas phase as well. During the ablation process, the analyte molecules are usually ionized by being protonated or deprotonated with the nearby matrix molecules. The most common MALDI ionization format is for analyte molecules to carry a single positive charge. The ionized analytes will then be sent to traditional TOF (time of flight) analyzer to determine the m/z value.

3.3.14.1 Results and discussion of MALDI-TOF experiments.

For such analysis, usually, a sample of pure protein and a sample of the drug-protein incubated complex should be prepared. A total of 3 samples were prepared and delivered to NMSF for MALDI-TOF study, and details are given below:

Table 10: MALDI-TOF samples submitted to NMSF.

Sample name	Content	Preparation
GS-0	Pure Hsp90	Hsp90 (1 mg/mL) 3µL + 7 µL pure water
GS-1	Hsp90 + **108**	Hsp90 (1 mg/mL) 3µL + 7 µL **108** (700 µM)
GS-2	Hsp90 + **108**	Hsp90 (1 mg/mL) 3µL + 7 µL **108** (700 µM)

While GS-0 was prepared as a reference of the blank sample, GS-1 and GS-2 were two identical Hsp90-**108** complex samples. GS-1 and GS-2 were incubated at 37 ˚C for 24 hours to in accordance with samples prepared in the MST and native page gel tests. These samples were sealed, packed, and delivered under room temperature. Because the samples had to be in a queue for analysis, they were stored at -20 ˚C for 2 months before the first trial.

Then, these samples were analyzed by MALDI in positive-linear mode, with a mixture of DHB and CHCA matrices. However, no visible material was in the submitted vials, and 10µL of the recommended water/DMSO solvent was added. The first set of data was acquired with a sample to matrix ratio of 1:10; nothing is labelled in each spectrum as no peaks were observed (the “peaks” in the baseline are due to electronic noise) (Figure. 187)

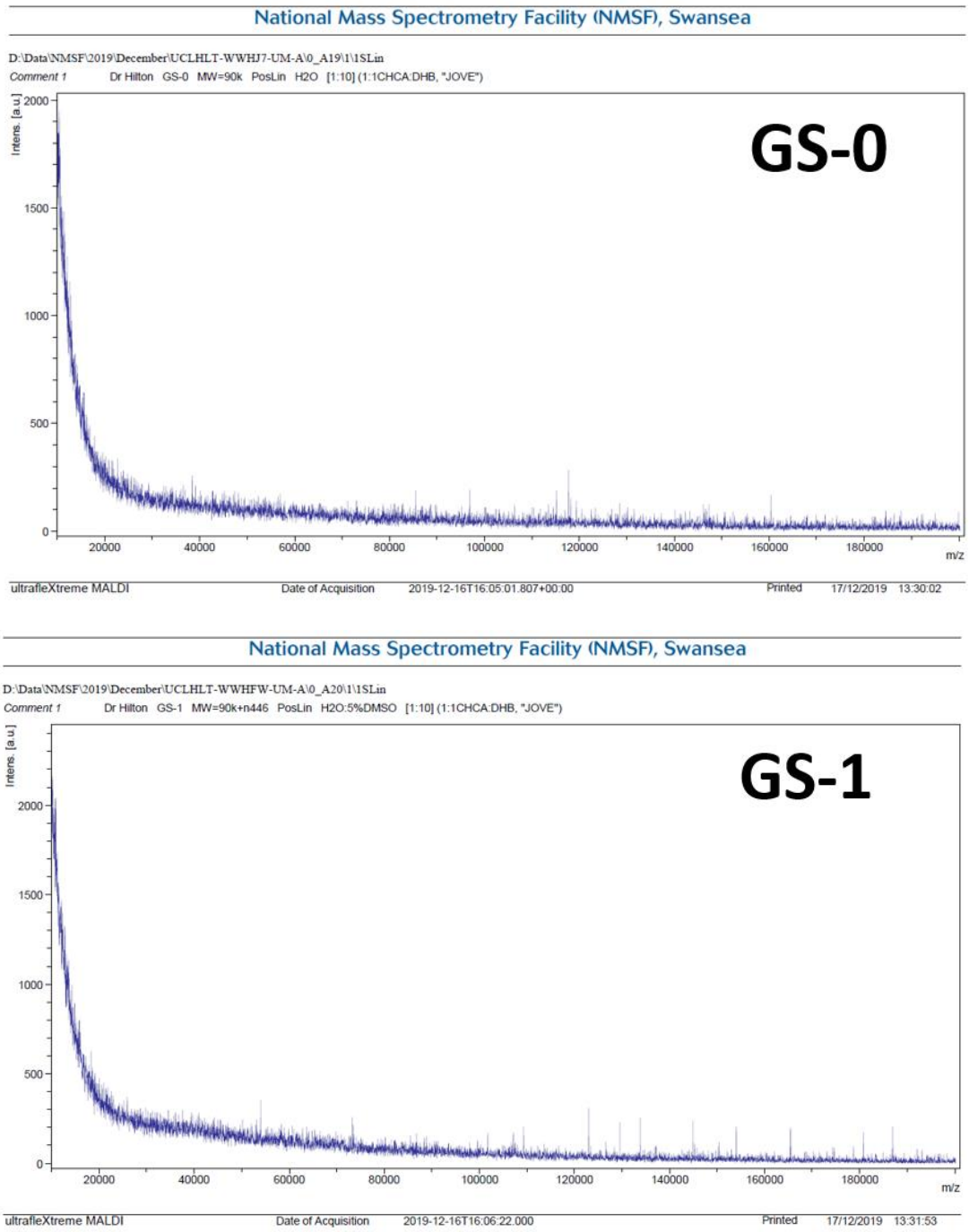


Figure 187: First MALDI-TOF trial using 1:10 matrix ratio showed no peaks.

The second set of data was for more concentrated 1:1 layer preparation; again nothing was observed for the unmodified sample, but both the modified samples indicated very broad ‘humps’ in the baseline at ~40kDa. The crystallization of these latter preparations was poor and indicated possibly too much sample was loaded in this trial (Figure. 188).

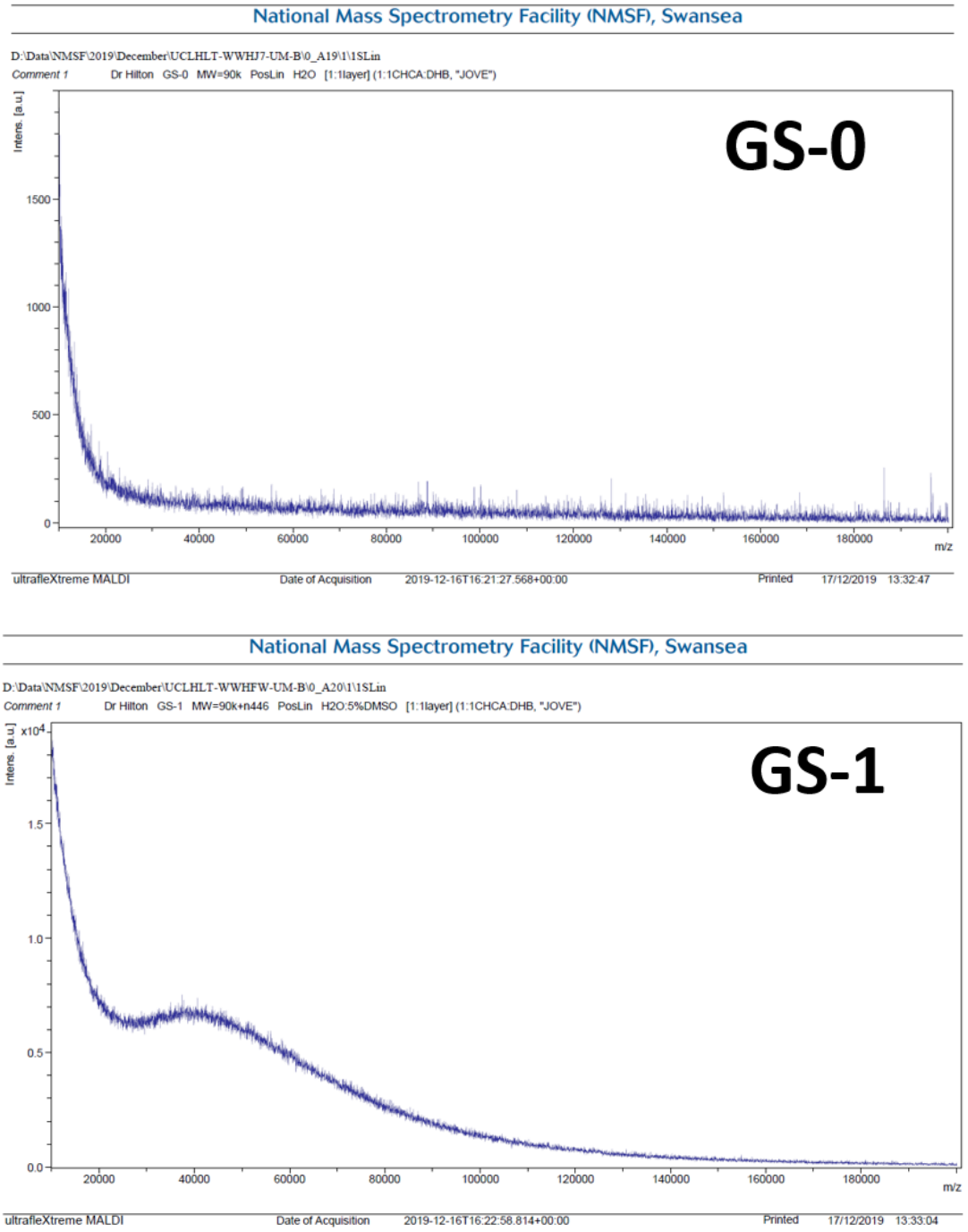


Figure 188: First MALDI-TOF trial using 1:1 matrix ratio showed no peaks

As a result, for completion of a final set of less concentrated 1:49 matrix preparations were analyzed, again, no ions were observed for all three samples (Figure. 189).

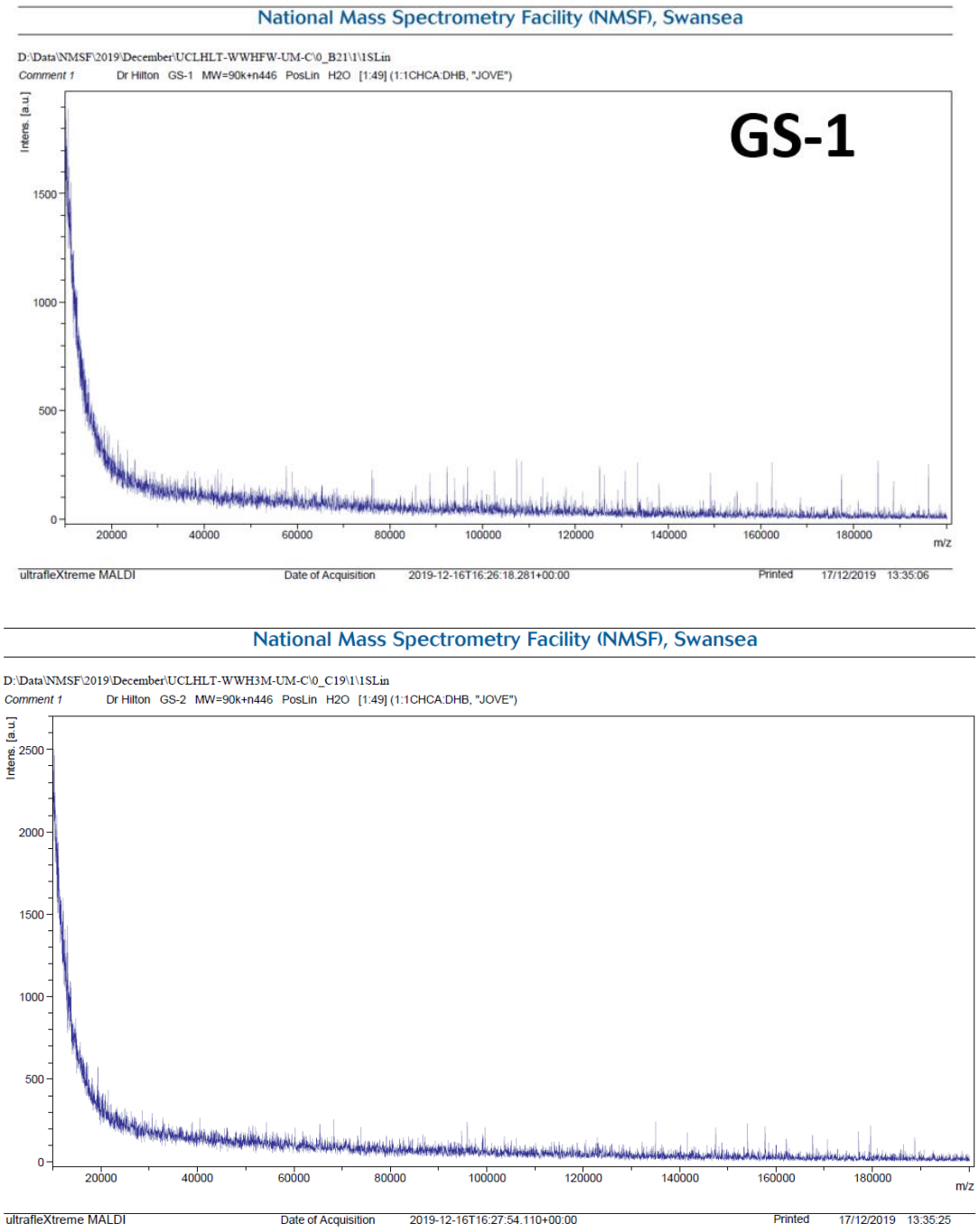


Figure 189: First MALDI-TOF trial using 1:49 matrix ratio showed no peaks.

It was suggested by the technician at NMSF that there were a few explanations for these results. Firstly, perhaps there was nothing in the vial. This was not likely due to the sample preparation was carefully monitored and recorded. Secondly, perhaps there was not enough sample in the vial. At the same time, nothing was visible to the naked eye in the vials, anything in the µg-ng range

should have been sufficient to get a reasonable spectrum, and pg amounts should still have indicated low-intensity species.

Thirdly, the samples had not been sufficiently cleaned up; salts and buffers used in protein purification procedures can suppress ionization; therefore, samples should be cleaned and then lyophilized or stored in pure solvent. Considering that the commercial protein sample purchased was dissolved in tris-buffer, we requested that the technician perform desalting and reanalyze the sample. The desalting process would remove any unwanted inorganic ions which may remove the noise peak and reveal the product peak.

After another 2 months' queueing, our samples were analyzed by MALDI in positive-linear mode, with a mixture of DHB and CHCA matrices again. Prior to analysis, the samples were cleaned up and concentrated via an established C4 Zip Tip protocol. Remarkably, there does seem to be a slight benefit from this protocol, in that possible doubly and singly protonated species are observed at *m/z* 43587 and 86997 in GS-0, respectively (Figure. 190).

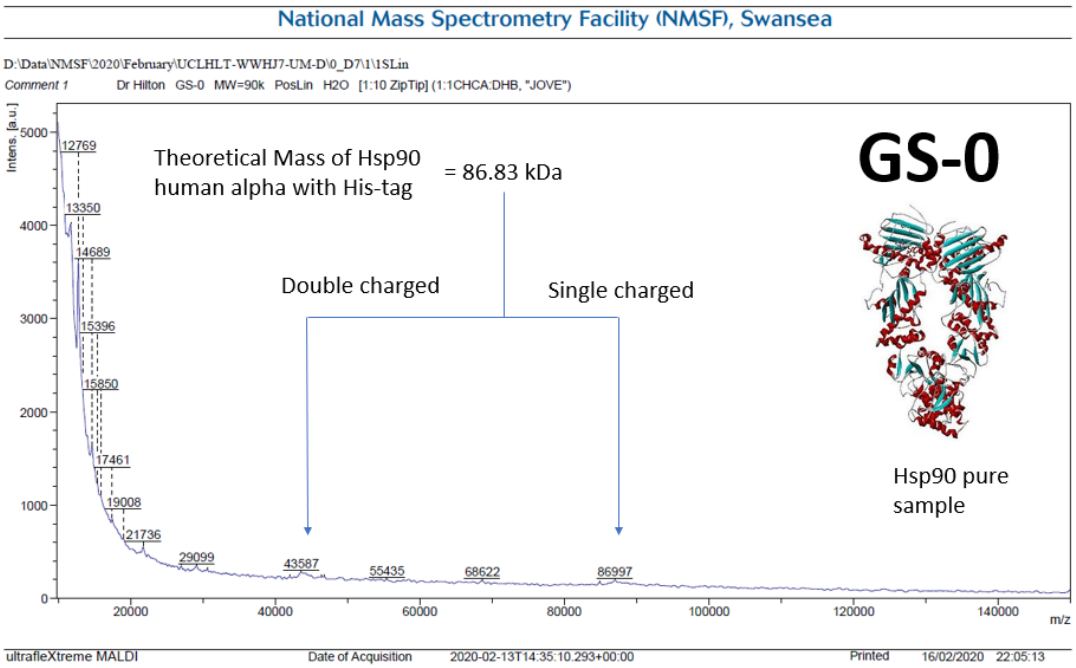


Figure 190: Second MALDI-TOF trial results of GS-0, desalted by Zip-Tip.

However, these species were still extremely low intensity and inadequate evidence for the desired unmodified protein. This could due to either the concentration of protein was too diluted, or the protein was washed off during

desalting operation (it is normal to lose some quantity of sample after wash). To find out a better blank reference for this trial, herein we refer to a spectrum carried out by Dr Xuexia Huang from Min's group (Figure 191).

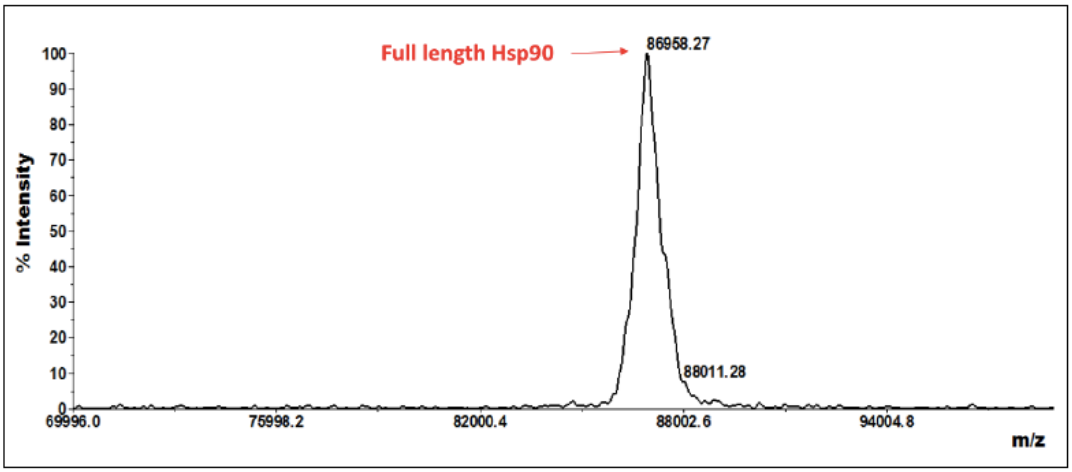


Figure 191: MALDI-TOF MS spectrum of human Hsp90, His-tagged Hsp90 was detected for a mass of 86958 (Supported by Dr Xuexia Huang).

Mass of Hsp90 in our MALDI-TOF result was 86997 Da, which was identical to the reference we showed in figure above (86958 Da in Figure. 191), also in accordance with its theoretical mass of 86.8 kDa (uniport p07900). Although the peak was small and weak, it correctly presented the actual mass of Hsp90. The result validated the purity of the unmodified protein.

Then, the same pre-desalting procedure was applied to **108**-Hsp90 complex sample GS-1 (Figure. 192) and GS-2 (Figure. 193), afterwards, their MALDI-TOF results showed similar extremely low-intensity species for both the modified samples with a noticeable shift to higher mass. Especially for GS-2, a relatively stable peak with mass > 86.8 kDa was formed:

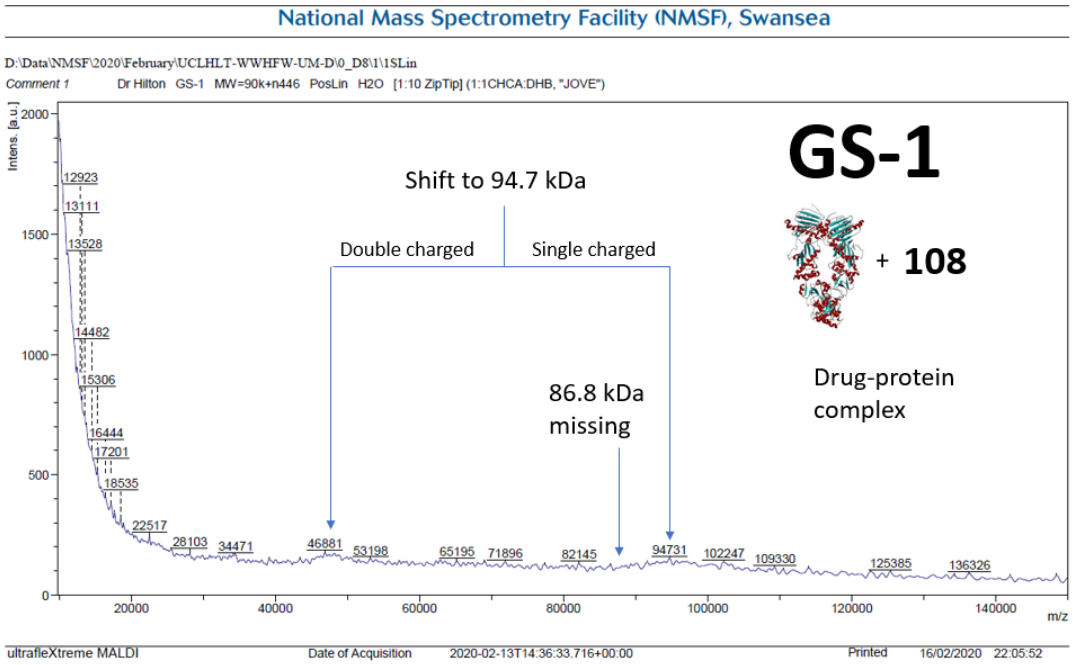


Figure 192: Second MALDI-TOF trial results of GS-1, desalted by Zip-Tip.

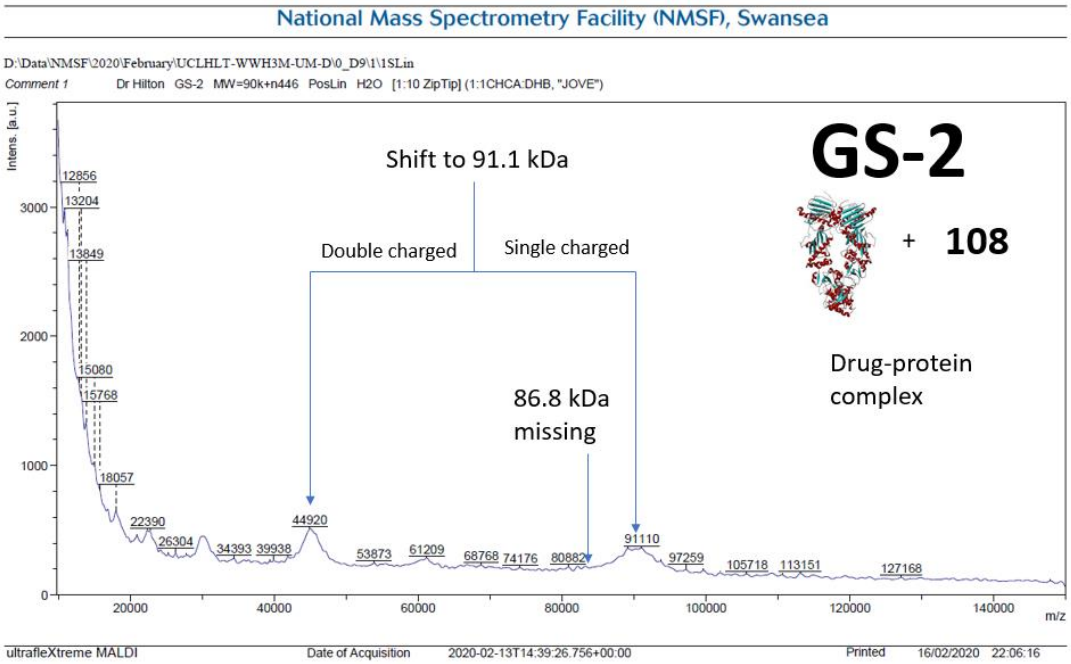


Figure 193: Second MALDI-TOF trial results of GS-2, desalted by Zip-Tip.

GS-2 (Figure 193) clearly showed a peak at 91.1 kDa and its double-charged form can be found at 44.9 kDa. The board peak shape suggesting that Zip Tip cleaning was beneficial; however, the peak still presented a low resolution.

Subsequently, an established a distinct magnetic bead cleans up protocol was followed, and the samples reanalyzed. Only GS-2 showed any ions (Figure. 194), and the intensity had dropped relative to the ZipTip data. This could be due to a majority of the sample being removed during the ZipTip protocol, as there was not enough sample to use separate aliquots for each cleanup. Had the magnetic beads been used first, that data may have had higher intensities.

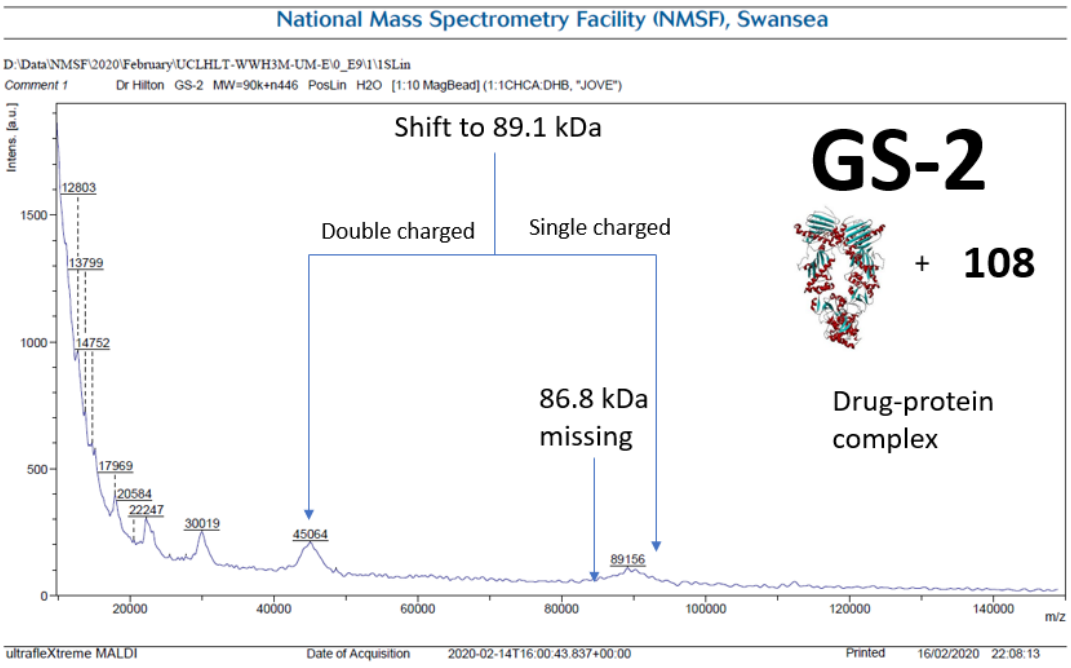


Figure 194: Second MALDI-TOF trial results of GS-2, desalted by magnetic bead.

MALDI-TOF results of GS-2 (Figure. 193 and 194) sample showed a more apparent peak at 89.1 kDa, and its double-charged form was found at 45054 Da. However, peak at such low intensity do not define the right mass added to the protein. We concluded that there was not adequate sample of enough purity to obtain reliable evidence for the modifications of the starting protein, but the data does indicate some modifications have taken place; otherwise, there should not always be found a pair of single/double charged data. Overall, MADLI-TOF analysis provides some evidence of covalent binding between GS-2 and **108**, while it failed to give any precise information about the changing of protein mass. Due to limited time and resources, there was no other sample be submitted to NMSF for further investigation. If so, a sample

with a higher concentration of protein should be submitted for better peak intensity.

3.3.15 Proteomic analysis 108-Hsp90 full-length protein complex for characterisation of Hsp90 CTD active cysteine residues.

As discussed previously, in both Hsp90 native-gel assay and MALDI-TOF analysis, we found some evidence of **108**-Hsp90 covalent modification. In order to find out the exact residue modified by **108**, a proteomic analysis using an LC-MS/MS tandem mass spectrometer was carried out.

Tandem mass spectrometry is known as MS/MS. It is a technique in an instrumental analysis where two or more mass analyzers are combined together using an extra reaction step to enhance their performances to study chemical samples. This method has wildly been used in the analysis of protein, peptides and biomolecule, including drug-protein complex (McNaught and Wilkinson 1997).

The molecules of analytes are ionized, and the first spectrometer (inbuilt special MS1) separates these ions by their mass-to-charge ratio (*m/z* or *m/Q*). Ions of a particular *m/z*-ratio coming from MS1 can be selected and then break into smaller fragment ions, e.g. by collision-induced dissociation, ion-molecule reaction, or photodissociation. These fragments are then introduced into the second mass spectrometer (MS2), which in turn separates the fragments by their m/z-ratio and detects them. The fragmentation step makes it possible to identify and separate ions that have very similar m/z-ratios in regular mass spectrometers due to their different subsequential fragmental mass (Figure. 195).

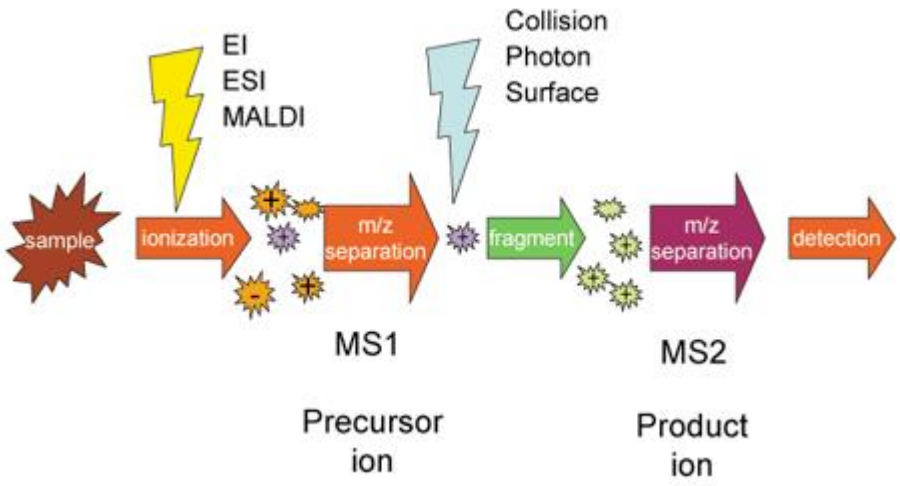


Figure 195: Mechanism of MS/MS spectrometry.

As a result, the peptide which may carry the **108** modification, MS1 would detect the full mass of peptide (precursor ion). However, the mass value may also refer to some other unmodified peptides with a totally different sequence. MS2 provides a solution for this case, as the suspicious mass will be fragmented again into different ions, fragmentation result will be unique due to different chemical structure.

To increase the accuracy of MS/MS results, usually, whole protein samples are digested into peptides range among several thousand Da (in most cases protein are digested by trypsin) and then sent to LC. For each second in LC, the machine will sample peptides form LC and pass them through MS/MS for analysis. This will result in a raw data of undetermined MS/MS peak pack, which not yet be interpreted into identified peptides. Process introduced above is called ‘data acquisition’ (Eng, Searle et al. 2011) (Figure 196, left arm).

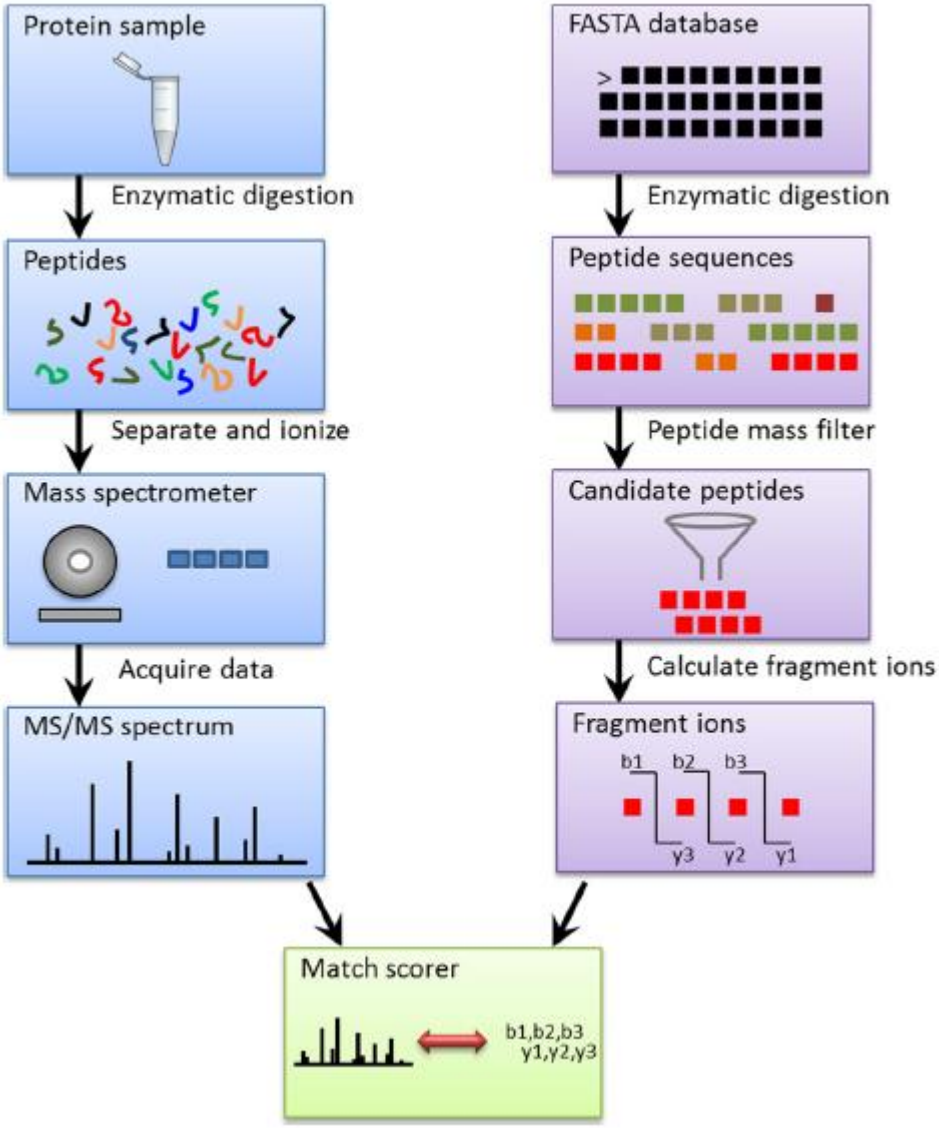


Figure 196: A typical workflow of proteomic analysis using tandem MS spectrometry.

After data acquisition, the MS/MS spectra are searched against a protein sequence database to identify the peptides and proteins represented in the acquired spectra. The search requires developed and validated algorithms.

Because sequence of protein is already known (sequence see experimental p306), sequence of its digested peptides can be calculated or can be found in a database such as the FASTA database (Martínez-Ruiz, Villanueva et al. 2005), which is generally used by different algorithms (decoy). Herein, we will introduce a simplified mechanism of searching algorithms adopted by most of the searching tool. In general, searching algorithms will read a collection of

MS/MS raw data from data acquisition, then query a sequence database to select peptides of the right mass, score these peptides against the experimental spectra, after which will return with peptide characterisations.

For each MS/MS spectrum, an experimental peptide mass will be derived from the precursor m/z ratio and estimated or measured precursor charge state. A database search tool (e.g. Commet) will select peptides from the sequence database that are of the same approximate mass as the experimental peptide mass for the query spectrum. The set of potential peptides that get scored against each spectrum are referred to as candidate peptides. The set of candidate peptides is influenced by factors such as the enzyme specificity setting, and any post-translational modifications as well as ligand modification being considered in the search (Parker, Mocanu et al. 2010)

A set of expected fragment ion masses are then calculated for each candidate peptide sequence. The activation method used to fragment the peptides defines which types of fragment ion masses are considered, *e.g.* b- and y-ions for collision-induced dissociation and c- and z-ions for electron transfer dissociation. These fragment ions are compared against those in the experimental spectrum, and a similarity score or closeness-of-fit metric is calculated. Each candidate peptide is scored, and the highest-rated peptides for each spectrum query are reported (Eng, Searle et al. 2011)

In other words, the computer will simulate digestion process and generates a set of predicted results. If the data acquired matches those simulated data, they will be reported with its peptide identity with a score which reflects the credibility. Such a method for determination of peptide is called ‘database search’ method (Barsnes and Vaudel 2018).

Another method to interpret the raw MS/MS data is known as de novo sequencing. This is different from database search method that starts with prediction and simulation, de-novo sequencing use a totally different algorithm to 'translate' all spectrums inside the raw MS/MS data set into all possible sequences, regardless of which protein it is (Edmen and Begg 1967). Database search is good enough for identification of any existing protein; however, it will not characterize any new sequence due to there will be no record in the database for it to search against. De novo sequencing is an assignment of fragment ions from a mass spectrum. It will calculate all the possible sequence for each precursor mass in spectrums and generates a grid (Figure 197). The main idea of *de novo* sequencing is to use the mass difference between two fragment ions to calculate the mass of an amino acid residue on the peptide backbone. The mass can usually uniquely determine the residue. For example, for a spectrum generated from a precursor, de novo sequencing will firstly figure out all possible sequences according to the mass of precursor. Then for each sequence, its theoretical fragment ions are calculated and searched against the spectrum. The more fragment ion mass it finds, the better the score. Obviously, only the correct aligned sequence will match both the precursor mass and the fragment ion mass (Figure. 197).

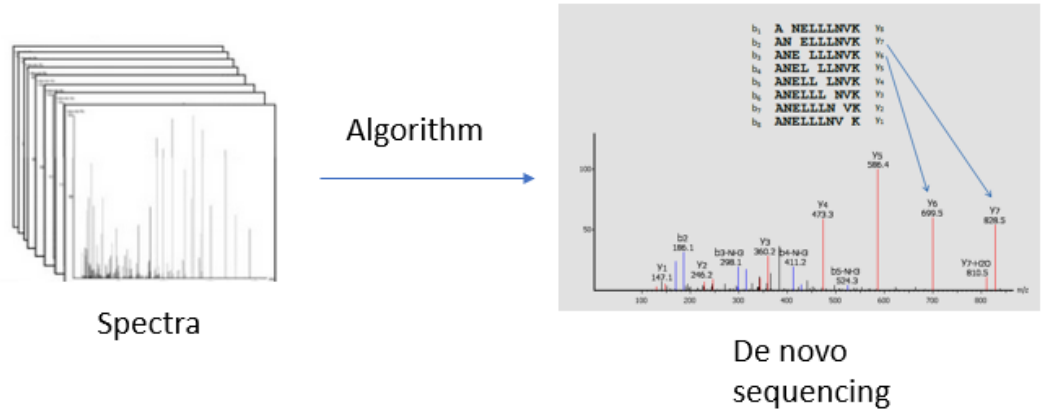


Figure 197: Mechanism of de novo sequencing.

As de novo sequencing interprets MS/MS data from a different angle, some searching engine will combine database search and de novo sequencing together to provide a more reliable result. This type of searching engine will match the identified sequence between both interpretation methods, picking

out the overlapped region and score again. In our project, we used such a combined method for better verification of **108**-Hsp90 binding site.

3.3.15.1 Method of sample preparation and data processing.

According to the basis of proteomic analysis, the input for MS/MS must be small molecule peptides; therefore, the drug-protein complex sampled from any source must be digested by an enzyme (e.g. trypsin) to become peptides. For the preparation of Hsp90-**108** protein-drug complex band, the suspected blue colour band located at native-page electrophoresis gel was cut (Fig. 196). The gel band was then digested by DigestPro automated digestion unit (Intavis Ltd.). The resulting peptides were fractionated using an Ultimate 3000 nano HPLC system in line with an LTQ-Orbitrap VelosPro mass spectrometer (Thermo Scientific). After this period, the raw MS/MS data set was generated and can be used for the database search.

Interpretation of MS/MS data was completed using Search GUI software (Barsnes and Vaudel 2018), and results were visualized through PeptideShaker software v 2.0.0. The advantage of Search GUI-PeptideShaker system is that it searches against multiple databases with various algorithms at the same time together with de novo sequencing, providing more accurate and reliable results compared with searching against one single database. Obviously, data analysed by multiple searching engine provides better interpretation and repeatability. Search GUI-PeptideShaker system currently supports MS/MS searching engines including X! Tandem, MS-GF+, MS Amanda, OMSSA, MyriMatch, Comet, Tide, Mascot, Andromanda and mzIdentML. At the time of writing, there are 328 publications in the area of proteomics cited PeptideShaker software since its birth in 2015 (Vaudel, Burkhart et al. 2015).

Database search using Search GUI was carried out with 2 inputs. First one was MS/MS raw spectrum data file obtained from the instrument. The second input was a decoy sequence with or without modifications. Despite the FASTA

sequence downloaded from Uniport databank (ID: P07900), user-defined modifications were also defined into the decoy file. This includes oxidation of methionine (general requirement, mass +15.99), carbamidomethylation of cysteine (removal of free disulfide bonds, mass + 57.02) and covalent addition of compound **108** to cysteine (mass +666.24). Tolerance of precursor (mass of full peptide) was set to 10 ppm, and tolerance of fragments was set to 0.5 Da (default).

When input files were ready, the next step was to select multiple searching engines and algorithms for de novo sequencing. After a few tests, we found that not every searching engine could process our raw MS/MS data due to format issue. In the end, X!Tandem. Comet and Andromeda were 3 searching engine selected which successfully give consistent results. Algorithms used for de novo sequencing were in-built Novor and DirecTag.

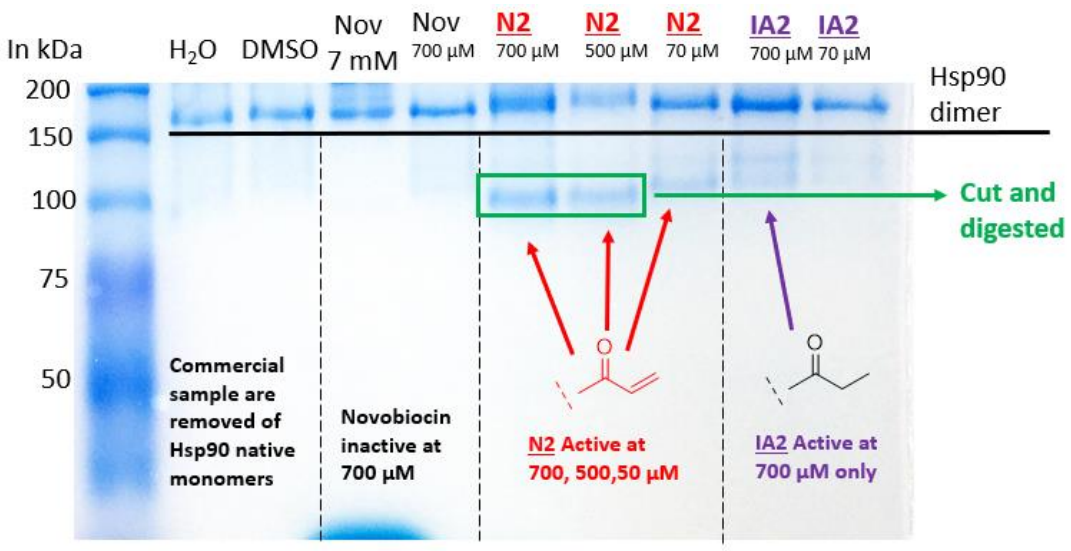


Figure 198: Gel band of interest was cut for trypsin digestion and MS/MS. The band was suspected to represent the Hsp90 monomer-**108** complex.

3.3.15.2 Results and discussion

Search for 108 exact mass 666.24 failed to find any hits.

After the searching result be visualized by PeptideShaker, we identified no **108** modification in the sample. In this entry (Figure 199), 147 peptides with > 98% confidence were verified. As previously we reviewed that Hsp90 Cys597/Cys598 plays an important role during Hsp90 dimerization, we found that the software validated the peptide with 100% confidence. Both cysteines were alkylated by pre-added alkylation reagent.

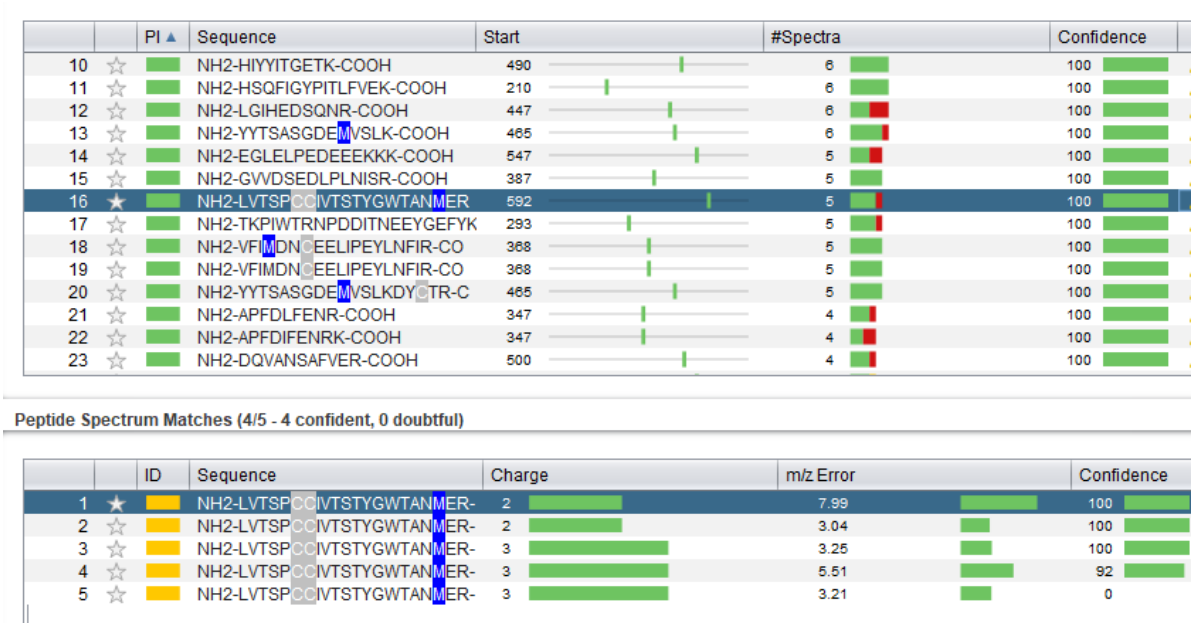


Figure 199: PeptideShaker identified NH2-LVSTPCCVTSTYGWTANMER-COOH sequence which contained the important switch point cysteine.

Identification of this specific sequence suggested that the digest of protein was a success (Figure 199). Since the peptide of interest did exist inside the MS/MS data pack, one of the reasons we did not find a modification of **108**, could be we do not know the exact mass which added to the protein. Usually, in published work, MOLDITOF analysis was used to determine this value, as

we discussed in the last chapter. However, in this case, MOLDITOF was not successful, making the search for modification difficult.

Search for 108 + proton, which is exact mass 666.24 + 1 confirmed its binding to Cys597/Cys598.

As discussed previously, search for **108** exact mass 666.24 found no hits indicated some potential variation of ligand mass might happen. We reviewed our preparation process and noted that all mass spectrum was recorded under ESI+ mode. **108** recorded by mass spec through ESI+ mode is usually protonated, showing $[M + H]^+$ = 667.24 *m/z* instead of [M] = 666.24. Considering this possibility, we carried out another database search following the same procedure except changed the mass variation from 666.24 to 666.24 + 1 = 667.24. Surprisingly, we found modified sequence NH2-LVTSP**C**< [**108** + H]> **C**<[**108** + H]> IVTSTYGWTANMER-COOH was marked with 100% confidence, and there was no other cysteine contained peptides marked as modified (Figure. 200).

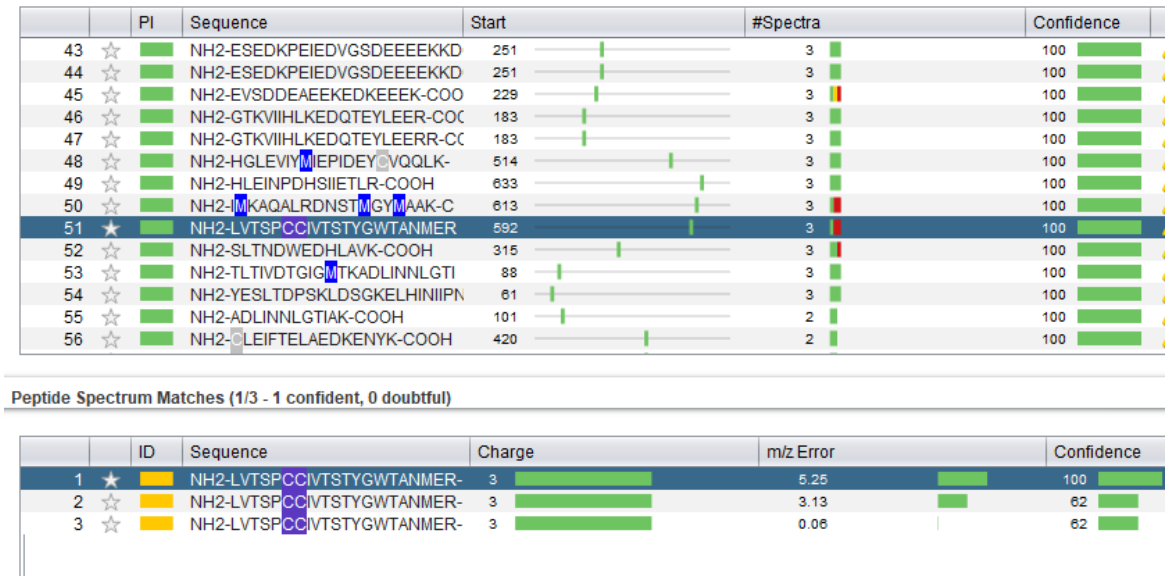


Figure 200: PeptideShaker identified NH2-LVSTPC<[**108** + H]> C<[**108** + H]> VTSTYGWTANMER-COOH modification. This modification (+ 667.24) is marked in purple.

For the modified sequence, 3 spectrums were matched, and one of them was in good quality, scored 100% confidence. In order to check whether 666.24 +1 Da search was a random match, we searched 666.24 -1 Da and 666.24 +2 Da and resulted in no matches.

The precursor mass of **108** modified peptide was 1222.87 *m/z* with 3 charges, represented a peptide with a mass of 3668.61 Da. Unmodified same peptides were also found in both entry with a mass of 1231.57 *m/z* with 2 charges represented a peptide with a mass of 2463.14 Da. Due to cysteines and methionine residue in unmodified peptides were methylated and oxidized, the mass of native unmodified target peptide is 2333.1 Da. Therefore, variation of mass between modified and unmodified target peptides is 3668.61 – 2333.1 = 1335.5, which is two times of [**108** + H] mass (1335.5 / 2 = 667.7)

When precursor peptides are subsequently fragmented, it does not fragment sequentially, in other words, the first fragmentation event does not start at the amino terminus and proceeds sequentially one residue at a time down the amino acid chain. The fragmentation events are somewhat random and not sequential (Biemann and Martin 1987). The typical pattern in terms of fragmentation of peptides is described as **a**, **b** and **y** ions (Figure. 201).

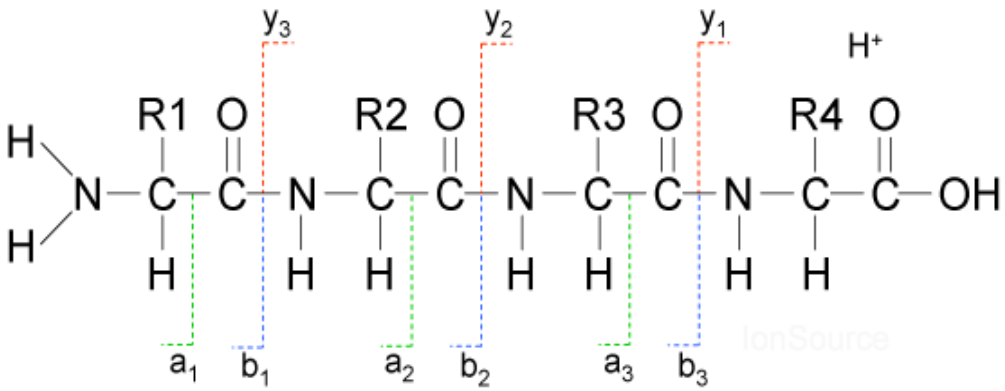


Figure 201: Example of MS fragmentation. Breakage of CO-NH generates **y** and **b** ions. While **y** ions point towards carboxy terminus, **b** ions point towards the amino terminus.

The most common peptide fragments observed in low energy collisions are **a, b** and **y** ions, as described in the figure above (Figure. 202). The **b** ions appear to extend from the amino terminus, sometimes called the N-terminus, and **y** ions appear to extend from the carboxyl terminus or C-terminus. While readily observed and diagnostic for **b** ions, **a** ion occurs at a lower frequency and abundance in relation to **b** ions (Biemann and Martin 1987).

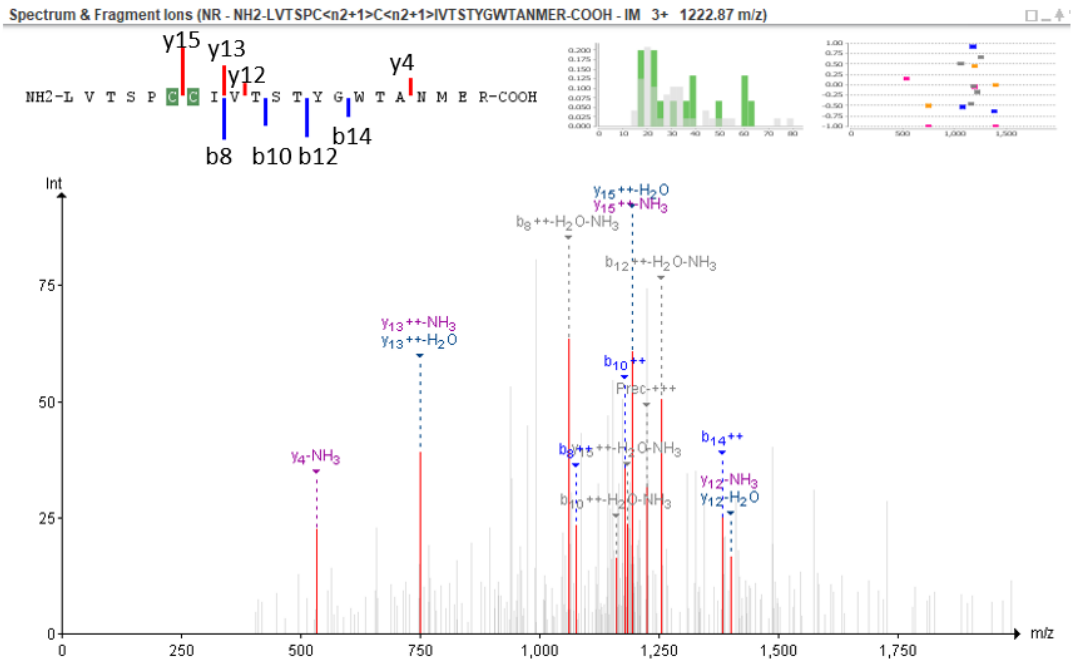


Figure 202: Spectrum and fragment Ions analysis of **108** modified peptide by Peptide shaker. Red represents y ions, and blue represent b ions.

Spectrum and fragment ions of a modified peptide are given above. **y** and **b** ions peaks showed in the figure are picked by software with high-resolution. For **y** ions, if Cys598 was modified, then the mass of [**108** +H] (667 Da) should be found. In the figure above, deviation between $y15^{++}$ and $y13^{++}$ represents mass of CI peptide + [**108** + H]. Herein, both y15 and y13 are double charged; their deviation mass can be calculated as 1200.02 * 2 – 758.35 * 2 = 883.34 Da, obviously higher than residue C + I alone. The mass of CI peptide, in this case, is [mass of peptide CI] – H2O. The loss of water comes from fragmentation upon terminus. CI peptide has a mass of 234.30. Herein, [mass of $y15^{++}$ - $y13^{++}$] – [mass of CI peptide – H_2O] = 883.34 – (234.30 – 18.00) = 667.04 Da Considering true mass of [**108** + H] = 667.24 Da, the result is within

the 0.5 Da tolerance. Moreover, since the protein-drug complex sample was prepared from a commercial pure Hsp90 product, there is no peptide from other protein source act as an interferer, making the peak assignment much more reliable. With the strong evidence from y15 and y13, the rest set of **y** ions simply helped to identify the peptide sequence. For **b** ions, deviations between b8, b10, b12, b14 contributes to the identification of sequence. Mass of $b8^{++}$ (mass = 2152.90 *m/z*) alone demonstrated that two drug ligands bound to the peptide. [mass of $b8^{++}$] – [mass of LVTSPCCI – H_2O] = 2152.90 – (835.78 – 18) = 1335.12 Da. 1335.12 / 2 = 667.56 Da, representing two units of ligand, 667.56 Da each were added to peptide covalently.

Sequence NH2-LVTSPCCIVTSTYGWTANMER-COOH containing two **108** modified Cys597/Cys598 residues was the only sequence identified as **108** modified sequence. The result is in accordance with literature (Retzlaff, Stahl et al. 2009) defined Hsp90 CTD switch point which located at Cys598. Our results provided reliable proteomic information about the binding pocket of novobiocin scaffold. 4'-substituents of novobiocin analogues no doubt can interact with that region.

Another proteomic analysis we have done was to confirm uncharacterized Hsp90 bands from electrophoresis gel page results (Figure. 203), made with lab-produced Hsp90 protein. We searched for literature but only found all literature either ignore it (Cruz, Zhang et al. 2013) or just claim it to be Hsp90 without any proof (Krishna, Reddy et al. 1997). This is due to most research of Hsp90 electrophoresis focus on SDS page rather than the native page. It is crucial to confirm such bands, especially the one located at 100 kDa, is not impurity from bacterial cells.

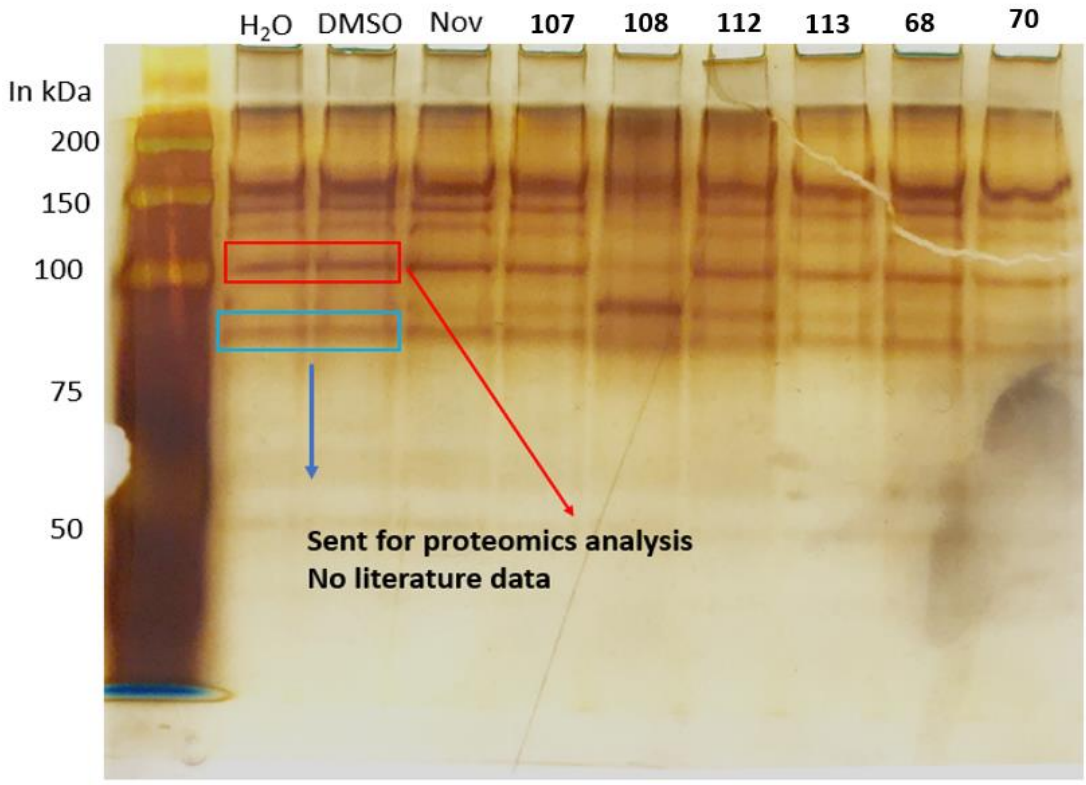


Figure 203: In the lab-produced Hsp90 sample, there are two uncharacterized bands located at 100 kDa and 75 – 100 kDa region. Both were sent for proteomic analysis in our project. The result confirms they are from Hsp90 alpha human.

The tandem MS/MS analysis was done by our collaborator from the University of Bristol and identified them both as Hsp90 with identical sequence towards Uniport database sequence (P07900). It can be proposed that monomeric Hsp90 has two forms and 100 kDa form was more active towards Hsp90 CTD inhibition as 100 kDa band decreased when Hsp90 CTD inhibitor was induced. By confirming two types of native form Hsp90 monomers, it will help to distinguish between Hsp90 dimeric and monomeric inhibition. A detailed discussion of the mechanism, please see the previous MST test chapter (p252).

Chapter 4: Conclusions and Future work

4.1 Conclusions

Hsp90 has been an important anti-cancer drug target over the last two decades. Its inhibition will subsequentially inhibit multiple onco-enzymic pathways. While Hsp90 NTD inhibitors failed to pass phase II & phase III clinical trials due to high toxicity and unwanted cancerous antiproliferative heat shock response, Hsp90 CTD inhibitors overcame the negative heat shock response and maintained general benefits from Hsp90 inhibition. Novobiocin is a lead structure act as Hsp90 CTD inhibitor. However, due to lack of high-resolution crystal of Hsp90 CTD, binding pocket of novobiocin towards Hsp90 CTD remained unsolved and significantly limited the development of more potent novobiocin based Hps90 CTD inhibitors. Knowledge of SAR of Hsp90 CTD inhibitors is urgently needed.

In previous research, we found that glycosylation of novobiocin on its 4' position dramatically increases its antiproliferative activities. As a result, the first part of this project chapter 2 aims to synthesis glycosylated novobiocin analogues, including 4',7'-dia glycosylated products substituted with different amide sidechains. These side chains, including 2-indole amide and Bi-aryl amide, which were reported to increase antiproliferative activity according to literatures. The synthesis started with the degradation of novobiocin with acetic anhydride solution then 10M HCl to afford 3-amino-4,7-dihydroxy-8-methyl-coumarin **37**. In the following step, relevant carboxylic acids were mixed with synthesized novobiocin coumarin core in 50% pyridine/DCM catalyzed by EDCI to complete amide bond formation. The EDCI catalyzed amide coupling yield in the correct form of core-indole **131** and core-Bi-aryl **42** amides. We also found that for crude indole amide **131** product, the purification can be simplified by precipitation using MeOH, without using time-wasting flash column. Glycosylation of these synthetic amides proved to be difficult as traditional Lewis-acid catalyzed glycosylation did not work effectively. This is due to poor reactivity of 4',7' hydroxy groups and poor solubility of amides. We developed two methods to solve this problem. For 4' mono glycosylation, acetyl protected sugar bromide donor was mixed with novobiocin and stirred in DMF would result in a neat reaction, giving expected 4'-glycosylated

products **64** and **65** without any side product. DMF, in this case, acts as a solvent and regio-selective catalyst. For 4',7'-diglycosylated product, we successfully developed a phase transfer catalyzed glycosylation method. A 0.15 M K_2CO_3 solution + 1 eqv. BTAB composed water phase, plus DCM/Chloroform + acetyl protected sugar donor bromide composed organic phase. Heated under 45 °C overnight, 4',7'-diglycosylated product **59** was afforded with reasonable yield (16-20%). It also should be emphasized that a significant amount of LC/MS analysis was used in this study. They provide the most valuable information on the side products, relative quantitative yield etc. so that to help to the optimization of reacting conditions.

The purification was also unexpected challenging due to the solubility of products, and similar retention time in the flash column. However, the problem also has been solved using a suitable solvent system and a reverse phase Biotage system. Nevertheless, acetyl protected novobiocin based 4'7'-diglycosides **59**, **68** and **70** failed to be deprotected. Deprotection with different reagent all led to unexpected cleavage of glycosidic bond, which has reported to cause by 3'-substituents effect on coumarin core. In the end, only 4'-mono glycosylated novobiocin **66** and **67** were prepared, the rest of the compounds were remained protected for bioassays. Unfortunately, antiproliferative test result concluded that 4'-deprotected sugar (glucose/galactose) moiety was decisive towards a better antiproliferative activity. 4'-glucose-novobiocin **66** significantly increase drug potency compared with 4'-acetylglucosyl-novobiocin **64**. Protected 4'7'-diglycosides **59** and **70** exhibited poor antiproliferative activities; however, indole amide side chain substitution **68** increase activity which is in accordance with the literature. Overall, although glycosylation of novobiocin moiety increases antiproliferative activity, difficulties of deprotection limited further development of glycoside type novobiocin based Hsp90 CTD inhibitors. There are few issues about this project worth further investigation in the future. Firstly, this project identified a new novobiocin based indo-amide acetyl-glycoside analogue **68** as a lead structure with IC_{50} = 130.9 μM. Due to our data strongly concluded that deprotection of acetyl-glycoside into glycoside would increase the antiproliferative activity significantly, it still worthy to synthesize the

deprotected **68** for SAR studies. The difficulty of acetyl deprotection may be solved by switching the acetyl protective group into other protection groups which requires no base for the deprotection (e.g. benzyl group) to avoid the unexpected cleavage of 4'-glycosidic bond. Secondly, lead structure **68** was not verified for its binding against Hsp90 via any protein-based assay in this project. However, we tested **68** using native page gel Hsp90-drug binding assay (developed in chapter 3) and the result suggested that **68** did not stop Hsp90 CTD dimerization. Whether **68** binds to Hsp90 CTD remained a question and further experiments such as Hsp90-ATPase assay, non-covalent MS study and Western Blot method may be applied to determine the actual binding site of **68**. Thirdly, we reported another important conclusion that 4'-glucose modification exhibited better antiproliferative activity than 4'-galactose which indicated sugar type may varies antiproliferative through a certain mechanism more than just Hsp90 binding affinity (e.g. glucosyl analogues may have different metabolism path way in cancer cells). As a result, more 4'-glycosyl-novobiocin analogues with other sugar types, such as fructose, worth to be synthesized and test in comparison with glucosyl and galactosyl analogues to determine the optimal moiety. In this project, we only reported MTT assay of novobiocin analogues against MCF-7 breast cancer cells. Cell lines which cover more types of cancer, e.g. lung, brain and pancreatic, may applied in the future MTT tests to investigate the effect of different sugar substitutions against different types of cancer.

Since our previous strategy of glycosylation failed to complete the goal of generating an ideal lead structure, an alternative plan was required for exploring the currently almost blank SAR of novobiocin 4'-substitution. Herein, we reviewed literature and learnt that targeted covalent drugs are new types of inhibitors which covalently bond to the reactive amino acid residue of proteins then enhanced the inhibition power. Among targetable amino acids, cysteine modified by Michael acceptors is proved to be the most successful story so far. Several Michael acceptor type EGFR covalent inhibitors were already approved by the FDA for treatment of different types of cancer. There is no reported Hsp90 CTD synthetic covalent inhibitor by far; however, literature already identified Cys597/Cys598 of Hsp90 human CTD plays a role

of the switch point. Mutation or deletion of this specific cysteine reduces its ATPase activity. Also, we found that some natural Michael acceptor product, e.g. zerumbone, was reported to inhibit CTD of Hsp90 with covalent inhibition style, but these studies of natural products never report which residue was modified. Considering these works from the literature, we designed 4'-Michael acceptors substituted novobiocin analogues, act as cysteine specific modifier to verify if cysteines located at Hsp90 CTD are targetable.

All synthesis was carried out in one step by merely mixing novobiocin with relevant acid chlorides. Through this method, cysteine reactive 4'-acrylate-novobiocin **108** and its reduced inactive from 4'- propionyl-novobiocin **109** were synthesized, together with other methyl acrylate substituted novobiocin **110**, **111**, vinylsulfone substituted novobiocin **112**. 4'-tosyl **107** and 4'-sulfonefluoride **113** substituted novobiocin was also synthesized to investigate any potential effect of 4'-electrophile substitution. In the meantime, we also synthesized a series of 4-substituted coumarin, including 4-tosyl coumarin **115**, 4-amino coumarin **116** and 4-thiol coumarin **117**, plus purchased coumarin **122**, 4-hydroxy coumarin **141** to study the 4-substituted effect on coumarin scaffold.

Synthetic novobiocin analogues were tested for antiproliferative activities using 24h MTT assay. All compounds were tested against MCF-7 breast cell line for potency screening. Compared with novobiocin itself, acrylate type analogues **108**, **109**, **110**, **111** and vinylsulfone **112** mildly increase antiproliferative activity range from 2.5-fold to 2.7-fold. These compounds showed improved activity probably due to enhanced lipophilic membrane permeation. Electrophile type **107** and **113**, especially sulfonefluoride **113**, exhibited superior activity which increased 11.9-fold compared with original novobiocin scaffold. **108**, **109**, **110**, **111**, **112** were also tested against multiple cancer cell lines for drug sensitivity. Unlike novobiocin showed very different sensitivity towards different cell lines, synthetic analogues all maintained a mild and robust IC_{50} value range from 500 µM – 700 µM, suggesting that hydrophobic substitution of novobiocin would contribute to better robustness. For coumarins, coumarin **122** and **141** were expected and indeed non-toxic towards cancer cells. While **116** had mild cytotoxicity (933 µM), the potency of

117 and **115** were 1.78 and 1.53 more potent than **116** respectively. As a benefit of 4'-tosyl substitution was already proved in the previous study, 4-thiol substitution upon coumarin was a new identified potent modification and may benefit novobiocin scaffold as well if applied.

In MTT assays, cysteine reactive acrylate analogue **108** showed the same potency as its inactive form **109**, announced the failure of covalent modification under cellular level. To investigate whether **108** and **109** have a varied effect at the protein level, a drug-Hsp90 dimerization inhibition assay, carried out by native page acrylamide gel electrophoresis, was developed. According to literature, native dimeric Hsp90 can be disassociated into monomer by Hsp90 CTD inhibitors like novobiocin; as a result, a new monomer band will be generated on the native page when the sample was incubated with Hsp90 CTD inhibitors. It also worth to note that the recombinant full-length Hsp90 protein used for dimerization assay must not be C-terminus His-tagged, as this modification will negatively affect the process of Hsp90 CTD inhibition. Surprisingly, 700 μM **108** creates a much stronger Hsp90 monomeric band than **109**. When the ligand solution was diluted into 70 μM, **108** still generated a strong Hsp90 monomeric band; however, **109** totally lost inhibitory effect. Obviously, the inhibitory effect of **108** is superior to **109**, showing potential covalent inhibition power. To find out why **108** lost its potential covalent power, we reacted **108** with common cellar detoxing reagent glutathione as well as common cellar enzyme esterase. We found that glutathione can reduce acrylate into inactive form and esterase can cleave ester bond of **108**. This means that the cysteine reactive warhead of **108** could already be detoxed before it reaches cellar Hsp90, therefore **108** and **109** shares similar activity under cellar level but varies under protein level. In order to resist the detoxing process of cancer cells, we modified acrylate group into a more hydrophobic cinnamic group. MTT assay showed that 4'-cinnamic-novobiocin was approximately 4-fold more potent than its inactive form 4'-hydrocinnamic-novobiocin. This suggested that potential covalent inhibition may have happened, and benzylic hydrophobic addition helps to resist the detoxing process. Moreover, using our developed drug-Hsp90 CTD inhibition assay, we identified that anti-proliferative potent analogue **107** and **113** could not

dissolute Hsp90 dimers into monomers, suggesting that such modification may change its binding pocket in Hsp90 or protein target. Future work would worth to explore the binding site of **113,** which may be different from binding pockets of other novobiocin analogues. Hopefully, a new binding pocket may be identified, then new types of Hsp90 inhibitor could be synthesized and add new knowledge to Hsp90 inhibition drug design. **117** and disulfiram were identified as strong Hsp90 CTD inhibitors. Thus, 4'-sulfur modification on novobiocin might be a promising modification in terms of drug potency, and disulfiram here was firstly reported as an Hsp90 CTD inhibitor. Additionally, while commercial Hsp90 protein only showed one dimer band on the native page, our lab-produced Hsp90 sample contains two extra protein band located at 100 kDa and 75 – 80 kDa area respectively. We later verified these additional bands through proteomic analysis. They were not impurity proteins from bacteria; they are Hsp90 monomeric form firstly validated by our group which have not been mentioned by any literature. 100 kDa Hsp90 monomeric band notably faded when incubated with Hsp90 CTD inhibitors, suggesting that Hsp90 CTD inhibitors inhibit not only dimeric Hsp90 but also Hsp90 monomers. It differs from the traditionally reported mechanism in which kinetics of dimeric protein is not being considered.

To determine the binding affinity Kd of novobiocin, **108** and **109** against Hsp90 full-length dimeric protein. We developed a native MST method. It worth to note that due to relatively high molecular weight of Hsp90 dimer, the incubation of labelling dye towards Hsp90 sample required a much longer period of 1-4 days in the fridge, instead of following the 30 mins duration decided by the user manual. Novobiocin, **108** and **109** were measured to had binding affinity Kd of 8.53 mM, 2.78 mM, 0.95 mM respectively, dramatically less active than those micromolar values measured by antiproliferative MTT assay. To explain this phenomenon, we reviewed our native gel results and reanalyzed the process. We found that noncovalent inhibition between Hsp90 full-length protein and Hsp90 CTD inhibitors has two steps. MST only quantified the binding affinity of Hsp90 CTD inhibitors towards Hsp90 dimer, but unable to quantify the binding affinity against Hsp90 monomer. Both MST and native page gel electrophoresis measures the same native process of drug-protein

interaction. The worse binding affinity of Hsp90 CTD inhibitors indicates that inhibition against Hsp90 dimer is much more difficult than inhibition against Hsp90 monomer; obviously, steric hindrance effect of protein quaternary structure account for this difficulty. Although it was not quantified in this thesis, we previously have concluded that 100 kDa Hsp90 monomer is sensitive to Hsp90 CTD inhibition, as for **108** even at 70 μM, it faded 100 kDa, Hsp90 monomer band. Monomeric Hsp90 inhibition no doubt has better binding affinity than dimeric Hsp90 inhibition. The traditional equation was modified to take consideration of dimeric inhibition.

E_{Hsp90} = Hsp90 monomer
E_{dimer} = Hsp90 180 kDa dimer
I = Inhibitor

$$E_{dimer} + I \underset{}{\overset{K_{di}}{\rightleftharpoons}} E_{Hsp90} + I \overset{K_i}{\rightleftharpoons} E_{Hsp90} \cdot I \xrightarrow{K_2\ (K_{inact})} E\text{-}I$$

non-covalent non-covalent covalent

Finally, we applied MADLI-TOF and tandem MS/MS proteomic analysis to investigate the covalent binding site of **108** against Hsp90 dimer. MADLI-TOF MS failed to capture a reasonable mass peak of Hsp90, and Hsp90 + **108** due to the amount of protein-drug complex sample almost washed off by desalting pre-treatment. However, it existed weak signals on the spectrum indicated that there are some modifications already happened. Then, we digested the Hsp90-**108** native page gel band of interest and sent for tandem MS/MS analysis. The raw MS/MS data were searched against 3 databases and 2 de novo sequencing algorithms. The result was visualized by Pepteideshaker software, and Cys597/Cys598 were marked as two **108** modified residues with 100% confidence. The result announced that covalent drug strategy could be successfully applied to Hsp90 as an anticancer drug target.

Overall, our research identified a few lead scaffolds as Hsp90 CTD inhibitors (**113**, **108**, **117**, disulfiram). More importantly, we experimentally proved that

the success story of targeted covalent drug strategy could be applied to Hsp90 target. It opens the door for Hsp90 covalent drug development. Identification of targetable cysteine residue Cys597/Cys598 not only contributes to covalent drug design but also revealed some information about novobiocin Hsp90 CTD binding pocket which remained unsolved due to lack of high-resolution crystal structure. Also, the discovery of dimeric Hsp90 inhibition will promote a better understanding of Hsp90 machinery and drug design. Since all our protein assays were completed using full-length protein, the results were more reliable compared with same assays be done using Hsp90 CTD protein only. There are more issues worth investigation in the future, and they are discussed on next page.

The most challenging part of this project is the validation of covalent modification. Although our data provided with some proof about Hsp90 covalent modification, however, some infill research works is needed for absolute confirmation about the covalent binding.

1. Literature has reported that the Hsp90 CTD 'switch point' sequence containing C597/C598 does not exist in fungi and bacteria. Sequence alignment between human Hsp90 and yeast Hsp90 showed the double CC residues were mutated into double AA (moreover, yeast Hsp90 does not have any cysteine at all). Such mutation decreases Hsp90 ATPase activity. As a result, our synthetic covalent drug **108** should exhibit precisely the same pattern as **109** if they are incubated with yeast Hsp90 for native page method. **108** is expected to lose covalent power against yeast Hsp90 in which does not consist of any cysteine residue. The comparison of native page method yeast Hsp90 and human Hsp90 will be a convincing indirect proof of **108** covalent modification power.
2. Previously, we characterized C597/C598 as active cysteine residues modified by covalent drug **108**. A straightforward way to validate this issue is to incubate **108** with C597A/C598A mutated type human Hsp90 and test to find various properties. Through this mutation, **108** is expected to lose its activity at 70 μM on native page method; therefore, proved the importance of cysteine modification.
3. Due to limited time and resources, MADLI-TOF analysis of **108**-Hsp90 complex was unsuccessful, mainly due to incomplete method development. Lots of experimental factors, including sample preparation ratio, incubation duration, purification method, etc. need to be optimized. Also, the concentration of protein must be increased to resist the loss of quantity during the desalting process. The importance of MALDI-TOF is not only a straightforward proof of covalent binding but also helps to identify how many inhibitor molecules will bind to the target, which would use as an accurate reference for proteomic

database search. Our data so far still cannot confirm whether there are more than 2 cysteine residues be modified.

Besides the aspect of covalent binding, our study provides no information about the mode of action. It worth to study this issue through various kind of bioassays. Including, ATPase assays for direct visualization of enzymic catalytic inhibition effect, immune-blot method (Western-Blot) to characterize up & down-regulated client proteins, cell cycle assay to investigate the disruption phase caused by drug administration, etc. Also, our previous research reported that acrylate type covalent modifiers could be detoxed through cellar metabolism. It worth to find out the mechanism and maps the metabolism pathway using a radioactive label tracing technique.

Another meaningful impact of our research is we opened the door for Hsp90 CTD covalent drug design. Acrylate type Michael acceptors in this project are the basic design of this type. Taking advantage of literature, there are more cysteine covalent warhead can be applied. There are few promising warhead optimizations can be done in the future.

1. **Acrylamides**

   As discussed in chapter **3.3.11**, all EGFR covalent inhibitors are acrylamide type due to the instability of acrylates. Our MTT tests also demonstrated the instability of acrylate. Acrylamides are less active then acrylate; therefore, they present a ‘slow-steady’ inhibition pattern in cell culture. Since we already verified the importance of cinnamic scaffold of which the hydrophobic benzene addition increases its stability. Based on this information, 4’-acrylamide novobiocin **142** and 4’-cinnamic-amide novobiocin **143** are promising analogues (Figure. 204).

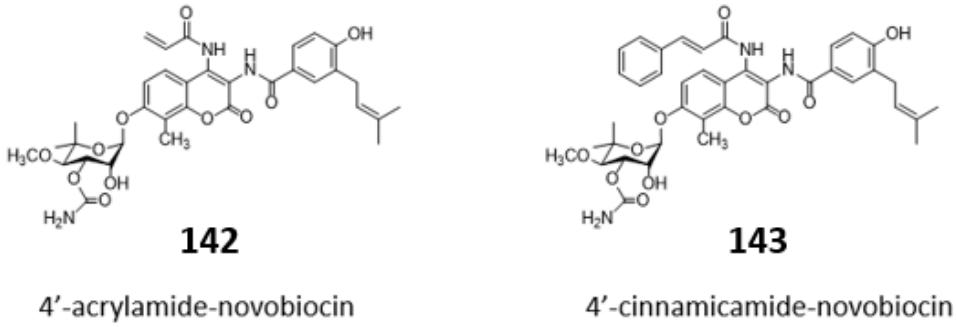


Figure 204: Acrylamide type proposed Hsp90 CTD covalent inhibitors.

2. Thiol and sulfides

In our search, we identified that sulfur-based structures are potential Hsp90 CTD inhibitors that are disrupting Hsp90 dimerization process. 4-thiol coumarin demonstrated the importance of thiol, especially when similar structure 4-hydroxy coumarin and 4-amino coumarin has no effect on Hsp90 dimerization. Disulfiram is also an Hsp90 CTD inhibitor, and its inhibitory power is dominated by its disulfide bond. Altogether, a potential covalent modification may happen as both thiol and disulfide bond are cysteine oxidation reagent. 4'-OH linker of novobiocin can be substituted with 4'-thiol, generating relevant analogues with potential better potency (Figure 205).

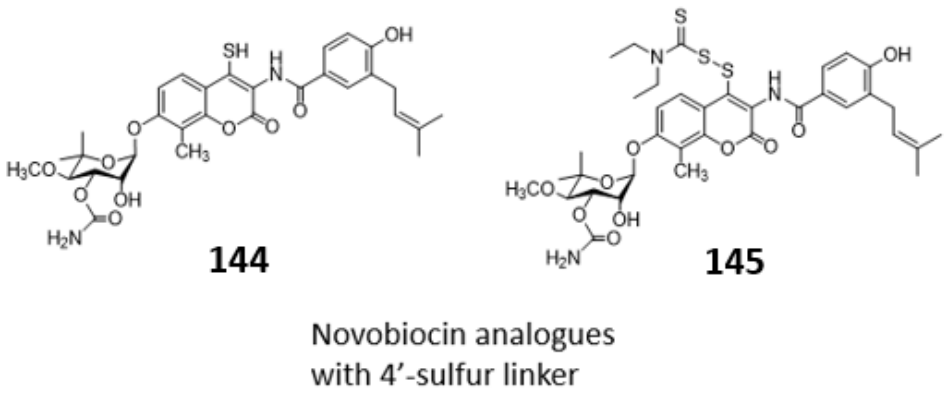


Figure 205: 4'-thiol-novobiocin analogues. **144** is 4'-thiol novobiocin and **145** is designed based on adding disulfiram moiety to novobiocin.

Optimization of non-covalent scaffold is also critical. According to literature, covalent inhibition always happens after non-covalent inhibition, which means, regardless of how perfectly designed the covalent warhead is, without a high-affinity non-covalent backbone carriage, the covalent binding will less likely to happen. Obviously, the original novobiocin scaffold only presents about 700 μM antiproliferative activity on average, and Kd of dimeric Hsp90 inhibition is even less reactive at mM level. Fortunately, novobiocin lead optimization work already be done by B.S Blagg, as reviewed in the introduction section. Without applying the covalent drug design at 4'-position, substitution of novobiocin amide side chain with 2-indole or benzyl-triazole moiety along with the substitution of 7'-noviose sugar with methylpiperdine will increase the non-covalent antiproliferative activity to 0.1 – 0.5 μM. Taking advantage of these SAR results, better covalent type scaffold can be designed (Figure. 206) and has the potential to increase their activity to nanomole level.

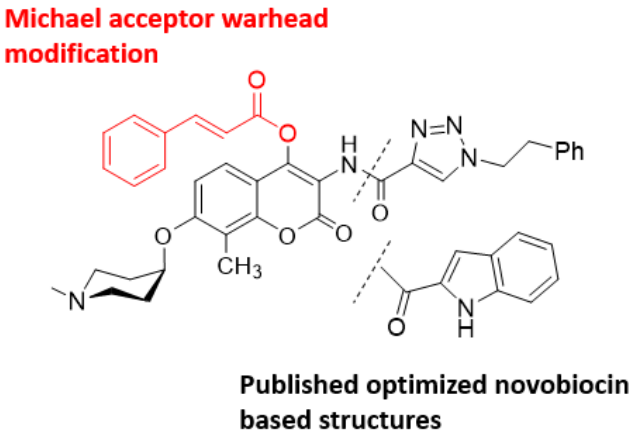
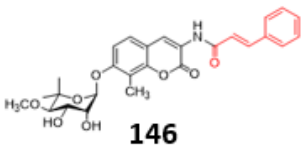


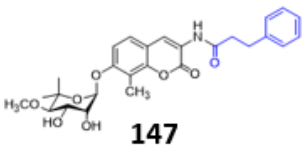
Figure 206: Antiproliferative activity optimized novobiocin backbone combined with covalent warhead may increase activity.

Another design worth future investigation is the position of the covalent warhead. In this project, we only evacuated 4'-Michael acceptors due to straight forward and time-saving synthesis. Nevertheless, the covalent warhead does not have to be assembled to 4'-hydroxy of novobiocin. Blagg in 2008 reported 3'-cinnamic-amide novobiocin analogue **146** exhibited much higher antiproliferative activity then its hydrolyzed form 3'-hydrocinnamic-amide analogue **147**, against several cancer cell lines (Figure. 207).



146

Cysteine reactive



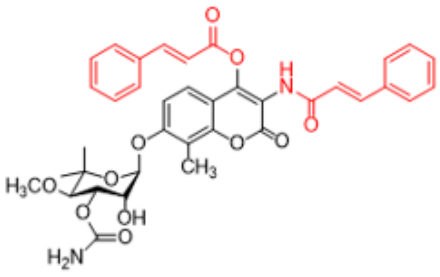
147

Cysteine unreactive

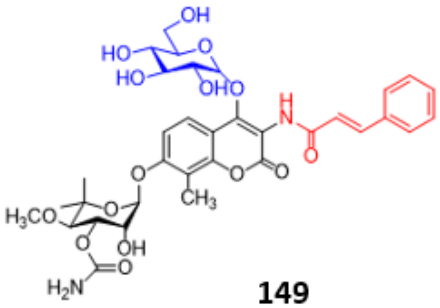
Cell lines	Activity (μM)	Activity (μM)
SkBr-3	2.6	10.2
PL-45	3.0	9.8
LNCaP	4.5	92.8
PC-3	3.9	22.0

Figure 207: Novobiocin 3'-amide covalent warhead modification may lead to covalent inhibition (Burlison, Avila et al. 2008).

The antiproliferative activity test showed the importance of cysteine reactive 3'-acrylamide moiety. Although date back to 2007, there was no concept of covalent drugs, nowadays, these data can be viewed as an indirect proof of covalent modification. Therefore, duo 3',4'- Michael acceptor substituted moiety (see **148**) can be promising, making the Michael addition happens with higher chance. Or, since covalent warhead can be assembled at 3' position, 4' position can be substituted with validated glucose moiety to increase non-covalent power (**149** in Figure. 208).



148



149

Figure 208: Designed novobiocin covalent inhibitors with duo Michael acceptors or with 3'-Michael acceptor-4'-glucose substitution.

Chapter 5: Experimental

5.1 General Methods

5.1.1 Chemical Synthesis

All reactions requiring the use of anhydrous conditions were carried out under an atmosphere of nitrogen, and all glassware was pre-dried in an oven 220 °C and cooled under nitrogen prior to use. Stirring was by internal magnetic follower unless otherwise stated. All reactions were followed by TLC, and organic phases extracted were dried with anhydrous magnesium sulphate.

Diethyl ether, tetrahydrofuran, dichloromethane, methanol, acetonitrile, triethylamine, dimethylformamide were purchased as anhydrous solvents from Sigma-Aldrich chemical company.

5.1.2 Thin-layer chromatography (TLC)

TLC analysis was carried out with Merck aluminium-backed plates coated with silica gel 60 F_{254}. Sampled TLC plates were visualized using either ultraviolet light (short-wave = 254 nm, long-wave = 365 nm) or ceric ammonium molybdate stain.

5.1.3 Purification

The flash column chromatography purification mentioned in this thesis was carried out by manually assembled flash column using silica gel manufactured by Sigma-Aldrich with pore size 60 A, 230-400 mesh particle size, 40-63 micrometres particle size, technical grade.

.

An automated reverse phase purification system Biotage was applied to purification of final step compounds. For this Biotage method, the pre-packed reverse phase column “Biotage® Sfär C18 Duo 100 Å 30 µm 120 g” (Product

code: FSUD-0401-0120) was selected and applied to the Biotage Isolera One machine. The solvent selected was methanol/water system. Solvents are analytical-pure degree. A typical run of Biotage requires the setting of the initial segment, gradient and final stage, which defines the percentage component of the solvent system and the duration of each run. MeOH/Water mobile phase was applied to all Biotage methods. There are 3 methods being developed, and details are listed below:

Biotage method 1	
Initial segment	0% Methanol for 3 CV
Gradient	To 100% Methanol for 10 CV
Final stage	100% Methanol for 3 CV

Biotage method 2	
Initial segment	30% Methanol for 3 CV
Gradient	To 100% Methanol for 14 CV
Final stage	100% Methanol for 3 CV

Biotage method 3	
Initial segment	30% Methanol for 3 CV
Gradient	To 100% Methanol for 16 CV
Final stage	100% Methanol for 3 CV

5.2 Analytical Methods

5.2.1 Nuclear magnetic resonance (NMR)

^{1}H NMR and ^{13}C NMR were recorded on either a Bruker AM400 spectrometer operation at 400 MHz for ^{1}H and 101 MHz for ^{13}C or a Bruker AM500 spectrometer operation at 500 MHz for proton and 126 MHz for carbon

Chemical shifts (δ_H and δ_C) are quoted as parts per million downfield from 0. The multiplicity of a ^{1}H NMR signal is designated by one of the following abbreviations: s = singlet, d = doublet, t = triplet, q = quartet, quin = quintet, sept = septet, br = broad and m = multiplet. Coupling constants (*J*) are expressed in Hertz.

5.2.2 Liquid chromatography-mass spectrometry (LC-MS)

LC-MS samples were prepared in 1 mL vial in MeOH (HPLC grade) solution. The analysis was carried out on a SHIMADZU LC-MS-2020. Two LC/MS methods were used in the system.

"Method A": (7 min run, UV: 254 nm + 280 nm)

MS range: 155-2000 m/z; Column: XTerra, C18 2.5 μm, 4.6 mm; LC: Solvent A 0.1% FA in Water; Solvent B 0.1 FA in Acetonitrile. Flow rate 0.05 mL/min

Gradient: Start at 10% B, 4 min 95%B, 5.2 min 10% B, and end at 7 min

"Method B": (12 min run, UV: 254 nm + 280 nm)

MS range: 155-2000 m/z; Column: XTerra, C18 2.5 μm, 4.6 mm; LC: Solvent A 0.1% FA in Water, Solvent B 0.1 FA in acetonitrile Flow rate 0.05 mL/min.

Gradient: Start at 0%B, 7 min 95% B, 10.2 min 10%B, end at 12 min.

High-resolution mass spectra were carried out at the School of Pharmacy. Mass spectra carried out at using either an Agilent Micromass Q-TOF premier Tandem Mass Spectrometer from Micromass utilising electrospray. All

samples were run under Electrospray ionization mode using 50% acetonitrile in water and 0.1% formic acid as solvent.

5.2.3 Proteomic mass spectrometry analysis.

The gel bands were subjected to in-gel tryptic digestion using a DigestPro automated digestion unit (Intavis Ltd.). The resulting peptides were fractionated using an Ultimate 3000 nano HPLC system in line with an LTQ-Orbitrap VelosPro mass spectrometer (Thermo Scientific). In brief, peptides in 1% (v/v) formic acid was injected onto an Acclaim PepMap C18 nano-trap column (Thermo Scientific). After washing with 0.5% (v/v) acetonitrile 0.1% (vol/vol) formic acid peptides were resolved on a 250 mm × 75 μm Acclaim PepMap C18 reverse-phase analytical column (Thermo Scientific) over an 80 min organic gradient, (1-50% solvent B over 55min., 50-90%B over 0.5 min., held at 90%B for 5.5 min and then reduced to 1%B over 0.5 min.) with a flow rate of 300 nl min−1. Solvent A was 0.1% formic acid, and Solvent B was aqueous 80% acetonitrile in 0.1% formic acid. Peptides were ionized by nano-electrospray ionization at 2.1 kV using a stainless-steel emitter with an internal diameter of 30 μm (Thermo Scientific) and a capillary temperature of 250°C. Tandem mass spectra were acquired using an LTQ- Orbitrap Velos mass spectrometer controlled by Xcalibur 2.1 software (Thermo Scientific) and operated in data-dependent acquisition mode. The Orbitrap was set to analyze the survey scans at 60,000 resolution (at m/z 400) in the mass range m/z 300 to 2000, and the top twenty multiply charged ions in each duty cycle selected for MS/MS in the LTQ linear ion trap. Charge state filtering, where unassigned precursor ions were not selected for fragmentation, and dynamic exclusion (repeat count, 1; repeat duration, the 30s; exclusion list size, 500) were used. Fragmentation conditions in the LTQ were as follows: normalized collision energy, 40%; activation q, 0.25; activation time 10ms; and minimum ion selection intensity, 500 counts.

The raw data files were processed and quantified using Proteome Discoverer software v1.4 (Thermo Scientific) and searched against the Uniprot Human

database (150786 sequences, downloaded October 2019) using the SEQUEST algorithm. Peptide precursor mass tolerance was set at 10ppm, and MS/MS tolerance was set at 0.8 Da. Search criteria included carbamidomethylation of cysteine (+57.0214) as a fixed modification and oxidation of methionine (+15.9949) as a variable modification. Searches were performed with full tryptic digestion, and a maximum of 2 missed cleavage events was allowed. The reverse database search option was enabled, and all peptide data were filtered to satisfy the false discovery rate (FDR) of 1%

5.2.4 Hsp90 binding assay – native polyacrylamide gel electrophoresis.

Qualitative assay

Hsp90 protein (1 mg/mL, 2.50 µg, Abcam, UK, product code ab48801 or lab produced recombinant Hsp90 protein) was incubated with relevant compounds. The total mixture volume was 10 µL. Suitable stock solutions were prepared so that the volume added of each of the inhibitors at different concentrations was 7.5 µL. The same volume of distilled water and DMSO was added to the control without inhibitor. Samples were incubated overnight at 37 °C before analysis by native polyacrylamide gel electrophoresis (8% polyacrylamide native page gel was manually prepared following the common standard procedure). When the reaction was complete, the native page gel was dyed by instant blue to reveal the test results. Any band of interest were carefully cut and sent for proteomic analysis.

5.2.5 Antiproliferative activity test.

L-glutamine, sodium pyruvate and fetal bovine serum were supplied from Gibco Life Technologies Inc (Paisley, UK). Dulbecco′s Modified Eagle′s Medium, dimethyl sulfoxide and MTT [3-(4,5-Dimethyl-2-thiazolyl)-2,5-diphenyl-2H-tetrazolium bromide] were purchased from Sigma Aldrich

(Gillingham, Dorset, UK). The anti-proliferative activities were performed using the MTT reduction method. Human cancer cell lines derived from ovarian (A2780, ref. no. ECACC 93112519), breast (MCF7, ref. no. ATCC HTB-22), brain cancer (U-87 MG, ref. no. ATCC HTB-14), pancreatic (MIA PaCa-2, ref. no. ATCC CRM-CRL-1420) and lung (A549, ref. no. ATCC CRM-CCL-185) were obtained from the European Collection of Cell Cultures (ECACC, Salisbury, UK). Briefly, human cancer cells were grown in Dulbecco′s Modified Eagle Medium (DMEM) contain 4500 mg/L glucose, 4 mM L-glutamine, 1 mM sodium pyruvate, and 1500 mg/L sodium bicarbonate supplemented with 10 % fetal bovine serum. All cells were grown in an incubator at 37 °C with 95% humidity and 5% CO_2. Cell lines were plated in a flat-bottom 96-wells plate at 1 x 10^5 cells per well density for 24 h prior to compound treatment.Then, the media was removed, and then incubated with fresh media containing 1 to 1 x10^{-8} mM of each compound (n = 5) for another 24 h. After this period, the MTT solution (0.5 mg/ml) was added to each well for 2 h at 37 °C. After which medium was removed, 200 µL dimethyl sulfoxide was added, and the plate was incubated under shaker stirring (KL-2, Edmund Bühler GmbH™, Fisher Scientific, UK) at 100 rpm for 20 min at 37 °C. The absorbance was measured using a SPECTROstar Omega (BMG Labtech Ltd., Aylesbury, UK) at 570 nm. Untreated and lysed cells (with Triton X-100, 10 mg/mL, 200 µL per well) were used as control (0% and 100 % cell death, respectively). The IC_{50} values were calculated by a dose-response analysis using the Origin 6.0 ® software taking into account the ratio between the measured absorbance values of treated and untreated cells.

5.2.6 Microscale thermophoresis (MST) test – Determination of ligand & Hsp90 binding kD.

MST tests were performed on a Nano temper monolith NT.115 microscale thermophoresis machine. The test capillaries used were Nano temper Monolith™ NT 115 series capillaries Cat# MO – K022. The target protein was pre-labelled by Monolith His-tag labelling kit RED-tris-NTA 2nd generation dye. Target Hsp90 protein was purchased from Abcam 48801.

Following procedure describes the labelling procedure for one experiment with 16 capillaries.

Preparation of 1 × PBS – T: 8.0 mL aq. Deionized water was added to the vials containing 5 × PBS – T and the 1 × PBS – T was prepared.

Preparation of 5 µM dye solution: The original pack of dye was suspended into 50 µL 1 × PBS – T and 5 µM dye was obtained.

These two solutions are stored under -20 °C and can be directly used by each batch of the test.

Preparation of labelled Abcam Hsp90 48801: 2 µL of 5 µM dye was mixed with 98 µL 1 × PBS – T to obtain a 100 nM dye solution. 2.5 µL of Hsp90 protein (1mg/mL Abcam UK 48801) was mixed with 138.4 µL PBS (common PBS solution) to obtain a 200 nM protein solution. Then 100 µL of 200 nM protein solution was mixed with 100 µL of 100 nM dye. The resulted solution was stored at room temperature in the dark for 24 h. The vial was centrifuged for 10 min at 15000 g. Then, the protein must be kept in 4 °C fridge for 3-5 days to stabilize the labelled complex. The labelled protein produced by such a procedure can be stored under -80 °C for any time to be used.

MST test can be carried out following the exact same procedure provided by the MST inbuilt software. However, when the preparation of 16 samples of dilution was complete, the samples were kept under 37 °C for 24 hours before it was dipped by MST capillaries prior to MST testing process.

5.2.7 Lab produced recombinant Hsp90 human alpha

The protein shares the same sequence as the one in Uniport database P07900 with an added His-tag at N-terminus. The sequence is given below:

MGSSHHHHHH SSGLVPRGSH MPEETQTQDQ PMEEEEVETF AFQAEIAQLM SLIINTFYSN KEIFLRELIS NSSDALDKIR YESLTDPSKL DSGKELHINL IPNKQDRTLT IVDTGIGMTK ADLINNLGTI AKSGTKAFME ALQAGADISM IGQFGVGFYS AYLVAEKVTV ITKHNDDEQY AWESSAGGSF TVRTDTGEPM GRGTKVILHL KEDQTEYLEE RRIKEIVKKH SQFIGYPITL FVEKERDKEV SDDEAEEKED KEEEKEKEEK ESEDKPEIED VGSDEEEEKK DGDKKKKKKI KEKYIDQEEL NKTKPIWTRN PDDITNEEYG EFYKSLTNDW EDHLAVKHFS VEGQLEFRAL LFVPRRAPFD LFENRKKKNN IKLYVRRVFI MDNCEELIPE YLNFIRGVVD SEDLPLNISR EMLQQSKILK VIRKNLVKKC LELFTELAED KENYKKFYEQ FSKNIKLGIH EDSQNRKKLS ELLRYYTSAS GDEMVSLKDY CTRMKENQKH IYYITGETKD QVANSAFVER LRKHGLEVIY MIEPIDEYCV QQLKEFEGKT LVSVTKEGLE LPEDEEEKKK QEEKKTKFEN LCKIMKDILE KKVEKVVVSN RLVTSPCCIV TSTYGWTANM ERIMKAQALR DNSTMGYMAA KKHLEINPDH SIIETLRQKA EADKNDKSVK DLVILLYETA LLSSGFSLEDPQTHANRIYR MIKLGLGIDE DDPTADDTSA AVTEEMPPLE GDDDTSRMEE VD

5.2.8 Preparation of E. *coli* culture

Frozen recombinant human Hsp90 *E. coli* stock was defrosted on ice and then added into LB medium (25 mL) containing of ampicillin (50 µg/ml). This media was incubated in the shaker set at 200 rpm. And 37 °C overnight, after which the media was transferred into LB media (1 L) also containing ampicillin (50 µg/ml). The media was cultured in the shaker at the same conditions for 4 hours, and OD_{600} (optical density at 600 nm wavelength UV) was monitored. When OD_{600} value reached the range of 0.4 - 0.6 Au, the temperature was set to 20 °C for 1 h. Isopropyl β- d-1-thiogalactopyranoside (IPTG, 1 M, 1 mL; final concentration: 1 mM) was added into the media. This media was kept in the shaker overnight at room temperature. The cells were harvested by centrifuging at 3850 g aT 4 °C for 15 minutes.

5.2.9 Protein expression

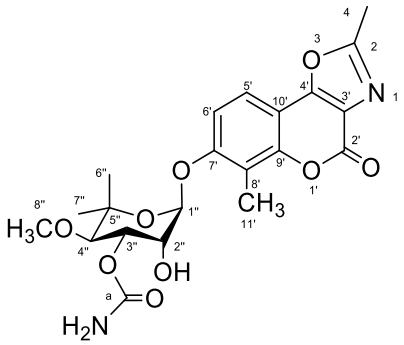
E. coli cells (1 g) were suspended in cold PBS buffer (NaCl 8.0 g/L, KCl 0.2 g/L, Na_2HPO_4 1.42 g/L and KH_2PO_4 0.24 g/L, pH 7.4,10 mL). Cells in solution were then sonicated for 90 seconds with a 30-seconds pause every 30 seconds. This procedure was repeated for another two times to make sure sonication was complete. During the sonication, the tube that contained cells was kept on ice. The supernatant was isolated by the centrifugation at 10770 g and 4 °C for 1 hour.

5.2.10 Protein purification using an affinity column

HisTrap HP column (GE Healthcare, 1 ml) was prewashed with deionized water (5 CV) and equilibrated with at least binding buffer (20 mM sodium phosphate, 500 mM NaCl, 20 mM imidazole, pH 7.4, 5 CV). The supernatant from cell lysis was slowly loaded onto the column. Binding buffer (10 CV) was eluted to wash off unbound proteins, after which Hsp90 complexes were eluted with about elution buffer (20 mM sodium phosphate, 500 mM NaCl, 500 mM imidazole, pH 7.4; 2.5 mL). Fractions obtained that contained Hsp90 protein were combined and preserved in PBS buffer (NaCl 8.0 g/L, KCl 0.2 g/L, Na_2HPO_4 1.42 g/L and KH_2PO_4 0.24 g/L, pH 7.4, 0.1 mL).

5.3 Synthetic procedure of target compounds

Synthesis of (3R,4S,5R,6R)-6-((2,6-dimethyl-4-oxo-4,9 -dihydro-3a*H*-chromeno [3,4-*d*- oxazol-7-yl)oxy)-5-hydroxy-3-methoxy -2,2-dimethyltetrahydro-2H-pyran-4-yl carbamate 36



A mixture of pyridine (200 mL), acetic anhydride (40 mL) and novobiocin sodium salt (20.0 g, 0.0315 mol) was added into a 1000 mL r.b flask, stirred and heated under reflux for 4 hours. When the system was thoroughly cooled to room temperature, 5N-HCl was added to it dropwise until the pH value of the system reached 1 (monitored by pH indicator paper).

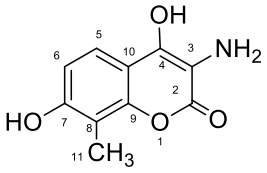
Then the reaction mixture was kept below 25 °C inside the ice bath until the brown syrup precipitated. The aqueous phase which above the syrup was removed and the mixture were washed with a small amount of diethyl ether to afford a crude grey product. The crude product could either be washed further with 20 mL of diethyl ether to afford a light grey powder or washed by water and ethyl acetate to afford a white powder. (9.5 g, 69%): Rf =0.13, petroleum ether/ethyl acetic 1:2 v/v. ^{1}H NMR (400 MHz, CD_3OD): δ = 7.51 (1H, d, *J* = 8.7 Hz, H5’), 7.12 (1H, d, *J* = 8.8 Hz, H6’), 5.79 (1H, d, *J* = 2.3 Hz, H1’’), 5.47 (1H, dd, *J* = 3.3, 9.9 Hz, H3’’), 5.42 (1H, dd, *J* = 2.3, 3.0 Hz, H2’’), 3.49 (3H, s, H), 3.40 (1H, dd, *J* = 3.4, 7.0 Hz, H4’’), 2.59 (3 H,s, H11’), 2.33 (3H, s, H4’), 2.13 (3H, s, H8’), 1.46 (3H, s, H6’’), 1.32 (3H, s, H7’’). ^{13}C NMR (126 MHz, CD_3OD): δ = 171.5 (C2’), 165.3 (C2), 158.7 (C9’’), 158.5 (C7’’), 153.4 (C9’), 123.5 (C10’), 120.5 (C8’), 116.7 (C5’), 112.4 (C1’’), 107.1 (C6’),

97.4 (C4'), 80.2 (C3'), 72.0 (C5''), 70.6 (C2''), 62.0 (C3''), 54.8 (C8''), 23.3 (C6''), 20.7 (C7''), 13.9 (C4), 9.1 (C11').

ESI-MS calculated for 447.17 found 447.17 ($M + H^+$)

This is a literature compound and the data is in accordance with literature (Simon, Huang et al. 2017).

Synthesis of 3-amino-4,7-dihydroxy-8-methylcoumarin 37

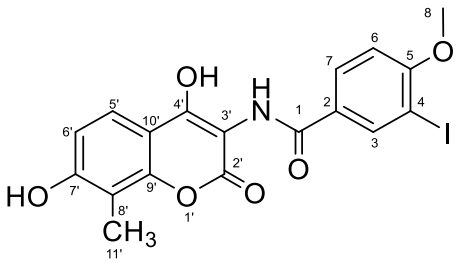


A mixture of compound **36** (5.50 g, 0.0120 mol), anhydrous methanol (75 mL) and 10% HCl/methanol (150 mL) was dissolved in a 500mL r.b. flask, stirred and heated to reflux for 2 hours. When a clear black solution appeared, the mixture was evaporated in the vacuum until precipitation started. Then the mixture was stored in 4°C fridge overnight, and the precipitation was filtered and washed with cold methanol to afford a pale-yellow powder. The filtrate was evaporated again, and the crystallizing procedure was repeated twice to afford more pale-yellow powder products. (2.50 g, 99%). Rf = 0.15, chloroform: methanol 4:1 v/v. ^{1}H NMR (400 MHz, MeOD): δ = 7.77 (1 H, d, *J* = 8.29 Hz, H5), 6.95 (1 H, d, *J* = 8.29 Hz, H6), 2.28 (3 H, s, H11). ^{13}C NMR (126 MHz, $MeOD_4$): δ = 164.20 (C2), 162.39 (C7), 161.79 (C4), 153.69 (C9), 122.64 (C5), 113.73 (C6), 113.51 (C8), 108.15 (C10), 96.27 (C3), 9.21 (C11).

ESI-MS calculated for 230.18 found 230.18 ($M + Na^+$)

This is a literature compound and the data is in accordance with literature (Simon, Huang et al. 2017).

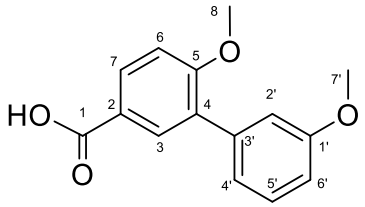
Synthesis of *N*-(4,7-dihydroxy-8-methyl-2-oxo-2H-chromen-3-yl)-3-iodo-4-methoxybenzamide 38



Compound 3-amino-4,7-dihydroxy-8-methylcoumarin (**37**, 20 mg, 0.096 mmol) was added into a 100 mL r.b flask that already contained a mixture of anhydrous dichloromethane and pyridine (1:1, 10 mL). Then N-(3-dimethylaminopropyl)-N'- ethylcarbodiimide hydrochloride (EDCI) (15 mg, 0.096 mmol) and 3-Iodo-4-methoxy-benzoic acid (26.80 mg, 0.0096 mmol) were weighed and dissolved in a mixture anhydrous dichloromethane and pyridine (1:1, 5 mL) inside a sealed conical flask. The solution in the conical flask was absorbed by a syringe and be injected drop by drop to the r.b. flask in which 3-amino-4,7-dihydroxy-8-methylcoumarin was contained and stirred under the protection of nitrogen gas. After 14 h the solvent was removed under vacuum and then 100 mL DCM was added to the dark-brown residue, resulted in a dark-brown suspension. After filter through a funnel, the solid inside the suspension was removed and the rest of the solution was concentrated under vacuum to afford a dark-brown solid. The crude product was then purified by flash column using 3:1 ethyl acetate: hexane system to afford a yellow solid (12.50 mg 31%). ^{1}H NMR (400 MHz, DMSO): δ = 10.38 (1H, s, 4-OH), 9.34 (1H, s, 7-OH), 8.43 (1H, s, H3), 8.02 (1H, d, *J* = 9.35 Hz, H7), 7.57 (1H, d, *J* = 8.46 Hz, H5'), 7.12 (1H, d, *J* = 8.61 Hz, H6), 6.87 (1H, d, *J* = 8.76 Hz, H6'), 3.92 (3H, s, H8), 2.17 (3H, s, H11'). ^{13}C NMR (126 MHz, DMSO): δ = 165.34 (C5), 158.87 (C2'), 158.15 (C7'), 157.22 (C9'), 152.14 (C4'), 135.92 (C4) 128.93 (C7), 127.39 (C2), 127.55 (C5'), 112.34 (C6), 110.97 (C6'), 97.09 (C3'), 86.00 (C4), 55.64 (C8), 8.29 (C11').

HRMS calculated for 489.9764 found 489.9755 ($M + Na^+$)

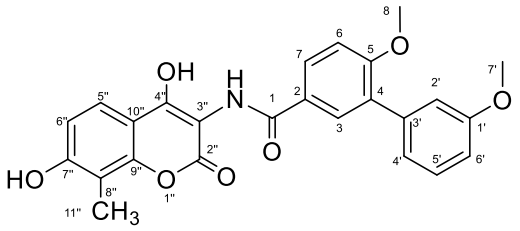
Synthesis of 3',6-dimethoxy-3-biphenylcarboxylic acid 41



3-Methoxyphenylboronic acid **39** (1.00 g, 6.579 mmol) and 3-Indo-4-methoxy benzoic acid **40** (1.82 g 6.579 mmol) were dissolved in a mixture of 10mL 2M K_2CO_3 aq. And 15 mL 1,4-dioxane. The whole mixture was stirred for 30 mins until 1,1-bis (diphenylphosphine) ferrocene]-dichloropalladium (II) (158 mg 3% eqv.) was added. After the solution is stirred at room temperature for 2 days, the solvent inside was removed under vacuum to afford a black syrup. 40mL acetone was added to the syrup, and this black solution was continuously passed through Celite 577 fine and a thin silica pad. Drops of HCl/methanol were added into the black solution until its pH reaches 2. Then the solution was dried under vacuum, and the crude product was purified by flash column using 3:7 acetone/DCM to afford a brown solid (432 mg 45%). ^{1}H NMR (400 MHz, $CDCl_3$): δ = 8.13 (1H, d, *J* = 8.21 Hz, H7), 8.11 (1H, s, H3), 7.36 (1H, t, *J* = 7.65, H5'), 7.13 (1H, d, *J* = 8.21Hz, H6), 7.10 (1H, s, H2'),7.0 (1H, d, *J* = 8.21 Hz, H6'), 6.93 (1H, d, *J* = 8.30, H4'), 3.92 (3H, s, H8), 3.87 (3H, s, H7'). ^{13}C NMR (126 MHz, $CDCl_3$): δ = 170.61 (**C**OOH), 160.86 (C5), 159.31 (C1'), 138.1 (C3'), 133.03 (C7), 131.28 (C3), 130.93 (C5'), 129.1 (C2), 122.01(C4), 121.56 (C4'), 115.29 (C5), 112.98 (C2'), 110.69 (C6'), 55.86 (C8), 55.32 (C7').

HRMS calculated for 281.0790 found 281.0788 ($M + Na^+$)

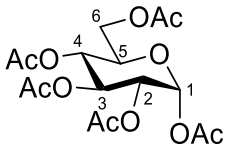
Synthesis of 6,3'-dimethoxy-biphenyl-3-carboxylic acid (4,7-dihydroxy-8-methyl-2-oxo-2H-chromen-3-yl)-amide 42



Compound 3-amino-4,7-dihydroxy-8-methylcoumarin (**37**, 200 mg, 0.96 mmol) was added into a 100 mL r.b flask that already contained a mixture of anhydrous dichloromethane and pyridine (1:1, 15 mL). Then EDCI (172 mg, 0.96 mmol) and 3',6-dimethoxy-3-biphenylcarboxylic acid (**41**, 248 mg, 0.96 mmol) was weighed and dissolved in a mixture anhydrous dichloromethane and pyridine (1:1, 10 mL) inside a sealed conical flask. The solution in the conical flask was absorbed by a syringe and be injected drop by drop to the r.b flask in which 3-amino-4,7-dihydroxy-8-methylcoumarin was contained and stirred under the protection of nitrogen gas. After 6 hours the solvent was removed under vacuum and then 100 mL DCM was added to the dark-brown residue, resulted in a dark-brown suspension. After filter through a funnel, the solid inside the suspension was removed, and the rest of the solution was concentrated under vacuum to afford a dark-brown solid. The crude product was then purified by flash column using 2:1 ethyl acetate: hexane system to afford a yellow solid. The yellow solid was then stirred in a methanol-K_2CO_3 solution for 1 hour to afford the product (115 mg 27%). ^{1}H NMR (400 MHz, DMSO): δ = 8.01 (1H, d, *J* = 11.90Hz, H7), 7.99 (1H, s, H3), 7.55 (1H, d, *J* = 8.55 Hz, H5''), 7.33 (1H, t, *J* = 8.19, H5'), 7.15 (1H, d, *J* = 8.38 Hz, H6), 7.11 (1H, d, *J* = 8.38 Hz, H4'), 7.09 (1H, s, H2'), 6.91 (1H, d, *J* = 8.19 Hz, H6'), 6.72 (1H, d, *J* = 7.98 Hz, 6''), 3.83 (3H, s, H8), 3.79 (3H, s, H7'),2.12 (1H, s, H11''). ^{13}C NMR (126 MHz, DMSO): δ = 165.34 (C5), 162.51 (C1'), 158.87 (C2''), 158.13 (C7''), 157.22 (C9''), 152.14 (C4''), 139.03 (C3'), 130.28 (C5'), 128.93 (C7), 128.79 (C3), 127.39 (C2), 122.55 (C5''), 121.74 (C4'), 115.24 (C2'), 115.14 (C6'), 112.34 (C6), 110.9 (C6''), 97.09 (C3''), 55.72 (C8), 55.04 (C7'), 8 .29 (C11'').

HRMS calculated for 470.1216 found 470.1238 ($M + Na^+$)

Synthesis of 1,2,3,4,6-penta-*O*-acetyl-α-*D*-glucose 55

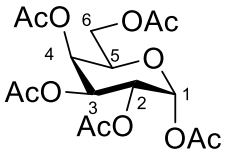


Acetic anhydride (20 mL) was added to a solution of α-*D*-glucose (2.50 g, 0.014 mol) in anhydrous pyridine (30 mL). The resulting mixture was stirred overnight at room temperature. Once the reaction was complete, the pyridine was reduced under vacuum. The resulting viscous residue was taken up in chloroform (50 mL) and the organic solution washed successively with 5 N HCl (20 ml), saturated $NaHCO_3$ (2 x 10 mL) and water (3 x 20 mL). The organic phase was dried over magnesium sulphate and filtered. The rest of solvent was removed under reduced pressure to afford the title compound **55** as a colourless solid (4.87 g, 90%); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 6.36 (1 H, d, *J* = 3.7 Hz, H1), 5.49 (1 H, t, *J* = 10.0 Hz, H3), 5.16 (1 H, t, *J* = 10.0 Hz, H4), 5.12 (1 H, dd, *J* = 3.3, 10.0 Hz, H2), 4.29 (1 H, dd, *J* = 4.2, 12.9 Hz, H6), 4.13-4.17 (1 H, m, H5), 4.12 (1 H, dd, *J* = 2.9, 10.0 Hz, H6), 2.20 (3 H, s, $OCOCH_3$), 2.11 (3 H, s, $OCOCH_3$), 2.05 (9 H, s, $OCOCH_3$). ^{13}C NMR (126 MHz, $CDCl_3$): δ = 169.65 (5 × -$OCOCH_3$), 89.10 (C1), 69.88 (C5), 69.82 (C3), 69.25 (C2), 67.97 (C4), 61.20 (C6), 20.43 (5 × -$OCOCH_3$).

ESI-MS: calculated for $C_{16}H_{22}O_{11}Na^+$, 413.12; found 413.12 ($M + Na^+$)

This is a literature compound and the data is in accordance with literature (Patel, Fuente et al. 2011).

Synthesis of 1, 2,3,4,6-penta-*O*-acetyl-*D*-galactose 56

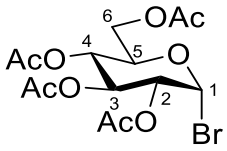


Acetic anhydride (20 mL) was added to a solution of α-*D*-galactose (2.50 g, 0.014 mol) in anhydrous pyridine (30 mL). The resulting mixture was stirred overnight at room temperature. Once the reaction was complete, the pyridine was reduced under vacuum. The resulting viscous residue was taken up in chloroform (50 mL) and the organic solution washed successively with 5N HCl (20 mL), saturated $NaHCO_3$ (2 x 10 mL) and water (3 x 20 mL). The organic phase was dried over magnesium sulphate and filtered. The rest of solvent was removed under reduced pressure to afford the title compound **56** as a colourless solid (4.87 g, 90%);^{1}H NMR (400 MHz, $CDCl_3$) δ=6.34 (1H, d *J* = 1.6 Hz, H1), 5.46 (1H ,dd, *J* = 1.2 Hz, H4), 5.33-5.27 (2H, m, H2, H3), 4.31 (1H, d, *J* = 6.6 Hz, H5), 4.15-4.00 (2H, m, H2 and H6), 2.12, 2.08, 2.00, 1.98, 1.96, (15H, CH_3CO × 5). ^{13}C NMR (126 MHz, $CDCl_3$): δ = 169.77 (5 × -$OCOCH_3$), 89.50 (C1), 69.94 (C5), 69.84 (C3), 69.12 (C2), 68.00 (C4), 61.19 (C6), 20.62 (5 × -$OCOCH_3$).

ESI-MS: calculated for $C_{16}H_{22}O_{11}Na^+$, 413.12; found 413.12 (M + Na^+)

This is a literature compound and the data is in accordance with literature (Patel, Fuente et al. 2011).

Synthesis of α-1-Bromine-2,3,4,6-tetra-*O*-acetyl-*D*-glucose 57

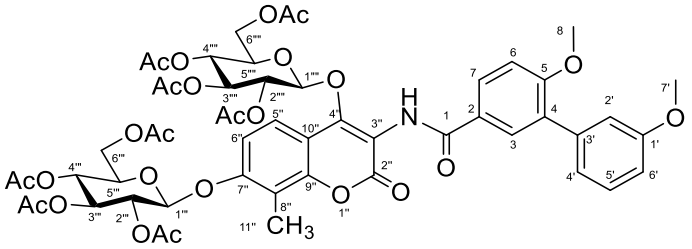


1,2,3,4,6-penta-*O*-acetyl-α-*D*-glucose (**55**, 2.00g, 5.12 mmol) was dissolved in anhydrous dichloromethane (DCM, 5 mL) under nitrogen. Hydrogen bromide solution (HBr, 10 mL, 33% wt. in acetic acid, 20 mL) was added into the system and stirred for 2 hours. The flask was dipped into an ice bath before the system was added with chloroform (100 mL) and cold deionized water (50 mL). The system was left stirring under 0 °C for 15 mins. The organic layer inside was washed with saturated $NaHCO_3$ (30 mL × 2) and deionized water (50 mL). The organic layer was dried over sodium sulphate and removed by vacuum to afford a white solid (1.81 g, 86%). ^{1}H NMR (400 MHz, $CDCl_3$): δ = 6.63 (1 H, d, *J* = 4.1 Hz, H1), 5.58 (1 H, t, *J* = 9.6 Hz, H3), 5.18 (1 H, t, *J* = 9.8 Hz, H4), 4.86 (1 H, dd, *J* = 4.0, 10.0 Hz, H2), 4.35 (1 H, dd, J = 4.2, 12.5 Hz, H6), 4.32 (1 H, m, H5), 4.15 (1 H, dd, *J* = 2.0, 12.5 Hz, H7), 2.10 (12 H, s, -OCOCH$_3$). ^{13}C NMR (126 MHz, $CDCl_3$): δ = 169.93 (5 × -OCOCH$_3$), 86.61 (C1), 72.20 (C5), 70.66 (C2), 70.29 (C3), 67.25 (C4), 61.00 (C6), 20.61 (5 ×-OCOCH$_3$).

ESI-MS: calculated for $C_{14}H_{19}BrO_9Na^+$, 433.02; found 433.02 (M + Na^+)

This is a literature compound and the data is in accordance with literature (Patel, Fuente et al. 2011).

Synthesis of β-4',7'-*O*-dia(2,3,4,6-tetra-acetyl glucosyl) 3N-(4,7-dihydroxy-8-methylcoumarin)-1H-6,3'-dimethoxy-biphenyl-3-carboxamide 59



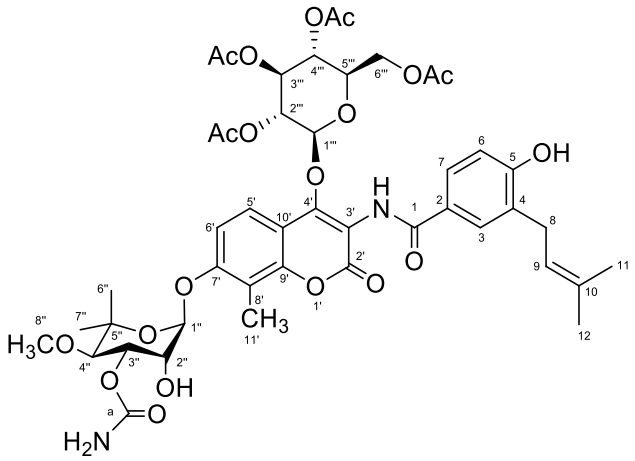
6,3'-Dimethoxy-biphenyl-3-carboxylic acid (4,7-dihydroxy-8-methyl -2-oxo-2H-chromen-3-yl)-amide (**42** 100 mg, 0.286 mmol) was dissolved in a 100 mL beaker that already contained 30 mL 0.15 M potassium carbonate water solution. α-1-Bromine-2,3,4,6-tetra-*O*-acetyl-*D*-glucose (**57**, 950 mg, 2.29mol) and benzyltributylammonium bromide (100 mg, 0.289 mmol) were dissolved in 20 mL chloroform inside a 250 mL r.b flask. When the stirring of compound **42** was complete, the yellow turbid liquid inside the beaker was transferred into the r.b flask and 20 mL of 0.15 M potassium carbonate water solution was used to wash the beaker. Once the separation of two layers in the r.b flask was complete, the whole system was put into 45°C oil bath and stirred for 9 hours. After this period, the water phase inside was removed, and the organic phase was washed by sat. $NaHCO_3$ (20 mL × 1), brine (20 mL × 1) and deionized water (20 mL × 2). The brown organic phase was then being concentrated under vacuum to obtain a solid brown residue. The residue was purified by flash column (ethyl acetate: hexane 3:2 R_f = 0.24) and preparative HPLC to afford a white solid (64 mg 22%). ^{1}H NMR (500 MHz, $CDCl_3$): δ = 8.03 (1H, d, *J* = 8.82 Hz, H7), 7.96 (1H, s, H3), 7.66 (1H, d, *J* = 9.01 Hz, H5''), 7.59 (1H, s, NH), 7.37 (1H, t, *J* = 8.01, H5'), 7.13 (2H, m, H6, H4'), 7.07 (1H, s, H2'), 7.03 (1H, d, *J* = 9.02 Hz, H6''), 6.94 (1H, d, *J* = 7.51 Hz, H6'), 5.46 (1H, d, *J* = 7.75 Hz, H1''''), 5.37 (3H, m, H2''', H2'''', H3''''), 5.23 (2H, m, H3''', H4''''), 5.14 (1H, d, *J* = 7.51 Hz, H1'''), 5.08 (1H, t, *J* = 9.82 Hz, H4''''), 4.35 (1H, q, *J* = 5.43, H6'''), 4.33 (1H, d, 12.7 Hz, H6'''), 4.09 (1H, q, *J* = 5.57 Hz, H6''''), 3.95

(5H, m, H7, H6’’’, H5’’’), 3.87 (3H, s, H7’), 3.70 (1H, m, H5’’’), 2.27 (1H, s, H11’’), 2.19-1.93 (24H, × 8 -OAc). ^{13}C NMR (126 MHz, $CDCl_3$): δ = 170.23, 70.17, 170.04, 169.82, 169.60, 169.41, 169.34, 169.19 (× 8 -C=O), 166.62 (C5), 161.24 (C1’), 160.10 (C2’’), 159.38 (C7’’), 157.69 (C9’’), 155.81 (C4’’), 138.63 (C3’), 130.24 (C5’), 129.21 (C7), 129.13 (C3), 127.39 (C2), 122.13 (C5’’), 121.89 (C4’), 115.83 (C2’), 115.55 (C6’), 112.75 (C6), 111.50 (C6’’), 111.27 (C9’’), 99.06 (C1’’’’), 98.03 (C1’’’), 72.53 (C5’’’’), 72.46 (C5’’’), 72.33 (C2’’’’), 72.15 (C2’’’), 71.01 (C3’’’’), 70.86 (C3’’’), 68.16 (C4’’’’), 68.09 (C4’’’), 61.82 (C6’’’’), 61.75 (C6’’’), 55.96 (C8), 55.30 (C7’), 20.43-20.75 (× 8 -CO-$\underline{C}H_3$), 8.39 (C11’’).

LC/MS under Method B, LC Rt = 4.37 min. ES^- (m/z) calculated 1106, found at [M - H^+] 1106.

HRMS calculated for 1130.3117 found 1130.3144 (M + Na^+)

Synthesis of β-4'-*O*-(2,3,4,6-tetra-acetyl glucosyl) novobiocin 64



Method 1: Mercuric Cyanide promoted Glycosylation

A mixture of novobiocin sodium (2.000 g, 3.159 mmol), 1-bromine-2,3,4,6-tetra-*O*-acetyl-*D*-glucose (**57,** 1.323g, 3.218 mmol), activated 4 Å molecular sieves (approx. 6.0 g) and anhydrous tetrahydrofuran (100 mL) were added into a 250 mL under nitrogen. Mercury (II) cyanide (0.9010 g, 3.575 mmol) was added into the system, and then the reaction mixture was stirred for 7 days under nitrogen gas protection at room temperature in the dark. The reaction mixture was then carefully filtered with Celite® 577 to remove any solid material. The filtrate was successively washed with 2 M potassium iodide (25 mL × 2), saturated $NaHCO_3$ (10 mL × 2) and dried over with sodium sulphate. The excessive solvent remained in the organic phase was reduced under vacuum to afford a light brown crude product, which was purified by flash column chromatography (ethyl acetate: petroleum ether: methanol= 12: 4: 1, v/v/v, R_f = 0.3) to afford a white powder (0.76 g, 26%).

Method 2: DMF solvent catalyzed S_N2 attack

Novobiocin Sodium (500 mg 0.315 mmol) and 1-bromine-2,3,4,6-tetra-*O*-acetyl-*D*-glucose **57** (195 mg 0.475 mmol) and few 4 Å molecular sieves were added into a 5 mL anhydrous DMF solution. After being stirred for 12 hours, the solution was filtered through Celite® 577 fine and then directly loaded to Biotage method 1 for purification. The product came out at around 9 CV, in a form of white solid (50 mg 7%).

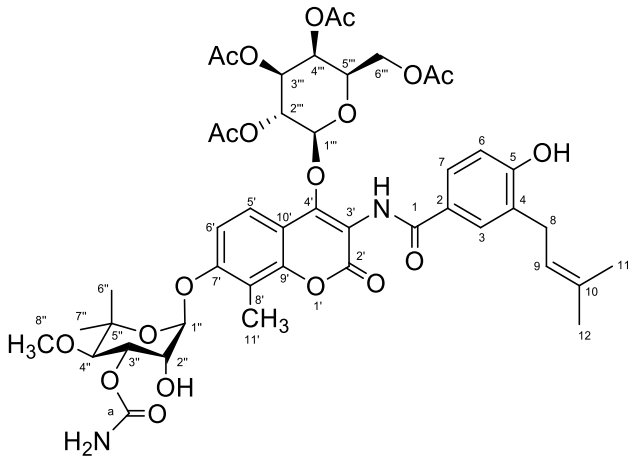
^{1}H NMR (500 MHz, CD_3OD): δ = 7.81 (1H, d, *J* = 2.0 Hz, H3), 7.77 (1H, dd, *J* = 2.4, 8.4 Hz, H7), 7.71 (1H, d, *J* = 9.6 Hz, H5'), 7.27 (1H, d, *J* = 9.2 Hz, H6'), 6.90 (1H, d, *J* = 8.4 Hz, H6), 5.63 (1H, d, *J* = 3.2 Hz, H1''), 5.60 (1H, d, *J* = 7.6 Hz, H1'''), 5.39 (1H, m, H3''), 5.36 (1H, m, H9), 5.30 (1H, t, *J* = 8.0 Hz, H3'''), 5.26 (1H, t, *J* = 9.2 Hz, H2'''), 5.12 (1H, t, *J* = 9.2 Hz, H4'''), 4.27 (1H, t, *J* = 2.5 Hz, H2''), 4.08 (1H, dd, *J* = 4.1, 13.2 Hz, H6'''), 3.78 (1H, dd, *J* = 2.4, 12.4 Hz, H6'''), 3.63 (1H, m, H5'''), 3.61 (1H, m, H4''), 3.58 (3H, s, H8''), 3.33 (2H, m, H8), 2.34 (3H, s, H11'), 1.99 (3H, s, CH_3CO), 1.98 (3H, s, CH_3CO), 1.92 (3H, s, CH_3CO), 1.89 (3H, s, CH_3CO), 1.78 (6H, s, H11, H12), 1.32 (3H, s, H6''), 1.17 (3H, s, H7''). ^{13}C-NMR (126 MHz, CD_3OD): δ = 169.64-170.53 (4×CH_3CO), 168.64 (Ca), 161.48 (C2′), 159.66 (C4′), 159.35 (C5), 158.09 (C7′), 157.71 (C1), 150.82 (C9′), 132.34 (C10), 129.43 (C3), 128.60 (C4), 127.18 (C7), 123.70 (C2), 122.13 (C5′), 121.95 (C9), 114.27 (C6), 113.89 (C8′), 110.77 (C10′), 110.45 (C6′), 107.53 (C3′), 99.82 (C1‴), 98.55 (C1″), 81.24 (C4″), 78.74 (C5″), 72.50 (C5‴), 72.16 (C2‴), 71.61 (C3‴), 71.34 (C3″), 69.42 (C2″), 67.67 (C4‴), 60.57 (C8″), 60.18 (C6‴), 27.88 (C8), 27.66 (C6″), 24.68 (C11), 21.85 (C7″), 19.16 (4×CH_3CO), 16.66 (C12), 7.30 (C11′). HMBC indicated that H1‴ is coupled to C4′ not C5.

LC/MS under Method A, LC Rt= 6.12 min, ES^- (*m/z*) calculated 941, found at [m - H^+] 941. LC/MS under Method B, LC Rt= 3.76 min, ES^- (*m/z*) calculated 941, found at [m - H^+] 941.

HRMS calculated for 943.3324 found 943.3327 (M + H^+)

This is a literature compound and the data is in accordance with literature (Patel, Fuente et al. 2011).

Synthesis of β-4'-*O*-(2,3,4,6-tetra-acetyl galactosyl) novobiocin 65



A mixture of novobiocin sodium (2.00 g, 3.16 mmol), 1-bromine-2,3,4,6-tetra-*O*-acetyl-D-galactose (**58**, 1.32 g, 3.22 mmol), activated 4 Å molecular sieves (approx 6.0 g) and anhydrous tetrahydrofuran (100 mL) were added into a 250mL under nitrogen. Mercury (II) cyanide (0.901 g, 3.58 mmol) was added into the system, and then the reaction mixture was stirred for 7 days at room temperature in the dark. The reaction mixture was then carefully filtered through Celite® 577. The filtrate was successively washed with 2M potassium iodide (25 mL × 2), saturated $NaHCO_3$ (10 mL × 2) and dried over with sodium sulphate. The excessive solvent remained in the organic phase was removed under vacuum to afford a light brown crude product. The crude product was purified by flash column chromatography (ethyl acetate: petroleum ether: methanol= 12: 4: 1, v/v/v, Rf = 0.3) to afford a white powder (1.26 g, 43%). ^{1}H NMR (500 MHz, CD_3OD): δ = 7.68 (1H, d, *J* = 1.8 Hz, H3), 7.65 (1H, dd, *J* = 1.4, 7.8 Hz, H7), 7.55 (1H, d, *J* = 9.6 Hz, H5′), 7.13 (1H, d, *J* = 8.7 Hz, H6′), 6.76 (1H, d, *J* = 8.6 Hz, H6), 5.50 (1H, d, *J* = 2.5 Hz, H1″), 5.45 (1H, d, *J* = 8.1 Hz, H1‴), 5.33 (1H, dd, *J* = 7.9, 10.3 Hz, H2‴), 5.24 (1H, m, H3″), 5.21 (1H, m, H4‴), 5.22 (1H, m, H9), 5.01 (1H, dd, *J* = 3.3, 10.3 Hz, H3‴), 4.14 (1H, t, *J* = 2.5 Hz, H2″), 3.89 (1H, dd, *J* = 7.6, 10.3 Hz, H6‴), 3.72 (1H, m, H5‴), 3.56 (1H, *J* = 5.1, 10.5 Hz, H6‴), 3.47 (1H, m, H4′), 3.45 (3H, s, H8″), 3.20 (2H, m, H8), 2.21 (3H, s, H11′), 2.02 (3H, s, CH_3CO), 1.83 (3H, s, CH_3CO), 1.75 (3H, s, CH_3CO), 1.71 (3H, s, CH_3CO), 1.63 (6H, s, H11, H12), 1.25 (3H, s, H6″), 1.05 (3H, s, H13). ^{13}C-NMR (126 MHz, CD_3OD): δ = 171.50 (4×CH_3CO), 169.96 (Ca), 163.01 (C2′), 161.12 (C4′), 160.81 (C5), 159.52 (C7′), 159.13

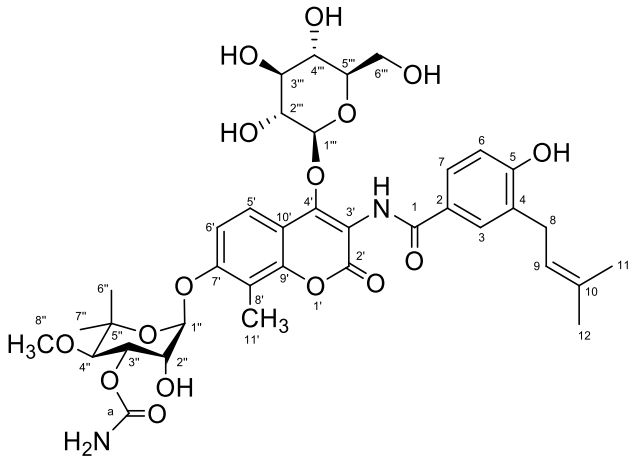
(C1), 152.22 (C9′), 133.71 (C10), 130.89 (C3), 130.12 (C4), 128.48 (C7), 125.01 (C2), 123.43 (C5′), 123.31 (C9), 115.71 (C6), 115.29 (C8′), 112.1 (C10′), 111.78 (C6′), 108.45 (C3′), 101.72 (C1‴), 100.17 (C1″), 82.62 (C4″), 80.11 (C5″), 73.01 (C4‴), 72.72 (C5‴), 71.98 (C3‴), 70.82 (C2″), 70.29 (C2‴), 68.01 (C3″), 62.03 (C8″), 61.30 (C6‴), 29.29 (C8), 29.01 (C6″), 26.03 (C11), 23.27 (C7″), 20.46 (4×CH_3CO), 18.01 (C12), 8.72 (C11′). HMBC indicated that H1‴ is coupled to C4′.

LC/MS under Method A, LC Rt= 6.12 min, ES^- (m/z) calculated 941, found at [M - H^+] 941. LC/MS under Method B, LC Rt = 3.76 min, ES^- (m/z) calculated 941, found at [M - H^+] 941.

HRMS calculated for 943.3324 found 943.3327 (M + H^+)

This is a literature compound and the data is in accordance with literature (Patel, Fuente et al. 2011).

Synthesis of β-4'-*O*-glucosyl novobiocin 66

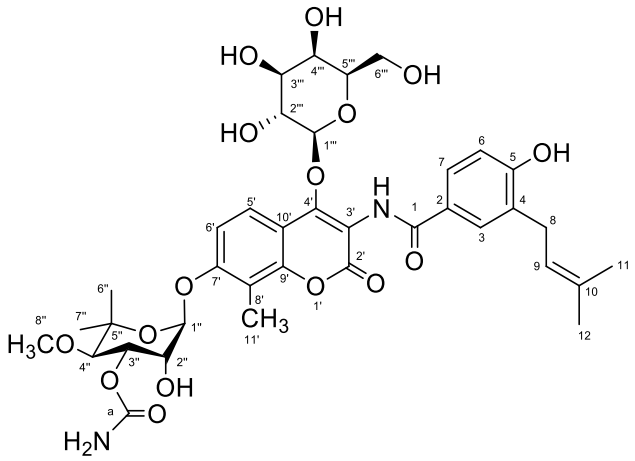


Potassium carbonate (14.63 mg, 0.106 mmol) was added into an anhydrous methanol solution (10 mL) of β-4'-*O*-2,3,4,6-tetra-acetyl glucosyl novobiocin (**64**, 100 mg, 0.105 mmol), contained in a 25 mL r.b flask. After the mixture be stirred for 1.5 hours, the solvent was removed under reduced pressure. The crude product was purified by the Biotage method 1 and then freeze dried to afford a white powder (37 mg, 42%). ^{1}H NMR (500 MHz, CD_3OD): δ = 7.79 (1H, d, *J* = 9.1 Hz, H5', 7.76 (1H, s, H3), 7.74 (1H, d, *J* = 7.8 Hz, H7), 7.58 (1H, d, *J* = 9.0 Hz, H6'), 7.16 (1H, d, *J* = 9.6 Hz, H6), 5.51 (1H, d, *J* = 2.7 Hz, H1''), 5.31 (1H, t, *J* = 6.6 Hz, H9), 5.26 (1H, dd, *J* = 3.3, 9.9 Hz, H3''), 4.99 (1H, d, *J* = 8.7 Hz, H1'''), 4.17 (1H, t, *J* = 2.7 Hz, H2''), 3.85 (1H, dd, *J* = 2.1, 11.7 Hz, H6'''), 3.65 (1H, dd, *J* = 5.3, 12.0 Hz, H6'''), 3.52 (1H, m, H5'''), 3.52 (1H, m, H4''), 3.50 (1H, m, H4'''), 3.49 (3H, s, H8''), 3.48 (1H, m, H3'''), 3.42 (1H, m, H2'''), 3.24 (2H, m, H8), 2.25 (3H, s, H11'), 1.68 (6H, s,H11, H12), 1.28 (3H, s, H6''), 1.09 (3H, s, H7''). ^{13}C-NMR (126 MHz, CD_3OD): δ = 170.11 (Ca), 161.73 (C2′), 159.86 (C4′), 159.64 (C5), 158.44 (C7′), 158.02 (C1), 151.12 (C9′), 132.63 (C10), 129.71 (C3), 128.88 (C4), 127.46 (C7), 123.96 (C2), 122.46 (C5′), 122.23 (C9), 114.55 (C6), 113.41 (C8′), 110.78 (C6′), 111.07 (C10′), 107.69 (C3′), 100.73 (C1‴), 99.22 (C1″), 81.18 (C4″), 78.14 (C5″), 76.88 (C3‴), 76.44 (C4‴), 72.48 (C3″), 69.92 (C2″), 69.81 (C2‴), 60.76 (C8″), 60.88 (C6‴), 60.76 (C5‴), 28.28 (C8), 28.02 (C7″), 24.97 (C11), 21.78 (C6″), 16.96 (C12), 7.61 (C11′). HMBC indicated that H1‴ is coupled to C4′.

LC/MS under method B, LC Rt = 2.96 min, ES^- (m/z) calculated 773.27, found at [M - H^+] 773.45

HRMS calculated for 797.2716 found 797.2738 ($M + Na^+$)

Synthesis of β-4'-*O*-galactosyl novobiocin 67



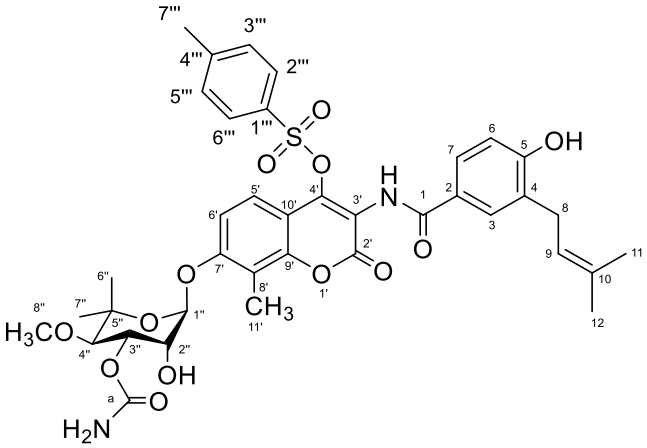
Potassium carbonate (14.63 mg 0.106 mmol) was added into an anhydrous methanol solution (10 mL) of β-4'-*O*-2,3,4,6-tetra-acetyl galactosyl novobiocin (**65**, 100 mg, 0.105 mmol), contained in a 25 mL r.b flask. After the mixture be stirred for 1.5 h, the solvent inside the flask was removed under reduced pressure. The crude product was purified by the Biotage method 1 and then freeze-dried to afford a white powder (37 mg, 42%). ^{1}H NMR (500 MHz, CD_3OD): δ= 7.71 (1H, d, *J* = 8.7 Hz, H5'), 7.67 (1H, d, *J* = 1.8 Hz, H3), 7.62 (1H, dd, *J* = 2.3, 8.4 Hz, H7), 7.02 (1H, d, *J* = 8.6 Hz, H6'), 6.71 (1H, d, *J* = 8.3 Hz, H6), 5.45 (1H, d, *J* = 2.3 Hz, H1″), 5.23 (1H, m, H9), 5.25 (1H, m, H3''), 5.13 (1H, s, H1'''), 4.13 (1H, t, *J* = 2.6 Hz, H2''), 4.23 (1H, m, H6'''), 3.43 (1H, m, H6'''), 4.20 (1H, t, *J* = 4.3 Hz, H5'''), 3.50 (1H, m, H4''), 3.84 (1H, t, *J* = 4.5 Hz, H4'''), 3.46 (3H, s, H8''), 3.73 (1H, m, H3'''), 3.55 (1H, m, H2'''), 3.22 (2H, m, H8), 2.21 (3H, s, H11'), 1.64 (6H,s, H11, H12), 1.25 (3H, s, H6''), 1.07 (3H, s, H7''). ^{13}C-NMR (126 MHz, CD_3OD): δ = 170.25 (Ca), 163.04 (C2′), 161.12 (C4′), 160.84 (C5), 159.52 (C7′), 159.13 (C1), 152.22 (C9′), 133.71 (C10), 130.89 (C3), 130.12 (C4), 128.48 (C7), 125.01 (C2), 123.43 (C5′), 123.31 (C9), 115.71 (C6), 115.29 (C8′), 111.78 (C6′), 112.14 (C10′), 108.45 (C3′), 102.81 (C1‴), 99.94 (C1″), 82.33 (C4″), 79.78 (C5″), 76.12 (C4‴), 73.28 (C2‴), 72.51 (C3‴), 71.72 (C3″), 71.03 (C2″), 65.86 (C5‴), 64.51 (C6‴), 62.21 (C8″), 29.29 (C8), 29.11 (C7″), 26.03 (C11), 22.25 (C6″), 18.01 (C12), 8.85 (C11′). HMBC indicated that H1‴ is coupled to C4′.

LC/MS under Method B, LC Rt= 2.96 min, ES^- (m/z) calculated 773, found at [M - H^+] 773.

HRMS calculated for 797.2716 found 797.2738 ($M + Na^+$)

This is a literature compound and the data is in accordance with literature (Patel, Fuente et al. 2011).

Synthesis of 4’ – tosyl – novobiocin 107



Chemical Formula: $C_{38}H_{42}N_2O_{13}S$
Exact Mass: 766.24
Molecular Weight: 766.82

Novobiocin sodium salt **0** (300 mg, 0.472 mmol) was slowly added into a r.b flask containing 5 mL anhydrous DMF and stirred until it was fully dissolved. 4-Toluenesulfonyl chloride (100 mg, 0.519 mmol) was added into the solution, and the reaction was continued for a period of 2 hours under the protection of nitrogen at room temperature. When the reaction was complete, the resulted liquid was filtered to remove any unreacted novobiocin. Then the filtrate was directly purified using Biotage method 2 to afford a light-yellow solid (125 mg 35%). Malting point: Decomposed at 162.5 °C

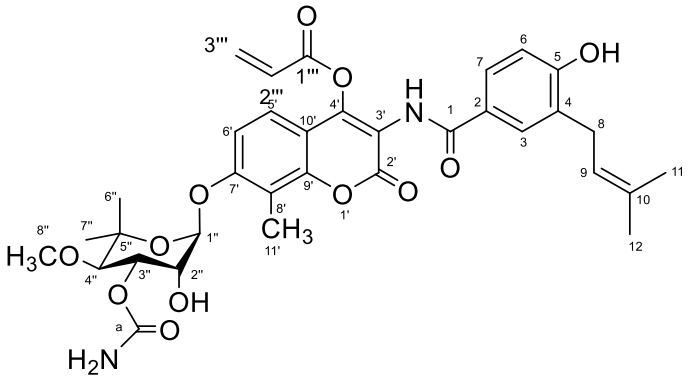
^{1}H NMR (500 MHz, DMSO): δ = 10.08 (1H, s, C5-O**H**, showed HMBC coupling with C5, C4), 9.48 (1H, s, N**H**CO, showed HMBC weak coupling with C1 and strong with C4’ and C2’), 7.77 (2H, d, *J* = 8.21 Hz, H3’’’ and H5’’’), 7.51 (1H, d, *J* = 5.86 Hz, H7), 7.51 (1H, s, H3), 7.42 (1H, d, *J* = 8.80 Hz, H5’), 7.26 (1H, d, *J* = 9.38 Hz, H6), 7.17 (2H, d, *J* = 8.21 Hz, H2’’’ and H6’’’), 6.78 (1H, d, *J* = 8.80 Hz, H6’), 5.58 (1H, d, *J* = 1.76 Hz, H1’’), 5.29 (1H, t, *J* = 6.80 Hz, H9), 5.17 (1H, dd, *J1* = 3.27 Hz, *J2* = 9.82 Hz, H4’’), 4.09 (1H, m, H2’’), 3.48 (1H, t, *J* = 9.87 Hz H3’’), 3.25 (2H, d, *J* = 7.67 Hz H8), 2.26 (3H, s, H11’’), 2.17 (3H, s, H7’’’), 1.73 (6H, s + s, H11, H12), 1.29 (3H, s, H7’’), 1.05 (3H, s, H6’’). ^{13}C NMR (126 MHz, DMSO): δ = 165.41 (C2’), 158.38 (C1), 157.35 (C7), 156.25 (C4’), 151.61 (C9’), 149.79 (C5), 145.86 (C1’’’), 132.44 (C10), 131.38 (C3), 130.00 (C4’’’), 129.67 (C2’’’, C6’’’), 127.58 (C4), 127.26 (C3’’’, C5’’’), 126.96

(C5'), 123.11 (C2), 122.59 (C7), 122.50 (C9), 114.84 (C8'), 113.90 (C6'), 110.92 (C6), 110.31 (C10'), 98.46 (C1''), (80.63) C3'', 78.19 (C5''), 70.24 (C4''), 68.57 (C2''), 28.36 (C6''), 28.08 (C8), 25.57 (C12), 22.74 (C6''), 20.98 (C7'''), 17.71 (C11), 8.28 (C11'), 60.97 (-OMe).

LC/MS under Method B, LC Rt = 4.07 min. ES^- (*m/z*) calculated 765.24, found at $[M - H^+]$ = 764.71

HRMS calculated for 767.2495 found at $[M + H^+]$ 767.2486

Synthesis of 4’ – acrylate – novobiocin 108



Chemical Formula: $C_{34}H_{38}N_2O_{12}$
Exact Mass: 666.24
Molecular Weight: 666.68

Novobiocin sodium salt (300 mg, 0.472 mmol) was slowly added into a r.b flask containing 5 mL anhydrous DMF and stirred until it was fully dissolved. Acryloyl chloride (50 mg, 0.552 mmol) was added into the solution and the reaction was continued for a period of 2 hours under protection of nitrogen at room temperature. When the reaction was complete, the resulted liquid was filtered to remove any unreacted novobiocin. Then the filtrate was directly purified using Biotage method 2 to afford a white solid (62 mg 20%). Malting point: Decomposed at 160.0 °C

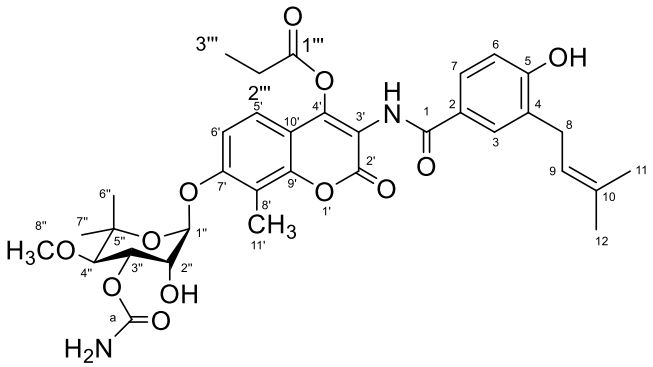
^{1}H NMR (500 MHz, DMSO): δ = 10.07 (1H, s, C5-O**H**, showed HMBC coupling with C5, C4), 9.64 (1H, s, N**H**CO, showed HMBC coupling with C1 and C4’), 7.64 (1H, s, H3), 7.62 (1H, d, *J* = 10.51 Hz, H7), 7.47 (1H, d, *J* = 9.05 Hz, H5’), 7.28 (1H, d, *J* = 9.05 Hz, H6’), 6.84 (1H, d, *J* = 8.33 Hz, H6), 6.57 (1H, d, *J* = 17.37 Hz, H3’’’ trans), 6.40 (1H, dd, *J1* = 10.06 Hz, *J2* = 17.37 Hz H2’’’), 6.19 (1H, d, *J* = 10.06 Hz, H3’’’ cis), 5.57 (1H, d, *J* = 2.13 Hz, H1’’), 5.30 (1H, t, *J* = 6.80 Hz, H9), 5.17 (1H, dd, *J1* = 2.90 Hz, *J2* = 9.78 Hz, H4’’), 4.11 (1H, m, H2’’), 3.48 (1H, d, *J* = 7.49 Hz H3’’), 3.25 (2H, d, *J* = 8.15 Hz H8), 2.28 (3H, s, H11’’), 1.70 (6H, s + s, H11, H12), 1.27 (3H, s, H7’’), 1.06 (3H, s, H6’’). ^{13}C NMR (126 MHz, DMSO): δ = 171.26 (C-carbamate), 165.41 (C1), 161.48 (C1’’’), 160.13 (C2’), 158.39 (C5), 157.32 (C7’), 153.16 (C4’), 149.98 (C9’)

135.33 (C3’’’), 131.70 (C10), 131.51 (C3), 130.88 (C7), 127.36 (C2’’’), 127.19 (C4), 123.88 (C2), 123.42 (C9), 122.4 (C5’), 114.27 (C8’), 114.15 (C6’), 110.72 (C6’), 109.58 (C10’), 109.55 (C3’), 98.47 (C1’’), 80.63 (C3’’), 78.14 (C5’’), 70.24 (C4’’), 68.63 (C2’’), 28.29 (C7’’), 27.96.01 (C8), 25.48 (C12), 22.68 (C6’’), 17.59 (C11), 8.21 (C11’).

LC/MS under Method B, LC Rt = 3.73 min. ES^- (m/z) calculated 665.24, found at $[M - H^+]$ = 665.00

HRMS calculated for 665.2347 found at $[M - H^+]$ 665.2336

Synthesis of 4’ – propionyl – novobiocin 109



Chemical Formula: $C_{34}H_{40}N_2O_{12}$
Exact Mass: 668.26
Molecular Weight: 668.70

Novobiocin sodium salt (300 mg, 0.472 mmol) was slowly added into a r.b flask containing 5 mL anhydrous DMF and stirred until it was fully dissolved. Propionyl chloride (51 mg, 0.551 mmol) was added into the solution, and the reaction was continued for a period of 2 hours under the protection of nitrogen at room temperature. When the reaction was complete, the resulted liquid was filtered to remove any unreacted novobiocin. Then the filtrate was directly purified using Biotage method 2 to afford a white solid (68 mg 22%). Malting point: 143.6 °C

^{1}H NMR (500 MHz, DMSO): δ = 10.07 (1H, s, C5-O**H**, showed HMBC coupling with C7), 7.85 (1H, d, *J* = 8.17 Hz, H7), 7.82 (1H, s, H3), 7.45 (1H, d, *J* = 9.99 Hz, H5’), 7.18 (1H, d, *J* = 9.99 Hz, H6’), 6.98 (1H, d, *J* = 8.32 Hz, H6), 5.54 (1H, d, *J* = 1.86 Hz, H1’’), 5.30 (1H, t, *J* = 7.08 Hz, H9), 5.15 (1H, m, H4’’), 4.08 (1H, s, H2’’), 3.47 (1H, m, H3’’), 3.29 (2H, d, *J* = 8.15 Hz H8), 2.26 (3H, s, H11’’), 2.16 (2H, q, *J* = 8.30 Hz, H2’’’), 1.70 (6H, s + s, H11, H12), 1.25 (3H, s, H7’’), 1.03 (3H, s, H6’’), 0.87 (3H, t, *J* = 8.30 Hz, H3’’’).

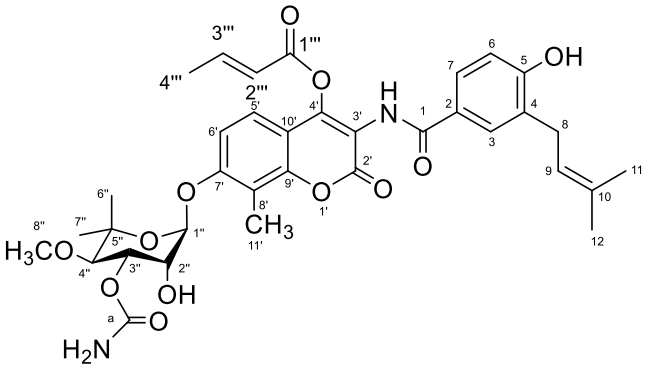
^{13}C NMR (126 MHz, DMSO): δ = 172.53 (C1’’’), 162.04 (C2’), 161.36 (C5), 159.47 (C1), 156.26 (C4’), 157.07 (C7’), 149.88 (C9’), 132.29 (C10), 131.43 (C3), 129.89 (C4), 128.30 (C7), 122.36 (C2), 122.53 (C5’), 121.76 (C9), 115.01 (C6), 112.66 (C8’), 110.83 (C6’), 110.02 (C10’), 98.41 (C1’’), 80.79

(C3''), 78.12 (C5''), 70.35 (C4), 68.68 (C2''), 61.98 (-OMe), 28.42 (C2'''), 28.31 (C8), 28.03 (C7''), 25.72 (C12), 23.36 (C6''), 17.55 (C11), 9.69 (C3'''), 8.43 (C11').

LC/MS under Method B, LC Rt = 3.35 min. ES^- (*m/z*) calculated 667.26, found at $[M - H^+]$ = 666.90

HRMS calculated for 667.2503 found at $[M - H^+]$ 667.2507

Synthesis of 4’ – crotonoyl – novobiocin 110



Chemical Formula: $C_{35}H_{40}N_2O_{12}$
Exact Mass: 680.26
Molecular Weight: 680.71

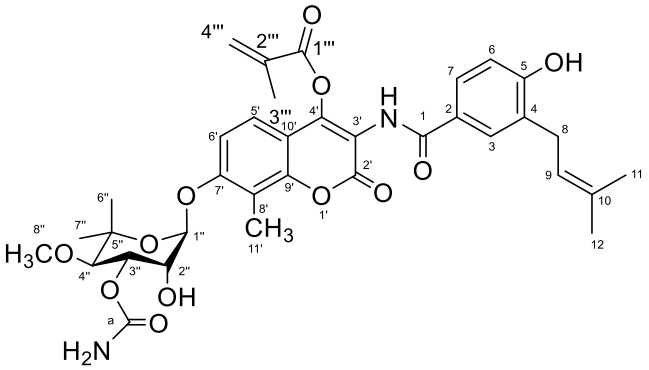
Novobiocin sodium salt (300 mg, 0.472 mmol) was slowly added into a r.b flask containing 5 mL anhydrous DMF and stirred until it was fully dissolved. Crotonyl chloride (50 mg, 0.472 mmol) was added into the solution and the reaction was continued for a period of 2 hours under protection of nitrogen at room temperature. When the reaction was complete, the resulted liquid was filtered to remove any unreacted novobiocin. Then the filtrate was directly purified using Biotage method 2 to afford a white solid (72 mg 22%). Malting point: 145.2 °C

^{1}H NMR (500 MHz, DMSO): δ = 10.09 (1H, s, C5-O**H**, showed HMBC coupling with C5, C4), 9.61 (1H, s, N**H**CO), 7.83 (1H, d, *J* = 8.40 Hz, H7), 7.80 (1H, s, H3), 7.43 (1H, d, *J* = 8.74 Hz, H5’), 7.18 (2H, m, H6’, H3’’’), 6.96 (1H, d, *J* = 8.51 Hz, H6), 5.55 (1H, m, H1’’), 5.29 (1H, m, H9), 5.15 (1H, m, H4’’), 4.08 (1H, s, H2’’), 3.47 (4H, m, H3’’, OMe), 3.26 (2H, d, *J* = 7.33 Hz H8), 2.27 (3H, s, H11’’), 1.90 (3H, d, *J* = 7.00, H4’’’), 1.69 (6H, s + s, H11, H12), 1.26 (3H, s, H7’’), 1.03 (3H, s, H6’’). ^{13}C NMR (126 MHz, DMSO): δ = 162.06 (C2'), C4’ (161.54), 161.12 (C1’’’), 160.9 (C1),159.30 (C5), 158.40 (C9’), 149.99 (C7’), 149.86 (C3’’’), 131.83 (C10), 131.51 (C3), 128.31 (C4), 127.35 (C7), 122.40 (C9), 121.74 (C5), 120.46 (C2), 119.98 (C2’’’) 115.01 (C6), 113.45 (C8’), 110.84 (C6’), 109.98 (C10’), 98.33 (C1’’), 80.63 (C3’’), 70.24 (C4’’), 68.59 (C2’’), 61.02 (OMe), 28.33 (C7’’), 27.99 (C8), 25.51 (C12), 22.71 (C6’’), 17.72 (C4’’’), 17.65 (C11), 8.25 (C11’).

LC/MS under Method B, ES^- (*m/z*) calculated 679.26, found at [M - H^+] = 679.14.

HRMS calculated for 679.2503 found at [M - H^+] 679.2507

Synthesis of 4’ – methacryloyl – novobiocin 111



Chemical Formula: $C_{35}H_{40}N_2O_{12}$
Exact Mass: 680.26
Molecular Weight: 680.71

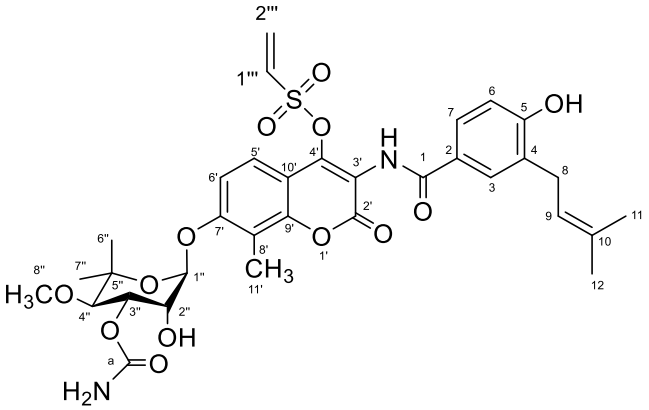
Novobiocin sodium salt (300 mg, 0.472 mmol) was slowly added into a r.b flask containing 5 mL anhydrous DMF and stirred until it was fully dissolved. Methacryloyl chloride (50 mg, 0.521 mmol) was added into the solution and the reaction was continued for a period of 2 hours under protection of nitrogen at room temperature. When the reaction was complete, the resulted liquid was filtered to remove any unreacted novobiocin. Then the filtrate was directly purified using Biotage method 2 to afford a white solid (64 mg 20%). Malting point: Decomposed at 144.8 °C

^{1}H NMR (500 MHz, DMSO): δ = 10.07 (1H, s, C5-O**H**, showed HMBC coupling with C5, C4), 9.64 (1H, s, N**H**CO), 7.83 (1H, d, *J* = 8.50 Hz, H7), 7.79 (1H, s, H3), 7.46 (1H, d, *J* = 9.27 Hz, H5’), 7.19 (1H, d, *J* = 9.27 Hz, H6’), 6.97 (1H, d, *J* = 8.25 Hz, H6), 5.68 (1H, s, H4’’’ trans), 5.55 (1H, m, H1’’), 5.40 (1H, s, H4’’’ cis), 5.28 (1H, m, H9), 5.16 (1H, m, H4’’), 4.01 (1H, m, H2’’), 3.47 (4H, m, H3’’, OMe), 3.27 (2H, d, *J* = 7.80 Hz H8), 2.27 (3H, s, H11’’), 1.78 (3H, s, H3’’’), 1.70 &1.68 (6H, s + s, H11, H12), 1.25 (3H, s, H7’’), 1.04 (3H, s, H6’’). ^{13}C NMR (126 MHz, DMSO): δ = 166.99 (C1’’’), 162.83 (C2'), 162.11 (C4’), 161.11 (C5), 159.29 (C1), 158.40 (C9’), 150.05 (C7’), 138.88 (C2’’’), 131.79 (C10), 131.57 (C3), 128.31 (C4), 127.35 (C7), 122.34 (C9), 121.91 (C5’), 121.72 (C4’’’), 120.83 (C2), 115.02 (C6), 113.42 (C8’), 110.84 (C6’), 109.95 (C10’), 98.33 (C1’’), 80.63 (C3’’), 70.24 (C4’’), 68.59 (C2’’), 60.95 (OMe), 28.31 (C7’’), 27.97 (C8), 25.50 (C12), 22.69 (C6’’), 18.47 (C3’’’), 17.64 (C11), 8.26 (C11’).

LC/MS under Method B, ES^- (m/z) calculated 679.26, found at [M - H^+] = 679.01.

HRMS calculated for 679.2503 found at [M - H^+] 679.2507

Synthesis of 4’ – ethenesulfonyl – novobiocin 112



Chemical Formula: $C_{33}H_{38}N_2O_{13}S$
Exact Mass: 702.21
Molecular Weight: 702.73

Novobiocin sodium salt (300 mg, 0.472 mmol) was slowly added into a r.b flask containing 5 mL anhydrous DMF and stirred until it was fully dissolved. 2-Chloro-1-ethanesulfonyl chloride (85 mg, 0.521 mmol) was added into the solution and the reaction was continued for a period of 2 hours under protection of nitrogen at room temperature. When the reaction was complete, the resulted liquid was filtered to remove any unreacted novobiocin. Then the filtrate was directly purified using Biotage method 2 to afford a white solid (43 mg 13%). Malting point: Decomposed at 160.1 °C

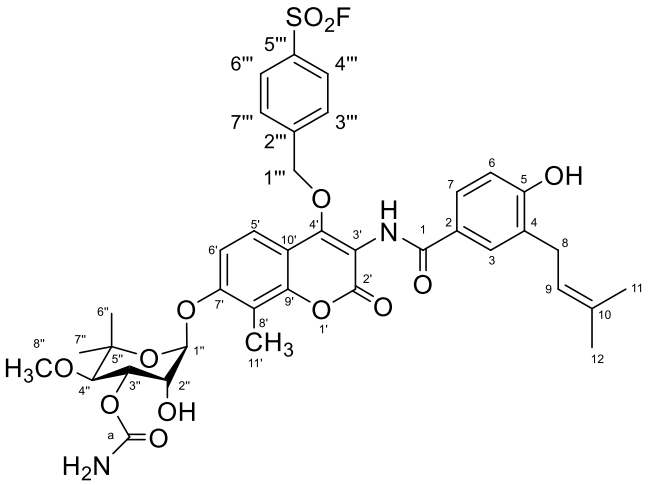
^{1}H NMR (126 MHz, DMSO): δ = 10.14 (1H, s, C5-O**H**, showed HMBC coupling with C5, C4), 9.64 (1H, s, N**H**CO, showed HMBC coupling with and C2’), 7.70 (1H, s, H3), 7.68 (1H, d, *J* = 9.60 Hz, H7), 7.63 (1H, d, *J* = 9.11 Hz, H5’), 7.30 (1H, d, *J* = 9.11 Hz, H6’), 7.18 (1H, dd, *J1* = 10.00 Hz, *J2* = 16.40 Hz H1’’’), 6.89 (1H, d, *J* = 8.20 Hz, H6), 6.41 (1H, d, *J* = 16.41 Hz, H2’’’ trans), 6.25 (1H, d, *J* = 10.03 Hz, H2’’’ cis), 5.59 (1H, d, *J* = 2.49 Hz, H1’’), 5.32 (1H, t, *J* = 7.10 Hz, H9), 5.16 (1H, m, H4’’), 4.09 (1H, s, H2’’), 3.50 (1H, m, H3’’), 3.28 (2H, d, *J* = 6.98 Hz H8), 2.26 (3H, s, H11’’), 1.71 (6H, s + s, H11, H12), 1.29 (3H, s, H7’’), 1.06 (3H, s, H6’’), 3.48 (3H,s, -OMe). ^{13}C NMR (500 MHz, DMSO): δ = 165.32 (C2’), 159.36 (C4’), 158.55 (C5), 157.51 (C7), 156.24 (C1), 149.93 (C9), 133.27 (C2’’’), 133.02 (C10), 132.23 (C1’’’), 131.59 (C3), 129.70 (C7), 127.42 (C4), 123.64 (C2), 122.27 (C9), 122.40 (C5’), 114.87 (C6), 114.34

(C8), 113.57 (C3'), 110.98 (C6'),110.10 (C10') 98.50 (C1''), 80.79 (C3''), 78.21 (C5''), 70.23 (C4''), 68.83 (C2''), 61.98 (OMe), 28.42 (C7''), 28.31 (C8), 25.55 (C12), 22.78 (C6''), 17.67 (C11), 8.27 (C11').

LC/MS under Method B, LC Rt = 3.74 min. ES^- (*m/z*) calculated 701.21, found at $[M - H^+]$ = 701.05

HRMS calculated for 701.2017 found at $[M - H^+]$ 701.2017

Synthesis of 4’ – para -benzylsulfonyl fluoride – novobiocin 113



Chemical Formula: $C_{38}H_{41}FN_2O_{13}S$
Exact Mass: 784.23
Molecular Weight: 784.81

Novobiocin sodium salt (300 mg, 0.472 mmol) was slowly added to a flask containing anhydrous DMF (5 ml) and stirred until it was fully dissolved. 4 -(Bromomethyl) benzenesulfonyl fluoride (120 mg, 0.472 mmol) was added into the solution and the reaction was continued for a period of 2 hours under the protection of nitrogen at room temperature. When the reaction was complete, the resulted liquid was filtered to remove any unreacted novobiocin. Then the filtrate was directly purified using Biotage method 2 to afford a white solid (52 mg 14%). Malting point: Decomposed at 168.0 °C

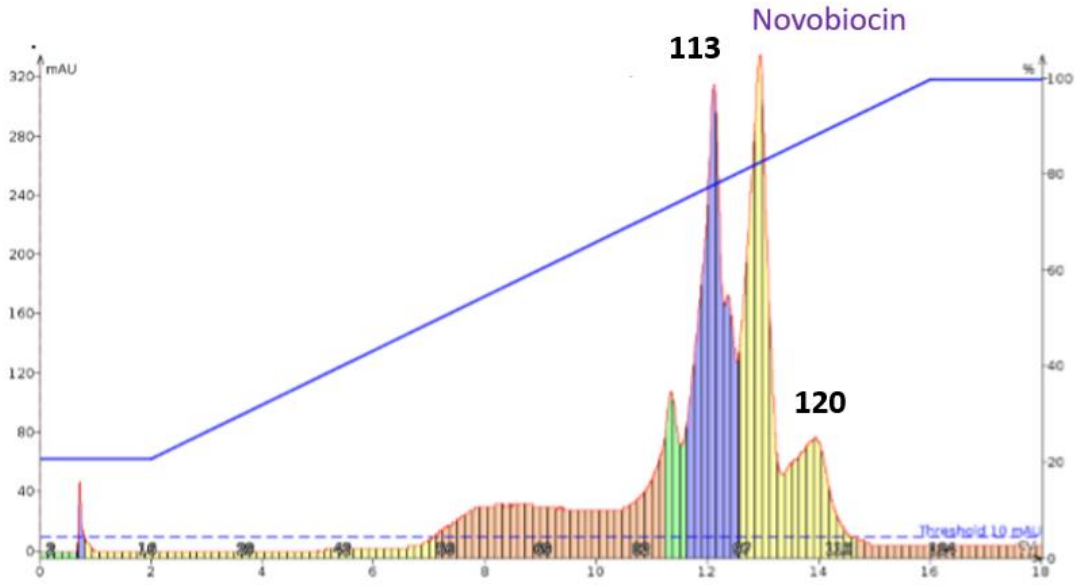
^{1}H NMR (500 MHz, DMSO): δ = 10.21 (1H, s, C5-O**H**, showed HMBC coupling with C5, C4), 9.79 (1H, s, N**H**CO, showed HMBC coupling with C1), 7.87 (2H, d, *J* = 8.43 Hz, H4’’’ and H6’’’), 7.68 (1H, d, *J* = 8.41 Hz, H7), 7.64 (1H, s, H3), 7.60 (1H, d, *J* = 8.85 Hz, H5’), 7.44 (2H, d, *J* = 8.42 Hz, H3’’’ and H7’’’), 7.00 (1H, d, *J* = 9.08 Hz, H6’), 6.88 (1H, d, *J* = 8.48 Hz, H6), 5.47 (1H, m, H1’’), 5.29 (1H, m, H9), 5.10 (1H, m, H4’’), 4.02 (1H, m, H2’’), 3.48 (6H, m, H3’’, H1’’’, OMe), 3.26 (2H, d, *J* = 8.16 Hz H8), 1.92 (3H, s, H11’’), 1.70 (6H, s + s, H11, H12), 1.23 (3H, s, H7’’), 0.974 (3H, s, H6’’). ^{13}C NMR (126 MHz, DMSO): δ = 188.64 (C2’), 167.68 (C1), 167.31 (C4’), 160.33 (C7’), 158.85 (C5), 151.95 (C9’), 141.64 (C2’’’), 131.96 (C10), 131.86 (C7’’’, C3’’’), 131.81 (C5’’’), 129.66 (C3), 127.83 (C4’’’,C6’’’), 127.62 (C7), 125.65 (C5’), 122.42 (C2), 122.19 (C9),

114.31 (C6), 113.62 (C8'), 113.39 (C10)', 110.25 (C6), 98.45 (C1''), 80.52 (C5''), 78.55 (C3''), 70.15 (C4''), 68.64 (C2''), 61.10 (OMe), 41.31 (C1'''), 28.54 (C7''), 27.81 (C8), 25.54 (C12), 22.51 (C6''), 17.67 (C11), 7.83 (C11').

LC/MS under Method B, LC Rt = 4.09 min. ES^- (*m/z*) calculated 783.23, found at $[M - H^+]$ = 782.80.

HRMS calculated for 783.2235 found at $[M - H^+]$ 783.2246

Characterization of **120**

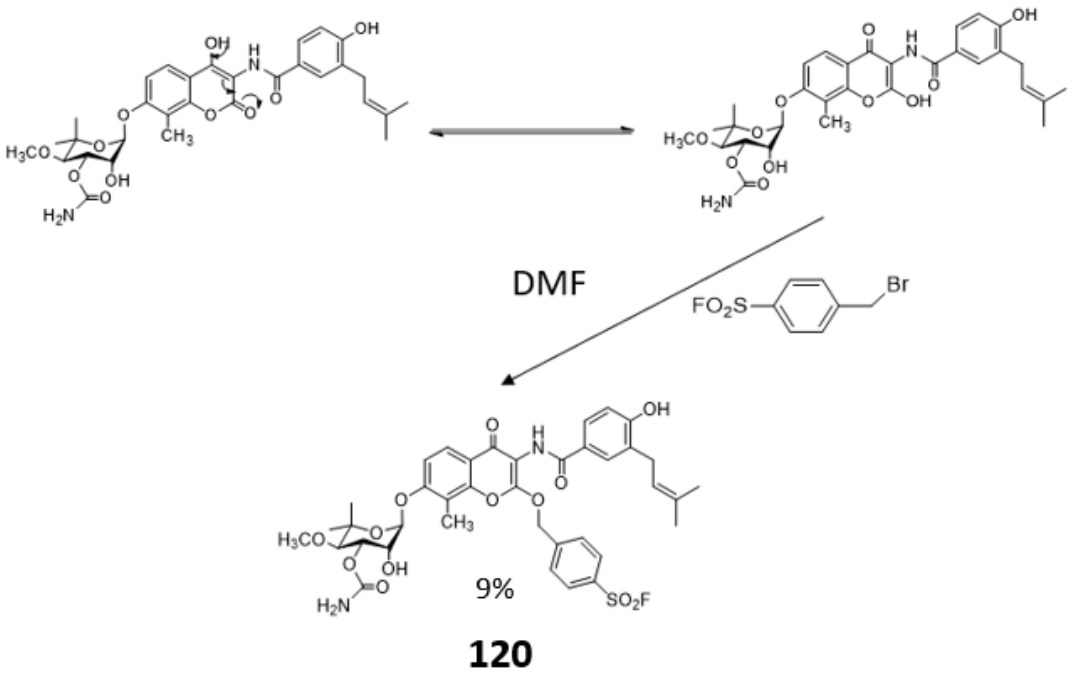


We also separated and tested a side product of **113** (mentioned as compound **120**), this compound has a higher retention time than **113** in Biotage purification. This compound has the same molecular mass according to LC-MS, and it is an isomeric form of **113**.

Due to NMR of this compound still showed C5-O**H** and amide N**H** like all other compounds, none of these two positions is substituted. In ^{1}H NMR, H1''' have varied shifts, found at 3.48 ppm (Mixed peak, assigned using DEPT-135 and HSQC) in **113** and 5.59 ppm in **120** (assigned using the same method). Another significant difference between these two compounds is that the 188.64 ppm peak of **113**'s C2' is missing in **120**, suggesting that C2' modification may have happened. Moreover, in our published paper, we

reported that novobiocin sodium salt underwent 4'-OH to 2'-OH tautomerism and resulted in both 4'-substituted and 2'-substituted products. H1''' of 4'-substituted products also shifted from 3.3 ppm to 5.5 ppm in 2'- substituted product (Simon, Huang et al. 2017). In ^{13}C spectrums of our products, HMBC of H1''' in **113** showed strong coupling with C4' and weak but visible coupling with C2'. H1''' in **120** only showed coupling with C2' (ppm changed to 160 ppm instead of 188 ppm)

As a result, we conclude that **120** is the 2'-substituted isomeric form of **113**.



Yield for **120** is 33 mg, 9 %. Malting point: Decomposed at 215.3 °C

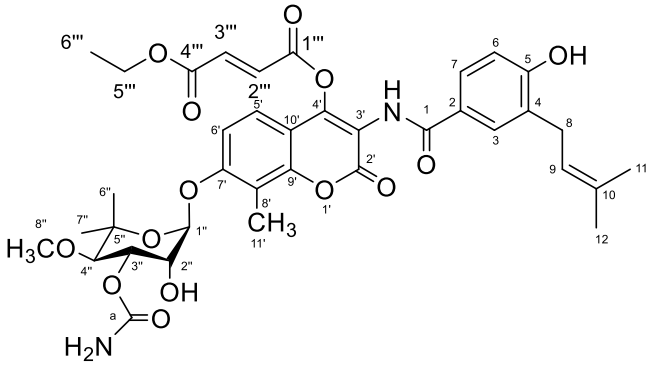
^{1}H NMR (500 MHz, DMSO): δ = 10.08 (1H, s, C5-O**H**, showed HMBC coupling with C5, C4), 9.79 (1H, s, N**H**CO, showed HMBC coupling with C1), 8.07 (2H, d, *J* = 8.30 Hz, H4''' and H6'''), 7.75 (3H, m, H7''', H3''', H5'), 7.63 (1H, s, H3), 7.60 (1H, d, *J* = 8.27 Hz, H7), 7.20 (H, d, *J* = 9.17 Hz, H6'), 6.83 (1H, d, *J* = 8.37 Hz, H6), 5.68 (1H, s, H1'''), 5.57 (1H, d, *J* = 2.31, H1''), 5.30 (1H, m, H9), 5.16 (1H, m, H4''), 4.09 (1H, m, H2''), 3.48 (4H, m, H3'', OMe), 3.26 (2H, d, *J* = 7.18 Hz H8), 2.24 (3H, s, H11''), 1.69 (6H, s + s, H11, H12), 1.26 (3H, s, H7''), 1.05 (3H, s, H6''). ^{13}C NMR (126 MHz, DMSO): δ = 167.25

(C1), 160.92 (C2'), 160.34 (C4'), 158.45 (C5), 157.12 (C7'), 150.54 (C9'), 145.41 (C2'''), 131.49 (C5'''), 130.50 (C10), 129.48 (C3), 128.60 (C4''',C6'''), 128.51 (C3''', C7'''), 127.34 (C4), 127.04 (C7), 123.74 (C2), 122.42 (C9), 121.99 (C5'), 114.21 (C6), 112.99 (C8'), 110.50 (C10'), 110.38 (C6), 105.67 (C3'), 98.37 (C1''), 80.64 (C3''), 78.11 (C5''), 72.44 (C1'''), 70.26 (C4''), 68.64 (C2''), 60.97 (OMe), 28.38 (C8), 28.00 (C7), 25.52 (C12), 22.68 (C6''), 17.66 (C11), 8.26 (C11').

LC/MS under Method B, LC Rt = 4.72 min. ES^- (*m/z*) calculated 783.23, found at $[M - H^+]$ = 782.79.

HRMS calculated for 783.2235 found at $[M - H^+]$ 783.2231

Synthesis of 4’ – ethyl fumaroyl – novobiocin 136



Chemical Formula: $C_{37}H_{42}N_2O_{14}$
Exact Mass: 738.26
Molecular Weight: 738.74

Novobiocin sodium salt (300 mg, 0.472 mmol) was slowly added into a r.b flask containing 5 mL anhydrous DMF and stirred until it was fully dissolved. Ethyl fumaroyl chloride (77 mg, 0.472 mmol) was added into the solution and the reaction was continued for a period of 2 hours under protection of nitrogen at room temperature. When the reaction was complete, the resulted liquid was filtered to remove any unreacted novobiocin. Then the filtrate was directly purified using Biotage method 2 to afford a brown oil (55 mg 16%).

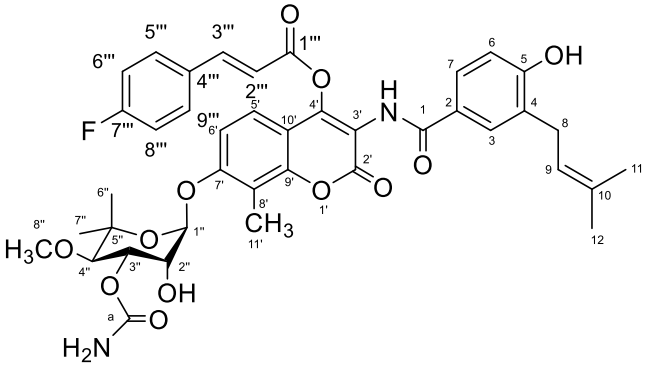
^{1}H NMR (500 MHz, DMSO): δ = 10.72 (1H, s, C5-O**H**), 9.24 (1H, s, N**H**CO), 7.84 (1H, d, *J* = 8.47 Hz, H7), 7.80 (1H, s, H3), 7.42 (1H, d, *J* = 7.46 Hz, H5’), 7.186 (2H, m, H6’, H3’’’), 6.91 (1H, d, *J* = 8.58 Hz, H6), 6.57 (1H, d, *J* = 15.47 Hz, H2’’’), 5.55 (1H, m, H1’’), 5.28 (1H, m, H9), 5.15 (1H, m, H4’’), 4.18 (2H, q, *J* = 7.06, H5’’’), 4.08 (1H, m, H2’’), 3.47 (4H, m, H3’’, OMe), 3.27 (2H, d, *J* = 7.13 Hz H8), 2.27 (3H, s, H11’’), 1.70 & 1.67 (6H, s + s, H11, H12), 1.25 (6H, m, H7’’, H6’’’), 1.03 (3H, s, H6’’). ^{13}C NMR (126 MHz, DMSO): δ = 164.69 (C1’’’), 169.61 (C4’’’), 161.93 (C2’), 161.23 (C5), 159.11 (C1), 158.02 (C4’), 157.36 (C9’), 150.04 (C7’), 135.69 (C3’’’), 131.82 (C10), 130.15 (C2’’’), 129.80 (C3), 128.35 (C4), 127.39 (C7), 122.59 (C5’), 122.51 (C9), 121.68 (C2), 114.16 (C6), 113.50 (C8’), 110.96 (C6’), 110.00 (C10), 98.40 (C1’’), 80.68 (C3’’), 78.15 (C5’’), 70.27 (C4’’), 68.68 (C2’’), 60.83 (C5’’’), 60.95 (OMe), 28.43

(C8), 28.30 (C7''), 15.55 (C12), 22.68 (C6''), 17.60 (C11), 14.37 (C6'''), 8.74 (C11')..

LC/MS under Method B, LC Rt = 3.83 min. ES^- (m/z) calculated 737.26, found at $[M - H^+]$ = 737.88.

HRMS calculated for 737.2558 found at $[M - H^+]$ 737.2535

Synthesis of 4’ – para – fluoro - cinnamoyl – novobiocin 137



Chemical Formula: $C_{40}H_{41}FN_2O_{12}$
Exact Mass: 760.26
Molecular Weight: 760.77

Novobiocin sodium salt (300 mg, 0.472 mmol) was slowly added into a r.b flask containing 5 mL anhydrous DMF and stirred until it was fully dissolved. 4 – Fluorocinnamoyl chloride (87 mg, 0.472 mmol) was added into the solution and the reaction was continued for a period of 2 hours under protection of nitrogen at room temperature. When the reaction was complete, the resulted liquid was filtered to remove any unreacted novobiocin. Then the filtrate was directly purified using Biotage method 2 to afford a white solid (49 mg 14%). Malting point: Decomposed at 160.0 °C

^{1}H NMR (500 MHz, DMSO): δ = 10.74 (1H, s, C5-O**H**), 7.84 (1H, d, *J* = 8.72 Hz, H7), 7.79 (1H, s, H3), 7.61 (2H, m, H5’’’, H9’’’), 7.45 (2H , m, H5’, H3’’’), 7.22 (3H, m, H6’, H6’’’, H8’’’), 6.94 (1H, d, *J* = 8.72 Hz, H6), 6.74 (1H, d, *J* = 15.61 Hz, H2’’’), 5.54 (1H, m, H1’’), 5.23 (1H, m, H9), 5.16 (1H, m, H4’’), 4.08 (1H, s, H2’’), 3.49 (4H, m, H3’’, OMe), 3.23 (2H, d, *J* = 7.31 Hz H8), 2.27 (3H, s, H11’’), 1.64 (6H, s, H11, H12), 1.25 (3H, s, H7’’), 1.03 (3H, s, H6’’).

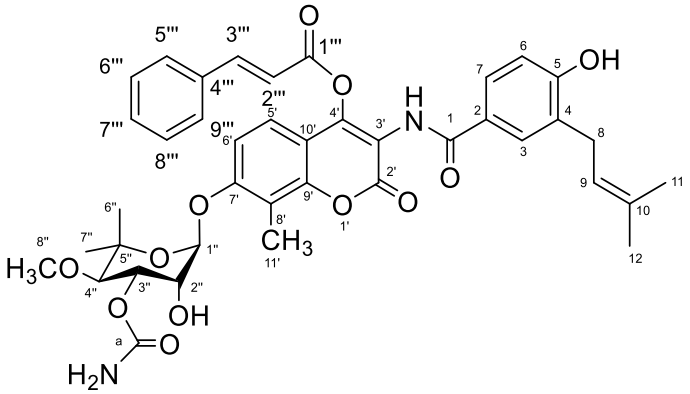
^{13}C NMR (126 MHz, DMSO): δ = 168.53 (C2’), 164.13 (C1’’’), 163.5 (C4’), 163.44 (C7’’’), 161.13 (C5), 160.8 (C1), 157.13 (C7’), 149.88 (C9’), 139.57 (C3’’’), 131.82 (C5’’’, C9’’’), 131.74 (C10), 131.51 (C4’’’), 130.20 (C3), 128.33 (C4), 128.28 (C7), 122.51 (C9), 122.47 (C5’), 121.65 (C2), 115.86 (C6’’’, C8’’’), 115.77 (C8’), 115.00 (C2’’’), 114.15 (C6), 110.89 (C6’), 98.34 (C1’’), 80.66

(C3''), 70.30 (C4''), 68.67 (C2''), 28.40 (C7''), 28.02 (C8), 25.54 (C12), 22.67 (C6''), 17.67 (C11), 8.98 (C11').

LC/MS under Method B, LC Rt = 3.98 min. ES^- (*m/z*) calculated 759.26, found at $[M - H^+]$ = 759.00

HRMS calculated for 759.2565 found at $[M - H^+]$ 759.2582

Synthesis of 4’ – cinnamoyl – novobiocin 138



Chemical Formula: $C_{40}H_{42}N_2O_{12}$
Exact Mass: 742.27
Molecular Weight: 742.78

Novobiocin sodium salt (300 mg, 0.472 mmol) was slowly added into a r.b flask containing 5 mL anhydrous DMF and stirred until it was fully dissolved. Cinnamoyl chloride (70 mg, 0.473 mmol) was added into the solution, and the reaction was continued for a period of 2 hours under the protection of nitrogen at room temperature. When the reaction was complete, the resulted liquid was filtered to remove any unreacted novobiocin. Then the filtrate was directly purified using Biotage method 2 to afford a white solid (40 mg 11%). Malting point: Decomposed at 160.0 °C

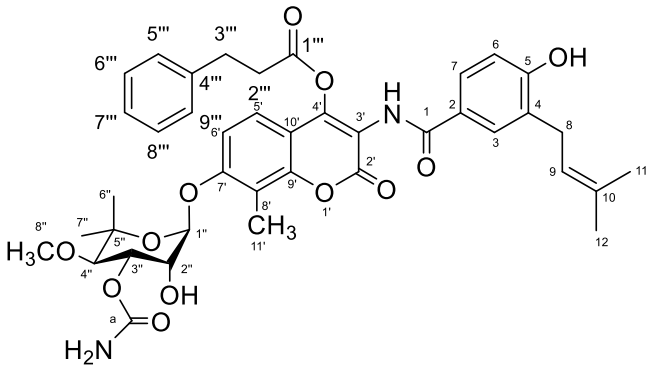
^{1}H NMR (500 MHz, DMSO): δ = 10.06 (1H, s, C5-O**H**), 9.74 (1H, s, N**H**CO), 7.85 (1H, d, *J* = 8.85 Hz, H7), 7.80 (1H, s, H3), 7.43 (7H, m, H5’, H3’’’, all H on cinnamic benzyl ring), 7.20 (1H , d, *J* = 9.09, H6’), 6.95 (1H, d, *J* = 8.85 Hz, H6), 6.80 (1H, d, *J* = 15.89 Hz, H2’’’), 5.55 (1H, m, H1’’), 5.25 (1H, m, H9), 5.16 (1H, m, H4’’), 4.09 (1H, s, H2’’), 3.47 (4H, m, H3’’, OMe), 3.24 (2H, d, *J* = 7.55 Hz H8), 2.29 (3H, s, H11’’), 1.64 (6H, s, H11, H12), 1.25 (3H, s, H7’’), 1.04 (3H, s, H6’’). ^{13}C NMR (126 MHz, DMSO): δ = 168.38 (C2’), 164.75 (C1’’’), 114.66 (C4’), 162.09 (C5), 161.19 (C1), 157.43 (C7’), 150.04 (C9’), 140.71 (C3’’’), 135.4 (C4’’’), 134.41 (C10), 131.82 (C3), 130.20 (C7), 128 -130 (cinnamic ring carbons), 128.80 (C4), 122.38 (C5’), 122.16 (C9), 120.53 (C2’’’), 118.78 (C2), 115.00 (C6), 114.20 (C8’), 110.88 (C6’), 109.82 (C10’),

98.41 (C1''), 80.63 (C3''), 78.14 (C5''), 70.25 (C4''), 68.70 (C2''), 60,95 (-OMe), 28.44 (C7''), 28.31 (C8), 25.44 (C12), 22.68 (C6''), 17.60 (C11), 8.27 (C11').

LC/MS under Method B, LC Rt = 3.97 min. ES^- (*m/z*) calculated 741.27, found at $[M - H^+]$ = 741.10.

HRMS calculated for 741.2659 found at $[M - H^+]$ 741.2673

Synthesis of 4’ – hydrocinnamoyl – novobiocin 139



Chemical Formula: $C_{40}H_{44}N_2O_{12}$
Exact Mass: 744.29
Molecular Weight: 744.79

Novobiocin sodium salt (300 mg, 0.472 mmol) was slowly added into a r.b flask containing 5 mL anhydrous DMF and stirred until it was fully dissolved. Hydrocinnamoyl chloride (85 mg, 0.521 mmol) was added into the solution and the reaction was continued for a period of 2 hours under protection of nitrogen at room temperature. When the reaction was complete, the resulted liquid was filtered to remove any unreacted novobiocin. Then the filtrate was directly purified using Biotage method 2 to afford a white solid (37 mg 11%). Malting point: Decomposed at 142.7 °C

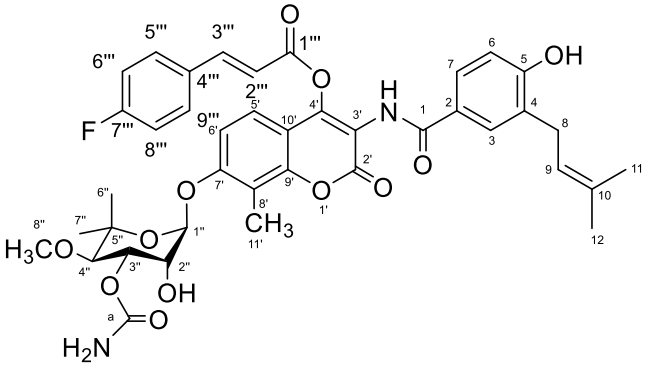
^{1}H NMR (500 MHz, DMSO): δ = 10.80 (1H, s, C5-O**H**), 7.86 (1H, d, *J* = 8.46 Hz, H7), 7.83 (1H, s, H3), 7.45 (7H, d, H5’), 7.17 (3H , m, H6’, H6’’’, H8’’’), 7.12 (1H, t, *J* = 7.15, H7’’’), 7.36 (2H, d, *J* = 6.96, H5’’’, H9’’’), 7.00 (1H, d, *J* = 8.39 Hz, H6), 5.54 (1H, m, H1’’), 5.28 (1H, m, H9), 5.15 (1H, m, H4’’), 4.08 (1H, s, H2’’), 3.57 (4H, m, H3’’, OMe), 3.28 (2H, d, *J* = 7.39 Hz H8), 2.65 (2H, t, *J* = 7.87, H3’’’), 2.44 (2H, t. *J* = 7.31, H2’’’), 2.26 (3H, s, H11’’), 1.67 (6H, s, H11, H12), 1.24 (3H, s, H7’’), 1.02 (3H, s, H6’’). ^{13}C NMR (126 MHz, DMSO): δ = 173.78 (C1’’’), 162.07 (C2’), 161.19 (C5), 159.27 (C4’), 158.02 (C1), 157.12 (C7’), 149.91 (C9’), 140.81 (C4’’’), 132.31 (C10), 131.89 (C3), 129.80 (C7), 128.37 (C6’’’, C8’’’), 128,26 (C5’’’, C9’’’), 128.21 (C4), 125.85 (C7’’’), 122.47 (C5’), 122.34 (C9), 121.71 (C2), 114.30 (C6), 113.42 (C8’), 110.86 (C6), 109.97 (C10), 98.35 (C1’’), 80.66 (C3’’), 78.13 (C5’’), 70.25 (C4’’), 68.71 (C2’’),

36.71 (C2'''), 30.62 (C3'''), 28.39 (C7''), 28.28 (C8), 25.54 (C12), 22.74 (C6''), 17.67 (C11), 8.25 (C11').

LC/MS under Method B, LC Rt = 3.98 min. ES^- (*m/z*) calculated 743.29, found at $[M - H^+]$ = 743.10

HRMS calculated for 743.2816 found at $[M - H^+]$ 743.2817

Synthesis of 4’ – para – fluoro - cinnamoyl – novobiocin 137



Chemical Formula: $C_{40}H_{41}FN_2O_{12}$
Exact Mass: 760.26
Molecular Weight: 760.77

Novobiocin sodium salt (300 mg, 0.472 mmol) was slowly added into a r.b flask containing 5 mL anhydrous DMF and stirred until it was fully dissolved. 4 – Fluorocinnamoyl chloride (87 mg, 0.472 mmol) was added into the solution and the reaction was continued for a period of 2 hours under protection of nitrogen at room temperature. When the reaction was complete, the resulted liquid was filtered to remove any unreacted novobiocin. Then the filtrate was directly purified using Biotage method 2 to afford a white solid (49 mg 14%). Malting point: Decomposed at 160.0 °C

^{1}H NMR (500 MHz, DMSO): δ = 10.74 (1H, s, C5-O**H**), 7.84 (1H, d, *J* = 8.72 Hz, H7), 7.79 (1H, s, H3), 7.61 (2H, m, H5’’’, H9’’’), 7.45 (2H , m, H5’, H3’’’), 7.22 (3H, m, H6’, H6’’’, H8’’’), 6.94 (1H, d, *J* = 8.72 Hz, H6), 6.74 (1H, d, *J* = 15.61 Hz, H2’’’), 5.54 (1H, m, H1’’), 5.23 (1H, m, H9), 5.16 (1H, m, H4’’), 4.08 (1H, s, H2’’), 3.49 (4H, m, H3’’, OMe), 3.23 (2H, d, *J* = 7.31 Hz H8), 2.27 (3H, s, H11’’), 1.64 (6H, s, H11, H12), 1.25 (3H, s, H7’’), 1.03 (3H, s, H6’’).

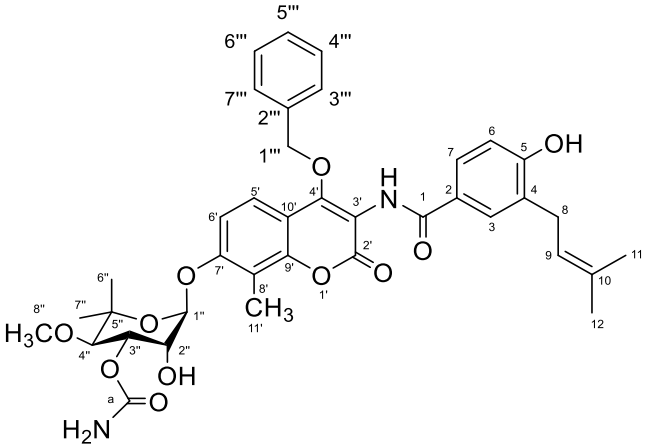
^{13}C NMR (126 MHz, DMSO): δ = 168.53 (C2’), 164.13 (C1’’’), 163.5 (C4’), 163.44 (C7’’’), 161.13 (C5), 160.8 (C1), 157.13 (C7’), 149.88 (C9’), 139.57 (C3’’’), 131.82 (C5’’’, C9’’’), 131.74 (C10), 131.51 (C4’’’), 130.20 (C3), 128.33 (C4), 128.28 (C7), 122.51 (C9), 122.47 (C5’), 121.65 (C2), 115.86 (C6’’’, C8’’’), 115.77 (C8’), 115.00 (C2’’’), 114.15 (C6), 110.89 (C6’), 98.34 (C1’’), 80.66

(C3''), 70.30 (C4''), 68.67 (C2''), 28.40 (C7''), 28.02 (C8), 25.54 (C12), 22.67 (C6''), 17.67 (C11), 8.98 (C11').

LC/MS under Method B, LC Rt = 3.98 min. ES^- (*m/z*) calculated 759.26, found at $[M - H^+]$ = 759.00

HRMS calculated for 759.2565 found at $[M - H^+]$ 759.2582

Synthesis of 4’ – benzyl – novobiocin 140



Chemical Formula: $C_{38}H_{42}N_2O_{11}$
Exact Mass: 702.28
Molecular Weight: 702.76

Novobiocin sodium salt (300 mg, 0.472 mmol) was slowly added into a r.b flask containing 5 mL anhydrous DMF and stirred until it was fully dissolved. Benzyl bromide (81 mg, 0.472 mmol) was added into the solution and the reaction was continued for a period of 2 hours under protection of nitrogen at room temperature. When the reaction was complete, the resulted liquid was filtered to remove any unreacted novobiocin. Then the filtrate was directly purified using Biotage method 2 to afford a white semi-solid (27 mg 8%).

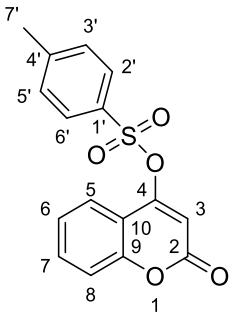
^{1}H NMR (500 MHz, DMSO): δ = 10.16 (1H, s, C5-O**H**, showed HMBC coupling with C5, C4), 9.59 (1H, s, N**H**CO, showed HMBC coupling with C1), 7.73 (1H, m, H7), 7.61 (1H, d, *J* = 9.05 Hz, H5’), 7.39 (5H, m, 4’-benzyl H), 7.15 (1H, d, *J* = 9.03 Hz, H6’), 6.68 (1H, d, *J* = 8.18 Hz, H6), 5.53 (1H, d, *J* = 2.36, H1’’), 5.31 (1H, m, H9), 5.14 (1H, m, H4’’), 4.07 (1H, m, H2’’), 3.47 (4H, m, H3’’, OMe), 3.27 (2H, d, *J* = 7.11 Hz H8), 2.23 (3H, s, H11’’), 1.69 (6H, s + s, H11, H12), 1.26 (3H, s, H7’’), 1.03 (3H, s, H6’’). ^{13}C NMR (126 MHz, DMSO): δ = 166.65 (C1), 161.09 (C2’), 160.44 (C4’), 158.45 (C5), 157.01 (C9’), 150.10 (C7’), 137.60 (C10), 129.54 (C3), 128 -129 (4’-benzyl carbons), 128.59 (C2’’’), 128.10 (C4), 127.10 (C7), 124.00 (C2), 122.76 (C9), 122.37 (C5’), 114.37 (C6), 113.9 (C8), 110.76 (C10’), 110.31 (C6’), 105.29 (C3’), 98.37 (C1’’), 80.65

(C3''), 78.07 (C5''), 73.73 (C1'''), 70.27 (C4''), 68.62 (C2''), 28.32 (C7''), 27.99 (C8), 25.52 (C12), 22.78 (C6''), 17.65 (C11), 8.25 (C11').

LC/MS under Method B, LC Rt = 3.98 min. ES^- (*m/z*) calculated 701.28, found at $[M - H^+]$ = 701.19

HRMS calculated for 701.2711 found at $[M - H^+]$ 701.2701

Synthesis of 4 – tosyl – coumarin 115



Chemical Formula: $C_{16}H_{12}O_5S$
Exact Mass: 316.04
Molecular Weight: 316.33

4 - hydroxy coumarin (500 mg, 3.08 mmol) and 4-Toluenesulfonyl chloride (587mg, 3.08 mmol) were dissolved in 50 mL DCM, and a rapid reaction was observed. The mixture was stirred for 1 hour, and the solvent was removed under vacuum to afford a white solid. The white solid was furtherly washed with 10 mL * 3 methanol to obtain the pure product (821 mg, 84%).

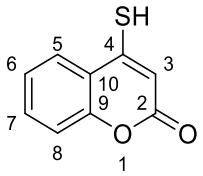
^{1}H NMR (400 MHz, DMSO): δ = 7.92 (1H, d, *J* = 8.45 Hz, H2', H6'), 7.66 (1H, d, *J* = 8.05 Hz, H5), 7.59 (1H, t, *J* = 8.87 Hz, H7), 7.41 (2H, d, *J* = 7.89 Hz, H3', H5'), 7.30 (2H, d + t, *J* = 8.85 Hz H6, H8), 6.33 (1H, s, H3), 2.48 (3H, s, H7'). ^{13}C NMR (126 MHz, $CDCl_3$) *δ* = 160.8 (C4), 157.8 (C2), 153.4 (C9), 146.9 (C1'), 133.3 (C4'), 131.5 (C3' & C5'), 130.4 (C2' & C6'), 128.4 (C7), 124.5 (C6), 123.1 (C5), 116.9 (C10), 114.9 (C8), 103.5 (C3), 21.8 (C7').

LC/MS under Method B, LC Rt= 3.60 min, ES^- (*m/z*) calculated 315.04, found at [M - H^+] 315.33

HRMS calculated for 315.0327 found at [M - H^+] 317.0326

This is a literature compound and the data is in accordance with literature (Majumdar and Ghosh 2002)

Synthesis of 4 – thiol – coumarin 117



Chemical Formula: $C_9H_6O_2S$
Exact Mass: 178.01
Molecular Weight: 178.21

4 – tosyl coumarin (500 mg, 3.08 mmol) and Sodium hydrosulfide (180 mg, 3.21 mmol) were mixed in 50 mL anhydrous MeOH and stirred for 4 hours. After this period, the solvent was removed under vacuum. The residue was added with 10 mL acetone, and the light-yellow precipitates (product) were filtered off. The filtrate was collected, and the process was repeated until no more precipitates were formed. All precipitates were collected to give a yield of 184 mg, 66%.

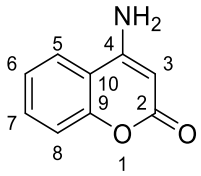
^{1}H NMR (400 MHz, DMSO): δ = 7.87 (1H, d, *J* = 7.40 Hz, H5), 7.73 (1H, t, *J* = 7.89 Hz H7), 7.52 (1H, d, *J* = 8.39 Hz, H8), 7.43 (1H, t, *J* = 7.89 Hz, H6), 6.74 (1H, s, H3). ^{13}C NMR (126 MHz, $CDCl_3$) *δ* = 158.02 (C2), 147.55 (C4), 145.55 (C9), 133.13 (C7), 128.08 (C10) 125.97 (C5), 125.00 (C6), 114.30 (C8), 117.05 (C3).

LC/MS under Method B, LC Rt= 1.54 min, ES^- (*m/z*) calculated 178.01, found at [M - H^+] 177.45

HRMS calculated for 177.0010 found at [M - H^+] 177.0013

This is a literature compound and the data is in accordance with literature (Majumdar and Ghosh 2002).

Synthesis of 4 – amino – coumarin 116



Chemical Formula: $C_9H_7NO_2$
Exact Mass: 161.05
Molecular Weight: 161.16

4 – Hydroxy coumarin (1000 mg, 6.21 mmol) and Ammonium acetate (480 mg, 6.21 mmol) were mixed and directly melted under 110 °C and stirred under reflux for 6 hours. After this period, the mixture was washed with 200 mL methanol. The precipitates were recrystallized using 1: 1 methanol: water and the product was formed. Yield = 861 mg, 86%.

^{1}H NMR (400 MHz, DMSO): δ = 7.99 (1H, d, *J* = 8.05 Hz, H5), 7.49 (1H, t, *J* = 7.34 Hz H7), 7.37 (2H, s, NH_2), 7.31 (2H, t, *J* = 8.13Hz, H6, H8), 5.22 (1H, s, H3). ^{13}C NMR (126 MHz, $CDCl_3$) *δ* = 161.63 (C4), 155.55 (C2), 153.57 (C9), 132.08 (C7), 123.19 (C6) 122.87 (C5), 116.73 (C8), 114.30 (C10), 83.73 (C3).

LC/MS under Method B, LC Rt= 1.54 min, ES^- (*m/z*) calculated 161.05, found at [M - H^+] 160.01.

HRMS calculated for 162.0555 found at [M + H^+] 162.0553

This is a literature compound and the data is in accordance with literature (Hua, Zhang et al. 2016).

Chapter 7 : References

Agarwal, C., et al. (2003). "Silibinin upregulates the expression of cyclin-dependent kinase inhibitors and causes cell cycle arrest and apoptosis in human colon carcinoma HT-29 cells." Oncogene **22**(51): 8271-8282.

Aguilar, B., et al. (2011). "Polyisoprenylation potentiates the inhibition of polyisoprenylated methylated protein methyl esterase and the cell degenerative effects of sulfonyl fluorides." Current cancer drug targets **11**(6): 752-762.

Andreia, M. (2011). "High throughput discovery of heteroaromatic-modifying enzymes allows enhancement of novobiocin selectivity." Chemical Communications **47**(38): 10569-10571.

Asmari, M., et al. (2018). "Thermophoresis for characterizing biomolecular interaction." Methods **146**: 107-119.

Audisio, D., et al. (2014). "Synthesis and antiproliferative activity of novobiocin analogues as potential hsp90 inhibitors." European journal of medicinal chemistry **83**: 498-507.

Bagatell, R. and L. Whitesell (2004). "Altered Hsp90 function in cancer: a unique therapeutic opportunity." Molecular Cancer Therapeutics **3**(8): 1021-1030.

Baird, I. M., et al. (2000). "Repeated dose study of sucralose tolerance in human subjects." Food and Chemical Toxicology **38**: 123-129.

Balendiran, G. K., et al. (2004). "The role of glutathione in cancer." Cell Biochemistry and Function: Cellular biochemistry and its modulation by active agents or disease **22**(6): 343-352.

Barf, T. and A. Kaptein (2012). "Irreversible protein kinase inhibitors: balancing the benefits and risks." Journal of medicinal chemistry **55**(14): 6243-6262.

Barsnes, H. and M. Vaudel (2018). "SearchGUI: A highly adaptable common interface for proteomics search and de novo engines." Journal of proteome research **17**(7): 2552-2555.

Bauer, R. A. (2015). "Covalent inhibitors in drug discovery: from accidental discoveries to avoided liabilities and designed therapies." Drug discovery today **20**(9): 1061-1073.

Berkov-Zrihen, Y., et al. (2012). "Acylation of novobiocin by carboxylic acid anhydrides: preparation and characterization of semi-synthetic novenamines." Tetrahedron **68**(10): 2306-2312.

Bhat, R., et al. (2014). "Structure–activity studies of (−)-epigallocatechin gallate derivatives as HCV entry inhibitors." Bioorganic & medicinal chemistry letters **24**(17): 4162-4165.

Biebl, M. M. and J. Buchner (2019). "Structure, function, and regulation of the Hsp90 machinery." Cold Spring Harbor perspectives in biology **11**(9): a034017.

Biemann, K. and S. A. Martin (1987). "Mass spectrometric determination of the amino acid sequence of peptides and proteins." Mass Spectrometry Reviews **6**(1): 1-75.

Blagosklonny, M. V., et al. (1996). "Mutant conformation of p53 translated in vitro or in vivo requires functional HSP90." Proceedings of the National Academy of Sciences **93**(16): 8379-8383.

Blume-Jensen, P. and T. Hunter (2001). "Oncogenic kinase signalling." Nature **411**(6835): 355-365.

Burlison, J. A., et al. (2008). "Development of novobiocin analogues that manifest anti-proliferative activity against several cancer cell lines." The Journal of organic chemistry **73**(6): 2130-2137.

Burlison, J. A., et al. (2006). "Novobiocin: redesigning a DNA gyrase inhibitor for selective inhibition of hsp90." Journal of the American Chemical Society **128**(48): 15529-15536.

Byrd, J. C., et al. (2016). "Acalabrutinib (ACP-196) in relapsed chronic lymphocytic leukemia." New England Journal of Medicine **374**(4): 323-332.

Caboni, P., et al. (2004). "Rotenone, deguelin, their metabolites, and the rat model of Parkinson's disease." Chemical research in toxicology **17**(11): 1540-1548.

Campos-Delgado, J., et al. (2008). "Bulk production of a new form of sp2 carbon: crystalline graphene nanoribbons." Nano letters **8**(9): 2773-2778.

Castillo, J. J., et al. (2016). "Inhibition of the bruton tyrosine kinase pathway in B-cell lymphoproliferative disorders." The Cancer Journal **22**(1): 34-39.

Chang, D.-J., et al. (2012). "Design, synthesis, and biological evaluation of novel deguelin-based heat shock protein 90 (HSP90) inhibitors targeting proliferation and angiogenesis." Journal of medicinal chemistry **55**(24): 10863-10884.

Chant, I. D., et al. (1995). "Analysis of heat-shock protein expression in myeloid leukaemia cells by flow cytometry." British journal of haematology **90**(1): 163-168.

Chen, W.-Y., et al. (2008). "Tubocapsenolide A, a novel withanolide, inhibits proliferation and induces apoptosis in MDA-MB-231 cells by thiol oxidation of heat shock proteins." Journal of Biological Chemistry **283**(25): 17184-17193.

Chung, H. S., et al. (2013). "Cysteine oxidative posttranslational modifications: emerging regulation in the cardiovascular system." Circulation research **112**(2): 382-392.

Collins, I. and P. Workman (2006). "New approaches to molecular cancer therapeutics." Nat Chem Biol **2**(12): 689-700.

Cross, D. A., et al. (2014). "AZD9291, an irreversible EGFR TKI, overcomes T790M-mediated resistance to EGFR inhibitors in lung cancer." Cancer discovery **4**(9): 1046-1061.

Cruz, I. N., et al. (2013). "Functional characterization of heat shock protein 90 targeted compounds." Analytical Biochemistry **438**(2): 107-109.

Dai, J., et al. (2019). "Penicisulfuranol A, a novel C-terminal inhibitor disrupting molecular chaperone function of Hsp90 independent of ATP binding domain." Biochemical pharmacology **163**: 404-415.

Davidson, S. M., et al. (2017). "Direct evidence for cancer-cell-autonomous extracellular protein catabolism in pancreatic tumors." Nature medicine **23**(2): 235-241.

Davis, R. E., et al. (2017). "A scaffold merging approach to Hsp90 C-terminal inhibition: synthesis and evaluation of a chimeric library." MedChemComm **8**(3): 593-598.

de Maria, P. D., et al. (2007). "Pig liver esterase (PLE) as biocatalyst in organic synthesis: From nature to cloning and to practical applications." Synthesis **2007**(10): 1439-1452.

Dimas, D. T., et al. (2018). "The prognostic significance of Hsp70/Hsp90 expression in breast cancer: a systematic review and meta-analysis." Anticancer research **38**(3): 1551-1562.

Dong, J., et al. (2014). "Sulfur (VI) fluoride exchange (SuFEx): another good reaction for click chemistry." Angewandte Chemie International Edition **53**(36): 9430-9448.

Donnelly, A. and B. S. Blagg (2008). "Novobiocin and additional inhibitors of the Hsp90 C-terminal nucleotide-binding pocket." Current medicinal chemistry **15**(26): 2702-2717.

Edmen, P. and G. Begg (1967). "A protein sequencer." Eur J Biochem **1**(8091.40).

Elfiky, A., et al. (2008). "BIIB021, an oral, synthetic non-ansamycin Hsp90 inhibitor: Phase I experience." Journal of Clinical Oncology **26**(15_suppl): 2503-2503.

Eng, J. K., et al. (2011). "A face in the crowd: recognizing peptides through database search." Molecular & Cellular Proteomics **10**(11).

Engelman, J. A., et al. (2007). "PF00299804, an irreversible pan-ERBB inhibitor, is effective in lung cancer models with EGFR and ERBB2 mutations that are resistant to gefitinib." Cancer research **67**(24): 11924-11932.

Evans, E. K., et al. (2013). "Inhibition of Btk with CC-292 provides early pharmacodynamic assessment of activity in mice and humans." Journal of Pharmacology and Experimental Therapeutics **346**(2): 219-228.

Fabbro, S. K., et al. (2015). "Panniculitis in patients undergoing treatment with the bruton tyrosine kinase inhibitor ibrutinib for lymphoid leukemias." JAMA oncology **1**(5): 684-686.

Fahey, J. W., et al. (2001). "The chemical diversity and distribution of glucosinolates and isothiocyanates among plants." Phytochemistry **56**(1): 5-51.

Farmer, K., et al. (2012). "KU-32, a novel drug for diabetic neuropathy, is safe for human islets and improves in vitro insulin secretion and viability." Experimental diabetes research **2012**.

Felts, S. J., et al. (2000). "The hsp90-related protein TRAP1 is a mitochondrial protein with distinct functional properties." Journal of Biological Chemistry **275**(5): 3305-3312.

Forsythe, H. L., et al. (2001). "Stable association of hsp90 and p23, but Not hsp70, with active human telomerase." Journal of Biological Chemistry **276**(19): 15571-15574.

Fry, D. W., et al. (1998). "Specific, irreversible inactivation of the epidermal growth factor receptor and erbB2, by a new class of tyrosine kinase inhibitor." Proceedings of the National Academy of Sciences **95**(20): 12022-12027.

Fukami, T. and T. Yokoi (2012). "The emerging role of human esterases." Drug metabolism and pharmacokinetics: DMPK-12-RV-042.

Furman, R. R., et al. (2014). "Ibrutinib resistance in chronic lymphocytic leukemia." The New England journal of medicine **370**(24): 2352.

Futami, J., et al. (2000). "Stabilization of human RNase 1 by introduction of a disulfide bond between residues 4 and 118." The Journal of Biochemistry **128**(2): 245-250.

Galanski, M. (2006). "Recent developments in the field of anticancer platinum complexes." Recent patents on anti-cancer drug discovery **1**(2): 285-295.

Garcia, J., et al. (2010). "Concise modular asymmetric synthesis of deguelin, tephrosin and investigation into their mode of action." Chemistry–A European Journal **16**(32): 9767-9771.

Garg, G., et al. (2016). Anticancer inhibitors of Hsp90 function: beyond the usual suspects. Advances in cancer research, Elsevier. **129:** 51-88.

Gazak, R., et al. (2007). "Silybin and silymarin-new and emerging applications in medicine." Current medicinal chemistry **14**(3): 315-338.

Gazdar, A. (2009). "Activating and resistance mutations of EGFR in non-small-cell lung cancer: role in clinical response to EGFR tyrosine kinase inhibitors." Oncogene **28**(1): S24-S31.

Ghalehshahi, H. G., et al. (2019). "Synthesis, CYP 450 evaluation, and docking simulation of novel 4-aminopyridine and coumarin derivatives." Archiv der Pharmazie **352**(3): 1800247.

Ghosh, A. K., et al. (2019). "Covalent inhibition in drug discovery." ChemMedChem **14**(9): 889.

Grad, I. and D. Picard (2007). "The glucocorticoid responses are shaped by molecular chaperones." Molecular and cellular endocrinology **275**(1): 2-12.

Grbovic, O., et al. (2006). "V600E B-Raf requires the Hsp90 chaperone for stability and is degraded in response to Hsp90 inhibitors." Proceedings of the National Academy of Sciences of the United States of America **103**(1): 57-62.

Grimster, N. P., et al. (2013). "Aromatic sulfonyl fluorides covalently kinetically stabilize transthyretin to prevent amyloidogenesis while affording a fluorescent conjugate." Journal of the American Chemical Society **135**(15): 5656-5668.

Hall, J. A., et al. (2016). "Novobiocin analogues that inhibit the MAPK pathway." J. Med. Chem **59**(3): 925-933.

Hanahan, D. and R. A. Weinberg "The Hallmarks of Cancer." Cell **100**(1): 57-70.

Hanahan, D. and R. A. Weinberg (2011). "Hallmarks of cancer: the next generation." cell **144**(5): 646-674.

Harris, P. J., et al. (1984). "An improved procedure for the methylation analysis of oligosaccharides and polysaccharides." Carbohydrate Research **127**(1): 59-73.

Harris, S. F., et al. (2004). "The crystal structure of the carboxy-terminal dimerization domain of htpG, the Escherichia coli Hsp90, reveals a potential substrate binding site." Structure **12**(6): 1087-1097.

Hartl, F. U. (1996). "Molecular chaperones in cellular protein folding." Nature **381**(6583): 571.

Hett, E. C., et al. (2015). "Rational targeting of active-site tyrosine residues using sulfonyl fluoride probes." ACS Chemical Biology **10**(4): 1094-1098.

Hinman, J., et al. (1957). "The Structure of Novobiocin1, 2." Journal of the American Chemical Society **79**(14): 3789-3800.

Hong, D. S., et al. (2013). "Targeting the molecular chaperone heat shock protein 90 (HSP90): lessons learned and future directions." Cancer treatment reviews **39**(4): 375-387.

Hua, C.-j., et al. (2016). "High quantum yield and pH sensitive fluorescence dyes based on coumarin derivatives: fluorescence characteristics and theoretical study." RSC advances **6**(54): 49221-49227.

Huang, X. (2018). Development of a combined chemical biology tool using synthetic probes and tandem mass spectrometry to elucidate heat shock protein 90 KDA's C-terminal-domain binding sites, University College London.

Ikawa, M., et al. (1944). "Studies on 4-Hydroxycoumarins. V. The Condensation of α, β-Unsaturated Ketones with 4-Hydroxycoumarin1." Journal of the American Chemical Society **66**(6): 902-906.

Iragavarapu, C., et al. (2015). "Novel ALK inhibitors in clinical use and development." Journal of hematology & oncology **8**(1): 17.

ITOH, H., et al. (1999). "A novel chaperone-activity-reducing mechanism of the 90-kDa molecular chaperone HSP90." Biochemical Journal **343**(3): 697-703.

Jackson, P. A., et al. (2017). "Covalent modifiers: A chemical perspective on the reactivity of α, β-unsaturated carbonyls with thiols via hetero-Michael addition reactions." Journal of medicinal chemistry **60**(3): 839-885.

Jackson, S. E., et al. (2004). "Hsp90: from structure to phenotype." Nature structural & molecular biology **11**(12): 1152.

Jaeger, A. M. and L. Whitesell (2019). "HSP90: enabler of cancer adaptation." Annual Review of Cancer Biology **3**: 275-297.

Jerabek-Willemsen, M., et al. (2014). "MicroScale Thermophoresis: Interaction analysis and beyond." Journal of Molecular Structure **1077**: 101-113.

Jhaveri, K., et al. (2012). "Advances in the clinical development of heat shock protein 90 (Hsp90) inhibitors in cancers." Biochimica et Biophysica Acta (BBA)-Molecular Cell Research **1823**(3): 742-755.

Jollow, D., et al. (1973). "Acetaminophen-induced hepatic necrosis. II. Role of covalent binding in vivo." Journal of Pharmacology and Experimental Therapeutics **187**(1): 195-202.

Jordan, P. and M. Carmo-Fonseca (2000). "Molecular mechanisms involved in cisplatin cytotoxicity." Cellular and Molecular Life Sciences CMLS **57**(8-9): 1229-1235.

K Hallenbeck, K., et al. (2017). "Targeting non-catalytic cysteine residues through structure-guided drug discovery." Current topics in medicinal chemistry **17**(1): 4-15.

Karas, M. and F. Hillenkamp (1988). "Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons." Analytical chemistry **60**(20): 2299-2301.

Karras, G. I., et al. (2017). "HSP90 shapes the consequences of human genetic variation." Cell **168**(5): 856-866. e812.

Kawazu, K. (1980). "Isolation of Vibsanines A, B, C, D, E and F from Viburnum odoratissimum." Agricultural and Biological Chemistry **44**(6): 1367-1372.

Khandelwal, A., et al. (2013). "Synthesis and structure–activity relationships of EGCG analogues, a recently identified Hsp90 inhibitor." The Journal of organic chemistry **78**(16): 7859-7884.

Kim, Y. M. and S. Yu (2003). "Palladium (0)-catalyzed amination, Stille coupling, and Suzuki coupling of electron-deficient aryl fluorides." Journal of the American Chemical Society **125**(7): 1696-1697.

King, J. L. and T. H. Jukes (1969). "Non-darwinian evolution." Science **164**(3881): 788-798.

Kirana, C., et al. (2003). "Antitumor activity of extract of Zingiber aromaticum and its bioactive sesquiterpenoid zerumbone." Nutrition and cancer **45**(2): 218-225.

Klenkar, J. and M. Molnar (2015). "Natural and synthetic coumarins as potential anticancer agents." Journal of Chemical and Pharmaceutical Research **7**(7): 1223-1238.

Klingler, F.-M., et al. (2015). "Approved drugs containing thiols as inhibitors of metallo-β-lactamases: strategy to combat multidrug-resistant bacteria." Journal of medicinal chemistry **58**(8): 3626-3630.

Knowles, J. and G. Gromo (2003). "Target selection in drug discovery." Nature Reviews Drug Discovery **2**(1): 63-69.

Kola, I. and J. Landis (2004). "Can the pharmaceutical industry reduce attrition rates?" Nature reviews Drug discovery **3**(8): 711-716.

Konstat-Korzenny, E., et al. (2018). "Artemisinin and its synthetic derivatives as a possible therapy for cancer." Medical Sciences **6**(1): 19.

Kozloff, M. F., et al. (2012). "Phase I trial of axitinib combined with platinum doublets in patients with advanced non-small cell lung cancer and other solid tumours." British journal of cancer **107**(8): 1277.

Krishna, P., et al. (1997). "Analysis of the native forms of the 90 kDa heat shock protein (hsp90) in plant cytosolic extracts." Plant molecular biology **33**(3): 457-466.

Krukenberg, K. A., et al. (2008). "Multiple conformations of E. coli Hsp90 in solution: insights into the conformational dynamics of Hsp90." Structure **16**(5): 755-765.

Krutak, J. J., et al. (1979). "Chemistry of ethenesulfonyl fluoride. Fluorosulfonylethylation of organic compounds." The Journal of Organic Chemistry **44**(22): 3847-3858.

Kusuma, B. R., et al. (2014). "Synthesis and biological evaluation of coumarin replacements of novobiocin as Hsp90 inhibitors." Bioorganic & medicinal chemistry **22**(4): 1441-1449.

Lake, B. (1999). "Coumarin metabolism, toxicity and carcinogenicity: relevance for human risk assessment." Food and Chemical Toxicology **37**(4): 423-453.

Lauwers, E., et al. (2018). "Hsp90 mediates membrane deformation and exosome release." Molecular cell **71**(5): 689-702. e689.

Le Bras, G., et al. (2007). "New novobiocin analogues as antiproliferative agents in breast cancer cells and potential inhibitors of heat shock protein 90." Journal of medicinal chemistry **50**(24): 6189-6200.

Lebeau, J., et al. (1991). "Constitutive overexpression of a 89 kDa heat shock protein gene in the HBL100 human mammary cell line converted to a tumorigenic phenotype by the EJ/T24 Harvey-ras oncogene." Oncogene **6**(7): 1125-1132.

Lee, S.-C., et al. (2015). "Synthesis and evaluation of a novel deguelin derivative, L80, which disrupts ATP binding to the C-terminal domain of heat shock protein 90." Molecular pharmacology **88**(2): 245-255.

Leonard, S. E., et al. (2009). "Mining the thiol proteome for sulfenic acid modifications reveals new targets for oxidation in cells." ACS chemical biology **4**(9): 783-799.

Li, C.-F., et al. (2008). "Heat shock protein 90 overexpression independently predicts inferior disease-free survival with differential expression of the α and β isoforms in gastrointestinal stromal tumors." Clinical Cancer Research **14**(23): 7822-7831.

Li, D., et al. (2016). "Natural product kongensin A is a non-canonical HSP90 inhibitor that blocks RIP3-dependent necroptosis." Cell chemical biology **23**(2): 257-266.

Li, J., et al. (2012). "The Hsp90 chaperone machinery: conformational dynamics and regulation by co-chaperones." Biochimica et Biophysica Acta (BBA)-Molecular Cell Research **1823**(3): 624-635.

Li, X., et al. (2014). "Discovery of a series of 2, 5-diaminopyrimidine covalent irreversible inhibitors of Bruton's tyrosine kinase with in vivo antitumor activity." Journal of medicinal chemistry **57**(12): 5112-5128.

Li, Y., et al. (2012). "Sulforaphane inhibits pancreatic cancer through disrupting Hsp90–p50Cdc37 complex and direct interactions with amino acids residues of Hsp90." The Journal of nutritional biochemistry **23**(12): 1617-1626.

Lipshutz, B. H., et al. (2012). "Catalysis in the service of green chemistry: Nobel prize-winning palladium-catalysed cross-couplings, run in water at room temperature: Heck, Suzuki-miyaura and Negishi reactions carried out in the absence of organic solvents, enabled by micellar catalysis." Platinum metals review **56**(2): 62.

Lonsdale, R. and R. A. Ward (2018). "Structure-based design of targeted covalent inhibitors." Chemical Society Reviews **47**(11): 3816-3830.

Ma, Y., et al. (2019). "Application of Dually Activated Michael Acceptor to the Rational Design of Reversible Covalent Inhibitor for Enterovirus 71 3C Protease." Journal of medicinal chemistry **62**(13): 6146-6162.

Majumdar, K. and S. Ghosh (2002). "Studies of bioactive heterocycles: facile thio-Claisen rearrangement of propargylthio [1] benzopyran-2-ones." Tetrahedron letters **43**(11): 2115-2117.

Mandler, M. D., et al. (2018). "Novobiocin enhances polymyxin activity by stimulating lipopolysaccharide transport." Journal of the American Chemical Society **140**(22): 6749-6753.

Marcu, M. G., et al. (2000). "The heat shock protein 90 antagonist novobiocin interacts with a previously unrecognized ATP-binding domain in the carboxyl terminus of the chaperone." Journal of Biological Chemistry **275**(47): 37181-37186.

Margarucci, L., et al. (2013). "Chemical proteomics-driven discovery of oleocanthal as an Hsp90 inhibitor." Chemical communications **49**(52): 5844-5846.

Marshall, M., et al. (1994). "An updated review of the clinical development of coumarin (1, 2-benzopyrone) and 7-hydroxycoumarin." Journal of cancer research and clinical oncology **120**(1): S39-S42.

Martínez-Ruiz, A., et al. (2005). "S-nitrosylation of Hsp90 promotes the inhibition of its ATPase and endothelial nitric oxide synthase regulatory activities." Proceedings of the National Academy of Sciences **102**(24): 8525-8530.

Matsumoto, K., et al. (2004). "Antinociceptive effect of 7-hydroxymitragynine in mice: Discovery of an orally active opioid analgesic from the Thai medicinal herb Mitragyna speciosa." Life sciences **74**(17): 2143-2155.

Matz, J. M., et al. (1995). "Characterization and regulation of cold-induced heat shock protein expression in mouse brown adipose tissue." American Journal of Physiology-Regulatory, Integrative and Comparative Physiology **269**(1): R38-R47.

Mayer, M. P., et al. (2009). "The Hsp90 mosaic: a picture emerges." Nature structural & molecular biology **16**(1): 2.

McNaught, A. D. and A. Wilkinson (1997). Compendium of chemical terminology, Blackwell Science Oxford.

Miller, K. D., et al. (2020). "Cancer statistics for adolescents and young adults, 2020." CA: A Cancer Journal for Clinicians.

Miller, V. A., et al. (2012). "Afatinib versus placebo for patients with advanced, metastatic non-small-cell lung cancer after failure of erlotinib, gefitinib, or both, and one or two lines of chemotherapy (LUX-Lung 1): a phase 2b/3 randomised trial." The lancet oncology **13**(5): 528-538.

Mimnaugh, E. G., et al. (1996). "Polyubiquitination and proteasomal degradation of the p185c-erbB-2 receptor protein-tyrosine kinase induced by geldanamycin." Journal of Biological Chemistry **271**(37): 22796-22801.

Mojarradi, H. (2010). Coupling of substances containing a primary amine to hyaluronan via carbodiimide-mediated amidation.

Mok, T., et al. (2017). "Osimertinib-Therapie bei EGFR-T790M-Punktmutation beim NSCLC." N Engl J Med **376**: 629-640.

Mollapour, M. and L. Neckers (2012). "Post-translational modifications of Hsp90 and their contributions to chaperone regulation." Biochimica et Biophysica Acta (BBA)-Molecular Cell Research **1823**(3): 648-655.

Moriarty, R. M. and S. Tyagi (2010). "Metal-free intramolecular aziridination of alkenes using hypervalent iodine based sulfonyliminoiodanes." Organic Letters **12**(2): 364-366.

Moulick, K., et al. (2011). "Affinity-based proteomics reveal cancer-specific networks coordinated by Hsp90." Nature chemical biology **7**(11): 818-826.

Murakami, A., et al. (1999). "Identification of zerumbone in Zingiber zerumbet Smith as a potent inhibitor of 12-O-tetradecanoylphorbol-13-acetate-induced Epstein-Barr virus activation." Bioscience, biotechnology, and biochemistry **63**(10): 1811-1812.

Mustafa, A., et al. (1956). "Reactions with Mercaptans. III. 1 Action of Aromatic Thiols on Coumarins, 4-Styrylcoumarins and 2-Styrylchromones." Journal of the American Chemical Society **78**(19): 5011-5016.

Nakamoto, H., et al. (2018). "Stimulation of the ATPase activity of Hsp90 by zerumbone modification of its cysteine residues destabilizes its clients and causes cytotoxicity." Biochemical Journal **475**(15): 2559-2576.

Nanbu, K., et al. (1996). "Expression of heat shock proteins HSP70 and HSP90 in endometrial carcinomas: correlation with clinicopathology, sex steroid receptor status, and p53 protein expression." Cancer **77**(2): 330-338.

Nardai, G., et al. (2000). "Reactive cysteines of the 90-kDa heat shock protein, Hsp90." Archives of biochemistry and biophysics **384**(1): 59-67.

Nathan, D. F. and S. Lindquist (1995). "Mutational analysis of Hsp90 function: interactions with a steroid receptor and a protein kinase." Molecular and cellular biology **15**(7): 3917-3925.

Neckers, L. and P. Workman (2012). "Hsp90 molecular chaperone inhibitors: are we there yet?" Clinical cancer research **18**(1): 64-76.

Nemoto, T., et al. (1995). "Mechanism of dimer formation of the 90-kDa heat-shock protein." European journal of biochemistry **233**(1): 1-8.

Novero, A., et al. (2014). "Ibrutinib for B cell malignancies." Experimental Hematology & Oncology **3**(1): 4.

Oh, S. H., et al. (2007). "Structural basis for depletion of heat shock protein 90 client proteins by deguelin." Journal of the National Cancer Institute **99**(12): 949-961.

Olbe, L., et al. (2003). "A proton-pump inhibitor expedition: the case histories of omeprazole and esomeprazole." Nature reviews drug discovery **2**(2): 132-139.

Palermo, C. M., et al. (2005). "Epigallocatechin gallate inhibits aryl hydrocarbon receptor gene transcription through an indirect mechanism involving binding to a 90 kDa heat shock protein." Biochemistry **44**(13): 5041-5052.

Parker, C. E., et al. (2010). "Mass spectrometry for post-translational modifications." Neuroproteomics **2010**: PMID: 21882444.

Patel, S. M., et al. (2011). "High throughput discovery of heteroaromatic-modifying enzymes allows enhancement of novobiocin selectivity." Chemical Communications **47**(38): 10569-10571.

Glycosylated analogues of novobiocin, discovered using a broad library of enzymes, have 100-fold improved activity against breast, brain, pancreatic, lung and ovarian cancers and ablated associated off-target activity leading to an up to 2.7 [times] 104 fold increase in selectivity.

Pearl, L. H. (2005). "Hsp90 and Cdc37–a chaperone cancer conspiracy." Current opinion in genetics & development **15**(1): 55-61.

Pearl, L. H., et al. (2008). "The Hsp90 molecular chaperone: an open and shut case for treatment." Biochemical Journal **410**(3): 439-453.

Pellati, F. and G. Rastelli (2016). "Novel and less explored chemotypes of natural origin for the inhibition of Hsp90." MedChemComm **7**(11): 2063-2075.

Pompella, A., et al. (2003). "The changing faces of glutathione, a cellular protagonist." Biochemical pharmacology **66**(8): 1499-1503.

Requejo, R., et al. (2010). "Cysteine residues exposed on protein surfaces are the dominant intramitochondrial thiol and may protect against oxidative damage." The FEBS journal **277**(6): 1465-1480.

Retzlaff, M., et al. (2009). "Hsp90 is regulated by a switch point in the C-terminal domain." EMBO reports **10**(10): 1147-1153.

Richardson, R. T., et al. (2006). "Nuclear autoantigenic sperm protein (NASP), a linker histone chaperone that is required for cell proliferation." Journal of Biological Chemistry **281**(30): 21526-21534.

Riebold, M., et al. (2015). "A C-terminal HSP90 inhibitor restores glucocorticoid sensitivity and relieves a mouse allograft model of Cushing disease." Nature medicine **21**(3): 276-280.

Ritossa, F. (1962). "A new puffing pattern induced by temperature shock and DNP in Drosophila." Cellular and Molecular Life Sciences **18**(12): 571-573.

Rosenhagen, M. C., et al. (2003). "The heat shock protein 90-targeting drug cisplatin selectively inhibits steroid receptor activation." Molecular Endocrinology **17**(10): 1991-2001.

Roth, G. J., et al. (1975). "Acetylation of prostaglandin synthase by aspirin." Proceedings of the National Academy of Sciences **72**(8): 3073-3076.

Ruiz, M. I. G., et al. (2008). "Integration of gene dosage and gene expression in non-small cell lung cancer, identification of HSP90 as potential target." PloS one **3**(3): e0001722.

Rutherford, S. L. and S. Lindquist (1998). "Hsp90 as a capacitor for morphological evolution." Nature **396**(6709): 336-342.

Sato, S., et al. (2000). "Modulation of Akt kinase activity by binding to Hsp90." Proceedings of the National Academy of Sciences **97**(20): 10832-10837.

Savi, P., et al. (2000). "Identification and biological activity of the active metabolite of clopidogrel." Thrombosis and haemostasis **84**(11): 891-896.

Schneider, T. H., et al. (2015). "Vinyl sulfone building blocks in covalently reversible reactions with thiols." New journal of chemistry **39**(7): 5841-5853.

Schwartz, P. A., et al. (2014). "Covalent EGFR inhibitor analysis reveals importance of reversible interactions to potency and mechanisms of drug resistance." Proceedings of the National Academy of Sciences **111**(1): 173-178.

Seidel, S. A., et al. (2013). "Microscale thermophoresis quantifies biomolecular interactions under previously challenging conditions." Methods **59**(3): 301-315.

Sequist, L. V., et al. (2010). "Neratinib, an irreversible pan-ErbB receptor tyrosine kinase inhibitor: Results of a phase II trial in patients with advanced non–small-cell lung cancer." Journal of clinical oncology **28**(18): 3076-3083.

Shao, L.-D., et al. (2017). "Design, synthesis, and biological activities of vibsanin B derivatives: a new class of HSP90 C-terminal inhibitors." Journal of Medicinal Chemistry **60**(21): 9053-9066.

Shiau, A. K., et al. (2006). "Structural analysis of E. coli hsp90 reveals dramatic nucleotide-dependent conformational rearrangements." Cell **127**(2): 329-340.

Sidera, K. and E. Patsavoudi (2014). "HSP90 inhibitors: current development and potential in cancer therapy." Recent patents on anti-cancer drug discovery **9**(1): 1-20.

Simon, B., et al. (2017). "Synthesis and characterization of photoaffinity labelling reagents towards the Hsp90 C-terminal domain." Organic & biomolecular chemistry **15**(7): 1597-1605.

Singh, J., et al. (2011). "The resurgence of covalent drugs." Nature reviews Drug discovery **10**(4): 307-317.

Skrott, Z., et al. (2017). "Alcohol-abuse drug disulfiram targets cancer via p97 segregase adaptor NPL4." Nature **552**(7684): 194-199.

Smith, P. F., et al. (2017). "A phase I trial of PRN1008, a novel reversible covalent inhibitor of Bruton's tyrosine kinase, in healthy volunteers." British journal of clinical pharmacology **83**(11): 2367-2376.

Sofis, M. J., et al. (2017). "KU32 prevents 5-fluorouracil induced cognitive impairment." Behavioural Brain Research **329**: 186-190.

Solit, D. B., et al. (2003). Hsp90 as a therapeutic target in prostate cancer. Seminars in oncology, Elsevier.

Sőti, C., et al. (2005). "Molecular chaperones as regulatory elements of cellular networks." Current opinion in cell biology **17**(2): 210-215.

Söti, C., et al. (2002). "A nucleotide-dependent molecular switch controls ATP binding at the C-Terminal domain of Hsp90 N-terminal nucleotide binding unmasks a C-Terminal binding pocket." Journal of Biological Chemistry **277**(9): 7066-7075.

Spero, L., et al. (1949). "Alkali-sensitive Glycosides of 3-Phenyl-4-hydroxycoumarin." Journal of the American Chemical Society **71**(11): 3740-3743.

Sreedhar, A. S., et al. (2004). "Inhibition of Hsp90: a new strategy for inhibiting protein kinases." Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics **1697**(1-2): 233-242.

Stamos, J., et al. (2002). "Structure of the epidermal growth factor receptor kinase domain alone and in complex with a 4-anilinoquinazoline inhibitor." Journal of Biological Chemistry **277**(48): 46265-46272.

Stepan, A. F., et al. (2013). "Metabolism-guided drug design." MedChemComm **4**(4): 631-652.

Stivarou, T., et al. (2016). "Targeting highly expressed extracellular HSP90 in breast cancer stem cells inhibits tumor growth in vitro and in vivo." Cancer biology & therapy **17**(8): 799-812.

Stockert, J. C., et al. (2018). "Tetrazolium salts and formazan products in Cell Biology: Viability assessment, fluorescence imaging, and labeling perspectives." Acta histochemica **120**(3): 159-167.

Strand, D. S., et al. (2017). "25 years of proton pump inhibitors: a comprehensive review." Gut and liver **11**(1): 27.

Subbarao Sreedhar, A., et al. (2004). "Hsp90 isoforms: functions, expression and clinical importance." FEBS letters **562**(1-3): 11-15.

Sun, G., et al. (2017). "Regioselective glycosylation of novobiocin alters activity." Carbohydrate research **452**: 116-121.

Sun, Y., et al. (2010). "Differential proteomics identification of HSP90 as potential serum biomarker in hepatocellular carcinoma by two-dimensional electrophoresis and mass spectrometry." International journal of molecular sciences **11**(4): 1423-1433.

Swinney, D. C. (2004). "Biochemical mechanisms of drug action: what does it take for success?" Nature reviews Drug discovery **3**(9): 801-808.

Swinney, D. C. (2006). "Biochemical mechanisms of new molecular entities (NMEs) approved by United States FDA during 2001-2004: mechanisms leading to optimal efficacy and safety." Current topics in medicinal chemistry **6**(5): 461-478.

Terracciano, S., et al. (2013). "Dimeric and trimeric triazole based molecules as a new class of Hsp90 molecular chaperone inhibitors." European journal of medicinal chemistry **65**: 464-476.

Terracciano, S., et al. (2016). "New dihydropyrimidin-2 (1 H)-one based Hsp90 C-terminal inhibitors." RSC advances **6**(85): 82330-82340.

Thompson, V. R. and A. P. DeCaprio (2013). "Covalent adduction of nitrogen mustards to model protein nucleophiles." Chemical research in toxicology **26**(8): 1263-1271.

Thress, K. S., et al. (2015). "Acquired EGFR C797S mutation mediates resistance to AZD9291 in non–small cell lung cancer harboring EGFR T790M." Nature medicine **21**(6): 560-562.

Trepel, J., et al. (2010). "Targeting the dynamic HSP90 complex in cancer." Nature reviews. Cancer **10**(8): 537.

Vaudel, M., et al. (2015). "PeptideShaker enables reanalysis of MS-derived proteomics data sets." Nature biotechnology **33**(1): 22-24.

Veber, D. F., et al. (2002). "Molecular properties that influence the oral bioavailability of drug candidates." Journal of medicinal chemistry **45**(12): 2615-2623.

Verba, K. A. and D. A. Agard (2017). "How Hsp90 and Cdc37 lubricate kinase molecular switches." Trends in biochemical sciences **42**(10): 799-811.

Walter, A. O., et al. (2013). "Discovery of a mutant-selective covalent inhibitor of EGFR that overcomes T790M-mediated resistance in NSCLC." Cancer discovery **3**(12): 1404-1415.

Walton-Diaz, A., et al. (2013). "Contributions of co-chaperones and post-translational modifications towards Hsp90 drug sensitivity." Future medicinal chemistry **5**(9): 1059-1071.

Waxman, D. J. and J. L. Strominger (1983). "Penicillin-binding proteins and the mechanism of action of beta-lactam antibiotics." Annual review of biochemistry **52**(1): 825-869.

Whitesell, L. and S. Lindquist (2009). "Inhibiting the transcription factor HSF1 as an anticancer strategy." Expert opinion on therapeutic targets **13**(4): 469-478.

Whitesell, L. and S. L. Lindquist (2005). "HSP90 and the chaperoning of cancer." Nature reviews. Cancer **5**(10): 761.

Wienken, C. J., et al. (2010). "Protein-binding assays in biological liquids using microscale thermophoresis." Nature communications **1**: 100.

Workman, P., et al. (2007). "Drugging the cancer chaperone HSP90." Annals of the New York Academy of Sciences **1113**(1): 202-216.

Wu, G., et al. (2019). "Identification of HSP90 as a direct target of artemisinin for its anti-inflammatory activity via quantitative chemical proteomics." Organic & biomolecular chemistry **17**(28): 6854-6859.

Wu, H., et al. (2017). "Irreversible inhibition of BTK kinase by a novel highly selective inhibitor CHMFL-BTK-11 suppresses inflammatory response in rheumatoid arthritis model." Scientific reports **7**(1): 1-10.

Wu, J., et al. (2016). "Second-generation inhibitors of Bruton tyrosine kinase." Journal of Hematology & Oncology **9**(1): 80.

Xiao, L., et al. (2006). "Effectiveness of hsp90 inhibitors as anti-cancer drugs." Mini reviews in medicinal chemistry **6**(10): 1137-1143.

Yang, N. J. and M. J. Hinner (2015). Getting across the cell membrane: an overview for small molecules, peptides, and proteins. Site-specific protein labeling, Springer**:** 29-53.

Yasuhiro, T., et al. (2017). "Anti-tumor efficacy study of the Bruton's tyrosine kinase (BTK) inhibitor, ONO/GS-4059, in combination with the glycoengineered type II anti-CD20 monoclonal antibody obinutuzumab (GA101) demonstrates superior in vivo efficacy compared to ONO/GS-4059 in combination with rituximab." Leukemia & lymphoma **58**(3): 699-707.

Ye, B.-X., et al. (2015). "Vibsanin B preferentially targets HSP90β, inhibits interstitial leukocyte migration, and ameliorates experimental autoimmune encephalomyelitis." The Journal of Immunology **194**(9): 4489-4497.

Yin, Z., et al. (2009). "(−)-Epigallocatechin-3-gallate is a novel Hsp90 inhibitor." Biochemistry **48**(2): 336-345.

Yodkeeree, S., et al. (2009). "Zerumbone enhances TRAIL-induced apoptosis through the induction of death receptors in human colon cancer cells: Evidence for an essential role of reactive oxygen species." Cancer research **69**(16): 6581-6589.

Yokoyama, Y., et al. (2015). "Goniothalamin enhances the ATPase activity of the molecular chaperone Hsp90 but inhibits its chaperone activity." The Journal of Biochemistry **157**(3): 161-168.

Yoshimura, N., et al. (2006). "EKB-569, a new irreversible epidermal growth factor receptor tyrosine kinase inhibitor, with clinical activity in patients with non-small cell lung cancer with acquired resistance to gefitinib." Lung Cancer **51**(3): 363-368.

Young, J. C., et al. (1997). "In vitro evidence that hsp90 contains two independent chaperone sites." FEBS letters **418**(1-2): 139-143.

Yu, M., et al. (2010). "MPC-3100, a fully synthetic, orally bioavailable Hsp90 inhibitor, in cancer patients." Journal of Clinical Oncology **28**(15_suppl): e13112-e13112.

Zhang, R., et al. (2019). "Improving cellular uptake of therapeutic entities through interaction with components of cell membrane." Drug delivery **26**(1): 328-342.

Zhang, S., et al. (2019). "High expression of HSP90 is associated with poor prognosis in patients with colorectal cancer." PeerJ **7**: e7946.

Zhang, Y., et al. (2014). "Sulphoxythiocarbamates modify cysteine residues in HSP90 causing degradation of client proteins and inhibition of cancer cell proliferation." British journal of cancer **110**(1): 71-82.

Zhao, H. and B. S. Blagg (2013). "Inhibitors of the Hsp90 C-terminus." Inhibitors of Molecular Chaperones as Therapeutic Agents: 259-301.

Zhao, H., et al. (2011). "Identification and initial SAR of silybin: an Hsp90 inhibitor." Bioorganic & medicinal chemistry letters **21**(9): 2659-2664.

Zhao, H., et al. (2011). "Engineering an antibiotic to fight cancer: optimization of the novobiocin scaffold to produce anti-proliferative agents." Journal of medicinal chemistry **54**(11): 3839-3853.

Zhao, H., et al. (2015). "Design, synthesis and biological evaluation of biphenylamide derivatives as Hsp90 C-terminal inhibitors." European journal of medicinal chemistry **89**: 442-466.

Zhao, H., et al. (2012). "3-Arylcoumarin derivatives manifest anti-proliferative activity through Hsp90 inhibition." ACS medicinal chemistry letters **3**(4): 327-331.