SYNTHESIS AND BIOLOGICAL EVALUATION OF HISTONE LYSINE METHYLTRANSFERASE INHIBITORS

A thesis submitted by

Fanny Cherblanc

In partial fulfilment of the requirements for the degree of

Doctor of Philosophy

Department of Chemistry
Imperial College London
South Kensington
London
SW7 2AZ
United Kingdom

September 2013

ABSTRACT

The epigenetic component of cancer is now well-established and epigenetic therapeutics targeting histone deacetylase and DNA methyltransferase enzymes have received FDA approval. Histone lysine methyltransferases (HKMTs) are a new and promising epigenetic target class for cancer treatment, but the search for HKMT inhibitors is still in its infancy.

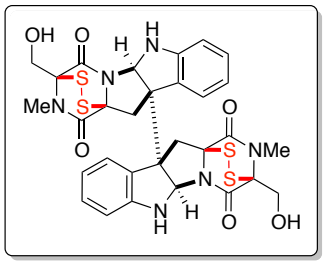


Figure 1 Chaetocin

First, chaetocin, a fungal metabolite belonging to the 3,6-epidithio-diketopiperazine (ETP) class, was previously described as a specific inhibitor of the HKMT SU(VAR)3-9. Like other ETP molecules however, it exhibits a broad cytotoxicity due to the presence of the disulfide bridge. The requirement of this moiety for HKMT inhibition has been questioned. We reasoned that access to semi-synthetic analogues devoid of such functionality would provide valuable insight into its reported HKMT inhibitory activity. Our results revealed a total loss in inhibitory potency upon modification of the disulfide bridge. In addition, we conducted biochemical and mass spectroscopy studies on the mechanism of action of chaetocin. Our results demonstrated the ETP functionality of chaetocin to be responsible for its reported HKMT inhibitory activity and suggest that chaetocin is a non-specific inhibitor.

The stereochemical course of ETP desulfurisation by triphenylphosphine has also been investigated by comparison of experimental and computationally simulated chiroptical spectra. Our data showed that all ETPs are desulfurised with retention of stereochemistry.

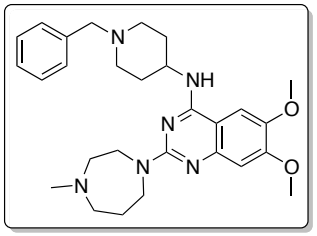


Figure 2 BIX-01294

Second, BIX-01294 has been identified as an inhibitor of G9a/GLP HKMTs. In light of the expected structural similarities of EZH2's peptide binding site with that of G9a, we felt we could use this molecule as a starting point for the reversal of H3K27-mediated gene silencing.

We anticipated that upon derivatisation of BIX-01294, it might be possible to expand the biological activities of the parent compound and gain EZH2 inhibitory activity. A library of analogues has been prepared in collaboration with co-workers in the Fuchter group and has been assessed in biological assays. Three compounds have been identified in a cell-based assay to lead to re-expression of EZH2 target genes and their biological activities have been further assessed.

TABLE OF CONTENTS

ABSTRACT 3

DECLARATION OF ORIGINALITY 7

COPYRIGHT DECLARATION 9

ACKNOWLEDGEMENTS 11

ABBREVIATIONS 15

CHAPTER I INTRODUCTION 19

CANCER AND DRUG DISCOVERY 21

EPIGENETICS AND THERAPY 23

EPIGENETIC MODIFICATIONS 23

EPIGENETIC THERAPY FOR CANCER 27

GLOBAL OBJECTIVES OF THE PROJECT 37

CHAPTER II DESIGN, SYNTHESIS AND BIOLOGICAL EVALUATION OF CHAETOCIN ANALOGUES 39

INTRODUCTION 41

RESULTS AND DISCUSSION 49

SYNTHESIS AND DERIVATISATION OF A MODEL ETP COMPOUND 49

ISOLATION OF CHAETOCIN AND SEMI-SYNTHESIS OF ANALOGUES 52

BIOLOGICAL EVALUATION OF CHAETOCIN AND ANALOGUES 59

CONCLUSIONS 95

CHAPTER III STEREOCHEMISTRY OF ETP DESULFURISATION 101

INTRODUCTION 103

RESULTS AND DISCUSSION 111

CHAETOCIN 111

DEHYDROGLIOTOXIN 117

GLIOTOXIN ANALOGUE 119

CONCLUSIONS 137

CHAPTER IV TOWARDS THE DISCOVERY OF NOVEL EZH2 INHIBITORS 139

INTRODUCTION 141

RESULT AND DISCUSSION **151**
SYNTHESIS OF BIX-01294 ANALOGUES 151
BIOLOGICAL EVALUATION 159
CONCLUSIONS **161**

CONCLUSIONS **163**

EXPERIMENTAL DETAILS **167**

COMPOUND SYNTHESIS **169**
MATERIALS AND METHODS 169
SYNTHETIC PROCEDURES AND COMPOUND CHARACTERISATION 171
CHIROPTICAL MEASUREMENTS 207
COMPUTATIONAL SIMULATIONS **208**
BIOLOGICAL EVALUATION **209**
BIOCHEMICAL ASSAYS 209
MASS SPECTROMETRY EXPERIMENTS 214
CELL-BASED ASSAY 217
CRYSTALLOGRAPHIC DATA **219**

REFERENCES **233**

DECLARATION OF ORIGINALITY

I, Fanny Cherblanc, hereby confirm that I conducted the research presented in this thesis under the supervision of Doctor Matthew J. Fuchter at the Department of Chemistry, Imperial College London, and that the work described herein is entirely my own unless otherwise stated.

London, 18th September 2013

Fanny Cherblanc

COPYRIGHT DECLARATION

The copyright of this thesis rests with the author and is made available under a Creative Commons Attribution Non-Commercial No Derivatives licence. Researchers are free to copy, distribute or transmit the thesis on the condition that they attribute it, that they do not use it for commercial purposes and that they do not alter, transform or build upon it. For any reuse or redistribution, researchers must make clear to others the licence terms of this work.

ACKNOWLEDGEMENTS

I would like to thank Dr. Matthew Fuchter for giving me the opportunity to perform my doctoral studies within his research group and for his guidance, support and his enthusiasm both for my research and for the group socials throughout the past four years.

I would like to express my gratitude to the analytical staff: Peter Haycock and Richard Sheppard (NMR), Jonathan Barton and Lisa Haigh (mass spectrometry) and Andrew White (crystallography) for their help and advice in all (challenging) situations.

The projects I have been involved with could not have existed without the many valuable collaborators we have had over the years: Dr. Laura Alcazar Fuoli and Dr. Elaine Bignell for their great help with the production of chaetocin; Dr. Nadine Chapman-Rothe, Elham Shamsaei and Prof. Robert Brown for designing and running the cellular assays; Dr. Katie Chapman for her guidance throughout the development of the AlphaLISA assay, her priceless advice and the comforting coffees when my experiments were going wrong; Prof. Henry Rzepa and Ya-Pei Lo for running the numerous calculations; Prof. Patrick Bultinck and Prof. Wouter Herrebout at the European Centre for Chirality for the VCD measurements; Dr. Marina Demetriades, Prof. Christopher Schofield, Dr. Pete DiMaggio and Aaron Borg for the mass spectrometry data which completed elegantly our work around chaetocin.

Four years is a long time, but the presence of a number of people on the 7th floor made them pass by much quicker… Thanks to my dear labmates for supporting me, my musical tastes and my variable temperament throughout the years: Jean-Noel, Marko, Paolo, Rob and Jez. Phil Collins still rocks! Thanks to the EZH2 team for keeping this challenging project going: Ken, Tom, Jojo, Sandeep and the *little* Thom.

I should not forget to thanks all the members of the Fuchter and Barrett Group who helped me throughout these years, at all levels including synthesis advice, grabbing (a) well-deserved beer(s) on a Friday evening (or any other day) at *late* HC or elsewhere, sharing a comforting coffee break, or just chatting through a door for an hour or while running a column… JN, Bhav, Max, Matthias, Fred, Flo, Jenny, Thais, Oki, Rob, Pete, Lauriane, Jojo, Thomas… Thank you guys (and *mes copines*), it was nice having you around, really!

Also, I want to give special thanks to those who kindly proofread this thesis: Thomas, Rob, Pete, Kate and Claire.

Finally I have to thank my parents for their special support throughout the years, and for providing me an amazing *place to live* while writing this thesis.

And last but not least, I deeply want to thank the one person you have shared my life for the past three years, for his limitless support throughout the PhD, his patience and everyday help during the writing, for reassuring me during my darkest moments but also for sharing my happiness… *Merci Chéri.*

"There has been an alarming increase in the number of things that I know nothing about."

Winnie the Pooh (A.A. Milne), 1926

ABBREVIATIONS

Δ	enantiomeric similarity index
Σ	similarity measure
Alpha	amplified luminescent proximity homogeneous assay
aq.	aqueous
br	broad
˚C	degrees Celsius
c	concentration
calc.	calculated
CH_2Cl_2	dichloromethane
CI	chemical ionisation
Cys	cysteine residue
d	doublet
d	*Drosophila melanogaster*
DDQ	2,3-dichloro-5,6-dicyano-1,4-benzoquinone
DIAD	diisopropyl azodicarboxylate
DIEA	*N,N*-diidopropylethylamine
DKP	diketopiperazine
DMA	dimethylacetamide
DMF	dimethylformamide
DMSO	dimethylsulfoxide
DNMT	DNA methyltransferase
DNA	deoxyribonucleic acid
DTT	dithiothreitol
DZNep	3-deazaneplanocin A
ECD	electronic circular dichroism
eq.	equivalent
ESI	electrospray ionisation
EtOH	ethanol
Et_2O	diethyl ether
EtOAc	ethylacetate
ETP	3,6-epidithio-diketopiperazine

EZH2	enhancer of zest homologue 2
FDA	food and drug administration
GLP	G9a like protein
H*x*K*y*me*z*	histone *x* lysine *y* methylated *z* times
HCl	hydrochloric acid
HDAC	histone deacetylase
HKMT	histone lysine methyltransferase
HMT	histone methyltransferase
HPLC	high performance liquid chromatography
HRMS	high resolution mass spectroscopy
HRP	horseradish peroxidase
HTS	high throughput screen
Hz	hertz
IC_{50}	half maximal inhibitory concentration
J	coupling constant
K	lysine residue
LC-MS	liquid chromatography-mass spectrometry
μ	micro (10^{-6})
m	multiplet
MBP	methyl-binding protein
MeCN	acetonitrile
MeOH	methanol
MLL	mixed lineage leukaemia
MS	mass spectroscopy
Ms	mesylate (methylsulfonate)
m/z	mass to charge ratio
n	normal or number of independent experiments
N	normality (acid or base concentration)
NaOH	sodium hydroxide
NBS	*N*-bromosuccinimide
NMR	nuclear magnetic resonance
NOESY	nuclear Overhauser effect spectroscopy
p	*para*

pX	-log(X)
PDB	protein data bank
Pd/C	palladium on charcoal
PMB	para methoxy benzyl
ppm	parts per million
PPY	(*R*)-(+)-4-pyrrolidinopyridinyl (pentamethylcyclopentadienyl)iron
PRC2	polycomb repressive complex 2
PRMT	protein arginine methyltransferase
quant.	quantitative
R	general substituent
RNA	ribonucleic acid
RT	room temperature
RT-PCR	real-time polymerase chain reaction
s	singlet
SAH	*S*-adenosylhomocysteine
SAHA	suberoylanilide hydroxamic acid
SAM	*S*-adenosylmethionine
SAR	structure-activity relationship
sat.	saturated
SD	standard deviation
SEM	standard error of the mean
SET	suvar3-9, enhancer-of-zeste, trithorax
SUV39	suppressor of variegation 3-9
t	triplet
TBAB	tetra-*n*-butylammonium bromide
TBS	*tert*-butyldimethylsilyl
TCEP	tris(2-carboxyethyl)phosphine
TFA	trifluoroacetic acid
THF	tetrahydrofuran
TLC	thin layer chromatography
TSG	tumour suppressor gene
V70	2,2'-azobis(4-methoxy-2,4-dimethylvaleronitrile
VCD	vibrational circular dichroism

Chapter I

Introduction

Cancer and drug discovery

Cancer is a leading cause of death worldwide. In 2008, it accounted for 13% of all deaths (7.6 million) and this situation is expected to worsen with an estimated 13.1 million people dying from cancer in 2030.[1] Although lifestyle choices have now been shown to have some significant influence (*e.g.* tobacco or alcohol use, high body mass index), inherited genetic factors and increasing age are uncontrollable variables that also strongly contribute to the incidence of cancer. The complex nature of the disease makes it one that is difficult to detect, treat and understand. Nevertheless, research in oncology has become one of the major focus areas in both academia and pharmaceutical companies, strongly illustrated by the high number of cancer-related clinical trials (in 2009, *ca.* 16,000 which represent about 40% of all trials registered on ClinicalTrials.gov).[2]

Cancer can be defined as a class of diseases in which cells have acquired the ability to proliferate abnormally and to invade surrounding normal tissue.[3] In advanced stages of the disease, cancer cells can metastasize, whereby they migrate from their initial location and spread to another site to develop a secondary tumour. This abnormal proliferation state is the outcome of a complex multi-step process whereby the cell acquires a number of capabilities ("hallmarks of cancer" as described by Hanahan and Weinberg)[4,5] allowing it to evade control by the regulatory machinery. Acquisition of these capabilities is primarily the result of deregulation of the expression of a number of key genes, which can be separated in two main families: proto-oncogenes and tumour suppressor genes. Proto-oncogenes are genes that, when mutated or overexpressed, contribute to the transformation of a normal cell into a cancer cell, thus becoming oncogenes. They are commonly involved in the regulation of cell growth or proliferation and are usually implicated in signal transduction pathways; receiving and transmitting growth stimulatory signals (that originate from the extracellular environment). The activation of proto-oncogenes into oncogenes can be achieved *via* diverse pathways leading to overexpression or hyper-active proteins, such as chromosomal translocation, point mutation or gene amplification. Once activated, oncogenes become overly active growth-promoting

genes and will drive the cell to proliferate independently from extracellular signalling.[6] In contrast, tumour suppressor genes (TSG) counterbalance these growth-promoting genes and are involved in numerous cell processes for regulating cell growth and death. Their deactivation is believed to be as, if not more, important than the activation of oncogenes for cancer formation.[6] However, the two alleles of a TSG have to be inactivated for the malignant phenotype to be expressed, since it is generally recessive. This rare situation can be achieved in various ways such as mutation, loss of heterozygosity or epigenetic silencing, as will be described below. Once both copies of the TSG are inactivated, the cell no longer receives anti-proliferative signals, thus meeting another requirement for carcinogenesis.[6]

The first approach to cancer chemotherapeutic treatment was the development of cytotoxic agents. Amongst these are the nitrogen mustards, which originate from the discovery that mustard gas used during World Wars I and II exhibits an anti-carcinogenic effect.[7] A number of cytotoxic drugs have received FDA approval over the years and are still the first line treatment for cancer patients today. Such drugs include Taxol (an anti-microtubule agent which blocks cell division), Cis- and Carboplatin (DNA crosslinking agents) and the anthracyclines (DNA intercalating agents). While these drugs have had an important impact on patients' lives, they all induce very severe side effects due to their lack of specificity for cancer cells over healthy cells. Next generation chemotherapies promise to address this major issue. The increasing understanding of molecular biology and the discovery of cancer-causing genes (oncogenes and TSG mentioned above) led to the opportunity of specifically targeting cancerous cells. Targeted chemotherapy, which includes both small molecule inhibitors and monoclonal antibodies, aims to affect these deregulated pathways specifically and consequently hopes to have only minor impact on healthy cells. The FDA approval of the Bcr-Abl kinase inhibitor Gleevec (Imatinib) in 2001 for the treatment of multiple cancers, most notably chronic myeloid leukaemia (CML), confirmed the viability of this approach and translated into impressive survival rates for patients.[8]

While cancer cells exhibit numerous genetic alterations as mentioned above, it can appear counterintuitive that selective inhibition of a single (or a limited number of) protein(s) would lead to cancer cell death. The concept of "oncogene addiction" has been proposed to rationalise this phenomenon, with dependence of some cancer

cells on one (or a few) gene(s) for the maintenance of the malignant phenotype.[9] This hypothesis implies that the development of cancer is not the result of a simple addition of the individual effects of oncogene activation and tumour suppressor gene inactivation. It suggests that the proteins encoded by these cancer-related genes are implicated in different, more essential, roles in cancer cells than in normal cells. In other words, the regulatory machinery is severely altered in cancer cells and their survival can become dependent on distinctive pathways. Identification of this "Achilles heel" in a given cancer cell type is an essential step for selecting the appropriate drug target to pursue.[9,10]

Epigenetics and therapy

The genetic origin of cancer is now well established, whereby alterations of critical genes regulating cell growth, proliferation, or death lead to the malignant phenotype. However, intense research over the past two decades has shed light onto the crucial role of the transcriptional regulatory pathways in the cell, known as epigenetic processes in cancer. As implied by its name ("Epi" comes from the Greek word for "over"), epigenetic processes control the regulation of gene expression without involving genetic alteration of the DNA sequence. It is now clearly established that epigenetic modifications play a critical role at every stage of carcinogenesis through deregulation of the expression of critical cancer-related genes.[11,12]

Epigenetic modifications

Epigenetic processes are modifications that influence the structure of chromatin, which consists of DNA and histone proteins (Figure 1). Within the nucleus, *ca.* 147 base pairs of DNA are wrapped around an octamer of histone proteins: two dimers of histones H3 and H4, complemented by two dimers of histones H2A and H2B.[13] Together, this complex makes up the fundamental repeating unit of chromatin called a nucleosome.[14] DNA methylation and post-translational modifications (PTM) of histones are examples of epigenetic modifications and will be further described below (Figure 1).

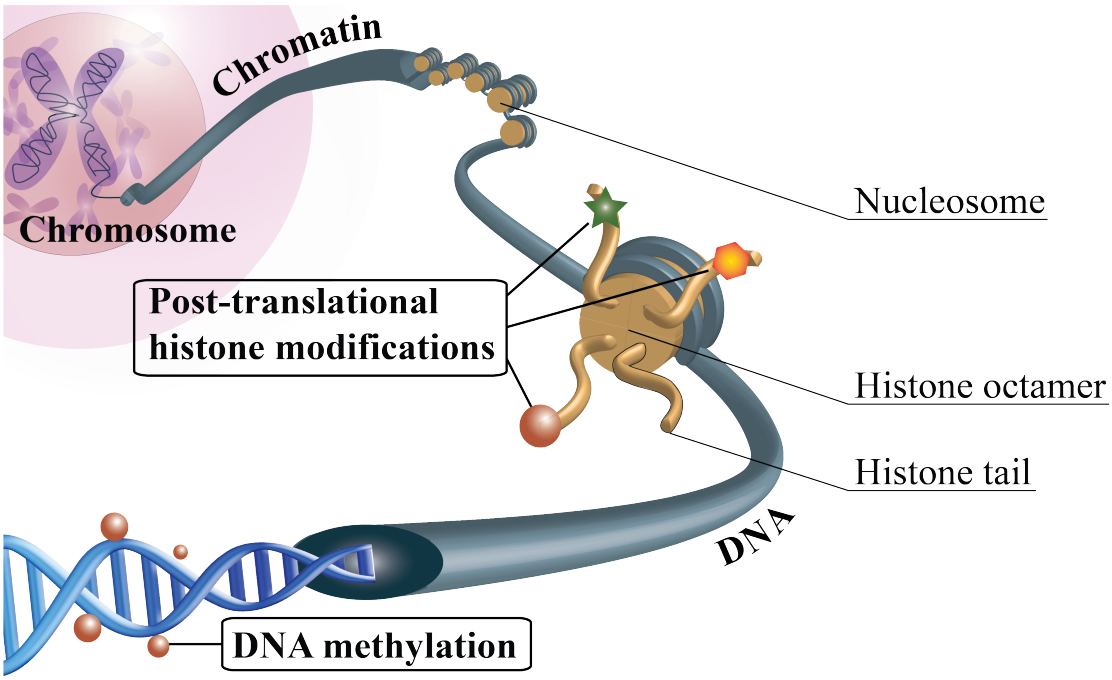
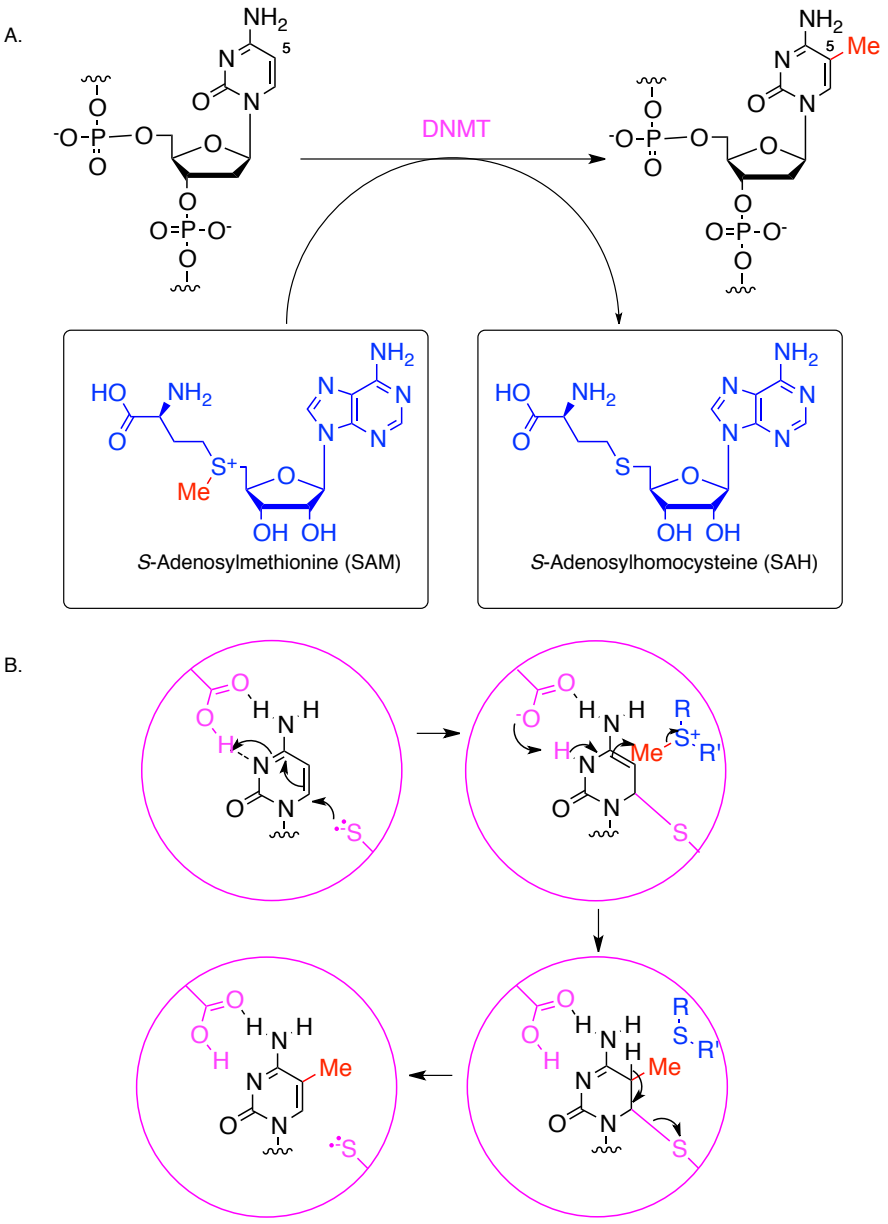


Figure 1 Chromatin and epigenetic modifications (Adapted from Cherblanc *et al.*[14])

DNA methylation

DNA methylation is carried out by DNA methyltransferases (DNMTs) which transfer a methyl group from the endogenous co-factor *S*-adenosylmethionine (SAM) to the *C*-5 position of specific cytosine nucleotide (Scheme 1).[15] DNA hyper-methylation at promoter proximal CpG-rich regions (Cytosine-phosphate-Guanidine, or CpG island) is generally associated with gene silencing. Two main mechanisms have been proposed to relate DNA methylation to transcriptional silencing.[16] Firstly, DNA methylation leads to a decrease in binding affinity between transcription factors and their target DNA sequence, which can be regarded as a direct mechanism for inhibition of transcription.[17] Secondly, methylated cytosines are recognition elements for methyl-CpG-binding proteins (MBPs), which are in turn able to recruit further transcriptional repressive complexes. [18,19]



Scheme 1 A. General scheme for DNA methylation of cytosine at *C*-5. B. Mechanism of the methylation reaction. The DNMT is depicted in pink, *S*-adenosylmethionine/*S*-adenosylhomocysteine cofactors in blue and the transferred methyl group in red. Adapted from Cherblanc *et al.*[20]

Histone modifications

Histone proteins not only act as a scaffold for DNA condensation but also as a dynamic structure that influences gene expression. In particular, the poorly structured *N*-terminal region of histone proteins (histone tail, Figure 1) is subject to a variety of post-translational modifications (PTMs) including acetylation, methylation, phosphorylation, ubiquitination, sumoylation and ADP ribosylation (see Figure 2 for selected examples).[13,21] Acetylation and methylation of histone residues are the most characterised modifications and will be described in further detail.

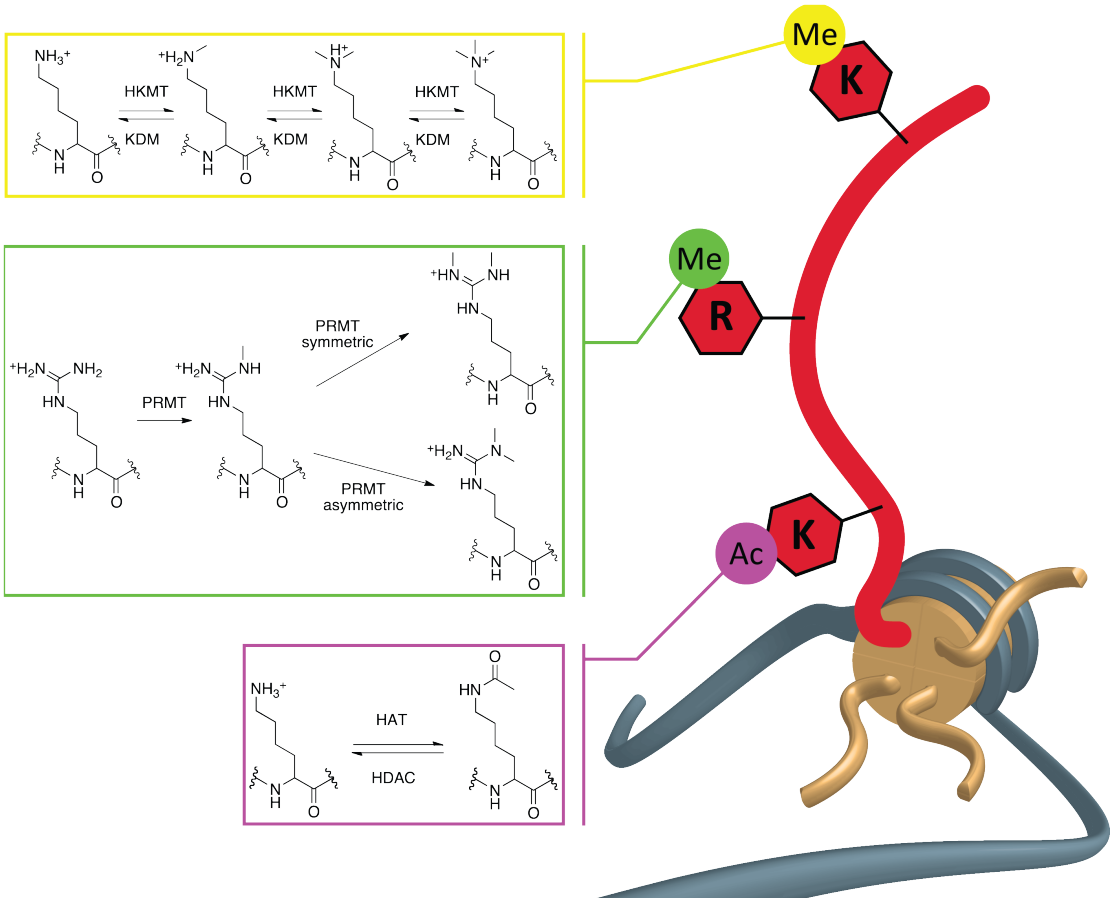


Figure 2 Selected examples of post-translational modifications of histones (Adapted from Cherblanc *et al.*[14,20])

Acetylation of specific lysine residues of histones is regulated by the counteracting action of histone acetyltransferases (HATs) and histone deacetylases (HDACs), which respectively add or remove an acetyl group to specific lysine residues (Figure 2). Similarly, methylation of specific lysine or arginine residues is mediated by histone lysine methyltransferases (HKMTs) and protein arginine methyltransferase (PRMTs) respectively (Figure 2). While lysine residues can be mono-, di- or trimethylated, arginine residues can be mono- or di-methylated, in a symmetrical or asymmetrical fashion. Demethylation of lysine residues is performed by lysine demethylases (KDMs). However, no direct arginine demethylase has yet been reported. Protein arginine deiminase (PADs) have been suggested to lead to demethylation of arginine *via* deimination of both arginine and monomethylarginine.[22]

Similarly to DNA methylation, these PTMs are key to epigenetic gene regulation by two main mechanisms. Firstly, the three-dimensional structure of chromatin can be controlled as a result of the presence of these epigenetic marks.[23]

Acetylation is thought to lead to a more "open" chromatin (*i.e.* accessible to the transcription machinery) *via* neutralisation of the positive charge on the lysine residue, which is suggested to lead to a lower electrostatic affinity of the histone tail for the negatively charged DNA backbone.[23] Secondly, these PTMs constitute docking sites for regulatory effector proteins. Each PTM is specifically recognised by a "reader" of histone modifications,[24] which includes bromodomain-containing proteins (bind to acetylated lysine residues), and chromo-, PHD-, or tudor domain-containing proteins (specific for methylation marks). Transcription is consequently repressed or activated by the recruitment of further transcriptional regulatory protein complexes and further chromatin modifications.[25]

Importantly, the downstream effects of a PTM on a histone tail will depend on the site of modification (position of the lysine or arginine residue) but also the nature and multiplicity of the mark (*e.g.* ac, me, me2, me3).[13] Most commonly, transcriptionally active chromatin presents elevated levels of acetylation in combination with specific tri-methylation marks at lysine 4 on histone 3 (H3K4), H3K36, and H3K79. Conversely, repressive chromatin is associated with reduced levels of acetylation together with elevated levels of tri-methylation at H3K9, H3K27 and H4K20.[13,26]

Further complexity in the deciphering function of single chromatin marks and their impact on gene expression stems from the crosstalk between the different modifications, often referred to as the "histone code" hypothesis.[27] Active gene transcription or silencing is a result of the combination of epigenetic marks, including histone PTMs and DNA methylation. Furthermore, the enzymes mediating and reading these marks are also interrelated (*e.g.* have the ability to recruit each other or are part of the same complexes) which results in an intricate network currently not fully understood.[25]

Epigenetic therapy for cancer

Deregulation of epigenetic marks have unsurprisingly been observed in a variety of cancer types.[11] Epigenetic aberrations in cancer can stem from a variety of alterations to the enzymes mediating the installation or removal of these marks, such as mutations, chromosomal translocations or overexpression.[28-33] The ensuing abnormal epigenetic modifications result in the aberrant expression of oncogenes or

conversely the anomalous silencing of TSGs.[34] For example, numerous TSGs have been reported to be silenced by DNA hyper-methylation in a variety of cancer types.[35] This includes genes involved in DNA repair, signal transduction, cell cycle regulation, apoptosis or angiogenesis.[33,36,37] Similarly, mutations of MLL1 (a HKMT specific for the methylation of H3K4, an activating mark) are suggested to be responsible for various types of leukaemia,[38] while overexpression of EZH2, the main enzyme responsible for the silencing H3K27me3 mark, has been reported in breast and prostate cancers.[39]

An important feature of epigenetic modifications, in contrast with genetic mutations, is their reversibility. Pharmacological disruption of the enzymes mediating these epigenetic marks is expected to lead to reversal of the aberrant epigenetic landscape of a cancer cell, resulting in a more normal state. The removal of anomalous silencing marks is for example accepted to lead to the re-expression of aberrantly silenced pathways and induce beneficial downstream effects such as triggering apoptosis of the cancer cells. Similarly to the oncogene addiction concept outlined above, it is possible that cancer cells become addicted to their aberrant epigenetic landscape and that its disruption may lead to cancer cell death selectively over healthy cells.[40] It is also envisaged that the combination of several epigenetic modifiers may be required to significantly restore the epigenetic state of a cancer cell to that of a normal cell. For example the synergistic effect of DNMT and HDAC inhibitors has already been reported.[41-43] Furthermore, the resulting re-expression of a large set of genes is also expected to re-sensitise cancer cells to conventional therapies (chemo or radiotherapy).[44]

Cognisant of the implication of epigenetic modifications in cancer and their potential reversibility, the enzymes mediating the installation or removal of these marks have become attractive targets for the development of new cancer treatment.[40]

Successful epigenetic therapies

The initial validation of the concept of epigenetic therapy came with the development of DNMT inhibitors,[45] which culminated in the FDA approval of two drugs for the treatment of myelodysplastic syndromes (MDS, pre-leukemic disease): Azacitidine and Decitabine (Figure 3A).[33] These two drugs, originally developed as classical cytostatic agents,[46] are azanucleosides which act as prodrugs: upon phosphorylation in the cellular environment and partial deoxygenation in the case of

Azacitidine (a riboside), the unnatural nucleotides become incorporated into DNA to subsequently trap DNMT *via* covalent linkage. They subsequently lead to DNA hypomethylation following cell division (Figure 3B).[47]

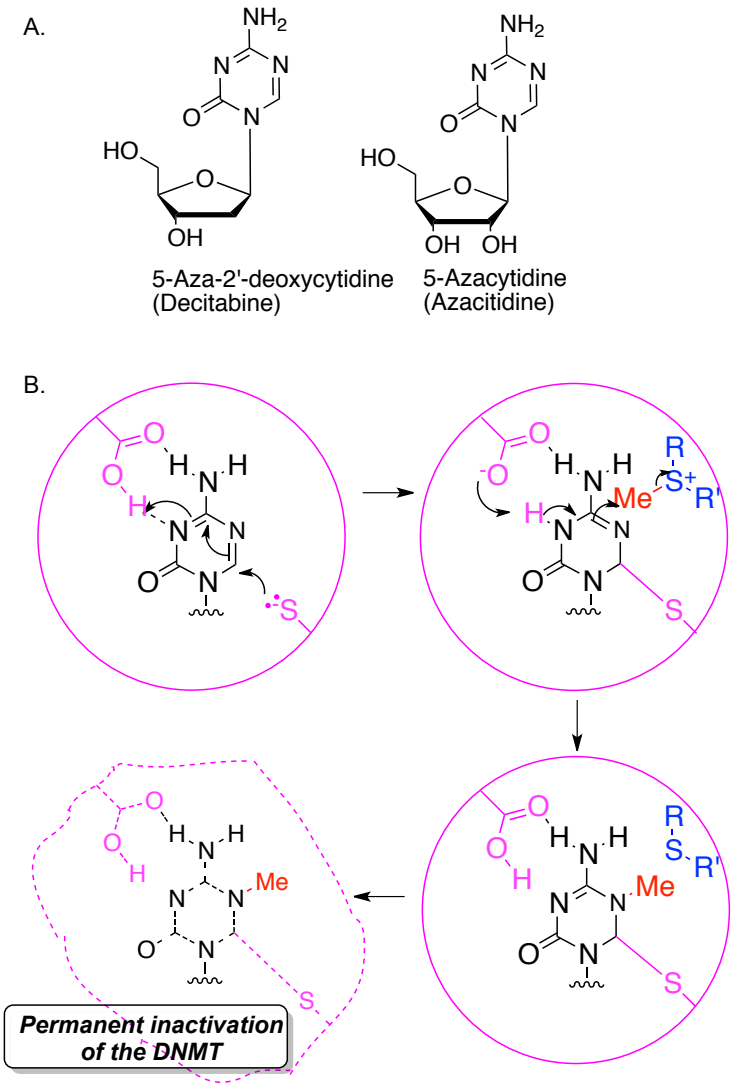


Figure 3 A. Structure of the two FDA-approved DNMT inhibitors. B. Mechanism of action of azanucleoside DNMT inhibitors. The DNMT is depicted in pink, *S*-adenosylmethionine/*S*-adenosylhomocysteine cofactors in blue and the transferred methyl group in red. For clarity only the cytosine part of the inhibitor is depicted in black. Adapted from Cherblanc *et al.*[20]

Further success in epigenetic therapeutic discovery was the approval of two HDAC inhibitors as a treatment for cutaneous T-cell lymphoma (CTCL).[48-51] Interestingly, both compounds, Vorinostat (SAHA) and Romidepsin (FK228) (Figure 4) were first identified as anti-tumour compounds reported to induce differentiation; their molecular targets, HDACs, were identified much later. A number of novel, isoform specific, HDAC inhibitors are currently in clinical trials.[52-54]

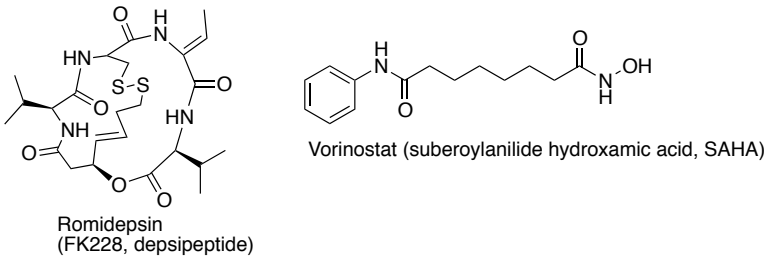


Figure 4 Structure of the FDA approved HDAC inhibitors

Given the success in developing inhibitors of DNMTs and HDACs, it is anticipated that enzymes regulating histone methylation would also constitute a promising class of drug targets. While the therapeutic potential of arginine methytransferases[22] and lysine demethylases[38,55] has been recognised, we will concentrate here on HKMTs, the most extensively characterised class and the main focus of this project.[56]

HKMTs as an emerging target class in cancer

An important number of studies in the past 10–15 years have gathered significant evidence on the pathogenic role of several HKMTs in a variety of cancer types. Table 1 presents selected examples of HKMTs and their alteration in cancer.

Enzymes involved in installing silencing (EZH2, G9a) or activating (DOT1L, MLL1) marks have both been linked to cancer pathogenesis. Furthermore, overexpression of several HKMTs, including EZH2, has been suggested to be a marker of the aggressiveness of the disease and poor prognosis.[57-59] A significant amount of biological data points to EZH2 as the most therapeutically relevant HKMT enzyme.[39] Chapter IV will describe our efforts at developing novel EZH2 inhibitors.

Table 1 Association of HKMTs with particular tumour types

Enzyme	Modification	Cancer type	Alteration
EZH2	H3K27me3	Bladder carcinoma	Overexpression
		Breast cancer	Overexpression
		Colorectal cancer	Overexpression
		Gastric cancer	Overexpression
		Prostate cancer	Overexpression
		Lymphomas	Mutation at Tyr641, Ala677 or Ala687
MLL1	H3K4	Leukaemia	Chromosomal translocation, amplification
G9a	H3K9me1/2	Hepatocellular carcinomas	Overexpression
		Gastric cancer	Up-regulation
SUV39H1	H3K9me2/3	Colon cancer	Overexpression
NSD1	H3K36me2/3	Neuroblastoma and gliomas	Silencing by promoter hyper-methylation
		Acute myeloid leukaemia	Chromosomal translocation
DOT1L	H3K79	Leukaemia	Chromosomal translocation

Adapted from Cherblanc *et al.,* [20] data from [59-61]

Structure of HKMTs

The HKMT family is composed of 52 closely related enzymes.[61] With the exception of DOT1L, the catalytic site of HKMTs is contained within the so-called SET domain (Suvar3-9, Enhancer of Zeste, Trithorax).[62,63] The active site of HKMTs is characterised by the presence of three interconnected domains: the cofactor SAM binds in a pocket on one side of the protein, the H3 substrate binds to a cleft on the opposite side and a narrow hydrophobic channel, where the side chain of the lysine to be methylated is positioned, connects the two binding sites (Figure 5).[64]

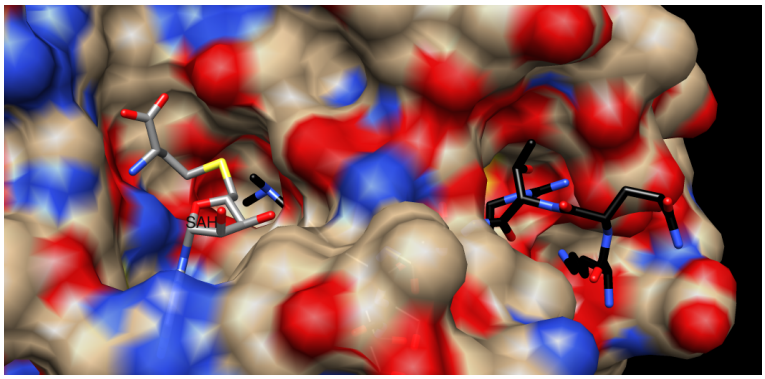


Figure 5 SAH (grey carbons) and H3 peptide (black carbon) with methylated lysine 9 co-crystallised with GLP. Generated from pdb file 2RFI, the surface represents the electrostatic potential.[65]

HKMTs have been reported to be specific for the degree of methylation they perform and the residue they modify (*e.g.* EZH2 only trimethylates H3K27, SET7/9 only monomethylates H3K4), although the latter is becoming a challenged view as non-histone targets have been identified for a number of HKMTs.[23,66] Depending on the enzyme involved, multiple methylation events have been shown to be performed in a processive (*i.e.* without release of the substrate, *e.g.* G9a)[67] or non-processive (*e.g.* SUV39H1)[68] manner. Crystallographic and mutagenesis studies of different HKMTs (DIM-5, G9a and SET7/9) have shed light on the mechanism controlling the specificity in the degree of methylation. The presence of either a phenylalanine or a tyrosine residue at a conserved position in the active site was shown to act as a switch between single methylation or multiple methylation capabilities.[69] The hydroxyl moiety of the tyrosine residue was suggested to prevent the necessary rotation of the monomethylated lysine (*via* hydrogen bonding) thus impeding the next methylation step.

Reported HKMT inhibitors

Based on the catalytic mechanism and the overall structure of HKMTs, it has been suggested than this enzyme class would be amenable to small molecule inhibition.[70] In particular, the presence of a small, well-defined, cofactor binding pocket has given confidence in the possibility of designing such inhibitors.[62] An increasing number of HKMTs have now been crystallised and this has opened the way for rational design.[71]

Rational design of cofactor analogues

The first HKMT inhibitors were designed as cofactor (SAM) analogues, including *S*-adenosylhomocysteine (SAH, the product of the methylation reaction), methylthioadenosine and the natural product sinefungin (SAM analogue found in bacteria) (Figure 6). Unsurprisingly, although the compounds exhibited inhibitory activity, they were not selective for a particular HKMT.[62,72]

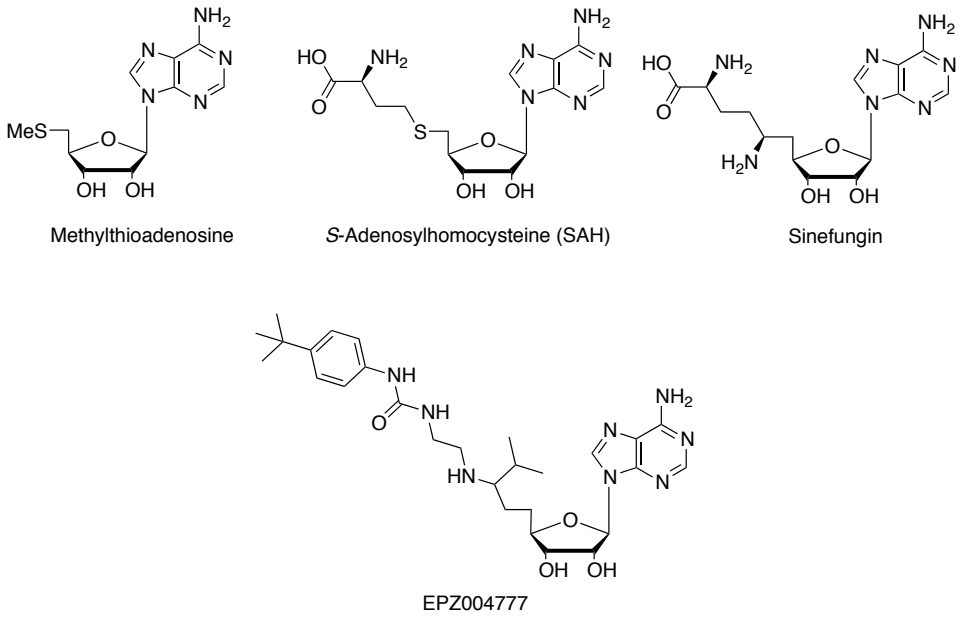


Figure 6 Structure of SAM analogues HKMT inhibitors

More recently, based on intensive crystallographic studies and iterative medicinal chemistry, Daigle *et al*. reported a highly potent (IC_{50} = 0.4 nM) and selective (>1000-fold) inhibitor of DOT1L.[73] DOT1L has been suggested as a potential therapeutic target in mixed lineage leukaemia (MLL) as a result of a chromosomal translocation of the MLL gene. This translocation leads to the recruitment of DOT1L to aberrant gene locations and subsequently to abnormal methylation patterns. As depicted in Figure 6, EPZ00477 retained the nucleoside core of SAM and was confirmed by kinetic analysis to bind in the cofactor binding pocket. Treatment of MLL cells with the compound resulted in the reduction of H3K79 methylation together with the inhibition of expression of key MLL fusion target genes. The compound induced selective killing of cells bearing the MLL fusion and this further translated into anti-tumour activity in a mouse MLL xenograft model.[73]

HKMT inhibitors identified by HTS

The current prevailing approach to identify hit compounds for a drug discovery project is through the use of a high throughput screen (HTS). Epigenetic drug discovery is no exception; a number of HKMT inhibitors were discovered by screening large libraries of compounds.

Chaetocin (**1**, Figure 7), a fungal metabolite, has been identified to inhibit SU(VAR)3-9, the drosophila orthologue of SUV39H1, in the low micromolar range (IC_{50} = 0.6 μM).[74] The biological activity of chaetocin and its synthetic analogues will be the focus of Chapter II.

Upon screening a preselected fraction of the Boehringer Ingelheim library, Jenuwein and co-workers identified two compounds that inhibited the H3K9 dimethylase G9a (Figure 7).[75] While BIX-01294 (**2**) and BIX-01338 both inhibited G9a in the low micromolar range (IC_{50} of 2.7 and 4.7 μM respectively), BIX-01338 was found to be competitive with the cofactor SAM whereas BIX-01294 was reported to be uncompetitive with SAM, suggesting that it binds to the enzyme-cofactor complex but not to the free enzyme. Furthermore, most likely in line with their modes of action, BIX-10338 inhibited all HKMTs tested in an unspecific manner while BIX-01294 inhibited G9a and only the closely related GLP (but with a lower potency of 38 μM). SAR studies have been reported around the BIX-01294 scaffold and will be described in more detail in Chapter IV. With the aim of improving BIX-01338 selectivity, Schreiber and co-workers synthesised a library of analogues based on its benzimidazole scaffold and discovered BRD9539 (Figure 7) which displayed moderate *in vitro* potency against G9a as well as EZH2 ($IC_{50} \approx 6$ μM for both enzymes) but was selective over SUV39H1, NSD2, DNMT1, 16 other chromatin-modifying enzymes, and 100 kinases.[76] BRD4770 (Figure 7), the methyl ester of BRD9539, was used as a prodrug for cell-based studies and was found, using mass spectrometry, to reduce cellular levels of H3K9 di- and trimethylation, while not affecting the methylation level of H3K27.[76] This suggested that EZH2 activity, which is responsible for the methylation of H3K27, was not inhibited in cells, in contrast with the data from the biochemical assay. Nevertheless, BRD4770 induced senescence in pancreatic cancer cells and represents a useful tool to study the role of G9a in these aggressive cancer types.

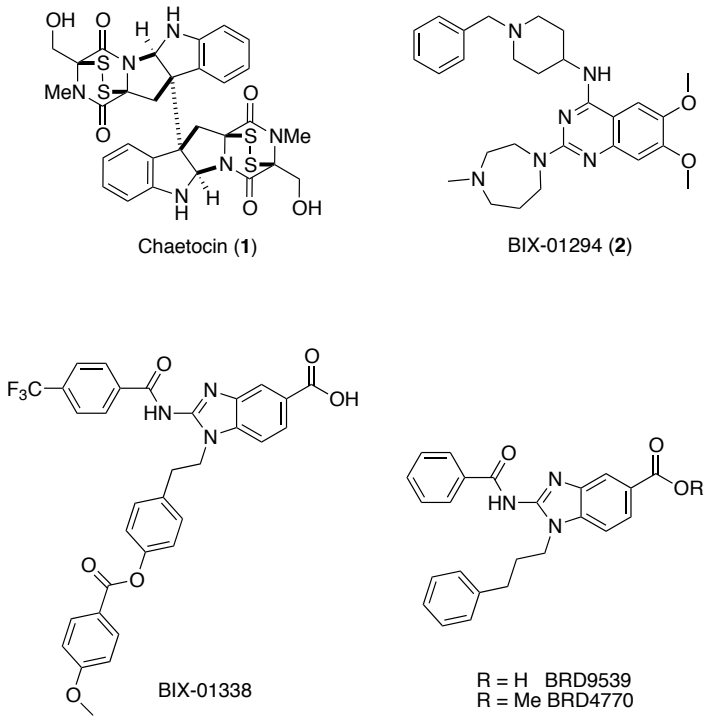


Figure 7 Structure of the HKMT inhibitors identified by HTS and derivatives (BRD compounds)

Recently, screening efforts also delivered hit compounds that have been developed into potent and selective EZH2 inhibitors by a number of academic groups and pharmaceutical companies.[77-81] These studies will be described with greater detail in Chapter IV.

Global objectives of the project

The initial purpose of the work presented in this thesis was to contribute to epigenetic cancer drug discovery efforts by developing and studying inhibitors of HKMTs.

We first endeavoured to further characterise the inhibitory activity of the natural product chaetocin (**1**, Figure 7), which has been described as a specific inhibitor of SU(VAR)3-9.[74] The preparation of analogues, and their biological activity will be described in Chapter II. Since the stereochemistry of one of our chaetocin analogues required further characterisation, our efforts to address this problematic assignment together with that of molecules of similar structure, will be presented in Chapter III.

Finally, we attempted to improve EZH2 inhibitors, based on the quinazoline scaffold of BIX-01294 (**2**, Figure 7),[75] recently discovered in the group. The design, synthesis and biological assessment will be described in Chapter IV.

Chapter II

Design, synthesis and biological evaluation of chaetocin analogues

Introduction

In 2005, Imhof and co-workers performed a high throughput screen (HTS) with the aim to discover new small molecule inhibitors against *Drosophila melanogaster* SU(VAR)3-9 (*d*SU(VAR)3-9), a histone lysine methyltransferase reported to specifically methylate lysine 9 of histone H3 (H3K9).[74,82] Of the *ca*. 3,000 compounds screened, the natural product chaetocin (**1**, Figure 7) was identified as the most potent inhibitor (IC_{50} of 0.6 μM against SU(VAR)3-9).[74] Moreover, it was shown that chaetocin could inhibit the human orthologue of *d*SU(VAR)3-9 (SUV39H1) with a similar IC_{50} value (0.8 μM), and other members of the SU(VAR)3-9 class of HKMTs including mouse G9a (IC_{50} = 2.5 μM) or *Neurospora crassa* DIM5 (IC_{50} = 3 μM). Chaetocin was shown however to be inactive against HKMTs not belonging to the SU(VAR)3-9 class, such as *d*E(z)-complex (the orthologue of human EZH2/PRC2 complex), PRSET7 or SET7/9. Chaetocin was therefore reported to be specific for the SU(VAR)3-9 class of enzymes.

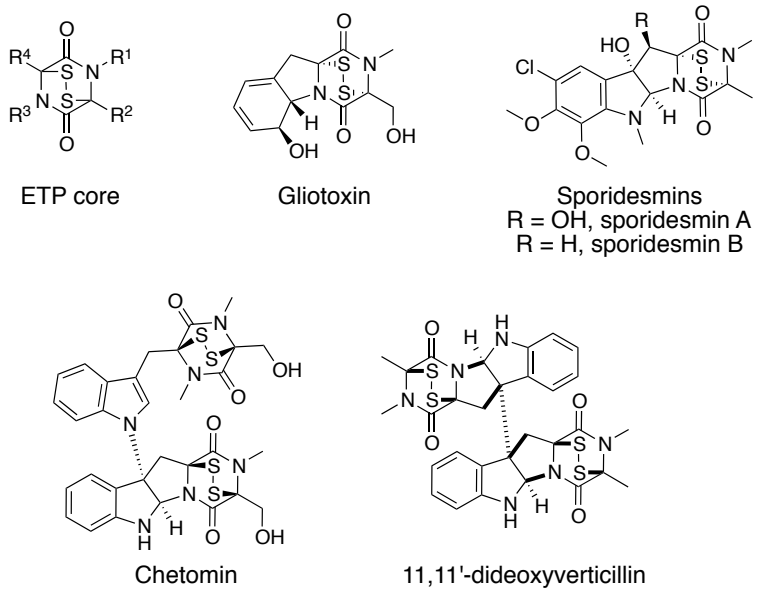


Figure 8 Examples of ETP containing natural products

Chaetocin is a fungal metabolite first isolated from the fermentation of *Chaetomium minutum*.[83] It belongs to the 3,6-epidithio-diketopiperazine (ETPs) class

of fungal toxins, which contain a characteristic diketopiperazine (DKP) scaffold bridged by two sulfur atoms (ETP core, Figure 8).[84,85] ETP compounds are believed to be primarily involved in the chemical defence and virulence mechanisms of fungi.[85] Their toxicity extends outside the fungi class and they have been reported to exhibit a variety of biological activities in higher organisms, including mammals, both beneficial and detrimental.[85] For example, sporidesmins (Figure 8), produced by *Pithomyces chartarum* on infected grasses, has been implicated in the development of facial eczema and liver disease in sheep.[86] Conversely, the toxicity of ETPs has also made them attractive for potential therapeutic uses: gliotoxin (Figure 8), the best-characterised ETP, has been shown to have antibacterial, antifungal, immunosuppressive, antiviral and anti-tumour activities.[87] A large number of structurally, functionally, and evolutionary unrelated enzymes have been shown to be inhibited by ETP compounds. Gliotoxin has, for example, been reported to target reverse transcriptase,[88] the transcription factor NF-κB,[89] poliovirus polymerase 3Dpol,[90] farnesyltransferase and geranylgeranyltransferase[91] amongst many others.[92] Several mechanisms have been proposed to account for the broad toxicity and various inhibitory activities of ETP compounds. These include protein cross-linking through reaction of the disulfide functionality with cysteine residues (Figure 9A), the generation of reactive oxygen species *via* redox cycling (Figure 9B), and the ejection of structurally important zinc ions from the protein (Figure 9C).[84,85,93,94] All these mechanisms rely on the chemical reactivity of the ETP disulfide bridge.

Similarly, chaetocin has also been shown to target several enzymes other than HKMTs, including thioredoxin reductase,[95] and to disrupt the protein-protein interaction between p300:HIF-1α (hypoxia-inducible factor).[93] Interestingly, both targets are involved in the sensing and scavenging of harmful reactive oxygen species (ROS) in cells and in each study, the disulfide bond has been suggested to be central to chaetocin's activity.

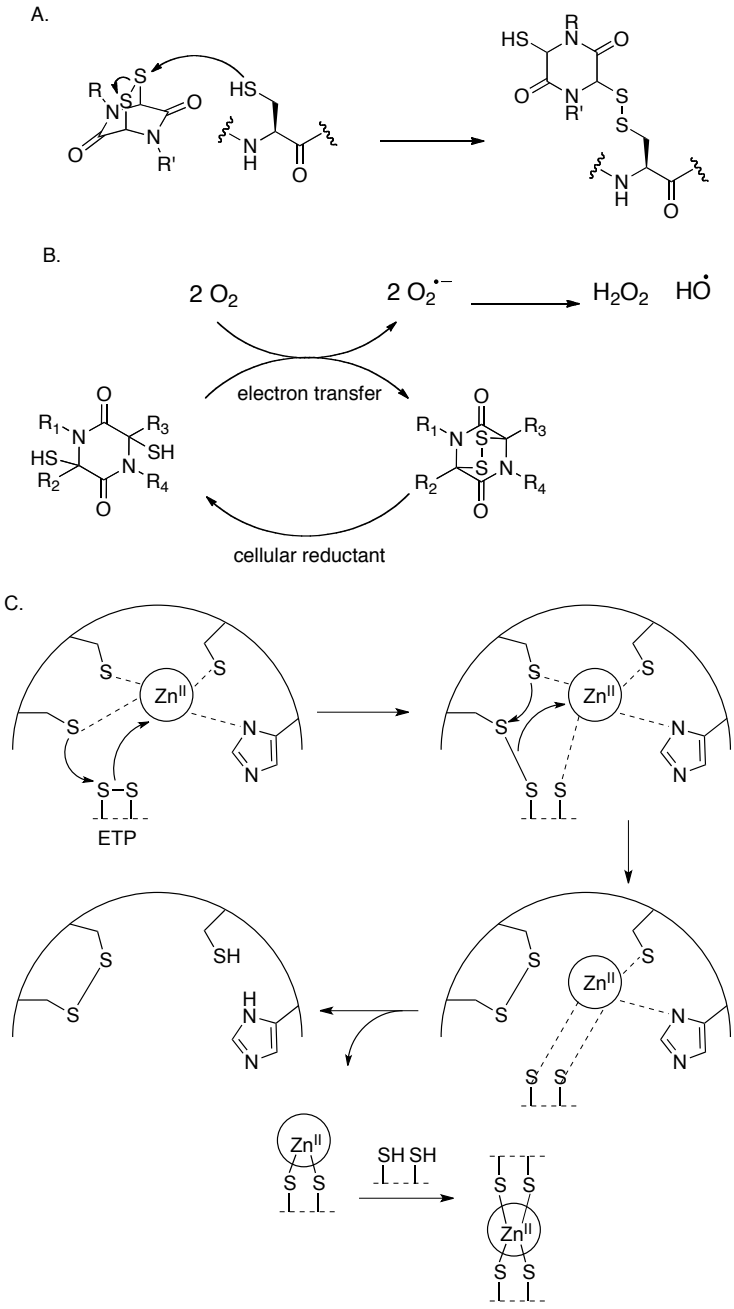


Figure 9 Suggested toxicity pathways for ETP compounds, adapted from [84,93]

Nevertheless, one important point in the initial report by Imhof and co-workers warranted further investigation on the HKMT inhibitory activity of chaetocin: Based on the fact that the reducing agent dithiothreitol (DTT) had no noticeable effect on the inhibitory activity of chaetocin against the biochemical activity of SU(VAR)3-9,[74] the authors have suggested that the disulfide bridge of this ETP molecule is not required for its HKMT inhibitory activity. Should this hypothesis prove correct, removal of the reactive disulfide functionality from chaetocin, while retaining the skeletal framework should result in the discovery of a novel HKMT

inhibitor with increased specificity and reduced off-target toxicity. However, most ETP analogues where the disulfide moiety has been removed or converted to an unreactive functionality have been found to be inactive compared to their parental ETP.[85,96]

Aims of the project and strategy of the inhibitor design

The initial aim of this project was to assess the requirement of the disulfide bridge of chaetocin for its HKMT inhibitory activity and to attempt to expand its SAR. A number of analogues were designed without the reactive disulfide bond while retaining the structure of chaetocin (**1**, Figure 7). The monosulfide analogue **3** (Figure 10) was particularly interesting as it differs from the parent molecule by only one sulfur atom and most importantly retains the bicyclic structure of the ETP core. This analogue was indeed crucial to probe whether the three-dimensional structure of chaetocin influenced its inhibitory activity. The methylthioether derivative **4** (Figure 10) was also designed to convert the reactive disulfide bond into an unreactive functionality while conserving all sulfur atoms. A fully desulfurised compound **5** (Figure 10) was also interesting to further assess the dependence of chaetocin's inhibitory activity on the presence of the sulfur atoms.

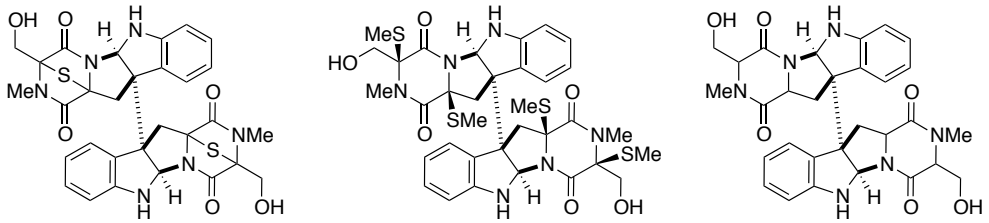
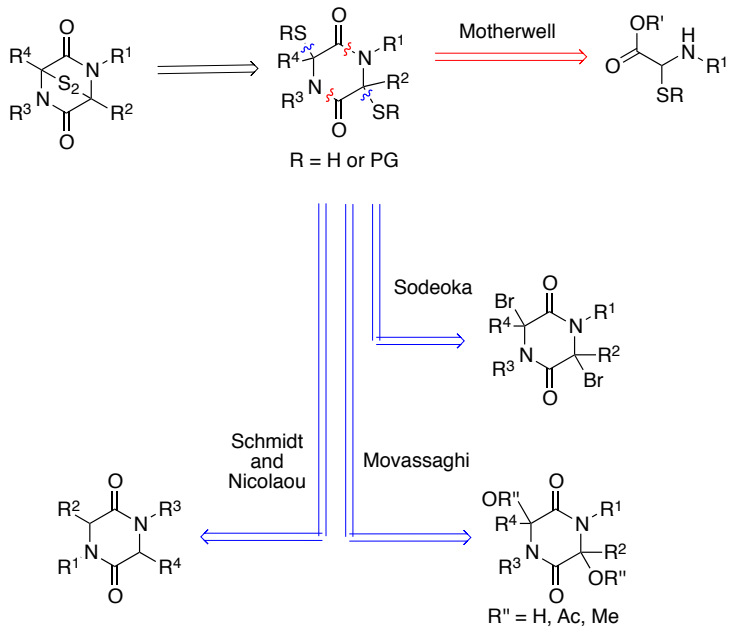


Figure 10 Selected analogues of chaetocin

While the total synthesis of chaetocin and its analogues would have been an attractive synthetic challenge, it was considered beyond the remit of this project. A semi-synthetic approach was preferred, whereby the natural product could be isolated from its natural source and then derivatised selectively around its disulfide functionality.[96] Isolation of chaetocin from one of the producing fungi, as previously reported,[97-99] was envisaged to allow access to significant amounts of the natural product in a more cost-effective manner than obtaining it from commercial sources.

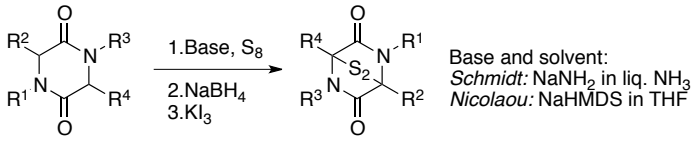
Background on ETP synthesis in the literature

In order to develop a reliable methodology for derivatisation at the disulfide bridge of chaetocin, a simplified model ETP compound was studied in order to probe the reactivity of the ETP moiety prior to studying the precious natural product. Over the years, the synthesis of ETP-containing natural or synthetic molecules has generated considerable interest due to their high structural complexity.[100] Here, we will give an overview of selected examples focused on the various methods developed to build the ETP core. Scheme 2 shows the most common retrosynthetic analyses.



Scheme 2 Retrosynthetic analyses to the ETP core[101-105]

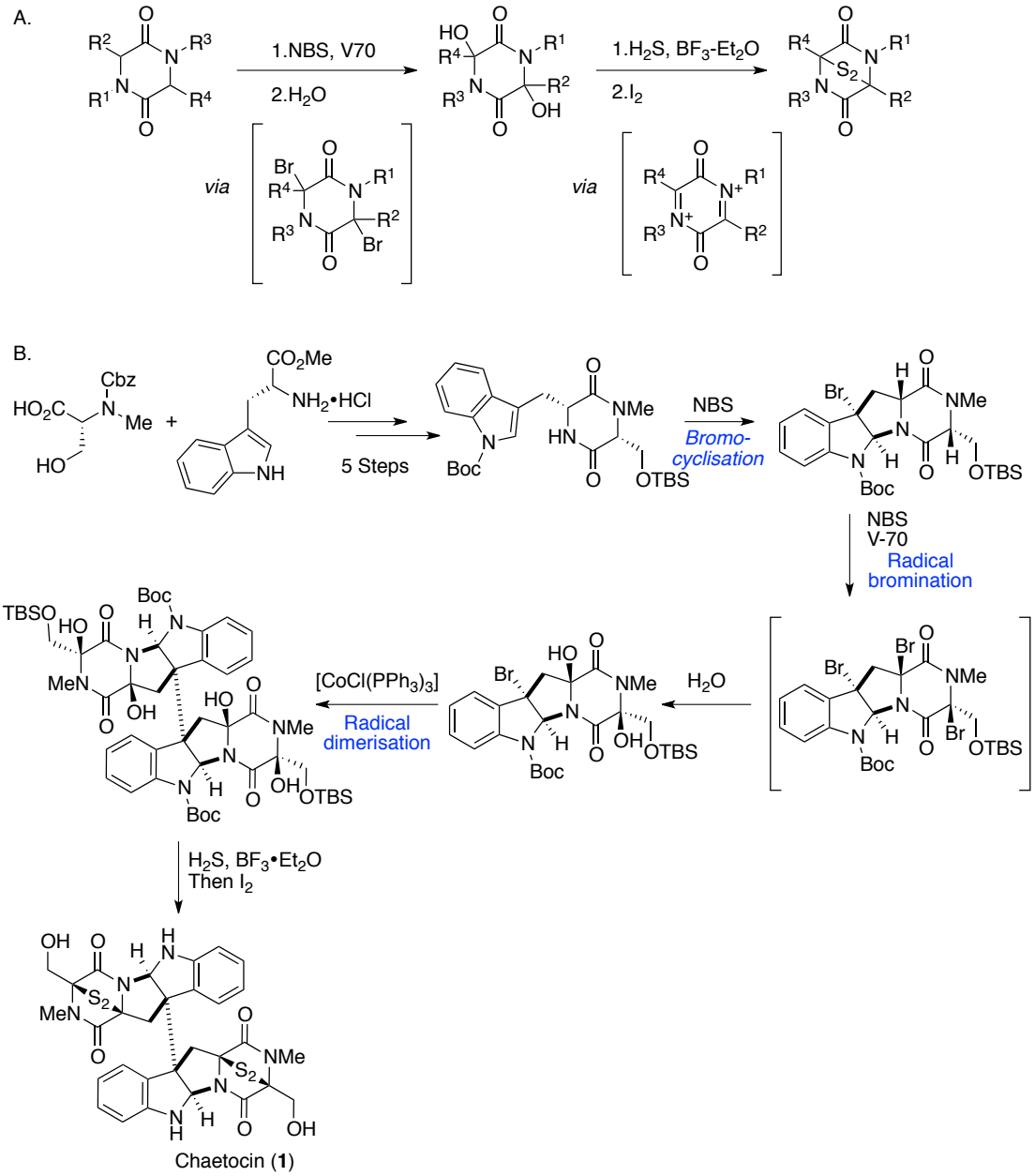
Schmidt and co-workers reported an efficient sulfenylation method of the DKP core by treatment with elemental sulfur and sodium amide in liquid ammonia (Scheme 3).[101] The methodology was further improved by Nicolaou and co-workers, who devised a more convenient procedure whereby NaHMDS was used as the base, and THF as the solvent.[100,102] Nicolaou and co-workers reported the synthesis of a large number of natural products using this methodology with variable yields (20 to 70%).[100] This approach was applied to the synthesis of a gliotoxin analogue and will be further described in Chapter III.



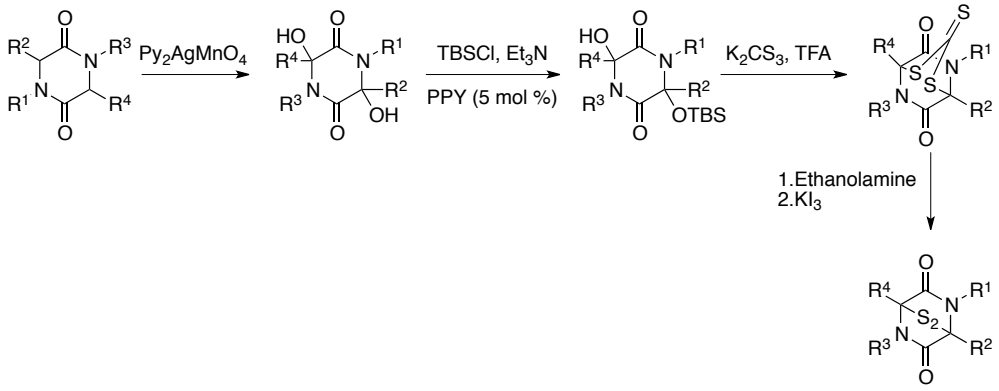
Scheme 3 Schmidt and Nicolaou's approach to the ETP core[101,102]

Sodeoka and co-workers accomplished the first total synthesis of chaetocin using also a late stage sulfenylation of the DKP ring, the general approach of which is depicted in Scheme 4A.[104] The DKP ring was subjected to radical bromination using *N*-bromosuccinimide (NBS) and V70 (a diazo-containing radical initiator); subsequent treatment with water led to the formation of an hemiaminal. The dithiol was then obtained by substitution reaction of the tetraol with H_2S in presence of BF_3•OEt_2 as Lewis acid, *via* a putative iminium intermediate. Final oxidation of the dithiol with iodine afforded the ETP core (44% yield from the hemiaminal in the case of chaetocin's scaffold).[104] The synthesis of chaetocin is depicted Scheme 4B.

A number of additional sulfenylating agents have been used for the formation of the ETP core, including potassium thioacetate (KSAc)[106], sodium tetrasulfide,[107] and more recently Movassaghi and co-workers pioneered the use of potassium trithiocarbonate (K_2CS_3) to generate the ETP moiety from monosilylated 3,6-dihydroxydiketopiperazine intermediate in the total synthesis of (+)-11,11'-dideoxyverticillin A (general approach in Scheme 5).[105]

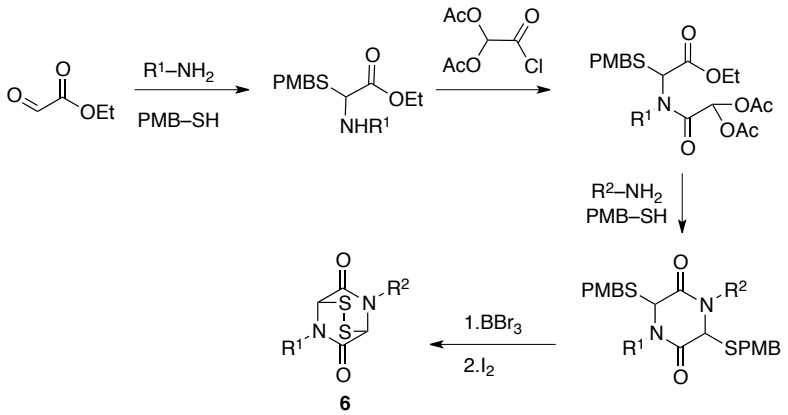


Scheme 4 Sodeoka's approach to the ETP core (A) and synthesis of chaetocin (**1**)[104]



Scheme 5 Movassaghi's approach to the ETP core[105]

While most efforts have been focused on the late stage sulfenylation of the preformed diketopiperazine core, Motherwell and co-workers developed a mild and convenient method to rapidly access the ETP core without involving the formation of the DKP, which are known to endow the molecule with solubility issues.[103] As depicted in Scheme 6, the route incorporates two three-component reactions in a four-step synthetic sequence. The alkyl group on the nitrogen atoms can be conveniently modified and this method is applicable to the synthesis of compound libraries.[93] This method has been employed for the synthesis of a model ETP compound (**6**, $R^1 = R^2 = Bn$, Scheme 6) and the synthesis will be described in greater details in the next section.



6

Scheme 6 Motherwell's approach to the ETP core[103]

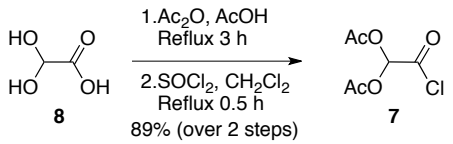
Results and Discussion

Synthesis and derivatisation of a model ETP compound

Synthesis of a model ETP compound

A simple benzyl-protected ETP (**6**, $R^1 = R^2 = Bn$, Scheme 6) was devised as a model compound to probe the reactivity of the disulfide bridge. The route developed by Motherwell and co-workers was considered the most convenient to access substantial amount of material for derivatisation studies and was followed to generate the ETP model compound **6** (Scheme 6) on a gram-scale.[93,103]

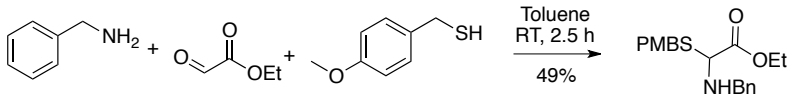
Acetyl chloride **7** was first synthesised following the literature procedure:[108] acetylation of glyoxylic acid monohydrate (**8**) with acetic anhydride followed by chlorination with thionyl chloride afforded the acetyl chloride 7 bearing the protected aldehyde in excellent yield (89%) (Scheme 7).



8 7

Scheme 7

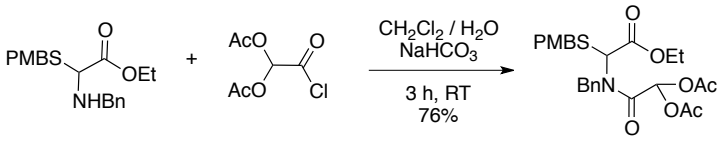
The second coupling partner ester **9** was obtained *via* a three component reaction optimised by Hilton *et al.* (Scheme 8).[109] Reaction of benzylamine (**10**), ethyl glyoxalate (**11**) and *para*-methoxybenzylmercaptan (**12**) in toluene afforded ester **9** in moderate yield (49%).



10 11 12 9

Scheme 8

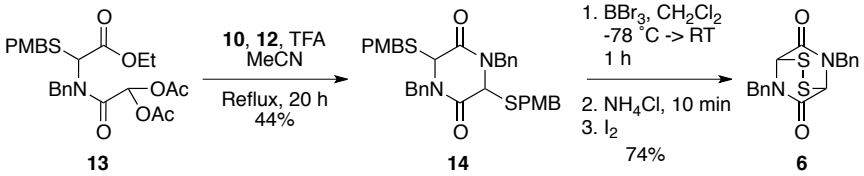
Subsequently, acetyl chloride **7** was coupled with amine **9** to afford amide **13** in good yield (76%) under Schotten – Baumann biphasic conditions (Scheme 9).[103]



9 **7** **13**

Scheme 9

The cyclisation to form diketopiperazine **14** proceeded in a moderate yield *via* a second three-component reaction in the presence of a catalytic amount of TFA to perform the *in situ* deprotection of the acetal. This step allowed the introduction of the second amide bond and installed the second PMB protected sulfur moiety. Importantly the desired *cis* isomer was obtained exclusively (Scheme 10), as previously described. Indeed, Aliev *et al.* reported that upon heating the reaction at reflux for over 16 h, the *cis* isomer was obtained solely.[103] Intrigued by this preference, they performed gas phase DFT calculations which supported this thermodynamically favoured configuration by revealing that the *cis* arrangement is ~9 kJ.mol^{-1} more stable than the *trans* configuration.[103] In the last step, the PMB protecting groups were cleaved with boron tribromide, followed by *in situ* oxidation of the di-thiol with iodine to afford ETP **6** in good yield (Scheme 10).

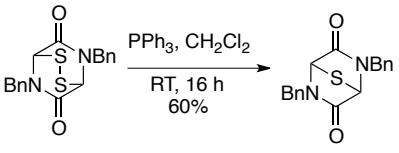


13 **14** **6**

Scheme 10

Derivatisation studies on the model ETP compound

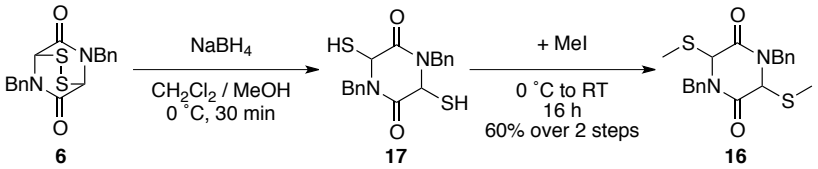
The preparation of monosulfide-bridged diketopiperazines by treatment of ETP compounds with triphenylphosphine is very well precedented.[110] Monosulfide **15** was prepared following recent literature procedure and was obtained in a moderate yield from ETP **6** (Scheme 11).[111]



6 **15**

Scheme 11

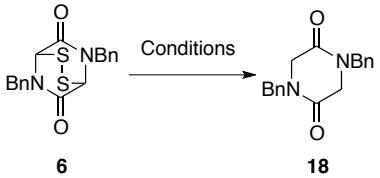
The methylthioether derivative **16** was prepared in a two-step one-pot reaction.[111] The disulfide bridge of ETP **6** was first reduced to the dithiol **17** with sodium borohydride and this was then directly alkylated with methyl iodide to afford the desired dithioether **16** in good yield (Scheme 12).



Scheme 12

Next, we assessed the feasibility of accessing a fully desulfurised compound. Raney nickel is generally the reagent of choice for the reduction of carbon–sulfur bonds and it has some precedent in the desulfurisation of ETP compounds,[112] however, various attempts at using Raney nickel in refluxing ethanol only furnished the expected desulfurised product **18** in very low yield (10%) or led to degradation when larger amount of reagent was used (Table 2). It was decided to explore other reagents for this reaction as shown in Table 2.

Table 2 Conditions for the desulfurisation of ETP **6**



Conditions	Comments or Yields
Raney nickel, Ethanol, Reflux[112]	10% and decomposition products if amount of Raney nickel doubled
NaH, t-AmOH, $Ni(OAc)_2$ (5/2/1 ratio) THF Reflux[113]	No product
$NiCl_2.6H_2O$, $NaBH_4$, Ethanol, Reflux[114,115]	75%

Nickel-containing complex reducing agent (NiCRA) has been developed by Caubère and co-workers as powerful desulfurisating agent with a greater chemo- and stereoselectivity than Raney nickel, as well as being much easier to handle.[113] Its

reducing abilities is believed to rely on single electron transfer (SET) from NaH whose basicity is decreased and SET ability increased by the presence of a sodium alkoxide and metal salts.[116] This methodology was surveyed to desulfurise ETP **6** (Table 2) but unfortunately no product was observed by NMR or MS analysis.

Finally, the use of *in situ* generated nickel boride, which also has some precedent in desulfurisation reactions, was attempted although the utility of this method for disulfides was reported to be diminished by the formation of sulfides as by-products.[115] The reaction of nickel chloride salts with sodium borohydride is believed to form a nickel boride complex Ni_2B (of observed dark colour), which is thought to be the active reducing agent. It was shown that the extruded sulfur atoms remained on the metal, explaining the requirement for a stoichiometric amount of nickel.[117] While the precise mechanism is unclear, it has been suggested by labelling experiment that the dihydrogen generated by the reduction of Ni(II) by $NaBH_4$ could be adsorbed and activated at the nickel boride surface followed by formation of transient nickel hydride intermediate.[117] Gratifyingly, the use of this reagent finally afforded the desulfurised product **18** in good yield (75%).[115]

Isolation of chaetocin and semi-synthesis of analogues

Production and isolation of chaetocin

The production of chaetocin was investigated by culturing a strain of the fungus (*Chaetomium virescens* var. *thielavioideum*) under various conditions.[83,97] First, a conidial suspension (spores) of the fungus was prepared in water and was plated on solid complete media and incubated for 5 days at 37 ˚C (Figure 11).[118]



Figure 11 Culture of *Chaetomium virescens* var. *thielavioideum*

From those plates a conidial inoculum was prepared in distilled water, which was then inoculated to a liquid complete media and cultured at room temperature for a week. As secondary metabolites (like chaetocin) are thought to be part of the defence mechanism of microorganisms, it was expected that the metabolite should be located outside the cell (*i.e.* in the supernatant). The fungal suspension was therefore filtered to separate the supernatant from the mycelia. Two different solvents (CH_2Cl_2 and EtOAc) were used to extract the supernatant. Unfortunately, HPLC-MS analysis did not show any trace of chaetocin in the extracts, or in the supernatant. Accordingly, we decided to extract the mycelia as well. HPLC-MS analysis showed traces of chaetocin in the mycelia extract. As the signal was very weak it was decided to culture the fungus on a larger scale and to switch to a solid media as it has been more commonly described in the literature.[83,96] *Chaetomium virescens* var. *thielavioideum* was then cultured on solid complete media (5 plates) for two weeks and then extracted with CH_2Cl_2. After purification by column chromatography, milligram quantities of a solid were obtained. ^{1}H NMR and LC-MS analysis were consistent with that of chaetocin, however, the amount of chaetocin isolated was so small that it was not possible to obtain a completely pure sample.

The incubation was thus repeated on a larger scale (ten 1 L bottles half filled with *ca.* 400 mL of solid complete media) (Figure 12).



Figure 12 Large scale production of chaetocin

After three weeks of incubation at room temperature, the cultures were extracted with CH_2Cl_2 and the extracts were purified by column chromatography and trituration in hexane. Pure chaetocin could also be obtained by precipitation (in a mixture of CH_2Cl_2 and hexane followed by sonication).

Pleasingly, approximately 200 – 400 mg of pure chaetocin were isolated, the structure of which was confirmed by NMR (^{1}H and ^{13}C), LC-MS and measurement of the specific rotation $[\alpha]_D^{20}$. For biological testing, chaetocin was further purified by HPLC as trace amount of an analogue with a higher number of sulfur atoms could be identified by MS. Indeed, Saito *et al.* described the isolation of chaetocin B (trisulfide bridge over one of the diketopiperazine ring) and chaetocin C (trisulfide bridge at both diketopiperazine rings) from the same fungal culture (Figure 13).[96]

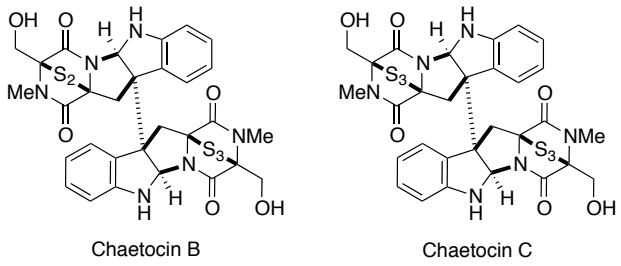
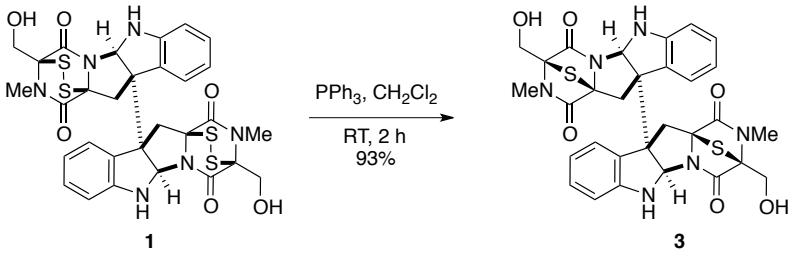


Figure 13 Structures of chaetocin B and C[96]

Derivatisation of chaetocin

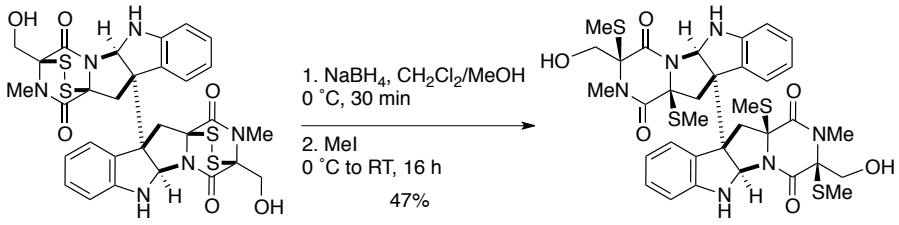
With a substantial amount of chaetocin in hand, the synthesis of the various analogues was undertaken. The monosulfide derivative **3** was first synthesised by treatment of chaetocin (**1**) with triphenylphosphine, as described for the ETP model compound **6**, in high yield (Scheme 13).[111] The stereochemistry of monosulfide **3** was assigned based on a range of chiroptical techniques (see Chapter III).



1 **3**

Scheme 13

The methylthioether derivative **4** was also prepared in moderate yield following an analogous method to that used for the synthesis of model compound **16** (Scheme 14).[111]

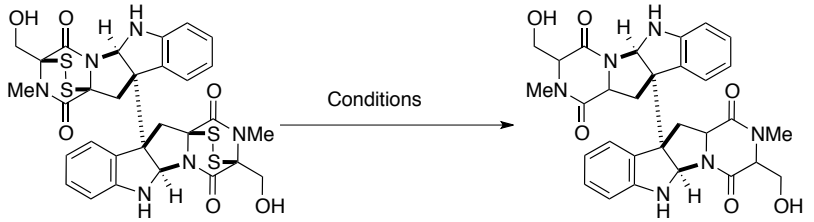


1 **4**

Scheme 14

The complete desulfurisation of chaetocin (**1**) was then attempted (Table 3). The optimal conditions developed for the desulfurisation of ETP **6**, a nickel boride complex formed *in situ* by reaction of $NiCl_2$ and $NaBH_4$,[115] were first tried but unfortunately, no desulfurised product **5** could be observed. Following the reported successful desulfurisation of gliotoxin, an aluminium amalgam was prepared,[119-121] but again, no product could be obtained. In both cases, a product of low molecular weight was observed, which lead to the suspicion that the bond connecting the two monomers of chaetocin was not stable to these conditions. The use of Raney nickel[111,120] was nevertheless attempted but did not afford the product in isolable quantity.

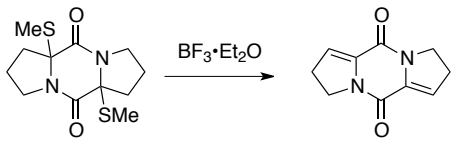
Table 3 Attempts for the complete desulfurisation of chaetocin



1 **5**

Conditions	Observations
$NiCl_2.6\ H_2O$, $NaBH_4$, $EtOH/H_2O$[115] (RT or 0 ˚C)	No expected product formation MS: Starting material + unidentified product (Mass ion = 308 g/mol)
Al-Hg amalgam, $EtOH/H_2O$[119-121], RT	No expected product formation MS: Starting material + unidentified product (Mass ion = 308 g/mol)
RaNi, EtOH (RT then reflux)	Traces of **5** by MS, Not isolable nor identifiable by ^1H NMR

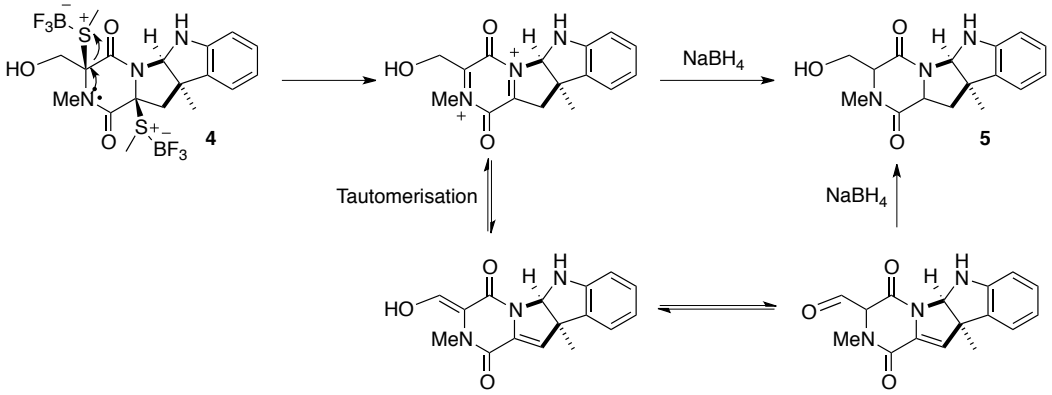
Another strategy to desulfurise the ETP core was proposed by Öhler and co-workers in their work on the synthesis of ETP-containing natural product.[122] It was shown that the treatment of the methylthioether **19** with boron trifluoride led to unsaturated compound **20** by elimination of methanethiol, most likely through the iminium species (Scheme 15).



19 **20**

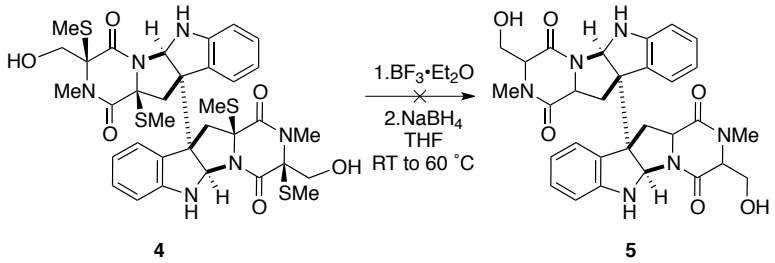
Scheme 15

It was hoped that a similar elimination would be possible for chaetocin methylthioether **4**, (Scheme 16, where only half of the molecule is presented for clarity). Reduction of the formed iminium functionalities (or the enamine and aldehyde tautomers) could then potentially lead to the expected desulfurised analogue **5**.



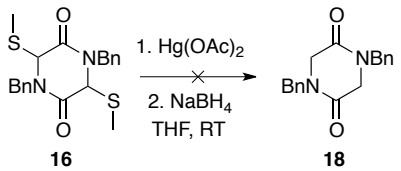
Scheme 16

Chaetocin methylthioether derivative **4** was treated with boron trifluoride etherate in THF (Scheme 17) until consumption of all the starting material. Half of the reaction mixture was treated with sodium borohydride, while the other half was directly quenched. Unfortunately, a mixture of unidentified decomposition products was obtained in both cases.



Scheme 17

$Hg(OAc)_2$ was also tested as Lewis acid tested for this desulfurisation reaction using the methylthioether model compound **16** (Scheme 18), but it did not yield the expected product **18** (starting material and decomposition products only).



Scheme 18

While we were conducting this work, Sodeoka and co-workers performed the total synthesis of the desired fully desulfurised analogue of chaetocin **5** and they assessed its biological activity against G9a, which will be discussed in the next

section.[104] The preparation of this analogue was thus considered unnecessary and further synthetic work was abandoned.

The low quantities and the peculiar reactivity of chaetocin brought some difficulties in the synthesis and isolation of analogues. Nevertheless, two structurally important derivatives were successfully prepared: the monosulfide analogue **3** and the methylthioether **4**. Biological assessment of these two compounds was thought to be sufficient to allow us to address the role of the disulfide bridge in chaetocin's HKMT activity. In addition, it was also decided to assess the biological activity of the ETP **6** to test whether the presence of the intact ETP core (including the disulfide bridge) would lead to some inhibitory activity towards HKMTs.

Biological evaluation of chaetocin and analogues

Preliminary data against SUV39H1

The inhibitory activity of chaetocin and its derivatives was first assessed by a commercially available chemiluminescent assay kit against SUV39H1. This assay relies on the recognition of the methylation mark at lysine 9 by a specific 'primary' antibody (Figure 14). This primary antibody is in turn recognised by a secondary antibody, specific for the primary antibody. The secondary antibody is conjugated with the Horseradish peroxidase (HRP) enzyme. HRP is able to catalyse the oxidation of luminol by hydrogen peroxide resulting in an excited product, which will produce light upon decay (an enhancer is also added to increase the signal). This light, or chemiluminescence, is the signal monitored. A number of washes between each step allows for a specific signal to be obtained with low background only when the methyl mark is present in the given well. The sensitivity of the assay allows monitoring of the reduction in the signal if the methylation reaction has been successfully inhibited.

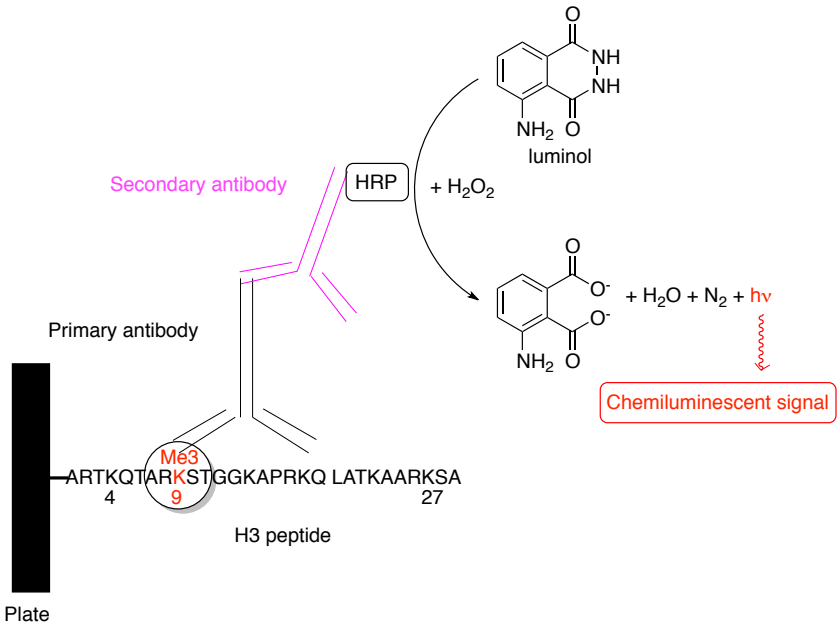


Figure 14 Principle of the chemiluminescent assay against SUV39H1

Chaetocin (**1**), the methylthioether derivative **4**, the monosulfide derivative **3** and ETP **6** were thus assayed against SUV39H1 (Figure 15).

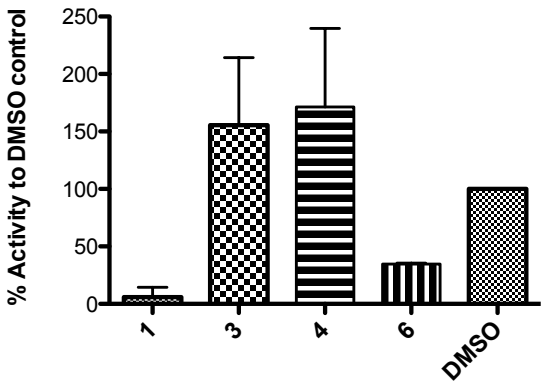


Figure 15 Inhibitory activities of compounds **1**, **3**, **4** and **6** at 20 μM in the SUV39H1 assay kit
The assay was run at a single concentration (20 μM) in duplicate, error bars represent SEM.

Interestingly, while chaetocin (**1**) did confirm a good inhibitory activity (95% inhibition), no activity could be observed for the derivatives of chaetocin **3** and **4** (an unexplained increased in activity was actually observed and this will be further discussed in Chapter IV). Remarkably, the simple ETP **6** showed a considerable inhibitory activity (74% inhibition). These results suggested that the disulfide bridge of the ETP core was necessary for activity and potentially sufficient for inhibition since the ETP **6**, lacking chaetocin's complex framework, retained significant inhibitory activity.

We were therefore interested in further characterising the inhibitory mechanism of ETP-containing molecules towards HKMTs (*e.g.* time-dependency, reversibility, denaturation). Unfortunately, the assay kit was unsuitable to further assess our derivatives. The peptide being covalently attached to the plate and the amount of reagent given limited, no variation of the assay conditions was possible.

Development of an AlphaLISA assay

In order to further explore the mechanism of inhibition of HKMTs by chaetocin and other ETP compounds, it was decided to develop an AlphaLISA assay, which would offer more flexibility in the assay settings to assess various parameters.

Description of the assay

The AlphaLISA assay is a homogeneous proximity immunoassay using Alpha donor and AlphaLISA acceptor beads (the acronym "Alpha" stands for amplified luminescent proximity homogeneous assay). As shown in Figure 16, a biotinylated

histone H3-derived peptide is used as substrate. After incubation with the cofactor (SAM) and the methyltransferase enzyme, the methylated substrate is detected by addition of a specific antibody (anti-H3K9meX). This antibody can either be bound to Protein A coated AlphaLISA acceptor beads (Protein A is known to bind strongly to antibodies) or alternatively directly coated on AlphaLISA acceptor beads. Streptavidin or *Strep*-Tactin coated AlphaLISA donor beads bind to the biotin tag of the peptide substrate thus bringing the two types of bead in close proximity.

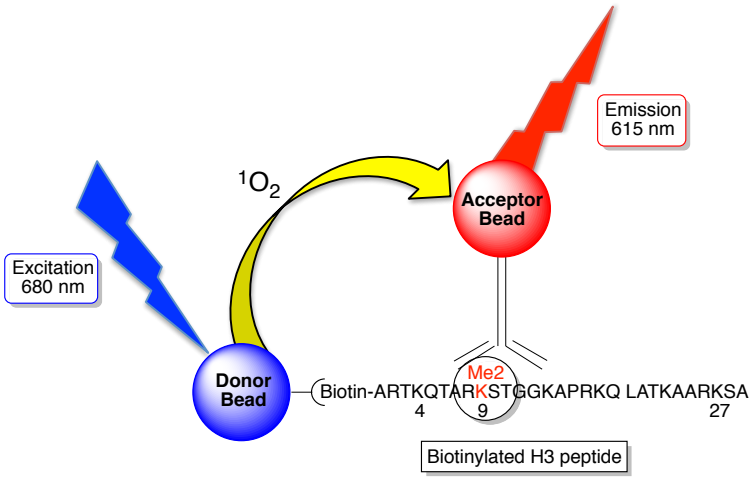


Figure 16 Schematic representation of the AlphaLISA assay principle (adapted from [123])

Donor beads contain a photosensitizer, phthalocyanine (blue-green-coloured aromatic macrocyclic compound) potentially coordinating a metal ion (Figure 17),[124,125] which converts ambient oxygen to singlet oxygen, upon illumination at 680 nm.

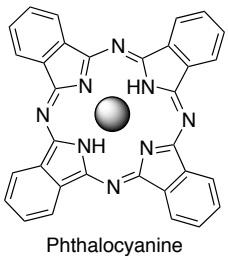


Figure 17 Structure of phthalocyanine
The grey ball represents the metal ion (*e.g.* Zn^{II}, Al^{III})

Singlet oxygen has a limited lifetime prior to falling back to ground state: its 4 µs half-life allows it to diffuse approximately 200 nm in solution. If the acceptor

beads are in close proximity with the donor beads (<200 nm, *i.e.* bound to the same peptide), energy will be transferred from the singlet oxygen to the acceptor bead. AlphaLISA acceptor beads contain thioxene and a europium chelate (Figure 18). The singlet oxygen emitted by the donor beads leads to the conversion of thioxene into a di-ketone derivative and to emission of light at 340 nm. This light will in turn excite the europium chelate will be then generating an intense light detectable within a narrow wavelength bandwidth centred around 615 nm, which will be the detected signal.

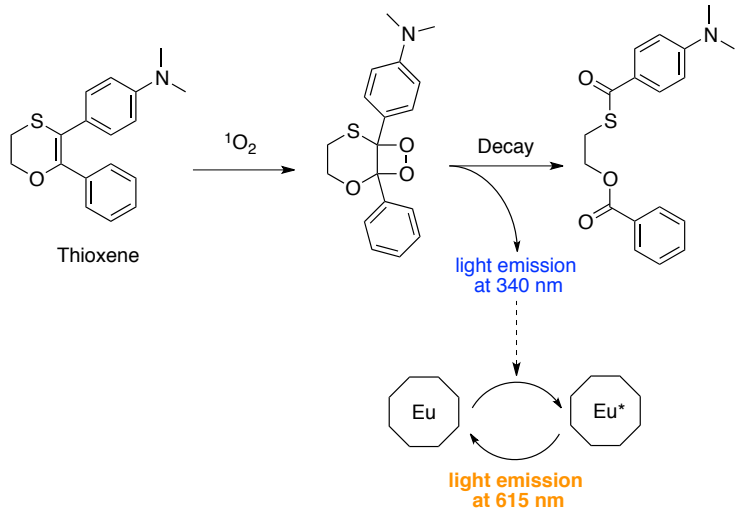


Figure 18 Chemistry of the AlphaLISA acceptor bead

If the acceptor beads are not in close proximity (*i.e.* not bound to the peptide *via* the antibody recognition), singlet oxygen falls to ground state and no signal is produced. Consequently, the intensity of light emission is proportional to the level of peptide modification and this allows quantification of the methylation.

Preliminary assay development with anti-H3K9me3 antibody and protein A coated beads

As aforementioned, SUV39H1 is a trimethylase enzyme and measuring its activity thus required the use of an anti-trimethyl-Histone H3 Lysine 9 (H3K9me3) antibody. The anti-H3K9me3 AlphaLISA acceptor Beads are not commercially available and it was thus necessary to use an anti-H3K9me3 antibody in combination with Protein A coated acceptor beads (Figure 19).

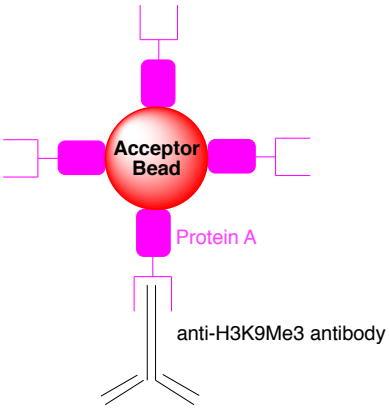


Figure 19 Protein A coated acceptor beads binding to anti-H3K9me3 antibody

First a Protein A – anti-H3K9me3 antibody titration was performed to investigate which concentration of the reagents would give the highest specific signal. The assay was performed using a trimethylated (H3K9me3) peptide and without the peptide as a negative control. The ratio of specific to non-specific signal (value obtained without peptide) was found to be around 2-fold for all the concentration systems tested (Figure 20).

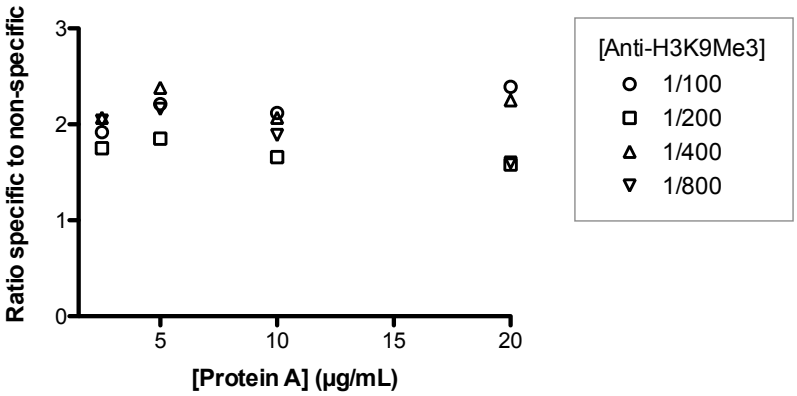


Figure 20 Ratio specific to non-specific signal at various concentrations of antibody and protein A
The assay was run as single points, with H3K9me3 peptide (1 µM), Protein A (2.5 – 20 µg/mL), anti-H3K9me3 (1/800 – 1/100).

Despite the fact that these results were not very satisfactory and gave only a very small window of opportunity to detect inhibition, we nevertheless proceeded with a peptide titration (using H3K9me3 peptide) to define the adequate concentration of peptide substrate to use to get the highest signal while avoiding the "hook effect".[126] This effect is commonly observed for bimolecular detection systems involving saturable reagents (*e.g.* the beads) used to bind specific partners (*e.g.* the H3 peptide). Saturation of the beads results in a characteristic signal decrease which is

observed after an initial concentration-dependent signal increase. Figure 21 gives a schematic explanation of the origin of this effect.

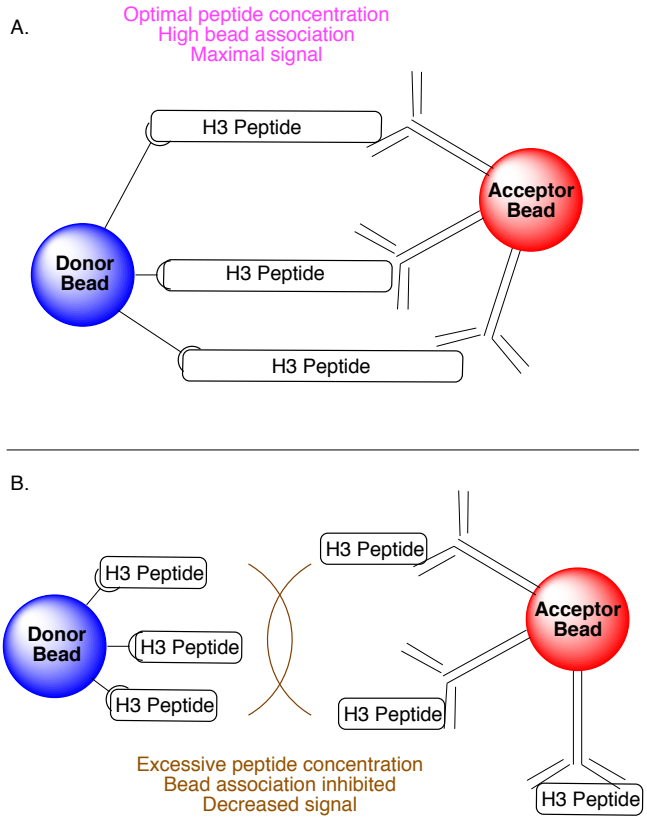


Figure 21 Schematic representation of the hook effect. A. Optimal concentration of peptide: maximal bead association leading to maximal signal. B. Hook effect: Excessive peptide concentration and decreased signal due to saturation of the binding sites and poor bead association[126]

In parallel, the unmethylated peptide was also titrated as a negative control to assess the specificity of the antibody for the trimethylation mark. To our surprise, the signals obtained for unmethylated and trimethylated peptides were similar (Figure 22), showing a total lack of specificity.

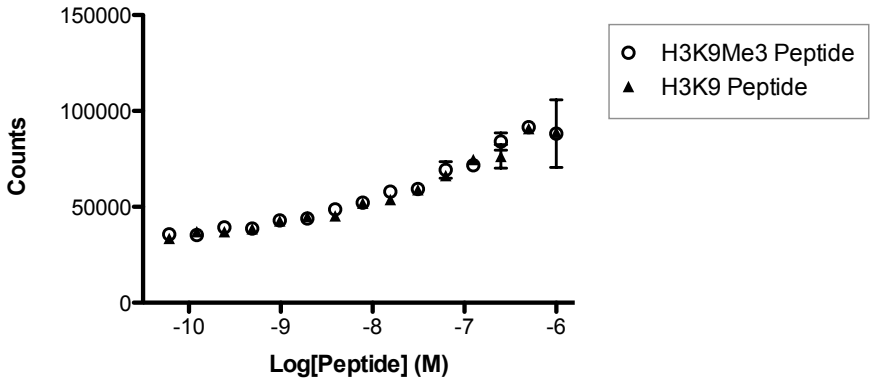


Figure 22 Peptide titration with both trimethylated (H3K9me3) and unmethylated peptide
The assay was run with Protein A (20 µg/mL), anti-H3K9me3 (1/100) and peptide (0.06 nM – 1 µM). The data shown is mean of duplicates; error bars represent SEM.

This technical issue is either attributable to the poor specificity of the anti-H3K9me3 antibody or alternatively to a reduced binding between the protein A and the antibody (leading to nonspecific binding).

The assay setting therefore had to be modified. Surveying the commercially available acceptor beads, anti-dimethyl-Histone H3 Lysine 9 (H3K9me2) AlphaLISA acceptor beads were considered. Firstly, the use of acceptor beads already coated with antibody would avoid the lengthy optimisation of a two-component system as is required for the antibody and Protein A setup (and considerably diminishes the risk of poor quality reagents). Secondly, G9a is responsible for the dimethylation of lysine 9 on histone 3 and is also a reported biological target of chaetocin.[67,74,104] In light of the fact that the two enzymes methylate the same residue and are highly similar in sequence and in structure,[63] it was reasoned that the mechanism of action of chaetocin toward this two enzymes would most likely be the same. Finally, SUV39H1 has been reported to trimethylate lysine 9 on the histone 3 in a non-processive way (*i.e.* releasing the peptide after each methylation step).[68] It seemed plausible therefore that the H3K9me2 mark would be detectable, in the right time window.

Consequently it was decided to use the anti-H3K9me2 AlphaLISA acceptor beads first with G9a to develop the assay conditions and then to verify whether detection could be achieved when using SUV39H1 as the methylating enzyme.

Preliminary assay development with anti-H3K9me2 coated acceptor beads

In light of the previous experience with the antibody – Protein A complex, the specificity of the antibody coated acceptor beads was first tested. The H3K9 dimethylated peptide was therefore titrated together with unmethylated and trimethylated peptides as controls (Figure 23).

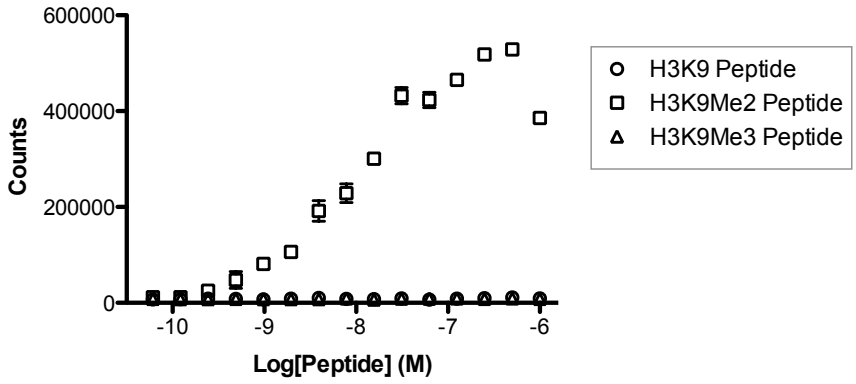


Figure 23 Peptide titration with tri-, di- and unmethylated H3 peptide
The assay was run with peptide (0.06 – 1000 nM) and anti-H3K9me2 coated acceptor beads (20 µg/mL). The data shown is mean of duplicates; error bars represent SEM.

The signal obtained was much stronger (up to half a million counts) than with the previous conditions, and importantly the specificity toward the dimethylated peptide was remarkable. The "hook effect" can be observed for peptide concentrations over 500 nM (Figure 23). For the subsequent experiments, the H3 peptide was therefore used at a concentration below 500 nM. For IC_{50} determination (see next section), the peptide was used at 100 nM in accordance with our titration results and as suggested by the manufacturer.

Next, in light of chaetocin's reactivity, we felt it was important to check whether the natural product and its analogues would interfere with the assay components. H3K9me2 peptide was then used with increasing amount of compounds (Figure 24).

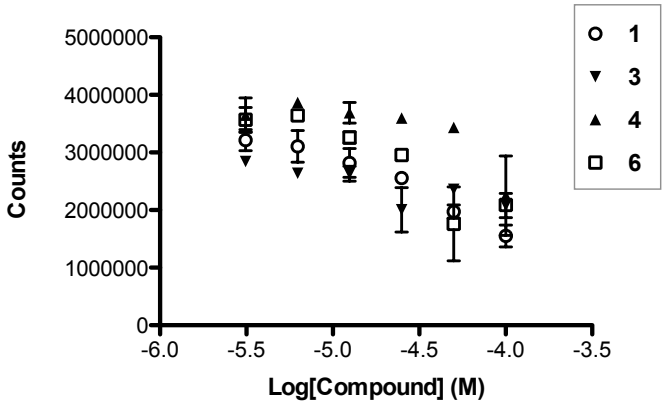


Figure 24 Influence of compounds on the assay reading
The assay was run with H3K9me2 peptide (500 nM) and compounds **1**, **3**, **4** and **6** (3.12 – 100 µM) while keeping the final amount of DMSO constant (1%). The data shown is mean of duplicates; error bars represent SEM.

As shown in Figure 24, a decrease in signal was observed for all compounds at higher concentration. This is commonly observed in biochemical assays, especially when very high concentrations of compounds are used (up to 100 µM here). It is nonetheless important to keep this in mind if a weak inhibition is observed at very high compound concentration.

Assay optimisation for G9a

Since the use of the anti-H3K9me2 coated acceptor beads gave suitable preliminary results (specificity for the dimethylated peptide and limited interference with our compounds), we undertook the optimisation of the other assay variables first using G9a as the methylating enzyme.

We initially sought to establish the optimal enzyme concentration and reaction time for the assay. A time course experiment was therefore performed at various enzyme concentrations. As the AlphaLISA assay is a stopped assay (non continuous), it was necessary to run several reactions in parallel and to stop them at a given time point. For practical reasons, the reactions were actually started with various delays by addition of the substrates at different time points to allow all reactions to be stopped at the same time, which allows the whole plate to be read at once (Figure 25).

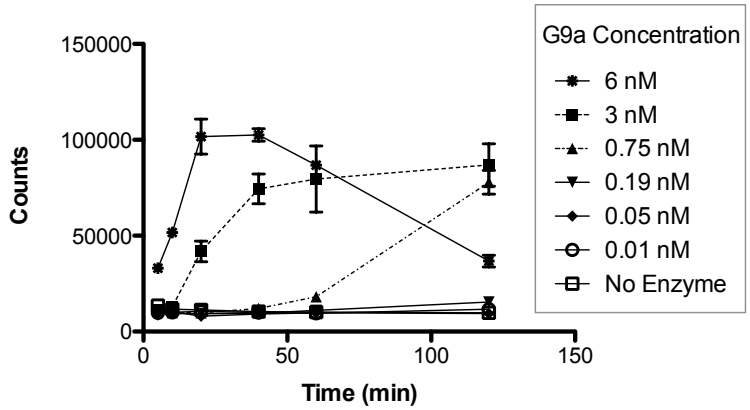


Figure 25 Time course at different G9a concentrations
The assay was run with H3K9 peptide (100 nM), SAM (20 µM) and G9a (0.01 – 6 nM). The data shown is mean of duplicates; error bars represent SEM.

The graph corresponding to the G9a concentration of 6 nM has the expected shape, with a decrease of the signal after a certain time, indicating the formation of the (undetected) trimethylated lysine (G9a has been reported to perform also trimethylation *in vitro*).[123,127,128] In addition, in order to remain in the linear range, the incubation time should be less than 20 min at this enzyme concentration.

Next, the concentration of the cofactor SAM was investigated to determine its Michaelis-Menten constant, K_M. K_M is defined as the substrate concentration that gives half maximal velocity of an enzymatic reaction. It represents the affinity of an enzyme for its substrate (with the lower the K_M, the higher the affinity). Screening using biochemical assays is usually performed at substrate concentration close to their K_M to allow unbiased approach to determining inhibitor potency. A time course at different SAM concentrations was thus performed. The signal obtained (proportional to the amount of product formed) was plotted against time, and the slope obtained by linear regression was therefore proportional to the reaction rate for each cofactor concentration investigated (Figure 26).

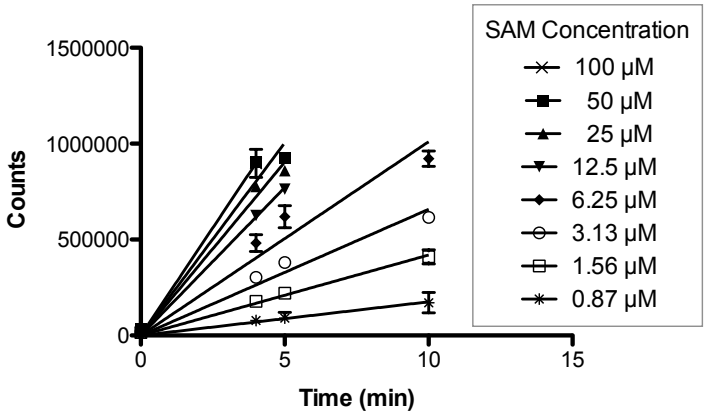


Figure 26 Time course with G9a at different SAM concentrations
The assay was run with G9a (6 nm), H3 peptide (100 nM) and SAM (0.87 – 100 µM). The data shown is mean of duplicates; error bars represent SEM.

The reaction rates obtained were then plotted against the concentration of SAM (Figure 27). Fitting of the data by non-linear regression into Michaelis-Menten equation (Equation 1) gave the expected K_M with a value of 8 µM (SE of 0.55), in close agreement with the manufacturer's data (K_M = 12 µM).[123]

$$v = \frac{\partial[P]}{\partial t} = \frac{V_{max} \times [S]}{K_M + [S]}$$

Equation 1 Michaelis-Menten equation

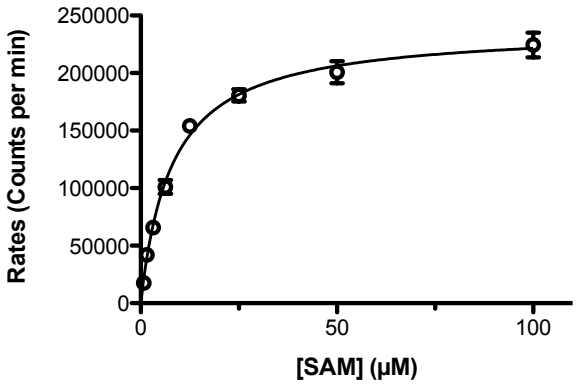


Figure 27 Michaelis-Menten curve: reaction rates plotted against SAM concentrations

A SAM concentration of 15 µM (close to K_M and recommended by the manufacturer[123]) was finally chosen for the subsequent experiments.

Finally, a new time course was run at this cofactor concentration with two concentrations of G9a (6 nM and 1 nM) to determine an adequate reaction time under these assay conditions. Also, the possibility of miniaturising the assay was

investigated: the volumes of all reagents were halved and the concentration of the beads was also halved (Figure 28).

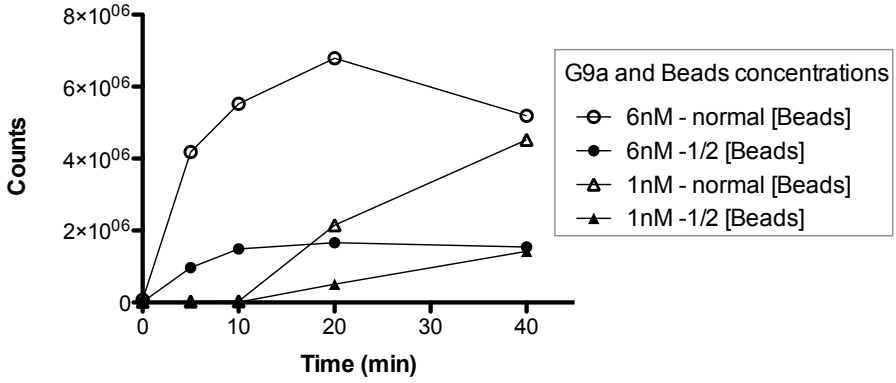


Figure 28 Time course with G9a at [SAM] = K_M
The assay was run as single points with H3 peptide (100 nM), G9a (1 or 6 nM) and SAM (15 µM). Two beads concentrations were investigated: normal (20 µg/mL for both acceptor and donor beads) and halved (10 µg/mL)

From this experiment, it was noticed that the linear range was too narrow for the high enzyme concentration (less than 10 min) hence it was decided to use G9a at the lower concentration (1 nM) for the subsequent experiments. Concurrently, as halving the volume of reagents did not seem to affect the assay and as the signal monitored with half the concentration of beads suggested in the supplier's protocol[123] was still very high (greater than a million counts), it was decided to use these conditions. A final time course was run under these optimised conditions to define the best incubation time (Figure 29).

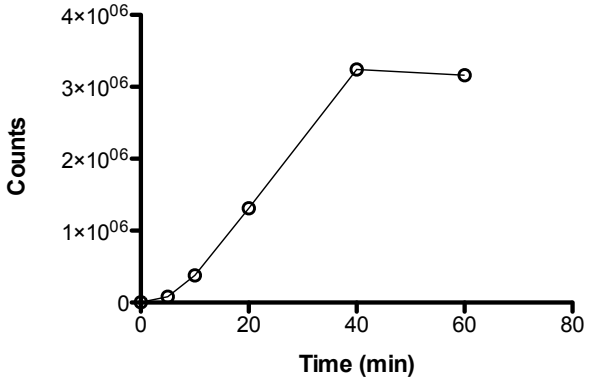


Figure 29 Time course under the selected conditions
The assay was run as single points with H3 peptide (100 nM), G9a (1 nM) and SAM (15 µM). Beads were used at 10 µg/mL.

From this experiment, a 40-minute incubation time was selected for the IC_{50} measurement against G9a in order to remain in the linear range and get the highest signal.

Assay optimisation for SUV39H1

A similar approach as that described for G9a has been followed to optimise the assay conditions using SUV39H1.

First, a time course at various enzyme concentrations was run. According to the supplier's datasheet, the specific activity (which could be defined as the amount of product formed by an enzyme in a given amount of time under given conditions per milligram of total protein, expressed in pmol/min/µg) of SUV39H1 is much lower than that of G9a (0.05 – 0.7 versus 10 pmol/min/µg, depending on the enzyme batch). In other words, the amount of active enzyme per total amount of protein is much lower in the case of SUV39H1. Consequently, a larger amount of SUV39H1 compared to G9a would likely be required for good signal production. A wider range of enzyme concentrations was thus explored (0.4 – 50 nM) as well as longer reaction times (up to two hours) (Figure 30).

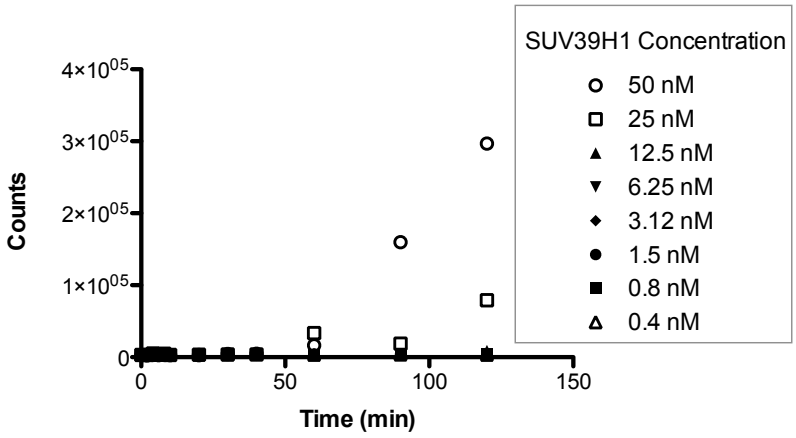


Figure 30 Time course at different SUV39H1 concentrations
The assay was run as single points with H3 peptide (100 nM), SAM (15 µM) and SUV39H1 (0.4 – 50 nM).

No signal could be detected for short time points or low enzyme concentration. It was apparent that the detection of the *transient* H3K9me2 mark would not be a problem as initially suspected, thanks to the weak activity of SUV39H1 (a trimethylase). Indeed, during the 2 h time course with 50 nM of SUV39H1, the dimethylation mark signal had not started to decrease (Figure 30) which constitute

evidence that further methylation of the dimethylated peptide has not significantly started.

The low signal detected could result from low enzyme activity or alternatively insufficient concentration of cofactor. We therefore decided to determine the K_M of SAM for SUV39H1 using a similar method as used for G9a and using SUV39H1 at high concentration (100 nM) (Figure 31). The value derived from the Michaelis-Menten curve for K_M was 42.3 µM (SEM 14.5).

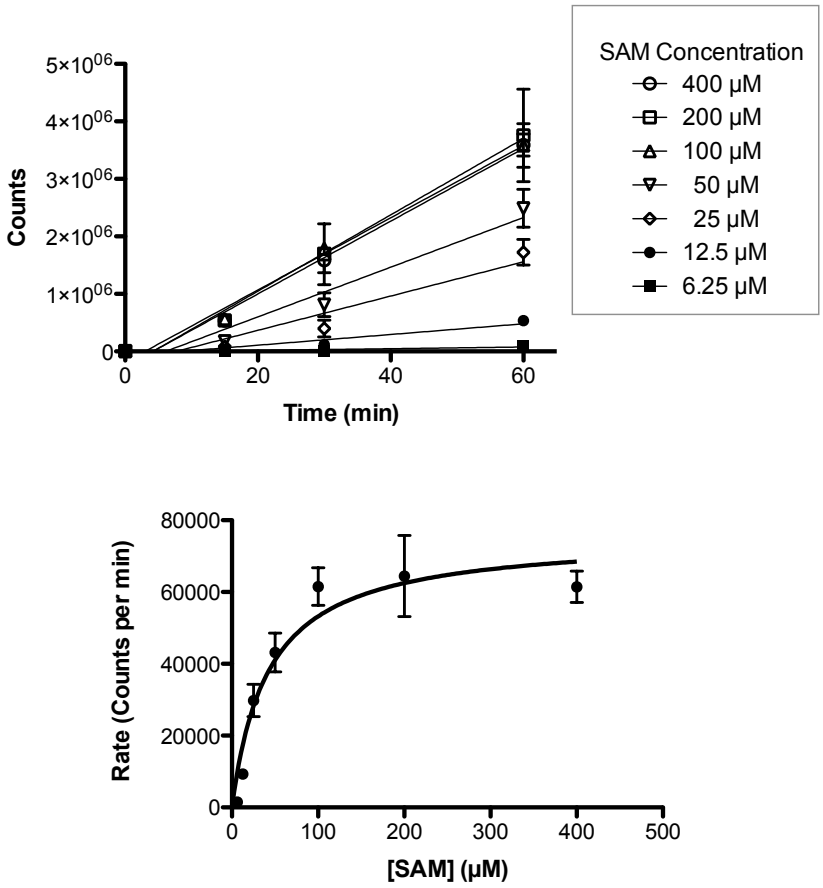


Figure 31 Michaelis-Menten kinetic analysis for determining the K_M of SAM for SUV39H1
The assay was run with SUV39H1 (100 nm), H3 peptide (100 nM) and SAM (6.25 – 400 µM). The data shown is mean of duplicates; error bars represent SEM.

Consequently, an additional time course was run with SAM concentration closer to its K_M (40 µM) to determine the enzyme concentration and the reaction time to use in the assay. As expected with a higher SAM concentration, a much higher signal was obtained (Figure 32). From this experiment, an enzyme concentration of

50 nM and a reaction time of 1.5 h were chosen for the IC_{50} measurement against SUV39H1.

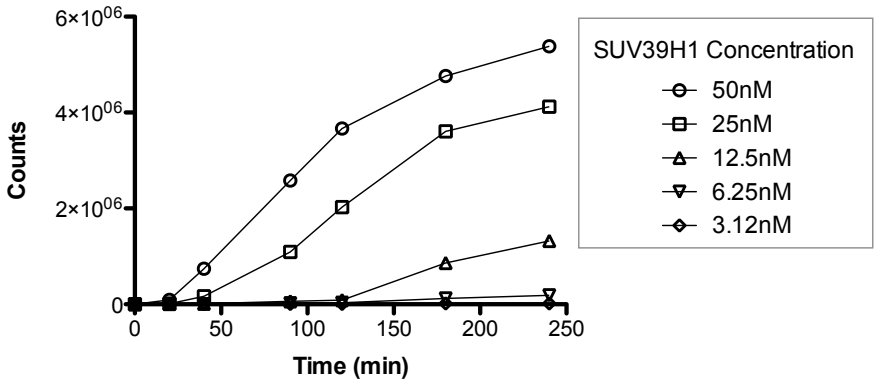


Figure 32 Time course at various concentration of SUV39H1 and [SAM] = K_M
The assay was run as single points with H3 peptide (100 nM), SAM (40 μM) and SUV39H1 (3.12 – 50 nM).

IC_{50} values determination

With the established optimised conditions for the assay, we next evaluated the inhibitory activity of our compounds against both G9a and SUV39H1.

Determination of IC_{50} values against G9a

Chaetocin (**1**), the monosulfide analogue **3**, the methylthioether analogue **4** and ETP **6** were thus assayed against G9a. SAH, the endogenous HKMT inhibitor of the cell, was used as a control. The inhibition curves are presented in Figure 33. pIC_{50} values were derived for the active compounds and their average is presented in Table 4. No IC_{50} value could be derived for the methylthioether analogue **4** since no inhibition was observed.

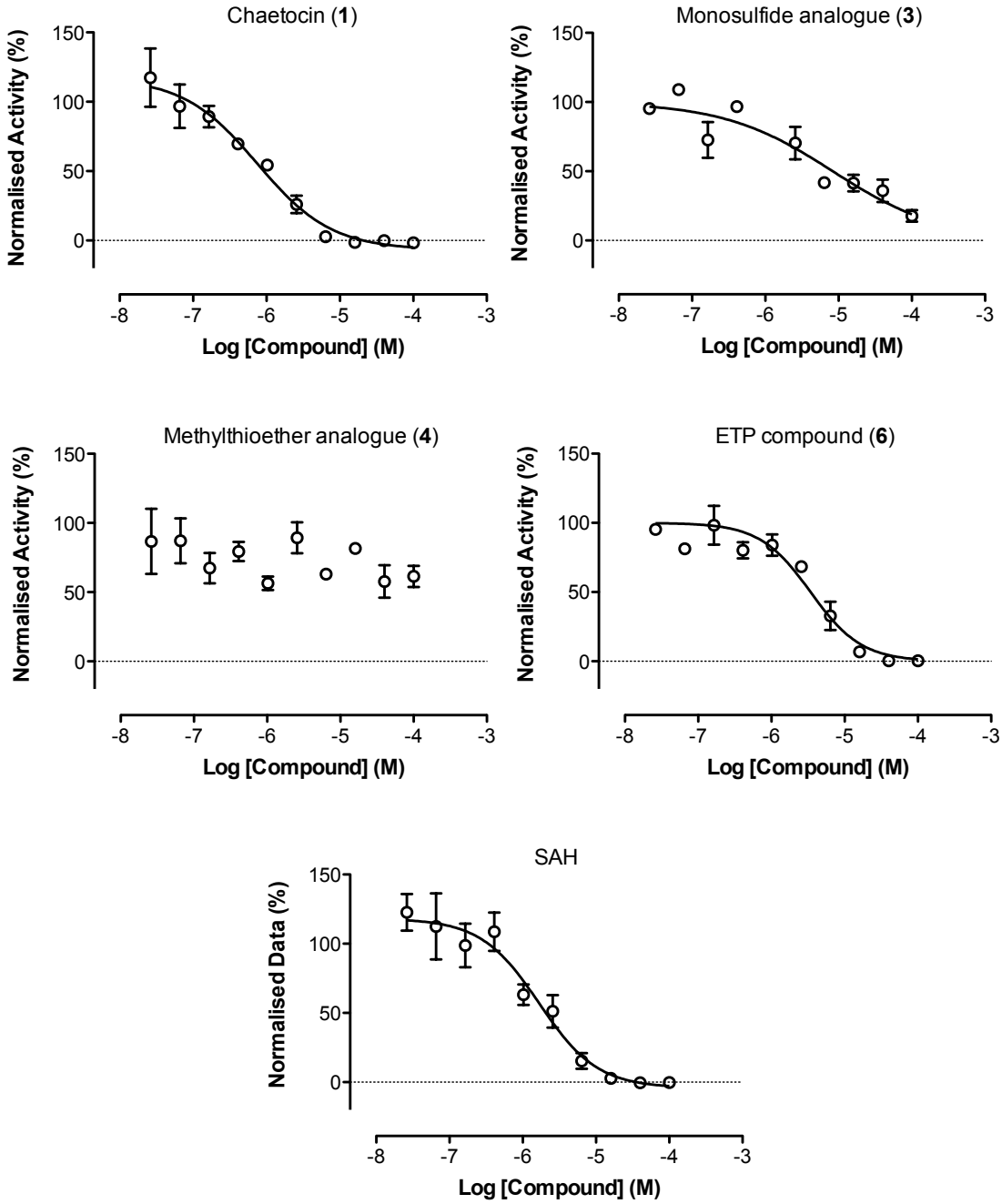


Figure 33 Concentration-inhibition curves for chaetocin and analogues against G9a. Adapted from Cherblanc *et al.*[129]
The assay was run with H3 peptide (100 nM), SAM (15 µM) and G9a (1 nM), with 30 min pre-incubation time and 40 min reaction time. Representative data set from one experiment, the data shown is mean of duplicates; error bars represent SEM.

The concentration-inhibition curve obtained for monosulfide analogue **3** does not present the expected sigmoidal shape (low Hill slope value) and even at a concentration of 25 µM, 36% enzyme activity on average remains (Table 4). The IC_{50} value derived from such a curve should be interpreted with caution. It should be noted that epi*mono*thiodiketopiperazines compounds such as analogue **3** are known to be

relatively unstable due to their highly strained sulfur containing ring.[130] The observed weak inhibition and low Hill slope value for this compound are most likely linked to its chemical reactivity. In addition, as mentioned earlier, interference with the assay setting at high concentration was observed with all the compounds tested, which could also partially account for the weak inhibition detected here.

Table 4 pIC_{50} values and characteristics of inhibition of G9a by compounds **1**, **3**, **6** and SAH. Adapted from Cherblanc *et al.*[129]

	1	**3**	**6**	**SAH**
pIC_{50} ± SEM (*n*=5)	5.59 ± 0.18	4.97 ± 0.29	5.31 ± 0.10	5.98 ± 0.10*
IC_{50} (μM)	2.6	10.6	4.9	1.0
Hill Slope ± SEM	-1.20 ± 0.17	-0.60 ± 0.09	-0.85 ± 0.12	-1.21 ± 0.08
% Activity at 25 μM (*n*=4)	8	36	19	0‡

IC_{50} values were determined from the average of the pIC_{50} values obtained in each independent experiment (run as duplicates), *n* (number of independent experiments); **n*=4 for SAH, ‡*n*=3 for SAH;

These results and the subsequent data on G9a inhibition have been accepted for publication in the Journal of Medicinal Chemistry in October 2013.[129]

Determination of IC_{50} values against SUV39H1

To corroborate the results obtained against G9a and for a direct comparison with the original literature report,[74] chaetocin (**1**), methylthioether analogue **4** and ETP **6** were also assayed against SUV39H1. The inhibition curves are presented in Figure 34.

The inhibitory activity of chaetocin (**1**) was also confirmed against SUV39H1: it was found to have significant inhibitory potency (IC_{50} = 0.11 μM) on purified human SUV39H1 in excellent agreement with the data of Greiner *et al.* (IC_{50} = 0.8 μM) against *d*SU(VAR)3-9, the *Drosophila* orthologue of SUV39H1.[74] Similar to the observations against G9a, the methylthioether **4** was found to be inactive while the structurally simplified ETP **6** was found to have significant inhibitory activity (IC_{50} = 3.2 μM).

These results and the following data generated against SUV39H1 have been published in Nature Chemical Biology in March 2013.[131]

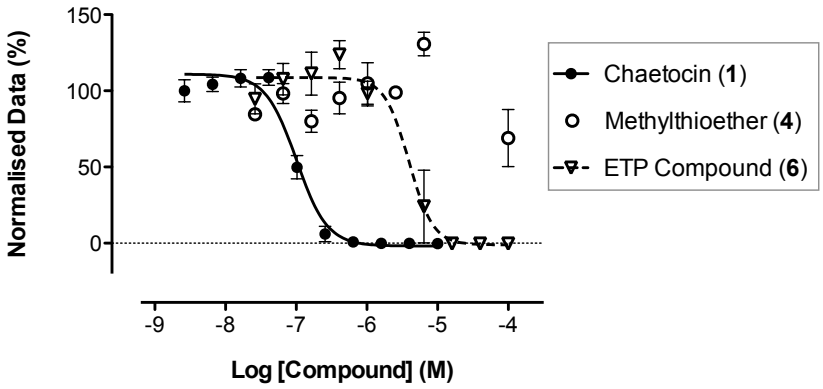


Figure 34 Inhibition curves for compound **1**, **4** and **6** against SUV39H1. Adapted from Cherblanc *et al.*[131] The assay was run with H3 peptide (100 nM), SAM (40 µM) and SUV39H1 (50 nM), with 30 min pre-incubation time and 90 min reaction time. Data shown are collated data of n=3 independent experiments run as duplicates, the data shown is mean of replicates; error bars represent SEM.

These data were in good agreement with our preliminary results against SUV39H1 and further confirmed that the disulfide bridge of chaetocin is required for its HKMT inhibitory activity, against both SUV39H1 and G9a. The simple ETP compound **6** confirmed its activity against both enzymes, highlighting the fact that the ETP core appears to be sufficient for inhibition of HKMTs.

Mechanistic studies

Cognisant of the requirement of the disulfide moiety, we next sought to establish the mechanism of action by which chaetocin mediates its inhibitory activity against HKMTs. As aforementioned, a number of mechanisms have been suggested to underlie the biological activities of ETP-containing natural products (Figure 9).

Reactive oxygen species generation test

ETP-containing compounds have been reported to generate reactive oxygen species (ROS) by redox cycling in presence of reducing agent (Figure 9B).[132,133] We therefore decided to assess the presence of ROS under our assay conditions (typically with 1 mM DTT in the buffer) by evaluating the presence of *in situ* generated H_2O_2.[134] The protocol we used has been developed by Johnston *et al.* and it is a colorimetric assay which relies on the ability of the HRP enzyme to catalyse the oxidation of phenol red by hydrogen peroxide into an oxidised product which absorbs light at 610 nm (Figure 35).[135]

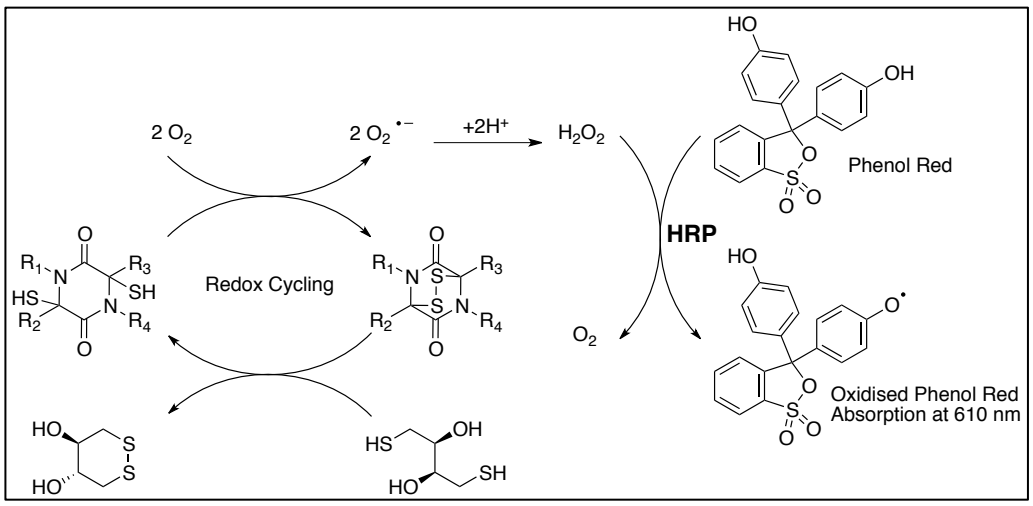


Figure 35 Assay principle for the detection of H_2O_2 production and detection *via* redox cycling

The data were normalised to the hydrogen peroxide control and IRC-083864 (a quinone[136] previously shown to undergo redox cycling[137]) was used as a positive control. As expected the quinone did lead to a strong absorbance signal, which was efficiently abolished by addition of catalase (also known as hydroperoxidases, enzymes which catalyse the decomposition of hydrogen peroxide to water and oxygen)[138] (Figure 36).

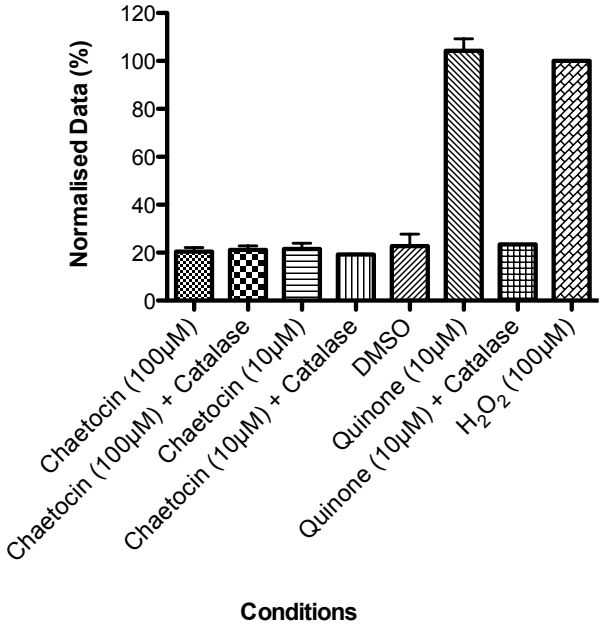


Figure 36 H_2O_2 generation test. Adapted from Cherblanc *et al.*[129]
The assay was performed as described (see experimental section),[134] and was run as duplicates, error bars represent SEM. Data were normalised to H_2O_2 control

No generation of H_2O_2 was detected at 10 or at 100 µM of chaetocin (Figure 36), which indicates that the *in vitro* inhibitory activity of chaetocin towards HKMTs is not dependent on the generation of ROS.

Influence of pre-incubation time on inhibition

Since cross-linking with cysteine residues is another mechanism of action proposed for ETP-containing natural products (Figure 9A), we reasoned that such effect would most likely be dependent on the pre-incubation time allowed to the compound to react with the enzyme.

It was thus decided to check whether chaetocin's inhibition was time-dependent or not. The IC_{50} of chaetocin was derived after various pre-incubation times and under different reducing conditions. Previously, Greiner *et al.* reported that the HKMT activity of chaetocin was not dependent on the presence of increasing concentrations of DTT,[74] and we therefore felt it was important to further assess the influence of DTT on the supposed time-dependency.

The time-dependency of G9a inhibition by chaetocin (**1**) was first assessed both in the presence and absence of DTT. Inhibition was measured after a variety of pre-incubation time-points (0, 5, 15, 30 min), where the enzyme and inhibitor had been pre-mixed (Figure 37). Inhibition by SAH and ETP **6** was also measured at 0 min and 30 min pre-incubation, with and without DTT, as controls.

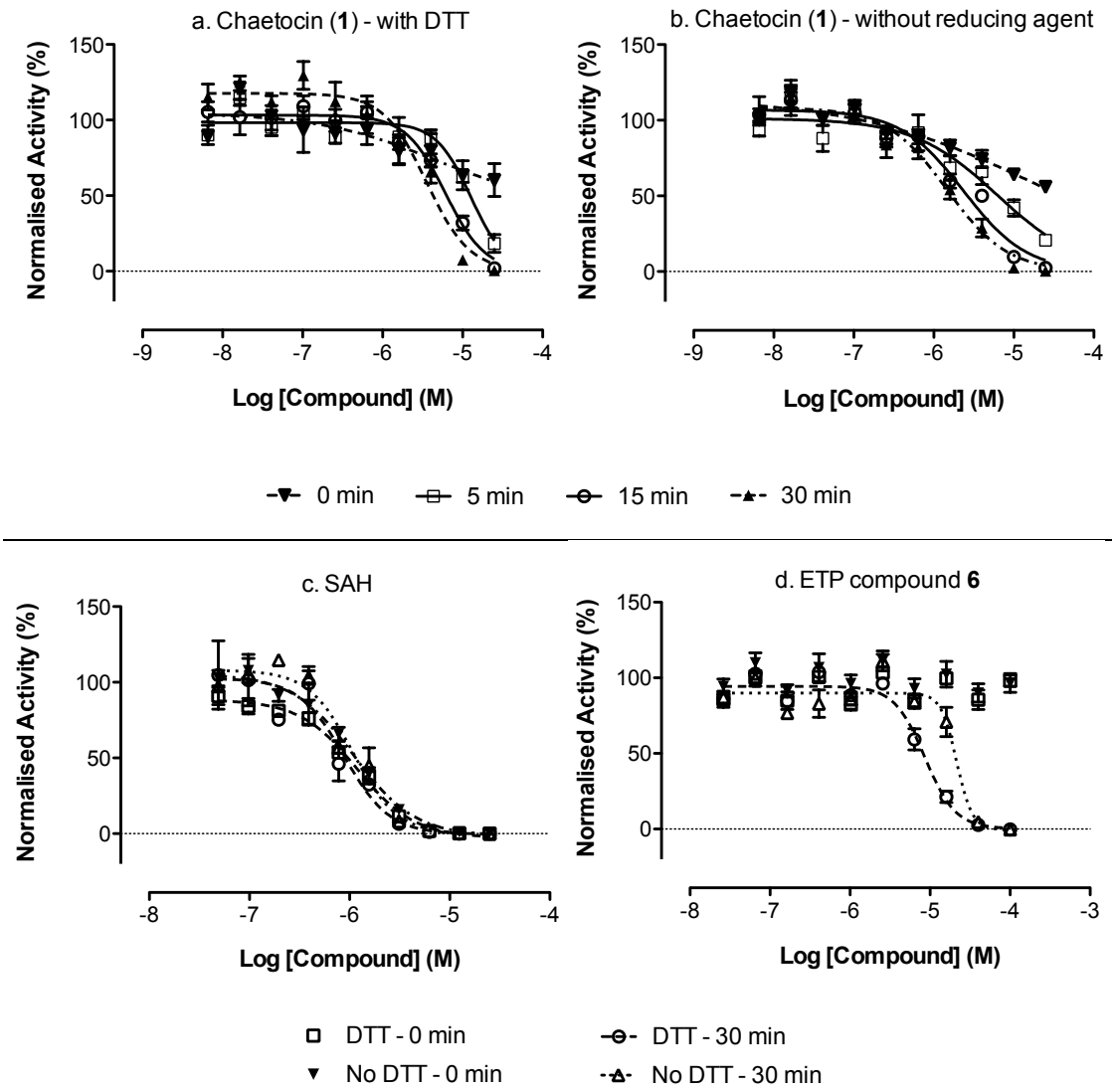


Figure 37 Influence of pre-incubation on the inhibitory activity of chaetocin (**1**), ETP **6** and SAH against G9a. Adapted from Cherblanc *et al.*[129]
a) and b) the effect on chaetocin (**1**) potency with 5, 15, 30 min or without pre-incubation time before measuring G9a activity inclusive and exclusive of DTT (1 mM in the buffer); c) the effect of pre-incubation of SAH or d) compound **6** with G9a, at 0 or 30 min pre-incubation, inclusive and exclusive of DTT. Collated data from n=3 independent experiments run as duplicates, the activity was normalised to DMSO control and the data shown is mean of replicates and error bars represents SEM.

Strikingly, the potency of chaetocin significantly increased with longer pre-incubation with G9a, displaying only weak inhibition when the substrate was added to the enzyme prior to the inhibitor (*i.e.* at 0 min pre-incubation). The difference in pIC_{50} values for chaetocin between 0 and 30 min of pre-incubation time was found to be statistically significant, with and without DTT (Table 5). ETP **6** showed analogous behaviour with little inhibition detected without pre-incubation, with and without DTT (Figure 37d). Notably, SAH did not exhibit such behaviour (Figure 37c).

Interestingly, the presence of DTT as a reducing agent did not have a statistically significant influence on the HKMT inhibitory activity of chaetocin, which corroborates the initial report by Greiner *et al.*.[74]

Table 5 Time dependency effect of chaetocin on pIC_{50} values against G9a with (A) and without DTT (B). Adapted from Cherblanc *et al.*[129]

Table A	With DTT			
Pre-incubation time	0 min	5 min	15 min	30 min
pIC_{50} (n=3) ± SEM	4.43 ± 0.34	4.90 ± 0.05	5.22 ± 0.05	5.45 ± 0.07
IC_{50} (μM)	37.2	12.7	6.1	3.5
Difference in pIC_{50} with 0 min	—	0.29 (ns)	0.60 (*)	0.81 (**)

Table B	Without reducing agent			
Pre-incubation time	0 min	5 min	15 min	30 min
pIC_{50} (n=3) ± SEM	4.71 ± 0.15	5.21 ± 0.11	5.63 ± 0.06	5.84 ± 0.08
IC_{50} (μM)	19.5	6.2	2.3	1.5
Difference in pIC_{50} with 0 min	—	0.60 (**)	0.97 (***)	1.19 (***)

pIC_{50} values were derived for 0, 5, 15 and 30 min pre-incubation and are average of n=3 independent experiments run as duplicates. P values were calculated by one-way ANOVA followed by Bonferroni's multiple comparison test. ***: $P < 0.001$ (Extremely significant), **: $0.001 < P < 0.01$ (Very significant), *: $0.01 < P < 0.05$ (Significant), ns: $P > 0.05$ (Not significant).

The time dependency of chaetocin inhibition was also assessed against SUV39H1 (Figure 38). In the presence of DTT, no time dependency could be observed while there was a slight increase in potency with longer pre-incubation time in the absence of reducing agent.

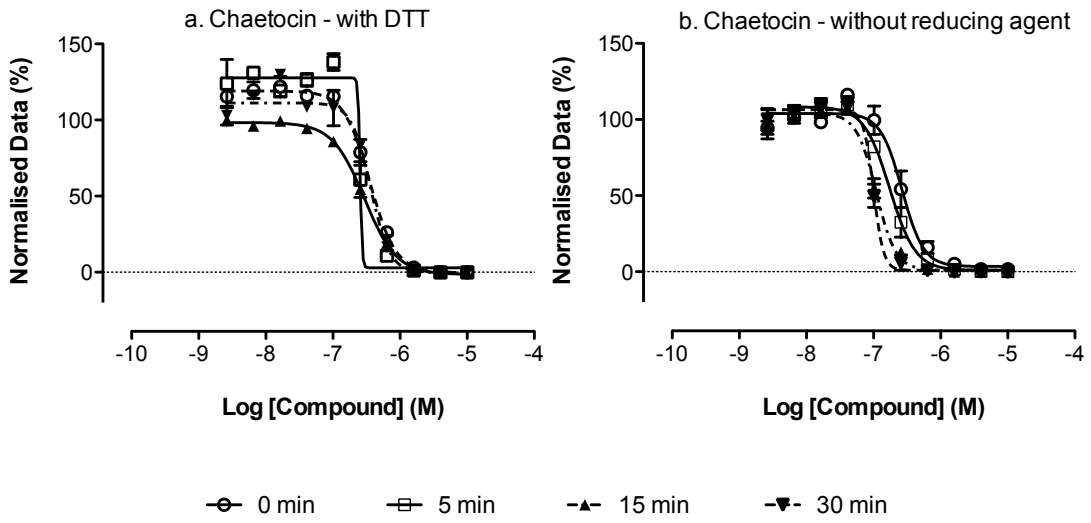


Figure 38 Influence of chaetocin pre-incubation on SUV39H1 activity. Adapted from Cherblanc *et al.*[131]
The effect on chaetocin (**1**) potency with 5, 15, 30 min or without pre-incubation time before measuring G9a activity a) inclusive and b) exclusive of DTT (1 mM in the buffer); Collated data from *n*=3 independent experiments run as duplicates, the activity was normalised to DMSO control and the data shown is mean of replicates and error bars represents SEM.

Table 6 Time dependency effect of chaetocin on pIC_{50} values against SUV39H1 without reducing agent

	Without reducing agent			
Pre-incubation time	0 min	5 min	15 min	30 min
pIC_{50} (*n*=3) ± SEM	6.33 ± 0.14	6.74 ± 0.08	6.86 ± 0.02	6.96 ± 0.05
IC_{50} (μM)	0.45	0.18	0.14	0.11
Difference in pIC_{50} with 0 min	—	0.41 (ns)	0.54 (*)	0.63 (*)

pIC_{50} values were derived for 0, 5, 15 and 30 min pre-incubation and are average of *n*=3 independent experiments run as duplicates. P values were calculated by one-way ANOVA followed by Bonferroni's multiple comparison test. ***: P < 0.001 (Extremely significant), **: 0.001 < P < 0.01 (Very significant), *: 0.01 < P < 0.05 (Significant), ns: P > 0.05 (Not significant).

Although more apparent in the case of G9a than SUV39H1, chaetocin's inhibitory activity was found to be increased with longer pre-incubation times, which is consistent with the cross-linking mechanism. Moreover, this effect appeared to be intensified in absence of reducing agent.

Reversibility of chaetocin's inhibition

Since the activity of chaetocin was found to be dependent on the presence of the chemically reactive disulfide ETP bridge and exhibited time dependence, we next wanted to assess whether such inhibitory effects were reversible. A "dilution test"[139] was used to assess the reversibility of G9a inhibition by chaetocin. In this test, G9a, at

100-fold over the typical assay concentration, was incubated for 30 minutes with the inhibitor, at a concentration of 10-fold the IC_{50}. The mixture was then diluted 100-fold into the reaction buffer containing the substrates. At this final dilution, the concentration of the enzyme has returned to the standard assay concentration, and the concentration of the inhibitor is 10 times lower than the IC_{50} (which should correspond to only 9% inhibition). For reversible inhibition, the enzyme should recover its activity. Controls experiments were run with the compound at 0.1 times the respective IC_{50} without dilution to illustrate that no significant inhibition is expected at this concentration.

In the presence of the reducing agent DTT, the inhibition by chaetocin (**1**) was reversible (Figure 39a), with G9a activity fully recovered. In the absence of such a reducing agent however, G9a did not recover its activity, which is strongly suggestive of irreversible inhibition under these conditions. ETP **6** displayed analogous behaviour with the reversibility of its inhibition being dependent on the presence of DTT (Figure 39c). As control, SAH inhibition was found to be reversible with and without DTT (Figure 39b). These data clearly show that the reversibility of G9a inhibition for ETP-containing compounds is different depending on the assay conditions. The fact that a reducing agent is able to impact the reversibility suggests that the bonding of the ETP compound to the enzyme is most likely *via* a mixed disulfide bridge (Figure 9A).

Unfortunately, it has not been possible to assess the reversibility of the inhibition of SUV39H1 by chaetocin. The technical issue stemmed from the commercial enzyme, which is formulated in a buffer containing 3 mM of DTT. To assess the effect of DTT on the reversibility of chaetocin's inhibition, it was first necessary to remove (or reduce) the amount of DTT present in the stock solution. Indeed, as the concentration of enzyme used for the reversibility test is initially 100-fold higher than the concentration used to measure the IC_{50} (50 nM), the stock enzyme (5 μM) had to be used directly (no dilution into the buffer with and without DTT), and consequently the amount of DTT would not be negligible as it was the case for G9a.

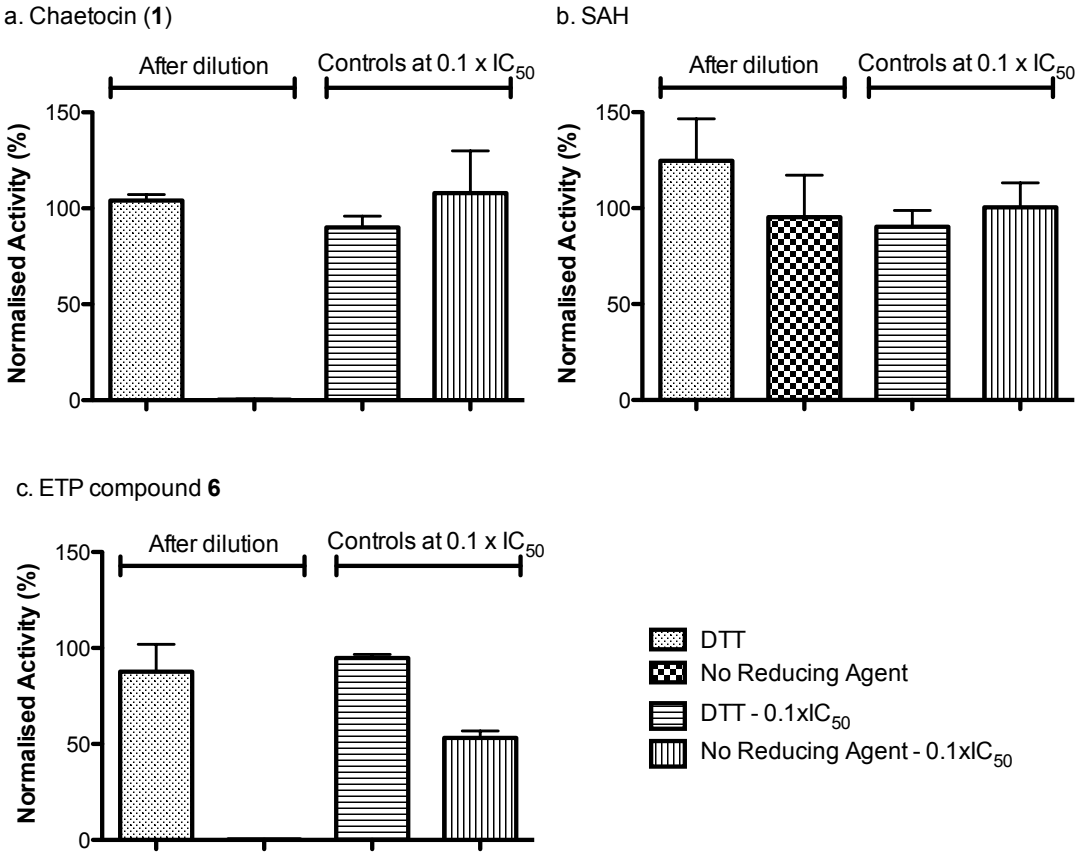


Figure 39 Dilution test to assess the reversibility of inhibition of G9a, with and without DTT. Adapted from Cherblanc *et al.*[129]
Data were normalised to DMSO controls. Experiments were run as duplicates. Data shown are mean values of n=3 independent experiments, error bars represent SEM. Controls experiments were run with the compound at 0.1 times the respective IC_{50} without dilution.

Two techniques were attempted to remove DTT. First the ZipTip technology was used: commercially available pipette tips with bed of chromatography media fixed at its end which are commonly used for concentrating and purifying samples, including desalting. The second technique attempted was through the use of a dialysis membrane. Unfortunately, both techniques led to the deactivation of the enzyme. Is it unclear whether the absence of DTT was responsible for this deactivation or whether the enzyme was simply not stable to these processes.

Enzyme concentration dependence for SUV39H1

In order to reduce the concentration of DTT for the reversibility test, we also attempted to reduce the enzyme concentration while increasing the cofactor SAM concentration to obtain a suitably high signal. Although it was not possible to decrease SUV39H1 concentration enough for the reversibility assay, it was noticed that the IC_{50} value appeared to be dependent on the enzyme concentration. The IC_{50} of

chaetocin was thus measured at a high (50 nM) and low (5 nM) concentration with a high concentration of SAM (4 mM) to increase the rate of the methylation reaction (Figure 40). A long reaction time (4.5 h) was required for the low enzyme concentration in order to obtain an acceptable signal.

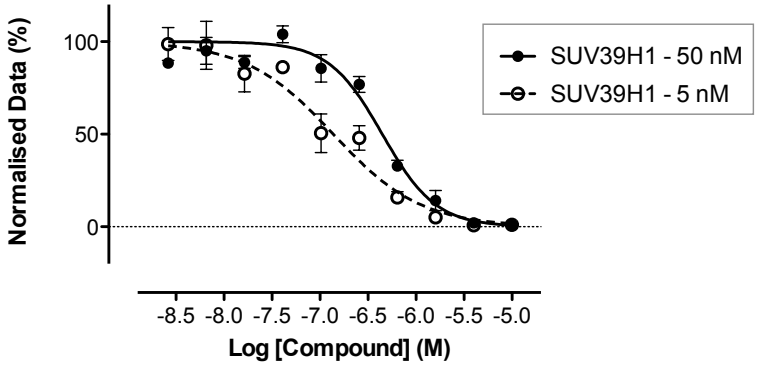


Figure 40 Influence of enzyme concentration on chaetocin's inhibition of SUV39H1. Adapted from Cherblanc *et al.*[131]
The assay was run with SAM at a concentration of 4 mM and SUV39H1 at 5 or 50 nM. Collated data of n=4 independent experiments, run in duplicates, the data shown is mean of replicates and error bars represent SEM.

A slight increase (2.4-fold) in potency was observed with the lower enzyme concentration compared to the higher one (IC_{50} of 0.34 μM for 50 nM SUV39H1 and 0.14 μM for 5 nM, statistically significant by t-test). This effect was not consistently observed in the case of G9a.

This type of behaviour was reported to be characteristic of a specific type of promiscuous inhibitors:[140] it was proposed that the actual active form of these promiscuous inhibitors might be an aggregate of many individual molecules. However, it was reported that the use of detergent would abolish the aggregate formation process and lead to a drop in the inhibitor potency.[141] The detergent Tween20 has been used throughout the studies on SUV39H1 hence it is unlikely that aggregate formation is responsible for the behaviour observed. It is therefore likely that this observation results from the nonspecific inhibition of the enzyme by chaetocin.

Competition test with SAM

While taken together our data were suggesting that the interaction of chaetocin with the two HKMTs under investigation was not a conventional inhibition (*e.g.* time-

dependency, irreversibility under non-reducing conditions), Greiner *et al.* reported however that chaetocin was competitive with the cofactor SAM. This implied a well-defined binding into the cofactor pocket and was not consistent with our emerging data. It was thus decided to first assess the influence of SAM concentration on chaetocin inhibitory activity against SUV39H1, at high (4 mM) versus usual (40 μM) SAM concentration. As shown (Figure 41), increasing the amount of SAM 100-fold did not influence chaetocin's inhibition.

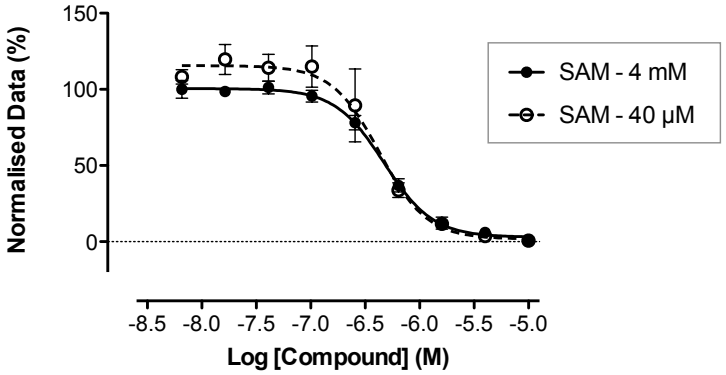


Figure 41 Inhibition curve of chaetocin against SUV39H1 at high and low SAM concentration. . Adapted from Cherblanc *et al.*[131]
The assay was run with SUV39H1 (50 nM) and SAM (40 μM or 4 mM). The data shown are collated data from n=3 independent experiments, run as duplicates. Shown as the mean of replicates and error bars represent SEM.

To further validate this piece of data, pIC_{50} values were then derived at a wide range of SAM concentrations against both SUV39H1 and G9a. The pIC_{50} values obtained were then plotted against $[SAM]/K_M$ (Figure 42) as this has been reported recently to assess SAM-competition.[81]

For both SUV39H1 and G9a, there was no variation in pIC_{50} values when SAM concentration was varied and this was further confirmed by linear regression (Figure 42, slope equivalent to zero). These data strongly indicate that chaetocin is not acting in a competitive manner with SAM for the inhibition of both enzymes. Statistical analysis of the pIC_{50} values (by one-way ANOVA followed by Bonferroni's test) also showed that the difference between them was not statistically significant.

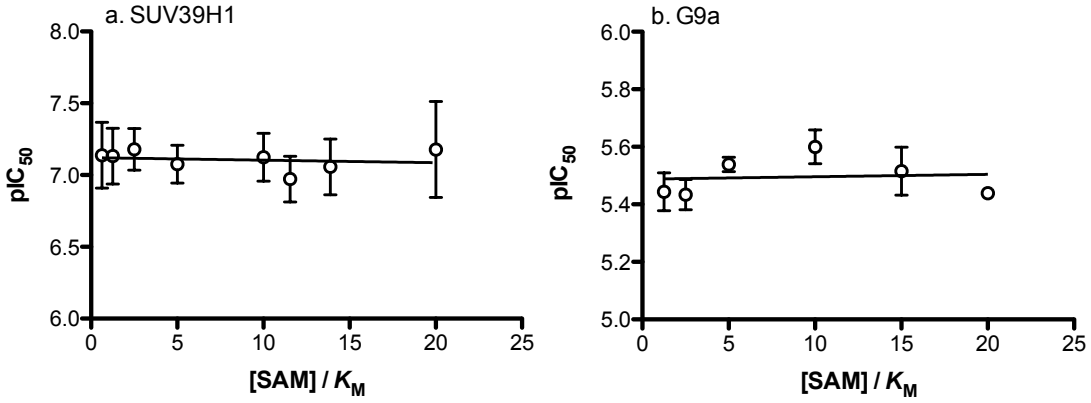


Figure 42 Influence of SAM concentration on pIC_{50} values against a) SUV39H1 and b) G9a. . Adapted from Cherblanc *et al.*[129,131]
a: The pIC_{50} values were derived from 4-point curves at chosen concentrations around the IC_{50} region (0.2 – 0.05 µM), SAM concentration ranged from 800 µM to 6 µM, the data represent the averaged pIC_{50} values of the n=6 independent experiments run as duplicates, the error bars represent SEM.
b: The pIC_{50} values were derived from 6-point curves (32 – 0.03 µM) , SAM concentration ranged from 160 µM to 10 µM, the data represent the averaged pIC_{50} values of the n=3 independent experiments run as duplicates, the error bar represents SEM.
For a) and b), the line corresponds to the linear regression of the data points.

Additionally, pIC_{50} values for the ETP compound **6** were also derived against G9a at high and low SAM concentrations (Figure 43) and as for chaetocin, there was no significant difference between the two values (difference of the means: 0.07 µM), as assessed by t-test.

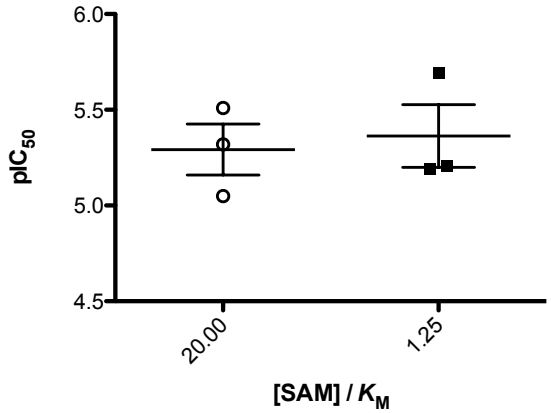


Figure 43 Influence of SAM concentration on ETP compound **6** pIC_{50} values against G9a
The pIC_{50} values were derived from 6-point curves (100 – 0.1 µM) , SAM concentration of 160 µM and 10 µM, the horizontal bars represent the averaged pIC_{50} values of the n=3 independent experiments run as duplicates, the error bar represents the SEM. Each data point corresponds to the averaged pIC_{50} value derived from one experiment run as duplicates.

It could be argued that the time-dependent and potentially covalent nature of the inhibition by chaetocin may mask the competition with the cofactor SAM. It

should be noted however that this experiment was run in the presence of DTT (1 mM in the assay buffer), conditions under which G9a inhibition by chaetocin is fully reversible (See Figure 39). Also, a constant pre-incubation time of 30 min was applied in each case to avoid time-dependent potency changes. Indeed, as shown in Figure 37 and Table 5, an important (and statistically significant) improvement in IC_{50} against G9a was observed when one compares 0 or 5 to 30 min pre-incubation, but no significant change between 15 and 30 min. It can therefore be considered that at 30 min steady-state phase has been reached. Also, at 30 min reaction time, it was shown that G9a has regained its full activity in the presence of DTT (Figure 39). Therefore even if the binding and dissociation were slow, there should have been sufficient time for SAM to displace chaetocin. Consequently, the absence of effect of SAM on chaetocin's inhibition observed here is strongly supportive that the compound is not competitive with the cofactor.

Denaturation test

The absence of competition with SAM was yet another indication that chaetocin did not display a conventional inhibitor behaviour and was most likely inhibiting the target HKMTs by nonspecific effects. A common mechanism of nonspecific inhibition is protein denaturation.[140] ETP natural products have previously been reported to alter the folding of proteins by NMR spectroscopy and circular dichroism.[142,143] We therefore decided to test biochemically whether chaetocin acts as denaturant by assessing its inhibitory potency in the presence of another denaturant, urea. This test relies on the idea that if a compound acts as denaturant, its potency should increase in the presence of another one, here, urea.[140] Urea was titrated to determine the optimal concentration at which G9a still displays acceptable enzymatic activity (Figure 44) and a concentration of 0.7 M was chosen for the denaturation test.

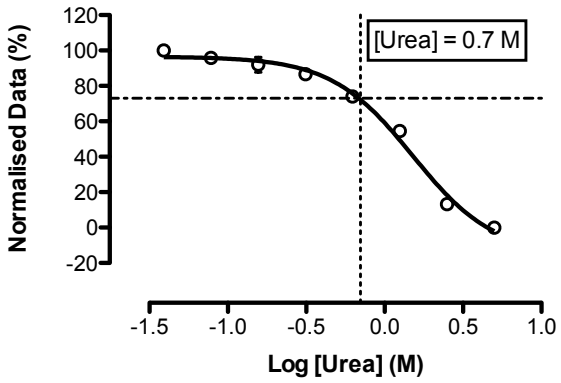


Figure 44 Urea titration against G9a. Adapted from Cherblanc *et al.*[129]
Urea was used in the concentration range: 0.29 – 5 M, the assay was run as single points.

The inhibitory potency of chaetocin was found to be (statistically) significantly increased in the presence of urea, while this was not observed for SAH (Figure 45, Table 7). This indicates that chaetocin inhibits G9a at least partially by nonspecific denaturation. It should be noted that these data were generated in presence of DTT as the signal obtained in the absence of DTT and in the presence of urea was highly variable, most likely due to poor stability of the enzyme under these conditions.

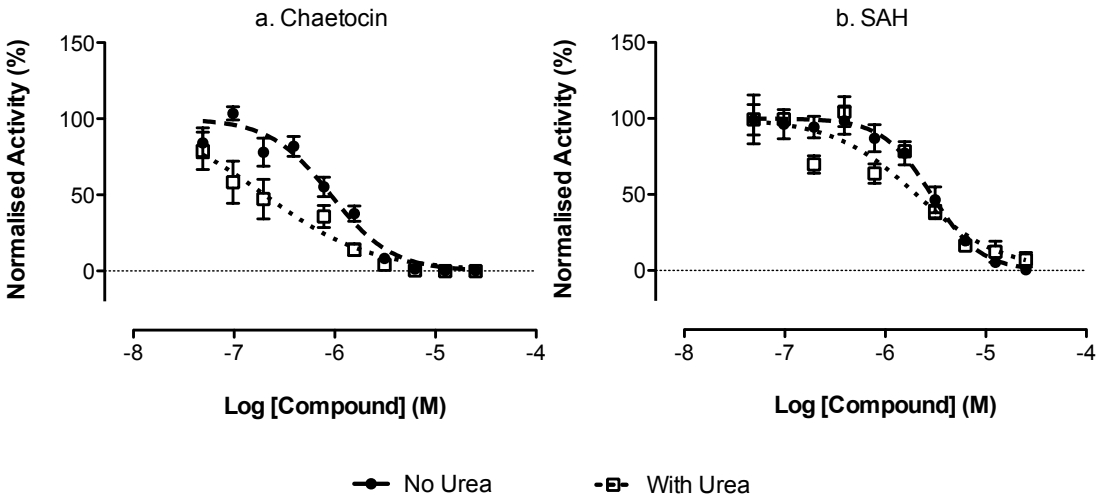


Figure 45 Influence of urea on chaetocin (a) or SAH (b) inhibition of G9a. Adapted from Cherblanc *et al.*[129]
The assay was run in the presence of urea (0.7 M) and DTT (1 mM). The data shown are collated data of n=3 independent experiments, run as duplicates, and error bars represent SEM.

Table 7 pIC_{50} values against G9a with and without urea. Adapted from Cherblanc *et al.*[129]

Compound	pIC_{50} With Urea	pIC_{50} Without Urea	P value	Summary
Chaetocin	6.248 ± 0.06	5.970 ± 0.04	0.0061	**
SAH	5.721 ± 0.16	5.481 ± 0.03	0.2920	ns

Mean pIC_{50} values are reported as average of the pIC_{50} derived for n=3 independent experiments, run in duplicate (±SEM). P values were determined by paired t-test.**: P value < 0.01, ns: not statistically significant

Mass spectrometry studies

To further validate the results obtained in the reversibility studies, we decided to obtain mass spectrometry data to verify whether chaetocin was forming a covalent adduct with G9a. A denaturing mass spectrometry experiment was first run by Dr. Marina Demetriades in Prof. Christopher Schofield lab (University of Oxford). To assess the influence of the reducing agent (as observed in the reversibility test), the mass spectrometry experiment was run in presence and absence of DTT (Figure 46). G9a and chaetocin were mixed and pre-incubated for 30 min (bottom panels) or not pre-incubated (top panels) and the mixture was denatured by treatment with formic acid just prior to ESI-MS analysis.

In the absence of reducing agent, an adduct of mass closely matching G9a–chaetocin was detected, with and without the 30 min pre-incubation time (right panels, Figure 46). Importantly, in the presence of DTT, such adduct was not observed, (left panels, Figure 46) which strongly suggests that chaetocin binds to G9a via mixed disulfide linkage, in agreement with the reversibility studies.

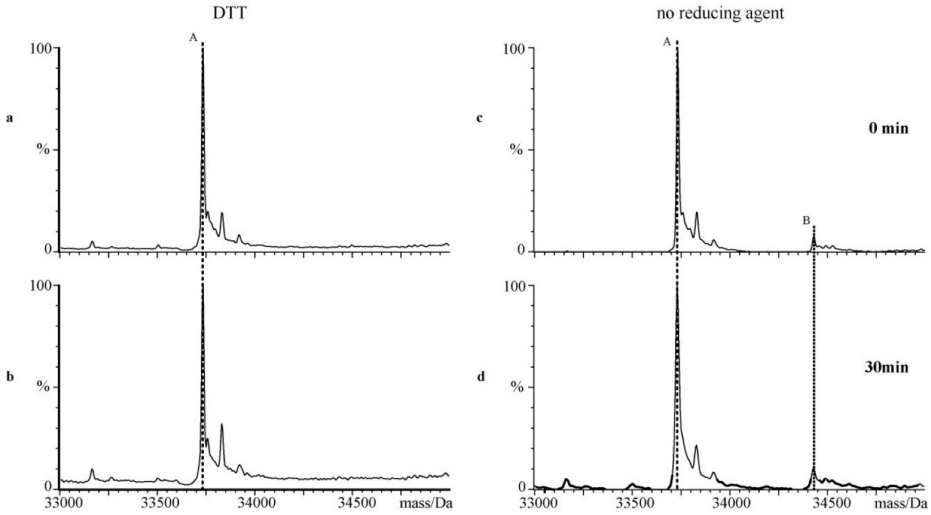


Figure 46 Binding of chaetocin to G9a with and without DTT under denaturing conditions
Deconvoluted MS spectra of G9a (1 eq.) with chaetocin (2 eq.) and with a-b) DTT (10 eq.), c-d) no reducing agent under denaturing conditions at 60 V cone voltage without incubation (top panels) and after 30 min incubation (bottom panels). Peaks correspond A: G9a, B: G9a–Chaetocin

Further validation of these data was obtained by Dr. Peter DiMaggio and Aaron Borg (Department of Chemical Engineering, Imperial College London). An elegant strategy was designed to identify the residues involved in disulfide bonding with chaetocin. The assay principle is depicted in Figure 47. G9a was incubated with chaetocin for 30 min and was then denatured by treatment with guanidine hydrochloride, followed by labelling of the cysteine residues not involved in bonding with chaetocin with *N*-ethylmaleimide. Treatment with DTT lead to the reduction of the disulfide bridges (mixed (between G9a and chaetocin) or native) and the released cysteine residues were labelled with iodoacetamide. The sample was then digested and the peptides analysed by LC-MS/MS and compared to untreated control samples.

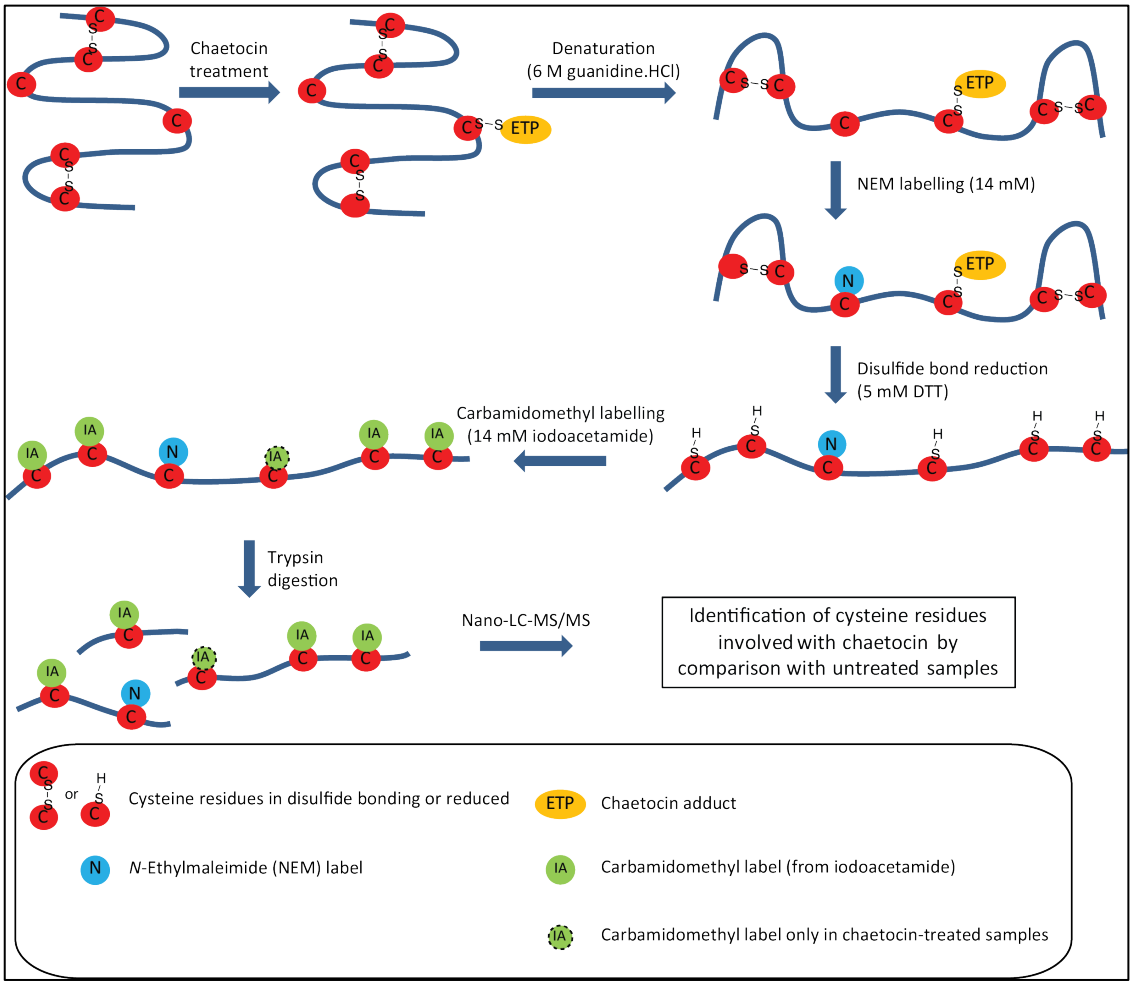


Figure 47 Principle of the mass spectrometry experiment designed to identify the cysteine residues involved in disulfide linkage with chaetocin

Out of the 12 quantitated cysteine residues in G9A (913-1193), five exhibited a greater than 2-fold change in carbamidomethyl (from iodoacetamide treatment) labelling upon chaetocin treatment (Figure 48): Cys994 (4.8 fold increase), Cys1017, Cys1021 and Cys1023 (3.1 fold increase) and Cys1115 (2.1 fold increase).

The cysteine residues involved in disulfide linkage with chaetocin have been highlighted in red in Figure 49. It is apparent that only one of the cysteine residues involved in bonding with chaetocin is actually present in the active site of G9a (Cys1115) and it is the least significant change (2.1-fold). Notably, the other cysteine residues are not in the active site of G9a but in the pre-SET domain and are believed to coordinate structural zinc ions.

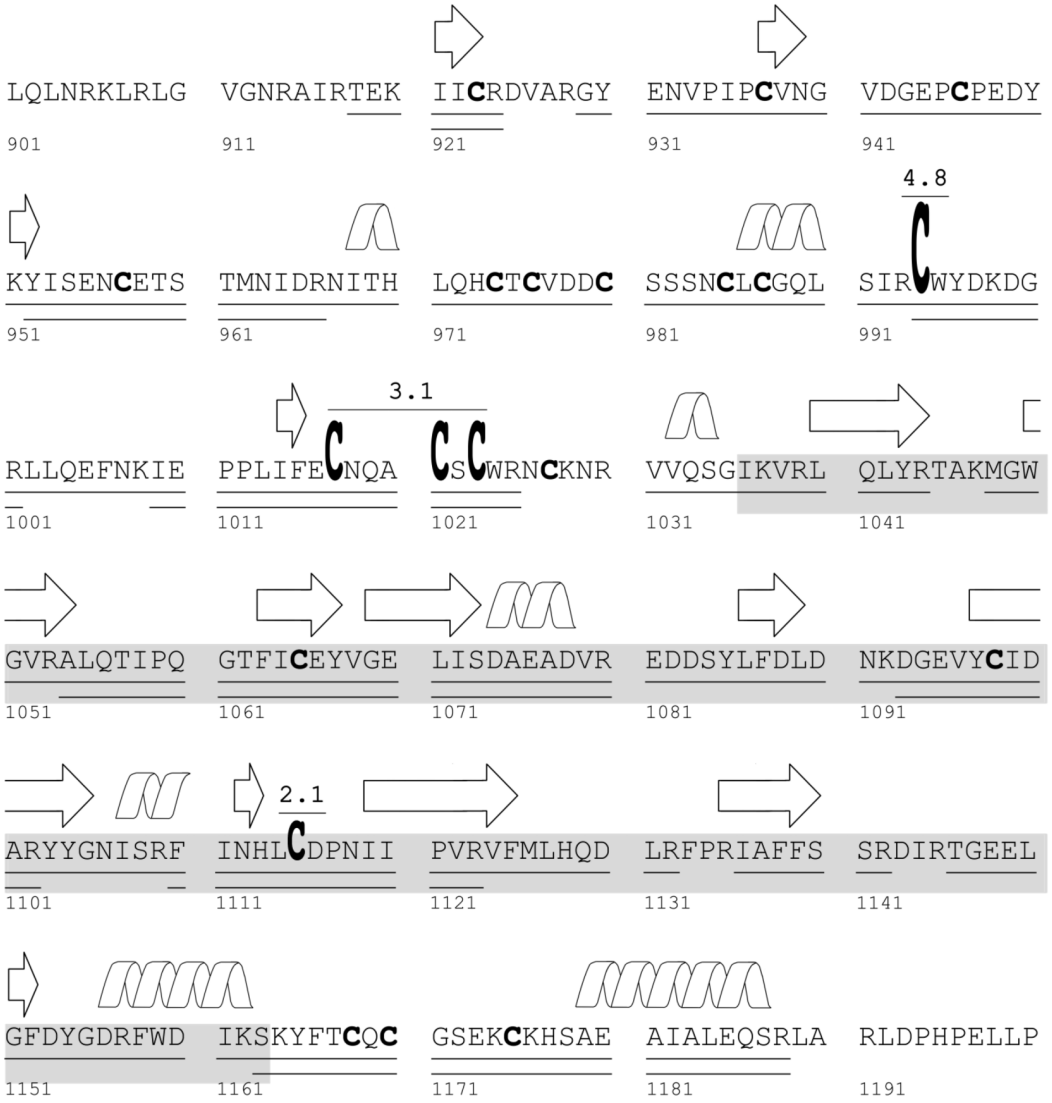


Figure 48 G9A (913-1193) sequence highlighting cysteine residues involved in chaetocin binding. Taken from Cherblanc *et al.*[129]
All 20 cysteine residues are highlighted in boldface, and those cysteine residues involved in the formation of disulfide bonds with chaetocin are scaled according to their fold increase in carbamidomethyl labelling upon chaetocin treatment (the numbers above these residues represent the fold change observed relative to control). Singly and doubly underlined sequences indicate identified and quantitated peptides from the LC-MS/MS data, respectively. Shaded sequences highlight the SET domain, which is flanked on the left by the pre-SET domain and on the right by the post-SET domain. The arrows and twists represent β sheet and α helical secondary structural elements, respectively, as identified by RSCB PDB accession number 2O8J.

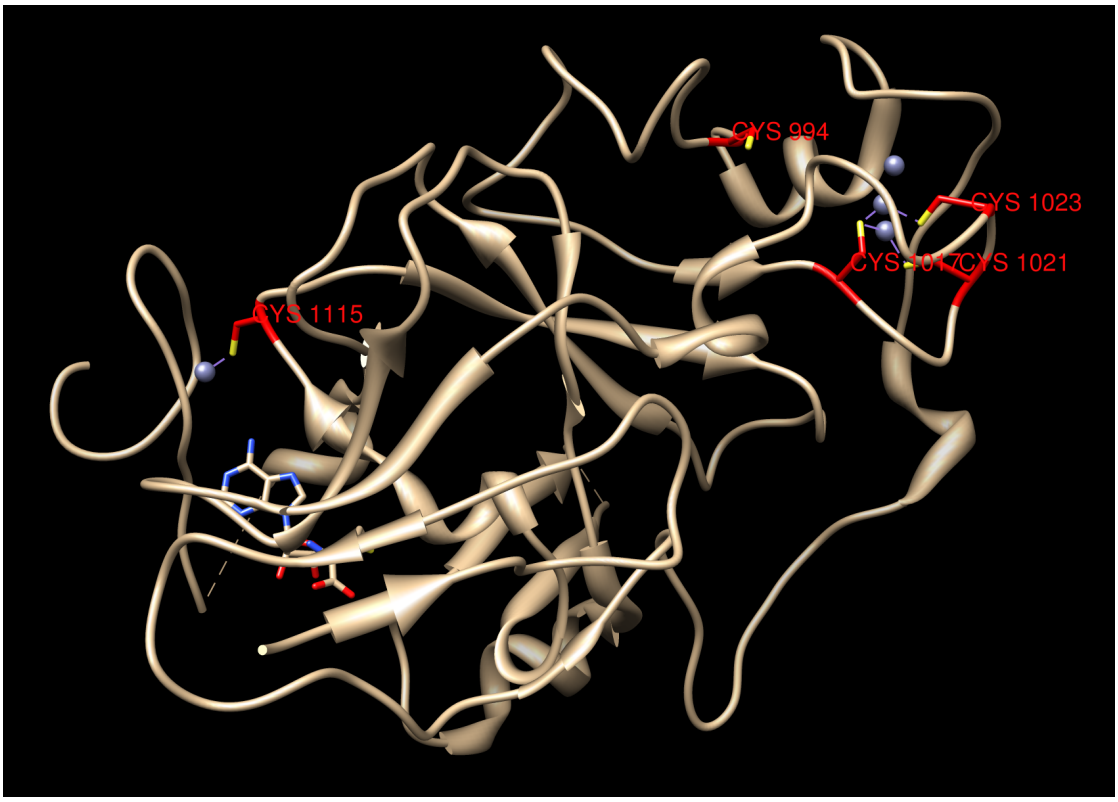


Figure 49 Cysteine residues of G9a involved in bonding with chaetocin.
Generated from the pdb structure 2O8J (G9a co-crystallised with SAH); the 5 cysteine residues identified by MS experiment are labelled and coloured in red. The grey balls represent zinc ions.

This data demonstrated that several cysteine residues reacted with chaetocin to form mixed disulfides. This clearly illustrated that chaetocin does not bind to a defined site of G9a and further confirmed that the interaction between the compound and the protein is nonspecific.

Conclusions

Importance of the disulfide moiety for HKMT inhibition

Measuring the inhibitory activity of chaetocin and its structurally important analogues clearly shed light onto the requirement of the disulfide bridge of ETP compounds for their HKMT inhibitory activity, in definite opposition with the hypothesis postulated by Greiner *et al.*.[74] In particular, the structurally simple ETP-containing molecule **6** retained substantial inhibitory activity, while the unreactive thioether derivative **4** displayed no activity whatsoever. As aforementioned, Sodeoka and co-workers have performed the first total synthesis of chaetocin (**1**), its unnatural enantiomer (*ent*-**1**) as well as sulfur-free analogues (**5** and *ent*-**5**).[84,104] Their synthetic chaetocin sample had similar inhibitory potency against G9a to the one we measured (IC_{50} of 2.4 and 1.7 μM for chaetocin and *ent*-chaetocin respectively). More importantly, the sulfur-free analogues, **5** and *ent*-**5**, were found to be totally inactive. These data are in very good agreement with our results on the requirement of the disulfide bridges for HKMT inhibition.

The attempts described here toward the development of a SAR around chaetocin structure led us to uncover a structure–*reactivity* relationship as proposed by Rishton,[144] whereby the *activity* observed is actually only a reflection of the *reactivity* of the compound evaluated toward the target. Here, the unreactive thioether (or the sulfur-free analogue synthesised by Sodeoka and co-workers) does not display any inhibitory activity while chaetocin, bearing the reactive disulfide linkage, does. This absence of SAR in a series of analogues from an original hit identified by HTS usually reflects that this series is not fit-for-purpose in terms of further development from a drug discovery perspective.[145] Consequently, chaetocin is not a suitable starting point for the discovery of a specific HKMT inhibitor.

Mechanism of action

Having established that the inhibitory activity of chaetocin was entirely dependent on the reactive ETP disulfide, we next sought to identify the molecular mechanisms responsible for the inhibitory effect of this class of molecules.

Firstly, the biological activities of ETP compounds have been reported to rely on their ability to redox cycle and generate reactive oxygen species (ROS).[85] Since

the generation of H_2O_2 under the assay conditions was not observed, we considered that this mechanism is unlikely to account for the HKMT inhibitory activity of ETP compounds in our cell free system.

Secondly, another mechanism proposed for inhibition by ETP compounds is cross linking with cysteine residues, potentially followed by formation of internal disulfide bridge.[94,146] The time-dependent nature of the HKMT inhibition gave the first hint of such a mechanism. In particular, the time-dependency was accentuated in absence of reducing agent (here DTT). The reversibility test corroborated nicely this observation. Indeed, the inhibition of G9a by ETP-containing compounds (chaetocin (**1**) or ETP compound **6**) was clearly irreversible in absence of reducing agent while fully reversible in presence of DTT. These biochemical observations suggested that cross-linking between the ETP functionality and the HKMT *via* a disulfide linkage (which can be reduced in presence of DTT) was central to the inhibition mechanism.

Thirdly, the mass spectrometric data, obtained by an orthogonal approach, further validated the biochemical conclusions. We first identified an adduct formation between chaetocin and G9a in the absence of reducing agent, while such adduct could not be observed in the presence of DTT, once again pointing at a disulfide linkage. Furthermore, by in depth LC-MS/MS analysis, we identified the cysteine residues involved in such bonding. Importantly, this data shows that more than one cysteine residues are involved in disulfide bond formation upon chaetocin treatment. It is therefore apparent that chaetocin does not exhibit specificity towards a single cysteine residue, which would be expected for a targeted covalent inhibitor.[147] Additionally, our SAM competition studies are in opposition with the conclusions of Greiner *et al.* who suggested that chaetocin binds to the co-factor binding cleft.[74] Our biochemical and mass spectrometric data refute such competitive binding and, in conjunction with literature data, strongly suggest that chaetocin does not bind to a defined pocket. Indeed, the fact that a structurally simple ETP compound **6** is able to significantly inhibit HKMT activity, not only supports the critical role of the ETP core for inhibition, but also suggests much of the rest of the complex structure of chaetocin to be superfluous. In addition to this, Sodeoka and co-workers showed that the two enantiomers of chaetocin they prepared inhibited the target equally effectively,[104] which clearly exposes the fact that the three dimensional structure of chaetocin is not involved in the binding to the HKMT target. On top of this, a number of structurally unrelated ETP compounds (including gliotoxin, chetomin and

11,11'-dideoxyverticillin, Figure 8) were shown to also inhibit G9a.[148,149] In combination, these results clearly show that the framework and three dimensional structure of chaetocin (or of any ETP compounds) are not essential for binding or inhibitory activity, the reactive disulfide bridge is the only required moiety. This prevents chaetocin from being classed as a targeted covalent inhibitor. Indeed, for a covalent inhibitor to be viable from a drug development perspective, the reactivity of the compound needs to be accompanied by specific shape complementarity with the biological target:[147] a variety of interactions between target and inhibitor (like with the more conventional reversible inhibitor) needs to precede the covalent attachment. In the case of ETP-containing compounds, this does not appear to be the case: our mass spectrometry experiment revealed that five different cysteine residues reacted with chaetocin. Furthermore, in opposition with other natural products containing electrophilic moieties,[150] the disulfide bridge of the ETP compound seems too reactive to be tuned towards HKMT selectivity: our data and that of other groups clearly show that upon modification of the disulfide bridge, all biological activity is abolished, which also points out the absence of binding interactions.

Finally, protein denaturation by ETPs has been observed previously by a number of techniques including circular dichroism and NMR studies.[142,143] The increased inhibitory activity of chaetocin observed in the presence of urea is also supportive that broad nonspecific protein denaturation is at least in part responsible for the observed inhibitory activity of ETP compounds. The adduct formation process, as identified by mass spectrometry, is likely to be responsible for the protein denaturation over time.

Observed HKMT selectivity

Our MS data showed that cysteine 994, 1017, 1021 and 1023 exhibit the largest reactivity towards chaetocin. These residues are located in the pre-SET domain of G9a and have been shown to be implicated in the binding of zinc atoms, which is thought to have a structural function in stabilising the SET domain.[151]

Interestingly, we observed some adduct formation with cysteine 1115, which is part of a highly conserved region in the SET domain. This region is involved in hydrogen bonding interactions with SAM and adduct formation with this cysteine may have some direct inhibitory effect.[151] While specific interaction with this cysteine residue would yield a promising covalent inhibitor, it has to be pointed out that only a

2.1-fold increase in labelling was observed at this residue, which is not as significant as for the other cysteine residues (3.1 and 4.8-fold). In addition, we did not observe any increased labelling at any of the other SET domain cysteine residues. This, together with the competition studies showing that chaetocin is not competitive with SAM, suggest that bonding with this cysteine is probably not the most significant interaction of chaetocin with G9a.

Based on our MS data, we propose that the sensitivity of a given HKMT to chaetocin's inhibition is related to the presence of a cysteine rich pre-SET and possibly post-SET domain (although we observe no binding of chaetocin to the post-SET domain in G9a), through mixed disulfide formation with the ETP functionality. Notably, the three conserved cysteine residues in the post-SET domain within DIM5 have been shown to be essential for HKMT activity.[152] Indeed, while Greiner *et al.* claimed that chaetocin was selective for HKMT in the SUV39 family, namely dSU(VAR)3-9, SUV39H1, G9a, DIM5,[74] our data suggests instead that the apparent selectivity is solely related to the presence of a pre-SET (and potentially post-SET) domain containing conserved cysteine clusters coordinating zinc ions.[63] Importantly, SET7/9 and PR-SET7, which do no have any cysteine clusters, were not inhibited by chaetocin. The *Drosophila* enhancer-of-zest (*d*E(z)) complex was not targeted by chaetocin,[74] although it has a pre-SET domain (with ~15 cysteine residues), but no post-SET domain. We also assayed chaetocin against the human orthologue EZH2 (in the assay kit described in chapter IV) and discovered that the natural product could also inhibit this enzyme (80% inhibition at 20 μM). Both *d*E(z) and EZH2 are active methyltransferases only as part of the Polycomb complex and as such can be causing difficulties in biochemical assays (*e.g.* purification, complex reconstitution).[39,153] We believe that Greiner's result on *d*E(z) should be interpreted with caution in light of the discrepancy with our results against EZH2. The reliability on the presence of a cysteine rich domain for ETP inhibition was again recently supported as gliotoxin and other ETP natural products have been reported to have a similar effect, inhibiting G9a and SU(VAR)3-9 but not SET7/9.[148,149] Therefore instead of a protein binding site conferring an authentic HKMT selectivity to chaetocin, we believe any apparent *in vitro* selectivity is related to a given protein's sensitivity to the ETP reactive disulfide, most likely concurrent with the presence of accessible cysteine residues.

Furthermore, ETP compounds have been shown to target numerous structurally, functionally, and evolutionary unrelated enzymes.[84,85] This could have appeared as an early hint of non-specific effects. For example, chaetocin has already been shown to inhibit thioredoxin reductase by acting as a competitive substrate,[95] and also to disrupt the binding of p300 to hypoxia-inducible factor-1α (HIF-1α).[93] In the case of gliotoxin, the most studied member of the ETP family, the list of biological effects and identified targets is extraordinarily long. Waring *et al*. previously reviewed this area 16 years ago and various new activities have been identified since then, including inhibition of G9a and inactivation of muscle creatine kinase.[92,94,149] Notably, most of the studies looking at various ETP compounds show that they share the same targets, with very little (if any) selectivity, as exemplified by the disruption of the binding of p300 to HIF-1α which is efficiently achieved by chetomin, chaetocin, gliotoxin and a wide range of small synthetic ETPs.[93,143] This, in addition to the data aforementioned, stresses the fact that it is highly unlikely that the ETP-containing compounds bind in a specific way to their biological target.

Toxicity and epigenetic cell-based assay of ETP-containing compounds

With this in mind, cellular evaluation of the effects of ETP compound appears highly challenging. Nevertheless, the effect of chaetocin on histone methylation in cells has been observed in several studies.[74,154-156] Although one plausible explanation for this result is indeed direct inhibitory effects on HKMTs, an alternative possibility is that the observations are a result of the toxicity of ETP-containing molecules,[85] together with their important off-target effects. The complex effect of compound treatment on histone methylation may instead be an indirect readout of on-going cell death. In particular, the release of histones from nucleosomes during apoptosis is well established,[157] and a variety of changes in histone modifications have been observed during the apoptotic process.[158] These general effects on histones during apoptosis make it difficult to separate histone changes due to HKMT inhibition from changes due to cell death effects. In addition, the concentrations used in cell-based assay of ETP-containing compounds are much lower (50 – 200 nM) due to cytotoxicity than the IC_{50} values derived from biochemical assays (0.5 – 3 μM) which is highly unusual and strongly suggests that the effects observed in cells are inconsistent with inhibition

of the target enzyme and are probably mediated by other mechanisms, most likely unspecific apoptosis-related effects.[139]

Final remarks on chaetocin and ETP-containing compounds

Structurally diverse inhibitors are critically needed to improve medicinally relevant chemical space, and HTS remains one of the most efficient ways to identify a new molecular starting point. Nevertheless, the drug discovery community and particularly academia should be more thorough in the validation of hits against new targets. Mechanism of action as well as basic SAR studies should be routinely performed to assess the authenticity of a new inhibitor to ensure that only fit-for-purpose molecules are used in further biological testing. Indeed, the use of chaetocin as a specific and selective H3K9-methylase inhibitor in a cell-based assay is highly arguable in light of all the other enzymes it targets and its known reactivity. We strongly believe that chaetocin has no potential to be developed into a selective HKMT inhibitor.

However, it is possible than ETP compounds find applicability in a non target-selective approach. In particular, chaetocin was shown to induce apoptosis selectively in myeloma and leukaemia cell lines *via* imposition of oxidative stress, to which they are more sensitive than normal cell lines.[132,133] Chaetocin was later shown to produce oxidative damage in myeloma cells by inhibiting the antioxidant enzyme thioredoxin reductase,[95] by acting as a competitive substrate. Further studies will hopefully reveal if chaetocin or a related ETP-containing compound has the potential to be developed into an anti-myeloma or leukaemia therapeutic.

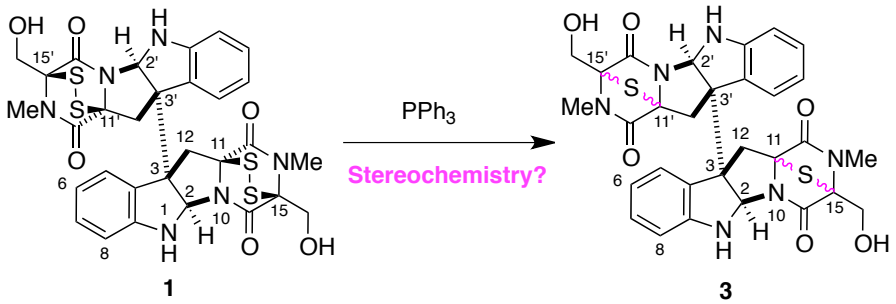
Chapter III

Stereochemistry of ETP desulfurisation

Introduction

It is crucial to know the stereochemistry of compounds used for biological applications, as enantiomers can sometimes display strongly different activities. Since natural biomolecules (*e.g.* enzymes) are chiral, the two enantiomers of a molecule will interact differently with a given target.[159] In most cases, only one of the enantiomers displays the desired biological effect while the opposite enantiomer may be inactive or even responsible for adverse effects, as illustrated by the unfortunate example of thalidomide.[160]

As outlined in Chapter II, analogues of the natural product chaetocin (**1**) have been prepared and biologically evaluated to characterise further the inhibitory activity of this class of compounds against HKMTs. Amongst these, monosulfide **3**, obtained from desulfurisation of chaetocin (**1**), was a key analogue as it retains the bicyclic structure of the ETP core. Should the product be obtained with retention of configuration, monosulfide **3** would have a similar three-dimensional structure to that of the parent natural product. The stereochemistry of monosulfide **3** was however not straightforward to assign (Scheme 19).

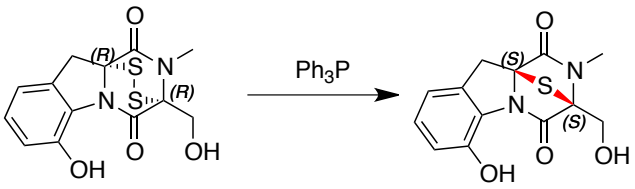


Scheme 19

Since our various attempts at growing a crystal of analogue **3** suitable for X-ray diffraction were unsuccessful, we decided to carefully review stereochemical assignments for similar desulfurisation reactions in the literature.

The desulfurisation of organic linear disulfides into the corresponding sulfides by treatment with phosphines has been reported as early as 1935,[161] and a number of reagents have been proposed to mediate this transformation.[162,163] Safe and Taylor reported the first example of desulfurisation of ETP-containing natural products in 1969.[110] In their work, dehydrogliotoxin (**21**) was converted to its monosulfide

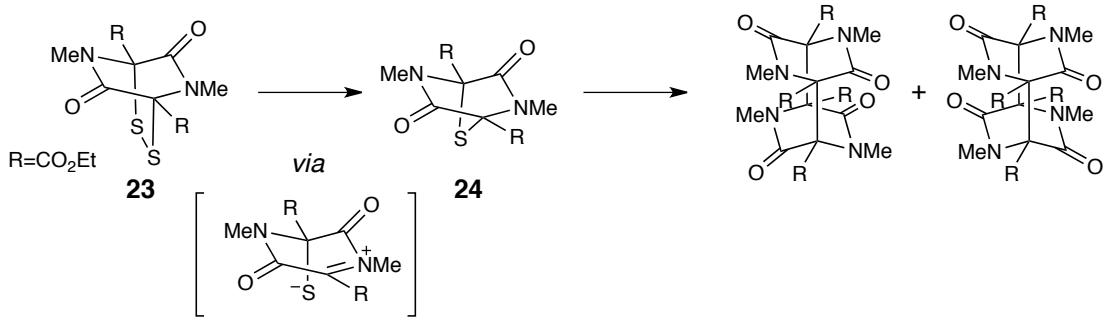
counterpart (**22**) by treatment with triphenylphosphine (Scheme 20). Comparison of the electronic circular dichroism (ECD) spectra of the monosulfide derivative **22** and that of the parental natural product **21** suggested opposite stereochemistry.[110,164] Indeed, the sign of the Cotton effects of the desulfurised product's ECD curve was opposite to those of the natural product.[164] The reaction was thus reported to lead to inversion of stereochemistry at the bridgehead carbon atoms.



Scheme 20

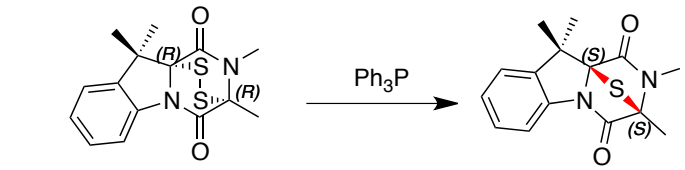
However, Sammes argued that the inversion of configuration was "mechanistically unfeasible", and that the curves were not comparable,[165] based on detailed circular dichroism studies performed on the natural product gliotoxin (Figure 8).[166] In an ECD spectrum, each Cotton effect corresponds to the differential absorption of left- and right-polarised light by a given electronic transition of a chromophore. A chromophore is defined as a molecular moiety, responsible for one or more electronic transitions, associated with absorption bands in the UV-visible range (typically a functional group or a combination of several groups with a more or less extended π electron system).[167] In the case of gliotoxin, it has been showed that the chromophoric disulfide moiety interferes with that of the diene, thus leading to a Cotton effect of the opposite sign to those found for similar diene system.[166] Importantly, the ECD curve of a sulfur-free analogue of gliotoxin did exhibit a Cotton effect of the expected sign for this type of diene, thus corroborating the proposed chromophore interaction theory. Sammes postulated that a similar scenario could be happening between dehydrogliotoxin and its monosulfide analogue: as the disulfide moiety is no longer present to interact with the other chromophores in the monosulfide derivative, it can be anticipated that an inversion of the sign of some Cotton effects could be observed without being related to a change in the product's stereochemistry.[165] Hence, the assignment of the stereochemistry of monosulfide **22** may not be highly reliable.

Sato and Hino, in subsequent studies on the reaction of a synthetic ETP derivative **23** with triphenylphosphine, observed monosulfide **24** *en route* to dimeric species for which they postulated an S_N1-type mechanism (Scheme 21). This implied that the monosulfide should be obtained with retention of configuration at the bridgehead carbon atoms.[168]



Scheme 21

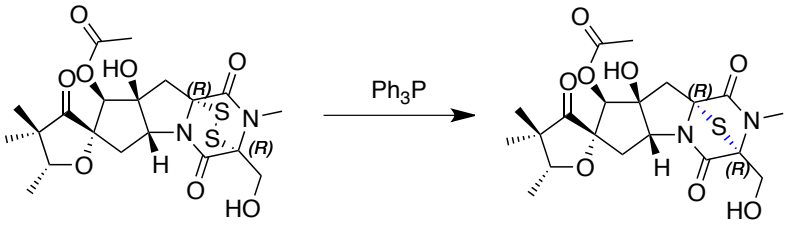
The problem of the stereochemical course of this reaction appeared to have found its definite solution in 1979, when Ottenheijm and co-workers published a detailed study of the triphenylphosphine-mediated desulfurisation of a synthetic analogue of natural product gliotoxin **25** (Scheme 22).[169]



Scheme 22

The (*R,R*) ETP **25** was desulfurised to give the (*S,S*) monosulfide **26**, with inversion at both stereocentres. The stereochemistry of the product was unambiguously assigned by X-ray diffraction using the anomalous dispersion of the sulfur atom.[170] ECD curves of the desulfurised product and parental ETP were also found to have opposite signs of most Cotton effects.[169] The authors therefore drew the conclusion that comparison of the ECD curves was actually a suitable way to determine relative stereochemistry.

Nevertheless, a few years later, Barbier and co-workers reported their study on the stereochemical course of the desulfurisation of sirodesmin PL (**27**) into its monosulfide derivative **28** (Scheme 23). Based on chemical derivatisation evidence together with X-ray analysis of a diacetyl derivative, the authors showed that the reaction proceeded with retention of configuration.[171]

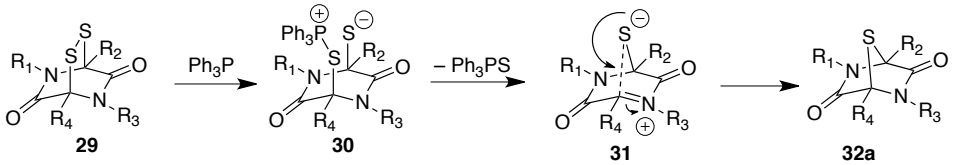


Scheme 23

Importantly, they also reported that the ECD curves of the natural product **27** and its desulfurised derivative **28** exhibited Cotton effects of opposite signs although the reaction was found to proceed with retention of stereochemistry. They therefore warned against the correlative use of ECD data as an approach to determine the absolute configuration of the bridgehead carbon atoms in molecules with similar framework.

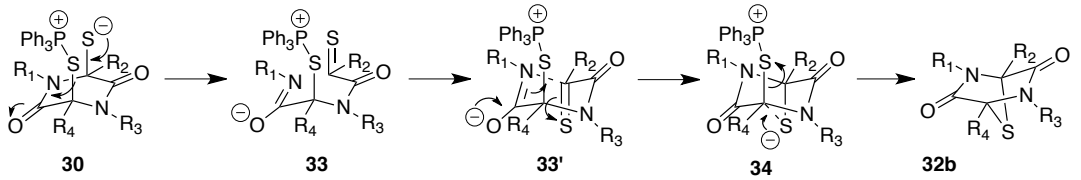
Three different mechanisms have been proposed to account for the observed stereochemistries.[168,169,171]

The first mechanism was proposed by Sato and Hino and is depicted in Scheme 24: nucleophilic attack of triphenylphosphine lead to the cleavage of the disulfide bond of ETP **29** to give phosphonium salt **30**; the lone pair of the nitrogen of the diketopiperazine was then suggested to displace triphenylphosphine sulfide to give iminium **31**, which then undergoes ring closing by nucleophilic attack of the thiolate onto the iminium to give monosulfide **32a** with retention of stereochemistry.[168]



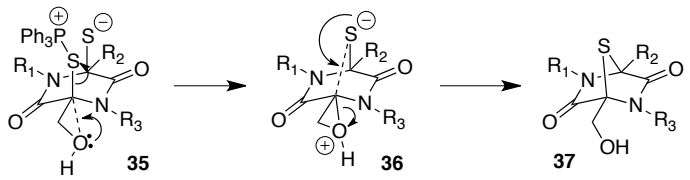
Scheme 24

To account for the observed inversion of stereochemistry, Ottenheijm and co-workers suggested instead that phosphonium salt **30** is stable to elimination (Scheme 25), and that the thiolate centre can epimerise *via* ring opening of the diketopiperazine ring, conformational rearrangement and ring closing (**30→33→33'→34**), followed by S_N2-type displacement of the triphenylphosphine sulfide leaving group by the thiolate to give monosulfide **32b** with inversion of stereochemistry.



Scheme 25

Finally, Barbier and co-workers proposed a mechanism leading to retention of stereochemistry but invoking the participation of the substituent of the diketopiperazine ring in order to explain the discrepancies observed with Ottenheijm's results.[171] In particular, when the substituent is a hydroxymethyl like in sirodesmin PL (**27**, Scheme 26), it is plausible that the triphenylphosphine sulfide leaving group is displaced by the hydroxyl group *via* S_N2-type displacement (**35→36**), followed by an additional S_N2-type displacement of the oxirane by the thiolate anion (**36→37**). This double inversion mechanism would eventually result in net retention of stereochemistry.



Scheme 26

In light of these conflicting results, it was not possible to assign the stereochemistry of chaetocin monosulfide **3** by analogy with known examples. As chaetocin presents a similar pendant alcohol as sirodesmin PL (**27**), it was expected that this natural product would most likely be desulfurised with retention of

configuration but we felt additional proof was needed. It was thus decided to record a variety of chiroptical spectra and to compare the experimentally obtained data with their simulated counterparts for the possible stereoisomers.

Chiroptical spectroscopy refers to the optical techniques using refraction, absorption or emission of anisotropic radiation that can be employed for identifying the relative and absolute stereochemistry of a substance. In this work, three different techniques have been used: optical rotation (at a fixed wavelength, here at 589 nm (D line of sodium), referred to as $[\alpha]_D$), electronic circular dichroism (ECD) and vibrational circular dichroism (VCD). Optical rotation and ECD have been used extensively for the determination of stereochemistry primarily *via* correlative approach, whereby the acquired spectra are compared to those of a related molecule of known stereochemistry. Examples of such correlative approaches based on ECD spectroscopy were outlined above and will be further discussed in this chapter. More recently, comparing experimental chiroptical spectra to *ab initio* simulations has become an increasingly common approach for the determination of absolute configuration and a number of examples have been recently reported.[167,172-174] The increased computer power allowed for more advanced quantum chemical calculations and enhanced the reliability of the predicted spectra.[172] Although the use of VCD in stereochemical assignment for natural products is less common, it is emerging as a highly useful technique.[175] VCD is an extension of ECD to the infrared and near-infrared regions of the spectrum where vibrational transitions occur within the ground electronic state of the molecule.[172] This technique does not require the presence of a chromophore in the molecule, which represents a major advantage over ECD. In addition, VCD has the important benefit of addressing all 3N-6 vibrational modes (*i.e.* stretching, bending, scissoring, rocking and twisting, N=number of atoms: 3 degrees of freedom for each atom minus the displacements and rotations of the whole molecule), thereby offering more information-rich spectra (large number of bands), while ECD is based on the relatively few electronic transitions.[176]

The simulations of optical rotation, ECD and VCD spectra were carried out by Prof. Rzepa (Chemistry Department, Imperial College London) assisted by an undergraduate student, Ya-Pei Lo, using Gaussian 09.[177] The simulated spectra were visually compared with the experimental data in terms of sign, magnitude and wavelength of the maxima. The VCD measurements were performed by the European

Centre for Chirality and, in addition to visual inspection, a mathematical algorithm was used to assess the adequacy of the manual stereochemical assignments in light of previous successful assignments (implemented in the CompareVOA program[178]).[179] This algorithm allowed us to establish a confidence level for the absolute configuration assignment made by visual interpretation. Briefly, a similarity measure Σ, expressing the degree of agreement between theory and experiment, is computed for both enantiomers and lies between 0 and 1. The enantiomeric similarity index Δ gives the difference between the values of Σ for both enantiomers of a given stereoisomer. Hence, Δ is a measure for the discriminative power of a VCD analysis and therefore a high quality VCD spectrum is characterised by a high absolute value of Δ. These quantities, Σ and Δ, are then compared to a database of previous successful assignments to calculate a confidence level.

Results and Discussion

Chaetocin

As aforementioned, we needed to assign the absolute configuration of chaetocin monosulfide **3** and for this purpose we decided to use three different chiroptical techniques and to compare the experimental data to the simulated ones. Given chaetocin's symmetry, there are three different diastereoisomers one could envisage after desulfurisation: one where both ETP centres (*i.e.* the stereochemistry at C11/C15 and C11'/C15'; see chaetocin numbering Scheme 19) had undergone retention (*S,S S,S*–structure **3a**, Figure 50); one where one ETP centre had undergone retention and one inversion (*R,R S,S*–structure **3b**, Figure 50) or one where both ETP centres had undergone inversion (*R,R R,R*–structure **3c**, Figure 50). Desulfurisation of chaetocin could result in hybrid structure **3b** if two mechanisms, one leading to inversion and one leading to retention of stereochemistry, were in competition.

We first started our investigation with optical rotation measurement and calculations. The measured optical rotation for analogue **3** was +484 (*c* 0.0039, $CHCl_3$, 25 ˚C). The optical rotations for the three possible stereoisomers were calculated at the ωB97XD/6-311G(d,p) level of theory and were +467 for **3a**, +357 for **3b**, and +259 for **3c**. The predicted values showed a significant difference, which gave us confidence for an efficient assignment. In particular, the calculated value for **3a** was very close to the experimental measurement, which suggested the reaction had proceeded with retention of stereochemistry at the bridgehead carbon atoms.

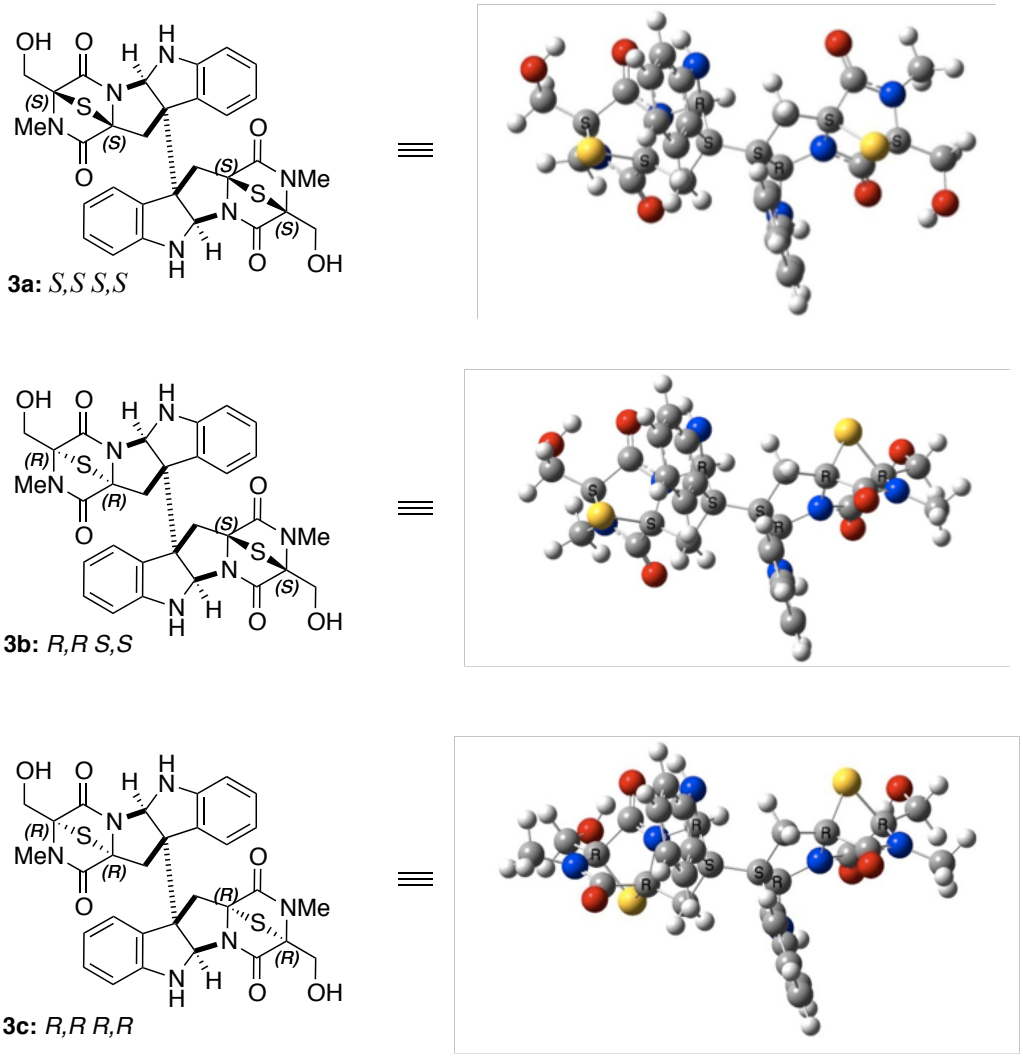


Figure 50 Possible stereoisomers for chaetocin desulfurised analogue calculated using Gaussian 09[177] at the ωB97XD/6-311++G(d,p) level of theory. Adapted from Cherblanc *et al.*[180].

To further confirm the assignment, we undertook the ECD studies of monosulfide **3**. We first wanted to confirm the ability of the computational method to predict ECD spectra for the chaetocin framework, and we therefore simulated the ECD spectrum of the parental molecule **1** at the ωB97XD/6-311++G(d,p) level of theory. The simulated spectrum of chaetocin was in good agreement with the experimental spectrum with a similar pattern of Cotton effects (large positive (235 – 240 nm), small negative (265 – 274 nm) then small positive (285 – 305 nm) (Figure 51). This analysis was in strong agreement with the established stereochemistry of the natural product and provided high confidence on the validity of our approach.

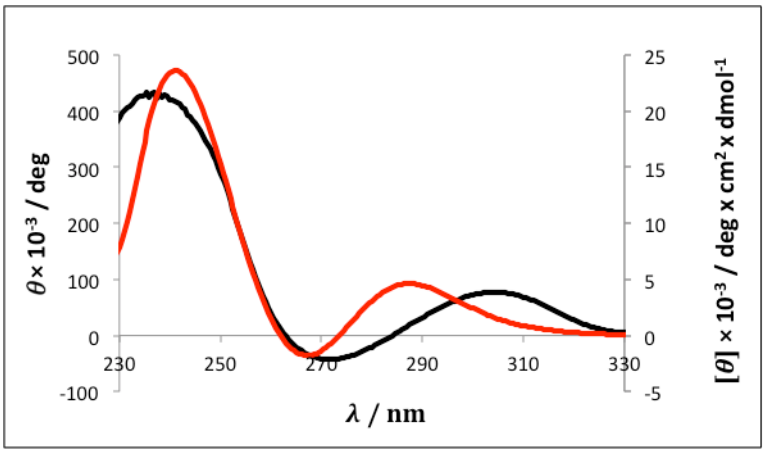


Figure 51 ECD spectra of chaetocin (**1**)
Black line: experimental results for **1** in methanol (left *y* axis); Red line: calculated spectra (right *y* axis). The calculations were performed at **ω**B97XD/6-311++G(d,p) level of theory; the predicted curves were shifted by +20 nm, convoluted line width 0.30 eV.

The ECD spectra for monosulfide **3** were likewise computed at the ωB97XD/6-311++G(d,p) level of theory for the three possible stereoisomers and compared with the experimental ECD curve (Figure 52). The experimental ECD curve of chaetocin monosulfide **3** exhibited one negative (222 nm) then two positive (250, 305 nm) Cotton effects. The ECD curve of **3c** displays a negative Cotton effect around 280 nm, in clear opposition with the experimental data. This diastereoisomer could thus be excluded with high confidence based on this observation. The ECD curve for **3b** had a clear pattern of negative (222 nm), positive (246 nm), negative (272 nm), positive (289 nm) Cotton effects and therefore could be excluded with good confidence due to the presence of an additional negative Cotton effect not observed in the experimental spectrum. Finally, the predicted ECD curve for **3a** had one negative (233 nm) followed by two positive (247, 288 nm) Cotton effects, in good agreement with the experimental data.

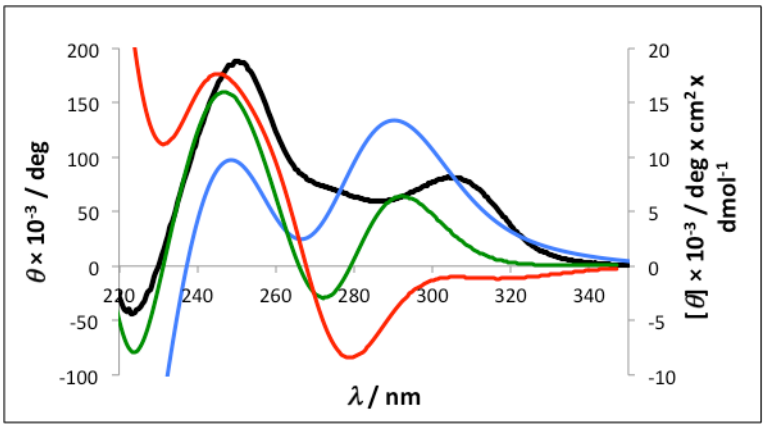


Figure 52 ECD spectra of chaetocin monosulfide **3**. Adapted from Cherblanc *et al*.[180]
Black line: experimental results for **3** in methanol (left *y* axis); Blue: calculated *S,S S,S*–structure **3a**, Green: calculated *R,R S,S*–structure **3b**, Red: calculated *R,R R,R*–structure **3c** (right *y* axis); The calculations were performed at **ω**B97XD/6-311++G(d,p) level of theory; the predicted curves were shifted by +20 nm and convoluted with a line width 0.24 eV.

The ECD analysis appeared consistent with the reaction proceeding with retention of ETP stereochemistry, in agreement with the conclusions drawn from the optical rotation study.

Lastly, a VCD analysis was performed to further confirm our assignment. The experimental and theoretical IR and VCD spectra of chaetocin monosulfide **3** are shown in Figure 53 and Figure 54. Both IR and VCD spectra were recorded at the European Centre for Chirality and our collaborators also performed the detailed analysis described here.

Simulated IR/VCD spectra are commonly scaled with a factor which usually lies between 0.96 and 1.00[181] in a way that the simulated IR spectrum gives acceptable visual agreement with the experimental one. This was done for all three diastereoisomers separately and a decent agreement for the IR spectra is apparent for each one (Figure 53).

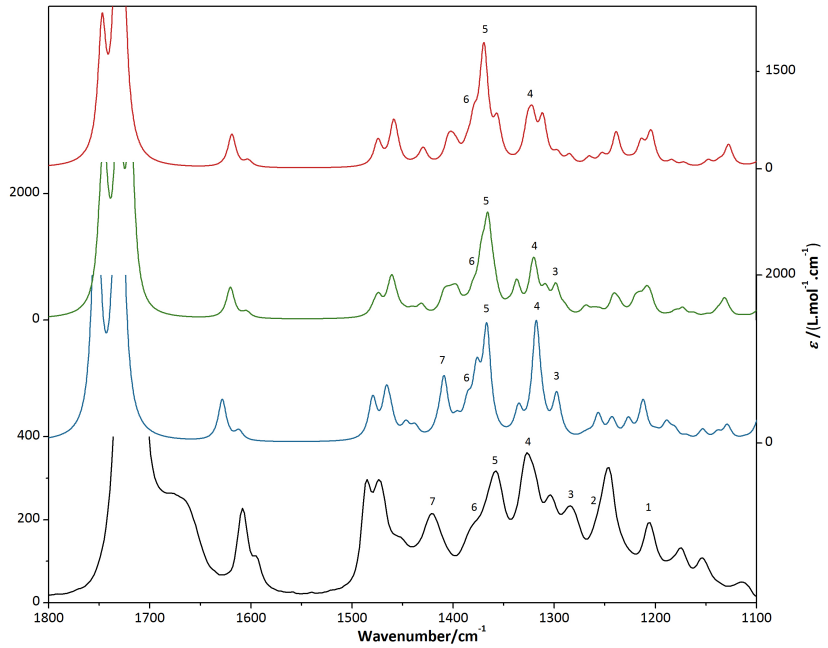


Figure 53 Comparison of the experimental IR spectra of compounds **3** (black) with calculated (ωB97XD/6-311G(d,p)) spectra for **3a** (blue), **3b** (green) and **3c** (red). Adapted from Cherblanc *et al.*[180]

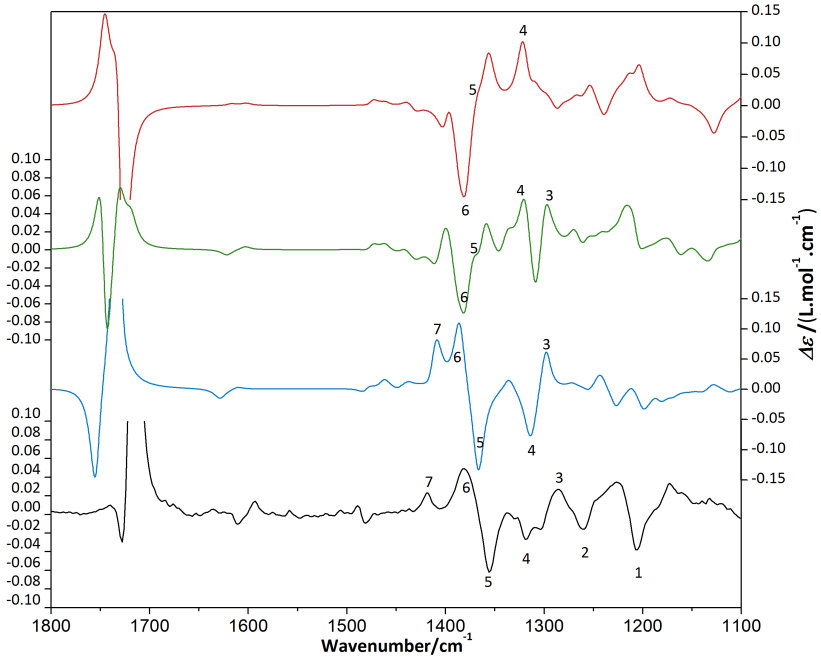


Figure 54 Comparison of the experimental VCD spectra of compounds **3** (black) with calculated (ωB97XD/6-311G(d,p)) spectra for **3a** (blue), **3b** (green) and **3c** (red). Adapted from Cherblanc *et al.*[180]

Notably, the spectrum of **2a** was the only one that reproduced the strong absorption at ν=1320 cm^{-1} (number 4). The simulated VCD spectra were scaled using the same scale factors determined for the IR spectra. Seven important peaks in the VCD spectrum, numbered 1–7 in Figure 53 and Figure 54 were correlated with peak in the

IR spectrum. Features 1 and 2 are not predicted for any of the three diastereoisomers. Only structure **3a** shows a good prediction of the remaining features (3–7). A poor agreement is observed for **3b** and **3c**. Feature 6 for example is negative for **3b** and **3c** while it is positive in the experimental spectrum. Importantly, a mirror-image relationship (negative vs. positive features) is meaningless here as it would imply that the enantiomer of the predicted stereoisomer was obtained, which is obviously impossible since this would involve the unnatural stereochemistry of the chaetocin core at C2/C3 and C2'/C3' (Scheme 19). Structure **3a** clearly showed the most satisfactory agreement and the visual VCD analysis was therefore also supportive that desulfurisation occurred with retention of stereochemistry. We subsequently used the algorithm mentioned earlier to assess the quality of the manual assignment in light of previous successful assignments (Figure 55).[179] The results of the analysis for compound **3a**, **3b** and **3c** are shown in Table 8. For **3b** and **3c**, Δ was found negative which means that the enantiomeric form of **3b** and **3c** have a higher similarity score than respectively **3b** and **3c**. As mentioned above however, this is chemically impossible as it implies inversion of stereochemistry at C2/C3 and C2'/C3', and **3b** and **3c** can thus be excluded. More importantly, the confidence level for the assignment for **3a** is computed to be very high (94%) in comparison with previous successful assignment (Figure 55, Table 8).

Table 8 Calculated quantities describing the similarity between calculated and experimental IR and VCD spectra.

Molecule	Optimal scaling factor	IR similarity (%)	Σ for best enantiomer	Δ	Confidence level (%)
3a	0.966	84.9	59.4	39.8	94%
3b	0.960	81.2	74.7	*–57.1*	99%
3c	0.961	84.7	54.1	*–12.9*	50%

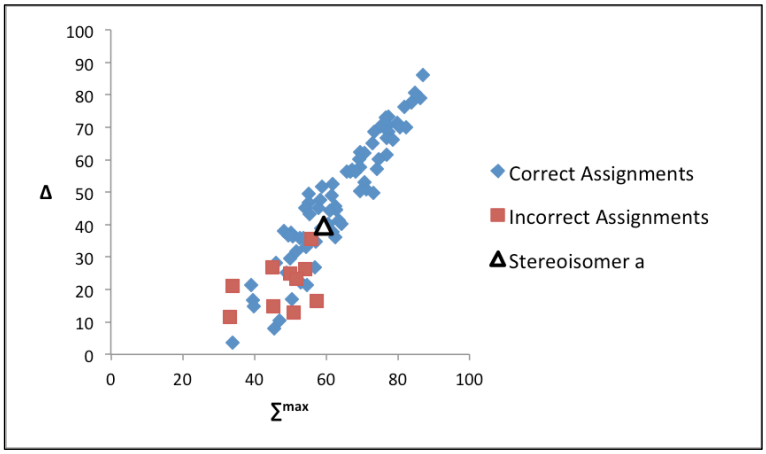


Figure 55 Positioning of **3a** in the Compare VOA database of previous assignments.

Since VCD analysis was in agreement with optical rotation and ECD analysis, the stereochemistry of the monosulfide chaetocin **3** could thus be assign with high confidence to that of structure **3a**, that is, retention of configuration of the bridgehead carbon atoms. This work has been published in Chemistry – A European Journal in September 2011.[180]

This assignment was in agreement with that of the monosulfide analogue of Sirodesmin PL **28** (Scheme 23), which also bears a pendent alcohol on the ETP core that has been suggested by Barbier and co-workers to be involved in the mechanism leading to retention of configuration (Scheme 26). Dehydrogliotoxin **21** (Scheme 20) however did not seem to follow the same trend and we thus decided to reconsider its stereochemical assignment.

Dehydrogliotoxin

Dehydrogliotoxin (**21**) has been reported to be desulfurised with inversion of configuration at the bridgehead carbon atoms to give the (*S,S*) monosulfide **22** (Scheme 20). As mentioned above, this assignment was solely based on the observation that the ECD spectrum of the desulfurised product **22** was approximately the mirror-image of that of the parent ETP **21**.[164] As the use of this correlative approach for absolute configuration determination has been cautioned,[165] we felt it was important to verify this assignment. The ECD spectrum of dehydrogliotoxin (**21**) was therefore simulated at the ωB97XD/6-311++G(d,p) level of theory and was compared to the reported experimental spectrum (Figure 56).[164] The two spectra were

in good agreement, displaying a strong negative Cotton effect around 230 nm and a positive one near 275 – 280 nm, therefore corroborating the accepted (*R*,*R*) stereochemistry for ETP **21** and the ability of the computational method to efficiently predict the ECD curve for this framework.

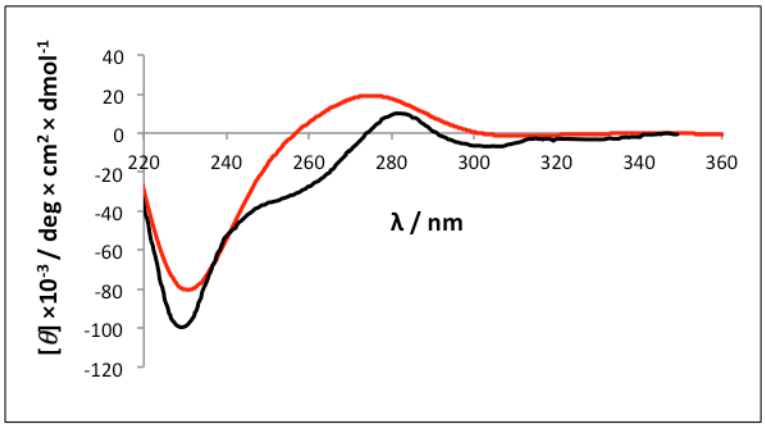


Figure 56 ECD spectra of dehydrogliotoxin (**21**)
Black line: literature spectra of **21** in 1,4-dioxane[164]; Red line: calculated spectra calculated at the ωB97XD/6-311++G(d,p) level of theory, scaled by 1.8 and shifted by +15 nm, convoluted with a line width 0.25 eV.

The ECD spectrum of dethiodehydrogliotoxin (**22**), was then simulated at the ωB97XD/6-311++G(d,p) level of theory for the reported (*S,S*) stereochemistry but this time however, the spectrum did not match at all (Figure 57). Strikingly, the simulated spectra for (*R,R*)-dethiodehydrogliotoxin (mirror-image of the (*S,S*) stereoisomer) was in very good agreement with the reported literature spectra.[164] Our data suggests that the assignment of stereochemistry by Safe and Taylor was therefore incorrect and in fact dethiodehydrogliotoxin (**22**) has the opposite stereochemistry (*R,R*) to that of structure **22** in Scheme 20 (*i.e. ent*-**22**). Consequently, the desulfurisation of dehydrogliotoxin (**21**) proceeded also with retention at both carbon centres, in opposition with Safe and Taylor conclusions.[164] The origin of this erroneous assignment will be discussed later. This work was also included in our publication in Chemistry – A European journal.[180]

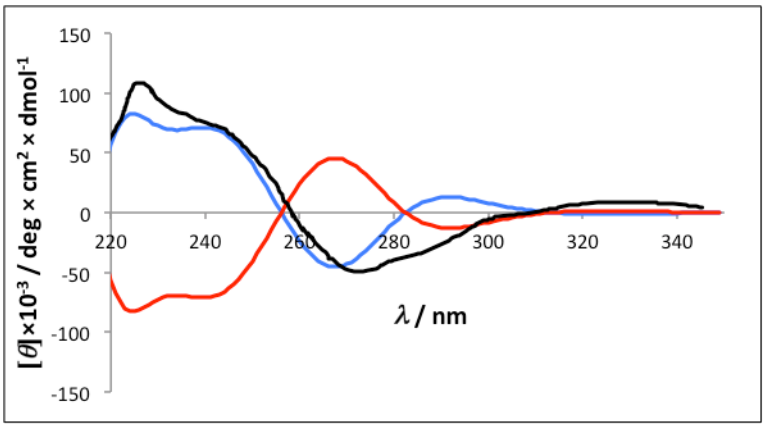


Figure 57 ECD spectra of dethiodehydrogliotoxin (**22**). Adapted from Cherblanc *et al*.[180]
Black: literature spectra[164] of **22** recorded in 1,4-dioxane; Red: calculated *S,S*-dethiodehydrogliotoxin (**22**); Blue: calculated *R,R*-dethiodehydrogliotoxin (*ent*-**22**). The calculations were done at ωB97XD/6-311++G(d,p) level of theory; scaled by 5 and shifted by +15 nm, convoluted with a line width 0.25 eV.

Gliotoxin analogue

At this stage, there remained only one example where ETP desulfurisation was reported to occur with inversion of configuration: the desulfurisation of synthetic gliotoxin analogue **25** into monosulfide **26**, reported by Ottenheijm and co-workers (Scheme 22).[169] As aforementioned, the stereochemical assignment of monosulfide **26** was based on anomalous dispersion X-ray crystallography. Ottenheijm and co-workers validated their stereochemistry assignment with a suitable R-factor test. This test is a simple and powerful approach that gives a measure of the agreement between the crystallographic model and the experimental X-ray diffraction data.

Although Ottenheijm and co-workers reported a mechanistic scenario to support the apparent inversion of configuration (Scheme 25),[169] we decided to computationally study the mechanistic course of this reaction, in collaboration with Prof. Henry Rzepa, by first calculating the energies along the retention pathways initially suggested by Sato and Hino (Figure 58, Scheme 24).[164,168,177]

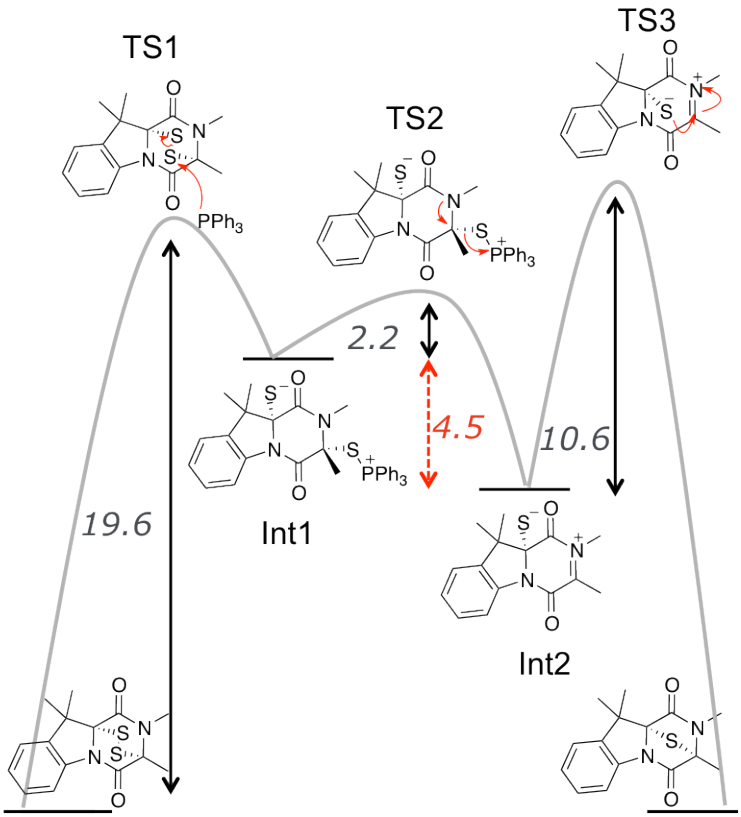
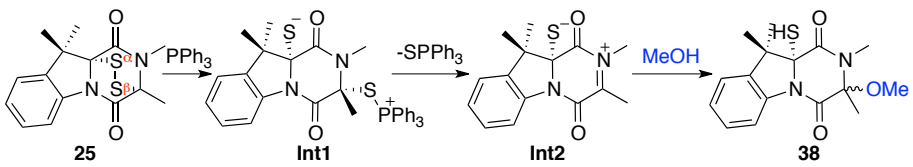


Figure 58 Energy barriers (in kcal/mol) for the desulfurisation of gliotoxin analogue **25** along the pathway leading to retention of stereochemistry (Scheme 24).
Indicated transition states and energy barriers were all located at the ωB97Xd/6-311G(d,p) level with inclusion of a continuum solvent model for THF (scale 1.5 cm = 4 kcal/mol). Triphenylphosphine or triphenylphosphine sulfide ($SPPh_3$) were included in the energy calculation throughout.

The initial bimolecular step involves nucleophilic attack of the triphenylphosphine on the β sulfur atoms of ETP **25** (TS1, Figure 58 and Scheme 27) with a calculated barrier $\Delta G_{298}^{\ddagger}$ of 19.6 kcal/mol relative to **25**. Attack on the α sulfur atom has a higher barrier of 27.7 kcal/mol due to the hindrance from the methyl groups. Consequently, only attack of the β sulfur corresponds to a room temperature reaction. Importantly, this regiochemical outcome is corroborated by the experimental trapping of a key intermediate (**Int2**) by methanol and isolation of compound **38** reported by Ottenheijm and co-workers (Scheme 27).[169]

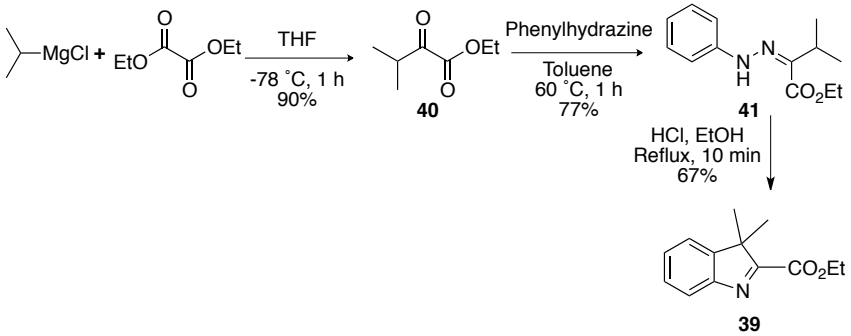


Scheme 27

The second unimolecular step *via* TS2 is the elimination of triphenylphosphine sulfide from **Int1** to form the zwitterionic intermediate **Int2**. The barrier for elimination of Ph_3PS from **Int1** is predicted to be very low; $\Delta G_{298}^{\ddagger}$ = 2.2 kcal/mol (TS2) relative to **Int1.** This energy is much lower than the one for TS1 and hence this second step is not rate determining. At this stage, it is worth pointing out that the inversion mechanism proposed by Ottenheijm and co-workers involved ring opening of the diketopiperazine, bond rotation and subsequent ring closure, prior to elimination of Ph_3PS (Scheme 25).[169] The very low activation barrier for Ph_3PS elimination from **Int1** (2.2 kcal/mol) strongly suggests that this intermediate is highly transient and therefore the mechanism proposed by Ottenheijm appears extremely unlikely. **Int2** itself is also computed to be 4.5 kcal/mol lower than **Int1** (Figure 58). The mechanism then involves re-closure of the zwitterionic species **Int2** to reform the sulfur bridge, with a barrier of 10.6 kcal/mol (TS3, Figure 58). Here again, the activation barrier is significantly lower in energy than the one leading to TS1. This suggests that, for this pathway, attack by triphenylphosphine (TS1) is rate determining, with the remainder of the pathway being energetically downhill.

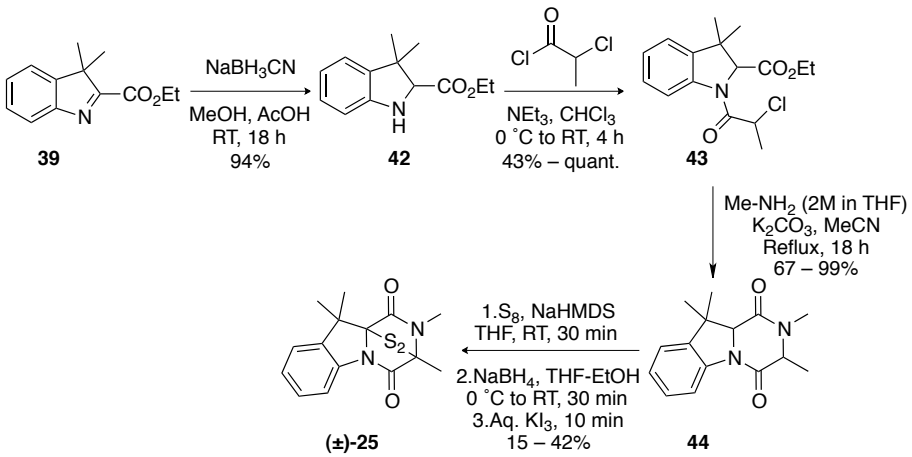
The energy profile of this reaction sequence strongly suggests that desulfurisation of ETP compound **25** occurs with retention of configuration. The mechanism proposed by Ottenheijm and co-workers leading to highly stereospecific inversion of both stereogenic centres requires that PPh_3S would be stable from elimination in **Int1**,[169] which does not seem realistic in light of the calculated energies.

This computational study thus brought into further questioning the observations of Ottenheijm and co-workers. Consequently, we decided to prepare ETP **25** and to investigate its desulfurisation in our own hands. Racemic ETP **25** was prepared in seven steps from commercially available material. Indolenine **39** was prepared following literature procedure (Scheme 28):[182] Addition of *i*PrMgCl to diethyl oxalate gave the product of monoalkylation **40** in high yield (90%). This was condensed with phenyl hydrazine to afford arylhydrazone **41** in good yield (77%), which was then cyclised by refluxing in ethanolic HCl to give the desired indolenine **39** in good yield (67%) in a Fischer-type reaction.



Scheme 28[183]

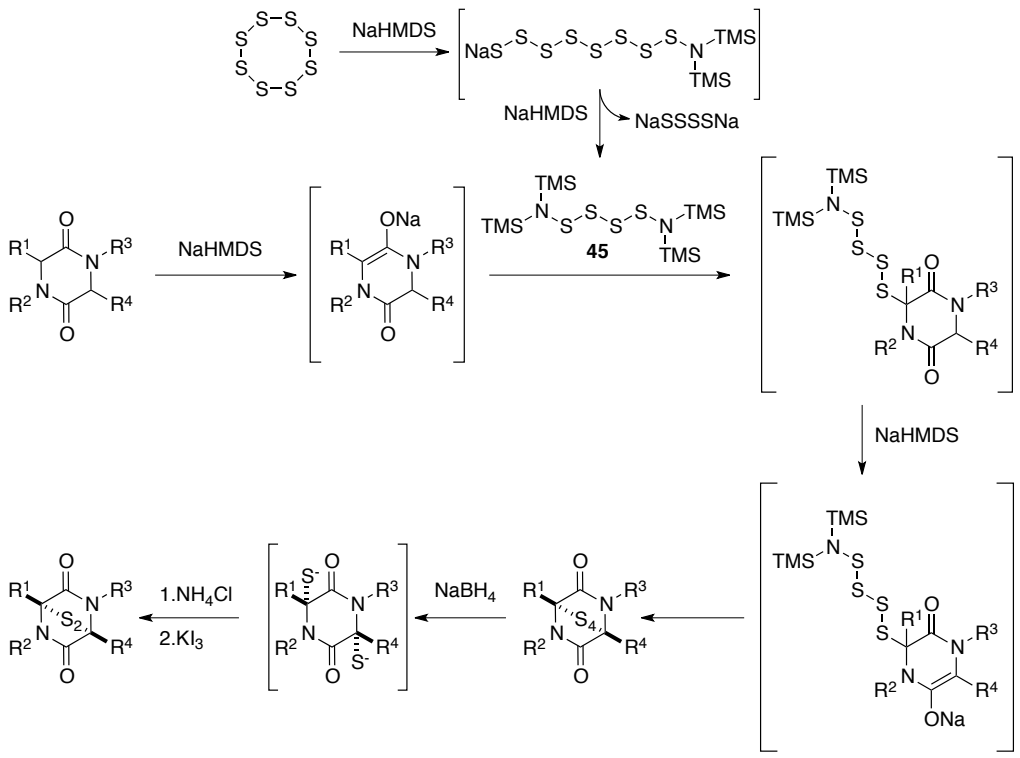
Indolenine **39** was reduced to indoline **42** in excellent yield (94%) by treatment with $NaBH_3CN$ (Scheme 29).[184] $NaBH_4$ was used as a safer alternative for scaled-up synthesis and afforded the product in similar yield. Indoline **42** was directly treated with 2-chloropropanoyl chloride to afford the acylated product **43** (Scheme 29).[185] Refluxing acylated indoline **43** with methylamine in THF/MeCN gave diketopiperazine **44**.[185] Both compound **43** and **44** were obtained as a mixture of diastereoisomers but separation was not necessary, as the two chiral centres will lose their stereochemical information in the final step.



Scheme 29[183]

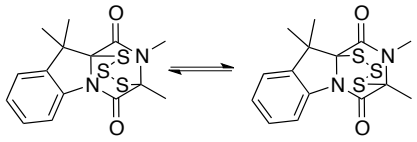
The sulfenylation of diketopiperazine **44** was performed using the method recently described by Nicolaou and co-workers (Scheme 29 and 30).[100,102] The reaction of NaHMDS and elemental sulfur (S_8) is postulated to give *N,N'*-tetrathiobis-trimethylsilyl (**45**, Scheme 30). Deprotonation of the diketopiperazine ring with NaHMDS formed an enolate which reacted with the tetrasulfide species **45** to lead to the formation of the first sulfur-carbon bond. Subsequent alpha deprotonation and

intramolecular attack of the enolate to the tetrasulfide moiety led to the formation of the second sulfur-carbon bond and the installation of the tetrasulfide bridge. The latter was then reduced with $NaBH_4$ to the dithiolate and final oxidation with KI_3 led to the desired ETP **25**.



Scheme 30 Postulated mechanism for the general formation of the ETP core from the corresponding diketopiperazine. Adapted from [100,102]

ETP (±)-**25** was thus obtained in low to medium yield (15 – 42%). Attempts at improving the yield of this reaction by varying the conditions (temperature, base, order of addition, reaction time) did not lead to significant improvement. A trisulfide analogue (**46**, Scheme 31) was consistently observed in 5 – 10% yield. In addition to a slight change in chemical shifts from the disulfide analogue, the NMR spectra of trisulfide **46** exhibited duplicated signals. This has been observed for a number of natural and synthetic products containing the epitrisulfide functionality and it is commonly accepted that the compound actually exists as two conformers (Scheme 31).[96,186,187] The trisulfide **46** could also be desulfurised to the monosulfide **26**, *via* the disulfide **25**, to limit loss of material.[96]



46

Scheme 31 The two conformers of the trisulfide derivative **46**

The enantiomers of ETP **25** were then separated using semi-preparative HPLC on a chiral stationary phase. The first ETP enantiomer eluted from the HPLC column had an optical rotation and an ECD spectrum in agreement with those of the starting material used by Ottenheijm and co-worker in their desulfurisation studies, assigned to have (*R,R*) stereochemistry by qualitative comparison with the ECD spectra of gliotoxin.[169,188] Since we and others showed that correlative methods can lead to important errors,[165,179,180] we sought to confirm this assignment by *ab initio* calculations as described above.

The optical rotation of (*R,R*)-**25** was first investigated. At the M062X/6-311++G(d,p) level of theory, the optical rotation of (*R,R*)-**25** was predicted to be −498. Ottenheijm *et al.* had reported a value of −502 (*c* 2.345, $CHCl_3$),[188] and we recorded a value of −335 (*c* 2.00, $CHCl_3$). The strong negative sign of the optical rotation was in agreement with the configuration of this stereoisomer to be (*R,R*). The ECD spectra were likewise experimentally measured and computationally simulated at the M062X/6-31++G(d,p) level of theory for (*R,R*)-**25**. Figure 59 shows the experimental ECD spectra recorded by us and the one reported by Ottenheim et al.,[188] together with the predicted one. The two experimental spectra are comparable with three negative Cotton effects around 230, 250 and 280 nm. The Cotton effects at 230 and 250 nm are observed in the simulated spectra but the one at 280 nm is of the opposite sign. While the simulation of (*R,R*)-**25** is undoubtedly closer to the experimentally obtained spectra than the simulated spectrum for its enantiomer ((*S,S*)-**25**, mirror image of that shown in Figure 59), the calculations were unable to correctly predict all three negative Cotton effects.

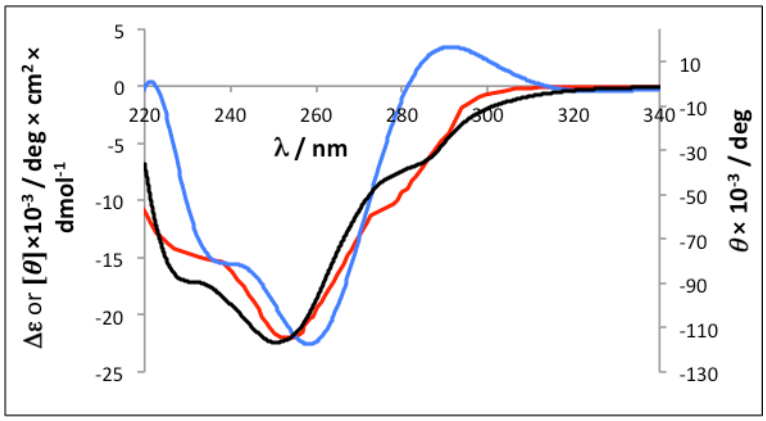


Figure 59 ECD spectra of ETP **25**. Adapted from Cherblanc *et al.*[183]
Black: Experimental data for **25** in dichloromethane (right *y* axis); Red: Data from Ottenheijm and co-workers[188] for **25** (left *y* axis, reported as Δε), Blue: Calculated data for (*R,R*)-**25** at M062X/6-31++G(d,p) (solvent=dichloromethane) (left *y* axis, rescaled by 0.385, shifted by +25 nm, and convoluted with a line width 0.24 eV).

VCD analysis was therefore undertaken to further validate the assignment of ETP **25**. The experimentally measured and computationally simulated IR and VCD spectra for (*R,R*)-**25** are depicted in Figure 60. Analogously to our work on chaetocin, a scaling factor (0.968) derived from the IR similarity measure was used to rescale the VCD spectrum.[179] Visual examination of the data showed an excellent agreement between calculated and experimental spectra: most bands observed in the IR and/or VCD spectra were predicted by the calculations. The agreement was further confirmed by the similarity measures Σ (describing the level of agreement between the experimental VCD spectra with the theoretical data derived for each enantiomer). The resulting values were Σ_{RR}=75.5 and Σ_{SS}=6.9. Combination of the enantiomeric similarity index Δ (the difference between the Σ value of each enantiomer) and the spectral database in the CompareVOA program[179] computed a confidence level of 100% for the (*R,R*) stereochemistry. Since all three chiroptical techniques used were supportive of the (*R,R*) assignment of ETP **25**, we were therefore confident in the stereochemistry reported by Ottenheijm and co-workers.[169,188]

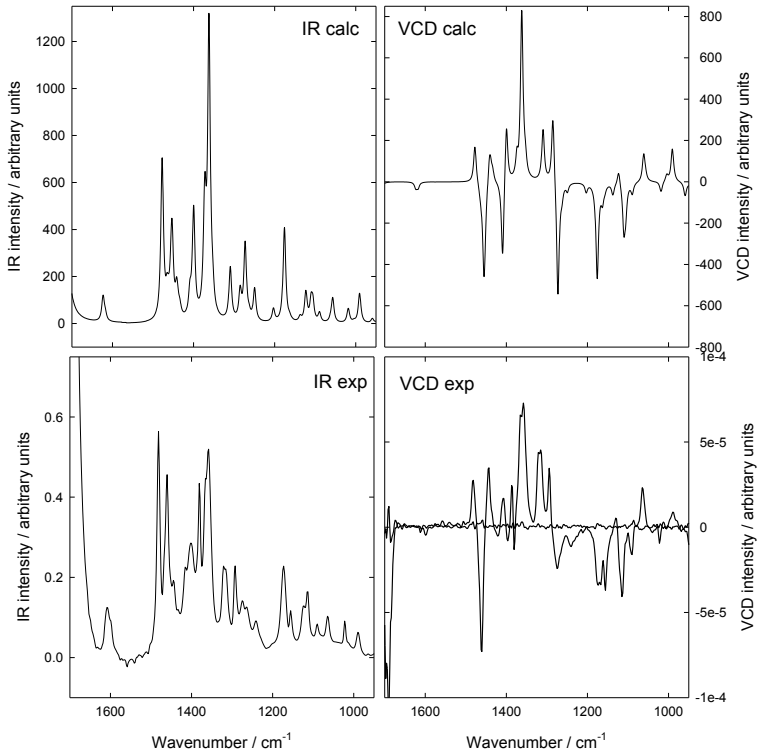


Figure 60 IR and VCD Spectra of ETP **25**. Taken from Cherblanc *et al.*[183]
The top panels show the calculated IR and VCD spectra for (*R,R*)-**25** obtained at M062X/6-311++G(d,p) (solvent=chloroform). The bottom panels show the experimental IR and VCD spectra for **25** obtained for a solution in $CDCl_3$. The theoretical spectra were obtained by using a scale factor of 0.968. The thin black line given in the bottom right panel refers to the noise spectrum supplementing the measured VCD data.

ETP (*R,R*)-**25** was desulfurised with triphenylphosphine in dioxane as described by Ottenheijm and co-workers.[169] A 93:7 enantiomeric ratio (e.r.) of the monosulfide product **26** (Scheme 22) was observed by chiral HPLC. Similarly, ETP (*S,S*)-**25** afforded a 2:98 e.r. of monosulfide **26**; the major isomer being the enantiomer to that obtained with (*R,R*)-**25**. Semi-preparative chiral HPLC was also performed on racemic mixtures of product **26** (obtained from the desulfurisation of racemic ETP **25**) to access larger and enantiopure quantities of monosulfide **26**.

The enantiomer of **26** of shorter retention time by chiral HPLC had an optical rotation in agreement with that reported by Ottenheijm and co-workers [−47.5 (*c* 1.12, CH_2Cl_2) and −53 (*c* 1.13, CH_2Cl_2),[169] respectively]. At the M062X/6-311++G(d,p) level of theory, the calculated optical rotation of (*R,R*)-**26** was −169. Although this predicted value was qualitatively different from the experimental ones, the strong negative sign of the rotation suggests that this enantiomer is of a (*R,R*) configuration. Since this enantiomer was the major product obtained by desulfurisation of (*R,R*)-**25**, this indicates that the reaction proceeded with retention of configuration, in opposition with Ottenheijm's conclusions.

To further validate our results, the ECD spectra of the supposed (*R,R*)-**26** were recorded (Figure 61). Our spectrum was in good agreement with that of Ottenheijm and co-workers,[169] with two positive Cotton effects (around 220 and 250 – 260 nm) followed by a negative one (280 – 290 nm) and finally a positive one (320 nm). The ECD spectrum of (*R,R*)-**26** was simulated at the M062X/6-31++G(d,p) level of theory but it did not compare well with the experimental data (Figure 61). The ECD spectrum for (*S,S*)-**26** would be the mirror image of the spectrum depicted in Figure 61 and would not be in good agreement with the experimental data either.

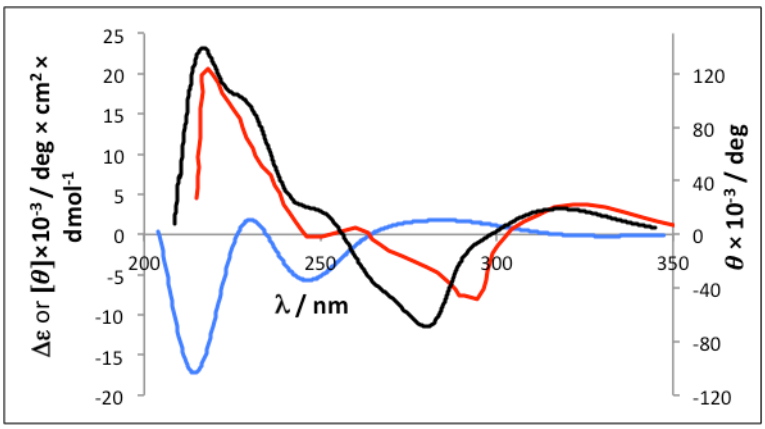


Figure 61 ECD spectra of monosulfide **26**. Adapted from Cherblanc *et al.*[183]
Black: Measured data for **26** (enantiomer 1) (right *y* axis); Red: Data from Ottenheijm and co-workers[169] for **26** (left *y* axis, reported as Δε), Blue: Calculated data for (*R,R*)-**26** at M062X/6-31++G(d,p) (solvent=dichloromethane) (left *y* axis, rescaled by 0.167, shifted by +25 nm, and convoluted with a line width 0.24 eV).

We therefore decided to turn our attention to VCD spectroscopy as this technique gave satisfactory results for the parent molecule **25**, although the simulated ECD spectrum was also in moderate agreement with the experimental spectra. The IR and VCD spectra of supposed (*R,R*)-**26** and the corresponding simulated spectra are shown in Figure 62.

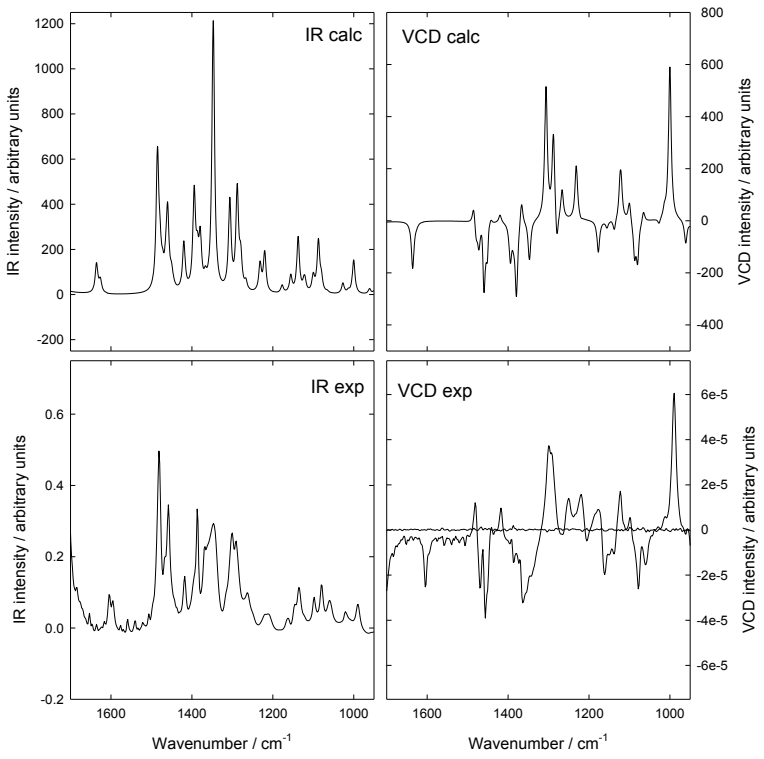
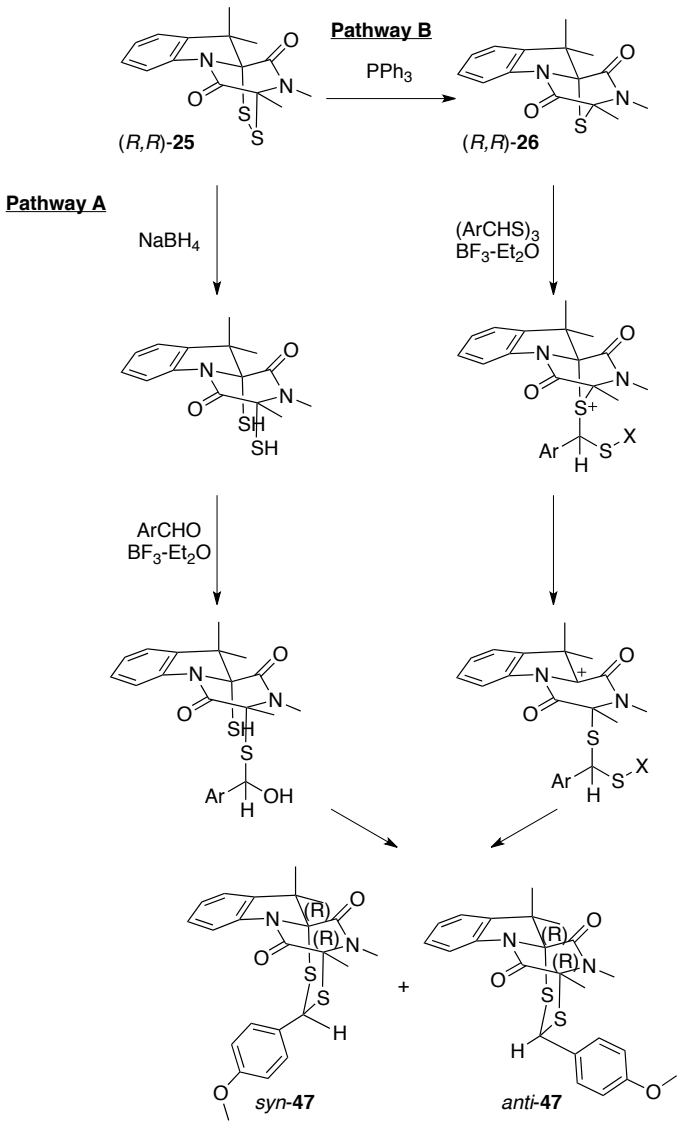


Figure 62 IR and VCD Spectra of monosulfide **26**. Taken from Cherblanc *et al.*[183]
The top panels show the calculated IR and VCD spectra for (*R,R*)-**26** obtained at M062X/6-311++G(d,p) (solvent=chloroform). The bottom panels show the experimental IR and VCD spectra for **26** obtained for a solution in $CDCl_3$. The theoretical spectra were obtained by using a scale factor of 0.976. The thin black line given in the bottom right panel refers to the noise spectrum supplementing the measured VCD data.

Gratifyingly, experimental and calculated spectra were in excellent agreement. The similarity measures were calculated for both enantiomers: Σ_{RR}=66.5 and Σ_{SS}=4.2. The derived enantiomeric similarity index Δ in combination with the spectral database in CompareVOA[179] computed a confidence level of 99% for the (*R,R*) stereochemistry for **26**. Since this enantiomer was obtained by desulfurisation of (*R,R*)-**25**, the VCD results, together with the optical rotation comparison, strongly advocates that the reaction proceeded with retention of configuration.

While VCD and optical rotation were clearly supportive of this assignment, the results obtained for the ECD comparison were unsatisfying and we therefore felt that an additional method would strengthen the reliability of our assignment. We reasoned that a non-spectroscopic approach would be particularly relevant and we therefore decided to use the method of chemical derivatisation that Barbier and co-workers employed to determine the stereochemistry of the desulfurisation of Sirodesmin PL **27**.[171] The natural product **27** (Scheme 23), of (*R,R*) configuration, as determined by X-ray crystallography analysis, was stereospecifically converted to a

thioacetal. In parallel, monosulfide **28**, isolated from the desulfurisation of the parent ETP **27**, was converted to the same thioacetal. Therefore, if the bridgehead carbon atoms in the thioacetal have the same stereochemistry when obtained from ETP **27** or from monosulfide **28**, then the desulfurisation reaction proceeded with retention of configuration at these centres. The validity of this approach has been confirmed by Barbier and co-worker by X-ray crystallography of an acetylated derivative of monosulfide **28** and appeared therefore as a suitable method for determining the relative stereochemistry of parent ETP and corresponding monosulfide. A summary of this approach, applied to the substrates of interest (ETP **25** and monosulfide **26**) for the formation of thioacetal **47**, is shown in Scheme 32.



Scheme 32 Derivatisation strategy applied to ETP **25** and monosulfide **26**. Ar– = (CH_3OPh)–; –X = –$[CH(p\text{-}CH_3OPh)S]_2$–$BF_3$. Adapted from[171,183]

We therefore prepared thioacetal **47** from ETP (*R,R*)-**25** and monosulfide **26** of suggested (*R,R*) stereochemistry. The reactions were first attempted on the racemic starting material. ETP (±)- and (*R,R*)-**25** were reduced to the dithiol by treatment with sodium borohydride, and converted directly into thioacetal **47** by treatment with anisaldehyde in presence of boron trifluoride. Thioacetal **47** was obtained in modest yield as a mixture of diastereoisomers, with the *syn* product (with respect to the anisaldehyde and polycyclic residues[171], Scheme 32) as a major product (43% from (±)-**25**, 13% from (*R,R*)-**25**) and only traces of the *anti* product (observed by TLC and LCMS of the crude but not isolated from (±)-**25**, 6% from (*R,R*)-**25**). The assignment of the *syn* diastereoisomers was based on a NOESY NMR experiment, which showed a correlation between the signal at 5.10 ppm (CH of the thioacetal carbon) and the signal at 3.27 ppm (CH_3 on the nitrogen atom of the diketopiperazine ring) (Figure 63). This correlation was absent from the NOESY NMR of the *anti* diastereoisomer.

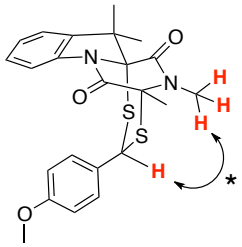
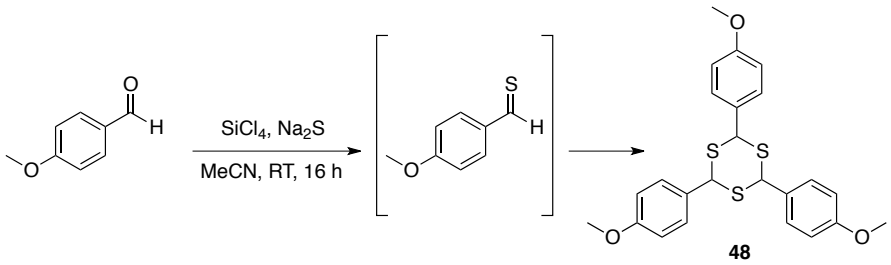


Figure 63 NOESY correlation observed for the *syn* diastereoisomer of **47**

Concomitantly, thioacetal **47** could be obtained from the racemic and enantiopure (proposed to be (*R,R*) stereochemistry) monosulfide **26** by treatment with the trithiane derivative of anisaldehyde **48** (Scheme 32). This trithiane **48** was prepared following a recent literature procedure by reaction of anisaldehyde with $SiCl_4$ and Na_2S in acetonitrile (Scheme 33).[189]



Scheme 33

Interestingly, in $CDCl_3$ both the α and β conformer of trithiane **48** (Figure 64) could be observed with the ratio of α increasing with temperature. In d_6-DMSO only the β isomer, the most stable chair conformation, was observed.[190] The acidity of chloroform may be responsible for the inter-conversion observed *via* opening of the 6-membered ring.

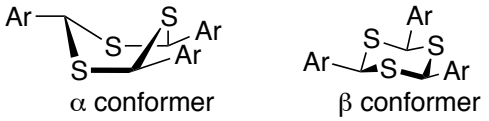


Figure 64 The α and β conformers of the trithiane of anisaldehyde **48**, Ar– = (CH_3OPh)–.[190]

The thioacetal **47** was also obtained as a mixture of diastereoisomers readily separable by column chromatography (*anti* : *syn* ratio of 70:30 from (±)-**26** and 40:60 from (*R,R*)-**26**) in medium to low yields (*anti*: 56% from (±)-**26**, 17% from (*R,R*)-**26**; *syn*: 25% from (±)-**26**, 25% (*R,R*)-**26**).

The *syn* diastereoisomers isolated had identical spectra (^{1}H-NMR, ^{13}C-NMR, HSQC and NOESY) to the ones obtained as major products from reaction of (±)- and (*R,R*)-**25**. Importantly, the optical rotations of the *syn* thioacetal **47** obtained from ETP (*R,R*)-**25** and monosulfide (*R,R*)-**26** were comparable: −66 (*c* 0.20, $CHCl_3$) from ETP **25** *via* pathway A and −40 (*c* 0.20, $CHCl_3$) from monosulfide **26** *via* pathway B. Final confirmation was obtained by analytical HPLC on a chiral stationary phase. Indeed, the thioacetal products obtained from enantiopure compounds by the two different pathways were found to be the same enantiomer as only one peak was observed upon spiking (mixing) (Figure 65). As a control, Figure 66 shows the trace obtained by spiking the thioacetal *syn*-**47** obtained by reaction of (±)-**25** and the one obtained by reaction of (±)-**26**, and it is clearly apparent that the two enantiomers are very well separated (approximately 40 min).

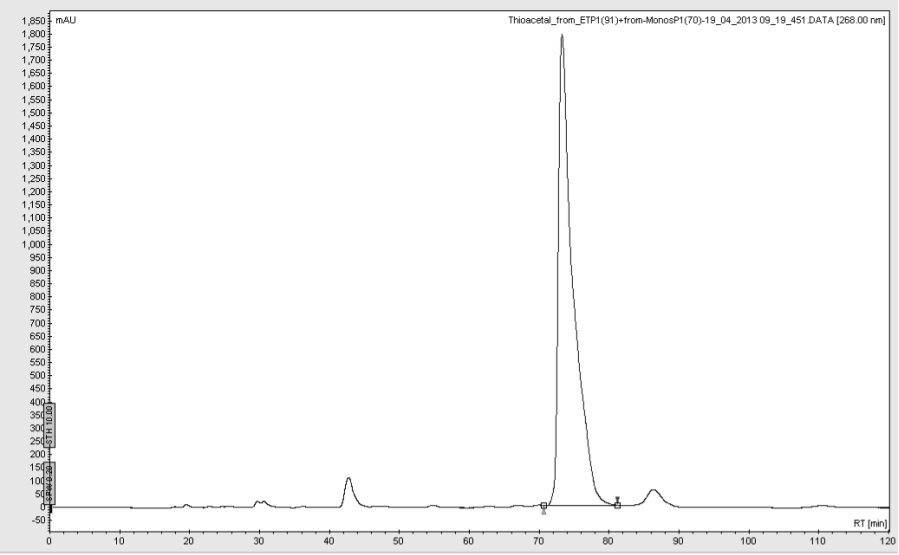
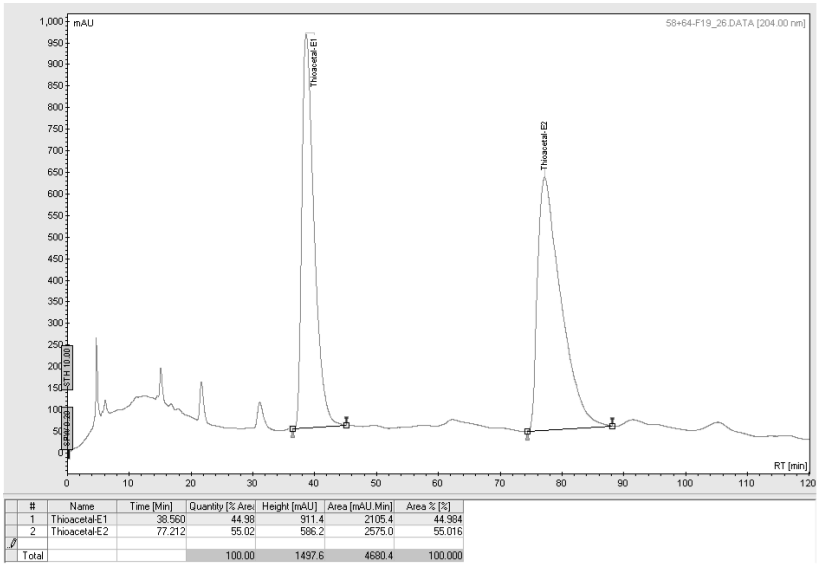


Figure 65 Thioacetal **47** obtained by reaction of (*R,R*)-**25** spiked with thioacetal **47** obtained by reaction of (*R,R*)-**26**, (*syn* diastereoisomer) (Hexane – Isopropanol, 98:2)



#	Name	Time [Min]	Quantity [% Area]	Height [mAU]	Area [mAU.Min]	Area % [%]
1	Thioacetal-E1	38.560	44.98	911.4	2105.4	44.984
2	Thioacetal-E2	77.212	55.02	586.2	2575.0	55.016
Total			100.00	1497.6	4680.4	100.000

Figure 66 Thioacetal **47** obtained by reaction of (±)-**25** spiked with thioacetal **47** obtained by reaction of (±)-**26**, (*syn* diastereoisomer) (Hexane – Isopropanol, 98:2)

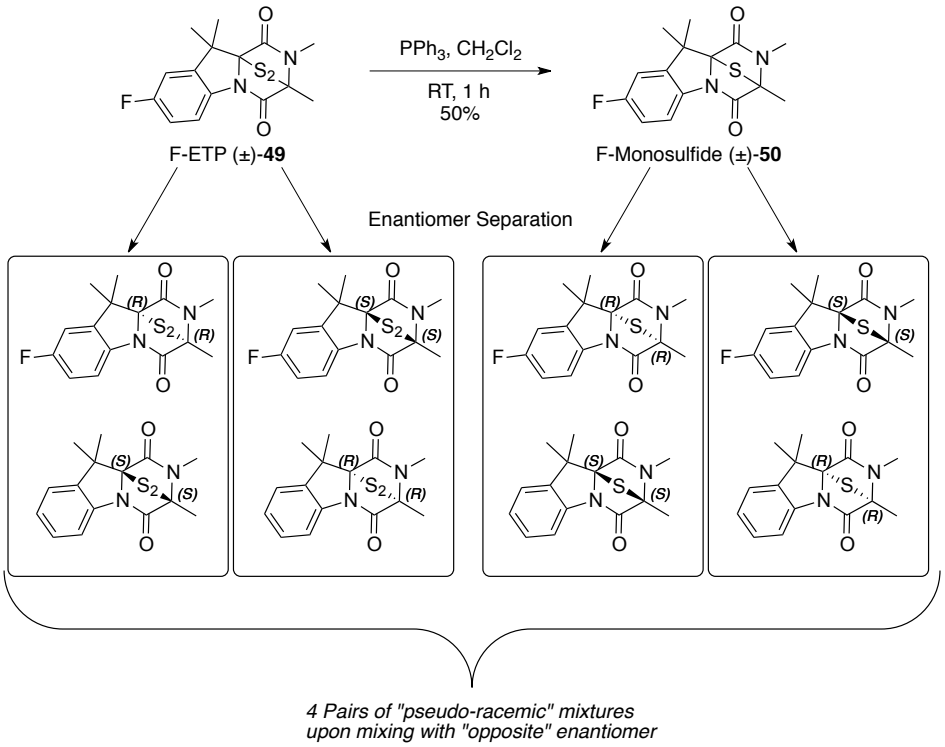
The same enantiomer of thioacetal **47** was obtained *via* either pathway and we could therefore conclude that the stereochemistry of ETP **25** and monosulfide **26** was the same, *i.e.* (*R,R*).

Our data, obtained by two orthogonal methods, strongly advocates that desulfurisation of ETP (*R,R*)-**25** occurred with retention of stereochemistry at the bridgehead carbon atoms contrary to the original report.[169] This work has been accepted for publication in the Journal of Organic Chemistry in October 2013.[183]

Nevertheless, we are still not in a position to identify the origin of Ottenheijm's misassignment, validated by an X-ray crystal structure. We still felt that obtaining a crystal structure of our samples would definitely ensure the accuracy of the results obtained by chemical derivatisation and by analysis of chiroptical spectra.

We thus tried to crystallise both starting material (*R,R)*-**25** and monosulfide (*R,R)*-**26** using various techniques such as slow evaporation from various solvents including dichloromethane as reported,[169] chloroform, methanol; solvent diffusion whereby the sample was dissolved in dichloromethane and a layer of n-octane was carefully added on top of it; vapour diffusion (combinations of dichloromethane, water, methanol, diethylether) but we could not obtained any crystals of suitable quality for X-ray analysis. The *racemic* starting material ETP **25** did crystallise from dichloromethane in the triclinic crystal system, of space group $P\bar{1}$ (centrosymmetric, achiral) and the unit cell contained two independent molecules (See crystallographic data). Ottenheijm and co-workers also reported the crystal structure of the racemic but not of the enantiopure ETP **25**.[188] Interestingly, racemic ETP **25** crystallised (from methanol–water) in a different crystal system (monoclinic, space group C2/c). The C-S-S torsion angles and the S–S bond length are however comparable in both structures.

Encouraged by the crystallisation of the racemic mixture of ETP **25**, we envisaged that mixing the (*R,R*)-**25** with the "opposite" enantiomer of a highly similar compound could potentially trigger the crystallisation of this "pseudo-racemic" mixture (Scheme 34). X-ray crystallography could then confirm the stereochemistry of the two different molecules. We selected a fluorinated analogue of ETP **25** (**49**, abbreviated F-ETP, Scheme 34**)**, with the fluorine atom on the phenyl ring, far from the ETP moiety in order to keep the disruption of the three-dimensional structure and hopefully the packing to a minimum. Upon desulfurisation of F-ETP **49** into the corresponding monosulfide **50** and mixing with "opposite" enantiomer, we reasoned that this strategy may allow us to obtain crystal structures of both enantiopure ETP **25** and monosulfide **26** (Scheme 34).



Scheme 34 Strategy of the "pseudo-racemic" mixture preparation

The synthesis of F-ETP **49** was carried out following the same route described for ETP **25** (Scheme 28 and 29) except using 4-fluorophenylhydrazine to form the dihydroindole core (See Experimental Details, compounds **49** to **55**). The yields were similar throughout except the sulfenylation step tended to give better yield. This observation is most likely related to the electron-withdrawing inductive effect of the fluorine atom, which would decrease the pK_a of the hydrogen α to the amide moiety. Racemic F-ETP **49** crystallised readily in the flask and an X-ray crystal structure was obtained (see crystallographic data). Once again, the unit cell contained two independent molecules, and F-ETP **49** crystallised in the same triclinic crystal system, space group $P\bar{1}$ (centrosymmetric), as ETP **25**. Consequently, the enantiomers of a fraction of F-ETP **49** were separated using semi-preparative HPLC on a chiral stationary phase. Their stereochemistry was assigned by comparison with ETP **25** and based on the sign of their optical rotation and their order of elution. As extra precaution, the optical rotation were also calculated (at the ωB97XD/6-311+G(d,p) level of theory) and the optical rotation of (*R,R*)-**49** was predicted to be –448, in close agreement with the value we measured (–419 (*c* 1.05, $CHCl_3$)). Consequently, ETP (*R,R*)-**25** was mixed in a 1:1 ratio with F-ETP (*S,S*)-**49** (and the opposite enantiomers

too, Scheme 34). Various solvent and techniques were attempted for crystallisation, but unfortunately no suitable crystals could be obtained.

F-ETP (±)-**49** was desulfurised to the monosulfide **50** by treatment with triphenylphosphine as described before in medium yield (50%) (Scheme 34). A crystal structure of the *racemic* fluorinated monosulfide **50** could be obtained (monoclinic crystal system, $P2_1/n$ space group, see crystallographic data). The angles and bond lengths observed were comparable to the one described by Tijhuis and co-workers for monosulfide **26** although the crystal system was different (orthorhombic).[170] In particular, we observed a much more pronounced boat conformation compared to the parent ETP **49** due to the presence of the highly strained monosulfide bridge.

Encouraged by the crystallisation of the racemic fluorinated monosulfide **50**, the enantiomers were also separated using semi-preparative HPLC on a chiral stationary phase. The stereochemical assignment was done by comparison with monosulfide **26** and the optical rotation was calculated (at the ωB97XD/6-311+G(d,p) level of theory). The optical rotation of (*R,R*)-**50** was predicted to be −124, and we measured −32.2 (*c* 1.01, $CHCl_3$). The same difference in predicted and experimental values has been observed for monosulfide **26** (calculated: –131, measured: −47.5 (*c* 1.12, CH_2Cl_2)). The strong negative sign of the predicted value and the agreement between the two experimental values (fluorinated and non-fluorinated monosulfide) were in strong support of a correct assignment. The samples were then mixed with the "opposite" enantiomer of the non-fluorinated monosulfide (Scheme 34). Unfortunately again, after a number of attempts, no suitable crystals were obtained.

In summary, although a wide range of methods were attempted, we were unable to crystallise our sample of monosulfide (*R,R*)-**26**. It is worth noting that Ottenheijm and co-workers also struggled to obtain crystals, as judged by their comment: "after numerous attempts **one** suitable crystal could be prepared for the X-ray analysis".[169]

Nevertheless, even if we had managed to obtain a crystal structure confirming the (*R,R*) stereochemistry of monosulfide **26**, we would still be unable to identify the origin of its incorrect stereochemical assignment by Ottenheijm and co-workers.[169] As mentioned earlier, they did obtain an X-ray crystal structure of monosulfide (*S,S*)-**26** and we can only speculate on the origin of this data (*e.g.* error in the labelling of the

ETP starting materials or in the monosulfide samples). Also, it should be noted that re-analysis (optical rotation or ECD) of the crystal used for the crystallographic studies was not performed; this would have ensured that its absolute configuration matched that of the bulk material.

Conclusions

Through the comparison of simulated and experimental chiroptical spectra, we have managed to unambiguously assign the stereochemistry of an analogue of the natural product chaetocin, monosulfide **3**, the biological activity of which is described elsewhere in this thesis (Chapter II).

This led us to re-assess the stereochemical course of the desulfurisation of ETP-containing compounds. Indeed, the various assignments prior to our work were often contradictory (see Scheme 20 to 23) and no unified mechanism could be proposed. In particular, the pendent alcohol present in a number of natural products (including sirodesmin PL (**27**) and chaetocin (**1**)) was proposed by Barbier *et al.* to potentially influence the course of the reaction and to favour a double inversion mechanism, ultimately leading to net retention of stereochemistry at the bridgehead carbon atoms (Scheme 26).[171] This was proposed in light of Ottenheijms's assignment[169] (inversion of stereochemistry) and was in agreement with our findings but was inconsistent with the assignment proposed for dehydrogliotoxin (**21**), which was proposed to be desulfurised with inversion of stereochemistry based on the comparison of the ECD curves of the product and that of the ETP starting material. Our studies on the monosulfide derivative of dehydrogliotoxin **22** allowed us to shed light on the misassignment of its stereochemistry, by simulating the ECD spectra of the two enantiomers and comparison with the reported one. Finally, gliotoxin analogue **25** remained the last example in the literature of an inversion of stereochemistry upon desulfurisation. Computational and chiroptical studies coupled with chemical derivatisation allowed us to re-assign this compound: all our data strongly support the fact that ETP **25** is also desulfurised with retention of stereochemistry.

We thus propose that all chiral ETP compounds are desulfurised with retention of stereochemistry at the bridgehead carbon atoms, likely via a unified mechanism depicted in Scheme 24 and Figure 58.

In broader terms, it is important to highlight that some of the incorrect stereochemical assignments of desulfurised ETP compounds in the literature stemmed from the use of correlative methods, particularly ECD spectroscopy. Our study,

coupled with literature examples, clearly shows that the sign of a given Cotton effect is not characteristic for ETP or monosulfide stereochemistry. In this case, although differing by only one sulfur atom, the ECD spectrum of an ETP compound is not comparable with that of its monosulfide analogue because different molecular orbitals are involved for the transitions in each molecule. In addition, the interference between the different chromophores present in a given molecule (*e.g.* the disulfide and the diene functionalities in the study on gliotoxin[166]) appear to have a significant impact on the resulting ECD spectra and any changes in chromophore may lead to important changes in the spectrum which are not necessarily related to the absolute configuration of the molecule.[165] Overall, correlative methods whereby the stereochemistry of a compound is assigned by comparison of its ECD curve to that of a related (but different) molecule should definitely be discarded, regardless of how structurally similar the two molecules are.

In conclusion, the exploitation of chiroptical data for stereochemistry assignment should not be performed solely by analogy with similar molecules but should always be accompanied by simulations, allowed by the computational power available today. It is also good practice to use several techniques where possible to validate an assignment with confidence.[191] Indeed, in the case of monosulfide **26**, the simulated ECD spectrum did not correlate well with the experimental one, while optical rotation and VCD gave a clear answer.

Chapter IV

Towards the discovery of novel EZH2 inhibitors

Introduction

As outlined in Chapter I, epigenetic therapy holds important promise for the future of targeted chemotherapy in cancer.[40] Research in epigenetics over the past two decades has uncovered a number of new targets for the drug discovery community, including the various enzyme classes described above (*e.g.* DNMTs, HDACs, HKMTs).[192]

HKMTs have only recently been recognised as new promising targets, and of those enzymes, Enhancer of Zest Homologue 2 (EZH2) is possibly one of the most therapeutically relevant for cancer therapy.[39] EZH2 is the catalytic subunit of the Polycomb Repressive Complex 2 (PRC2) and it trimethylates H3K27, thereby inducing the silencing of genes involved in critical processes such as differentiation, and proliferation.[193] To date, only EZH2 (and in some cases the closely related EZH1) was found able to perform trimethylation of H3K27.[194] Mono and dimethylation of H3K27 have been reported to be also mediated by G9a (primarily a H3K9 dimethylase) both *in vitro* and *in vivo*, but trimethylation was not observed.[195] This absence of redundancy in the enzyme able to install H3K27me3 is distinctive amongst the histone methylating enzymes. For example, H3K9 can be methylated by several enzymes including G9a, GLP and SUV39H1/2.[13] This unique feature of the H3K27me3 mark has important implications in terms of its pharmacologic disruption: the inhibition of EZH2 (and EZH1) should be sufficient to lead to a decrease of H3K27me3, since no other enzyme is able to install this mark.

Over the past decade, a large number of studies have demonstrated that EZH2 is a highly promising target for epigenetic therapy. Firstly, EZH2 was found to be overexpressed in numerous tumour types including breast and prostate cancer, where its level of expression has been linked to the aggressiveness of the disease and poor patient outcome (See also Table 1, Chapter I).[57,58,196,197] Furthermore, genetic disruption of EZH2 (siRNA) was found in a number of independent studies to result in a reduction of cancer cell growth.[39,198-200]

Secondly, a number of mutations in cancer have recently been identified that affect levels of the H3K27me3 mark. Inactivating mutations of the H3K27 demethylase UTX were found at significant levels in various cancer types.[201,202] The only other H3K27me3 demethylase known, JMDJ3, was found to be expressed at low

level in cancer cells.[203] Inactivation of H3K27 demethylases would result in higher level of H3K27me3. More recently, mutations in EZH2 have also been discovered. Recurrent mutations of tyrosine 641 of EZH2 were found in 7% and 22% of follicular lymphomas and diffuse large B-cell lymphoma respectively.[204] While this was first believed to be an inactivating mutation, further studies uncovered an unprecedented coordinated activity of wild type and mutant EZH2. It was shown that the mutant enzyme displayed poor ability to methylate an unmethylated lysine residue (reaction from H3K9 to H3K9me) but enhanced catalytic efficiency for the subsequent methylation reactions (H3K9me to me2 and then me3).[205] The mutant enzyme thus relies on the presence of wild-type EZH2 to perform the first methylation step. A similar mutation of EZH2 (A677G) in lymphoma cell lines was found to lead to distinct substrate specificity.[206] This EZH2 mutant demonstrated nearly equal efficiency for all three substrates (H3K27, H3K27me1 or H3K27me2). An additional point mutation in EZH2 (A687V) was also recently found to alter the pattern of substrate efficiency and to greatly favour methylation of the H3K27me1 state over other H3K27 forms.[207] It has not yet been confirmed if this mutation would also lead to H3K27 hyper-*tri*methylation (*via* a potential coupling with wild-type EZH2 activity). Nonetheless, both mutations (Tyr641 or A677) lead to increased level of H3K27me3 in lymphoma cell lines, which further implicates EZH2 as an oncogene and strongly suggests that H3K27me3 is implicated in the aberrant silencing of tumour suppressor genes in lymphoma.

However, a number of studies have reached contrasting conclusions, highlighting the potential role of EZH2 as a tumour suppressor gene in malignant myeloid diseases (*i.e.* disorders of haematopoietic stem cells or progenitors of red blood cells, granulocytes, and platelets, such as myelodysplastic syndromes (MDSs), myeloproliferative disorder (MPD), and acute myeloid leukaemia (AML)).[208] Three important studies reported that deletions or inactivation of the EZH2 gene are common in MDSs and in certain MPDs. The loss of the methyltransferase activity has also been demonstrated in some cases.[209-211] Furthermore, inactivating mutations of EZH2 in MDSs and some types of MPD have been shown to predict poor survival.[212,213] Likewise, inactivating mutations of EZH2 concurrent with reduced level of the H3K27me3 mark have also been found to be sufficient to cause T-acute lymphoblastic leukaemia (T-ALL).[214] Subsequently, mutations in PRC2 complex (EZH2, EED or SUZ12) were discovered to be very frequent in T-ALL, and

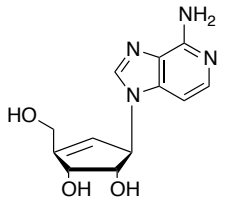
alterations were found at even higher levels in a pediatric subtype.[215,216] These studies clearly established EZH2/PRC2 as an important tumour suppressor gene in T-ALL as well as in a range of malignant myeloid diseases and that the loss of EZH2 activity (*i.e.* H3K27 trimethylation) might lead to the de-repression of oncogenic pathways.

Taken together, these results show the crucial importance of H3K27 methylation in a wide range of cancers. On top of this, the histone residue itself has been found to be recurrently mutated in aggressive brain tumours, including a lysine to methionine change found in 78% of the cases in specific pediatric gliomas.[217,218] In a recent study, it was reported that histones bearing this K27M mutation were able to inhibit EZH2 and lead to lower H3K27me3 levels in cells.[219]

It is therefore of the highest importance to further study the regulation of H3K27 methylation in cancer. The development of modulators of this mark holds promise for the treatment of a variety of cancer types. However, as described above, changes in H3K27 methylation level may have contrasting effects depending on the cancer in question, and it will thus be crucial to prove the validity of targeting the enzyme installing the mark (EZH2 or EZH1) or removing it (UTX, JMJD3) in a tissue specific manner. Likewise, the success of clinical trials will most likely depend upon the identification of reliable biomarkers (*e.g.* mutations) and subsequent patient stratification.

This project proposed to address such need. In particular, our main objective is to identify molecules that reduce H3K27 trimethylation and concomitantly lead to re-expression of silenced EZH2 target genes. As mentioned above, the re-activation of pathways aberrantly silenced by EZH2 is expected to induce apoptosis of cancer cells.

At the time this project was initiated, no direct EZH2 inhibitor had been publicly disclosed. One compound, 3-Deazaneplanocin A (DZNep, Figure 67) had been reported to selectively inhibit H3K27me3 and H4K20me3 in cells as well as to induce apoptosis in cancer cells by depleting cellular level of PRC2 proteins (at least in part by proteasome-mediated protein degradation), while not affecting their mRNA level.[220]



DZNep

Figure 67 Structure of DZNep

However, DZNep was previously identified as a SAH hydrolase inhibitor and would thus be expected to indirectly inhibit (through increasing concentration of SAH) all of the SAM-dependent enzymes (HKMT, PRMT, DNMT, and various non-epigenetic methyltransferases).[221,222] Jones and co-workers subsequently demonstrated that DZNep globally inhibited histone methylation and does not display any HKMT selectivity.[223] While these studies further confirmed that disruption of the PRC2 complex effectively induced apoptosis in cancer cells, this compound was not a suitable starting point for our project in light of its poor selectivity and obscure mechanism of action.

Rational design of SAM analogues that display EZH2 selectivity would be one potential option. Although SAM binding pockets are highly conserved between HKMTs,[63] careful crystallographic studies can guide the design of inhibitors and specificity could be achieved, as recently illustrated by the development of a highly specific DOT1L inhibitor (Figure 6, Chapter I).[73] However, in the case of EZH2, this approach was impeded by the absence of a crystal structure to lead the medicinal chemistry effort in a rational way.

We instead focused our interest on a small molecule inhibitor reported to inhibit another HKMT, the H3K9 dimethylase G9a. BIX-01294 (**2**, Figure 7) was identified through HTS (125,000 compounds) to inhibit the HKMT G9a.[75] The specificity of the compound was assessed against a number of histone methyltransferases, including an arginine methyltransferase (PRMT1) and additional lysine methyltransferases specific for other residues such as SET7/9 (H3K4me) or the hyper active SUV39H1 (H320R) mutant, but excluding EZH2. BIX-01294 specifically inhibited G9a (IC_{50} = 1.7 μM) and to a lesser extent the closely related GLP (G9a-like protein) (IC_{50} = 38 μM). No inhibition of the other histone methyltransferases was observed at the tested concentrations (up to 45 μM). BIX-01294 was shown to act in an uncompetitive manner with SAM, which

suggested that it could only bind to the enzyme-cofactor complex. Furthermore, when cells were treated with BIX-01294, a reduction of 20% of H3K9me2 mark was observed by mass spectrometry, together with a comparable increase of the unmethylated state. The other methylation states of H3K9 (H3K9me and H3K9me3) were not significantly altered, nor were the other lysine residues (H3K27, H3K36, H4K20).[75] Analyses of microarrays of G9a null mouse ES cells compare to wild-type identified a number of G9a target genes and six of those were investigated by chromatin immunoprecipitation (ChIP) experiments. Upon treatment with BIX-01294, ChIP confirmed a pronounced decrease of the H3K9me2 mark at their promoter regions.[75] However, the reduction of the repressive chromatin mark H3K9me2 only resulted in increased gene expression for two target genes out of the six tested.[75] Finally, it was shown that the level of H3K9me2 was fully restored after washing of the inhibitor. BIX-01294 was thus able to transiently modify the H3K9me2 repressive mark in cells but its effect on gene re-expression was limited.

Two years later, the crystal structure of the catalytic SET domain of GLP in complex with BIX-01294 and SAH was reported.[65] GLP was preferred since a crystal structure of GLP in complex with SAH and an H3 peptide substrate was already available. Moreover, the two enzymes are highly similar: GLP shares around 80% sequence identity with G9a in their SET domains. Furthermore, when assayed under different reaction conditions, BIX-01294 inhibited GLP better than G9a (IC_{50} = 1.9 μM for G9a, IC_{50} = 0.7 μM for GLP).[65] Importantly, it was shown that the inhibitor bound in the histone peptide groove thereby preventing the substrate from binding (Figure 68).

Based on the structural information provided by the co-crystal structure, SAR studies have been performed around the BIX-01294 scaffold. A number of analogues with increasing potency towards G9a (and GLP) and physicochemical properties have been reported. Figure 69 summarises the achievement of different groups.

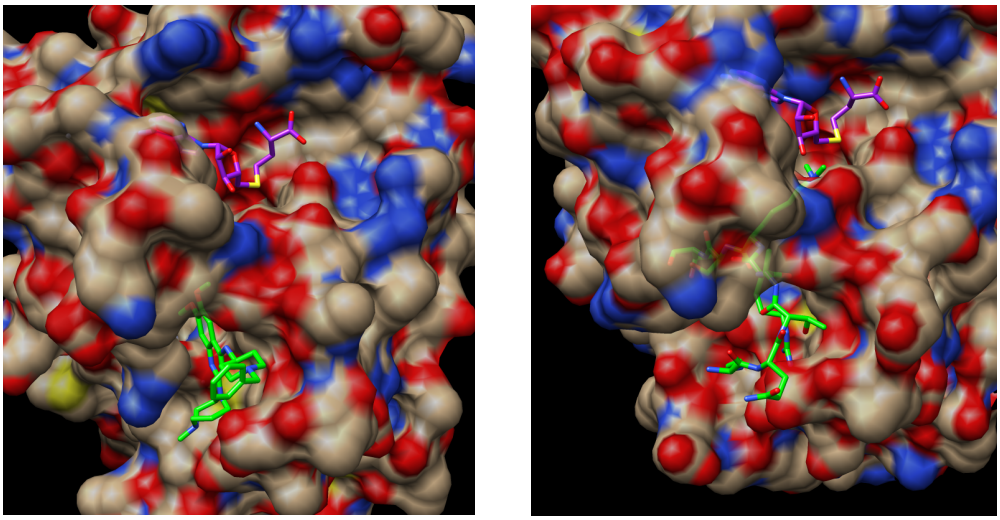


Figure 68 Ternary complex of GLP with SAH and BIX-01294 (**2**) (PDB: 3FPD, left) or H3 peptide (PDB:2RFI, right). BIX-01294 and the H3 peptide are shown with green carbons and SAH with purple carbons for clarity.

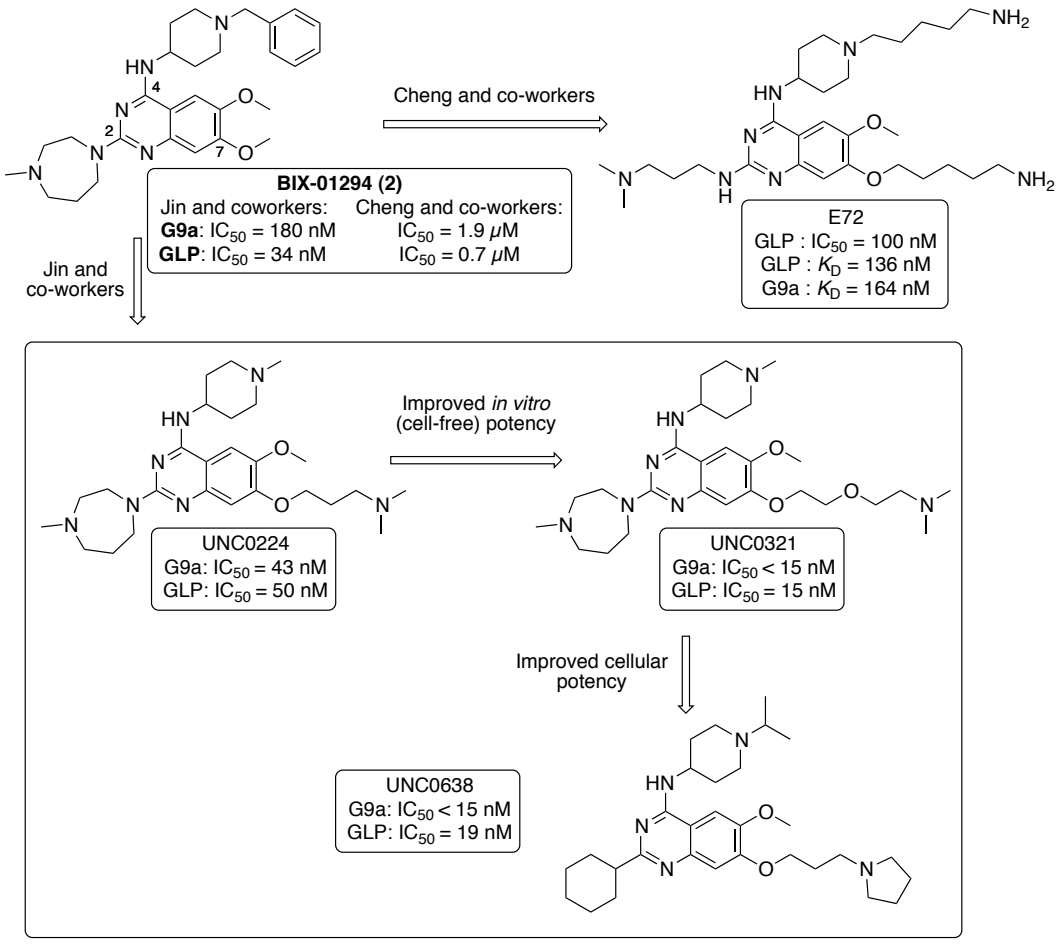


Figure 69 SAR development around BIX-01294 scaffold against G9a and GLP[224-228]

Cheng and co-workers designed a number of analogues bearing one or two lysine mimics (*e.g.* E72, Figure 69).[224] The lysine mimic was shown to extend into the lysine-binding channel and rationalised the observed increased affinity of the

compound for the enzyme. Jin and co-workers reported an extensive SAR study summarised in Figure 69.[225-228] Efforts to improve the cellular potency of this class of compounds through increasing lipophilicity led to the discovery of a couple of optimised analogues, including the highly potent UNC0638 (Figure 69).[225,228]

From the in-depth work of Jin and co-workers, the BIX-01294 scaffold proved to be amenable for the rapid synthesis of a large number of analogues to generate a suitable SAR study. Upon careful manipulation of the physicochemical properties of the analogues it was possible to generate cell-penetrant, highly active tool compounds for cell-based studies.

Whilst the BIX-01294 studies focused on G9a (and GLP) inhibition, there were a number of reasons we felt we could use this molecule as a starting point for the reversal of H3K27-mediated gene silencing.

Although the main activity of G9a is the dimethylation of H3K9me2, G9a has also been shown to contribute to H3K27 methylation.[195] Furthermore, sequence alignment clearly showed that the H3 sequence around lysine 9 and lysine 27 are appreciably similar (Figure 70),[24] and therefore the binding pockets accommodating the respective histone region in G9a and EZH2 would likely be similar. We therefore anticipated that upon derivatisation of the BIX-01294 scaffold, it might be possible to expand the biological activities of the parent compound and gain EZH2 inhibitory activity.

H3K9	KQTARKSTGGKA
H3K27	TKAARKSAPATG

Figure 70 Sequence of histone peptides[24]

It is worth stating that H3K9 dimethylation, mediated by G9a, is a transcriptionally repressive mark. In light of the epigenetic crosstalk mentioned above, it may be necessary to develop agent targeting several histone-modifying enzymes to eventually reach a more permissive chromatin state. Consequently, molecules leading to a reduction of both H3K9 and H3K27 methylation may be more efficient at killing cancer cells than compounds targeting a single enzyme. Dual or polypharmacology might be a desirable feature in the field of epigenetics.

In late 2012, a number of pharmaceutical companies disclosed the structure of their EZH2 inhibitors (Figure 71). Epizyme developed the first highly potent EZH2-selective inhibitor, EPZ005687, which displays a K_i of 24 nM and a selectivity greater than 500-fold against 15 methyltransferases (50-fold against EZH1).[77] The compound reduced H3K27 trimethylation in cells without altering the methylation state at other histone residues. Furthermore, it killed lymphoma cells bearing an EZH2 mutation (Tyr641 or Ala677) selectively, while having little effect on cell lines containing wild-type EZH2. These data suggested that EZH2 mutant-bearing lymphoma cells critically depend on EZH2 activity for proliferation and survival.[77] Very recently, Epizyme reported another EZH2 inhibitor, EPZ-6438, which showed dose-dependent regression of malignant rhabdoid tumours (rare and highly malignant childhood neoplasm, originally of the kidney).[229] Shortly after, GSK, Novartis and Jin and co-workers disclosed the structure of their EZH2 inhibitors (GSK343/126, EI1 and UNC1999 respectively, Figure 71).[78-81] Highly similar to EPZ005687, the molecules contain an indazole or indole core bearing a pyridone side chain connected via an amide linkage at position 4 (indole numbering). Side chains at position 6 principally differentiate the molecules. All compounds display nanomolar potency and high selectivity in biochemical assays and were competitive with SAM. Jin and co-workers reported that UNC1999 was inhibiting both EZH2 and EZH1, which was expected to inhibit further H3K27 methylation compared with EZH2 specific inhibitors. In addition, their compound was orally bioavailable and therefore could be employed conveniently in animal studies. Importantly, all the compounds reported were able to kill lymphoma cell harbouring the Tyr641 point mutation but had little effect on cells with wild-type EZH2.

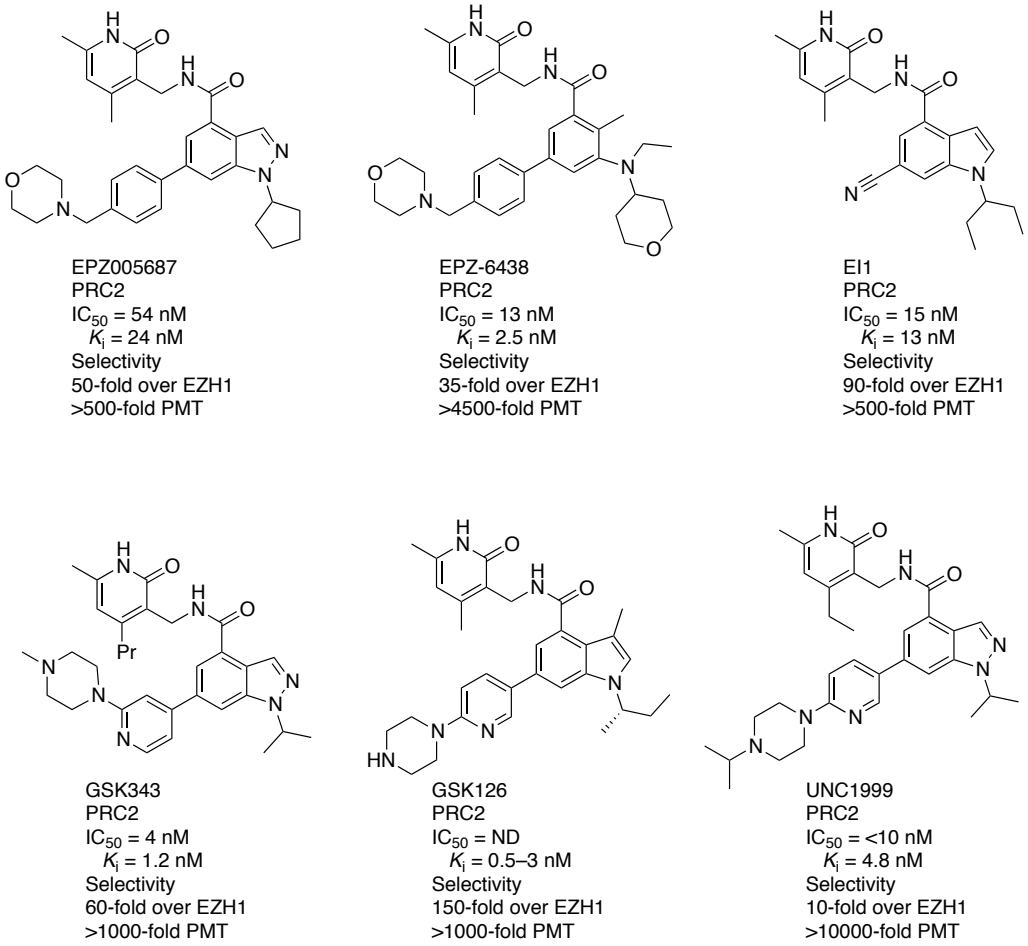


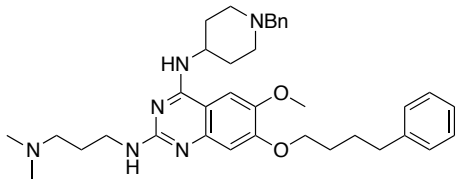
Figure 71 Current specific inhibitors of EZH2

In light of the fact that overexpression of wild-type EZH2 is observed in a variety of cancers, there is still a need for small molecules able to kill cancer cells harbouring high levels of wild-type EZH2. Indeed, the discovery of potent inhibitors displaying selectivity for cancer cells with mutated EZH2 was a great advance, but the use of such compounds will be limited to a very defined patient group, (*e.g.* the Tyr641 mutation is only present in ~ 7% of follicular lymphomas and 22% of diffuse large B-cell lymphoma).

This chapter will describe our approach for the identification and development of small molecules leading to the reversal of wild-type EZH2-mediated gene silencing in breast cancer cells, using BIX-01294 (**2**) as a chemical start point.

The biological evaluation of preliminary BIX-01294 analogues was performed by collaborators in Prof. Robert Brown's lab (Imperial College London). The initial analogue library used consisted of ~30 analogues provided by our collaborators Thota

Ganesh and James Snyder at Emory university, as well as compounds made by co-workers in house. The compounds were first tested in an *in vitro* cell-free assay kit against EZH2 (using the PRC2 complex, specific for H3K27) and SUV39H1 (specific for H3K9), as described in Chapter II. None of the BIX-01294 analogues tested led to a reproducible decrease in EZH2 activity. In general, a flat response to many of the analogues was observed. Unexpectedly, some compounds led to an apparent increased activity for EZH2 and SUV39H1. A few compounds showed some inhibitory activity toward SUV39H1, and compound **56** (Figure 72) exhibited the most significant effect (65% inhibition at 20 μM).



56

Figure 72 Structure of analogue **56**

The mechanism leading to the observed increased activity was unclear but we felt it was most likely an artefact of the assay setting, questioning the reliability of the commercial assay kit for this compound series. As an alternative assay, Prof. Brown and co-workers developed a cell-based assay (MDA-MB-231 breast cancer line) designed to monitor the ability of compounds to reverse H3K27me3 dependent gene silencing using a real-time reverse transcription polymerase chain reaction (RT-PCR) to examine mRNA levels of candidate genes.[230]

Result and discussion

Synthesis of BIX-01294 analogues

In order to initiate the project, a focused library of analogues had to be synthesised to expand the biological activities of the BIX-01294 scaffold and hopefully gain EZH2 inhibition (and possibly specificity) as well as to develop a SAR around the quinazoline core. Three positions were to be explored: the nature of the amines at position 2 and 4 as well at the substituent at position *O*-7 (Figure 69). Unfortunately, as mentioned above, in the absence of EZH2 crystal structure, no structural information was available to guide our initial design.

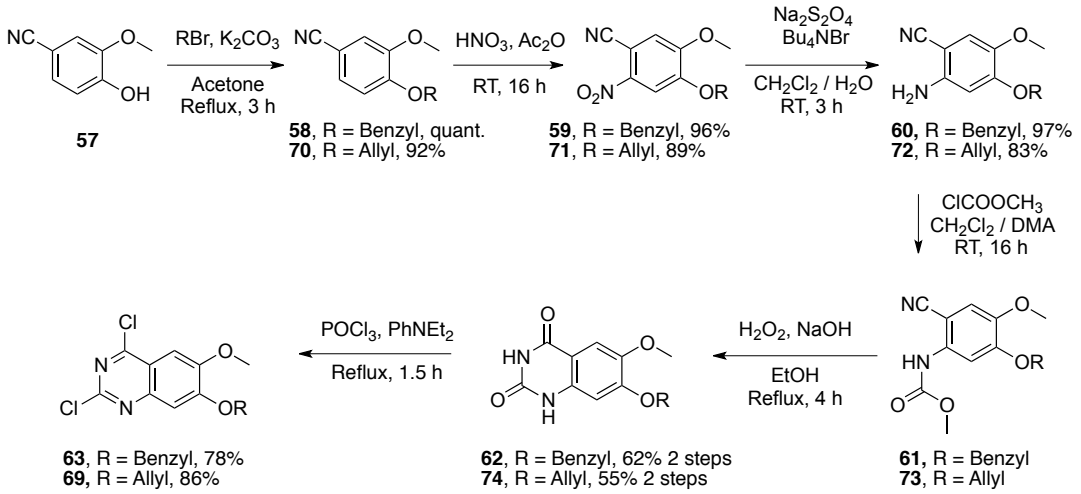
Derivatisation at position *O*-7

Since mainly analogues with a lysine mimic at *O*-7 have been reported (Figure 69), we felt that further exploration of this position using other side chains was warranted. Additionally, in light of the preliminary results obtained for compound **56** and the questioned reliability of the assay, we felt that re-synthesis of this compound in house was desirable to confirm the effect seen.

To access derivatives with various ether side chains at this position, the core of the quinazoline had to be synthesised with a robust protecting group on the hydroxy moiety at position 7. For this purpose, the route developed by Liu *et al.* was adapted to access the quinazoline scaffold with a benzyl protecting group at position 7.[226]

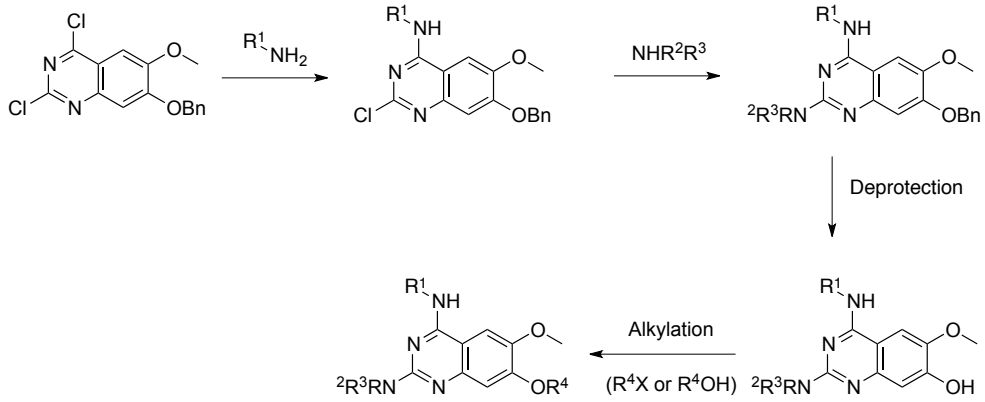
Protection of commercially available phenol **57** (Scheme 35) proceeded in quantitative yield with benzylbromide. Nitration of protected phenol **58** with nitric acid in acetic anhydride afforded nitro compound **59** in excellent yield (96%) as a single regioisomer. Reduction of the nitro moiety of compound **59** was initially performed with iron dust,[226] but an alternative method using sodium dithionite turned out to be more convenient and afforded aniline **60** as a cleaner product by a simple acid/base extraction in excellent yield (97%).[231] Quinazolinedione **62** was obtained in two steps from aniline **60** by formation of a methyl carbamate intermediate **61** and oxidative hydration of the nitrile group followed by spontaneous cyclisation in good

yield (62% over two steps). Chlorination of quinazolinedione **62** with phosphorus oxychloride afforded dichloroquinazoline **63** in excellent yield (97%).



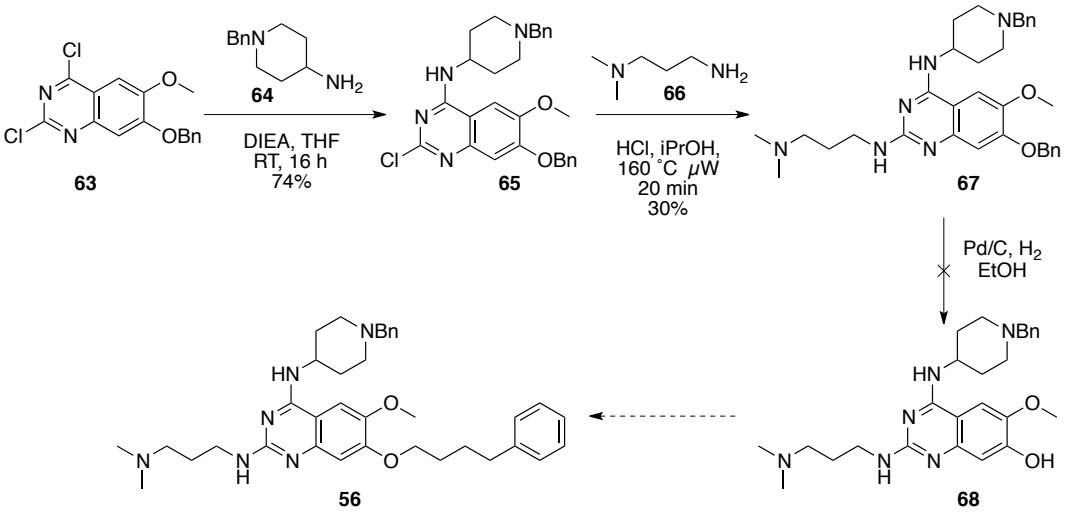
Scheme 35

From the common dichloro intermediate **63**, fully substituted compounds were four steps away: substitution at positions 2 and 4 with various amines, deprotection and finally *O*-alkylation under either Mitsunobu conditions or by treatment with alkyl halides (Scheme 36).



Scheme 36

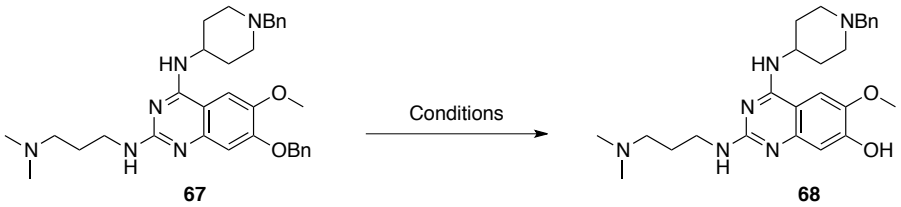
The attempted synthesis of compound **56** is depicted in Scheme 37.



Scheme 37

Treatment of dichloro compound **63** with the corresponding amine **64** under mild conditions resulted in selective substitution of the more reactive chloride, affording compound **65** in good yield (74%). The second displacement with amine **66** required more forcing conditions (microwave irradiation at 160 ˚C) and the di-substituted quinazoline **67** was obtained in low yield (30%), due to difficulties in the purification of this highly polar compound. Unexpectedly, the debenzylation of 2,4-diaminoquinazoline **67** turned out to be quite problematic (Table 9).

Table 9 Debenzylation conditions

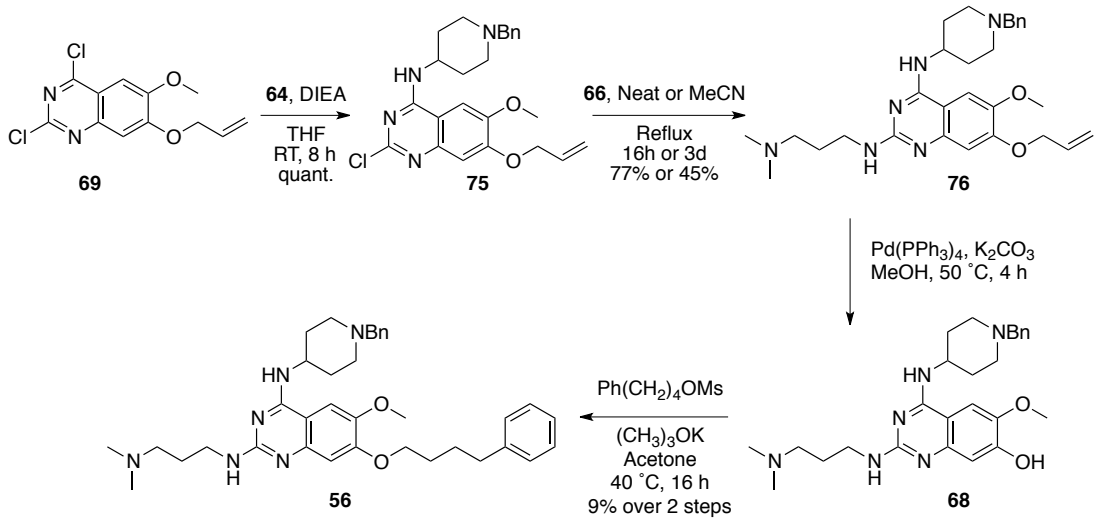


Conditions	Observations
5% Pd/C, H_2, Ethanol, RT[226]	Unidentified products by ^{1}H NMR and MS
5% $PdBaSO_4$, H_2, THF, RT[232]	Unidentified products by ^{1}H NMR and MS
BF_3•Et_2O, EtSH, CH_2Cl_2, RT[233]	Starting Material

Chang *et al.* reported the debenzylation of monochlorinated compound **65** under standard conditions of hydrogenation (palladium on charcoal and hydrogen) and reported a 3:1 ratio of mono and di-debenzylated product, with the *O*-benzyl group being cleaved first. It was therefore assumed that careful monitoring of the reaction progress would allow for the selective cleavage of the *O*-benzyl group of

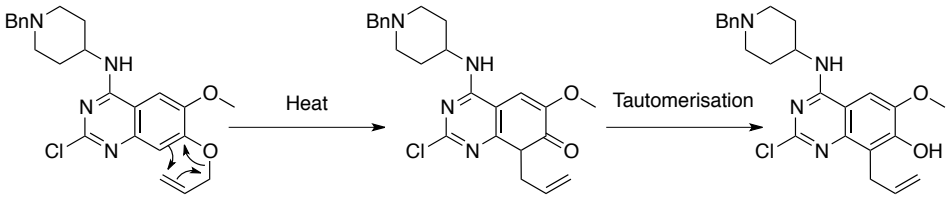
compound **67**. However, similar reaction conditions for the debenzylation of compound **67** not only failed to selectively deprotect compound **68,** but resulted in a complex mixture of products in which the main product could not be identified. A milder reductive method using a poisoned catalyst ($Pd/BaSO_4$) was attempted but afforded results similar to the first hydrogenation.[232] Alternatively, a Lewis acid mediated method (BF_3•Et_2O and EtSH)[233] was explored but no reaction was observed. This difficult debenzylation lead to the consumption of all the benzyl protected compound **67**.

Consequently, it was decided that an alternative protecting group would be used. The allyl group was chosen for this purpose as numerous methods are known for its cleavage, and the *N*-benzyl group should not be affected under such deprotection conditions. Allyl-protected dichloroquinazoline **69** was thus prepared following the same route with comparable yields (Scheme 35).



Scheme 38

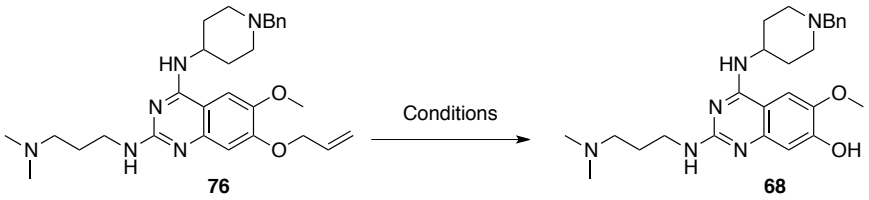
Mono-displacement of the most reactive chloride with amine **64** afforded substituted chloroquinazoline **75** in quantitative yield (Scheme 38). For the second chloro displacement however, microwave irradiation resulted in a Claisen rearrangement (Scheme 39). The two main products obtained by this method were the rearranged mono- and di-substituted quinazoline.



Scheme 39

Milder conditions (in refluxing acetonitrile) afforded the expected product **76** in very good yield (80%), without any rearrangement, but a long reaction time was required (3 days). Alternatively, refluxing compound **75** in neat amine **66** (b.p. 133 ˚C) afforded the disubstituted quinazoline **76** in a shorter reaction time but with lower yield (45%) due to the formation of a minor amount of rearranged products. Unfortunately, deprotection of the allyl group of **76** also proved quite challenging. Several standard conditions were attempted (Table 10).

Table 10 Conditions for cleavage of the allyl protecting group of **76**

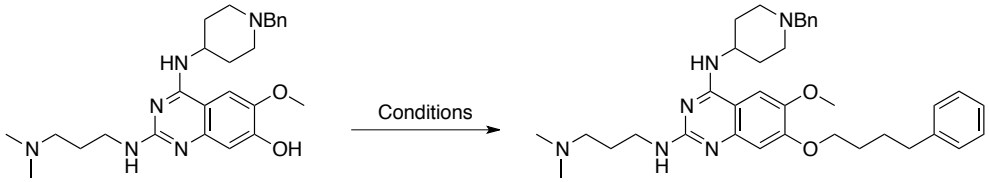


Conditions	Observations
$RhCl(PPh_3)_3$, DABCO, Ethanol reflux[234]	Starting material
$Pd(PPh_3)_4$, $LiBH_4$, THF, RT[235]	Starting material
$Pd(PPh_3)_4$, Et_3SiH, AcOH, Toluene, RT[236]	No product detectable by NMR, traces in LC-MS
$NaBH_4$, I_2, THF, RT[237]	Starting material
$Pd(PPh_3)_4$, Morpholine, THF, RT[238]	Product detected by NMR / LC-MS, not isolated
$Pd(PPh_3)_4$, K_2CO_3, MeOH, 50˚C[239]	Product isolated and confirmed by NMR and LC-MS

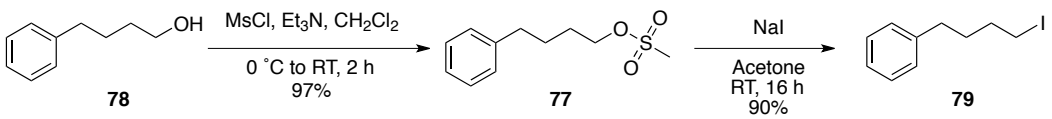
Palladium and rhodium complexes are believed to isomerise the double bond and lead to the formation of a vinyl ether *via* a π-allyl species, which can then be cleaved under various conditions.[234,239,240] Vutukuri *et al.* developed a mild deprotection method involving potassium carbonate in methanol in the presence of palladium which was shown to be selective for the deprotection of aryl allyl ethers in

the presence of alkyl allyl ethers.[239] The authors hypothesised that the carbonate anion acts as the nucleophile to react with the Pd-π-allyl species, although this has not been demonstrated. This methodology gave the most promising results for the deallylation of compound **76**. However purification of free phenol **68** was not possible due to its high polarity and low solubility (^{1}H NMR and HRMS were obtained on the crude material to confirm the identity of the product). A number of conditions were then attempted to alkylate crude phenol **68** (Table 11). Mesylate **77**, prepared from alcohol **78** in high yield (97%, Scheme 40), was first used as electrophile. The use of a mild base (Cs_2CO_3) resulted in traces amount of product and increasing the strength of the base ($(CH_3)_3OK$) did not lead to a significant increase in yield (from 2 to 9%). Mitsunobu conditions were attempted as an alternative method but no reaction was observed. Finally, an attempt was made to change the electrophile for iodide **79** (prepared from mesylate **77**, Scheme 40), but this did not lead to the desired product.

Table 11 Conditions for alkylation of phenol **68**



Electrophile	Conditions	Observations
Mesylate **77**	Cs_2CO_3, DMF, 80 ˚C	2% yield
Mesylate **77**	$(CH_3)_3OK$, Acetone, 40 ˚C	9% yield
Alcohol **78**	DIAD, PPh_3, THF, RT	No reaction
Iodide **79**	K_2CO_3, Acetone, RT, then $(CH_3)_3OK$, 60 ˚C	No reaction (with K_2CO_3) then mixtures of unidentified products

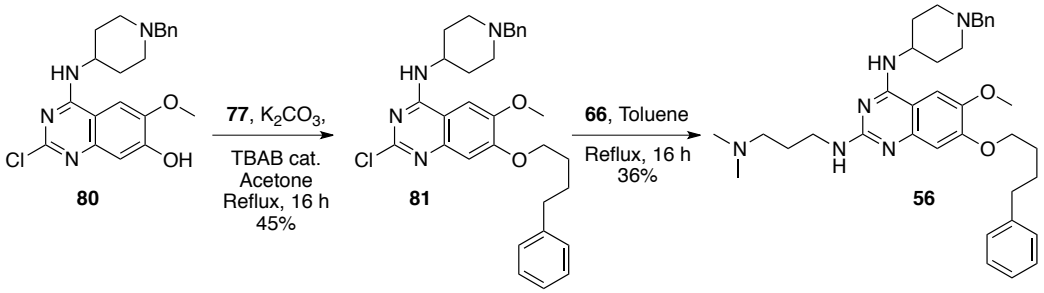


Scheme 40

None of the conditions attempted afforded the final product **56** in an acceptable yield. It is still unclear if the low yields obtained are due to the poor

nucleophilicity or purity of phenol **68** or to the instability of the different electrophiles used (to competing elimination).

It was therefore decided that an alternative route was needed to efficiently prepare a large number of derivatives. As it was likely that most of the issues encountered in the deprotection – alkylation steps were a results of the high polarity and rich electron density of the compounds we were dealing with, we envisaged that deprotection of the quinazoline core at an earlier stage might result in better yield. Chang *et al.* deprotected the benzyl on the quinazoline core before the introduction of the amine at position 2. Accordingly, a postdoctoral associate in the group, Dr. Thomas Moss, performed the benzylic ether deprotection of compound **65** with palladium on charcoal and could isolate phenol **80** in good yield (79%), which confirmed the significant influence of the second amino side chain. A sample of free phenol was supplied to complete the synthesis of butylphenyl analogue **56** (Scheme 41). Phenol **80** was alkylated by treatment with mesylate **77** to afford compound **81** in medium yield (44%) and displacement of the second chlorine with the corresponding amine **66** led to final compound **56** in 36% yield. Performing the deprotection and alkylation at position 7 at an earlier stage allowed the synthesis to be completed with higher yields and the intermediates were easier to handle and purify due to their limited polarity.

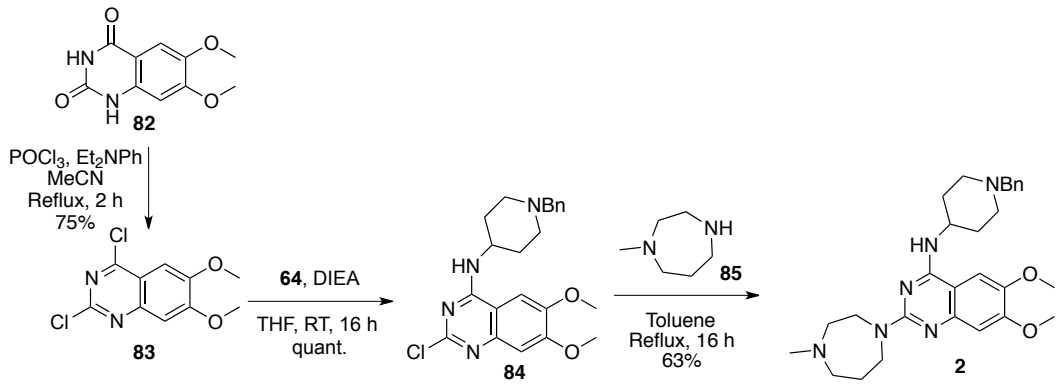


Scheme 41

A number of derivatives have been prepared by co-workers in the Fuchter group using this route (Figure 73). Derivatisation at position 7 was performed either by alkylation with corresponding halides or under Mitsunobu conditions, depending on the availability of the starting materials.

Dimethoxyquinazoline analogues

Analogues with methoxy groups at position 6 and 7 were prepared in three steps from commercially available bis-methylated dicarbonyl compound **82** (Scheme 42, applied to the synthesis of BIX-01294 (**2**)). The dichloroquinazoline core **83** was obtained by treatment with $POCl_3$ and two subsequent displacements with corresponding amines afforded the final compounds. Dr. Tomas Moss prepared large quantities of dichloroquinazoline **83** and kindly supplied a sample for the synthesis of BIX-01294 (**2**) which was to be used as a control molecule (Scheme 42). The first displacement with amine **64** afforded compound **84** in quantitative yield. The second displacement with the corresponding amine **85** was performed in toluene at reflux and afforded BIX-01294 (**2**) in medium yield (63%) for biological testing.



Scheme 42

A large number of derivatives have been synthesised by co-workers in the Fuchter group as well as at a contract synthesis company using this route, in order to access a large number of derivatives for SAR development. Figure 73 summarises the structure of the analogues synthesised in the Fuchter group for biological evaluation.

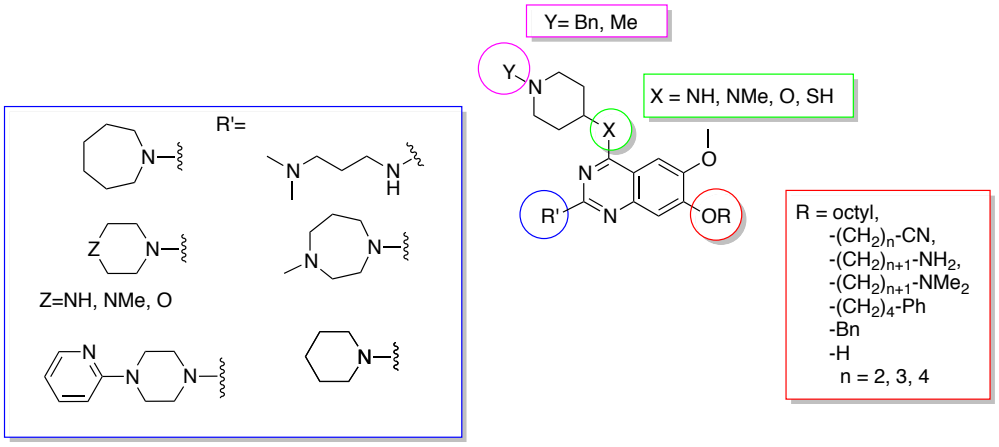


Figure 73 Analogues prepared in the Fuchter group based on BIX-01294 scaffold for SAR exploration

Biological evaluation

Biochemical assay against EZH2 and SUV39H1

The library of new analogues was tested in the biochemical assay kit as described in Chapter II, against EZH2 and SUV39H1. However, the data obtained frequently contained significant well-to-well variability. Furthermore, the batch of newly synthesised analogue **56** was assessed but its inhibitory activity against SUV39H1 could not be reproduced (this effect was observed for a number of compounds).

While the origin of the discrepancy of the data remains obscure, the commercial assay kit was therefore considered non reliable for our needs and its use was abandoned.

Cell-based assay

The focused library of BIX-01294 analogues was screened in order to identify compounds which induce the reactivation of PRC2 silenced genes. Cells were treated with compounds at a concentration of 10 μM and mRNA levels of two EZH2 target genes (KRT17 and FBXO32) were monitored by real-time RT-PCR.

Three compounds were found to up-regulate KRT17 and FBXO32 RNA levels, and were identified as hit compounds (Figure 74). The RT-PCR data for these compounds is displayed in Table 12 (Entry 1–3).

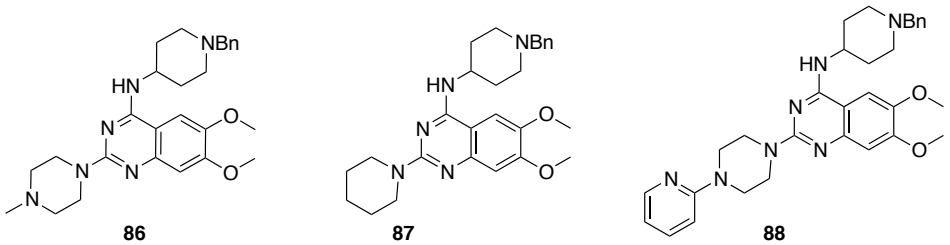


Figure 74 Structure of the hits identified in the cell-based assay

Compound **56** was also assessed in this assay. Since cell death was observed at 10 and 5 μM, the concentration was dropped to 1 μM for RT-PCR measurement. It is of particular importance in such cell-based assay to monitor cell death as it can lead to changes in histone marks and can thus interfere with the assay reading (See discussion in Chapter II). As seen in Table 12 (Entry 5), no effect on gene expression was observed at 1 μM.

Table 12 Effects on gene expression of compounds in cell-based screen

Entry	Compound	KRT17	FBXO32
1	**86**	4.1	3.7
2	**87**	4.3	29.4
3	**88**	7.0	33.3
4	**2**	1.1	3.3
5	**56***	1.2	1.1

Average qRT-PCR data for single concentration (10 μM, except *, 1 μM) dose treated in triplicate, for 48 h. RNA levels for target genes are normalised against the housekeeping gene GAPDH and shown is the fold increase compared to the untreated sample.

The G9a/GLP inhibitor, BIX-01294 (**2**) was also tested in this assay (Entry 4). Although closely related to our hits from a chemical structure perspective, such G9a/GLP selective compound displayed a different effect on gene expression profile and did not significantly up-regulate KRT17, but did up-regulate FBXO32. Within our focused library, the majority of compounds had little or no effect on both KRT17 and FBXO32.

AlphaLISA assay against G9a

In light of the obvious similarity of the structures of the identified hits with that of known G9a inhibitors, we sought to assess whether our compounds retained some G9a inhibitory activity. The hit compounds were tested in the G9a AlphaLISA assay described in Chapter II, and BIX-01294 (**2**) was used as control. As shown in Table 13, the hit compounds exhibited some moderate G9a activity, ~ 5 to 10 μM.

Table 13 G9a inhibitory activity of the hits and negative compounds

	86	**87**	**88**	**2**
Mean pIC_{50} (SD)	-5.244 (0.038)	-5.017 (0.096)	-5.050 (0.067)	-5.775 (0.033)
IC_{50} (μM)	5.70	9.62	8.91	1.68

pIC_{50} are average of three independent experiments run as duplicate (standard deviation). Assay conditions: G9a at 1 nM, H3 peptide at 100 nM, SAM at 15 μM, 30 min reaction time.

Conclusions

The synthetic route towards BIX-01294 analogues allowing late-stage derivatisation at position *O*-7 has been investigated. While the synthesis of compound **56** presented some challenges, an alternative route was developed to efficiently access analogues with diverse side-chains at position 7 (Figure 73). The library of compounds generated within the Fuchter group was initially evaluated in biochemical assays, however the results were found to be irreproducible. Evaluation of the compounds in a cell-based assay identified three hits able to reverse EZH2-mediated gene silencing.

Assessment of the three hits in the G9a biochemical assay described in Chapter II revealed that the hits retained some modest G9a inhibitory activity (5 to 10 μM). We can therefore hypothesise that our compounds lead to gene re-expression by targeting a panel of HKMTs with some selectivity toward the ones involved in repressive chromatin remodeling (G9a and EZH2). As mentioned above, the substrate binding site of G9a and EZH2 may be significantly similar and our initial data is consistent with the dual inhibition of G9a and EZH2, possibly with binding of the compound in the peptide binding groove. Further work is on-going in the Fuchter and Brown groups to test this hypothesis. For example, mass spectrometry experiments are on-going to monitor global changes in histone marks and the dynamics of the H3K27 trimethylation mark upon compound treatment. Additionally, probes for affinity based protein profiling have been designed based on our hit compounds in order to pull-out the proteins involved in hit binding.

Conclusions

The importance of epigenetic regulation is now well established and it is crucial to develop tools to study in depth these important mechanisms. The work presented in this thesis initially aimed at developing the SAR around two histone lysine methyltranseferase inhibitors previously reported: chaetocin (**1**) and BIX-01294 (**2**).

Our work on chaetocin shed light onto the poor characteristics of this compound as a specific HKMT inhibitor. The absolute requirement of the reactive disulfide bridge led to the establishment of a very steep SAR, therefore impeding the progression of such compound through the next steps of drug development. This also stressed the importance of the proper validation of hits discovered through high throughput screen, tedious but fundamental task often underrated especially in academia. We also established the mode of action of chaetocin: biochemical and mass spectrometric data are concordant to suggest that the ETP core reacts with available cysteines to form covalent adducts which ultimately leads to protein inactivation. Although covalent inhibitors can be interesting in drug discovery, we do not believe that the reactivity of ETP compounds can be tuned in order to achieve selectivity and as such, we do not envision that chaetocin, or any ETP-containing compounds, could be used as specific HKMT inhibitors.

The preparation of the monosulfide analogue of chaetocin (**3**) took us a bit off-track of our initial medicinal chemistry effort. The establishment of the stereochemistry of monosulfide **3** was required, however. In light of the inconsistencies in the literature, we proposed an alternative approach relying on the comparison of experimental and simulated data for three chiropotical properties: optical rotation, ECD and VCD. The method was also used, in parallel with chemical derivatisation, to correct the assignment of two additional compounds, and led us to conclude that all ETPs are desulfurised by triphenylphosphine with retention of configuration at the carbon bridgehead atoms. This approach is an additional and potent tool for absolute configuration determination but is relatively recent in the literature. We foresee that such method is bound to expand in the near future due to the ease with which it can be performed (should the spectrometer be available) and the increasing computational power leading to ever more accurate and reliable simulations.

The project presented in the last chapter of this thesis aimed at expanding the biological activities of the known G9a inhibitor, BIX-01294 (**2**), and especially to

achieve EZH2 inhibition. Chemical derivatisation of the original lead was anticipated to lead to this additional activity, but the absence of crystallographic data precluded rational design. The cell-free assay available at the time turned out to be unreliable and a more sophisticated cell-based assay was therefore developed by our collaborators and used as the primary assay. Three analogues have been identified to re-express EZH2 target genes and were also found to retain some G9a inhibitory activity. We suspect that our hit compounds act as dual inhibitor and a number of additional studies are ongoing in our labs to validate our assumption and increase the potency of our compounds. Undeniably, EZH2 represents a very promising epigenetic target as illustrated by the number of pharmaceutical companies working on this HKMT and the proof of concept should soon be available since clinical trials of EZH2 inhibitors are now commencing.

Experimental Details

Compound synthesis

Materials and methods

All reagents and solvents were supplied from commercial sources, and used as supplied unless otherwise indicated. Reactions requiring anhydrous conditions were conducted in heat-dried glassware (heat gun), under an inert atmosphere (argon), and using anhydrous solvents. CH_2Cl_2, toluene, triethylamine, pyridine and MeOH were distilled over CaH_2. THF and Et_2O were distilled over Na/benzophenone. Toluene was distilled over Na. *N*,*N*-dimethylformamide, acetonitrile and ethanol were obtained as dry or anhydrous from Aldrich Chemical Company, Other solvents and all reagents were obtained from commercial suppliers (Fluka; Aldrich Chemical Company; Lancaster Chemicals) and were used as obtained if purity was >98%.

Reactions were performed under an atmosphere of N_2 unless otherwise stated. All organic extracts were dried over magnesium sulfate.

Specific rotation $[\alpha]_D^T$ was recorded on a Perkin-Elmer 241 polarimeter at 589 nm (Na D-line) with a path length of 0.5 dm. Concentrations (*c*) are quoted in g/100 mL and specific rotations are quoted in units of $deg.dm^{-1}.cm^3.g^{-1}$ at the indicated temperature (in ˚C).

Infrared spectra were recorded neat on Perkin-Elmer Paragon 1000 Fourier transform spectrometer with automated background subtraction. Only selected absorptions are reported and are of strong or medium strength unless stated otherwise and given in wavenumbers (cm^{-1}).

^1H NMR spectra were recorded on a Bruker DRX-400 spectrometer operating at 400 MHz. ^{13}C NMR spectra were recorded on a Bruker DRX-400 spectrometer operating at 100 MHz. Chemical shifts (δ) are quoted in units of parts per million (ppm) downfield from tetramethylsilane and are referenced to a residual solvent peak. ($CDCl_3$ (δ_H: 7.26, δ_C: 77.0)). Coupling constants (*J*) are quoted in units of Hertz (Hz). The following abbreviations are used within ^1H NMR analysis: s = singlet, d = doublet, t = triplet, q = quartet, quint = quintuplet, sept = septuplet, m = multiplet. Spectra recorded at 400 MHz (^1H NMR) and 100 MHz (^{13}C NMR) were carried out by the Imperial College London Department of Chemistry NMR Service. Low and

high resolution mass spectrometry (EI, CI, FAB) were recorded by the Imperial College London Department of Chemistry Mass Spectrometry Service using a Micromass Platform II and Micromass AutoSpec-Q spectrometer.

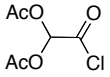
All flash column chromatography was carried out on BDH silica gel 60, particle size 0.040 – 0.063 mm unless otherwise stated.

Thin layer chromatography (TLC) was performed on pre-coated aluminium backed (Merck Kieselgel 60 F_{254}), and visualised with ultraviolet light (254 nm) or potassium permanganate ($KMnO_4$), vanillin, ninhydrin or phosphomolybdic acid (PMA) stains as deemed appropriate.

Synthetic procedures and compound characterisation

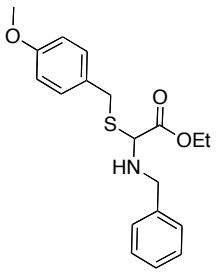
Chapter II Design, synthesis and biological evaluation of chaetocin analogues

2-Chloro-2-oxoethane-1,1-diyl diacetate (7)[108]



Compound **7** was prepared as follows:[108] Glyoxylic acid (21.5 mL of a 75% w/w in water, 175.0 mmol), acetic anhydride (175.0 mL, 1.8 mol) and glacial acetic acid (40.0 mL, 0.7 mol) were mixed and heated under reflux for 3 h. The solvents were rotary evaporated and the remaining volatiles were removed as an azeotrope with toluene. The residual brown oil was dissolved in a mixture of CH_2Cl_2 (120 mL) and $SOCl_2$ (45 mL, 0.6 mol) and heated under reflux for 0.5 h. The mixture was rotary evaporated to afford acyl chloride **7** as a brown oil that was used in the next step without further purification (30.5 g, 89%): IR (neat) 1767, 1373, 1185, 1042 cm^{-1}; ^1H NMR (400 MHz, $CDCl_3$) δ 6.91 (s, 1H), 2.21 (s, 6H); ^{13}C (100 MHz, $CDCl_3$) δ 168.1 (2C), 167.5, 87.8, 20.3 (2C); MS (CI) *m/z* 195 $(M+H)^+$, 212 $(M+NH_4)^+$; HRMS (CI) *m/z* calc. for $C_6H_{11}NO_5Cl$ 212.0326, found: 212.0331.

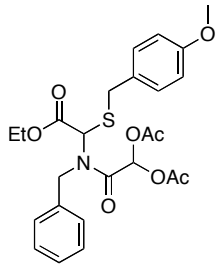
Ethyl 2-(benzylamino)-2-((4-methoxybenzyl)thio)acetate (9)[109]



Compound **9** was prepared as follows:[109] Benzylamine (13.8 mL, 126 mmol) was added to a solution of ethyl glyoxalate (25.0 mL of a 50% solution in toluene, 126 mmol) in toluene (200 mL). After 2 min, *p*-methoxybenzylmercaptan (17.5 mL, 126 mmol) was added and the resulting white solution was allowed to stir for 2.5 h at room temperature. The solvent was rotary evaporated and the resulting oil was purified by flash column chromatography (EtOAc : Hexanes, 1 : 2) to afford the title compound **9** as a yellow oil (21.2 g, 49%): IR (neat) 1729, 1611, 1511, 1243, 1172, 1032, 699 cm^{-1}; ^1H NMR (400 MHz, $CDCl_3$) δ 7.30 – 7.24 (m, 7H), 6.84 (d, *J* =

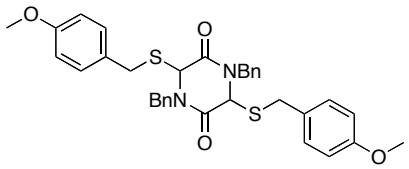
8.8 Hz, 2H), 4.35 (s, 1H), 4.22 (q, *J* = 7.2 Hz, 2H), 3.87 (d, *J* = 13.0 Hz, 1H), 3.81 (s, 3H), 3.76 (d, *J* = 3.2 Hz, 2H), 3.72 (d, *J* = 13.0 Hz, 1H), 1.32 (t, *J* = 7.2 Hz, 3H); ^{13}C (100 MHz, $CDCl_3$) δ 169.9, 158.7, 138.7, 130.0 (3C), 128.5 (2C), 128.4 (2C), 127.2, 114.0 (2C), 63.6, 61.4, 55.3, 49.2, 33.2, 14.1; MS (ESI) *m/z* 346 $(M+H)^+$; HRMS (ESI) *m/z* calc. for $C_{19}H_{24}NO_3S$ 346.1477, found: 346.1489.

2-(Benzyl(2-ethoxy-1-((4-methoxybenzyl)thio)-2-oxoethyl)amino)-2-oxoethane-1,1-diyl diacetate (13)[93]



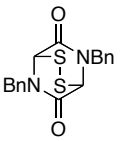
Compound **13** was prepared as follows:[93] Acyl chloride **7** (4.23 g, 21.8 mmol) in CH_2Cl_2 (30 mL) was added dropwise to a biphasic mixture of amine **9** (7.52 g, 21.8 mmol) in CH_2Cl_2 (100 mL) and $NaHCO_3$ (1.83 g, 21.8 mmol) in water (50 mL) at room temperature. The resulting biphasic solution was allowed to stir vigorously for 3 h. The organic phase was separated and the aqueous layer was extracted with CH_2Cl_2 (2 × 100 mL). The combined organic layers were dried ($MgSO_4$), filtered and rotary evaporated. Flash column chromatography (EtOAc : Hexanes, 1 : 2) afforded the title compound **13** as a colourless oil (8.31 g, 76%): IR (neat) 1770, 1743, 1684, 1513, 1246, 1192, 1028 cm^{-1}; ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.35 – 7.22 (m, 7H), 6.98 (s, 1H), 6.85 (d, *J* = 8.8 Hz, 2H), 5.91 (s, 1H), 4.93 (d, *J* = 17.6 Hz, 1H), 4.64 (d, *J* = 17.6 Hz, 1H), 4.09 – 3.99 (m, 2H), 3.81 (s, 3H), 3.78 (s, 2H), 2.04 (s, 3H), 2.02 (s, 3H), 1.21 (t, *J* = 7.1 Hz, 3H); ^{13}C (100 MHz, $CDCl_3$) δ 168.6, 168.3, 167.1, 165.7, 158.9, 135.7, 130.2 (2C), 128.7, 128.5 (2C), 127.6, 126.8 (2C), 114.0 (2C), 84.0, 62.2, 61.7, 55.2, 48.7, 35.5, 20.4, 20.2, 13.7; MS (ESI) *m/z* 504 $(M+H)^+$, 526 $(M+Na)^+$; HRMS (ESI) *m/z* calc. for $C_{25}H_{30}NO_8S$ 504.1692, found: 504.1674.

1,4-dibenzyl-3,6-bis((4-methoxybenzyl)thio)piperazine-2,5-dione (14)



Compound **14** was prepared as follows:[93,103] Benzylamine (1.3 mL, 11.7 mmol), followed by *p*-methoxybenzyl mercaptan (0.8 mL, 5.8 mmol) were added to a solution of ester **13** (2.00 g, 3.9 mmol) in MeCN (50 mL). The resulting mixture was allowed to stir for 2 min before adding TFA (85 μL, 1.1 mmol) and then heated at reflux for 20 h. The formed precipitate was recovered by filtration and rinsed with EtOAc to give the protected dithiodiketopiperazine **14** (1.02 g, 44%) as a white solid that was used in the next step without further purification: m.p. 172 – 174 ˚C; IR (neat) 1668, 1609, 1511, 1440, 1238, 1172, 1035 cm^{-1}; ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.36 (d, *J* = 8.6 Hz, 4H), 7.22 – 7.11 (m, 6H), 6.88 (d, *J* = 8.6 Hz, 4H), 6.65 (d, *J* = 7.2 Hz, 4H), 5.13 (d, *J* = 14.5 Hz, 2H), 4.39 (s, 2H), 3.99 (d, *J* = 14.0 Hz, 2H), 3.85 (s, 6H), 3.82 (d, *J* = 14.0 Hz, 2H), 3.57 (d, *J* = 14.5 Hz, 2H); ^{13}C (100 MHz, $CDCl_3$) δ 165.1 (2C), 159.2 (2C), 134.9 (2C), 130.8 (4C), 128.9 (2C), 128.6 (4C), 128.5 (4C), 127.8 (2C), 114.2 (4C), 57.5 (2C), 55.3 (2C), 46.1 (2C), 36.4 (2C); MS (ESI) *m/z* 599 $(M+H)^+$, 621 $(M+Na)^+$; HRMS (ESI) *m/z* calc. for $C_{34}H_{35}N_2O_4S_2$ 599.2038, found: 599.2024.

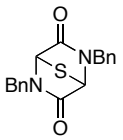
5,7-Dibenzyl-2,3-dithia-5,7-diazabicyclo[2.2.2]octane-6,8-dione (6)[93]



Compound **6** was prepared as follows:[93] BBr_3 (160 μL, 1.68 mmol) was added dropwise to a solution of diketopiperazine **14** (250 mg, 0.42 mmol) in CH_2Cl_2 (10 mL) at −78 ˚C. The solution was allowed to stir at this temperature for 45 min and then for 30 min at room temperature. After the addition of sat. aq. NH_4Cl (20 mL), the biphasic mixture was allowed to stir for 10 min and iodine was added portionwise until the red colour persisted. The resulting mixture was allowed to stir for 3 min, time after which solid sodium thiosulfate was added, until the colour of iodine disappeared. The mixture was allowed to stir for 10 min and then diluted with CH_2Cl_2 (10 mL) and water (20 mL). The organic layer was separated and the aqueous layer

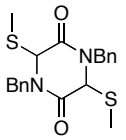
was extracted with CH_2Cl_2 (2 × 20 mL). The combined organic extracts were dried ($MgSO_4$), filtered and rotary evaporated. Purification by column chromatography (EtOAc : Hexanes, 1 : 3) furnished the title compound **6** as a white solid (110 mg, 74%): IR (neat) 1673, 1496, 1422, 1246, 1167, 730, 696 cm^{-1}; ^1H NMR (400 MHz, $CDCl_3$) δ 7.41 – 7.28 (m, 10H), 5.24 (s, 2H), 4.86 (d, J = 14.8 Hz, 2H), 4.50 (d, J = 14.8 Hz, 2H); ^{13}C (100 MHz, $CDCl_3$) δ 163.8 (2C), 134.1 (2C), 129.2 (4C), 128.7 (2C), 128.5 (4C), 64.7 (2C), 47.7 (2C); MS (CI) *m/z* 357 $(M+H)^+$, 374 $(M+NH_4)^+$; HRMS (CI) *m/z* calc. for $C_{18}H_{20}N_3O_2S_2$ 374.0997, found: 374.0996.

2,5-Dibenzyl-7-thia-2,5-diazabicyclo[2.2.1]heptane-3,6-dione (15)



Compound **15** was prepared as follows:[111] PPh_3 (40 mg, 0.15 mmol) was added to a solution of ETP **6** (50 mg, 0.14 mmol) in CH_2Cl_2 (10 mL). The mixture was allowed to stir for 16 h at room temperature. The solvent was then removed and column chromatography (CH_2Cl_2 : MeOH, 95 : 5) afforded the monosulfide derivative **15** (26.6 mg, 60%) as a white solid: m.p. 84 – 86 ˚C; IR (neat) 1699, 1357, 1190, 696 cm^{-1}; ^1H NMR (400 MHz, $CDCl_3$) δ 7.39 – 7.24 (m, 10H), 4.79 (d, J = 15.2 Hz, 2H), 4.72 (s, 2H), 4.27 (d, J = 15.2 Hz, 2H); ^{13}C (100 MHz, $CDCl_3$) δ 172.6 (2C), 134.6 (2C), 128.9 (4C), 128.3 (2C), 128.2 (4C), 66.4 (2C), 47.4 (2C); MS (CI) *m/z* 325 $(M+H)^+$, 342 $(M+NH_4)^+$; HRMS (CI) *m/z* calc. for $C_{18}H_{20}N_3O_2S$ 342.1276, found: 342.1275.

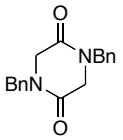
1,4-Dibenzyl-3,6-bis(methylthio)piperazine-2,5-dione (16)[93]



Compound **16** was prepared as follows:[111] $NaBH_4$ (9.4 mg, 0.24 mmol) was added to a solution of ETP **6** (11.2 mg, 0.03 mmol) in CH_2Cl_2 (3 mL) and MeOH (1 mL) at 0 ˚C. The reaction mixture was allowed to stir for 0.5 h. MeI (2.0 mL, 32 mmol) was added and the resulting mixture was allowed to stir for 16 h. The solution was then washed with 10% aq. HCl (2 mL). The aqueous layer was extracted with CH_2Cl_2 (2 × 3 mL). The organic layers were combined, dried ($MgSO_4$) and rotary evaporated to

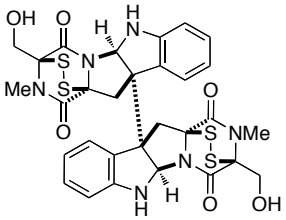
afford methylthioether **16** as a brown oil (7.4 mg, 60%) that was used in the next step without further purification: IR (neat) 1667, 1429, 1241, 1168, 727, 695 cm^{-1}; ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.37 – 7.26 (m, 10H), 5.36 (d, *J* = 15.0 Hz, 2H), 4.51 (s, 2H), 4.14 (d, *J* = 15.0 Hz, 2H), 2.33 (s, 6H); ^{13}C (100 MHz, $CDCl_3$) δ 164.3 (2C), 134.7 (2C), 129.1 (4C), 128.7 (4C), 128.4 (2C), 62.0 (2C), 46.9 (2C), 16.7 (2C); MS (CI) *m/z* 387 $(M+H)^+$, 404 $(M+NH_4)^+$; HRMS (CI) *m/z* calc. for $C_{20}H_{26}N_3O_2S_2$ 404.1466, found: 404.1471.

1,4-Dibenzylpiperazine-2,5-dione (18)[241]



Compound **18** was prepared as follows:[115] Nickel (II) chloride hexahydrate (267 mg, 1.12 mmol) was added to a solution of ETP **6** (40 mg, 0.11 mmol) in EtOH (5 mL). At 0 ˚C, a solution of $NaBH_4$ (127.5 mg, 3.37 mmol) in water (5 mL) was added dropwise to the reaction mixture which subsequently darkened. The mixture was allowed to stir at reflux for 16 h and was then extracted with CH_2Cl_2 (3 × 10 mL). The combined organic layers were washed with water (20 mL), dried ($MgSO_4$) and rotary evaporated. Column chromatography (CH_2Cl_2 : MeOH, 95 : 5) afforded the title compound **18** as a white solid (24.5 mg, 75%): IR (neat) 1648, 1483, 1333, 1254, 718, 699 cm^{-1}; ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.35 – 7.26 (m, 10H), 4.58 (s, 4H), 3.93 (s, 4H); ^{13}C (100 MHz, $CDCl_3$) δ 163.2 (2C), 134.9 (2C), 128.9 (4C), 128.5 (4C), 128.2 (2C), 49.3 (2C), 49.2 (2C); MS (ESI) *m/z* 295 $(M+H)^+$, 336 $(M+MeCN+H)^+$; HRMS (ESI) *m/z* calc. for $C_{18}H_{19}N_2O_2$ 295.1447, found: 295.1445.

Chaetocin (1) Production[96]

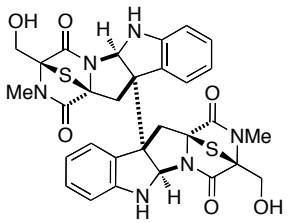


A conidial suspension of *Chaetomium virescens var. thielavioideum* was plated on solid complete media, containing 0.5% yeast extract (w/v), 1% vitamin solution, 1% (w/v) glucose as carbon source and 5 mM ammonium tartrate as

nitrogen source. The pH was adjusted to 6.5 with 10 M aq. NaOH. The plates were incubated during 5 days at 37 ˚C. A mycelia inoculum of *Chaetomium virescens var. thielavioideum* was prepared by dissecting agar surfaces of cultures grown on complete media and was inoculated into ten 1 L bottles half-filled with solid complete media.

After three weeks of incubation at room temperature, the culture media were extracted with CH_2Cl_2 (2 × 500 mL per bottle) for 24 h and 48 h with orbital shaking. The resulting mixtures were then filtered, dried over $MgSO_4$, and rotatory evaporated. The extracts were then purified by column chromatography (CH_2Cl_2 : MeOH, 98 : 2). The resulting solid was further purified by trituration in hexanes with sonication to afford 210 mg of chaetocin (**1**): $[\alpha]_D^{20}$ + 542 (*c* 0.82, $CHCl_3$); CD (MeOH) $\Delta\varepsilon^{20}$ (nm): +434 (237), 0 (263), –42 (272), 0 (284), +78 (304), +1 (350); IR (neat) 3383, 3336, 1670, 1066, 749 cm^{-1}; ^1H NMR (400 MHz, $CDCl_3$) δ 7.42 (d, *J* = 7.5 Hz, 2H), 7.25 (t, *J* = 7.5 Hz, 2H), 6.92 (t, *J* = 7.5 Hz, 2H), 6.74 (d, *J* = 7.5 Hz, 2H), 5.25 (s, 2H), 5.24 (br s, 2H), 4.25 (dd, *J* = 12.5, 6.0 Hz, 2H), 4.17 (dd, *J* = 12.5, 9.4 Hz, 2H), 3.83 (d, *J* = 15.0 Hz, 2H), 3.28 (dd, *J* = 9.4, 6.0 Hz, 2H), 3.08 (s, 6H), 2.74 (d, *J* = 15.0 Hz, 2H); ^{13}C (100 MHz, $CDCl_3$) δ 165.6 (2C), 162.8 (2C), 149.1 (2C), 130.4 (2C), 127.4 (2C), 125.1 (2C), 120.4 (2C), 110.7 (2C), 80.5 (2C), 75.7 (2C), 73.3 (2C), 60.6 (2C), 59.8 (2C), 39.2 (2C), 27.3 (2C); LCMS: R_t = 5.62 min, MS (ESI) *m/z* 697 $(M+H)^+$; HRMS (ESI) *m/z* calc. for $C_{30}H_{29}N_6O_6S_4$ 697.1031, found: 697.1021.

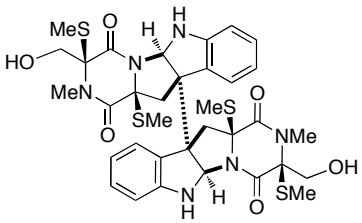
Chaetocin monosulfide derivative (3)[96]



Compound **3** was prepared as follows:[111] PPh_3 (28 mg, 0.11 mmol) was added to a solution of chaetocin (**1**, 20 mg, 0.03 mmol) in CH_2Cl_2 (5 mL) and the resulting mixture was allowed to stir for 2 h at room temperature. The solvent was rotatory evaporated and the resulting pink solid was purified by column chromatography (CH_2Cl_2 : EtOAc, 60 : 40) to afford monosulfide **3** (17 mg, 93 %) as a white solid: m.p. 210 – 212 ˚C (dec.); $[\alpha]_D^{20}$ + 484 (*c* 0.0039, $CHCl_3$); UV λ_{max} ($CHCl_3$) (A): 243.4 (0.49), 298 (0.21) nm; CD (MeOH) $\Delta\varepsilon^{20}$ (nm): 0 (217), –43.7 (223.5), 0 (230), +184.9

(248.5), +59.5 (285), +81.6 (304), +0.8 (350); IR (neat) 3388, 2921, 1712, 1468, 1322, 753 cm^{-1}; ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.47 (d, *J* = 7.8 Hz, 2H), 7.22 (td, *J* = 7.8, 0.7 Hz, 2H), 6.91 (td, *J* = 7.8, 0.7 Hz, 2H), 6.74 (d, *J* = 7.8 Hz, 2H), 5.46 (s, 2H), 4.89 (s, 2H), 4.17 (dd, *J* = 13.1, 5.2 Hz, 2H), 4.09 (dd, *J* = 13.1, 8.2 Hz, 2H), 3.63 (d, *J* = 14.9 Hz, 2H), 3.00 (s, 6H), 2.81 (dd, *J* = 8.2, 5.2 Hz, 2H), 2.49 (d, *J* = 14.9 Hz, 2H); ^{13}C (100 MHz, $CDCl_3$) δ 174.9 (2C), 173.5 (2C), 149.2 (2C), 130.2 (2C), 128.1 (2C), 124.9 (2C), 120.3 (2C), 110.6 (2C), 80.7 (2C), 79.4 (2C), 78.5 (2C), 63.6 (2C), 59.1 (2C), 31.6 (2C), 28.3 (2C); LCMS: R_t = 4.77 min; MS (ESI) *m/z* 633 $(M+H)^+$; HRMS (ESI) *m/z* calc. for $C_{30}H_{29}N_6O_6S_2$ 633.1590 found: 633.1598.

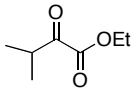
Chaetocin methylthioether derivative (4)[96]



Compound **4** was prepared as follows:[111] $NaBH_4$ (16 mg, 0.42 mmol) was added to a solution of chaetocin (**1**, 20 mg, 0.03 mmol) in CH_2Cl_2 / MeOH (6 mL / 2 mL) at 0 ˚C. After 30 min, MeI (1 mL, excess) was added dropwise to the solution. It was allowed to stir for 16 h and then treated with 1 N aq. HCl (5 mL). The organic layer was separated, the aqueous layer was extracted with CH_2Cl_2 (2 × 5 mL) and the combined organic layers were dried ($MgSO_4$), rotary evaporated and purified by column chromatography (CH_2Cl_2 : MeOH, 99 : 1) to afford the methylated compound **4** (10 mg, 47 %) as a white solid: m.p. 130 – 132 ˚C (dec.); $[\alpha]_D^{20}$ + 175 (*c* 0.002, $CHCl_3$); UV λ_{max} ($CHCl_3$) (A): 238.6 (1.07), 303.3 (0.44) nm, CD (MeOH) $\Delta\varepsilon^{20}$ (nm): +180 (258), +53 (280), +126 (306), 0 (332); IR (neat) 3369, 2922, 2852, 1656, 1420, 1385 cm^{-1}; ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.48 (d, *J* = 7.6 Hz, 2H), 7.20 (t, *J* = 7.6 Hz, 2H), 6.85 (t, *J* = 7.6 Hz, 2H), 6.63 (d, *J* = 7.6 Hz, 2H), 5.73 (s, 2H), 4.90 (br s, 2H), 4.65 (br s, 2H), 4.29 (d, *J* = 10.7 Hz, 2H), 3.78 (d, *J* = 10.7 Hz, 2H), 3.09 (s, 6H), 2.64 (d, *J* = 15.2 Hz, 2H), 2.44 (d, *J* = 15.2 Hz, 2H), 2.19 (s, 6H), 1.87 (s, 6H); ^{13}C (100 MHz, $CDCl_3$) δ 166.5 (2C), 165.4 (2C), 150.9 (2C), 129.7 (2C), 128.7 (2C), 125.1 (2C), 119.1 (2C), 110.1 (2C), 80.5 (2C), 77.2 (2C), 69.2 (2C), 64.6 (2C), 60.6 (2C), 43.7 (2C), 28.6 (2C), 15.8 (2C), 12.8 (2C); LCMS: R_t = 5.58 min; MS (ESI) *m/z* 779 $(M+Na)^+$; HRMS (ESI) *m/z* calc. for $C_{34}H_{40}N_6O_6S_4Na$ 779.1790, found: 779.1777.

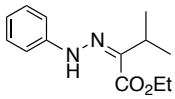
Chapter III Stereochemistry of ETP desulfurisation

Ethyl 3-methyl-2-oxobutanoate (40)[182]



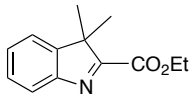
Compound **40** was prepared as follows:[182] Isopropyl magnesium chloride (2 M in THF, 50 mL, 100 mmol) was added dropwise to a solution of diethyl oxalate (6.8 mL, 50 mmol) in dry THF (50 mL) at −78 ˚C. The solution was allowed to stir at −78 ˚C for 1 h, then poured into a allowed to stir suspension of ice (40 g), Et_2O (50 mL) and conc. aq. HCl (3.6 mL). The mixture was extracted with Et_2O (3 × 50 mL). The organic layer was washed with water, dried ($MgSO_4$), filtered through celite and rotatory evaporated to afford the title compound **40** as a yellow oil (3.310 g, 46%) that was used in the next step without further purification: ^{1}H NMR (400 MHz, $CDCl_3$), δ 4.31 (q, *J* = 7.2 Hz, 2H), 3.25 (sept, *J* = 6.8 Hz, 1H), 1.35 (t, *J* = 7.2 Hz, 3H), 1.14 (d, *J* = 6.8 Hz, 6H); ^{13}C (100 MHz, $CDCl_3$) δ 198.2, 161.8, 62.2, 36.9, 17.1 (2C), 14.0; MS (CI) *m/z* 162 ($M+NH_4^+$); HRMS (CI) *m/z* calc. for $C_7H_{16}NO_3$ 162.1130, found: 162.1128.

(*Z*)-Ethyl 3-methyl-2-(2-phenylhydrazono)butanoate (41)[182]



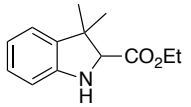
Compound **41** was prepared as follows:[182] A solution of keto-ester **40** (3.25 g, 22.6 mmol) and phenylhydrazine (3.4 mL, 33.9 mmol) in toluene was heated at 60 ˚C for 1 h under N_2. The solution was rotatory evaporated and the crude product was purified by column chromatography (EtOAc : Hexanes, 5 : 95) to afford hydrazone **41** as a yellow oil (4.11 g, 78%): ^{1}H NMR (400 MHz, $CDCl_3$) δ 12.02 (br s, 1H), 7.26 – 7.32 (m, 2H), 7.19 – 7.21 (m, 2H), 6.93 – 6.97 (m, 1H), 4.29 (q, *J* = 7.1 Hz, 2H), 3.02 (sept, *J* = 6.8 Hz, 1H), 1.36 (t, *J* = 7.1 Hz, 3H), 1.19 (d, *J* = 6.8 Hz, 6H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 163.6, 143.9, 133.0, 129.2 (2C), 121.4, 113.5 (2C), 60.4, 30.9, 21.2 (2C), 14.2; MS (ESI) *m/z* 235 ($M+H^+$); HRMS (ESI) *m/z* calc. for $C_{13}H_{19}N_2O_2$ 235.1447, found 235.1444.

Ethyl 3,3-dimethyl-3*H*-indole-2-carboxylate (39)[182]



Compound **39** was prepared as follows:[182] A solution of 34% HCl in EtOH was prepared by adding acetyl chloride to EtOH at 0 ˚C. Hydrazone ester **41** (4.11 g, 17.6 mmol) was added to the ethanolic HCl (150 mL) and the mixture was heated to reflux for 10 min. The solvent was then rotatory evaporated to afford a residue that was taken up in Et_2O (100 mL) and treated with 5% aq. Na_2CO_3. The organic phase was separated and the aqueous layer was extracted with Et_2O (3 × 50 mL), and the combined organic layers were washed with water until neutral, dried ($MgSO_4$) and concentrated under reduced pressure. Purification by column chromatography (EtOAc : Hexanes, 1 : 9) afforded indolenine **39** (2.57 g, 67%) as light yellow solid: ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.83 – 7.81 (m, 1H), 7.33 – 7.38 (m, 3H), 4.44 (q, *J* = 7.1 Hz, 2H), 1.51 (s, 6H), 1.44 (t, *J* = 7.1 Hz, 3H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 175.5, 161.6, 151.9, 147.5, 128.4, 128.0, 123.5, 121.3, 61.7, 54.2, 22.6 (2C), 14.2; MS (ESI) *m/z* 218 ($M+H^+$); HRMS (ESI) *m/z* calc. for $C_{13}H_{16}NO_2$ 218.1181, found 218.1184.

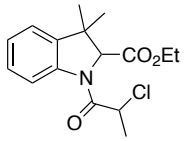
Ethyl 3,3-dimethylindoline-2-carboxylate (42)[182]



Compound **42** was prepared as follows:[184] To a solution of indolenine **39** (615 mg, 2.83 mmol) in 12 mL of MeOH and 2.5 mL of AcOH was added sodium cyanoborohydride (232 mg, 3.69 mmol). The reaction mixture was allowed to stir at room temperature for 18 h, time after which the reaction was quenched by addition of 12 N aq. HCl (0.5 mL). Removal of the solvent under reduced pressure gave a residue which was treated with 5% aq. Na_2CO_3 (15 mL) and the resulting solution was extracted with CH_2Cl_2 (3 × 50 mL). The combined organic extracts were dried ($MgSO_4$) and evaporated to give indoline **42** a colourless oil (640 mg, 98%) that was used in the next step without further purification: ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.06 (td, *J* = 7.8, 1.3 Hz, 1H), 7.01 (dd, *J* = 7.5, 1.3 Hz, 1H), 6.79 (td, *J* = 7.3, 0.9 Hz, 1H), 6.71 (*app*-d, *J* = 7.9 Hz, 1H), 4.32 – 4.22 (m, 2H), 4.19 (br s, 1H), 4.16 (s, 1H), 1.53 (s, 3H), 1.33 (t, *J* = 7.2 Hz, 3H), 1.16 (s, 3H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 172.2,

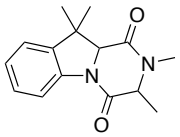
148.1, 137.2, 127.7, 121.9, 119.7, 110.4, 71.5, 60.9, 45.8, 27.7, 24.4, 14.3; MS (ESI) *m/z* 220 $(M+H)^+$; HRMS (ESI) *m/z* calc. for $C_{13}H_{18}NO_2$ 220.1338, found: 220.1337.

Ethyl 1-(2-chloropropanoyl)-3,3-dimethylindoline-2-carboxylate (43)



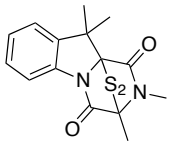
Compound **43** was prepared as follows:[185] 2-Chloropropionyl chloride (0.35 mL, 3.56 mmol) was added dropwise to a solution of triethylamine (0.46 mL, 3.29 mmol) and indoline **42** (600 mg, 2.70 mmol) in chloroform (50 mL) at 0 ˚C. The mixture was allowed to warm to room temperature and stir for a further 4 h. After that time, the reaction mixture was poured into ice-water, and extracted with CH_2Cl_2 (3 × 50 mL). The combined organic extracts were washed with brine, dried ($MgSO_4$), filtered and rotatory evaporated. The residue was purified by column chromatography (Hexanes : EtOAc, 95 : 5) to afford the title compound **43** as two diastereoisomers as colourless oil (combined mass 731 mg, 87%): IR (neat) 2972, 1747, 1673, 1598, 1483, 1409, 1199, 752 cm^{-1}; Diastereoisomer 1: ^1H NMR (500 MHz, $CDCl_3$) δ 8.22 (br s, 1H), 7.28 – 7.24 (m, 1H), 7.10 – 7.08 (m, 2H), 4.86 (s, 1H), 4.26 – 4.11 (m, 3H), 1.73 (br d, *J* = 6.4 Hz, 3H), 1.42 (s, 3H), 1.40 (s, 3H), 1.26 (t, *J* = 7.2 Hz, 3H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 169.8, 166.1, 141.2, 138.7, 128.3, 124.9, 121.6, 117.5, 72.0, 61.7, 51.6, 45.1, 32.0, 22.7, 20.3, 14.2; MS (CI) *m/z* 310 $(M+H^+)$; HRMS (ESI) *m/z* calc. for $C_{16}H_{21}NO_3Cl$ 310.1210, found: 310.1213; Diastereoisomer 2: ^1H NMR (500 MHz, DMSO) δ 8.05 (br d, *J* = 7.8 Hz, 1H), 7.28 – 7.24 (m, 2H), 7.10 (br t, *J* = 7.5 Hz, 1H), 5.09 (br s, 1H), 5.02 (br q, *J* = 6.2 Hz, 1H), 4.10 (q, *J* = 7.0 Hz, 2H), 1.78 (br d, *J* = 6.2 Hz, 3H), 1.33 (br s, 3H), 1.31 (br s, 3H), 1.17 (t, *J* = 7.0 Hz, 3H); ^{13}C NMR (100 MHz, DMSO) δ 169.2, 166.8, 141.4, 139.6, 127.6, 124.6, 122.1, 116.7, 71.0, 60.9, 51.5, 44.6, 31.2, 21.8, 21.2, 13.9; MS (CI) *m/z* 310 $(M+H^+)$; HRMS (ESI) *m/z* calc. for $C_{16}H_{21}NO_3Cl$ 310.1210, found: 310.1215.

2,3,10,10-tetramethyl-2,3,10,10a-tetrahydropyrazino[1,2-*a*]indole-1,4-dione (44)



Compound **44** was prepared as follows:[185] Potassium carbonate (218 mg, 1.7 mmol) and an excess of methylamine (5.5 mL of a 2M solution in THF, 11.0 mmol) were added to a solution of acylated indoline **43** (340 mg, 1.1 mmol) in MeCN (10 mL) in a sealed tube. After refluxing for 18 h, the reaction mixture was filtered, concentrated and the residue was purified by column chromatography (Hexanes : EtOAc, 1 : 1) to afford the title compound **44** as a mixture of non-separable diastereoisomers as a colorless oil (190 mg, 67%): IR (neat) 2969, 1666, 1602, 1483, 1426, 1402, 1290, 755 cm^{-1}; Diastereoisomer 1 (*cis)*: ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.03 (*app*-d, *J* = 7.6 Hz, 1H), 7.28 – 7.13 (m, 3H), 4.38 (d, *J* = 1.6 Hz, 1H), 4.16 (dq, *J* = 7.1, 1.6 Hz, 1H), 3.05 (s, 3H), 1.70 (s, 3H), 1.67 (d, *J* = 7.1 Hz, 3H), 1.16 (s, 3H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 164.6, 163.8, 140.4, 138.8, 127.5, 125.4, 121.6, 116.9, 68.5, 57.3, 44.5, 30.9, 24.0, 23.1, 18.0; Diastereoisomer 2 (*trans*): ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.04 (*app*-d, *J* = 7.8 Hz, 1H), 7.28 – 7.13 (m, 3H), 4.34 (s, 1H), 4.02 (q, *J* = 7.1 Hz, 1H), 3.01 (s, 3H), 1.70 (s, 3H), 1.54 (d, *J* = 7.1 Hz, 3H), 1.19 (s, 3H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 165.1, 164.8, 140.8, 138.7, 127.9, 125.6, 121.9, 116.9, 68.2, 60.2, 45.1, 31.7, 24.9, 24.3, 17.4; MS (ESI) m/z 259 ($M+H^+$); HRMS (ESI) *m/z* calc. for $C_{15}H_{19}N_2O_2$ 259.1447, found: 259.1462.

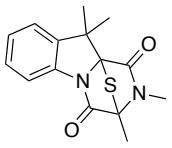
2,3,10,10-Tetramethyl-2,3-dihydro-1*H*-3,10a-epithiopyrazino[1,2-*a*]indole-1,4(10*H*)-dione 11-sulfide – ((±)-25)[242]



Compound **25** was prepared as follows:[100,102] NaHMDS (6.4 mL of a 0.6 M solution in PhMe, 3.84 mmol) was added dropwise to a solution of elemental sulfur (328 mg, 10.24 mmol) in dry THF (40 mL) at room temperature. The solution was allowed to stir for 1 min, time after which diketopiperazine **44** (330 mg, 1.28 mmol) in THF (6 mL) was added dropwise. The solution was allowed to stir for 1 min and additional NaHDMS (4.3 mL of a 0.6 M solution in PhMe, 2.56 mmol) was added. The resulting orange / light brown mixture was allowed to stir for 30 min at room temperature. The

solution was then quenched with aq. sat. NH_4Cl and extracted with CH_2Cl_2 (3 × 50 mL). The combined organic layers were dried ($MgSO_4$), filtered and concentrated to give a dark green residue which was taken up in a degassed mixture of THF – EtOH (1:1, 26 mL) and cooled to 0 ˚C. Sodium borohydride (1.210 g, 32 mmol) was then added in small portions and the resulting mixture was allowed to warm up to room temperature and to stir for 45 min. The solution was subsequently cooled to 0 ˚C and quenched by addition of aq. sat. NH_4Cl. The mixture was extracted with EtOAc (3 × 50 mL) and the combined organic extracts were then treated with an aqueous solution of KI_3 (1.4 M), until the dark purple colour persisted. The mixture was allowed to stir for 10 min and quenched with 5% aq. $Na_2S_2O_3$ solution until disappearance of the dark purple colour. The resulting mixture was extracted with EtOAc (3 × 50 mL) and the combined organic layers were dried ($MgSO_4$), filtered and concentrated. The residue was purified by column chromatography (Hexanes : EtOAc, 90 : 10) to afford the racemic title compound **25** as a yellow solid (100 mg, 24%) which was recrystallised from CH_2Cl_2 to afford white crystals: m.p. 104 – 106 °C; IR (neat) 1690, 1460, 1357, 1174 cm^{-1}; ^1H NMR (400 MHz, $CDCl_3$) δ 8.02 (d, *J* = 8.0 Hz, 1H), 7.37 – 7.33 (m, 1H), 7.28 – 7.25 (m, 2H), 3.12 (s, 3H), 2.10 (s, 3H), 1.79 (s, 3H), 1.59 (s, 3H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 164.3, 162.4, 139.1, 136.1, 128.8, 126.1, 121.3, 115.9, 82.6, 74.3, 49.2, 28.7, 27.3, 20.5, 18.2; MS (CI) *m/z* 321 $(M+H)^+$, 257 $(M–S_2+H)^+$; HRMS (CI) *m/z* calc. for $C_{15}H_{17}N_2O_2S_2$ 321.0731, found: 321.0745. The obtained enantiomers could be separated by chiral HPLC (OD+ semiprep column, Hexane : Isopropanol, 1 : 1). First peak: $[\alpha]_D^{20}$ −335 (*c* 2.00, $CHCl_3$), Second peak: $[\alpha]_D^{20}$ +394 (*c* 2.05, $CHCl_3$).

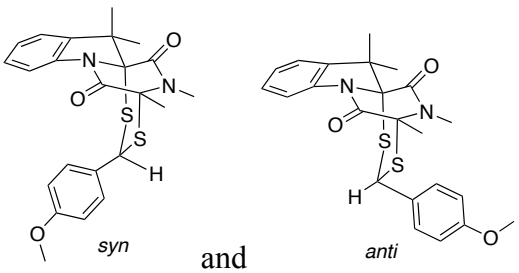
2,3,10,10-Tetramethyl-2,3-dihydro-1*H*-3,10a-epithiopyrazino[1,2-*a*]indole-1,4(10*H*)-dione - ((±)-26)[242]



To a solution of ETP **25** (33 mg, 0.10 mmol) in dioxane (8 mL) was added PPh_3 (33 mg, 0.16 mmol) and the resulting mixture was allowed to stir overnight at room temperature. The solvent was then removed under reduced pressure and the pink residue was purified by column chromatography (Hexanes : EtOAc, 100 : 0 to 95 : 5)

to afford the racemic title compound **26** as colourless oil (19 mg, 64%), which was recrystallised from CH_2Cl_2 to give a white solid: m.p. 58 – 60 ˚C; IR (neat) 1720, 1456, 1387, 1288, 1134 cm^{-1}; ^1H NMR (400 MHz, $CDCl_3$) δ 8.54 (*app*-d, *J* = 7.8 Hz, 1H), 7.25 (td, *J* = 7.8, 1.0 Hz, 1H), 7.20 (dd, *J* = 7.8, 1.0 Hz, 1H), 7.13 (td, *J* = 7.8, 1.0 Hz, 1H), 2.96 (s, 3H), 1.83 (s, 3H), 1.75 (s, 3H), 1.48 (s, 3H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 172.5, 172.0, 139.7, 138.1, 128.1, 124.7, 122.4, 113.6, 86.6, 75.1, 43.5, 27.2, 26.3, 25.7, 13.3; MS (CI) *m/z* 289 $(M+H)^+$, 306 $(M+NH_4)^+$; HRMS (CI) *m/z* calc. for $C_{15}H_{17}N_2O_2S$ 289.1011, found: 289.1026. The obtained enantiomers could be separated by chiral HPLC (OD+ semiprep column, Hexane : Isopropanol, 90 : 10): First peak: $[\alpha]_D^{20}$ −47.5 (*c* 1.12, CH_2Cl_2), Second peak $[\alpha]_D^{20}$ +34.4 (*c* 1.12, CH_2Cl_2).

2-(4-methoxyphenyl)-4,11,11,13-tetramethyl-4*H*-4,11a-(epiminomethano)[1,3,5]dithiazepino[5,4-*a*]indole-5,12(11*H*)-dione (47)



From ETP **25**: $NaBH_4$ (4.5 mg, 0.117 mmol) was added to a solution of racemic ETP **25** (15.0 mg, 0.047 mmol) in degassed MeOH : THF (1:1, 5 mL) at 0 ˚C. Once complete (monitored by TLC), a few drops of 0.1 N aq. HCl were added and the solvent was removed under reduced pressure. The residue was taken up with CH_2Cl_2 (5 mL), dried ($MgSO_4$) and the solvent was removed under reduced pressure. The residue obtained was dried under vacuum and then dissolved in dry CH_2Cl_2 (5 mL). *p*-anisaldehyde (7 μL, 0.056 mmol) and boron trifluoride etherate (9 μL, 0.073 mmol) were added and the reaction was allowed to stir for 16 h under N_2. The solvent was then removed under reduced pressure and the residue was purified by column chromatography (Hexanes : EtOAc, 80 : 20) to afford the title compound as a mixture of diastereoisomers as a colorless oil (Diastereoisomer 1: only traces observed by TLC and LCMS of the crude but not isolated; Diastereoisomer 2: 9 mg, 43%;). Diastereoisomer 2: IR (neat) 1682, 1606, 1511, 1483, 1372, 1257 cm^{-1}; ^1H NMR (400 MHz, $CDCl_3$) δ 8.18 (d, *J* = 7.9 Hz, 1H), 7.35 – 7.31 (m, 1H), 7.28 (*app*-d, *J* = 8.8 Hz, 2H), 7.19 (d, *J* = 4.2 Hz, 2H), 6.80 (*app*-d, *J* = 8.8 Hz, 2H), 5.10 (s, 1H), 3.76 (s, 3H),

3.27 (s, 3H), 1.98 (s, 3H), 1.72 (s, 3H), 1.34 (s, 3H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 165.5, 162.2, 160.4, 139.5, 137.5, 130.6 (2C), 128.4, 126.8, 125.9, 121.5, 117.4, 114.2 (2C), 82.1, 68.8, 55.3, 52.0, 49.2, 28.0, 27.6, 22.3, 21.8; A NOESY experiment showed a correlation between the signal at 5.1 ppm (CH of the thioacetal carbon) and the signal at 3.27 ppm (CH_3 on the nitrogen atom of the ETP ring) thus suggesting that this diastereoisomer is the *syn*. MS (ES) *m/z* 441 $(M+H)^+$; HRMS (ES) *m/z* calc. for $C_{23}H_{25}N_2O_3S_2$ 441.1307, found: 441.1316.

The 2 enantiomers could be observed by chiral HPLC (OD+ analytical column, Hexane : Isopropanol, 98 : 2, 1 mL/min) around 60 min and 100 min. The reaction was repeated on a single enantiomer of ETP **25** (first peak by HPLC separation) and the product was also obtained as mixture of diastereoisomers (13% of diastereoisomer 2, and 6% of diastereoisomer 1). Diastereoisomer 2 was obtained as a single enantiomer eluting with a comparable retention time as the first peak of the racemic (around 70 min). $[\alpha]_D^{20}$ −66 (*c* 0.20, $CHCl_3$)

From monosulfide (**26**):

The trithiane derivative of *p*-anysaldehyde **48** was prepared as follows:[189,190] A mixture of Na_2S (390 mg, 5 mmol) and $SiCl_4$ (1.15 mL, 10 mmol) in MeCN (7.5 mL) was allowed to stir for 15 min. After that time, *p*-anysaldehyde (304 µL, 2.5 mmol) was added and the mixture was allowed to stir for 16 h at room temperature. Ice-cold water (50 mL) was then added and the resulting mixture was extracted with EtOAc (2 × 100 mL), dried ($MgSO_4$) and the solvent was removed under reduced pressure. The solid obtained was recrystallised from acetic acid to afford the trithiane derivative of *p*-anysaldehyde **48** as a pink solid (41 mg, 4%).

In DMSO, only the Beta conformer was observed[243]: ^{1}H NMR (400 MHz, *d*-DMSO) δ 7.36 (d, *J* = 8.7 Hz, 3 × 2H), 6.96 (d, *J* = 8.7 Hz, 3 × 2H), 5.68 (s, 3 × 1H), 3.75 (s, 3 × 3H); ^{13}C NMR (100 MHz, DMSO) δ 159.4 (3 × 1C), 130.2 (3 × 1C), 129.1 (3 × 2C), 114.4 (3 × 2C), 56.4 (3 × 1C), 55.2 (3 × 1C); In $CDCl_3$, alpha/beta ratio[243] varied as a function of the temperature (1:10 at 25 ˚C versus 1:5 at 40 ˚C). Alpha conformer: ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.80 (d, *J* = 8.8 Hz, 2H), 7.38 (d, *J* = 8.7 Hz, 2 × 2H), 6.97 (d, *J* = 8.8 Hz, 2H), 6.88 (d, *J* = 8.7 Hz, 2 × 2H), 5.69 (s, 1H), 5.40 (s, 2 × 1H), 3.86 (s, 3H), 3.80 (s, 2 × 3H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 174.4, 160.0 (2 × 1C), 130.4 (2 × 1C), 130.3, 130.1 (2 × 1C), 129.2 (2 × 2C), 114.4 (2 × 2C), 114.3 (1 × 2C), 55.4,

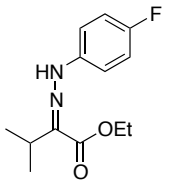
55.3 (2 × 1C), 52.8, 51.2 (2 × 1C); MS (CI) *m/z* 153 $(1/3M+H)^+$; HRMS (CI) *m/z* calc. for C_8H_9OS 153.0369, found: 153.0422.

A solution of the racemic monosulfide **26** (21 mg, 0.073 mmol) in CH_2Cl_2 (10 mL) was treated with the trithiane derivative of *p*-anysaldehyde **48** (17 mg, 0.036 mmol) and boron trifluoride etherate (15 μL, 0.117 mmol). The mixture was allowed to stir at 40 ˚C for 5 h and then additional boron trifluoride etherate (2 × 30 μL, 0.234 mmol) was added. The mixture was subsequently absorbed on silica gel and purified by column chromatography (Hexanes : EtOAc, 85 : 15) and title compound **47** was obtained as a mixture of two diastereoisomers as colourless oils (56% of diastereoisomer 1 and 25% of diastereoisomer 2, 7:3): Diastereoisomer 1: IR (neat) 1682, 1606, 1510, 1481, 1459, 1372, 1254 cm^{-1}; ^1H NMR (500 MHz, $CDCl_3$) δ 8.23 (d, *J* = 8.0 Hz, 1H), 7.37 – 7.33 (m, 1H), 7.24 (*app*-d, *J* = 8.7 Hz, 2H), 7.27 – 7.21 (m 2H), 6.78 (*app*-d, *J* = 8.7 Hz, 2H), 4.98 (s, 1H), 3.76 (s, 3H), 3.14 (s, 3H), 2.00 (s, 3H), 1.74 (s, 3H), 1.33 (s, 3H); ^{13}C NMR (125 MHz, $CDCl_3$) δ 166.0, 161.7, 160.4, 140.0, 136.6, 130.6 (2C), 128.2, 126.7, 126.5, 122.2, 116.8, 114.2 (2C), 77.8, 71.3, 55.3, 51.6, 50.4, 28.5, 27.6, 22.7, 19.5; A NOESY experiment did not show any correlation between the signal at 5.1 ppm (CH of the thioacetal carbon) and the signal at 3.37 ppm (CH_3 on the nitrogen atom of the ETP ring) thus suggesting that this diastereoisomer is the *trans*. MS (ES) *m/z* 441 $(M+H)^+$; HRMS (ES) *m/z* calc. for $C_{23}H_{25}N_2O_3S_2$ 441.1307, found: 441.1308;

Diastereoisomer 2 matched the data of the compound obtained by reaction of the ETP **25** (IR, ^1H, ^{13}C, HSQC and NOESY). MS (ES) *m/z* 441 $(M+H)^+$; HRMS (ES) *m/z* calc. for $C_{23}H_{25}N_2O_3S_2$ 441.1307, found: 441.1318; The two enantiomers could be observed by chiral HPLC (OD+ analytical column, Hexane : Isopropanol, 98 : 2, 1 mL/min) at 52 min and 80 min.

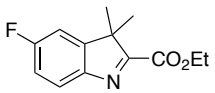
The reaction was repeated on a single enantiomer of monosulfide **26** (first peak by HPLC separation) and diastereoisomer 2 was obtained as a single enantiomer $[\alpha]_D^{20}$ –40 (*c* 0.20, $CHCl_3$) eluting with the comparable retention time as the first peak of the racemic (around 70 min), and importantly with the same retention time as the thioacetal obtained from ETP **25** (as confirmed by spiking).

Ethyl 2-(2-(4-fluorophenyl)hydrazono)-3-methylbutanoate (51)



Compound **51** was prepared as follows:[182] Keto-ester **40** (2.000 g, 13.87 mmol) was added to a solution of the hydrochloride salt of 4-fluorophenylhydrazine (3.384 g, 20.81 mmol) and triethylamine (2.9 mL, 20.81 mmol) in toluene and the resulting mixture was heated at 60 ˚C for 1 h. The solvent was rotatory evaporated and the crude product purified by column chromatography (EtOAc : Hexanes, 5 : 95) to afford hydrazone **51** as a yellow oil (3.250 g, 93%): IR (neat) 1727, 1683, 1551, 1510, 1205, 1155 1092 cm^{-1}; ^{1}H NMR (400 MHz, $CDCl_3$) δ 12.01 (br s, 1H), 7.13 – 7.10 (m, 2H), 7.01 – 6.96 (m, 2H), 4.28 (q, *J* = 7.1 Hz, 2H), 3.01 (sept, *J* = 6.8 Hz, 1H), 1.36 (t, *J* = 7.1 Hz, 3H), 1.18 (d, *J* = 6.8 Hz, 6H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 163.7, 158.0 (d, *J* = 237 Hz, 1C), 140.3, 133.0, 115.8 (d, *J* = 22 Hz, 2C), 114.4 (d, *J* = 8 Hz, 2C), 60.5, 30.8, 21.2 (2C), 14.2; ^{19}F NMR (377 MHz, $CDCl_3$) δ −123.05; MS (ESI) *m/z* 253 ($M+H^+$), 236; HRMS (ESI) *m/z* calc. for $C_{13}H_{18}FN_2O_2$ 253.1352, found 253.1358.

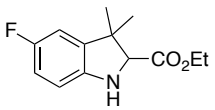
Ethyl 5-fluoro-3,3-dimethyl-3*H*-indole-2-carboxylate (52)



Compound **52** was prepared as follows:[182] A solution of 34% HCl in EtOH was prepared by adding acetyl chloride to EtOH at 0 ˚C. Hydrazone ester **51** (3.250 g, 12.88 mmol) was added to the ethanolic HCl (150 mL) and the mixture was heated to reflux for 10 min. The solvent was then rotatory evaporated to afford a residue that was taken up in Et_2O (100 mL) and treated with 5% aq. Na_2CO_3. The organic phase was separated and the aqueous layer was extracted with Et_2O (3 × 50 mL), and the combined organic layers were washed with water until neutral, dried ($MgSO_4$) and rotatory evaporated. Purification by column chromatography (EtOAc : Hexanes, 1 : 9) afforded indolenine **52** (1.662 g, 55%) as a pale pink solid: m.p. 90 – 92 ˚C; IR (neat) 1710, 1595, 1543, 1461, 1305, 1198, 1093, 1057 cm^{-1}; ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.77 (dd, *J* = 8.4, 4.7 (J_{H-F}) Hz, 1H), 7.09 – 7.02 (m, 2H), 4.44 (q, *J* = 7.1 Hz, 2H), 1.51 (s, 6H), 1.44 (t, *J* = 7.1 Hz, 3H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 175.3, 163.1 (d, *J* = 250.3 Hz, 1C), 161.3, 149.9 (d, *J* = 8.6 Hz, 1C) , 147.9, 124.6 (d, *J* = 9.7 Hz, 1C),

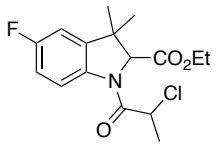
115.1 (d, J = 24.0 Hz, 1C), 109.1 (d, J = 24.9 Hz, 1C), 61.8, 54.7, 22.6 (2C), 14.2; ^{19}F NMR (377 MHz, $CDCl_3$) δ −111.64; MS (ESI) *m/z* 236 ($M+H^+$); HRMS (ESI) *m/z* calc. for $C_{13}H_{15}FNO_2$ 236.1087, found 236.1107.

Ethyl 5-fluoro-3,3-dimethylindoline-2-carboxylate (53)



Compound **53** was prepared as follows:[184] To a solution of indolenine **52** (1.445 g, 6.14 mmol) in 28 mL of MeOH and 7 mL of AcOH was added sodium cyanoborohydride (502 mg, 7.98 mmol). The reaction mixture was allowed to stir at room temperature for 18 h. 12 N aq. HCl (2 mL) was then added. The solvent was rotatory evaporated and the obtained residue which was taken up in 5% aq. Na_2CO_3 (50 mL). The solution was extracted with CH_2Cl_2 (3 × 50 mL) and the combined organic extracts were dried ($MgSO_4$) and rotatory evaporated to afford title compound **53** as a yellow oil (1.721 g, quant.) that was used in the next step without further purification: ^{1}H NMR (400 MHz, $CDCl_3$) δ 6.77 – 6.70 (m, 2H), 6.62 (dd, J = 8.3, 4.3 ($J_{H\text{-}F}$) Hz, 1H), 4.31 – 4.22 (m, 2H), 4.16 (s, 1H), 4.08 (br s, 1H), 1.50 (s, 3H), 1.33 (t, J = 7.1 Hz, 3H), 1.14 (s, 3H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 172.0, 157.8 (d, J = 234.5 Hz, 1C), 143.9, 139.2 (d, J = 8.2 Hz, 1C), 113.8 (d, J = 23.7 Hz, 1C), 111.1 (d, J = 8.7 Hz, 1C), 109.4 (d, J = 24.1 Hz, 1C), 71.9, 61.1, 46.1, 27.5, 24.2, 14.3; ^{19}F NMR (377 MHz, $CDCl_3$) δ −124.62; MS (ESI) *m/z* 238 $(M+H)^+$; HRMS (ESI) *m/z* calc. for $C_{13}H_{17}FNO_2$ 238.1243, found: 238.1248.

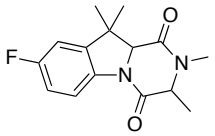
Ethyl 1-(2-chloropropanoyl)-5-fluoro-3,3-dimethylindoline-2-carboxylate (54)



Compound **54** was prepared as follows:[185] 2-Chloropropionyl chloride (0.774 mL, 7.98 mmol) was added dropwise to a solution of triethylamine (1.027 mL, 7.37 mmol) and crude indoline **53** (1.721 g, 6.14 mmol) in chloroform (100 mL) at 0 ˚C. The mixture was allowed to warm to room temperature and stir for a further 4 h. After that time, the reaction mixture was poured into ice-water, and extracted with CH_2Cl_2 (3 × 50 mL). The combined organic extracts were washed with brine, dried ($MgSO_4$),

filtered and rotatory evaporated. The residue was purified by column chromatography (Hexanes : EtOAc, 95 : 5) to afford the title compound **54** as two diastereoisomers as colourless oils (combined mass 1.862 g, 92% over 2 steps): Diastereoisomer 1: IR (neat) 1741, 1674, 1487,1400, 1259, 1199, 1168 cm^{-1}; ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.16 (br s, 1H), 6.94 (td, *J* = 9.0, 2.9 Hz, 1H), 6.80 (br d, *J* = 7.2 Hz, 1H), 4.88 (s, 1H), 4.26 – 4.18 (m, 3H), 1.71 (br d, *J* = 6.1 Hz, 3H), 1.42 (s, 3H), 1.39 (s, 3H), 1.26 (t, *J* = 7.1 Hz, 3H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 169.6, 166.6, 160.2 (d, *J* = 248.6 Hz, 1C), 140.9 (d, *J* = 4.8 Hz, 1C), 137.2, 118.5 (d, *J* = 7.7 Hz, 1C), 114.5 (d, *J* = 23.3 Hz, 1C), 109.0 (d, *J* = 22,8 Hz, 1C), 72.2, 61.9, 51.3, 45.0, 31.7, 22.4, 20.3, 14.2; ^{19}F NMR (377 MHz, $CDCl_3$) δ −117.35; MS (ESI) *m/z* 328 ($M+H^+$); HRMS (ESI) *m/z* calc. for $C_{16}H_{20}FNO_3Cl$ 328.1116, found: 328.1118; Diastereoisomer 2: IR (neat) 1740, 1665, 1485,1399, 1257, 1196, 1167 cm^{-1}; ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.17 (br s, 1H), 6.94 (td, *J* =8.7, 2.5 Hz, 1H), 6.81 (br s, 1H), 4.75 (m, 2H), 4.18 (q, *J* = 7.1 Hz, 2H), 1.76 (br d, *J* = 4.5 Hz, 3H), 1.38 (br s, 6H), 1.24 (br unresolved t, 3H); ^{13}C NMR (100 MHz, DMSO) δ 168.8, 166.8, 160.1 (d, *J* = 240.6 Hz, 1C), 141.4, 137.4, 118.8, 114.4 (d, *J* = 24.2 Hz, 1C), 109.1, 73.3, 61.7, 51.3, 44.8, 31.5, 21.8, 21.1, 14.1; ^{19}F NMR (377 MHz, $CDCl_3$) δ −117.27; MS (ESI) *m/z* 328 ($M+H^+$); HRMS (ESI) *m/z* calc. for $C_{16}H_{20}FNO_3Cl$ 328.1116, found: 328.1117.

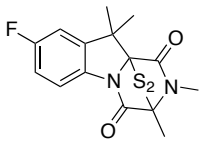
8-Fluoro-2,3,10,10-tetramethyl-2,3,10,10a-tetrahydropyrazino[1,2-*a*]indole-1,4-dione (55)



Compound **55** was prepared as follows:[185] Potassium carbonate (1.209 mg, 9.1 mmol) and an excess of methylamine (31 mL of a 2M solution in THF, 61.0 mmol) were added to a solution of acylated indoline **54** (2.000 g, 6.1 mmol) in MeCN (20 mL) in a sealed tube. After refluxing for 18 h, the reaction mixture was filtered, concentrated and the residue was purified by column chromatography (Hexanes : EtOAc, 1 : 1) to afford the title compound **55** as a mixture of non-separable diastereoisomers as a colorless oil (1.554 g, 92%): IR (neat) 1665, 1611, 1482, 1418, 1400 cm^{-1}; Diastereoisomer 1 (*cis*): ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.99 (dd, *J* = 8.6, 4.7 (J_{H-F}) Hz, 1H), 6.94 (*app*-td, *J* = 8.7 (J_{H-F} and J_{H-H}), 2.7 Hz, 1H), 6.89 (dd, *J* = 8.1 (J_{H-F}), 2.6 Hz, 1H), 4.40 (d, *J* = 1.7 Hz, 1H, *cis coupling*, confirmed by NOESY), 4.17 (qd, *J* = 6.7,

1.7 Hz, 1H, *cis coupling*, confirmed by NOESY), 3.05 (s, 3H), 1.68 (s, 3H), 1.67 (d, J = 6.7 Hz, 3H), 1.16 (s, 3H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 164.6, 163.9, 160.7 (d, J = 244.4 Hz, 1C), 143.0 (d, J = 7.8 Hz, 1C), 135.0, 118.5 (d, J = 8.8 Hz, 1C), 114.2 (d, J = 22.7 Hz, 1C), 109.5 (d, J = 23.8 Hz, 1C), 69.1, 57.6, 45.1, 31.3, 24.6, 23.2, 18.4; ^{19}F NMR (377 MHz, $CDCl_3$) δ −115.58; Diastereoisomer 2 (*trans*): ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.00 (dd, J = 8.7, 4.7 ($J_{H\text{-}F}$) Hz, 1H), 6.94 (*app*-td, J = 8.8 ($J_{H\text{-}F}$ and $J_{H\text{-}H}$), 2.5 Hz, 1H), 6.89 (dd, J = 8.1 ($J_{H\text{-}F}$), 2.7 Hz, 1H), 4.36 (s, 1H), 4.01 (q, J = 7.0 Hz, 1H), 3.01 (s, 3H), 1.68 (s, 3H), 1.54 (d, J = 7.0 Hz, 3H), 1.19 (s, 3H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 164.8, 164.6, 160.6 (d, J = 244.3 Hz, 1C), 143.0 (d, J = 7.4 Hz, 1C), 134.7, 118.1 (d, J = 8.7 Hz, 1C), 114.3 (d, J = 23.4 Hz, 1C), 109.4 (d, J = 23.4 Hz, 1C), 68.3, 60.0, 45.3, 31.7, 24.8, 24.1, 17.5; MS (ESI) m/z 277 ($M+H^+$), 225, 214, 206; HRMS (ESI) *m/z* calc. for $C_{15}H_{18}FN_2O_2$ 277.1352, found: 277.1359.

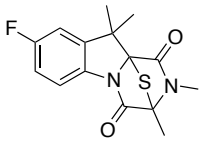
8-Fluoro-2,3,10,10-tetramethyl-2,3-dihydro-1*H*-3,10a-epithiopyrazino[1,2-*a*]indole-1,4(10*H*)-dione 11-sulfide (49)



Compound **49** was prepared as follows:[100,102] NaHMDS (9.2 mL of a 0.6 M solution in PhMe, 5.48 mmol) was added dropwise to a solution of elemental sulfur (468 mg, 14.62 mmol) in dry THF (73 mL) at room temperature. The solution was allowed to stir for 1 min, time after which diketopiperazine **55** (500 mg, 1.83 mmol) in THF (9 mL) was added dropwise. The solution was allowed to stir for 1 min and additional NaHDMS (6.1 mL of a 0.6 M solution in PhMe, 3.66 mmol) was added. The resulting orange / light brown mixture was allowed to stir for 30 min at room temperature. The solution was then quenched with aq. sat. NH_4Cl and extracted with CH_2Cl_2 (3 × 50 mL). The combined organic layers were dried ($MgSO_4$), filtered and concentrated to give a dark green residue which was taken up in a degassed mixture of THF – EtOH (1:1, 36 mL) and cooled to 0 ˚C. Sodium borohydride (1.730 g, 45.70 mmol) was then added in small portions and the resulting mixture was allowed to warm up to room temperature and to stir for 45 min. After that time, the solution was cooled to 0 ˚C and quenched by addition of aq. sat. NH_4Cl. The mixture was extracted with EtOAc (3 ×

50 mL) and the combined organic extracts were then treated with an aqueous solution of KI_3 (1.4 M), until the dark purple colour persisted. The mixture was allowed to stir for 10 min and quenched with 5% aq. $Na_2S_2O_3$ solution until disappearance of the dark purple colour. The resulting mixture was extracted with EtOAc (3 × 50 mL) and the combined organic layers were dried ($MgSO_4$), filtered and rotatory evaporated. The residue was purified by column chromatography (Hexanes : EtOAc, 90 : 10) to afford the racemic title compound **49** as a yellow solid (262 mg, 42%) which was recrystalised from CH_2Cl_2 to afford yellow crystals: m.p. 138 – 140 ˚C; IR (neat) 1686, 1486, 1357, 1316, 1259, 1164 cm^{-1}; ^1H NMR (400 MHz, $CDCl_3$) δ 7.95 (dd, *J* = 8.8, 4.5 ($J_{H\text{-}F}$) Hz, 1H), 7.01 (*app*-td, *J* = 8.8 ($J_{H\text{-}F}$ and $J_{H\text{-}H}$), 2.6 Hz, 1H), 6.95 (dd, *J* = 7.9 ($J_{H\text{-}F}$), 2.6 Hz, 1H), 3.10 (s, 3H), 2.07 (s, 3H), 1.74 (s, 3H), 1.57 (s, 3H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 164.0, 162.2, 161.0 (d, *J* = 247.2 Hz, 1C), 141.2 (d, *J* = 7.7 Hz, 1C), 132.03, 117.0 (d, *J* = 9.4 Hz, 1C), 115.3 (d, *J* = 23.5 Hz, 1C), 109.1 (d, *J* = 24.7 Hz, 1C), 82.8, 74.2, 49.2, 28.5, 27.3, 20.4, 18.1; ^{19}F NMR (377 MHz, $CDCl_3$) δ −115.00; MS (CI) *m/z* 356 $(M+NH_4)^+$, 339 $(M+H)^+$, 275 $(M\text{-}S_2+H)^+$; HRMS (CI) *m/z* calc. for $C_{15}H_{16}FN_2O_2S_2$ 339.0637, found: 339.0631. The two enantiomers could be separated by chiral HPLC (OD+ semiprep column, Hexane : Isopropanol, 1 : 1): First peak: $[\alpha]_D^{20}$ −419 (*c* 1.05, $CHCl_3$), Second peak $[\alpha]_D^{20}$ +300 (*c* 1.05, $CHCl_3$).

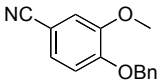
8-Fluoro-2,3,10,10-tetramethyl-2,3-dihydro-1*H*-3,10a-epithiopyrazino[1,2-*a*]indole-1,4(10*H*)-dione (50)



Compound **50** was prepared as follows:[111] PPh_3 (187 mg, 0.71 mmol) was added to a solution of F-ETP **49** (100 mg, 0.29 mmol) in CH_2Cl_2 (10 mL) and the resulting mixture was allowed to stir for 1 h at room temperature. The solvent was rotatory evaporated and the orange residue was purified by column chromatography (Hexanes : EtOAc, 97 : 3 to 93 : 7) to afford the title compound **50** as a white solid (45.5 mg, 50%): m.p. 126 – 128 ˚C; IR (neat) 1720, 1487, 1301, 1259, 1192 cm^{-1}; ^1H NMR (400 MHz, $CDCl_3$) δ 7.48 (dd, *J* = 8.6, 4.5 ($J_{H\text{-}F}$) Hz, 1H), 6.94 (*app*-td, *J* = 8.6 ($J_{H\text{-}F}$ and $J_{H\text{-}H}$), 2.5 Hz, 1H), 6.90 (dd, *J* = 8.1 ($J_{H\text{-}F}$), 2.5 Hz, 1H), 2.96 (s, 3H), 1.83 (s, 3H), 1.74 (s, 3H), 1.46 (s, 3H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 172.5, 171.7, 160.3 (d,

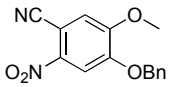
J = 242.3 Hz, 1C), 141.6 (d, *J* = 8.0 Hz, 1C), 134.3, 114.7 (d, *J* = 24.5 Hz, 1C), 114.4 (d, *J* = 9.7 Hz, 1C), 110.0 (d, *J* = 24.1 Hz, 1C), 86.8, 75.1, 43.7, 27.1, 25.8, 25.7, 13.2; MS (ESI) *m/z* 307 $(M+H)^+$, 240; HRMS (ESI) *m/z* calc. for $C_{15}H_{16}FN_2O_2S$ 307.0917, found: 307.0925. The obtained enantiomers could be separated by chiral HPLC (OD+ semiprep column, Hexane : Isopropanol, 90 : 10). First peak: $[\alpha]_D^{20}$ −32.2 (*c* 1.01, $CHCl_3$), Second peak $[\alpha]_D^{20}$ +30.7 (*c* 1.03, $CHCl_3$).

4-Benzyloxy-3-methoxybenzonitrile (58)[226]



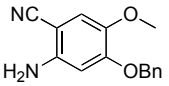
Compound **58** was prepared as follows:[226] To an ice-cooled solution of 4-hydroxy-3-methoxybenzonitrile (**57**, 5.00 g, 33.5 mmol) in DMF (60 mL) were slowly added K_2CO_3 (6.95 g, 50.3 mmol) followed by benzylbromide (4.4 mL, 36.8 mmol). The solution was allowed to stir overnight at room temperature. Sat. aq. NaCl (100 mL) was then added and the resulting precipitate was filtered, washed with water and dried under vacuum to afford the title compound **58** as a white solid (7.74 g, 96%) that was used in the next step without further purification: IR (neat) 2223, 1595, 1512, 1267, 1241, 1132, 984 cm^{-1}; ^1H NMR (400 MHz, $CDCl_3$) δ 7.43 – 7.31 (m, 5H), 7.21 (dd, *J* = 8.2, 1.9 Hz, 1H), 7.09 (d, *J* = 1.9 Hz, 1H), 6.90 (d, *J* = 8.2 Hz, 1H), 5.20 (s, 2H), 3.90 (s, 3H); ^{13}C (100 MHz, $CDCl_3$) δ 151.9, 149.6, 135.7, 128.7 (2C), 128.2, 127.2 (2C), 126.2, 119.2, 114.3, 113.2, 104.1, 70.8, 56.1; MS (CI) *m/z* 257 $(M+NH_4)^+$; HRMS (CI) *m/z* calc. for $C_{15}H_{17}N_2O_2$ 257.1290, found: 257.1292.

4-Benzyloxy-5-methoxy-2-nitrobenzonitrile (59)



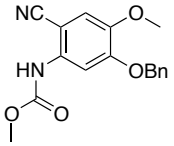
Compound **59** was prepared as follows:[226] Benzonitrile **58** (7.00 g, 29.3 mmol) was dissolved in acetic anhydride (70 mL) and cooled to 0 ˚C. Nitric acid (4.9 mL of 69 wt.% solution in water, 76.7 mmol) was added slowly and the reaction was allowed to stir overnight at room temperature. The mixture was then poured into ice-water and the resulting precipitate was filtered, washed with water and dried under vacuum to afford the title compound **59** (8.04 g, 96%) as a white solid that was used in the next step without further purification: m.p. 128 – 130 ˚C; IR (neat) 2228, 1572, 1533, 1517, 1336, 1288, 1223, 1057 cm^{-1}; ^1H NMR (400 MHz, $CDCl_3$) δ 7.85 (s, 1H), 7.46 – 7.37 (m, 5H), 7.21 (s, 1H), 5.27 (s, 2H), 4.01 (s, 3H); ^{13}C (100 MHz, DMSO) δ 153.4, 150.8, 142.1, 135.4, 128.6 (2C), 128.4, 128.1 (2C), 116.7, 115.7, 109.6, 99.8, 70.8, 57.1; MS (CI) *m/z* 302 $(M+NH_4)^+$; HRMS (CI) *m/z* calc. for $C_{15}H_{16}N_3O_4$ 302.1141, found: 303.1141.

2-Amino-4-benzyloxy-2-methoxybenzonitrile (60)[226]



Compound **60** was prepared as follows:[231] To a solution of nitrobenzonitrile **59** (100 mg, 0.35 mmol) in CH_2Cl_2 (3 mL), tetrabutylammonium chloride monohydrate (68 mg, 0.24 mmol) was added, followed by the slow addition of a solution of sodium dithionite (367 mg, 2.11 mmol) in water (3 mL). The mixture was allowed to stir overnight at room temperature and then basified with 2 N aq. NaOH at 0 ˚C. The aqueous layer was extracted with CH_2Cl_2 (2 × 10 mL). A precipitate was obtained upon treatment of the combined organic layers with 1M HCl in Et_2O. The white solid was filtered and dissolved in water. The aqueous layer was basified with 2N NaOH and extracted with EtOAc. The organic layer was dried ($MgSO_4$) and concentrated to yield the title compound **60** (87 mg, 97%) as a white solid: IR (neat) 3440, 3352, 3243, 2205, 1641, 1620, 1505, 1239, 1119, 985 cm^{-1}; ^1H NMR (400 MHz, DMSO) δ 7.44 – 7.37 (m, 5H), 6.87 (s, 1H), 6.51 (s, 1H), 5.57 (s, 2H), 5.03 (s, 2H), 3.65 (s, 3H); ^{13}C (100 MHz, $CDCl_3$) δ 153.4, 148.2, 140.4, 136.2, 128.5 (2C), 128.1, 127.9 (2C), 116.6, 114.3, 100.1, 83.9, 69.5, 56.2; MS (CI) *m/z* 272 $(M+NH_4)^+$; HRMS (CI) *m/z* calc. for $C_{15}H_{18}N_3O_2$ 272.1399, found: 272.1402.

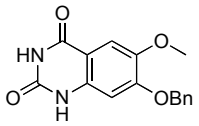
Methyl (5-(benzyloxy)-2-cyano-4-methoxyphenyl)carbamate (61)



Compound **61** was prepared as follows:[226] To a solution of aniline **60** (50 mg, 0.19 mmol) and DIEA (137 μL, 0.79 mmol) in DMA – CH_2Cl_2 (1 mL : 0.5 mL) at 0 ˚C was added methyl chloroformate (23 μL, 0.29 mmol). The mixture was allowed to stir overnight at room temperature. Water (5 mL) was then added and the resulting precipitate was recovered and dried under vacuum to afford the title compound **61** (47 mg, 79%) as a yellow solid that was used in the next step without further purification: IR (neat) 3329, 2223, 1716, 1595, 1510 1220, 1128, 996 cm^{-1}; ^1H NMR (400 MHz, DMSO) δ 9.55 (br s, 1H), 7.47 – 7.37 (m, 5H), 7.35 (s, 1H), 7.19 (s, 1H), 5.14 (s, 2H), 3.80 (s, 3H), 3.67 (s, 3H); ^{13}C (100 MHz, DMSO) δ 154.7, 151.7, 146.6, 136.0, 135.2, 128.5 (2C), 128.2, 128.0 (2C), 117.1, 114.5, 110.8, 109.5, 70.1, 56.1,

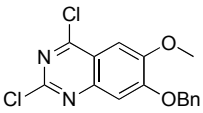
52.1; MS (CI) *m/z* 330 $(M+NH_4)^+$; HRMS (CI) *m/z* calc. for $C_{17}H_{20}N_3O_4$ 330.1454, found: 330.1462.

7-(Benzyloxy)-6-methoxyquinazoline-2,4(1*H*,3*H*)-dione (62)[226]



Compound **62** was prepared as follows:[226] Carbamate **61** (1.088 g, 3.5 mmol) was added to a solution of H_2O_2 (10 mL of 35 wt.% solution in water, 110 mmol) and NaOH (1.28 g, 32 mmol) in ethanol (50 mL) and the resulting mixture was allowed to stir for 4 h under reflux. Water (70 mL) was added and the solution was allowed to cool to room temperature and then acidified with 1 N aq. HCl. The resulting precipitate was recovered, washed with water and dried under vacuum to afford the title compound **62** (609 mg, 64%) as a white solid that was used in the next step without further purification: IR (neat) 1718, 1649, 1618, 1512, 1267, 1107, 838 cm^{-1}; ^1H NMR (400 MHz, DMSO) δ 11.12 (s, 1H), 10.94 (s, 1H), 7.50 – 7.38 (m, 5H), 7.30 (s, 1H), 6.80 (s, 1H), 5.16 (s, 2H), 3.80 (s, 3H); ^{13}C (100 MHz, DMSO) δ 163.1, 153.6, 151.8, 144.7, 138.7, 136.1, 128.5 (2C), 128.1, 128.0 (2C), 107.3, 106.7, 100.1, 70.0, 55.7; MS (ESI) *m/z* 299 $(M+H)^+$; HRMS (ESI) *m/z* calc. for $C_{16}H_{15}N_2O_4$ 299.1032, found: 299.1036.

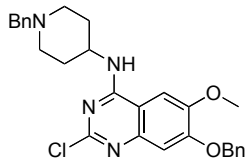
7-(Benzyloxy)-2,4-dichloro-6-methoxyquinazoline (63)[226]



Compound **63** was prepared as follows:[226] A mixture of quinazolinedione **62** (282.2 mg, 0.95 mmol), *N,N*-diethylaniline (166 μL, 1.04 mmol) and $POCl_3$ (2 mL) was allowed to stir under reflux for 3 h and then rotary evaporated. Sat. aq. $NaHCO_3$ (5 mL) was added and the resulting mixture was extracted with chloroform (2 × 10 mL). The combined organic layer was dried ($MgSO_4$), rotary evaporated and purified by column chromatography (EtOAc : Hexanes, 2 : 8) to afford the title compound **63** (248.4 mg, 78%) as a colourless oil: IR (neat) 1611, 1547, 1500, 1424, 1237, 1149, 847 cm^{-1}; ^1H NMR (400 MHz, $CDCl_3$) δ 7.50 – 7.35 (m, 5H), 7.36 (s, 1H), 7.30 (s, 1H), 5.31 (s, 2H), 4.06 (s, 3H); ^{13}C (100 MHz, $CDCl_3$) δ 160.1, 156.8,

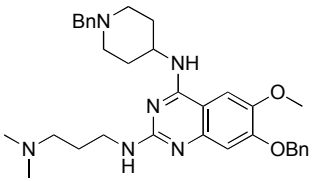
153.4, 151.9, 150.4, 134.8, 128.9 (2C), 128.6, 127.4 (2C), 117.9, 107.6, 102.8, 71.3, 56.5; MS (ESI) *m/z* 335 $(M+H)^+$; HRMS (ESI) *m/z* calc. for $C_{16}H_{13}N_2O_2Cl_2$ 335.0354, found: 335.0370.

7-(Benzyloxy)-N-(1-benzylpiperidin-4-yl)-2-chloro-6-methoxy-quinazolin-4-amine (65)



Compound **65** was prepared as follows:[226] A mixture of dichloroquinazoline **63** (265 mg, 0.79 mmol), 4-amino-1-benzylpiperidine (**64**, 484 μL, 2.37 mmol) and DIEA (206 μL, 1.18 mmol) in THF (2 mL) was allowed to stir overnight at room temperature. The solution was then rotary evaporated and purified by column chromatography (MeOH : CH_2Cl_2, 5 : 95) to afford the title compound **65** (285 mg, 74%) as a white solid: m.p. 108 – 110 ˚C; IR (neat) 1584, 1514, 1452, 1260, 1218 cm^{-1}; ^1H NMR (400 MHz, $CDCl_3$) δ 7.32 – 7.20 (m, 10H), 7.05 (s, 1H), 7.03 (s, 1H), 6.11 (d, *J* = 8.0 Hz, 1H), 5.14 (s, 2H), 4.24 – 4.20 (m, 1H), 3.81 (s, 3H), 3.48 (s, 2H), 2.83 (br *app*-d, *J* = 12.0 Hz, 2H), 2.13 (br *app*-td, *J* = 12.0, 1.5 Hz, 2H), 2.01 (br *app*-d, *J* = 12.0 Hz, 2H), 1.57 (qd, *J* = 12.0, 3.6 Hz, 2H); ^{13}C (100 MHz, $CDCl_3$) δ 159.2, 156.0, 153.8, 149.1, 147.6, 137.9, 135.3, 129.1 (2C), 128.6 (2C), 128.1 (3C), 127.0 (3C), 108.4, 107.0, 100.5, 70.7, 62.8, 56.1, 52.1 (2C), 48.4, 31.9 (2C); MS (ESI) *m/z* 489 $(M+H)^+$; HRMS (ESI) *m/z* calc. for $C_{28}H_{30}N_4O_2Cl$ 489.2057, found: 489.2047.

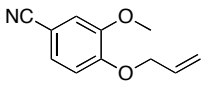
7-(Benzyloxy)-N^4-(1-benzylpiperidin-4-yl)-N^2-(3-(dimethylamino)propyl)-6-methoxyquinazoline-2,4-diamine (67)



Compound **67** was prepared as follows:[226] A mixture of chloroquinazoline **65** (367 mg, 0.75 mmol), *N,N*-dimethyl-1,3-propanediamine (**66**, 189 μL, 1.50 mmol) and 4 M HCl in dioxane (375 μL, 1.50 mmol) in isopropanol (4 mL) was allowed to

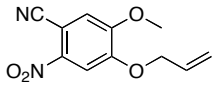
stir for 20 minutes at 160 ˚C in a microwave reactor. The solution was then diluted with 10 mL of CH_2Cl_2 and washed with sat. aq. $NaHCO_3$ (10 mL). The organic layer was dried ($MgSO_4$), rotary evaporated and purified by column chromatography (CH_2Cl_2 : 7M NH_3 in MeOH, 95 : 5) to afford the title compound **67** (124 mg, 30%) as an orange oil: IR (neat) 1581, 1505, 1450, 1249, 738 cm^{-1}; ^1H NMR (400 MHz, $CDCl_3$) δ 7.45 – 7.27 (m, 10H), 6.91 (s, 1H), 6.77 (s, 1H), 5.20 (br s, 2H+1H), 4.18 – 4.14 (m, 1H), 3.90 (s, 3H), 3.54 (s, 2H), 3.48 (*app*-q, *J* = 6.9 Hz, 2H), 2.86 (br *app*-d, *J* = 12.0 Hz, 2H), 2.37 (t, *J* = 6.9 Hz, 2H), 2.22 (s, 6H), 2.19 (br *app*-td, *J* = 12.0, 2.6 Hz, 2H), 2.06 (br *app*-dd, *J* = 12.0, 4.4 Hz, 2H), 1.77 (quint, *J* = 6.9 Hz, 2H), 1.60 (qd, *J* = 12.0, 4.4 Hz, 2H); ^{13}C (100 MHz, $CDCl_3$) δ 159.0, 158.6, 153.2, 148.3, 145.3, 138.2, 136.0, 128.9 (2C), 128.3 (2C), 128.0 (2C), 127.7, 127.0 (2C), 126.8, 106.5, 103.8, 102.1, 70.2, 62.9, 57.7, 56.1, 52.5 (2C), 48.2, 45.4 (2C), 40.2, 31.9 (2C), 27.7; MS (ESI) *m/z* 555 $(M+H)^+$; HRMS (ESI) *m/z* calc. for $C_{33}H_{43}N_6O_2$ 555.3448, found: 555.3441.

4-Allyloxy-3-methoxybenzonitrile (70)



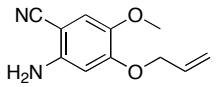
Compound **70** was prepared as follows:[226] To an ice-cooled solution of benzonitrile **57** (2.00 g, 13.4 mmol) in acetone (150 mL) were slowly added K_2CO_3 (2.22 g, 16.1 mmol), followed by allyl bromide (2.32 mL, 26.8 mmol). The solution was refluxed for 3 h, time after which it was quenched with water (100 mL) and extracted with Et_2O (2 × 100 mL). The combined organic layers were dried ($MgSO_4$), filtered and rotary evaporated to afford the title compound **70** (2.33 g, 92%) as a white solid that was used in the next step without further purification: m.p. 48 – 50 ˚C; IR (neat) 2223, 1595, 1512, 1241, 1135, 1003 cm^{-1}; ^1H NMR (400 MHz, $CDCl_3$) δ 7.24 (dd, *J* = 8.4, 1.6 Hz, 1H), 7.08 (d, *J* = 1.6 Hz, 1H), 6.89 (d, *J* = 8.4 Hz, 1H), 6.10 – 6.00 (m, 1H), 5.42 (*app*-dq, *J* = 17.4, 1.3 Hz, 1H), 5.36 (*app*-dq, *J* = 10.4, 1.3 Hz, 1H), 4.66 (*app*-dt, *J* = 5.4, 1.3 Hz, 2H), 3.89 (s, 3H); ^{13}C (100 MHz, $CDCl_3$) δ 151.9, 149.5, 132.0, 126.3, 119.3, 118.9, 114.2, 112.8, 104.0, 69.8, 56.2; MS (CI) *m/z* 207 $(M+NH_4)^+$; HRMS (CI) *m/z* calc. for $C_{11}H_{15}N_2O_2$ 207.1134, found: 207.1133.

4-Allylyloxy-5-methoxy-2-nitrobenzonitrile (71)



Compound **71** was prepared as follows:[226] Nitric acid (1.4 mL of 69 wt% solution in water, 21.14 mmol) was slowly added to a solution of benzonitrile **70** (2.00 g, 10.57 mmol) in acetic anhydride (50 mL) at 0 ˚C. The reaction was allowed to stir for 1 h at room temperature, time after which the mixture was poured into ice-water and the resulting precipitate was filtered, washed with water and dried under vacuum to afford the title compound **71** (2.216 g, 89%) as a pale yellow solid that was used in the next step without further purification: m.p. 156 – 158 ˚C; IR (neat) 2224, 1572, 1507, 1334, 1224, 1058 cm^{-1}; ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.79 (s, 1H), 7.21 (s, 1H), 6.11 – 6.01 (m, 1H), 5.49 (*app*-dq, *J* = 17.2, 1.2 Hz, 1H), 5.41 (*app*-dq, *J* = 10.4, 1.2 Hz, 1H), 4.75 (*app*-dt, *J* = 5.6, 1.2 Hz, 2H), 4.01 (s, 3H); ^{13}C (100 MHz, $CDCl_3$) δ 153.6, 151.1, 142.6, 130.8, 120.0, 115.6, 115.5, 109.2, 100.9, 70.5, 56.9; MS (CI) *m/z* 252 $(M+NH_4)^+$; HRMS (CI) *m/z* calc. for $C_{11}H_{14}N_3O_4$ 252.0984, found: 252.0982.

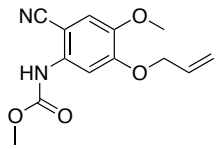
2-Amino-4-allyloxy-2-methoxybenzonitrile (72)



Compound **72** was prepared as follows:[231] To a solution of nitrobenzonitrile **71** (2.21 g, 9.39 mmol) in CH_2Cl_2 (50 mL) was added tetrabutylammonium bromide (2.1 g, 6.57 mmol), followed by the slow addition of a solution of sodium dithionite (9.8 g, 56.3 mmol) in water (50 mL). The mixture was allowed to stir for 3 h at room temperature and then basified with 2 N aq. NaOH at 0 ˚C. The organic layer was recovered and the aqueous layer extracted with CH_2Cl_2 (2 × 100 mL). A precipitate was obtained upon treatment of the combined organic layer with 1M HCl in Et_2O. The white solid was filtered and dissolved in water. The aqueous layer was basified with 2N NaOH and extracted with EtOAc (2 × 100 mL). The organic layer was dried ($MgSO_4$) and rotary evaporated to give the title compound **72** (1.32 g, 69%) as a white solid that was used in the next step without further purification: m.p. 50 – 52 ˚C; IR (neat) 3438, 3355, 3248, 2200, 1646, 1507, 1236 cm^{-1}; ^{1}H NMR (400 MHz, $CDCl_3$) δ 6.76 (s, 1H), 6.23 (s, 1H), 6.06 – 5.96 (m, 1H), 5.37 (*app*-dq, *J* = 17.2, 1.5 Hz, 1H), 5.29 (*app*-dq, *J* = 10.4, 1.5 Hz, 1H), 4.55 (*app*-dt, *J* = 5.6, 1.5 Hz, 2H), 4.20

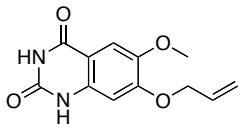
(br s, 2H), 3.76 (s, 3H); ^{13}C (100 MHz, $CDCl_3$) δ 153.4, 146.0, 141.9, 131.9, 118.6, 118.1, 113.6, 100.3, 86.3, 69.5, 56.4; MS (ESI) *m/z* 205 $(M+H)^+$, 246 $(M+MeCN+H)^+$; HRMS (ESI) *m/z* calc. for $C_{11}H_{13}N_2O_2$ 205.0977 found: 205.0974.

Methyl (5-(allyloxy)-2-cyano-4-methoxyphenyl)carbamate (73)



Compound **73** was prepared as follows:[226] To a solution of aniline **72** (1.382 g, 6.77 mmol) and DIEA (4.7 mL, 27.1 mmol) in DMA – CH_2Cl_2 (10 mL : 5 mL) at 0 ˚C was added methyl chloroformate (784 μL, 10.15 mmol). The mixture was allowed to stir overnight at room temperature. Water (30 mL) was added and the resulting precipitate was recovered, washed with water and dried under vacuum. It was re-dissolved in CH_2Cl_2 and washed with 1M aq. HCl, then with brine, dried ($MgSO_4$) and rotary evaporated to afford the title compound **73** (1.415 g, 80%) as a white solid that was used in the next step without further purification: IR (neat) 2223, 1734, 1600, 1524, 1213 cm^{-1}; ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.85 (br s, 1H), 7.03 (br s, 1H), 6.93 (s, 1H), 6.12 – 6.02 (m, 1H), 5.46 (*app*-dq, *J* = 17.2, 1.3 Hz, 1H), 5.36 (*app*-dq, *J* = 10.4, 1.3 Hz, 1H), 4.69 (*app*-dt, *J* = 5.4, 1.3 Hz, 2H), 3.86 (s, 3H), 3.81 (s, 3H); ^{13}C (100 MHz, $CDCl_3$) δ 153.6, 152.8, 145.1, 136.4, 131.7, 119.3, 116.7, 113.1, 104.4, 91.5, 69.9, 56.3, 52.8; MS (ESI) *m/z* 263 $(M+H)^+$; HRMS (ESI) *m/z* calc. for $C_{13}H_{15}N_2O_4$ 263.1032 found: 263.1033.

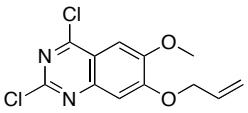
7-(Allyloxy)-6-methoxyquinazoline-2,4(1*H*,3*H*)-dione (74)



Compound **74** was prepared as follows:[226] Carbamate **73** (1.415 g, 5.4 mmol) was added to a solution of H_2O_2 (10 mL of 35 wt% solution in water, 110 mmol) and NaOH (2.13 g, 54 mmol) in ethanol (100 mL) and the resulting mixture was allowed to stir for 4 h under reflux. Water (100 mL) was added and the solution was allowed to cool to room temperature, then acidified with 1 N aq. HCl. The resulting precipitate was recovered, washed with water and dried to afford the title compound **74** (932 mg, 69%) as a white solid that was used in the next step without further purification: m.p.

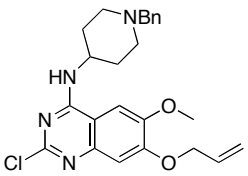
202 – 204 ˚C; IR (neat) 1709, 1649, 1621, 1267, 1206 cm^{-1}; ^{1}H NMR (400 MHz, DMSO) δ 11.08 (br s, 1H), 10.90 (br s, 1H), 7.27 (s, 1H), 6.68 (s, 1H), 6.10 – 6.00 (m, 1H), 5.44 (*app*-dq, *J* = 17.2, 1.5 Hz, 1H), 5.32 (*app*-dq, *J* = 10.4, 1.5 Hz, 1H), 4.60 (*app*-dt, *J* = 5.4, 1.5 Hz, 2H), 3.79 (s, 3H); ^{13}C (100 MHz, DMSO) δ 162.4, 153.7, 150.4, 145.1, 136.4, 132.6, 118.6, 107.4, 106.3, 98.9, 69.0, 55.7; MS (ESI) *m/z* 249 $(M+H)^+$, 290 $(M+MeCN+H)^+$; HRMS (ESI) *m/z* calc. for $C_{12}H_{13}N_2O_4$ 249.0875, found: 249.0870.

7-(Allyloxy)-2,4-dichloro-6-methoxyquinazoline (69)



Compound **69** was prepared as follows:[226] A mixture of quinazolinedione **74** (920 mg, 3.69 mmol), *N,N*-diethylaniline (649 μL, 4.07 mmol) and $POCl_3$ (20 mL) was allowed to stir under reflux for 1.5 h. The solution was slowly poured into a sat. aq. $NaHCO_3$ at 0 ˚C. The organic layer was separated and the aqueous layer was extracted with chloroform (2 × 50 mL). The combined organic layers were dried ($MgSO_4$), rotary evaporated and purified by column chromatography (EtOAc : Hexanes, 3 : 7) to afford the title compound **69** (908 mg, 86%) as a white solid: m.p. 92 – 94 ˚C; IR (neat) 1614, 1549, 1500, 1415, 1237, 1147 cm^{-1}; ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.30 (s, 1H), 7.20 (s, 1H), 6.13 – 6.03 (m, 1H), 5.52 (*app*-dq, *J* = 17.2, 1.2 Hz, 1H), 5.37 (*app*-dq, *J* = 10.4, 1.2 Hz, 1H), 4.75 (*app*-dt, *J* = 5.6, 1.2 Hz, 2H), 4.03 (s, 3H); ^{13}C (100 MHz, $CDCl_3$) δ 159.9, 156.6, 153.4, 151.7, 150.3, 131.1, 119.6, 117.7, 107.2, 102.6, 70.2, 56.4; MS (ESI) *m/z* 285 $(M+H)^+$; HRMS (ESI) *m/z* calc. for $C_{12}H_{11}N_2O_2Cl_2$ 285.0198, found: 285.0197.

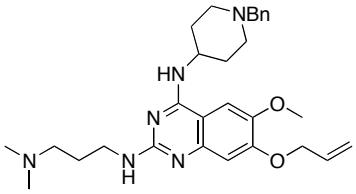
7-(Allyloxy)-*N*-(1-benzylpiperidin-4-yl)-2-chloro-6-methoxy-quinazolin-4-amine (75)



Compound **75** was prepared as follows:[226] A mixture of dichloroquinazoline **69** (102 mg, 0.36 mmol), 4-amino-1-benzylpiperidine (**64**, 214 μL, 1.07 mmol) and

DIEA (92 μL, 0.536 mmol) in THF (2 mL) was allowed to stir for 8 h at room temperature. The solution was then rotary evaporated and purified by column chromatography (MeOH : CH_2Cl_2, 4 : 96) to afford the title compound **75** (133.5 mg, 85%) as a white solid: m.p. 90 – 92 ˚C; IR (neat) 3313, 1584, 1510, 1452, 1253, 961 cm^{-1}; ^1H NMR (400 MHz, $CDCl_3$) δ 7.33 – 7.23 (m, 5H), 7.10 (s, 1H), 6.84 (s, 1H), 6.11 – 6.01 (m, 1H), 5.50 (d, *J* = 7.8 Hz, 1H), 5.44 (*app*-dq, *J* = 17.3, 1.2 Hz, 1H), 5.32 (*app*-dq, *J* = 10.5, 1.2 Hz, 1H), 4.67 (*app*-dt, *J* = 5.4, 1.2 Hz, 2H), 4.33 – 4.23 (m, 1H), 3.96 (s, 3H), 3.55 (s, 2H), 2.90 (br *app*-d, *J* = 12.0 Hz, 2H), 2.23 (*app*-dt, *J* = 12.0, 1.5 Hz, 2H), 2.12 (br *app*-d, *J* = 12.0 Hz, 2H), 1.64 (*app*-dq, *J* = 12.0, 3.5 Hz, 2H); ^{13}C (100 MHz, $CDCl_3$) δ 159.1, 156.1, 153.8, 149.2, 147.8, 138.2, 131.7, 129.1 (2C), 128.2 (2C), 127.1, 119.1, 108.4, 106.7, 99.9, 69.7, 62.9, 56.3, 52.2 (2C), 48.3, 32.2 (2C); MS (ESI) *m/z* 439 $(M+H)^+$; HRMS (ESI) *m/z* calc. for $C_{24}H_{28}N_4O_2Cl$ 439.1901, found: 439.1908.

7-(allyloxy)-N^4-(1-benzylpiperidin-4-yl)-N^2-(3-(dimethylamino)propyl)-6-methoxyquinazoline-2,4-diamine (76)



Procedure 1:[226]

A mixture of chloroquinazoline **75** (478 mg, 1.09 mmol) and *N,N*-dimethyl-1,3-propanediamine (**66**, 1.5 mL, 11.4 mmol) in MeCN (1.5 mL) was allowed to stir for 3 days at reflux. The solution was then diluted with CH_2Cl_2 (10 mL) and washed with sat. aq. $NaHCO_3$ (10 mL). The organic layer was separated and the aqueous layer was extracted with CH_2Cl_2 (2 × 50 mL) and the combined organic layers were washed with brine, dried ($MgSO_4$), rotary evaporated and purified by column chromatography (CH_2Cl_2 : 7M NH_3 in MeOH; 9 : 1) to afford the title compound **75** (426 mg, 77%).

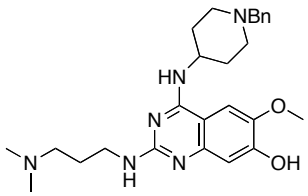
Procedure 2

Chloroquinazoline **75** (528 mg, 1.2 mmol) was refluxed in *N,N*-dimethyl-1,3-propanediamine (**66**, 10 mL, excess) for 16 h. The crude reaction mixture was then

purified by column chromatography (CH_2Cl_2 : MeOH, 95 : 5, with 1% NH_3 aq. solution) to afford the title compound **76** (274 mg, 45%).

Compound **76** was obtained as a light yellow solid: m.p. 56 – 58 ˚C; IR (neat) 1579, 1498, 1454, 1246, 979 cm^{-1}; ^1H NMR (400 MHz, $CDCl_3$) δ 7.33 – 7.20 (m, 5H), 6.85 (s, 1H), 6.75 (s, 1H), 6.15 – 6.06 (m, 1H), 5.43 (*app*-dq, *J* = 17.3, 1.2 Hz, 1H), 5.31 (*app*-dq, *J* = 10.5, 1.2 Hz, 1H), 5.13 (br d, *J* = 7.1 Hz, 1H), 5.09 (br s, 1H), 4.65 (*app*-dt, *J* = 5.4, 1.2 Hz, 2H), 4.22 – 4.12 (m, 1H), 3.90 (s, 3H), 3.54 (s, 2H), 3.48 (*app*-q, *J* = 7.0 Hz, 2H), 2.89 (br *app*-d, *J* = 12.0 Hz, 2H), 2.37 (t, *J* = 7.0 Hz, 2H), 2.22 (s, 6H), 2.19 (br *app*-td, *J* = 12.0, 2.1 Hz, 2H), 2.11 (br *app*-dd, *J* = 12.0, 3.2 Hz, 2H), 1.78 (quint, *J* = 7.0 Hz, 2H), 1.60 (*app*-dq, *J* = 12.0, 3.2 Hz, 2H); ^{13}C (100 MHz, $CDCl_3$) δ 159.0, 158.5, 153.5, 148.3, 145.5, 138.4, 132.3, 129.1 (2C), 128.2 (2C), 127.0, 118.6, 106.5, 103.5, 101.3, 69.5, 63.1, 57.8, 56.5, 52.5 (2C), 48.1, 45.6 (2C), 40.2, 32.3 (2C), 27.9; MS (ESI) *m/z* 505 $(M+H)^+$; HRMS (ESI) *m/z* calc. for $C_{29}H_{41}N_6O_2$ 505.3291, found: 505.3292.

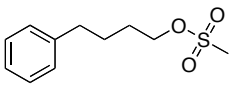
4-(1-benzylpiperidin-4-ylamino)-2-(3-(dimethylamino)propylamino)-6-methoxyquinazolin-7-ol (68)



Compound **68** was prepared as follows:[239] $Pd(PPh_3)_4$ (34 mg, 0.029 mmol) was added to a solution of allyl protected quinazoline **76** (294 mg, 0.580 mmol) in MeOH (50 mL). After 5 min, K_2CO_3 (400 mg, 2.900 mmol) was added and the solution was allowed to stir for 4 h at 50 ˚C. The solvent was rotatory evaporated to afford a brown oil which was subsequently treated with 1 N aq. HCl (50 mL). The solution was then basified with 2 N aq. NaOH and extracted with CH_2Cl_2 (2 × 30 mL). The organic layers were then combined, dried ($MgSO_4$) and rotary evaporated to afford the title compound **68** (crude, 286 mg). The obtained product was too polar to be purified by column chromatography and was thus used directly in the next step without further purification: ^1H NMR (400 MHz, MeOD) δ 7.45 (br s, 1H), 7.36 – 7.27 (m, 5H), 6.50 (br s, 1H), 4.26 – 4.20 (m, 1H), 3.98 (br s, 3H), 3.60 (s, 2H), 3.48 (br t, *J* = 7.0 Hz, 2H), 3.01 (br *app*-d, *J* = 12.0 Hz, 2H), 2.51 (*app*-t, *J* = 7.0 Hz, 2H), 2.34 (s, 6H), 2.19

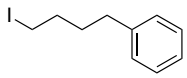
(br *app*-dt, J = 12.0, 1.8 Hz, 2H), 2.03 (br *app*-d, J = 12.0 Hz, 2H), 1.85 (quint, J = 7.0 Hz, 2H), 1.78 (*app*-qd, J = 12.0, 3.3 Hz, 2H); MS (ESI) *m/z* 465 $(M+H)^+$; HRMS (ESI) *m/z* calc. for $C_{26}H_{37}N_6O_2$ 465.2978 found: 465.2974.

4-phenylbutyl methane sulfonate (77)[244]



To a solution of 4-phenylbutan-1-ol (**78**, 200 μL, 1.35 mmol) in CH_2Cl_2 (5 mL) at 0 ˚C, was added Et_3N (282 μL, 2.03 mmol), followed by methanesulfonyl chloride (109 μL, 1.41 mmol). The resulting mixture was allowed to stir for 2 h and allowed to warm to room temperature. Sat. aq. NH_4Cl (5 mL) was added and the organic layer was rotatory evaporated. The aqueous layer was extracted with Et_2O (2 × 50 mL) and the combined organic layers were dried ($MgSO_4$) and rotary evaporated to afford the title compound **77** as a colourless oil (299 mg, 97%) that was used in the next step without further purification: ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.31 – 7.27 (m, 2H), 7.22 – 7.16 (m, 3H), 4.23 (t, J = 6.2 Hz, 2H), 2.98 (s, 3H), 2.66 (t, J = 6.7 Hz, 2H), 1.82 – 1.71 (m, 4H); ^{13}C (100 MHz, $CDCl_3$) δ 141.5, 128.4 (2C), 128.4 (2C), 126.0, 69.8, 37.3, 35.1, 28.6, 27.1; MS (CI) *m/z* 246 $(M+NH_4)^+$; HRMS (CI) *m/z* calc. for $C_{11}H_{20}NO_3S$ 246.1164, found: 246.1152.

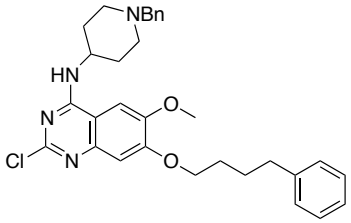
(4-Iodobutyl)benzene (79)[245]



Compound **79** was prepared as follows:[245] To a solution of NaI (284 mg, 1.927 mmol) in acetone (3 mL) was added sulfonate **77** (100 mg, 0.438 mmol). The resulting mixture was allowed to stir at room temperature for 16 h and then rotary evaporated. The residue was taken up in water (20 mL) and extracted with Et_2O (2 × 30 mL) and the combined organic layers were dried ($MgSO_4$) and rotary evaporated to afford the title compound **79** as a yellow oil that was used in the next step without further purification (103 mg, 90%). IR (neat) 1604, 1495, 1452, 1206, 1166 cm^{-1}; ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.32 – 7.28 (m, 2H), 7.22 – 7.18 (m, 3H), 3.21 (t, J = 7.2 Hz, 2H), 2.65 (t, J = 7.3 Hz, 2H), 1.91 – 1.84 (m, 2H), 1.79 – 1.71 (m, 2H); ^{13}C (100 MHz,

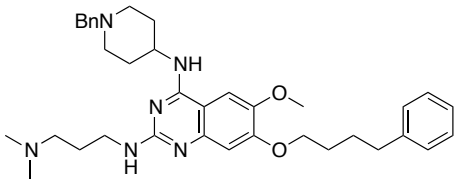
$CDCl_3$) δ 141.8, 128.4 (4C), 125.9, 34.7, 32.9, 32.1, 6.7; MS (EI) *m/z* 260 (M^+); HRMS (EI) *m/z* calc. for $C_{10}H_{13}I$ 260.0062, found: 260.0054.

_N_-(1-benzylpiperidin-4-yl)-2-chloro-6-methoxy-7-(4-phenylbutoxy)quinazolin-4-amine (81)



TBAB (4 mg, 0.013 mmol) and K_2CO_3 (26 mg, 0.187 mmol) were added to a solution of phenol **80** (50 mg, 0.125 mmol) in acetone (5 mL). Methane sulfonate **77** (34 mg, 0.138 mmol) was then added and the mixture was allowed to stir overnight under reflux. NaI (18.7 mg, 0.125 mmol) was then added and the reaction was allowed to stir under reflux for an extra 20 h. After this time, the mixture was filtered, washed with CH_2Cl_2 and rotary evaporated. Purified by column chromatography (CH_2Cl_2 : MeOH, 95 : 5) afforded the title compound **81** as a colourless oil (30 mg, 45%): IR (neat) 1621, 1581, 1508, 1439, 1250, 1219 cm^{-1}; ^1H NMR (400 MHz, $CDCl_3$) δ 7.33 – 7.18 (m, 10H), 7.08 (s, 1H), 6.87 (s, 1H), 5.59 (d, *J* = 7.9 Hz, 1H), 4.35 – 4.25 (m, 1H), 4.09 (t, *J* = 6.5 Hz, 2H), 3.94 (s, 3H), 3.58 (s, 2H), 2.93 (br *app*-d, *J* = 12.0 Hz, 2H), 2.68 (t, *J* = 7.5 Hz, 2H), 2.26 (*app*-dt, *J* = 12.0, 1.5 Hz, 2H), 2.12 (br *app*-d, *J* = 12.0 Hz, 2H), 1.95 – 1.88 (m, 2H), 1.84 – 1.76 (m, 2H), 1.69 (*app*-dq, *J* = 12.0, 3.5 Hz, 2H); ^{13}C (100 MHz, $CDCl_3$) δ 159.1, 156.0, 154.5, 149.2, 147.9, 141.9, 137.5, 129.2 (2C), 128.4 (2C), 128.3 (2C), 128.2 (2C), 127.3, 125.8, 107.8, 106.5, 99.9, 68.9, 62.8, 56.4, 52.2 (2C), 48.1, 35.4, 31.9 (2C), 28.2, 27.7; MS (ESI) *m/z* 531 $(M+H)^+$; HRMS (ESI) *m/z* calc. for $C_{31}H_{36}N_4O_2Cl$ 531.2527 found: 531.2528.

N^4-(1-benzylpiperidin-4-yl)-N^2-(3-(dimethylamino)propyl)-6-methoxy-7-(4-phenylbutoxy)quinazoline-2,4-diamine (56)



Procedure 1 : From phenol **68** (with Cs_2CO_3 in DMF)

Crude phenol **68 (**50 mg, 0.11 mmol) was dissolved in degassed DMF (2 mL). Cs_2CO_3 (52 mg, 0.16 mmol) was added followed by methane sulfonate **77** (27 mg, 0.12 mmol). The resulting solution was allowed to stir at room temperature for 16 h then heated to 80 ˚C for 1 h. The mixture was then diluted with water (10 mL) and extracted with CH_2Cl_2 (3 × 20 mL) and the combined organic layers were washed with water (3 × 20 mL), dried ($MgSO_4$) and rotary evaporated. Purification by column chromatography (CH_2Cl_2 : 7M NH_3 in MeOH, 95 : 5) afforded the title compound **56** (1.2 mg, 2%).

Procedure 2 : From phenol **68** (with *t*BuOK in acetone)

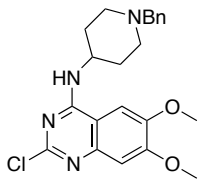
To a solution of crude phenol **68** (35 mg, 0.075 mmol) in acetone (3 mL) was added *t*BuOK (8.5 mg, 0.075 mmol) and methane sulfonate **77** (17.1 mg, 0.075 mmol) and the resulting solution was allowed to stir at 40 ˚C for 16 h. The mixture was then diluted with water (5 mL) and extracted with CH_2Cl_2 (3 × 10 mL). The combined organic layers were washed with water (3 × 20 mL), dried ($MgSO_4$) and rotary evaporated. Purification by column chromatography (CH_2Cl_2 : 7M NH_3 in MeOH, 95 : 5) afforded the title compound **56** (4 mg, 9%).

Procedure 3: From chloroquinazoline (**81**)

N,N-dimethyl-1,3-propanediamine (**66**, 36 μL, 0.280 mmol) was added to a solution of chloroquinazoline **81** (30 mg, 0.056 mmol) in toluene (2 mL). The resulting mixture was heated under reflux for 16 h, then diluted with CH_2Cl_2 (3 mL) and then washed with sat. aq. $NaHCO_3$. The organic layer was separated and the aqueous layer was extracted with CH_2Cl_2 (3 × 3 mL). The combined organic layers were dried ($MgSO_4$), rotary evaporated and purified by column chromatography (CH_2Cl_2 : 7M NH_3 in MeOH, 90 : 10) to afford the title compound **56** (12 mg, 36%).

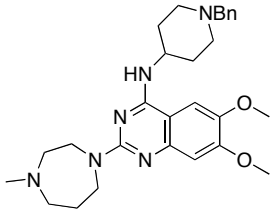
Compound **56** was obtained as a colourless oil: IR (neat) 2154, 1645, 1608, 1585, 1514, 1455, 1274 cm^{-1}; ^1H NMR (400 MHz, $CDCl_3$) δ 7.33 – 7.18 (m, 10H), 6.86 (s, 1H), 6.77 (br s, 1H), 5.41 (br s, 1H), 4.18 – 4.08 (m, 1H), 4.09 (t, *J* = 6.5 Hz, 2H), 3.91 (s, 3H), 3.55 (s, 2H), 3.51 – 3.47 (m, 2H), 2.91 (br *app*-d, *J* = 12.0 Hz, 2H), 2.69 (t, *J* = 7.4 Hz, 2H), 2.39 (t, *J* = 7.0 Hz, 2H), 2.24 (s, 6H), 2.19 (*app*-dt, *J* = 12.0, 1.5 Hz, 2H), 2.12 (br *app*-d, *J* = 12.0 Hz, 2H), 1.95 – 1.88 (m, 2H), 1.84 – 1.76 (m, 4H), 1.64 (*app*-dq, *J* = 12.0, 3.5 Hz, 2H); ^{13}C (100 MHz, $CDCl_3$) δ 158.7, 155.3, 155.2, 146.9 (2C), 141.9, 137.9, 129.1 (2C), 128.4 (2C), 128.3 (2C), 128.2 (2C), 127.1, 125.8, 103.5, 102.1, 99.9, 69.2, 62.9, 57.1, 57.0, 52.4 (2C), 49.4, 45.2 (2C), 39.6, 35.4 (2C), 31.4, 28.3, 27.6, 27.0; MS (ESI) *m/z* 597 $(M+H)^+$; HRMS (ESI) *m/z* calc. for $C_{36}H_{49}N_6O_2$ 597.3917 found: 597.3911.

_N_-(1-benzylpiperidin-4-yl)-2-chloro-6,7-dimethoxyquinazolin-4-amine (84)[226]



Compound **84** was prepared as follows:[226] A mixture of dichloroquinazoline **83** (300 mg, 1.15 mmol), 4-amino-1-benzylpiperidine (**64**, 0.700 mL, 3.45 mmol) and DIEA (0.300 mL, 1.72 mmol) in THF (5 mL) was allowed to stir at room temperature for 16 h. The solution was concentrated and purified by column chromatography (MeOH : CH_2Cl_2, 10 : 90) to afford the title compound **84** (478 mg, quant.) as a white solid: m.p. 116 – 118 ˚C; IR (neat) 2936, 1739, 1579, 1508, 1454, 1428, 1366 cm^{-1}; ^1H NMR (400 MHz, $CDCl_3$) δ 7.33 – 7.25 (m, 5H), 7.11 (s, 1H), 6.81 (s, 1H), 5.44 (d, *J* = 7.8 Hz, 1H), 4.33 – 4.23 (m, 1H), 3.97 (s, 3H), 3.95 (s, 3H), 3.55 (s, 2H), 2.90 (br *app*-d, *J* = 12.0 Hz, 2H), 2.23 (br *app*-td, *J* = 12.0, 1.6 Hz, 2H), 2.12 (br *app*-d, *J* = 12.0 Hz, 2H), 1.63 (br *app*-qd, *J* = 12.0, 3.6 Hz, 2H); ^{13}C (100 MHz, $CDCl_3$) δ 159.1, 156.2, 154.9, 149.0, 148.1, 138.1, 129.1 (2C), 128.2 (2C), 127.1, 107.3, 106.7, 99.6, 62.9, 56.3, 56.2, 52.2 (2C), 48.3, 32.2 (2C); MS (ESI) *m/z* 413 $(M+H)^+$; HRMS (ESI) *m/z* calc. for $C_{22}H_{26}N_4O_2Cl$ 413.1744 found: 413.1757.

***N*-(1-benzylpiperidin-4-yl)-6,7-dimethoxy-2-(4-methyl-1,4-diazepan-1-yl)quinazolin-4-amine) (2, BIX-01294)[226,227]**



Compound **2** was prepared as follows:[226] 1-Methyl-1,4-diazepane (**85**, 150 μL, 1.210 mmol) was added to a solution of chloroquinazoline **84** (100 mg, 0.242 mmol) in toluene (2 mL). The reaction mixture was allowed to stir under reflux for 16 h, diluted with CH_2Cl_2 (3 mL) and washed with sat. aq. $NaHCO_3$. The organic layer was separated and the aqueous layer was extracted with CH_2Cl_2 (3 × 3 mL). The combined organic layers were dried ($MgSO_4$), rotary evaporated and purified by column chromatography (CH_2Cl_2 : 7M NH_3 in MeOH, 90 : 10) to afford the title compound **2** as a yellow oil (75 mg, 63%): IR (neat) 1573, 1493, 1423, 1213 cm^{-1}; ^1H NMR (400 MHz, $CDCl_3$) δ 7.35 – 7.25 (m, 5H), 6.91 (s, 1H), 6.78 (s, 1H), 5.15 (d, *J* = 7.0 Hz, 1H), 4.16 – 4.09 (m, 1H), 4.01 – 3.98 (m, 2H), 3.94 (s, 3H), 3.92 (s, 3H), 3.89 (t, *J* = 6.5 Hz, 2H), 3.56 (s, 2H), 2.92 (br *app*-d, *J* = 12.0 Hz, 2H), 2.72 – 2.70 (m, 2H), 2.59 – 2.57 (m, 2H), 2.38 (s, 3H), 2.23 – 2.13 (m, 4H), 2.06 – 2.00 (m, 2H), 1.63 (br *app*-qd, *J* = 12.0, 3.6 Hz, 2H); ^{13}C (100 MHz, $CDCl_3$) δ 158.6, 157.9, 154.1, 149.5, 144.8, 138.4, 128.9 (2C), 128.1 (2C), 126.9, 105.9, 102.7, 100.7, 63.0, 58.9, 57.3, 56.2, 55.8, 52.5 (2C), 48.3, 46.7, 45.8, 45.7, 32.1 (2C), 27.8; MS (ESI) *m/z* 491 $(M+H)^+$; HRMS (ESI) *m/z* calc. for $C_{28}H_{39}N_6O_2$ 491.3134 found: 491.3140.

Chiroptical measurements

ECD spectroscopy

The experimental spectra were recorded on an Applied Photophysics Chirascan spectrometer in dichloromethane (Temperature: 22 °C, Wavelength: from 180 to 260 nm, Step: 0.5 nm, Band width: 1 nm, Time per point: 1 s)

VCD spectroscopy

VCD measurements were performed at the European Centre for Chirality by Prof. Wouter Herrebout.

Solutions of the enantiomerically pure compounds of interest and of the corresponding racemates were prepared in $CDCl_3$ (99.98%, Aldrich). All spectra were recorded using a demountable liquid cell equipped with BaF_2 windows and 100 μm spacers. All spectra were recorded at 4 cm^{-1} resolution for approximately 13 h, accumulating 40 000 scans. Background corrections for VCD were introduced by subtracting the spectra of the enantiomerically pure compounds and those obtained for the corresponding racemates.

Computational simulations

All calculations were performed by Prof. Henry Rzepa and Ya-Pei Lo (Chemistry Department, Imperial College London) using the procedures implemented in the Gaussian 09 program.[26]

*Chaetocin (**1**) and monosulfide analogue (**3**)*

Optical rotations (589 nm) and VCD were calculated at the ωB97XD/6-311G(d,p) level of theory and ECD at the ωB97XD/6-311++G(d,p) with a continuum solvation model for the solvent used for the experimental measurements.

*Dehydrogliotoxin (**21**) and dethiodehydrogliotoxin (**22**)*

ECD was calculated at the ωB97XD/6-311++G(d,p) level of theory, in 1,4-dioxane

*Gliotoxin analogue (**25**) and its monosulfide analogue (**26**)*

The mechanistic exploration was performed at the ωB97XD/6-311G(d,p) level of theory with a continuum solvation model for THF.

Optical rotations (589 nm), ECD and VCD were calculated at the M062X/6-311++G(d,p) level of theory with a continuum solvation model for the solvent used for the experimental measurements.

Biological evaluation

Biochemical assays

In vitro EZH2 and SUV39H1 screen with assay kits

Commercially available kits from BPS Bioscience (EZH2/EED/SUZ12/ RbAp48/AEBP2 cat. #52009, SUV39H1 Direct Activity Assay Kit #52006) were used according to the manufacturer's instructions.

The 96-well plates coated with Histone H3 peptide substrate were rehydrated with TBST buffer (1 x TBS, pH 8.0, containing 0.05% Tween-20) for 15 min. In the meantime, the enzyme was diluted with HMT assay buffer and was pre-incubated with SAM and the candidate inhibitor for 10 min in a different plate. The mixture was then added to the histone-coated plate and incubated for 1 h at room temperature. The plate was washed 3 times with TBST buffer and then incubated with blocking buffer for 10 min. The primary antibody (specific for methylated lysine residue) was diluted 100-fold with blocking buffer and was added to each well and the mixture was incubated for 1 h with slow shaking. The plate was then washed 3 times with TBST buffer and incubated in blocking buffer. The secondary HRP-labelled antibody (specific for the primary antibody) was diluted 1000-fold with blocking buffer and added to each well. After 30 min incubation the plate was washed again 3 times with TBST buffer and incubated in blocking buffer. Just before use, HRP chemiluminescent substrates A and B were mixed on ice and were added to each well. The plate was immediately read with the chemiluminescence module on a PHERAstar Plus (BMG Labtech, Software Version 2.21).

Materials for G9a and SUV39H1 AlphaLISA assay

Human G9a (euchromatic histone-lysine N-methyltransferase 2, GenBank Accession No. NM_006709, amino-acids 785 – 1210, with N-terminal GST tag, MW = 74.6 kDa, expressed in a Baculovirus infected Sf9 cell expression system, specific activity: 10 pmol/min/µg) and human SUV39H1 (GenBank Accession No. NM_003173, amino acids 82 – 412, with N-terminal GST tag, MW= 64 kDa, expressed in *E. Coli*, specific activity: 0.05 and 0.7 pmol/min/µg) used for all AlphaLISA experiments were

purchased from BPS Bioscience; Histone H3 (1 – 21) biotinylated peptides (H3K9me$_x$ x=0, 2, 3) were purchased from AnaSpec; Anti-methyl-histone H3 lysine 9 (H3K9me2) AlphaLISA Acceptor Beads, AlphaScreen Streptavidin Donor beads and AlphaLISA Epigenetics Buffer were purchased from PerkinElmer; *S*-(5'-Adenosyl)-L-methionine chloride (SAM) was purchased from Sigma and New England Biolabs (UK) Ltd, S-(5'-Adenosyl)-L-homocysteine (SAH), H_2O_2 solution (30% wt/wt), phenol red, and horseradish peroxidase (HRP), Hanks' balanced salt solution (HBSS), NaCl solution, Trizma (Tris) buffer, Tween20 were purchased from Sigma. White low-volume 384-well plates, 384-well flat bottomed clear polystyrene microtiter plates, V-shape 96-well clear plates were bought from Greiner Bio-One, Gloucester, UK. White low-volume 384-well plates were read on PHERAstar Plus with AlphaScreen module (excitation wavelength 680 nM, reading at 570 nM).

Data Analysis

Data was normalised to high (no inhibitor) and low (no enzyme) controls and plotted using GraphPad Prism V5.0.4. pIC_{50} ($-\log(IC_{50})$) values were derived from 10-point curve using a four parameter fit. Assays were run in duplicates in at least 3 independent experiments. The statistical significance was assessed by paired t-test or 1-way ANOVA followed by Bonferroni's multiple comparison test. P value summary: ***: $P < 0.001$ (Extremely significant), **: $0.001 < P < 0.01$ (Very significant), *: $0.01 < P < 0.05$ (Significant), ns: $P > 0.05$ (Not statistically significant)

Compound handling

Compounds were stored as 10 mM DMSO solutions under N_2 atmosphere (FluidX) or at –20 ˚C. Purity was assessed by NMR (>95% pure) and LCMS.

Representative dilution scheme from 10 mM DMSO solution (final percentage of DMSO kept constant):

Initial DMSO plate: 2-fold serial dilution in DMSO – [C] in mM										
10.00	5.00	2.50	1.25	0.63	0.31	0.16	0.08	0.04	0.02	DMSO

Intermediate plate: 200 fold dilution from DMSO plate in assay buffer (0.5 µL DMSO solution + 100 µL buffer) – [C] in µM										
50.00	25.00	12.50	6.25	3.13	1.56	0.78	0.39	0.20	0.10	DMSO

Final Concentration in Assay Plate: 2.5 µL from Intermediate plate added to well of total volume 5 µL (2 fold dilution)– [C] in µM										
25.00	12.50	6.25	3.13	1.56	0.78	0.39	0.20	0.10	0.05	DMSO

AlphaLISA inhibition assay

1 nM G9a (or 50 nM SUV39H1) was incubated with compounds (up to 100 µM, up to 1% DMSO), 100 nM biotinylated H3 peptide and 15 µM (or 40 µM for SUV39H1) SAM in assay buffer (50 mM Tris-HCl, pH 9.0, 50 mM NaCl, 0.01% Tween-20 and optional 1 mM DTT) with a total volume of 5 µL for 30 min (or 1.5 h for SUV39H1) at 25 ˚C in white low-volume 384-well plates. The reaction was quenched by addition of Anti-methyl-Histone H3 Lysine 9 (H3K9me2) AlphaLISA acceptor beads in epigenetic buffer (10 µg/mL) and incubated for 1 h. AlphaScreen streptavidin donor beads in epigenetic buffer (10 µg/mL) were added and incubated for 30 min. The plates were then read on Pherastar Plus (AlphaScreen module).

Representative plate layout for IC_{50} determination (I: Inhibitor, E: Enzyme)

	1	2	3	4	5	6	7	8	9	10	11	12
A	Dilution of Compound 1: higher (column 1) to lower (column 10)										NO I	NO E
B												
C	Dilution of Compound 2										NO I	NO E
D												
E	Dilution of Compound 3										NO I	NO E
F												
G	Dilution of Compound 4										NO I	NO E
H												
I	Dilution of Compound 5										NO I	NO E
J												
K	Dilution of Compound 6										NO I	NO E
L												
M	Dilution of Compound 7										NO I	NO E
N												
O	Dilution of Compound 8										NO I	NO E
P												
											High Control	***Low Control***

ROS generation test

The ROS generation test was performed as described:[134] Compounds (100 and 10 μM) in Hanks' balanced salt solution (HBSS) were incubated with 0.5 mM DTT for 15 min in 384-well flat bottomed, clear polystyrene microtiter plates. 100 μM H_2O_2 in HBSS was used as control (100%). Phenol red/HRP detection buffer (100 μg/ml phenol red and 60 μg/ml HRP in HBSS) was added and incubated for 5 minutes. The reaction was stopped with addition of 1 M NaOH. Absorbance was then read at 610 μM on a BMG NOVOstar and values were normalised to 100 μM H_2O_2 value.

Time-dependency test

Enzyme and inhibitor in assay buffer were pre-incubated for 5, 15, or 30 minutes prior to addition of the substrate in assay buffer (with and without DTT). For time 0 pre-incubation data, the substrate was added just before adding the inhibitor. The assay was otherwise performed as stated above.

Reversibility test by dilution[139]

G9a (100 nM) was incubated with compound at a concentration of 10 × IC_{50} (total volume 2 μL) for 30 min in assay buffer (50 mM Tris-HCl, pH 9.0, 50 mM NaCl, 0.01% Tween-20) with and without 1 mM DTT at room temperature in V-shape 96-well plates. The enzyme was also incubated with DMSO as a positive control. After pre-incubation, the samples were diluted 100-fold with peptide H3 and SAM in assay buffer (with and without DTT) to decrease all reaction components to their usual assay concentrations (as described previously) and the inhibitor concentration to 0.1 × IC_{50}. After 30 min incubation, 5 μL of each reaction mixture was transferred to a white low-volume 384-well plate and acceptor and donor beads were added as described in the inhibition assay. The percentage activity was determined relative to DMSO control.

SAM competition assay

IC_{50} values were determined as stated above at various concentrations of SAM (concentration range 10 – 160 μM for G9a, 6 – 800 μM for SUV39H1) and with G9a at a concentration of 0.5 nM or SUV39H1 at 50 nM. Time courses were performed at

the highest SAM concentration to ensure the system was still in the linear phase under all these conditions. The assay was run in presence of 1 mM DTT to ensure the reversibility of the inhibition by chaetocin. Enzyme and inhibitor were incubated for 30 min prior to addition of the substrates and the reactions were run for 30 min. For SUV39H1, pIC_{50} values were derived from 4-point curves at chosen concentrations around the IC_{50} region (0.2 – 0.05 μM). For G9a, the pIC_{50} values were derived from 6-point curves (32 – 0.03 μM)

Enzyme concentration dependency

IC_{50} values were determined as stated above at various concentrations of enzyme: 50 and 5 nM for SUV39H1. When SUV39H1 was used at the low concentration (5 nM), SAM was used at an increased concentration (4 mM) and the incubation time was increased to 4.5 h ensuring the reaction was in the linear phase.

Inhibition assay in presence urea (G9a)

A titration with increasing amounts of urea was performed to determine a concentration at which the overall enzyme activity was not too affected, typically about 70–80 % remaining activity (concentration range: 0.29 – 5 M for urea). The assays were then run as described above, with 0.7 M urea in assay buffer (1 mM DTT) and 30 min pre-incubation of enzyme and compound prior to addition of the substrates.

Mass spectrometry experiments

Expression and purification of G9a for mass spectrometry experiments – Cloning and baculoviral expression (Performed by Jim Reid, Domainex)

An *N*-terminally His-tagged DNA fragment, encoding a.a. 913 – 1193 of G9a, was subcloned into pFastBac for baculoviral expression using the Bac-to Bac system (Invitrogen). Recombinant G9a was expressed in Sf9 cells grown to mid-log phase in Insect-Xpress media supplemented with 5% FCS (Lonza). The cells were infected with an MOI of 2.5 and grown for further 72 h at 27 ˚C.

Protein purification – Harvested cells were resuspended in lysis buffer (50 mM HEPES pH 7.4, 300 mM NaCl, 5 mM $MgCl_2$, 10% Glycerol, 1 mM TCEP supplemented with EDTA-free protease inhibitors). The cells were lysed by sonication and the clarified lysate bound to Ni-NTA resin (Qiagen). The resin was washed with lysis buffer supplemented with 50 mM imidazole and eluted in 50 mM HEPES pH 7.4, 300 mM NaCl, 5 mM $MgCl_2$, 10% Glycerol, 1 mM TCEP, 250 mM imidazole. The protein was further purified to homogeneity by gel filtration using a Superdex S75 26/60 column (GE Healthcare) in 50 mM HEPES pH 7.4, 300 mM NaCl, 5 mM $MgCl_2$, 10% Glycerol, 1 mM TCEP. Fractions corresponding to dimeric G9a were pooled and concentrated.

Denaturing ESI - mass spectrometry analysis (Performed by Marina Demetriades in Prof. Chris Schofield Lab, Department of chemistry, Oxford University)

G9a was desalted using an Amicon Ultra-0.5, Ultracel-10 Membrane, 10 kDa (Amicon, U.K.) in 15 mM ammonium acetate (pH 7.5). The stock solution was diluted with the same buffer to a final concentration of 100 μM. Chaetocin was dissolved to 200 mM in DMSO and then diluted to 200 μM, in buffer. EDTA was dissolved in water to 1 M and further to 1 mM. All reagents were mixed to a final concentration of G9a 19 μM, chaetocin 38 μM, in a total volume of 12 μL. For the experiments where DTT was used, its final concentration in the mixture was 190 μM. ESI-MS analysis was performed at 0 min and 30 min incubation after adding 1 μL of 1% formic acid to 4 μL of the reaction mixture, just prior to analysis. Mass spectrometric data were acquired using a Q-TOF mass spectrometer (Q-TOF micro, Micromass, Altrincham, U.K.) interfaced with a NanoMate (Advion Biosciences,

Ithaca, NY) with a chip voltage of 1.70 kV and a delivery pressure 0.5 psi. The sample cone voltage was typically 60 V with a source temperature of 80 ˚C and with an acquisition/scan time of 1 s/1 s. Calibration and sample acquisition were performed in the positive ion mode in the range of 2000-4000 *m/z*. The pressure at the interface between the atmospheric source and the high vacuum region was fixed at 6.30 mbar. External instrument calibration was achieved using sodium formate. Data were processed with MassLynx 4.0 (Waters).

Mass spectrometry analysis to identify the cysteine residues involved in bonding with chaetocin (Performed by Aaron Borg and Peter DiMaggio, Department of Chemical Engineering, Imperial College London)

Recombinant G9a was dialysed into 100 mM NaCl, 1 mM TCEP (Sigma-Aldrich) and 50 mM HEPES pH 7.3 at 4 ˚C. Chaetocin was added to half of the G9a sample in a molar ratio of 2:1 and allowed to react at 25 ˚C for 30 min. The second half of the G9a sample was used as control experiment to which no chaetocin was added. To both the control and the treated samples, guanidine HCl (Promega) was added to a final concentration of 6 M to denature the G9a protein, and *N*-ethylmaleimide (NEM) (Sigma-Aldrich) was added to a final concentration of 14 mM and allowed to react for 30 min at room temperature to covalently label cysteine residues not engaged in disulfide bonding. The samples were then reduced with a final concentration of 5 mM DTT for 1 h at 51 ˚C, and iodoacetamide (Sigma-Aldrich) was added at a final concentration of 14 mM and allowed to react in the dark for 45 min to label the reduced cysteines. Proteomics grade trypsin (Promega) was added at a 1:20 ratio to digest the G9a protein for 6 h at 37 ˚C. After quenching digestion, ZipTips were used to reconstitute both samples into the correct loading solvent for nanoLC-MS.

The peptides were chromatographically resolved using a linear gradient on an Ultimate 3000 RSLCnano System (Dionex), with an Acclaim PepMap100, C18 stationary phase, 3 µm particle size, 100 Å pore size, 75 µm internal diameter × 15 cm length column (Thermo Fisher). The LC conditions comprised of a flow rate of 0.3 µL/min and a linear gradient starting at 1% B (5% H_2O, 95% MeCN, 0.1% Formic Acid and 99% A (0.1% FA, 100% H_2O) and increased to 95% B over 70 min. Real time tandem mass spectra were acquired on an LTQ Velos Pro linear ion trap (Thermo Scientific) with an 80 min acquisition time over a 240 – 1800 *m/z* scan range

and CID fragmentation collision energy of 35%. Tandem mass spectra were collected using top 5 data-dependent acquisition, with a dynamic exclusion list (repeat count of 2, repeat duration of 10 sec, exclusion list size 100 and exclusion duration of 100 sec) to provide sufficient MS/MS peptide coverage.

Initial peptide identification from the LC-MS/MS data was performed using a Sequest search in Proteome discoverer 1.3 (Thermo Fisher) against the Uniprot database with G9a appended. Reverse decoy false discovery rate values of 0.01/0.05 were used (strict/relaxed), allowing for 2 missed cleavages and a 2 Da precursor mass tolerance. Dynamic side chain modifications included in the search were carbamidomethyl (+57.021; from iodoacetamide treament), NEM (+125.048) and NEM in oxidised form (143.058) on cysteine residues, and oxidation (+15.995) on methionine residues. The accuracy of the tandem MS identifications reported by Sequest was manually assessed, and the 2 most abundant charge states observed for each cysteine-containing peptide were then quantitated by peak integration of the precursor ion intensity in the extracted ion chromatogram. This was done for all possible combinations of modifications on the cysteine and methionine residues, and the raw abundances were normalised across all observed modified states of the same peptide sequence to determine the relative percentage of NEM (indicating an exposed cysteine) and carbamidomethyl (indicating a cysteine involved in disulfide bond formation) labelling for each cysteine residue. The fold change in carbamidomethyl-to-NEM labelling (*i.e.* disulfide bond formation) upon chaetocin treatment was evaluated by taking ratio between the abundances of the treated and control samples

Cell-based assay

The cell-based assays were performed by Nadine Chapman-Rothe and Elham Shamsaei in Prof. Robert Brown Laboratories.

Cell culture and compound treatment: MDA-MB-231 breast cancer cell lines were cultured in DMEM-Medium supplemented with 10% FCS (#02.00.830, First Link (UK), 2 mM L-Glutamine (#25030-024, Invitrogen), 100 U/mL Penicillin and 100 µg/mL Streptavidin (#15070-063, Invitrogen) until they were ~60% confluent and then treated for either 48 h or 72 h with a compound. Compounds were diluted in DMSO at stocks of 10 mM.

Real-time measurements for cell-based screening: Following compound treatment for 48 h (in 6-well plates), media was removed and 1.5 mL of TRIzol (#155966018, Invitrogen) was added directly to lyse cells. Once cells were lysed, the protocol was continued with 1 mL (TRIzol, cell lysate) according to the manufacturer's instructions. Purified RNA was dissolved in 35 µL of UltraPure DNase/RNase-free distilled water (#10977049, Invitrogen). Reverse transcription was done using the SuperScript III First-Strand Synthesis System (#18080-051, Invitrogen) according to the manufacturer's instructions, using 7 µL of the purified RNA as starting material. For real-time measurements the 2x iQ SYBR Green Supermix (#170-8882, Bio-Rad), 200 nM Primers and 0.4 µL of cDNA /per 20 µL reaction was used. The measurement was done in low-white 96-well plates (#MLL9651, Bio-Rad) on a CFX96 Real-time System/C1000 Thermal Cycler (Bio-Rad) with the following protocol: 95 ˚C for 3 min; 95 ˚C for 10 s, 56 ˚C for 10 s, 72 ˚C for 30 s. 42 cycles followed by a melting curve from 72 ˚C to 95 ˚C in order to control for primer dimer or unwanted products. Each measurement was done in triplicate. For normalisation we have tested numerous 'house-keeping genes' such as b-actin, HPRT1 as well as TBP, but in our hands, for drug treatment purposes only GAPDH as well as RNA pol II seem to be a reliable constant normaliser. In order to account for preparation/handling differences during drug treatment and real-time measurement, we are using a second GAPDH (GAPDH_2) primer pair, and would count an experiment as valid if the difference between these primer pairs is not greater than +/–0.15 fold for each real-time run.

Crystallographic data

Crystallographic data for compound **25**

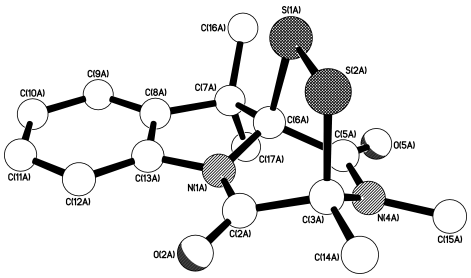


Table 1. Crystal data and structure refinement for **25.**

Identification code	MF1301
Formula	C15 H16 N2 O2 S2
Formula weight	320.42
Temperature	173 K
Diffractometer, wavelength	OD Xcalibur 3, 0.71073 Å
Crystal system, space group	Triclinic, P-1
Unit cell dimensions	a = 9.9436(5) Å α = 71.694(3)°
	b = 10.2460(3) Å β = 81.285(4)°
	c = 15.4717(6) Å γ = 89.134(3)°
Volume, Z	1478.35(11) $Å^3$, 4
Density (calculated)	1.440 Mg/m^3
Absorption coefficient	0.365 mm^{-1}
F(000)	672
Crystal colour / morphology	Colourless blocky needles
Crystal size	0.44 x 0.40 x 0.28 mm^3
θ range for data collection	3.07 to 29.59°
Index ranges	-13<=h<=8, -12<=k<=13, -21<=l<=19
Reflns collected / unique	12089 / 6834 [R(int) = 0.0211]
Reflns observed [F>4σ(F)]	5475
Absorption correction	Analytical
Max. and min. transmission	0.924 and 0.890
Refinement method	Full-matrix least-squares on F^2
Data / restraints / parameters	6834 / 0 / 381
Goodness-of-fit on F^2	1.029
Final R indices [F>4σ(F)]	R1 = 0.0414, wR2 = 0.0984

R indices (all data)	R1 = 0.0564, wR2 = 0.1084
Largest diff. peak, hole	0.358, -0.281 eÅ^{-3}
Mean and maximum shift/error	0.000 and 0.000

Table 2. Bond lengths [Å] and angles [°] for **25**.

S(1A)-C(6A)	1.8822(18)
S(1A)-S(2A)	2.0733(8)
S(2A)-C(3A)	1.881(2)
N(1A)-C(2A)	1.357(2)
N(1A)-C(13A)	1.416(2)
N(1A)-C(6A)	1.452(2)
C(2A)-O(2A)	1.214(2)
C(2A)-C(3A)	1.531(3)
C(3A)-N(4A)	1.448(3)
C(3A)-C(14A)	1.512(3)
N(4A)-C(5A)	1.365(3)
N(4A)-C(15A)	1.469(3)
C(5A)-O(5A)	1.215(2)
C(5A)-C(6A)	1.522(3)
C(6A)-C(7A)	1.563(3)
C(7A)-C(8A)	1.520(3)
C(7A)-C(16A)	1.531(3)
C(7A)-C(17A)	1.539(3)
C(8A)-C(9A)	1.384(3)
C(8A)-C(13A)	1.390(3)
C(9A)-C(10A)	1.387(3)
C(10A)-C(11A)	1.381(3)
C(11A)-C(12A)	1.388(3)
C(12A)-C(13A)	1.387(3)
S(1B)-C(6B)	1.8881(18)
S(1B)-S(2B)	2.0743(8)
S(2B)-C(3B)	1.886(2)
N(1B)-C(2B)	1.359(2)
N(1B)-C(13B)	1.417(2)
N(1B)-C(6B)	1.447(2)
C(2B)-O(2B)	1.214(2)
C(2B)-C(3B)	1.521(3)
C(3B)-N(4B)	1.449(3)
C(3B)-C(14B)	1.518(3)
N(4B)-C(5B)	1.372(3)
N(4B)-C(15B)	1.460(3)

C(5B)-O(5B)	1.213(2)
C(5B)-C(6B)	1.524(2)
C(6B)-C(7B)	1.565(3)
C(7B)-C(8B)	1.509(3)
C(7B)-C(16B)	1.530(3)
C(7B)-C(17B)	1.541(3)
C(8B)-C(13B)	1.383(3)
C(8B)-C(9B)	1.387(3)
C(9B)-C(10B)	1.386(3)
C(10B)-C(11B)	1.376(4)
C(11B)-C(12B)	1.401(3)
C(12B)-C(13B)	1.390(3)

C(6A)-S(1A)-S(2A)	98.04(6)
C(3A)-S(2A)-S(1A)	99.52(7)
C(2A)-N(1A)-C(13A)	127.48(16)
C(2A)-N(1A)-C(6A)	120.41(15)
C(13A)-N(1A)-C(6A)	109.85(14)
O(2A)-C(2A)-N(1A)	124.81(18)
O(2A)-C(2A)-C(3A)	123.68(18)
N(1A)-C(2A)-C(3A)	111.45(17)
N(4A)-C(3A)-C(14A)	114.77(18)
N(4A)-C(3A)-C(2A)	111.43(16)
C(14A)-C(3A)-C(2A)	110.72(19)
N(4A)-C(3A)-S(2A)	110.38(15)
C(14A)-C(3A)-S(2A)	105.91(14)
C(2A)-C(3A)-S(2A)	102.81(13)
C(5A)-N(4A)-C(3A)	118.05(16)
C(5A)-N(4A)-C(15A)	116.79(19)
C(3A)-N(4A)-C(15A)	119.57(19)
O(5A)-C(5A)-N(4A)	124.08(18)
O(5A)-C(5A)-C(6A)	122.59(19)
N(4A)-C(5A)-C(6A)	113.20(17)
N(1A)-C(6A)-C(5A)	112.23(16)
N(1A)-C(6A)-C(7A)	105.02(14)
C(5A)-C(6A)-C(7A)	118.53(15)
N(1A)-C(6A)-S(1A)	109.25(12)
C(5A)-C(6A)-S(1A)	101.53(12)
C(7A)-C(6A)-S(1A)	110.13(12)
C(8A)-C(7A)-C(16A)	113.10(16)
C(8A)-C(7A)-C(17A)	107.52(16)
C(16A)-C(7A)-C(17A)	110.01(16)

C(8A)-C(7A)-C(6A)	100.39(14)
C(16A)-C(7A)-C(6A)	113.65(16)
C(17A)-C(7A)-C(6A)	111.73(16)
C(9A)-C(8A)-C(13A)	119.12(18)
C(9A)-C(8A)-C(7A)	129.58(18)
C(13A)-C(8A)-C(7A)	111.23(16)
C(8A)-C(9A)-C(10A)	119.27(19)
C(11A)-C(10A)-C(9A)	120.50(19)
C(10A)-C(11A)-C(12A)	121.57(19)
C(13A)-C(12A)-C(11A)	116.91(19)
C(12A)-C(13A)-C(8A)	122.61(17)
C(12A)-C(13A)-N(1A)	128.81(17)
C(8A)-C(13A)-N(1A)	108.55(16)
C(6B)-S(1B)-S(2B)	97.67(6)
C(3B)-S(2B)-S(1B)	99.48(7)
C(2B)-N(1B)-C(13B)	128.60(15)
C(2B)-N(1B)-C(6B)	120.27(15)
C(13B)-N(1B)-C(6B)	110.05(14)
O(2B)-C(2B)-N(1B)	124.07(18)
O(2B)-C(2B)-C(3B)	123.96(18)
N(1B)-C(2B)-C(3B)	111.89(16)
N(4B)-C(3B)-C(14B)	114.86(17)
N(4B)-C(3B)-C(2B)	110.52(16)
C(14B)-C(3B)-C(2B)	110.70(18)
N(4B)-C(3B)-S(2B)	110.81(13)
C(14B)-C(3B)-S(2B)	105.93(15)
C(2B)-C(3B)-S(2B)	103.32(12)
C(5B)-N(4B)-C(3B)	118.06(15)
C(5B)-N(4B)-C(15B)	117.00(18)
C(3B)-N(4B)-C(15B)	120.68(19)
O(5B)-C(5B)-N(4B)	123.34(18)
O(5B)-C(5B)-C(6B)	123.24(18)
N(4B)-C(5B)-C(6B)	113.22(16)
N(1B)-C(6B)-C(5B)	112.25(14)
N(1B)-C(6B)-C(7B)	104.75(14)
C(5B)-C(6B)-C(7B)	118.66(15)
N(1B)-C(6B)-S(1B)	109.57(12)
C(5B)-C(6B)-S(1B)	101.21(12)
C(7B)-C(6B)-S(1B)	110.30(12)
C(8B)-C(7B)-C(16B)	113.50(16)
C(8B)-C(7B)-C(17B)	107.18(16)
C(16B)-C(7B)-C(17B)	109.67(17)

C(8B)-C(7B)-C(6B)	100.71(14)
C(16B)-C(7B)-C(6B)	113.87(17)
C(17B)-C(7B)-C(6B)	111.48(15)
C(13B)-C(8B)-C(9B)	119.76(19)
C(13B)-C(8B)-C(7B)	111.33(16)
C(9B)-C(8B)-C(7B)	128.74(19)
C(10B)-C(9B)-C(8B)	118.7(2)
C(11B)-C(10B)-C(9B)	120.8(2)
C(10B)-C(11B)-C(12B)	121.9(2)
C(13B)-C(12B)-C(11B)	116.0(2)
C(8B)-C(13B)-C(12B)	122.87(18)
C(8B)-C(13B)-N(1B)	108.59(15)
C(12B)-C(13B)-N(1B)	128.52(18)

Crystallographic data for compound **49**

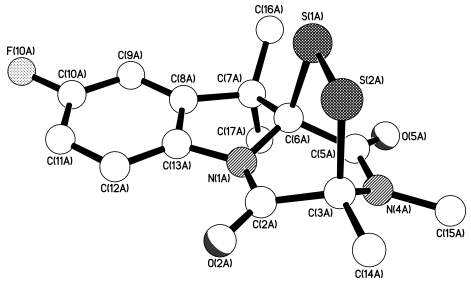


Table 1. Crystal data and structure refinement for **49**.

Identification code	MF1302
Formula	C15 H15 F N2 O2 S2
Formula weight	338.41
Temperature	173 K
Diffractometer, wavelength	OD Xcalibur 3, 0.71073 Å
Crystal system, space group	Triclinic, P-1
Unit cell dimensions	a = 7.5855(3) Å α = 106.977(4)°
	b = 14.5072(6) Å β = 99.642(3)°
	c = 14.9344(6) Å γ = 104.652(3)°
Volume, Z	1468.07(12) $Å^3$, 4
Density (calculated)	1.531 Mg/m^3
Absorption coefficient	0.382 mm^{-1}
F(000)	704
Crystal colour / morphology	Pale yellow blocks
Crystal size	0.38 x 0.25 x 0.15 mm^3
θ range for data collection	2.80 to 29.55°
Index ranges	-10<=h<=10, -18<=k<=20, -20<=l<=20
Reflns collected / unique	12227 / 6794 [R(int) = 0.0241]
Reflns observed [F>4σ(F)]	5616
Absorption correction	Analytical
Max. and min. transmission	0.951 and 0.897
Refinement method	Full-matrix least-squares on F^2
Data / restraints / parameters	6794 / 0 / 399
Goodness-of-fit on F^2	1.032
Final R indices [F>4σ(F)]	R1 = 0.0408, wR2 = 0.1039
R indices (all data)	R1 = 0.0521, wR2 = 0.1137
Largest diff. peak, hole	0.423, -0.340 $eÅ^{-3}$
Mean and maximum shift/error	0.000 and 0.001

Table 2. Bond lengths [Å] and angles [°] for **49.**

S(1A)-C(6A)	1.8944(19)
S(1A)-S(2A)	2.0825(7)
S(2A)-C(3A)	1.895(2)
N(1A)-C(2A)	1.364(2)
N(1A)-C(13A)	1.420(2)
N(1A)-C(6A)	1.447(2)
C(2A)-O(2A)	1.217(2)
C(2A)-C(3A)	1.521(3)
C(3A)-N(4A)	1.447(2)
C(3A)-C(14A)	1.516(3)
N(4A)-C(5A)	1.372(2)
N(4A)-C(15A)	1.464(2)
C(5A)-O(5A)	1.216(2)
C(5A)-C(6A)	1.517(3)
C(6A)-C(7A)	1.561(2)
C(7A)-C(8A)	1.520(3)
C(7A)-C(16A)	1.528(3)
C(7A)-C(17A)	1.538(3)
C(8A)-C(13A)	1.385(2)
C(8A)-C(9A)	1.388(3)
C(9A)-C(10A)	1.382(3)
C(10A)-F(10A)	1.363(2)
C(10A)-C(11A)	1.377(3)
C(11A)-C(12A)	1.389(3)
C(12A)-C(13A)	1.381(3)
S(1B)-C(6B)	1.8972(19)
S(1B)-S(2B)	2.0617(7)
S(2B)-C(3B)	1.904(2)
N(1B)-C(2B)	1.366(2)
N(1B)-C(13B)	1.420(2)
N(1B)-C(6B)	1.448(2)
C(2B)-O(2B)	1.211(2)
C(2B)-C(3B)	1.530(3)
C(3B)-N(4B)	1.441(2)
C(3B)-C(14B)	1.512(3)
N(4B)-C(5B)	1.374(2)
N(4B)-C(15B)	1.464(2)
C(5B)-O(5B)	1.217(2)
C(5B)-C(6B)	1.519(2)
C(6B)-C(7B)	1.557(2)
C(7B)-C(8B)	1.523(2)
C(7B)-C(16B)	1.525(3)
C(7B)-C(17B)	1.540(3)
C(8B)-C(9B)	1.380(3)
C(8B)-C(13B)	1.387(3)
C(9B)-C(10B)	1.382(3)
C(10B)-F(10B)	1.359(2)
C(10B)-C(11B)	1.373(3)
C(11B)-C(12B)	1.393(3)
C(12B)-C(13B)	1.385(3)

C(6A)-S(1A)-S(2A)	97.43(6)
C(3A)-S(2A)-S(1A)	99.19(6)
C(2A)-N(1A)-C(13A)	127.51(15)
C(2A)-N(1A)-C(6A)	119.89(15)
C(13A)-N(1A)-C(6A)	109.64(14)
O(2A)-C(2A)-N(1A)	124.33(18)
O(2A)-C(2A)-C(3A)	123.93(16)
N(1A)-C(2A)-C(3A)	111.64(15)
N(4A)-C(3A)-C(14A)	114.22(16)

```
N(4A)-C(3A)-C(2A)    112.29(15)
C(14A)-C(3A)-C(2A)   110.89(15)
N(4A)-C(3A)-S(2A)    110.58(12)
C(14A)-C(3A)-S(2A)   106.44(14)
C(2A)-C(3A)-S(2A)    101.52(13)
C(5A)-N(4A)-C(3A)    118.32(15)
C(5A)-N(4A)-C(15A)   117.71(16)
C(3A)-N(4A)-C(15A)   119.65(15)
O(5A)-C(5A)-N(4A)    124.12(17)
O(5A)-C(5A)-C(6A)    123.09(16)
N(4A)-C(5A)-C(6A)    112.59(15)
N(1A)-C(6A)-C(5A)    113.13(14)
N(1A)-C(6A)-C(7A)    104.11(14)
C(5A)-C(6A)-C(7A)    118.41(15)
N(1A)-C(6A)-S(1A)    110.26(12)
C(5A)-C(6A)-S(1A)    100.58(11)
C(7A)-C(6A)-S(1A)    110.38(12)
C(8A)-C(7A)-C(16A)   114.50(17)
C(8A)-C(7A)-C(17A)   106.82(15)
C(16A)-C(7A)-C(17A)  110.31(17)
C(8A)-C(7A)-C(6A)     99.96(14)
C(16A)-C(7A)-C(6A)   114.71(16)
C(17A)-C(7A)-C(6A)   109.88(15)
C(13A)-C(8A)-C(9A)   119.51(18)
C(13A)-C(8A)-C(7A)   110.25(16)
C(9A)-C(8A)-C(7A)    129.84(17)
C(10A)-C(9A)-C(8A)   116.85(18)
F(10A)-C(10A)-C(11A) 118.11(19)
F(10A)-C(10A)-C(9A)  117.94(19)
C(11A)-C(10A)-C(9A)  123.94(19)
C(10A)-C(11A)-C(12A) 119.09(19)
C(13A)-C(12A)-C(11A) 117.43(18)
C(12A)-C(13A)-C(8A)  123.15(17)
C(12A)-C(13A)-N(1A)  128.29(16)
C(8A)-C(13A)-N(1A)   108.56(15)
C(6B)-S(1B)-S(2B)     97.48(6)
C(3B)-S(2B)-S(1B)     98.52(6)
C(2B)-N(1B)-C(13B)   127.35(15)
C(2B)-N(1B)-C(6B)    120.06(15)
C(13B)-N(1B)-C(6B)   108.85(14)
O(2B)-C(2B)-N(1B)    124.49(17)
O(2B)-C(2B)-C(3B)    124.02(17)
N(1B)-C(2B)-C(3B)    111.36(15)
N(4B)-C(3B)-C(14B)   114.86(16)
N(4B)-C(3B)-C(2B)    111.69(15)
C(14B)-C(3B)-C(2B)   110.96(15)
N(4B)-C(3B)-S(2B)    112.30(13)
C(14B)-C(3B)-S(2B)   105.92(13)
C(2B)-C(3B)-S(2B)     99.98(12)
C(5B)-N(4B)-C(3B)    118.74(15)
C(5B)-N(4B)-C(15B)   117.61(15)
C(3B)-N(4B)-C(15B)   120.32(15)
O(5B)-C(5B)-N(4B)    123.92(17)
O(5B)-C(5B)-C(6B)    123.35(16)
N(4B)-C(5B)-C(6B)    112.57(15)
N(1B)-C(6B)-C(5B)    112.35(15)
N(1B)-C(6B)-C(7B)    104.26(14)
C(5B)-C(6B)-C(7B)    119.99(15)
N(1B)-C(6B)-S(1B)    111.01(12)
C(5B)-C(6B)-S(1B)    100.06(11)
C(7B)-C(6B)-S(1B)    109.17(13)
C(8B)-C(7B)-C(16B)   113.87(16)
C(8B)-C(7B)-C(17B)   106.88(15)
```

```
C(16B)-C(7B)-C(17B)  110.06(16)
C(8B)-C(7B)-C(6B)     99.66(14)
C(16B)-C(7B)-C(6B)   114.77(16)
C(17B)-C(7B)-C(6B)   110.98(16)
C(9B)-C(8B)-C(13B)   120.40(17)
C(9B)-C(8B)-C(7B)    129.48(17)
C(13B)-C(8B)-C(7B)   110.08(16)
C(8B)-C(9B)-C(10B)   116.78(18)
F(10B)-C(10B)-C(11B) 118.39(17)
F(10B)-C(10B)-C(9B)  118.07(19)
C(11B)-C(10B)-C(9B)  123.53(18)
C(10B)-C(11B)-C(12B) 119.74(18)
C(13B)-C(12B)-C(11B) 117.09(18)
C(12B)-C(13B)-C(8B)  122.41(18)
C(12B)-C(13B)-N(1B)  128.77(17)
C(8B)-C(13B)-N(1B)   108.73(15)
```

Crystallographic data for compound **50**

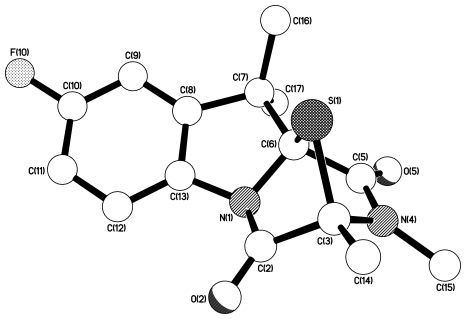


Table 1. Crystal data and structure refinement for **50**.

Identification code	MF1303
Formula	C15 H15 F N2 O2 S
Formula weight	306.35
Temperature	173 K
Diffractometer, wavelength	OD Xcalibur 3, 0.71073 Å
Crystal system, space group	Monoclinic, P2(1)/n
Unit cell dimensions	a = 12.1110(3) Å α = 90°
	b = 9.0355(3) Å β = 90.632(3)°
	c = 12.9502(4) Å γ = 90°
Volume, Z	1417.04(7) $Å^3$, 4
Density (calculated)	1.436 Mg/m^3
Absorption coefficient	0.245 mm^{-1}
F(000)	640
Crystal colour / morphology	Colourless blocks
Crystal size	0.51 x 0.42 x 0.33 mm^3
θ range for data collection	3.36 to 29.55°
Index ranges	-16<=h<=15, -11<=k<=11, -17<=l<=17
Reflns collected / unique	11162 / 3424 [R(int) = 0.0218]
Reflns observed [F>4σ(F)]	2958
Absorption correction	Analytical
Max. and min. transmission	0.940 and 0.915
Refinement method	Full-matrix least-squares on F^2
Data / restraints / parameters	3424 / 0 / 192
Goodness-of-fit on F^2	1.055
Final R indices [F>4σ(F)]	R1 = 0.0341, wR2 = 0.0845
R indices (all data)	R1 = 0.0412, wR2 = 0.0896
Largest diff. peak, hole	0.307, -0.242 $eÅ^{-3}$
Mean and maximum shift/error	0.000 and 0.001

Table 2. Bond lengths [Å] and angles [°] for **50.**

S(1)-C(3)	1.8447(13)
S(1)-C(6)	1.8591(14)
N(1)-C(2)	1.3926(17)
N(1)-C(13)	1.4197(16)
N(1)-C(6)	1.4856(16)
C(2)-O(2)	1.2096(16)
C(2)-C(3)	1.5510(19)
C(3)-N(4)	1.4717(17)
C(3)-C(14)	1.4981(19)
N(4)-C(5)	1.3551(17)
N(4)-C(15)	1.4519(17)
C(5)-O(5)	1.2164(16)
C(5)-C(6)	1.5493(18)
C(6)-C(7)	1.5405(18)
C(7)-C(8)	1.5223(19)
C(7)-C(16)	1.5337(19)
C(7)-C(17)	1.5407(19)
C(8)-C(9)	1.3867(19)
C(8)-C(13)	1.3934(19)
C(9)-C(10)	1.383(2)
C(10)-F(10)	1.3637(17)
C(10)-C(11)	1.375(2)
C(11)-C(12)	1.392(2)
C(12)-C(13)	1.3857(19)

C(3)-S(1)-C(6)	78.27(6)
C(2)-N(1)-C(13)	124.24(11)
C(2)-N(1)-C(6)	109.36(10)
C(13)-N(1)-C(6)	107.56(10)
O(2)-C(2)-N(1)	126.49(13)
O(2)-C(2)-C(3)	126.85(12)
N(1)-C(2)-C(3)	106.65(11)
N(4)-C(3)-C(14)	115.70(11)
N(4)-C(3)-C(2)	106.63(11)
C(14)-C(3)-C(2)	115.06(11)
N(4)-C(3)-S(1)	102.60(8)
C(14)-C(3)-S(1)	115.86(10)
C(2)-C(3)-S(1)	98.97(8)
C(5)-N(4)-C(15)	124.00(12)
C(5)-N(4)-C(3)	111.87(10)
C(15)-N(4)-C(3)	124.13(11)
O(5)-C(5)-N(4)	126.40(12)
O(5)-C(5)-C(6)	127.44(12)
N(4)-C(5)-C(6)	106.13(11)
N(1)-C(6)-C(7)	106.17(10)
N(1)-C(6)-C(5)	104.73(10)
C(7)-C(6)-C(5)	123.50(11)
N(1)-C(6)-S(1)	102.39(8)
C(7)-C(6)-S(1)	116.44(9)
C(5)-C(6)-S(1)	101.25(9)
C(8)-C(7)-C(16)	113.42(12)
C(8)-C(7)-C(6)	100.00(10)
C(16)-C(7)-C(6)	112.45(11)
C(8)-C(7)-C(17)	107.82(11)
C(16)-C(7)-C(17)	110.10(11)
C(6)-C(7)-C(17)	112.66(11)
C(9)-C(8)-C(13)	119.88(13)
C(9)-C(8)-C(7)	128.83(13)
C(13)-C(8)-C(7)	111.18(11)
C(10)-C(9)-C(8)	116.87(14)

F(10)-C(10)-C(11)	118.17(14)
F(10)-C(10)-C(9)	117.91(15)
C(11)-C(10)-C(9)	123.91(14)
C(10)-C(11)-C(12)	119.28(14)
C(13)-C(12)-C(11)	117.58(14)
C(12)-C(13)-C(8)	122.48(13)
C(12)-C(13)-N(1)	128.11(12)
C(8)-C(13)-N(1)	109.40(11)

References

(1) World-Health-Organisation http://www.who.int/mediacentre/factsheets/fs297/en/; Vol. 2013.

(2) Aggarwal, S. *Nat. Rev. Drug Discov.* **2010**, *9*, 427-428.

(3) Vogelstein, B.; Kinzler, K. W. *The genetic basis of human cancer*; 2nd ed.; McGraw-Hill Companies, 2002.

(4) Hanahan, D.; Weinberg, R. A. *Cell* **2000**, *100*, 57-70.

(5) Hanahan, D.; Weinberg, R. A. *Cell* **2011**, *144*, 646-674.

(6) Weinberg, R. A. *The biology of cancer*; Garland Science, Taylor and Francis Group, LLC, 2007.

(7) Papac, R. J. *Yale J. Biol. Med.* **2001**, *74*, 391.

(8) Capdeville, R.; Buchdunger, E.; Zimmermann, J.; Matter, A. *Nat. Rev. Drug Discov.* **2002**, *1*, 493-502.

(9) Weinstein, I. B. *Science* **2002**, *297*, 63-64.

(10) Weinstein, I. B.; Joe, A. K. *Nat. Clin. Pract. Oncol.* **2006**, *3*, 448-457.

(11) Jones, P. A.; Baylin, S. B. *Cell* **2007**, *128*, 683-692.

(12) Feinberg, A. P.; Ohlsson, R.; Henikoff, S. *Nat. Rev. Genet.* **2006**, *7*, 21-33.

(13) Kouzarides, T. *Cell* **2007**, *128*, 693-705.

(14) Cherblanc, F. L.; Davidson, R. W. M.; Di Fruscia, P.; Srimongkolpithak, N.; Fuchter, M. J. *Nat. Prod. Rep.* **2013**, *30*, 605-624.

(15) Baylin, S. B. *Nat. Clin. Pract. Oncol.* **2005**, *2*, S4-S11.

(16) Klose, R. J.; Bird, A. P. *Trends Biochem. Sci.* **2006**, *31*, 89-97.

(17) Watt, F.; Molloy, P. L. *Genes Dev.* **1988**, *2*, 1136-1143.

(18) Boyes, J.; Bird, A. *Cell* **1991**, *64*, 1123-1134.

(19) Wade, P. A. *Oncogene* **2001**, *20*, 3166-3173.

(20) Cherblanc, F.; Chapman-Rothe, N.; Brown, R.; Fuchter, M. J. *Future Med. Chem.* **2012**, *4*, 425-446.

(21) Strahl, B. D.; Allis, C. D. *Nature* **2000**, *403*, 41-45.

(22) Bedford, M. T.; Richard, S. *Mol. Cell* **2005**, *18*, 263-272.

(23) Bannister, A. J.; Kouzarides, T. *Cell Res.* **2011**, *21*, 381-395.

(24) Yun, M. Y.; Wu, J.; Workman, J. L.; Li, B. *Cell Res.* **2011**, *21*, 564-578.

(25) Suganuma, T.; Workman, J. L. *Annu. Rev. Biochem.* **2011**, *80*, 473-499.

(26) Lall, S. *Nat. Struct. Biol.* **2007**, *14*, 1110-1115.

(27) Jenuwein, T.; Allis, C. D. *Science* **2001**, *293*, 1074-1080.

(28) Zhu, P.; Martin, E.; Mengwasser, J.; Schlag, P.; Janssen, K. P.; Gottlicher, M. *Cancer Cell* **2004**, *5*, 455-463.

(29) Ropero, S.; Fraga, M. F.; Ballestar, E.; Hamelin, R.; Yamamoto, H.; Boix-Chornet, M.; Caballero, R.; Alaminos, M.; Setien, F.; Paz, M. F.; Herranz, M.; Palacios, J.; Arango, D.; Orntoft, T. F.; Aaltonen, L. A.; Schwartz, S.; Esteller, M. *Nat. Genet.* **2006**, *38*, 566-569.

(30) Moore, S. D. P.; Herrick, S. R.; Ince, T. A.; Kleinman, M. S.; Dal Cin, P.; Morton, C. C.; Quade, B. J. *Cancer Res.* **2004**, *64*, 5570-5577.

(31) Bryan, E. J.; Jokubaitis, V. J.; Chamberlain, N. L.; Baxter, S. W.; Dawson, E.; Choong, D. Y. H.; Campbell, I. G. *Int. J. Cancer* **2002**, *102*, 137-141.

(32) Dalgliesh, G. L.; Furge, K.; Greenman, C.; Chen, L. N.; Bignell, G.; Butler, A.; Davies, H.; Edkins, S.; Hardy, C.; Latimer, C.; Teague, J.; Andrews, J.; Barthorpe, S.; Beare, D.; Buck, G.; Campbell, P. J.; Forbes, S.; Jia, M. M.; Jones, D.; Knott, H.;

Kok, C. Y.; Lau, K. W.; Leroy, C.; Lin, M. L.; McBride, D. J.; Maddison, M.; Maguire, S.; McLay, K.; Menzies, A.; Mironenko, T.; Mulderrig, L.; Mudie, L.; O'Meara, S.; Pleasance, E.; Rajasingham, A.; Shepherd, R.; Smith, R.; Stebbings, L.; Stephens, P.; Tang, G.; Tarpey, P. S.; Turrell, K.; Dykema, K. J.; Khoo, S. K.; Petillo, D.; Wondergem, B.; Anema, J.; Kahnoski, R. J.; Teh, B.; Stratton, M. R.; Futreal, P. A. *Nature* **2010**, *463*, 360-363.

(33) Tsai, H. C.; Baylin, S. B. *Cell Res.* **2011**, *21*, 502-517.

(34) Blancafort, P.; Jin, J.; Frye, S. *Mol. Pharmacol.* **2013**, *83*, 563-576.

(35) Toyota, M.; Issa, J.-P. J. *Electrophoresis* **2000**, *21*, 329-333.

(36) Esteller, M. *Oncogene* **2002**, *21*, 5427-5440.

(37) Esteller, M. *Hum. Mol. Gen.* **2007**, *16*, R50-R59.

(38) Spannhoff, A.; Hauser, A. T.; Heinke, R.; Sippl, W.; Jung, M. *ChemMedChem* **2009**, *4*, 1568-1582.

(39) Simon, J. A.; Lange, C. A. *Mutat. Res., Fundam. Mol. Mech. Mutagen.* **2008**, *647*, 21-29.

(40) Kelly, T. K.; De Carvalho, D. D.; Jones, P. A. *Nat. Biotechnol.* **2010**, *28*, 1069-1078.

(41) Bots, M.; Johnstone, R. W. *Clin. Cancer Res.* **2009**, *15*, 3970-3977.

(42) Carew, J. S.; Giles, F. J.; Nawrocki, S. T. *Cancer Lett.* **2008**, *269*, 7-17.

(43) Yang, H.; Hoshino, K.; Sanchez-Gonzalez, B.; Kantarjian, H.; Garcia-Manero, G. *Leuk. Res.* **2005**, *29*, 739-748.

(44) Chapman-Rothe, N.; Brown, R. *Future Med. Chem.* **2009**, *1*, 1481-1495.

(45) Issa, J. P.; Kantarjian, H. M. *Clin. Cancer Res.* **2009**, *15*, 3938-3946.

(46) Sorm, F.; Piskala, A.; Cihak, A.; Vesely, J. *Experientia* **1964**, *20*, 202-203.

(47) Stresemann, C.; Lyko, F. *Int. J. Cancer* **2008**, *123*, 8-13.

(48) Mann, B. S.; Johnson, J. R.; Cohen, M. H.; Justice, R.; Pazdur, R. *Oncologist* **2007**, *12*, 1247-1252.

(49) Marks, P. A.; Breslow, R. *Nat. Biotechnol.* **2007**, *25*, 84-90.

(50) Whittaker, S. J.; Demierre, M.-F.; Kim, E. J.; Rook, A. H.; Lerner, A.; Duvic, M.; Scarisbrick, J.; Reddy, S.; Robak, T.; Becker, J. r. C.; Samtsov, A.; McCulloch, W.; Kim, Y. H. *J. Clin. Oncol.* **2010**, *28*, 4485-4491.

(51) Piekarz, R. L.; Frye, R.; Turner, M.; Wright, J. J.; Allen, S. L.; Kirschbaum, M. H.; Zain, J.; Prince, H. M.; Leonard, J. P.; Geskin, L. J.; Reeder, C.; Joske, D.; Figg, W. D.; Gardner, E. R.; Steinberg, S. M.; Jaffe, E. S.; Stetler-Stevenson, M.; Lade, S.; Fojo, A. T.; Bates, S. E. *J. Clin. Oncol.* **2009**, *27*, 5410-5417.

(52) Lane, A. A.; Chabner, B. A. *J. Clin. Oncol.* **2009**, *27*, 5459-5468.

(53) Ma, X.; Ezzeldin, H. H.; Diasio, R. B. *Drugs* **2009**, *69*, 1911-1934.

(54) Balasubramanian, S.; Verner, E.; Buggy, J. J. *Cancer Lett.* **2009**, *280*, 211-221.

(55) Lohse, B.; Kristensen, J. L.; Kristensen, L. H.; Agger, K.; Helin, K.; Gajhede, M.; Clausen, R. P. *Bioorg. Med. Chem.* **2011**, *19*, 3625-3636.

(56) Copeland, R. A. *Drug Discov. Today Ther. Strateg.* **2011**, DOI: 10.1016/j.ddstr.2011.1008.1001.

(57) Kleer, C. G.; Cao, Q.; Varambally, S.; Shen, R. L.; Ota, L.; Tomlins, S. A.; Ghosh, D.; Sewalt, R. G. A. B.; Otte, A. P.; Hayes, D. F.; Sabel, M. S.; Livant, D.; Weiss, S. J.; Rubin, M. A.; Chinnaiyan, A. M. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 11606-11611.

(58) Yu, J.; Yu, J.; Rhodes, D. R.; Tomlins, S. A.; Cao, X.; Chen, G.; Mehra, R.; Wang, X.; Ghosh, D.; Shah, R. B.; Varambally, S.; Pienta, K. J.; Chinniaiyan, A. M. *Cancer Res.* **2007**, *67*, 10657-10663.

(59) Albert, M.; Helin, K. *Sem. Cell Dev. Biol.* **2010**, *21*, 209-220.

(60) Varier, R. A.; Timmers, H. T. M. *BBA-Rev. Cancer* **2011**, *1815*, 75-89.

(61) Wigle, T. J.; Copeland, R. A. *Curr. Opin. Chem. Biol.* **2013**, *17*, 369-378.

(62) Copeland, R. A.; Solomon, M. E.; Richon, V. M. *Nat. Rev. Drug Discov.* **2009**, *8*, 724-732.

(63) Dillon, S.; Zhang, X.; Trievel, R.; Cheng, X. *Genome Biol.* **2005**, *6*, 227.

(64) Smith, B. C.; Denu, J. M. *BBA-Gene Regul. Mecha.* **2009**, *1789*, 45-57.

(65) Chang, Y.; Zhang, X.; Horton, J. R.; Upadhyay, A. K.; Spannhoff, A.; Liu, J.; Snyder, J. P.; Bedford, M. T.; Cheng, X. *Nat. Struct. Mol. Biol.* **2009**, *16*, 312-317.

(66) Zhang, X.; Wen, H.; Shi, X. *Acta Biochim. Biophys. Sin.* **2012**, *44*, 14-27.

(67) Patnaik, D.; Chin, H. G.; Estève, P.-O.; Benner, J.; Jacobsen, S. E.; Pradhan, S. *J. Biol. Chem.* **2004**, *279*, 53248-53258.

(68) Chin, H. G.; Patnaik, D.; Esteve, P. O.; Jacobsen, S. E.; Pradhan, S. *Biochemistry* **2006**, *45*, 3272-3284.

(69) Zhang, X.; Yang, Z.; Khan, S. I.; Horton, J. R.; Tamaru, H.; Selker, E. U.; Cheng, X. *Mol. Cell* **2003**, *12*, 177-185.

(70) Arrowsmith, C. H.; Bountra, C.; Fish, P. V.; Lee, K.; Schapira, M. *Nat. Rev. Drug Discov.* **2012**, *11*, 384-400.

(71) Wu, H.; Min, J. R.; Lunin, V. V.; Antoshenko, T.; Dombrovski, L.; Zeng, H.; Allali-Hassani, A.; Campagna-Slater, V.; Vedadi, M.; Arrowsmith, C. H.; Plotnikov, A. N.; Schapira, M. *Plos One* **2010**, *5*, e8570

(72) Cole, P. A. *Nat. Chem. Biol.* **2008**, *4*, 590-597.

(73) Daigle, S. R.; Olhava, E. J.; Therkelsen, C. A.; Majer, C. R.; Sneeringer, C. J.; Song, J.; Johnston, L. D.; Scott, M. P.; Smith, J. J.; Xiao, Y. H.; Jin, L.; Kuntz, K. W.; Chesworth, R.; Moyer, M. P.; Bernt, K. M.; Tseng, J. C.; Kung, A. L.; Armstrong, S. A.; Copeland, R. A.; Richon, V. M.; Pollock, R. M. *Cancer Cell* **2011**, *20*, 53-65.

(74) Greiner, D.; Bonaldi, T.; Eskeland, R.; Roemer, E.; Imhof, A. *Nat. Chem. Biol.* **2005**, *1*, 143-145.

(75) Kubicek, S.; O'Sullivan, R. J.; August, E. M.; Hickey, E. R.; Zhang, Q.; Teodoro, M. L.; Rea, S.; Mechtler, K.; Kowalski, J. A.; Homon, C. A.; Kelly, T. A.; Jenuwein, T. *Mol. Cell* **2007**, *25*, 473-481.

(76) Yuan, Y.; Wang, Q.; Paulk, J.; Kubicek, S.; Kemp, M. M.; Adams, D. J.; Shamji, A. F.; Wagner, B. K.; Schreiber, S. L. *ACS Chem. Biol.* **2012**, *7*, 1152-1157.

(77) Knutson, S. K.; Wigle, T. J.; Warholic, N. M.; Sneeringer, C. J.; Allain, C. J.; Klaus, C. R.; Sacks, J. D.; Raimondi, A.; Majer, C. R.; Song, J.; Scott, M. P.; Jin, L.; Smith, J. J.; Olhava, E. J.; Chesworth, R.; Moyer, M. P.; Richon, V. M.; Copeland, R. A.; Keilhack, H.; Pollock, R. M.; Kuntz, K. W. *Nat. Chem. Biol.* **2012**, *8*, 890-896.

(78) Konze, K. D.; Ma, A.; Li, F.; Barsyte-Lovejoy, D.; Parton, T.; MacNevin, C. J.; Liu, F.; Gao, C.; Huang, X.-P.; Kuznetsova, E.; Rougie, M.; Jiang, A.; Pattenden, S. G.; Norris, J. L.; James, L. I.; Roth, B. L.; Brown, P. J.; Frye, S. V.; Arrowsmith, C. H.; Hahn, K. M.; Wang, G. G.; Vedadi, M.; Jin, J. *ACS Chem. Biol.* **2013**, *8*, 1324-1334.

(79) McCabe, M. T.; Ott, H. M.; Ganji, G.; Korenchuk, S.; Thompson, C.; Van Aller, G. S.; Liu, Y.; Graves, A. P.; Iii, A. D. P.; Diaz, E.; LaFrance, L. V.; Mellinger, M.; Duquenne, C.; Tian, X.; Kruger, R. G.; McHugh, C. F.; Brandt, M.; Miller, W. H.; Dhanak, D.; Verma, S. K.; Tummino, P. J.; Creasy, C. L. *Nature* **2012**, *492*, 108-112.

(80) Qi, W.; Chan, H.; Teng, L.; Li, L.; Chuai, S.; Zhang, R.; Zeng, J.; Li, M.; Fan, H.; Lin, Y.; Gu, J.; Ardayfio, O.; Zhang, J.-H.; Yan, X.; Fang, J.; Mi, Y.; Zhang, M.; Zhou, T.; Feng, G.; Chen, Z.; Li, G.; Yang, T.; Zhao, K.; Liu, X.; Yu, Z.; Lu, C. X.; Atadja, P.; Li, E. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 21360-21365.

(81) Verma, S. K.; Tian, X.; LaFrance, L. V.; Duquenne, C.; Suarez, D. P.; Newlander, K. A.; Romeril, S. P.; Burgess, J. L.; Grant, S. W.; Brackley, J. A.; Graves, A. P.; Scherzer, D. A.; Shu, A.; Thompson, C.; Ott, H. M.; Aller, G. S. V.; Machutta, C. A.; Diaz, E.; Jiang, Y.; Johnson, N. W.; Knight, S. D.; Kruger, R. G.; McCabe, M. T.; Dhanak, D.; Tummino, P. J.; Creasy, C. L.; Miller, W. H. *ACS Med. Chem. Lett.* **2012**, *3*, 1091-1096.

(82) Rea, S.; Eisenhaber, F.; O'Carroll, N.; Strahl, B. D.; Sun, Z. W.; Schmid, M.; Opravil, S.; Mechtler, K.; Ponting, C. P.; Allis, C. D.; Jenuwein, T. *Nature* **2000**, *406*, 593-599.

(83) Hauser, D.; Weber, H. P.; Sigg, H. P. *Helv. Chim. Acta* **1970**, *53*, 1061-1073.

(84) Iwasa, E.; Hamashima, Y.; Sodeoka, M. *Isr. J. Chem.* **2011**, *51*, 420-433.

(85) Gardiner, D. M.; Waring, P.; Howlett, B. J. *Microbiol.-Sgm* **2005**, *151*, 1021-1032.

(86) Cheeke, P. R. *J. Anim. Sci.* **1995**, *73*, 909-918.

(87) Jordan, T. W.; Cordiner, S. J. *Trends Pharmacol. Sci.* **1987**, *8*, 144-149.

(88) Declercq, E.; Billiau, A.; Ottenheijm, H. C. J.; Herscheid, J. D. M. *Biochem. Pharmacol.* **1978**, *27*, 635-639.

(89) Pahl, H. L.; Krauss, B.; SchulzeOsthoff, K.; Decker, T.; Traenckner, E. B. M.; Vogt, M.; Myers, C.; Parks, T.; Warring, P.; Muhlbacher, A.; Czernilofsky, A. P.; Baeuerle, P. A. *J. Exp. Med.* **1996**, *183*, 1829-1840.

(90) Rodriguez, P. L.; Carrasco, L. *J. Virol.* **1992**, *66*, 1971-1976.

(91) Vigushin, D. M.; Mirsaidi, N.; Brooke, G.; Sun, C.; Pace, P.; Inman, L.; Moody, C. J.; Coombes, R. C. *Med. Oncol.* **2004**, *21*, 21-30.

(92) Waring, P.; Beaver, J. *Gen. Pharmacol.* **1996**, *27*, 1311-1316.

(93) Cook, K. M.; Hilton, S. T.; Mecinovic, J.; Motherwell, W. B.; Figg, W. D.; Schofield, C. J. *J. Biol. Chem.* **2009**, *284*, 26831-26838.

(94) Hurne, A. M.; Chai, C. L. L.; Waring, P. *J. Biol. Chem.* **2000**, *275*, 25202-25206.

(95) Tibodeau, J. D.; Benson, L. M.; Isham, C. R.; Owen, W. G.; Bible, K. C. *Antioxid. Redox Signal.* **2009**, *11*, 1097-1106.

(96) Saito, T.; Suzuki, Y.; Koyama, K.; Natori, S.; Iitaka, Y.; Kinoshita, T. *Chem. Pharm. Bull.* **1988**, *36*, 1942-1956.

(97) Sekita, S.; Yoshihira, K.; Natori, S. *Chem. Pharm. Bull.* **1980**, *28*, 2428-2435.

(98) Sekita, S.; Yoshihira, K.; Natori, S.; Udagawa, S.; Muroi, T.; Sugiyama, Y.; Kurata, H.; Umeda, M. *Can. J. Microbiol.* **1981**, *27*, 766-772.

(99) Udagawa, S.; Muroi, T.; Kurata, H.; Sekita, S.; Yoshihira, K.; Natori, S.; Umeda, M. *Can. J. Microbiol.* **1979**, *25*, 170-177.

(100) Nicolaou, K. C.; Lu, M.; Totokotsopoulos, S.; Heretsch, P.; Giguère, D.; Sun, Y.-P.; Sarlah, D.; Nguyen, T. H.; Wolf, I. C.; Smee, D. F.; Day, C. W.; Bopp, S.; Winzeler, E. A. *J. Am. Chem. Soc.* **2012**, *134*, 17320-17332.

(101) Ohler, E.; Poisel, H.; Tataruch, F.; Schmidt, U. *Chem. Ber.* **1972**, *105*, 635-641.

(102) Nicolaou, K. C.; Giguere, D.; Totokotsopoulos, S.; Sun, Y. P. *Angew. Chem., Int. Ed.* **2012**, *51*, 728-732.

(103) Aliev, A. E.; Hilton, S. T.; Motherwell, W. B.; Selwood, D. L. *Tetrahedron Lett.* **2006**, *47*, 2387-2390.

(104) Iwasa, E.; Hamashima, Y.; Fujishiro, S.; Higuchi, E.; Ito, A.; Yoshida, M.; Sodeoka, M. *J. Am. Chem. Soc.* **2010**, *132*, 4078–4079.

(105) Kim, J.; Ashenhurst, J. A.; Movassaghi, M. *Science* **2009**, *324*, 238-241.

(106) Trown, P. W. *Biochem. Biophys. Res. Commun.* **1968**, *33*, 402-407.

(107) Poisel, H.; Schmidt, U. *Angew. Chem., Int. Ed.* **1971**, *10*, 130-131.

(108) Rewcastle, G. W.; Sutherland, H. S.; Weir, C. A.; Blackburn, A. G.; Denny, W. A. *Tetrahedron Lett.* **2005**, *46*, 8719-8721.

(109) Hilton, S. T.; Motherwell, W. B.; Selwood, D. L. *Synlett* **2004**, 2609-2611.

(110) Safe, S.; Taylor, A. *J. Chem. Soc. D* **1969**, 1466-1467.

(111) Chai, C. L. L.; Elix, J. A.; Huleatt, P. B.; Waring, P. *Bioorg. Med. Chem.* **2004**, *12*, 5991-5995.

(112) Tataruch, F.; Schmidt, U. *Monatsh. Chem.* **1973**, *104*, 1530-1533.

(113) Becker, S.; Fort, Y.; Vanderesse, R.; Caubere, P. *J. Org. Chem.* **1989**, *54*, 4848-4853.

(114) Truce, W. E.; Roberts, F. E. *J. Org. Chem.* **1963**, *28*, 961-964.

(115) Truce, W. E.; Perry, F. M. *J. Org. Chem.* **1965**, *30*, 1316-1317.

(116) Caubère, P. *Angew. Chem., Int. Ed.* **1983**, *22*, 599-613.

(117) Back, T. G.; Baron, D. L.; Yang, K. *J. Org. Chem.* **1993**, *58*, 2407-2413.

(118) Pontecorvo, G.; Roper, J. A.; Hemmons, L. M.; Macdonald, K. D.; Bufton, A. W. J. *Adv. Genet.* **1953**, *5*, 141-238.

(119) Cheswort.W *Clays Clay Miner.* **1971**, *19*, 337-339.

(120) Johnson, J. R.; Buchanan, J. B. *J. Am. Chem. Soc.* **1953**, *75*, 2103-2109.

(121) Dutcher, J. D.; Johnson, J. R.; Bruce, W. F. *J. Am. Chem. Soc.* **1945**, *67*, 1736-1745.

(122) Öhler, E.; Tataruch, F.; Schmidt, U. *Chem. Ber.* **1973**, *106*, 165-176.

(123) Paquet, V.; Roy, M.; Pedro, L.; Gauthier, N.; Labonté, A.; Rodenbrock, A.; Pinard, G.; Beaudet, L.; Rodriguez-Suarez, R. "AlphaLISA G9a Histone H3-Lysine N-methyltransferase Assay," PerkinElmer, Inc., 2010.

(124) DeRosa, M. C.; Crutchley, R. J. *Coord. Chem. Rev.* **2002**, *233*, 351-371.

(125) Ishii, K. *Coord. Chem. Rev.* **2012**, *256*, 1556-1568.

(126) Minor, L. K. *Handbook of Assay Development in Drug Discovery*; Taylor & Francis, 2006, 58-59.

(127) Collins, R. E.; Tachibana, M.; Tamaru, H.; Smith, K. M.; Jia, D.; Zhang, X.; Selker, E. U.; Shinkai, Y.; Cheng, X. *J. Biol. Chem.* **2005**, *280*, 5563-5570.

(128) Osipovich, O.; Milley, R.; Meade, A.; Tachibana, M.; Shinkai, Y.; Krangel, M. S.; Oltz, E. M. *Nat. Immunol.* **2004**, *5*, 309-316.

(129) Cherblanc, F. L.; Chapman, K. L.; Reid, J.; Borg, A. J.; Sundriyal, S.; Alcazar-Fuoli, L.; Bignell, E.; Demetriades, M.; Schofield, C.; DiMaggio, P. A.; Brown, R.; Fuchter, M. J. *J. Med. Chem.* **2013**, *In Press*, DOI: 10.1021/jm401063r.

(130) Pedras, M. S. C.; Abrams, S. R.; Séguin-Swartz, G. *Tetrahedron Lett.* **1988**, *29*, 3471-3474.

(131) Cherblanc, F. L.; Chapman, K. L.; Brown, R.; Fuchter, M. J. *Nat. Chem. Biol.* **2013**, *9*, 136–137.

(132) Isham, C. R.; Tibodeau, J. D.; Jin, W.; Xu, R. F.; Timm, M. M.; Bible, K. C. *Blood* **2007**, *109*, 2579-2588.

(133) Chaib, H.; Nebbioso, A.; Prebet, T.; Castellano, R.; Garbit, S.; Restouin, A.; Vey, N.; Altucci, L.; Collette, Y. *Leukemia* **2012**, *26*, 662-674.

(134) Johnston, P. A.; Soares, K. M.; Shinde, S. N.; Foster, C. A.; Shun, T. Y.; Takyi, H. K.; Wipf, P.; Lazo, J. S. *Assay Drug Dev. Techn.* **2008**, *6*, 505-518.

(135) Pick, E.; Keisari, Y. *J. Immunol. Methods* **1980**, *38*, 161-170.

(136) Brezak, M. C.; Valette, A.; Quaranta, M.; Contour-Galcera, M. O.; Jullien, D.; Lavergne, O.; Frongia, C.; Bigg, D.; Kasprzyk, P. G.; Prevost, G. P.; Ducommun, B. *Int. J. Cancer* **2009**, *124*, 1449-1456.

(137) Scott, K. *personnal communication, unpublished data* **2012**.

(138) Chelikani, P.; Fita, I.; Loewen, P. C. *Cell. Mol. Life Sci.* **2004**, *61*, 192-208.

(139) Copeland, R. A. In *Evaluation of Enzyme Inhibitors in Drug Discovery: A Guide for Medicinal Chemists and Pharmacologists*; John Wiley & Sons, Inc.: New-Jersey, 2005.

(140) McGovern, S. L.; Caselli, E.; Grigorieff, N.; Shoichet, B. K. *J. Med. Chem.* **2002**, *45*, 1712-1722.

(141) Ryan, A. J.; Gray, N. M.; Lowe, P. N.; Chung, C.-w. *J. Med. Chem.* **2003**, *46*, 3448-3451.

(142) Kung, A. L.; Zabludoff, S. D.; France, D. S.; Freedman, S. J.; Tanner, E. A.; Vieira, A.; Cornell-Kennon, S.; Lee, J.; Wang, B.; Wang, J.; Memmert, K.; Naegeli, H.-U.; Petersen, F.; Eck, M. J.; Bair, K. W.; Wood, A. W.; Livingston, D. M. *Cancer cell* **2004**, *6*, 33-43.

(143) Block, K. M.; Wang, H.; Szabo, L. Z.; Polaske, N. W.; Henchey, L. K.; Dubey, R.; Kushal, S.; Laszlo, C. F.; Makhoul, J.; Song, Z. H.; Meuillet, E. J.; Olenyuk, B. Z. *J. Am. Chem. Soc.* **2009**, *131*, 18078-18088.

(144) Rishton, G. M. *Drug Discov. Today* **1997**, *2*, 382-384.

(145) Rishton, G. M. *Drug Discov. Today* **2003**, *8*, 86-96.

(146) Waring, P.; Sjaarda, A.; Lin, Q. H. *Biochem. Pharmacol.* **1995**, *49*, 1195-1201.

(147) Singh, J.; Petter, R. C.; Baillie, T. A.; Whitty, A. *Nat. Rev. Drug Discov.* **2011**, *10*, 307-317.

(148) Takahashi, M.; Takemoto, Y.; Shimazu, T.; Kawasaki, H.; Tachibana, M.; Shinkai, Y.; Takagi, M.; Shin-ya, K.; Igarashi, Y.; Ito, A.; Yoshida, M. *J. Antibiot. (Tokyo)* **2012**, *65*, 263-265.

(149) Sun, Y.; Takada, K.; Takemoto, Y.; Yoshida, M.; Nogi, Y.; Okada, S.; Matsunaga, S. *J. Nat. Prod.* **2012**, *75*, 111-114.

(150) Gersch, M.; Kreuzer, J.; Sieber, S. A. *Nat. Prod. Rep.* **2012**, *29*, 659-682.

(151) Qian, C.; Zhou, M. M. *Cell. Mol. Life Sci.* **2006**, *63*, 2755-2763.

(152) Zhang, X.; Tamaru, H.; Khan, S. I.; Horton, J. R.; Keefe, L. J.; Selker, E. U.; Cheng, X. *Cell* **2002**, *111*, 117-127.

(153) Müller, J.; Hart, C. M.; Francis, N. J.; Vargas, M. L.; Sengupta, A.; Wild, B.; Miller, E. L.; O'Connor, M. B.; Kingston, R. E.; Simon, J. A. *Cell* **2002**, *111*, 197-208.

(154) Lakshmikuttyamma, A.; Scott, S. A.; DeCoteau, J. F.; Geyer, C. R. *Oncogene* **2010**, *29*, 576-588.

(155) Cherrier, T.; Suzanne, S.; Redel, L.; Calao, M.; Marban, C.; Samah, B.; Mukerjee, R.; Schwartz, C.; Gras, G.; Sawaya, B. E.; Zeichner, S. L.; Aunis, D.; Van Lint, C.; Rohr, O. *Oncogene* **2009**, *28*, 3380-3389.

(156) Bernhard, W.; Barreto, K.; Saunders, A.; Dahabieh, M. S.; Johnson, P.; Sadowski, I. *FEBS Lett.* **2011**, *585*, 3549-3554.

(157) Wu, D.; Ingram, A.; Lahti, J. H.; Mazza, B.; Grenet, J.; Kapoor, A.; Liu, L.; Kidd, V. J.; Tang, D. *J. Biol. Chem.* **2002**, *277*, 12001-12008.

(158) Füllgrabe, J.; Hajji, N.; Joseph, B. *Cell Death Differ.* **2010**, *17*, 1238-1243.

(159) McConathy, J.; Owens, M. J. *Prim. Care Companion J. Clin. Psychiatry* **2003**, *5*, 70.

(160) Stephens, T. D.; Bunde, C. J. W.; Fillmore, B. J. *Biochem. Pharmacol.* **2000**, *59*, 1489-1499.

(161) Schönberg, A. *Chem. Ber.* **1935**, *68*, 163-164.

(162) Harpp, D. N.; Gleason, J. G. *J. Am. Chem. Soc.* **1971**, *93*, 2437-2445.

(163) Harpp, D. N.; Gleason, J. G.; Snyder, J. P. *J. Am. Chem. Soc.* **1968**, *90*, 4181-4182.

(164) Safe, S.; Taylor, A. *J. Chem. Soc. C* **1971**, 1189-1192.

(165) Sammes, P. G. In *Fortschritte der Chemie Organischer Naturstoffe / Progress in the Chemistry of Organic Natural Products*; Zechmeister, L., Herz, W., Grisebach, H., Kirby, G. W., Eds.; Springer Vienna: 1975; Vol. 32, p 51-118.

(166) Ziffer, H.; Weiss, U.; Charney, E. *Tetrahedron* **1967**, *23*, 3881-3886.

(167) Berova, N.; Bari, L. D.; Pescitelli, G. *Chem. Soc. Rev.* **2007**, *36*, 914-931.

(168) Sato, T.; Hino, T. *Tetrahedron* **1976**, *32*, 507-513.

(169) Herscheid, J. D. M.; Tijhuis, M. W.; Noordik, J. H.; Ottenheijm, H. C. J. *J. Am. Chem. Soc.* **1979**, *101*, 1159-1162.

(170) Noordik, J. H.; Beurkens, P. T.; Ottenheijm, H. C. J.; Herscheid, J. D. M.; Tijhuis, M. W. *Cryst. Struct. Comm.* **1978**, *7*, 669-677.

(171) Ferezou, J. P.; Quesneauthierry, A.; Cesario, M.; Pascard, C.; Barbier, M. *J. Am. Chem. Soc.* **1983**, *105*, 5402-5406.

(172) Felippe, L. G.; Batista Jr, J. M.; Baldoqui, D. C.; Nascimento, I. R.; Kato, M. J.; He, Y.; Nafie, L. A.; Furlan, M. *Org. Biomol. Chem.* **2012**, *10*, 4208-4214.

(173) Bruhn, T.; Schaumlöffel, A.; Hemberger, Y.; Bringmann, G. *Chirality* **2013**, *25*, 243-249.

(174) Berova, N.; Polavarapu, P. L.; Nakanishi, K.; Woody, R. W. *Comprehensive Chiroptical Spectroscopy, Instrumentation, Methodologies, and Theoretical Simulations*; Wiley, 2011.

(175) Nafie, L. A. *Nat. Prod. Commun.* **2008**, *3*, 451-466.

(176) Barron, L. D.; Buckingham, A. D. *Chem. Phys. Lett.* **2010**, *492*, 199-213.

(177) Gaussian 09 Revision A.1; Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K. R., A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.;

Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Gaussian, Inc., Wallingford CT **2009.**

(178) Debie, E.; Bultinck, P.; Nafie, L. A.; Dukor, R. K. *CompareVOA* **2010**, *BioTools, Jupiter, Florida (USA).*

(179) Debie, E.; De Gussem, E.; Dukor, R. K.; Herrebout, W.; Nafie, L. A.; Bultinck, P. *ChemPhysChem* **2011**, *12*, 1542-1549.

(180) Cherblanc, F.; Lo, Y.-P.; De Gussem, E.; Alcazar-Fuoli, L.; Bignell, E.; He, Y.; Chapman-Rothe, N.; Bultinck, P.; Herrebout, W. A.; Brown, R.; Rzepa, H. S.; Fuchter, M. J. *Chem.–Eur. J.* **2011**, *17*, 11868-11875.

(181) Alecu, I. M.; Zheng, J.; Zhao, Y.; Truhlar, D. G. *J. Chem. Theory Comput.* **2010**, *6*, 2872-2887.

(182) Rose, C.; Vargas, F.; Bourgeat, P.; Schwartz, J.-C.; Bishop, P. B.; Bambal, R. B.; Ganellin, C. R.; Leblond, B.; Moore, A. N. J.; Chan, S.; Lihua, Z.; Appl., P. I., Ed. 1996; Vol. WO9635805 (A2).

(183) Cherblanc, F.; Lo, Y.-P.; Bultinck, P.; Herrebout, W. A.; Rzepa, H. S.; Fuchter, M. J. *J. Org. Chem.* **2013**, *In Press*, DOI: 10.1021/jo401316a.

(184) Watthey, J. W. H.; Stanton, J. L.; Desai, M.; Babiarz, J. E.; Finn, B. M. *J. Med. Chem.* **1985**, *28*, 1511-1516.

(185) Zhao, H.; He, X.; Thurkauf, A.; Hoffman, D.; Kieltyka, A.; Brodbeck, R.; Primus, R.; Wasley, J. W. F. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 3111-3115.

(186) Rahman, R.; Safe, S.; Taylor, A. *J. Chem. Soc. C* **1969**, *0*, 1665-1668.

(187) Kim, J.; Movassaghi, M. *J. Am. Chem. Soc.* **2010**, *132*, 14376-14378.

(188) Ottenheijm, H. C. J.; Herscheid, J. D. M.; Tijhuis, M. W.; Nivard, R. J. F.; Declercq, E.; Prick, P. A. J. *J. Med. Chem.* **1978**, *21*, 799-804.

(189) Salama, T. A.; El-Ahl, A.-A. S.; Elmorsy, S. S.; Khalil, A.-G. M.; Ismail, M. A. *Tetrahedron Lett.* **2009**, *50*, 5933-5936.

(190) Takikawa, Y.; Shimada, K.; Makabe, T.; Takizawa, S. *Chem. Lett.* **1983**, *12*, 1503-1506.

(191) Polavarapu, P. L. *Chirality* **2008**, *20*, 664-672.

(192) Copeland, R. A.; Olhava, E. J.; Scott, M. P. *Curr. Opin. Chem. Biol.* **2010**, *14*, 505-510.

(193) Margueron, R.; Reinberg, D. *Nature* **2011**, *469*, 343-349.

(194) Hansen, K. H.; Bracken, A. P.; Pasini, D.; Dietrich, N.; Gehani, S. S.; Monrad, A.; Rappsilber, J.; Lerdrup, M.; Helin, K. *Nat. Cell Biol.* **2008**, *10*, 1291-1300.

(195) Wu, H.; Chen, X. Z.; Xiong, J.; Li, Y. F.; Li, H.; Ding, X. J.; Liu, S.; Chen, S.; Gao, S. R.; Zhu, B. *Cell Res.* **2011**, *21*, 365-367.

(196) Tsang, D. P. F.; Cheng, A. S. L. *J. Gastroen. Hepatol.* **2011**, *26*, 19-27.

(197) Bryant, R. J.; Cross, N. A.; Eaton, C. L.; Hamdy, F. C.; Cunliffe, V. T. *Prostate* **2007**, *67*, 547-556.

(198) Wilson, B. G.; Wang, X.; Shen, X.; McKenna, E. S.; Lemieux, M. E.; Cho, Y.-J.; Koellhoffer, E. C.; Pomeroy, S. L.; Orkin, S. H.; Roberts, C. W. M. *Cancer Cell* **2010**, *18*, 316-328.

(199) Varambally, S.; Dhanasekaran, S. M.; Zhou, M.; Barrette, T. R.; Kumar-Sinha, C.; Sanda, M. G.; Ghosh, D.; Pienta, K. J.; Sewalt, R. G. A. B.; Otte, A. P. *Nature* **2002**, *419*, 624-629.

(200) Bracken, A. P.; Pasini, D.; Capra, M.; Prosperini, E.; Colli, E.; Helin, K. *EMBO J.* **2003**, *22*, 5323-5335.

(201) Van Haaften, G.; Dalgliesh, G. L.; Davies, H.; Chen, L.; Bignell, G.; Greenman, C.; Edkins, S.; Hardy, C.; O'Meara, S.; Teague, J.; Butler, A.; Hinton, J.; Latimer, C.; Andrews, J.; Barthorpe, S.; Beare, D.; Buck, G.; Campbell, P. J.; Cole, J.; Forbes, S.; Jia, M.; Jones, D.; Kok, C. Y.; Leroy, C.; Lin, M.-L.; McBride, D. J.; Maddison, M.; Maquire, S.; McLay, K.; Menzies, A.; Mironenko, T.; Mulderrig, L.; Mudie, L.; Pleasance, E.; Shepherd, R.; Smith, R.; Stebbings, L.; Stephens, P.; Tang,

G.; Tarpey, P. S.; Turner, R.; Turrell, K.; Varian, J.; West, S.; Widaa, S.; Wray, P.; Collins, V. P.; Ichimura, K.; Law, S.; Wong, J.; Yuen, S. T.; Leung, S. Y.; Tonon, G.; DePinho, R. A.; Tai, Y.-T.; Anderson, K. C.; Kahnoski, R. J.; Massie, A.; Khoo, S. K.; Teh, B. T.; Stratton, M. R.; Futreal, P. A. *Nat. Genet.* **2009**, *41*, 521-523.

(202) Gui, Y. T.; Guo, G. W.; Huang, Y.; Hu, X. D.; Tang, A. F.; Gao, S. J.; Wu, R. H.; Chen, C.; Li, X. X.; Zhou, L.; He, M. H.; Li, Z. S.; Sun, X. J.; Jia, W. L.; Chen, J. N.; Yang, S. M.; Zhou, F. J.; Zhao, X. K.; Wan, S. Q.; Ye, R.; Liang, C. Z.; Liu, Z. S.; Huang, P. D.; Liu, C. X.; Jiang, H.; Wang, Y.; Zheng, H. C.; Sun, L.; Liu, X. W.; Jiang, Z. M.; Feng, D. F.; Chen, J.; Wu, S.; Zou, J.; Zhang, Z. F.; Yang, R. L.; Zhao, J.; Xu, C. J.; Yin, W. H.; Guan, Z. C.; Ye, J. X.; Zhang, H.; Li, J. X.; Kristiansen, K.; Nickerson, M. L.; Theodorescu, D.; Li, Y. R.; Zhang, X. Q.; Li, S. G.; Wang, J.; Yang, H. M.; Wang, J.; Cai, Z. M. *Nat. Genet.* **2011**, *43*, 875-U884.

(203) Agger, K.; Cloos, P. A. C.; Rudkjær, L.; Williams, K.; Andersen, G.; Christensen, J.; Helin, K. *Genes Dev.* **2009**, *23*, 1171-1176.

(204) Morin, R. D.; Johnson, N. A.; Severson, T. M.; Mungall, A. J.; An, J. H.; Goya, R.; Paul, J. E.; Boyle, M.; Woolcock, B. W.; Kuchenbauer, F.; Yap, D.; Humphries, R. K.; Griffith, O. L.; Shah, S.; Zhu, H.; Kimbara, M.; Shashkin, P.; Charlot, J. F.; Tcherpakov, M.; Corbett, R.; Tam, A.; Varhol, R.; Smailus, D.; Moksa, M.; Zhao, Y. J.; Delaney, A.; Qian, H.; Birol, I.; Schein, J.; Moore, R.; Holt, R.; Horsman, D. E.; Connors, J. M.; Jones, S.; Aparicio, S.; Hirst, M.; Gascoyne, R. D.; Marra, M. A. *Nat. Genet.* **2010**, *42*, 181-U124.

(205) Sneeringer, C. J.; Scott, M. P.; Kuntz, K. W.; Knutson, S. K.; Pollock, R. M.; Richon, V. M.; Copeland, R. A. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 20980-20985.

(206) McCabe, M. T.; Graves, A. P.; Ganji, G.; Diaz, E.; Halsey, W. S.; Jiang, Y.; Smitheman, K. N.; Ott, H. M.; Pappalardi, M. B.; Allen, K. E.; Chen, S. B.; Della Pietra, A.; Dul, E.; Hughes, A. M.; Gilbert, S. A.; Thrall, S. H.; Tummino, P. J.; Kruger, R. G.; Brandt, M.; Schwartz, B.; Creasy, C. L. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 2989-2994.

(207) Majer, C. R.; Jin, L.; Scott, M. P.; Knutson, S. K.; Kuntz, K. W.; Keilhack, H.; Smith, J. J.; Moyer, M. P.; Richon, V. M.; Copeland, R. A.; Wigle, T. J. *FEBS Lett.* **2012**, *586*, 3448-3451.

(208) Hock, H. *Genes Dev.* **2012**, *26*, 751-755.

(209) Nikoloski, G.; Langemeijer, S. M. C.; Kuiper, R. P.; Knops, R.; Massop, M.; Tonnissen, E. R. L. T. M.; van der Heijden, A.; Scheele, T. N.; Vandenberghe, P.; de Witte, T.; van der Reijden, B. A.; Jansen, J. H. *Nat. Genet.* **2010**, *42*, 665-667.

(210) Makishima, H.; Jankowska, A. M.; Tiu, R. V.; Szpurka, H.; Sugimoto, Y.; Hu, Z.; Sauntharajah, Y.; Guinta, K.; Keddache, M. A.; Putnam, P.; Sekeres, M. A.; Moliterno, A. R.; List, A. F.; McDevitt, M. A.; Maciejewski, J. P. *Leukemia* **2010**, *24*, 1799-1804.

(211) Ernst, T.; Chase, A. J.; Score, J.; Hidalgo-Curtis, C. E.; Bryant, C.; Jones, A. V.; Waghorn, K.; Zoi, K.; Ross, F. M.; Reiter, A.; Hochhaus, A.; Drexler, H. G.; Duncombe, A.; Cervantes, F.; Oscier, D.; Boultwood, J.; Grand, F. H.; Cross, N. C. P. *Nat. Genet.* **2010**, *42*, 722-726.

(212) Bejar, R.; Stevenson, K.; Abdel-Wahab, O.; Galili, N.; Nilsson, B.; Garcia-Manero, G.; Kantarjian, H.; Raza, A.; Levine, R. L.; Neuberg, D.; Ebert, B. L. *N. Engl. J. Med.* **2011**, *364*, 2496-2506.

(213) Guglielmelli, P.; Biamonte, F.; Score, J.; Hidalgo-Curtis, C.; Cervantes, F.; Maffioli, M.; Fanelli, T.; Ernst, T.; Winkelman, N.; Jones, A. V.; Zoi, K.; Reiter, A.; Duncombe, A.; Villani, L.; Bosi, A.; Barosi, G.; Cross, N. C. P.; Vannucchi, A. M. *Blood* **2011**, *118*, 5227-5234.

(214) Simon, C.; Chagraoui, J.; Krosl, J.; Gendron, P.; Wilhelm, B.; Lemieux, S.; Boucher, G.; Chagnon, P.; Drouin, S.; Lambert, R.; Rondeau, C.; Bilodeau, A.; Lavallee, S.; Sauvageau, M.; Hebert, J.; Sauvageau, G. *Genes Dev.* **2012**, *26*, 651-656.

(215) Ntziachristos, P.; Tsirigos, A.; Van Vlierberghe, P.; Nedjic, J.; Trimarchi, T.; Flaherty, M. S.; Ferres-Marco, D.; da Ros, V.; Tang, Z. J.; Siegle, J.; Asp, P.; Hadler, M.; Rigo, I.; De Keersmaecker, K.; Patel, J.; Huynh, T.; Utro, F.; Poglio, S.; Samon, J.

B.; Paietta, E.; Racevskis, J.; Rowe, J. M.; Rabadan, R.; Levine, R. L.; Brown, S.; Pflumio, F.; Dominguez, M.; Ferrando, A.; Aifantis, I. *Nat. Med.* **2012**, *18*, 296-301.

(216) Zhang, J. H.; Ding, L.; Holmfeldt, L.; Wu, G.; Heatley, S. L.; Payne-Turner, D.; Easton, J.; Chen, X.; Wang, J. M.; Rusch, M.; Lu, C.; Chen, S. C.; Wei, L.; Collins-Underwood, J. R.; Ma, J.; Roberts, K. G.; Pounds, S. B.; Ulyanov, A.; Becksfort, J.; Gupta, P.; Huether, R.; Kriwacki, R. W.; Parker, M.; McGoldrick, D. J.; Zhao, D.; Alford, D.; Espy, S.; Bobba, K. C.; Song, G. C.; Pei, D. Q.; Cheng, C.; Roberts, S.; Barbato, M. I.; Campana, D.; Coustan-Smith, E.; Shurtleff, S. A.; Raimondi, S. C.; Kleppe, M.; Cools, J.; Shimano, K. A.; Hermiston, M. L.; Doulatov, S.; Eppert, K.; Laurenti, E.; Notta, F.; Dick, J. E.; Basso, G.; Hunger, S. P.; Loh, M. L.; Devidas, M.; Wood, B.; Winter, S.; Dunsmore, K. P.; Fulton, R. S.; Fulton, L. L.; Hong, X.; Harris, C. C.; Dooling, D. J.; Ochoa, K.; Johnson, K. J.; Obenauer, J. C.; Evans, W. E.; Pui, C. H.; Naeve, C. W.; Ley, T. J.; Mardis, E. R.; Wilson, R. K.; Downing, J. R.; Mullighan, C. G. *Nature* **2012**, *481*, 157-163.

(217) Schwartzentruber, J.; Korshunov, A.; Liu, X. Y.; Jones, D. T. W.; Pfaff, E.; Jacob, K.; Sturm, D.; Fontebasso, A. M.; Quang, D. A. K.; Tonjes, M.; Hovestadt, V.; Albrecht, S.; Kool, M.; Nantel, A.; Konermann, C.; Lindroth, A.; Jager, N.; Rausch, T.; Ryzhova, M.; Korbel, J. O.; Hielscher, T.; Hauser, P.; Garami, M.; Klekner, A.; Bognar, L.; Ebinger, M.; Schuhmann, M. U.; Scheurlen, W.; Pekrun, A.; Fruhwald, M. C.; Roggendorf, W.; Kramm, C.; Durken, M.; Atkinson, J.; Lepage, P.; Montpetit, A.; Zakrzewska, M.; Zakrzewski, K.; Liberski, P. P.; Dong, Z. F.; Siegel, P.; Kulozik, A. E.; Zapatka, M.; Guha, A.; Malkin, D.; Felsberg, J.; Reifenberger, G.; von Deimling, A.; Ichimura, K.; Collins, V. P.; Witt, H.; Milde, T.; Witt, O.; Zhang, C. D.; Castelo-Branco, P.; Lichter, P.; Faury, D.; Tabori, U.; Plass, C.; Majewski, J.; Pfister, S. M.; Jabado, N. *Nature* **2012**, *482*, 226-231.

(218) Wu, G.; Broniscer, A.; McEachron, T. A.; Lu, C.; Paugh, B. S.; Becksfort, J.; Qu, C. X.; Ding, L.; Huether, R.; Parker, M.; Zhang, J. Y.; Gajjar, A.; Dyer, M. A.; Mullighan, C. G.; Gilbertson, R. J.; Mardis, E. R.; Wilson, R. K.; Downing, J. R.; Ellison, D. W.; Zhang, J. H.; Baker, S. J.; Hosp, S. J. C. s. R. *Nat. Genet.* **2012**, *44*, 251-253.

(219) Lewis, P. W.; Muller, M. M.; Koletsky, M. S.; Cordero, F.; Lin, S.; Banaszynski, L. A.; Garcia, B. A.; Muir, T. W.; Becher, O. J.; Allis, C. D. *Science* **2013**, *340*, 857-861.

(220) Tan, J.; Yang, X. J.; Zhuang, L.; Jiang, X.; Chen, W.; Lee, P. L.; Karuturi, R. K. M.; Tan, P. B. O.; Liu, E. T.; Yu, Q. *Genes Dev.* **2007**, *21*, 1050-1063.

(221) Chiang, P. K.; Cantoni, G. L. *Biochem. Pharmacol.* **1979**, *28*, 1897-1902.

(222) Glazer, R. I.; Hartman, K. D.; Knode, M. C.; Richard, M. M.; Chiang, P. K.; Tseng, C. K. H.; Marquez, V. E. *Biochem. Biophys. Res. Commun.* **1986**, *135*, 688-694.

(223) Miranda, T. B.; Cortez, C. C.; Yoo, C. B.; Liang, G. N.; Abe, M.; Kelly, T. K.; Marquez, V. E.; Jones, P. A. *Mol. Cancer Ther.* **2009**, *8*, 1579-1588.

(224) Chang, Y. Q.; Ganesh, T.; Horton, J. R.; Spannhoff, A.; Liu, J.; Sun, A. M.; Zhang, X.; Bedford, M. T.; Shinkai, Y. C.; Snyder, J. P.; Cheng, X. D. *J. Mol. Biol.* **2010**, *400*, 1-7.

(225) Liu, F.; Barsyte-Lovejoy, D.; Allali-Hassani, A.; He, Y.; Herold, J. M.; Chen, X.; Yates, C. M.; Frye, S. V.; Brown, P. J.; Huang, J.; Vedadi, M.; Arrowsmith, C. H.; Jin, J. *J. Med. Chem.* **2011**, *54*, 6139-6150.

(226) Liu, F.; Chen, X.; Allali-Hassani, A.; Quinn, A. M.; Wasney, G. A.; Dong, A.; Barsyte, D.; Kozieradzki, I.; Senisterra, G.; Chau, I.; Siarheyeva, A.; Kireev, D. B.; Jadhav, A.; Herold, J. M.; Frye, S. V.; Arrowsmith, C. H.; Brown, P. J.; Simeonov, A.; Vedadi, M.; Jin, J. *J. Med. Chem.* **2009**, *52*, 7950-7953.

(227) Liu, F.; Chen, X.; Allali-Hassani, A.; Quinn, A. M.; Wigle, T. J.; Wasney, G. A.; Dong, A. P.; Senisterra, G.; Chau, I.; Siarheyeva, A.; Norris, J. L.; Kireev, D. B.; Jadhav, A.; Herold, J. M.; Janzen, W. P.; Arrowsmith, C. H.; Frye, S. V.; Brown, P. J.; Simeonov, A.; Vedadi, M.; Jin, J. A. *J. Med. Chem.* **2010**, *53*, 5844-5857.

(228) Vedadi, M.; Barsyte-Lovejoy, D.; Liu, F.; Rival-Gervier, S.; Allali-Hassani, A.; Labrie, V.; Wigle, T. J.; DiMaggio, P. A.; Wasney, G. A.; Siarheyeva, A.; Dong, A.; Tempel, W.; Wang, S.-C.; Chen, X.; Chau, I.; Mangano, T. J.; Huang, X.-p.;

Simpson, C. D.; Pattenden, S. G.; Norris, J. L.; Kireev, D. B.; Tripathy, A.; Edwards, A.; Roth, B. L.; Janzen, W. P.; Garcia, B. A.; Petronis, A.; Ellis, J.; Brown, P. J.; Frye, S. V.; Arrowsmith, C. H.; Jin, J. *Nat. Chem. Biol.* **2011**, *7*, 566-574.

(229) Knutson, S. K.; Warholic, N. M.; Wigle, T. J.; Klaus, C. R.; Allain, C. J.; Raimondi, A.; Porter Scott, M.; Chesworth, R.; Moyer, M. P.; Copeland, R. A.; Richon, V. M.; Pollock, R. M.; Kuntz, K. W.; Keilhack, H. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 7922-7927.

(230) Chapman-Rothe, N.; Shamsaei, E.; Cherblanc, F.; Curry, E.; Ganesh, T.; Srimongkolpithak, N.; Liu, J.; Snyder, J. P.; Fuchter, M. J.; Brown, R. *In preparation* **2013**.

(231) Jung, F. H.; Pasquet, G.; Van der Brempt, C. L.; Lohmann, J.-J. M.; Warin, N.; Renaud, F.; Germain, H.; De Savi, C.; Roberts, N.; Johnson, T.; Dousson, C.; Hill, G. B.; Mortlock, A. A.; Heron, N.; Wilkinson, R. W.; Wedge, S. R.; Heaton, S. P.; Odedra, R.; Keen, N. J.; Green, S.; Brown, E.; Thompson, K.; Brightwell, S. *J. Med. Chem.* **2006**, *49*, 955-970.

(232) Lane, J. W.; Chen, Y. Y.; Williams, R. M. *J. Am. Chem. Soc.* **2005**, *127*, 12684-12690.

(233) Fuji, K.; Ichikawa, K.; Node, M.; Fujita, E. *J. Org. Chem.* **1979**, *44*, 1661-1664.

(234) Corey, E. J.; Suggs, J. W. *J. Org. Chem.* **1973**, *38*, 3224-3224.

(235) BoisChoussy, M.; Neuville, L.; Beugelmans, R.; Zhu, J. P. *J. Org. Chem.* **1996**, *61*, 9309-9322.

(236) Ando, H.; Manabe, S.; Nakahara, Y.; Ito, Y. *J. Am. Chem. Soc.* **2001**, *123*, 3848-3849.

(237) Thomas, R. M.; Mohan, G. H.; Iyengar, D. S. *Tetrahedron Lett.* **1997**, *38*, 4721-4724.

(238) Navarro, I.; Basset, J. F.; Hebbe, S.; Major, S. M.; Werner, T.; Howsham, C.; Brackow, J.; Barrett, A. G. M. *J. Am. Chem. Soc.* **2008**, *130*, 10293-10298.

(239) Vutukuri, D. R.; Bharathi, P.; Yu, Z.; Rajasekaran, K.; Tran, M.-H.; Thayumanavan, S. *J. Org. Chem.* **2002**, *68*, 1146-1149.

(240) Weissman, S. A.; Zewge, D. *Tetrahedron* **2005**, *61*, 7833-7863.

(241) Jiang, H.; Newcombe, N.; Sutton, P.; Lin, Q. H.; Mullbacher, A.; Waring, P. *Aust. J. Chem.* **1993**, *46*, 1743-1754.

(242) Ottenheijm, H. C. J.; Herscheid, J. D. M.; Kerkhoff, G. P. C.; Spande, T. F. *J. Org. Chem.* **1976**, *41*, 3433-3438.

(243) Jerumanis, S.; Lalancette, J. M. *Can. J. Chem.* **1964**, *42*, 1928-1935.

(244) Romeiro, L. A. S.; da Silva Ferreira, M.; da Silva, L. L.; Castro, H. C.; Miranda, A. L. P.; Silva, C. L. M.; Noël, F.; Nascimento, J. B.; Araújo, C. V.; Tibiriçá, E.; Barreiro, E. J.; Fraga, C. A. M. *Eur. J. Med. Chem.* **2011**, *46*, 3000-3012.

(245) R. Dennison, P.; Gibson, A.; I. Gray, A.; L. Patrick, G. *J. Chem. Soc. Perk. Trans. 1* **1997**, *0*, 721-728.